Origins of the Large Differences in Stability of DNA and RNA Helices: C-5 Methyl and 2′-Hydroxyl Effects†

Shaohui Wang and Eric T. Kool*

Department of Chemistry, University of Rochester, Rochester, New York 14627

Received October 19, 1994; Revised Manuscript Received January 17, 1995®

ABSTRACT: Recent studies have shown that there can be large differences in the stability of double and triple helical nucleic acid complexes, depending on whether RNA or DNA strands are involved. These differences have been attributed to structural differences in the sugar−phosphate backbone of these two polymers. However, since there are in fact two structural features which distinguish DNA from RNA (the 2′-hydroxyl and C-5 methyl groups), the stability differences may arise from either or both of these factors. We have separated effects of the 2′-hydroxyl and C-5 methyl groups by synthesizing nucleic acid strands which contain all possible combinations with and without these groups. Studies of the stabilities of double and triple helices involving these strands show that in fact the C-5 methyl group of thymine and the 2′-OH group of ribose have equally large effects on stability. The two effects vary with secondary structure and can be reinforcing or even opposing in their influence on stability. Three types of complexes are specifically examined: bimolecular pyrimidine·purine duplexes, termolecular pyrimidine·purine·pyrimidine triplexes, and bimolecular triplexes formed from circular pyrimidine oligonucleotides with purine target strands. It is found in general that the two types of substitutional effects are independent of one another and that C-5 methyl groups are in all cases stabilizing, while 2′-OH groups can be stabilizing or destabilizing, depending on the type of complex. In addition, studies with partially methylated duplexes lend evidence that the largest contribution to stabilization by the methyl group arises from increased base stacking ability rather than from a favorable hydrophobic methyl−methyl contact. The results allow explanation of observed differences in complexes of the two natural biopolymers. The data also point out what are the preferred structural features for highest-affinity synthetic oligonucleotide probes targeted to single- or double-stranded nucleic acids.

It has long been observed that there are considerable differences in the chemical and physical properties of DNA and RNA. For example, from the standpoint of reactivity, RNA, with its additional 2′-OH group, is hydrolytically much less stable than DNA. In addition, the C-5 methyl group of thymine in DNA is useful in chemically distinguishing it from deaminated cytosine, thus allowing accurate enzymatic repair of this mutagenic lesion. In terms of the physical stability of duplexes composed of the two polymers, it was observed early on that for long, sequence-averaged strands, RNA·RNA duplexes are more thermally stable than DNA·DNA duplexes of the same sequence, with DNA·RNA hybrid duplexes being the weakest of the three (Chamberlin & Berg, 1964). This relative ordering of stability has been rationalized by citing the differences in the two polymer backbones. However, this ordering does not always hold true; for example, homopyrimidine·homopurine duplexes composed from pyrimidine RNA and purine DNA strands are exceptionally weak, while the reverse mixed duplexes formed from purine RNA and pyrimidine DNA strands are much more thermally stable (Martin & Tinoco, 1980). It is clear, then, that structural differences in the backbone alone cannot explain all the observations.

Recent studies using short, sequence-defined oligonucleotides and involving RNA, DNA, and hydrid RNA−DNA double and triple helices have confirmed that there are indeed large differences in the thermodynamic stability for these different complexes (Martin & Tinoco, 1980; Hall & McLaughlin, 1991; Roberts & Crothers, 1992; Han & Dervan, 1993; Escudé *et al.*, 1993; Ratmeyer *et al.*, 1994); yet these differences largely have been attributed to the difference in sugar−phosphate backbone structure. Since DNA and RNA polymers differ structurally both by the thymine C-5 methyl group and by the 2′-hydroxyl group, however, the substitution of one natural strand type for another involves making two structural changes simultaneously. Without independently measuring effects of these two groups in separate experiments, it is not possible to attribute accurately the observed effects to their structural origins.

For this reason we decided to test separately the effects of these two distinguishing groups. This was done in the context of double and triple helices consisting of homopurine and homopyrimidine strands. This arrangement presents the two groups in approximately equal numbers for direct comparison; in addition, it allows testing of double and triple helices using similar sequences. In the synthesis we used deoxyribo and ribo phosphoramidites derived from the natural DNA and RNA nucleosides, as well as commercially available 2′-deoxyuridine phosphoramidite; we also synthesized a 2′-*O*-silyl-protected ribothymidine phosphoramidite monomer, thus allowing access to all possible strand combinations with and without 2′-OH and C-5 methyl

† This work was supported by the Office of Naval Research. E.T.K. also acknowledges an award from the Young Investigator Program of the Arnold and Mabel Beckman Foundation, a Camille and Henry Dreyfus Teacher-Scholar Award, and an Alfred P. Sloan Foundation Fellowship.

