Pergamon

Bioorganic & Medicinal Chemistry, Vol. 5, No. 6, pp 1043–1050, 1997

PII: S0968-0896(97)00059-X
0968-0896/97 $17.00 + 0.00

Recognition of RNA by Triplex Formation: Divergent Effects of Pyrimidine C-5 Methylation

Shaohui Wang,† Yanzheng Xu and Eric T. Kool*

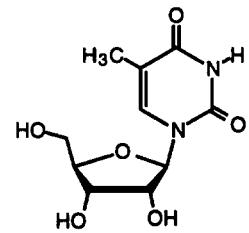
Department of Chemistry and Department of Biochemistry and Biophysics, University of Rochester, Rochester, NY 14627, U.S.A.

Abstract—In DNA triple helices, methylation at C-5 of thymine or cytosine is reported to have similar stabilizing effects for both bases. Here we show, however, that methylation of the same positions in RNA triplexes has distinctly different effects than in DNA. We have previously described the use of circular triplex-forming RNA oligonucleotides to recognize RNA sequences. Here it is shown that addition of C-5 methyl groups to uracils in these compounds very significantly increases not only affinity but also sequence selectivity in binding a purine-rich RNA target, as measured by thermal denaturation with various target RNAs. Surprisingly, however, addition of C-5 methyl groups to cytosines actually decreases affinity in binding RNA, while the same substitution in DNA is thermally stabilizing. Possible sources of this divergent behavior are discussed. A synthesis of 5-methylcytidine ribonucleoside 2′-*O*-silyl-3′-*O*-phosphoramidite is also described. © 1997 Elsevier Science Ltd.

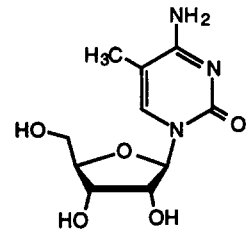
Introduction

A few recent studies have described a novel and effective approach to the recognition of RNA, in which linear or circular single-stranded oligonucleotides form cooperative triple helical complexes at a specific site in an RNA strand.[1–5] Such an approach can lead to high affinity, in part because the target sequence is bound on two sides by the oligonucleotide ligands. We have described the construction and RNA-binding properties of circular RNA oligonucleotides,[1] and more recently, circular chimeric oligonucleotides containing combinations of RNA, DNA, and 2′-O-methyl RNA nucleosides.[2]

Crothers and Roberts first described the sensitivity of triple helix stability to the backbone composition.[6] In a finding related to this, we have shown that circular pyrimidine-rich DNAs can bind DNAs tightly, but cannot form similar triplexes with RNA strands.[1] To obtain evidence as to the origin of the differences in DNA and RNA triplexes, we synthesized circular triplex-forming oligonucleotides with and without 2′-OH groups and with and without thymine(uracil) C-5 methyl groups.[7] The results showed that both 2′-OH groups and thymine methyl groups have strong effects on the stabilities of triple helices; the hydroxyl group was found to be either stabilizing or destabilizing, depending on the case. The thymine methyl group was found always to be stabilizing both to double and triple helical structures, adding ≈0.1–0.5 kcal/mol of stabilization per methyl group.



1
5-Methyluridine



2
5-Methylcytidine

Those previous studies led to the conclusion that, for highest-affinity recognition of RNA, a circular oligonucleotide composed of cytidine and ribothymidine (**1**) residues was optimum.[7] For example, a 34-nucleotide circular oligomer containing those residues was shown to bind a 12-base purine RNA complement with a free energy (37 °C) of −16.5 kcal/mol at neutral pH, while a standard linear Watson–Crick DNA complement binds the same sequence with a much lower free energy of −10.2 kcal/mol; thus the uracil-methylated circular ligand has a 30,000-fold advantage in association constant.

We describe here new studies of the effects of methyl groups on RNA-binding properties of circular oligonucleotides. We find that the addition of the C-5 methyl group to uracil (**1**) not only increases binding affinity but also greatly increases sequence specificity. In addition, since it has become common in DNA triplexes to add a C-5 methyl group to cytosine (m^5C) to increase binding affinity, we decided to make this substitution in RNA as well (**2**). We developed a synthesis of the corresponding m^5C phosphoramidite (**3**) and incorporated it into circular RNA. Interestingly, this substitution is found to be destabilizing to the RNA triplex,

Fax: (716) 473-6889. E-mail: etk@etk.chem.rochester.edu
†Current address: Department of Chemistry, University of Notre Dame, Notre Dame, IN 46556, U.S.A.

unlike the effect of m^5C in DNA triplexes and unlike the effect of 5-me-U in RNA triplexes.

Results

Effects of C-5 methyluracil on sequence specificity

A previous study established that the addition of C-5 methyl groups to uracils significantly increases the affinity of triplex formation between a circular RNA and an RNA complement.[7] To examine whether this substitution also affects the selectivity against mismatched targets, we studied the binding of circular RNA **5**, which contains 18 C-5 methylated uridines, to four different purine-rich RNA targets. These four sequences contain one variable nucleotide near the center of the sequence, giving rise to a complementary or mismatched base triad (U·A-U, U·G-U, U·C-U, or U·U-U) in the center of the complex. We compared the results for a circular RNA of the same sequence but lacking the 5-methyl groups on uracil (sequence **4**). Binding affinity was evaluated at pH 7.0 in a buffer containing 100 mM Na^+ and 10 mM Mg^{2+}, and was measured by thermal denaturation experiments monitored at 260 nm.