* To whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* March 15, 1995.

0006-2960/95/0434-4125$09.00/0

groups. We tested these structural substitutions in three different contexts, representing the most common hybridization types in practical use. Binding of single-stranded targets by duplex formation, binding of a naturally substituted DNA duplexes by third-strand triplex formation, and binding of purine-rich single strands by triplex formation were all examined in this light. We find that both the methyl and hydroxyl groups have substantial, and sometimes opposing, effects on helix stability.

EXPERIMENTAL PROCEDURES

Synthesis of Ribothymidine-Protected Phosphoramidite. The ribothymidine precursor was purchased from Sigma. The 5′-(dimethoxytrityl)-protected intermediate was synthesized using the published method (Maggio *et al.*, 1991); ^{1}H-NMR and mass spectral analysis agreed with published values.

2′-O-(tert-Butyldimethylsilyl)-5′-(dimethoxytrityl)ribothymidine. The 5′-(dimethoxytrityl) intermediate was silylated selectively at the 2′-position using the method of Ogilvie (Hakimelahi *et al.*, 1982). The product was isolated by silica gel chromatography (45% ethyl acetate−hexane) as a white foam (730 mg, 1.08 mmol) in 83% yield. ^{1}H NMR (DMSO-d_6): δ (ppm) 0.04 (d, 6H, J = 5.6, Si-Me), 0.84 (s, 9H, Si-t-Bu), 1.35 (s, 3H, Me), 3.18−3.24 (m, 2H, $H_{5'}$), 3.72 (s, 6H, OCH_3), 4.01 (m, 1H, $H_{4'}$), 4.05 (m, 1H, $H_{3'}$), 4.28 (t, 1H, $H_{2'}$), 5.12 (d, 1H, J = 5.0, OH), 5.82 (d, 1H, J = 4.4, $H_{1'}$), 6.88−7.39 (m, 13H, H-DMTr), 7.54 (s, 1H, H_6), 11.38 (s, 1H, NH); MS (FAB$^+$) *m/z* (relative intensity): 687 (8), 674 (M^+,56), 659 (8), 617 (46), 597 (20), 567 (30), 468 (6), 454 (30), 438 (26), 409 (23), 379 (19), 355 (100).

2′-O-(tert-Butyldimethylsilyl)-5′-(dimethoxytrityl)ribothymidine 3′-(Cyanoethyl N,N-diisopropylphosphoramidite). The phosphoramidite was synthesized using the general method of Usman (Scaringe *et al.*, 1990). The product was isolated by silica gel chromatography (20% CH_3CN−dichloromethane) as a white foam or yellow oil (350 mg, 0.4 mmol) in 80% yield. ^{1}H NMR (mixed diastereomers in DMSO-d_6): δ 0.00−0.07 (dd, 6H, Si-Me), 0.82−0.84 (ds, 9H, Si-t-Bu), 0.91−1.11 (m, 12H, $CH(CH_3)_2$), 1.38 (ds, 3H, CH_3), 2.52−2.77 (dt, 2H, CH_2CN), 3.52 (m, 2H, $H_{5'}$), 3.72 (s, 6H, OCH_3), 4.15 (m, 1H, $H_{4'}$), 4.21 (m, 1H, $H_{3'}$), 4.42−4.54 (m, 1H, $H_{2'}$), 5.83−5.93 (dd, 1H, $H_{1'}$), 6.87−7.36 (m, 13H, H-DMTr), 7.55 (s, 1H, H_6), 11.43 (s, 1H, NH); MS(FAB$^+$) *m/z* (relative intensity): 875 (M + H$^+$, 18), 774 (9), 749 (11), 571 (100), 555 (7).

Synthetic Methods for Linear and Circular Oligonucleotides. DNA oligomers were synthesized on an Applied Biosystems 392 automated synthesizer using the standard phosphoramidite method (Beaucage & Caruthers, 1991). The dU phosphoramidite was obtained from Glen Research. RNA oligonucleotides were prepared using *tert*-butyldimethylsilyl-protected phosphoramidites (Applied Biosystems) following the oligoribonucleotide synthesis procedure of Usman (Scaringe *et al.*, 1990). Tetrabutylammonium fluoride in THF (Aldrich) was dried over molecular sieves prior to use in the desilylation step (Hogrefe *et al.*, 1993). For synthesis of circular oligomers the linear precursors were 5′-phosphorylated on the synthesizer using a commercially available reagent purchased from Cruachem (Horn & Urdea, 1986). 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 (Borer, 1985).

The four circular oligomers were constructed from 34-nt ribo and deoxyribo linear precursors having the sequence 5′-pTTTCTTCACAC-TTCTTTCTTTTCCACACCTTTTC or 5′-pUUUCUUCACACUUCUUUCUUUUC-CACACCUUUUC. The two naturally substituted cases were described previously (Wang & Kool, 1994). For the two ribo cases a 3′-deoxynucleotide was used to ensure correct 5′−3′ junction substitution in the ligation. The same 12-nt oligomer 5′-dAAGAAAGAAAAG (55 μM) was used as a template for cyclization in all four cases; the 34-nt linear precursors were present at 50 μM concentration. The ligations were carried out in a buffer containing imidazole hydrochloride (200 mM, from a 0.5 M pH 7.0 stock) and $NiCl_2$ (100 mM). Solid BrCN was added with vortex mixing to give a final calculated concentration of 125 mM, and the reactions were allowed to proceed at 25 °C for 12 h. As the reaction proceeded, a light tan precipitate was observed, and previous studies have shown that this precipitate contains the majority of the nucleic acids. The solutions (including solids) were dialyzed against water and lyophilized. The resulting solid was loaded onto a preparative 20% denaturing PAGE gel for separation. The circular products were isolated by excision from the gel after visualization by UV shadowing; the circles migrated at a rate ~0.9 times that of the linear 34-nt precursors.

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. Products were visualized with Stains-all dye (Sigma). All 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 (1:1:1 for termolecular triplexes) molar ratio of linear or circular 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. The PIPES buffer was chosen to be compatible with Mg^{2+} ion in the solution, and because its pK_a has the lowest temperature dependence of the Good buffers (Good *et al.*, 1966). 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 DNA were prepared, they were heated to 90 °C and allowed to cool slowly to room temperature prior to the melting experiments.

The melting studies (see Figure 1 for representative plots) were carried out in Teflon-stoppered 1 cm path length quartz cells under nitrogen atmosphere on a Varian Cary 1 UV−vis spectophotometer equipped with thermoprogrammer. Absorbance (260 nm) was monitored while temperature was raised at a rate of 0.5 °C/min; slower heating rates do not affect the results with this instrument. Melting temperatures (T_m) are determined by computer fit of the first derivative of absorbance with respect to 1/T, and are reported at the weighted maximum of the first derivative (Gralla & Crothers, 1973). Uncertainty in T_m is estimated at ±0.5 °C based on repetitions of experiments.

Free energy values (see Figure 2) for the complexes were obtained from fits of the data using a two-state model with

Table 1: Melting Temperatures (T_m (°C)) and Free Energies ($-\Delta G°_{37}$ (kcal)) for Duplexes with All Possible Combinations of 2'-Hydroxyl and C-5 Methyl Groups

	purine strand			
	5'-d(AAGAAAGAAAAG)		5'-r(AAGAAAGAAAAG)	
pyrimidine strand	T_m (°C)[a,b]	$-\Delta G°_{37}$ (kcal)[b]	T_m (°C)[a,b]	$-\Delta G°_{37}$ (kcal)[b]
3'-d(UUCUUUCUUUUC)	32.4	7.6 (0.8)	37.1	8.7 (0.9)
3'-d(TTCTTTCTTTTC)	38.8	9.2 (0.9)	42.2	10.2 (1.0)
3'-r(UUCUUUCUUUUC)	19.9	4.9 (0.5)	46.5	11.4 (1.1)
3'-r(TTCTTTCTTTTC)	31.4	7.3 (0.7)	55.4	13.9 (1.4)

[a] Conditions: 3.0 μM total strand concentration, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM $MgCl_2$, 10 mM Na·PIPES, pH 7.0. [b] Error limits for individual measurements are estimated at ±0.5 °C in T_m and ±10% in free energy (error limits are given in parentheses).

linear sloping base lines (Petersheim & Turner, 1983). Previous experience with the sequences in this study (both duplexes and bimolecular triplexes) has shown that the free energies obtained by curve fitting agree to within 10−15% with values obtained by van't Hoff analysis ($1/T_m$ vs ln-[C_T]). Free energy values are reported at 37 °C, thus representing minimal extrapolations from the melting data for the duplexes. The free energies for the bimolecular triplexes should be considered somewhat less accurate since the T_m values in those cases are considerably higher. For the termolecular triplexes, the free energies for the third strand transition should be considered significantly less reliable because of the difficulty in fitting the upper base line due to the nearness of the duplex transitions. When the two transitions coincided, only T_m values were obtained. In general, precision in free energies is estimated at ±5−10%, based on repetitions of experiments.