The results are shown in Table 1. The unmethylated natural RNA circle **3** has moderate sequence selectivity, as measured by the difference in free energy between matched and mismatched complexes.[7] Selectivity in that case ranges from 1.5 to 4.3 kcal/mol (37 °C), which corresponds to ≈10- to 1000-fold selectivity in binding constant. Interestingly, the addition of C-5 methyl groups to the uracils (circle **5**) very significantly increases sequence selectivity. For the mismatched complexes the binding energy drops by 5.4–8.1 kcal/mol, corresponding to 6000- to 500,000-fold selectivity in equilibrium binding constant (Fig. 1). For both methylated and unmethylated cases, the U·G-U mismatch is considerably less destabilizing than the other two mismatches, although in the methylated case the discrimination is higher than that for all mismatches in the unmethylated case.

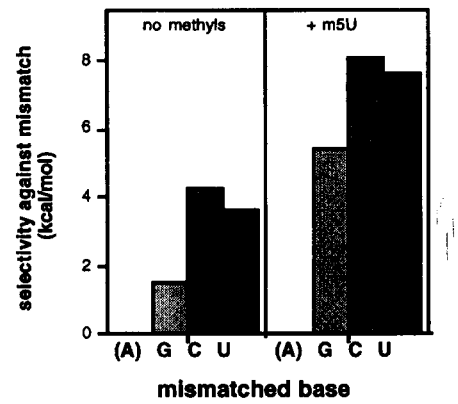


Figure 1. Effect of C-5 methyls on sequence selectivity of circular RNAs in binding matched and singly-mismatched target RNA strands at pH 7.0. On left, selectivity of unmethylated circle **4**; on right, selectivity of uracil-methylated circle **5**. Selectivity is defined as the free energy difference ($\Delta\Delta G°$) in binding the correct target and the mismatched one shown. Sequences and conditions are shown in Table 1.

Table 1. Effect of C-5 methyls on sequence selectivity. Melting temperatures (T_m) and free energies ($-\Delta G°_{37}$) for complexes of circular RNAs with matched and mismatched RNA target sequences at pH 7.0[a]

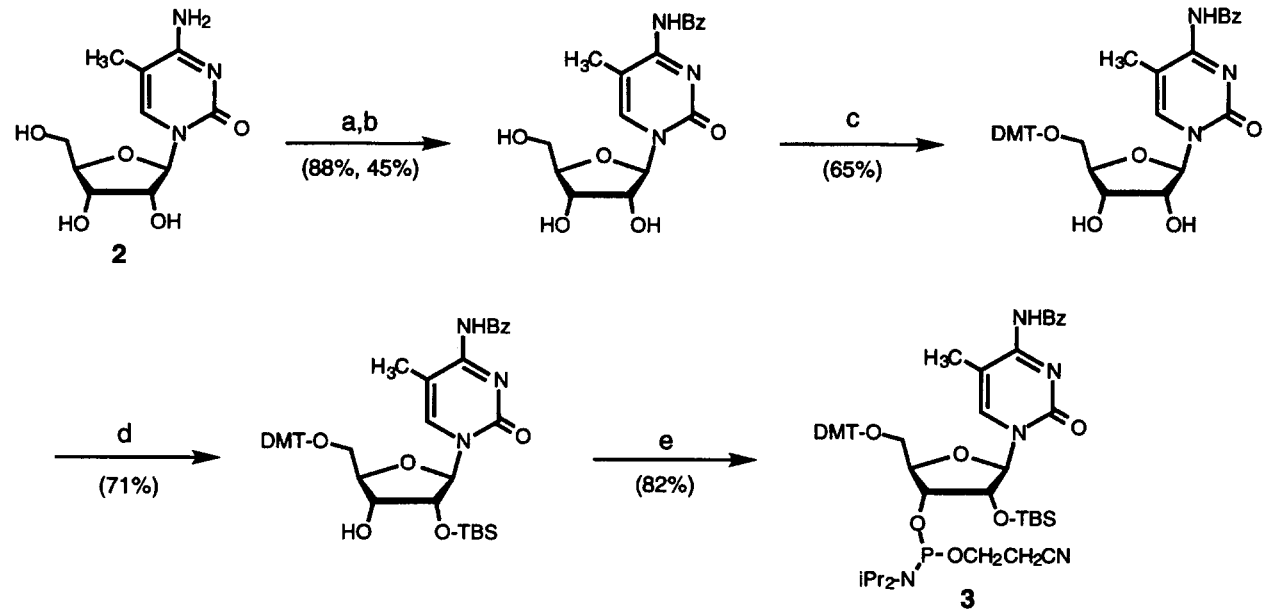
Target sequence	T_m[b,c] (°C)	$-\Delta G°_{37}$ (kcal/mol)
Without C-5 methyls:	A CU UCU UUC U UUU CC A / C ribo- C / A CU UCU UUC U UUU CC A	
5′-rAAGAAAGAAAG	51.2	15.4
5′-rAAGAAGGAAAG	44.1	13.9
5′-rAAGAACGAAAG	37.5	11.1
5′-rAAGAAUGAAAG	36.7	11.8
With ribothymidine:	A CTTCTT TCT TT TCC A / C ribo- C / A CTTCTT TCTTT TCC A	
5′-rAAGAAAGAAAG	59.2	20.3
5′-rAAGAAGGAAAG	51.1	14.9
5′-rAAGAACGAAAG	40.4	12.2
5′-rAAGAAUGAAAG	42.4	12.7

[a]Underlined C residues lack a 2′-OH for simplification of synthesis.
[b]Conditions: 3.0 μM total strand concentration, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM Na·PIPES buffer.
[c]Error limits for individual measurements are estimated at ± 0.5 °C in T_m and ± 5–10% in free energy.

Synthesis and RNA incorporation of 5-methylcytidine phosphoramidite

A number of previous studies have shown that methylation of cytosine significantly increases the stability of DNA triple helical complexes.[8–11] We therefore undertook to make this substitution in an all-RNA triplex to see whether it might increase binding affinity. The silyl-protected phosphoramidite derivative of 5-methylcytidine (**3**) was synthesized using methods similar to those for unmethylated cytidine (Scheme 1).[12] The N^4 amino group was benzoyl-protected, and then the 5′-dimethoxytrityl group was added, the 2′-OH was silylated selectively,[13] and the 3′-hydroxyl was phosphitylated.

The protected 5-methylcytidine phosphoramidite was coupled into RNA strands using the standard RNA cycle. Successful incorporation of this nucleoside was confirmed by enzymatic digestion of a strand into its component nucleotides, followed by HPLC analysis of nucleotide composition with reference to known standards. This showed that the 5-methylcytidine was indeed incorporated intact, with complete deprotection of the benzoyl group and with no detectable hydrolysis to 5-methyluridine (ribothymidine). A 34-nucleotide



Scheme 1. (a) BzCl, pyridine, rt; (b) NaOH, THF/MeOH/H_2O, 0 °C; (c) dimethoxytrityl chloride, pyridine, 0 °C; (d) pyridine, TBDMSCl, $AgNO_3$, THF, rt; (e) $CIP(NiPr_2)OCH_2CH_2CN$, collidine, N-methylimidazole, THF, rt.

Table 2. Effect of addition of C-5 methyls to uracil and to cytosine on binding properties of circular RNAs. Melting temperatures (T_m) and free energies ($-\Delta G°_{37}$) for bimolecular triplexes with no C-5 methyl groups, with C-5 methyls on uracils and C-5 methyls on cytosines at pH 5.5 and 7.0[a]

Circular RNA ligand	Purine target pH	DNA 5(d(AAGAAAGAAAAG)		RNA 5(-r(AAGAAAGAAAAG)	
		T_m (°C)[b,c]	$-\Delta G°_{37}$	T_m (°C)[b,c]	$-\Delta G°_{37}$
A CU UCU UUCU UUUCC A / C C / A CU UCU UUCU UUUCC A	7.0	48.5	11.8	51.2	12.6
	5.5	62.3	16.4	62.9	17.7
A CTTCTT TCT TT TCC A / C C / A CTTCTT TCTTT TCC A	7.0	59.9	16.0	59.2	16.5
	5.5	73.4	24.4	69.2	19.5
A mCTTmCTTTmCTTTTmCmC A / mC Cm / A mCTTmCTTTmCTTTTmCmC A	7.0	66.8	14.6	60.8	13.8
	5.5	79.0	18.0	71.4	15.4

[a]Underlined C residues lack a 2′-OH for simplification of synthesis. 5-Methylcytosines are designated mC.
[b]Conditions: 3.0 μM total strand concentration, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM Na·PIPES buffer.
[c]Error limits for individual measurements are estimated at ± 0.5 °C in T_m and ± 5–10% in free energy.

circular RNA (sequence **6**) containing both ribothymidine and 5-methylcytidine was synthesized using the same method as for the circle **5**;[7] the circularity was confirmed by nicking with base and examining products by gel electrophoresis relative to the linear precursor as before.[1]

RNA- and DNA-binding properties of fully C-5-methylated RNA

The circular RNA (**6**) carrying 30 C-5 methyl groups on all pyrimidines was then studied for its ability to bind a complementary RNA strand both at neutral and acidic pH. Comparison was made to the binding of RNA circle **5**, which has only methylated uracils, and naturally substituted RNA circle **4**, which has no C-5 methyls. Thermal denaturation studies were performed with the same ionic strength conditions described above.

The results are shown in Table 2 and Figure 2. As seen previously,[7] the natural unmethylated circular RNA binds its RNA complement with a T_m of 51.2 °C and a free energy (37 °C) of −12.6 kcal/mol at neutral pH. Addition of 18 methyls to the uracils in the circle is significantly stabilizing, raising the T_m another 8.0 °C and adding 3.9 kcal/mol of favorable binding free energy. However, at neutral pH the addition of 12 more methyl groups, this time to cytosine, has almost no effect on T_m, and binding free energy is actually less favorable by 2.7 kcal relative to the case with