RESULTS AND DISCUSSION

Thermal denaturation experiments were initially carried out at pH 7.0 with combinations of two complementary strands forming a pyrimidine−purine 12-bp duplex. We tested four pyrimidine strands having the same sequence but with (i) no methyls or hydroxyls, (ii) methyls but no hydroxyls, (iii) hydroxyls but no methyls, and (iv) both hydroxyls and methyls; these were hydridized to a complementary purine strand with hydroxyls and without. Results show (Table 1) that there is a large variation in melting temperatures (T_m) and free energies ($-\Delta G°_{37}$) for these duplexes, with ranges of 19−55 °C and −4.9 to −13.9 kcal/mol, respectively. The strongest duplex is the methylated RNA−RNA case, with methyls in the pyrimidine strand and hydroxyls in both strands. The weakest case is the RNA−DNA hybrid where the pyrimidine strand is RNA without methyl groups.

Representative thermal denaturation plots for the three types of complexes show (Figure 1) that the differences are qualitatively large. In the analysis below we chose to compare, for the most part, free energies of the complexes, rather than T_m values. It should be noted, however, that there is a good linear relationship between the T_m values and the free energies (see Figure 2), and so the same qualitative and quantitative results would be obtained by using either parameter.

Separating the Methyl and Hydroxyl Effects. The specific sources of these differences can be established by comparison of cases where only one type of group has been changed. For example, comparison of fully deoxyribose-substituted duplexes involving pyrimidine strands with and without nine methyls shows that these methyl groups are stabilizing, giving a ΔT_m of 6.4 (±0.7) °C and a $\Delta\Delta G°$ (37 °C) of 1.6 (±1.2) kcal/mol (Figure 3). Per methyl group, this is an average stabilization of 0.18 (±0.13) kcal when both strands lack hydroxyls. When the purine strand is substituted with hydroxyls, the methyl effect remains the same, at 0.18 (±0.15) kcal per methyl. When the pyrimidine strand itself contains hydroxyl groups, the methyl effect is somewhat larger, at 0.27 (±0.11) kcal/mol per methyl, and again is the same when the purine strand contains hydroxyls (0.28 ± 0.20 kcal/mol). Taken together, the comparisons show that methyl groups are always stabilizing in these duplexes.

Similar comparisons can be made for the effect of hydroxyls in one or both strands. Surprisingly, hydroxyls can be destabilizing, neutral, or stabilizing, depending on the context (Figure 3 and Table 1). For example, if the target purine strand is RNA, the effect of hydroxyls in the probe pyrimidine strand is 0.22 (±0.12) kcal/mol of stabilization per OH (if unmethylated). If the target strand is DNA, however, hydroxylation of the pyrimidine strand is *destabilizing* by the same amount (0.22 ± 0.09 kcal per OH if unmethylated). If the pyrimidine strand is methylated, hydroxyl substitution still has this opposing effect, with 0.31 (±0.14) kcal of stabilization with an RNA target and 0.16 (±0.09) kcal/mol of destabilization with a DNA target. Thus, when the purine strand is RNA, the methyl and hydroxyl groups in the pyrimidine strand reinforce each other as stabilizing factors; when the purine strand is DNA, the hydroxyls oppose the stabilization effect of the methyls. The preference of a hydroxyl-substituted strand for an RNA over a DNA complement is similar in both cases, at 0.54 (±0.10) and 0.55 (±0.13) kcal per base pair.

The hydroxyl effects can also be measured from the viewpoint of the purine strand (Table 1). We find that there are again large differences depending on whether the complement is DNA or RNA. Adding hydroxyl groups to the purine strand gives the large stabilization of 0.54 (±0.10) kcal/mol per base pair in binding a pyrimidine RNA complement, and it is the same if the complement is methylated (0.55 ± 0.13 kcal per pair). When the complementary pyrimidine strand is DNA, the hydroxyl effect is no longer destabilizing, but is instead relatively neutral, with a stabilization of 0.08 (±0.10) to 0.09 (±0.11) kcal per OH (and still independent of the methylation of the complement). In general, we find that the preference of a hydroxyl-substituted strand for RNA over DNA is 0.54−0.55 (±0.13) kcal per base pair. Presumably, this effect is due to the different conformational preferences that arise from 2'-OH substitution of the sugar. The hydroxylated strand preference for hybridization to another hydroxylated strand is independent of purine or pyrimidine base substitution, and also independent of the methyl effect.