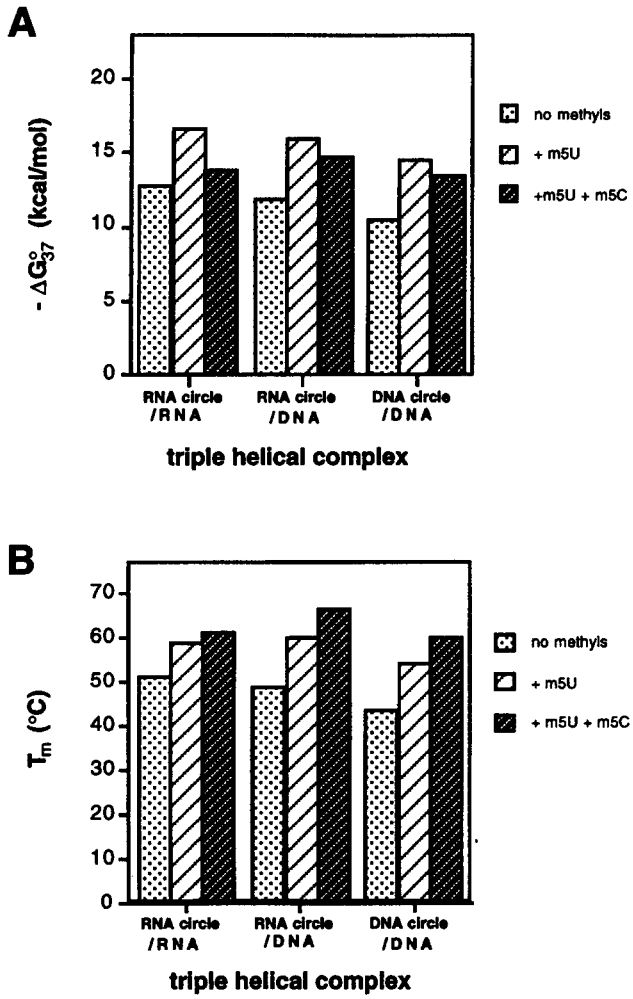


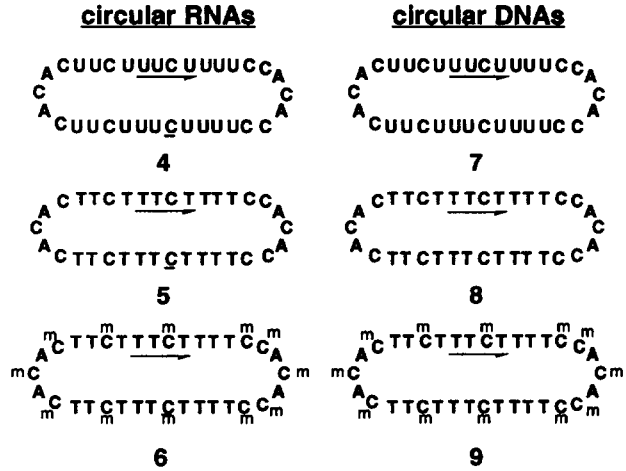
Figure 2. Effect of uracil and cytosine methylation on the free energies (A) and melting temperatures (B) of complexes between circular RNAs (**4–6**) and DNAs (**7–9**) and their 12-mer purine complements (sequence 5′-AAGAAAGAAAAG) at pH 7. Conditions and sequences are given in Table 2. The binding of circular DNAs to RNA complements is not shown because they do not form triplexes.[1,6]

methylated uracils alone. The results at acidic pH are much the same, but all complexes are stabilized relative to pH 7.0, consistent with the expected pH-dependent triple helical structure of the complexes. In general, then, the results show that methylation of cytosine in a triplex-forming RNA ligand is thermodynamically destabilizing.

Circular RNAs can bind DNA as well as RNA complements, forming RNA-DNA-RNA (R-D-R) triple helices.[1] We therefore examined the generality of this 5-methylcytosine destabilizing effect by testing the same three circular RNAs (**4–6**) for their ability to bind the DNA version of the same target sequence (Table 2). The results of those studies are similar to those for binding the RNA complement: addition of methyls to uracil is stabilizing, increasing both T_m and favorable free energy (and giving values almost the same as for binding the RNA strand). However, as before, the further addition of C-5 methyls to the cytosines causes thermodynamic destabilization, with free energy becoming 1.4 kcal/mol less favorable than the case with uracil C-5 methyls alone at neutral pH. One difference between binding RNA and DNA is evident, however: although the addition of the cytosine methyl groups is thermodynamically destabilizing in the R-D-R case, it is less so than for the complex with an RNA strand (the R-R-R case). Moreover, with addition of methyls to cytosines in the R-D-R case the T_m actually increases while the free energy becomes less favorable, indicating a lowering of cooperativity in the complex. In general, then, for binding an DNA strand, m^5C in a circular RNA is weakly stabilizing thermally but is destabilizing thermodynamically.

To compare the effect of cytosine methylation or the RNA circle with the effect on a DNA circle, we also synthesized a fully methylated DNA circle (sequence **9**, carrying both thymine and 5-methyldeoxycytosine) having the same sequence as the permethylated RNA circle **6**. The only difference between these two circles is the 33 2′-hydroxyls present on the RNA circle. Also available for comparison are previous results from DNA circles of the same sequence but which carry either no methyl groups (**7**), or methyls only on thymines[7] (**8**). The binding of the purine DNA complement was again measured for DNA circle **9** by thermal denaturation; the results are shown in Figure 2 in comparison to the effects for the RNA circles in this study. The data show significant thermal stabilization by m^5C in the circular DNA binding its DNA complement, although the free energy shows no effect or perhaps a slight destabilization.

The data confirm previous findings that cytosine methylation is thermally stabilizing for DNA (D-D-D-type) triple helices,[8,10] and also point out the contrasting behavior of cytosine methylation as compared to uracil methylation, as well as varied behavior depending on RNA content. As shown in Figure 2(A), the free energies of binding become more favorable in all cases on adding methyls to uracil, but adding methyls to



cytidine is destabilizing for the RNA circle, and energetically neutral for the DNA circle. The degree of destabilization appears to depend on the RNA content of the triplex, as cytosine methylation becomes less favorable for an R-D-R triplex and even less so for the R-R-R triplex.

Discussion

Origins of increased binding affinity with 5-methyluracil

The addition of a C-5 methyl group to uracil is significantly stabilizing in our study, and indeed, this substitution appears to be stabilizing in all double and triple helical complexes yet studied,[7] regardless of DNA or RNA composition. The origin of this effect is perhaps best explained by an increase in base-stacking ability of thymine relative to uracil.[7] This has been rationalized by an increase in polarizability due to the methyl substitution.[14] Differential DNA solvation effects may also play a role;[15] further studies are needed to clarify this effect.

Possible origins of elevated sequence specificity with 5-methyluracil

Interestingly, the present results show that not only is affinity on RNA recognition increased by methylation of uracil, but also sequence selectivity is very significantly enhanced as well. We have previously shown that circular DNAs have higher discrimination against mismatches than do standard linear DNAs; this positive effect was shown to be due in part to the stabilizing protonation of cytosine that occurs in such triplexes (but not with mismatched targets).[16] The selectivity enhancement is also due to the higher cooperativity (rigidity) of the circular complex relative to linear cases. The structural disruption of mismatches is apparently communicated more effectively throughout the complex and is thus more destabilizing.[16,17]

It seems possible, therefore, that the increase in sequence selectivity seen with methylation of uracils arises from a further increase in cooperativity on methylation of the circular RNA. The complex is clearly stabilized by the presence of the methyls (Table 2), presumably by enhanced base-stacking interactions, and this increased stacking likely adds rigidity to the structure. To our knowledge this is the first observation that such a C-5 substitution can have this desirable effect on specificity, and it is possible that it may be generally applicable to many nucleic acid complexes.

Possible origins of destabilization by 5-methylcytosine

The results demonstrate clearly that cytosine methylation is not universally stabilizing in triplexes. This has not to our knowledge been observed previously, possibly because until recently most triplex studies were carried out using DNAs as ligands. We find that increasing RNA content of a triplex is associated with an increasing destabilizing effect of the C-5 methyl group on cytosine. We do observe a thermal stabilization effect in cases with greater DNA strand content, but this disappears in an all-RNA triplex. Previous studies of that substitution in all-DNA triplexes have shown that the methylation does not enhance the basicity of cytosine,[18] and so the stabilization is thought to be due either to enhanced base-stacking interactions[14] (the same as for methylation of uracil) or to enhanced solvation effects for the complex.[15] However, it is difficult to rationalize why methylating cytosine causes destabilization in the same complex where methylation of uracil is stabilizing, and why the m^5C effect depends on the DNA/RNA content.

It seems likely that explanations for this divergent behavior must await new and in-depth structural characterization of DNA versus RNA triple helices. Since Crothers showed that triplex stability can vary greatly with RNA vs DNA content,[6] and Dervan presented evidence that the structure also varies significantly with RNA/DNA content,[19] it seems possible that our observed sensitivity of the m^5C effect to backbone composition may lie in this structural variation. It is also possible that C-G-C triads (or base steps involving them) may have a different geometry than U-A-U triads, and so the divergent effects of uracil vs cytosine methylation may also be studied by gaining greater structural information on the complexes.

It is worth noting that one related finding has been reported recently which supports the current observations.[20] A study of the binding of duplex DNA by 2′-O-methyl RNA oligonucleotides showed that C-5 methylation of the 2′-O-me-cytidine residues was unfavorable, while methylation of 2′-O-me-uridines was favorable for triplexes with those non-natural analogues. The authors of that study conjectured that the methyl group might lower the pK_a of cytosine,[20] as has been seen in propynyl-substituted triplex-forming DNAs;[21] however, this seems unlikely, since methyl substitution in a deoxynucleoside or an oligonucleotide does not significantly alter pK_a,[18] and the methyl group is electron-donating and thus would be expected, if anything, to increase pK_a; moreover, it does not explain why there is a difference between m^5C in DNA and in RNA. Thus, although the precise origins of the divergent effects of cytosine methylation cannot be explained adequately without further study, the practical result is clear: in the binding of RNA strands by RNA triplex-forming ligands, one can gain both stability and sequence selectivity by methylation of uracils in the ligand, but methylation of cytosines is counter productive.

Conclusions

Several conclusions can be made from these experimental results. First, addition of C-5 methyl groups to

uridine in triplex-forming RNAs can very significantly increase sequence specificity (discrimination against mismatches) in binding an RNA strand. We have previously shown that this substitution increases binding affinity as well.[7] Second, although addition of C-5 methyl groups to cytidine in DNA triplex-forming oligonucleotides is commonly used to increase affinity for binding DNA, the same substitution in an RNA triplex-forming ligand actually decreases affinity for either DNA or RNA strands. The findings are useful in the design of improved ligands for RNA, and thus may have utility in applications such as RNA-directed medical diagnostics and therapeutics.