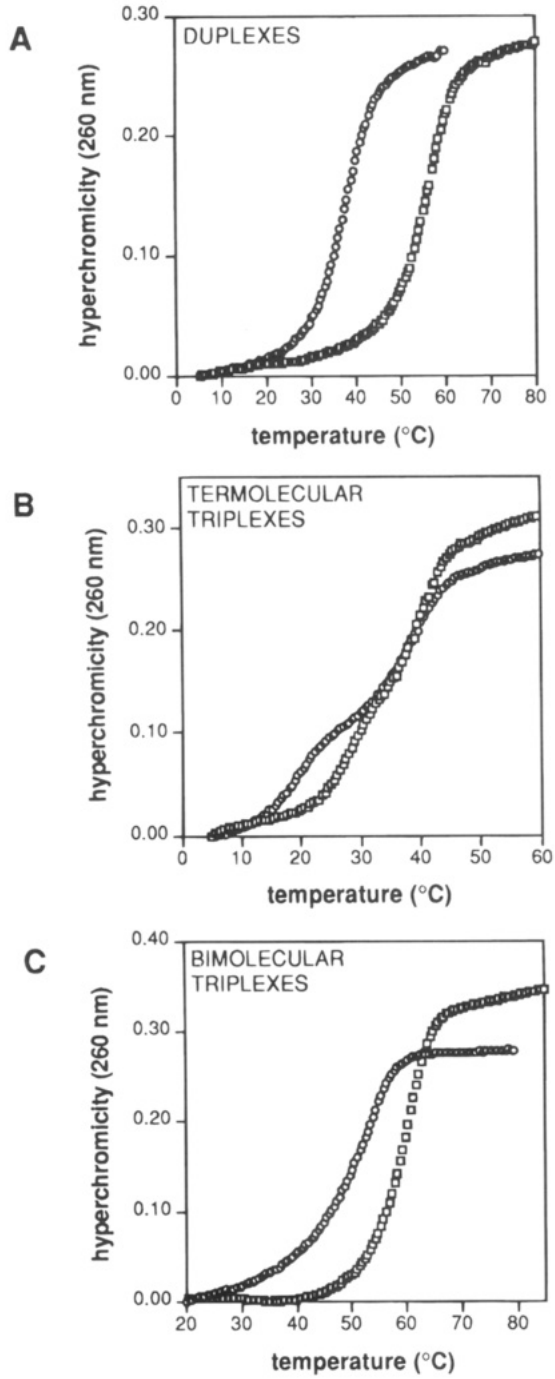

FIGURE 1: Representative thermal denaturation curves (monitored at 260 nm) for the three types of complexes in this study: (A) duplexes at pH 7.0: dCUUUUCUUUCUU + rAAGAAAGAAAAG (○); rCTTTTCTTTCTT + rAAGAAAGAAAAG (□). (B) termolecular triplexes at pH 5.5: dUUCUUUCUUUUC + [dAAGAAAGAAAAG·dCTTTTCTTTCTT] (○); dTTCTTTCTTTTC + [dAAGAAAGAAAAG·dCTTTTCTTTCTT] (□); (C) bimolecular triplexes at pH 7.0: ribo(U,C) circle + rAAGAAAGAAAAG (m); ribo(T,C) circle + rAAGAAAGAAAAG (○). Curves are normalized by the fraction increase (hyperchromicity) in A_{260} with increasing temperature. See Tables 1−3 for detailed experimental conditions.