Experimental

Synthesis of 2′-*O*-*t*-Butyldimethylsilyl-5′-dimethoxytrityl-ribothymidine-3′-(cyanoethyl-*N,N*-diisopropylphosphoramidite) (1). The ribothymidine phosphoramidite derivative was synthesized by the published method;[7,22] ^{1}H NMR and mass spectral analysis agreed with published values.

2′-*O*,N^4-bis-Trimethylsilyl-5-methylcytosine. A 5 g (40 mmol) quantity of 5-methylcytosine (USB) was suspended in 30 mL (142 mmol) of hexamethyldisilazane (Aldrich) and refluxed under nitrogen for 5 h. After cooling the clear solution was concentrated under vacuum to yield a white solid, which was used without further purification.

2′,3′,5′-tri-*O*-Benzoyl-5-methylcytidine. An 11 g (22 mmol) quantity of protected ribose and 6.0 g (22 mmol) protected cytosine were added to 200 mL 1,2-dichoroethane, followed by 4 mL of $SnCl_2$ (34 mmol, Aldrich). The suspension turned pink. After 5 h the reaction was quenched by addition of 150 mL satd $NaHCO_3$. A copious amount of sticky white precipitate formed. The organic phase was carefully separated, and the aqueous phase was extracted repeatedly with 1,2-dichloroethane until the compound was absent by TLC. The organic phase was dried over Na_2SO_4 concentrated, and purified on silica gel with 6% MeOH/EtOAc. The product (7.6 g, 13.3 mmol (60%) was obtained as an off-white foam.

5-Methylcytidine (2). A 7.6 g (13.3 mmol) quantity of the benzoylated nucleoside was dissolved in 200 mL dry methanol. Seven milliliters of 2 mM freshly prepared $NaOCH_3$ were added and the mixture stirred at room temperature. After 14 h the pH of the solution was adjusted to 6–7 by addition of HCl/MeoH (prepared by addition of CH_3COCl to MeOH). The solution was then concentrated and purified on silica gel with 3:7 MeOH:EtOAc to yield 3.1 g (12 mmol (91%)) of a white foam. The proton NMR spectrum matched the published one for this compound.[23]

N^4-Benzoyl-5-methylcytidine. A 1.3 g (5 mmol) quantity of the above nucleoside was dissolved in 25 mL dry pyridine, and then 3.5 mL (30 mmol) benzoyl chloride was added. After stirring at room temperature for 1.5 h, the reaction mixture was poured into 200 mL ice-water containing 10 mL satd $NaHCO_3$. The aqueous phase was extracted with EtOAc. The organic phase was dried over Na_2SO_4, concentrated, and washed with Et_2O. Three grams of white solid were collected by filtration. The solid was dissolved in 500 mL 5:4:1 THF:MeOH:H_2O and cooled to 0 °C. The solution was stirred for 30 min before treatment with pyridinium·Dowex resin to adjust the pH to 6–7. After filtration to remove the resin, the solution was concentrated and lyophilized. The crude material was purified on silica gel eluting with EtOAc to yield 0.827 g (2.25 mmol (45%)) of white foam. ^{1}H NMR (DMSO-d_6) δ 1.99 (s, 3H, Me), 3.59–4.06 (m, 5H, $H_{2'-5'}$), 5.09 (d, 1H, J = 5.5, OH), 5.23 (t, 1H, J = 4.1, OH), 5.47 (d, 1H, J = 4.7, OH), 5.78 (d, 1H, J =3.8, $H_{1'}$), 7.46–8.19 (m, 5H, H-Bz), 12.97 (s, 1H, NH); MS (FAB$^+$) m/z 361.1271, calcd 361.1274.

5′-*O*-Dimethoxytrityl-N^4-benzoyl-5-methylcytidine. A 737 mg (2.0 mmol) quantity of the above compound was dissolved in 14 mL pyridine and cooled to 0 °C under nitrogen; 821 mg dimethoxytrityl chloride (2.4 mmol, Aldrich) was added in three portions to the solution over 8 h. The mixture was treated with 5 mL of satd $NaHCO_3$ and extracted with CH_2Cl_2. The organic phase was dried over Na_2SO_4, concentrated and purified on silica gel eluting with 4% MeOH/$CHCl_3$. A yellowish foam was obtained (825 mg, 1.25 mmol (62%)). ^{1}H NMR (DMSO-d_6) δ 1.58 (s, 3H, Me), 3.21–3.27 (m, 2H, $H_{5',5}$), 3.72 (s, 6H, OCH_3), 4.03–4.20 (m, 3H, $H_{2'}$–$H_{4'}$), 5.19 (d, 1H, J = 5.8, OH), 5.59 (d, 1H, J = 5.4, OH), 5.80 (d, 1H, J = 3.3, $H_{1'}$), 6.88–8.14 (m, 14H, H-DMTr, H_6); MS (FAB$^+$) m/z 664.2679, calcd 664.2659.

2′-*O*-*t*-Butyldimethylsilyl-5′-*O*-dimethoxytrityl-N^4-benzoyl-5-methyl-cytidine. A 740 mg (1.1 mmol) quantity of the above compound was dissolved in 11 mL dry THF. To this was added sequentially 330 μL of pyridine, 436 mg (2.57 mmol) $AgNO_3$, and after 10 min, 440 mg (2.9 mmol) of *t*-butyldimethylsilyl chloride (Aldrich). After 5 h the reaction was quenched by addition of 12 mL satd $NaHCO_3$ and extracted with CH_2Cl_2. The organic phase was dried over Na_2SO_4, concentrated and purified on silica gel eluting with 1:3 EtOAc:hexane to yield 615 mg (0.79 mmol (71%)) of an off-white foam. ^{1}H NMR (DMSO-d_6) δ 0.07 (s, 6H, Si–Me), 0.85 (s, 9H, Si–*t*-Bu), 1.55 (s, 3H, CH_3), 3.25 (m, 2H, $H_{5',5}$), 3.73 (s, 6H, OCH_3), 4.05–4.07 (m, 1H, $H_{4'}$), 4.13–4.15 (m, 1H, $H_{3'}$), 4.30–4.32 (t, 1H, J = 4.0, $H_{2'}$), 5.19 (d, 1H, J = 6.0, OH), 5.80 (d, 1H, J = 3.6, $H_{1'}$), 6.89–8.14 (m, 14H, H-DMTr); MS (FAB$^+$) m/z 778.3537, calcd 778.3524.

2′-*O*-*t*-Butyldimethylsilyl-5′-*O*-dimethoxytrityl-5-methyl-cytidine-3′-(cyanoethyl-*N,N*-diisopropylphosphoramidite) (3). A 560 mg (0.72 mmol) quantity of the above protected nucleoside was dissolved in 2.16 mL dry THF, and 720 μL collidine (Aldrich) and 28.8 μL N-methylimidazole (Aldrich) were added. *N,N*-Diisopropyl-*O*-2-cyanoethyl-phosphonamidic chloride (Aldrich; 402 μL, 1.80 mmol) was syringed into the solution over 5 min. The reaction was stirred at room temperature for

1 h, 7 mL of satd $NaHCO_3$ was then added, followed by 7 mL EtOAc. The organic phase was dried over Na_2SO_4, concentrated, and purified on silica with 5% CH_3CN/CH_2Cl_2 to yield a white foam (580 mg, 0.59 mmol (82%)). The foam was dissolved in 2 mL 1,4 dioxane and lyophilized three times prior to use. ^1H NMR (mixed diastereomers in DMSO-d_6) δ 0.02–0.09 (dd, 6H, Si–Me), 0.83–0.85 (ds, 9H, Si–*t*-Bu), 1.06–1.13 (m, 12H, $CH(CH_3)_2$), 1.54–1.57 (ds, 3H, CH_3), 2.54–2.76 (dt, 2H, CH_2–CN), 3.47–3.55 (m, 2H, $H_{5',5}$), 3.72 (s, 6H, OCH_3), 4.00–4.02 (m, 1H, $H_{4'}$), 4.20 (m, 1H, $H_{3'}$), 4.48–4.56 (m, 1H, $H_{2'}$), 5.83–5.91 (dd, 1H, $H_{1'}$), 6.87–8.11 (m, 14H, H-DMTr, H_6); MS (FAB^+) *m/z* 978.4653, calcd 978.4602.

Synthetic methods for linear and circular oligonucleotides

DNA oligomers were synthesized on an Applied Biosystems 392 automated synthesizer using the standard phosphoramidite method. RNA oligonucleotides were prepared using *t*-butyl-dimethylsilyl-protected phosphoramidites (Applied Biosystems) following the oligoribonucleotide synthesis procedure of Usman.[24] Tetrabutylammonium fluoride in THF (Aldrich) was dried over molecular sieves prior to use in the desilylation step.[25] For synthesis of circular oligomers the linear precursors were 5′-phosphorylated on the synthesizer using a commercially available reagent purchased from Cruachem.[26] Oligonucleotides were purified by preparative 20% denaturing polyacrylamide gel electrophoresis and quantitated by absorbance at 260 nm. Extinction coefficients for the oligomers were calculated by the nearest neighbor method.[27]

Intact incorporation of the m^5C nucleoside into RNA was confirmed by digestion of an oligonucleotide (including circle **6**) with nuclease P1 (Gibco BRL), followed by HPLC analysis of the mononucleotide products. The chromatogram showed only the expected products in the expected ratio; the m^5C nucleotide peak was identified by coinjection with an authentic sample and with comparison to authentic m^5U monophosphate to rule out deamination.

The circular DNAs **7** and **8** were described previously.[7] The circular DNA **9** was synthesized following exactly the same procedure, using a linear precursor having the sequence 5′-dpTTTmCTTmCAmCAmCTTmC-TTTmCTTTTmCmCAmCAmCmCTTTTmC, where mC is deoxycytidine (phosphoramidite obtained from Glen Research). The circular RNAs **4** and **5** were synthesized following the previously described methods.[7] For the RNA circles a 3′-deoxynucleotide was used to ensure correct 5′–3′ junction substitution in the ligation. The circular RNA **6** was constructed from a linear precursor having the sequence 5′-r(pTTT-mCTTmCAmCAmCTTmCTTTmCTTTTmCmCAmC-AmCmCTTTTmC), where T is m^5U and mC is m^5C (except for the 3′-terminal one, which is a deoxynucleoside). The same 12-nt oligomer 5′-dAAGAAA-GAAAAG was used as a template for cyclization for all circles, and the procedures for BrCN-mediated ligation and purification were as described previously.[1,7] The circularity of the chemical ligation products was confirmed by partial digestion by S1 nuclease (DNA circles) or partial alkaline hydrolysis (RNA circles) followed by PAGE gel analysis.[7] Products were visualized with Stains-all dye (Sigma). All circular compounds gave a single initial cleavage product which migrates with the mobility of the linear 34-mer precursors.

Thermal denaturation experiments

Solutions for the thermal denaturation studies contained a 1:1 molar ratio of pyrimidine oligonucleotide and complementary 12-nt purine oligomer (1.5 μM each). Solutions were buffered with 10 mM Na·PIPES (1,4-piperazine-bis(ethanesulfonate), Sigma) at pH 5.5 or 7.0. Also present in the denaturation solutions were 100 mM NaCl and 10 mM $MgCl_2$. The buffer pH is that of a 1.4 × stock solution containing buffer and salts at 25 °C. After the solutions with RNA were prepared they were heated to 90 °C and allowed to cool slowly to room temperature prior to the melting experiments. The melting studies were carried out as described previously.[1,7] Uncertainty in T_m is estimated at ± 0.5 °C based on repetitions of experiments. Free energy values (see Fig. 2) for the complexes were obtained from fits of the data using a two-state model with linear sloping baselines,[28] as described previously.[1,7] In general, precision in free energies is estimated at ± 5–10%, based on repetitions of experiments.

Acknowledgements

We thank the National Institutes of Health (GM46625) for support.

References

1. Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1994**, *22*, 2326.
2. Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1995**, *23*, 1157.
3. Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3806.
4. Reynolds, M. A.; Arnold, L. J.; Almazan, M. T.; Beck, T. A.; Hogrefe, R. I.; Metzler, M. D.; Stoughto, S. R.; Tseng, B. Y.; Trapane, T. L.; Ts'o, P. O. P. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2433.
5. Kandimalla, E. R.; Agrawal, S.; Vekataraman, G.; Sasisekharan, V. *J. Am. Chem. Soc.* **1995**, *117*, 6416.
6. Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463.
7. Wang, S.; Kool, E. T. *Biochemistry* **1995**, *34*, 4125.
8. Lee, J. S.; Woodsworth, M. L.; Latimer, L. J. P.; Morgan, A. R. *Nucleic Acids Res.* **1984**, *12*, 6603.
9. Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *111*, 3059.
10. Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625.

11. Kool, E. T. In *Delivery Systems for Antisense Oligonucleotide Therapeutics*; S. Akhtar, Ed.; CRC Press: Boca Raton, 1995; pp 123–149.

12. Usman, N.; Ogilvie, K. K.; Jiang, M.-Y.; Cedergren, R. J. *J. Am. Chem. Soc.* **1987**, *109*, 7845.

13. Hakimelahi, G. H.; Proba, Z. A.; Ogilvie, K. K. *Can. J. Chem.* **1982**, *60*, 1106.

14. Sowers, L. C.; Shaw, B. S.; Sedwick, W. D. *Biochem. Biophys. Res. Commun.* **1987**, *148*, 790.

15. Escudé, C.; François, J.-C.; Sun, J.; Ott, G.; Sprinzl, M.; Garestier, T.; Hélène, C. *Nucleic Acids Res.* **1993**, *21*, 5547.

16. Wang, S.; Friedman, A.; Kool, E. T. *Biochemistry* **1995**, *34*, 9774.

17. Kool, E. T. *J. Am. Chem. Soc.* **1991**, *113*, 6265.

18. (a) Sober, H. A. *Handbook of Biochemistry*; CRC Press: Cleveland, 1970. (b) Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625. (c) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995.

19. Han, H.; Dervan, P. B. *Nucleic Acids Res.* **1994**, *22*, 2837.

20. Shimizu, M.; Koizumi, T.; Inoue, H.; Ohtsuka, E. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1029.

21. Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S. R. *Tetrahedron Lett.* **1992**, *33*, 5307.

22. Maggio, A.-F.; Boyer, V.; Aubertin, A.; Obert, G.; Kirn, A.; Imbach, J.-L. *Nucleosides Nucleotides* **1991**, *10*, 1431.

23. *The Aldrich Library of ^{13}C and ^{1}H NMR Spectra*; Pouchert, C. J.; Behnke, J., Eds.; Aldrich Chemical: Milwaukee, 1993; Vol. 3, p 389.

24. Scaringe, S. A.; Francklyn, C.; Usman, N. *Nucleic Acids Res.* **1990**, *18*, 5433.

25. Hogrefe, R. I.; McCaffrey, A. P.; Borozdina, L. U.; McCampbell, E. S.; Vaghefi, M. M. *Nucleic Acids Res.* **1993**, *21*, 4739.

26. Horn, T.; Urdea, M. *Tetrahedron Lett.* **1986**, *27*, 4705.

27. Borer, P. N. In *Handbook of Biochemistry and Molecular Biology*, 3rd Ed.; Fasman, G. D., Ed.; CRC Press: Cleveland, 1985; Vol. 1, p 589.

28. Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256.

(*Received in U.S.A.* 27 *August* 1996; *accepted* 18 *February* 1997)