We also can analyze the preferences of nonhydroxylated (DNA) strands by making similar comparisons. A pyrimidine DNA strand prefers hybridization to a hydroxylated (RNA) strand over a DNA complement by 0.08 (±0.10) to 0.09 (±0.11) kcal per base pair; this is again independent of methylation of the pyrimidines. Interestingly, a purine DNA strand reverses this selectivity, preferring hybridization

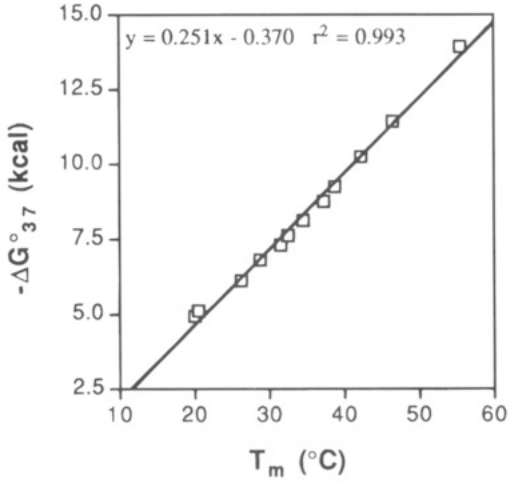

FIGURE 2: Correlation of free energy and melting temperature data from this study. The good linear relationship indicates that either set of data would produce the same general conclusions. Data are taken from Table 1.

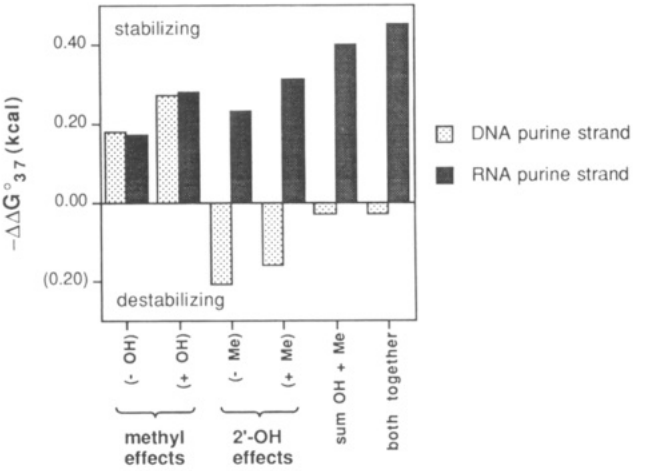

FIGURE 3: Effects of substituting nine C-5 methyl or twelve 2′-OH groups on a pyrimidine probe strand in the duplex-type binding of a purine DNA or RNA complement. Shown are average free energy increments ($-\Delta\Delta G°_{37}$) for adding a single substituent to the complex 5′-CUUUUCUUUCUU·5′-AAGAAAGAAAAG. The last two entries are the mathematical sum of the two individual effects and the experimental effect of making both changes together, given as indicators of the independence of the two types of substitution. Free energies for individual duplexes were measured from melting curves carried out at a total strand concentration of 3.0 μM in a buffer containing 100 mM NaCl, 10 mM $MgCl_2$, and 10 mM Na·PIPES (piperazinebis(ethanesulfonate)), pH 7.0.

instead to an unhydroxylated (DNA) complement, with a selectivity ranging from 0.23 (±0.08) kcal per base pair (without methyls) to 0.16 (±0.10) kcal per pair (with methyls).

We also find that in naturally-substituted RNA/DNA pyrimidine−purine duplexes there is a significant preference for the purines to be in the RNA strand, in agreement with previous studies (Martin & Tinoco, 1986; Roberts & Crothers, 1992; Ratmeyer *et al.*, 1994). Comparing a (DNA pyrimidine + RNA purine) duplex to a (RNA pyrimidine + DNA purine) duplex, we find a preference for the first case amounting to 0.44 (±0.09) kcal/mol per base pair. Once again, however, two structural changes have been made in that comparison. When we isolate the hydroxyl and methyl effects, we find that contributions from both groups make up this 0.44 kcal preference. When methylated pyrimidines

are present in both cases, the purine/RNA preference is 0.24 (±0.10) kcal per base pair; when pyrimidines are unmethylated, the preference is 0.32 (±0.09) kcal. This 0.24−0.32 kcal value can be considered the pure preference for purines to be on a hydroxylated (RNA) strand. Thus, the effect of the methyl groups is the remainder, or 0.12−0.20 kcal/mol per base pair, an effect per group that is nearly as large as the hydroxyl effect. This effect arises from the fact that only DNA has methylated pyrimidines, and thus there is a significant energetic preference for a pyrimidine-rich strand to contain DNA bases.

Similar experiments in the duplex series (not shown) were also carried out at pH 5.5. Values for T_m and $\Delta G°$ are the same, within experimental error, except in two cases: when the two ribose-substituted pyrimidine strands are hybridized to a DNA purine target, the binding is pH-dependent, suggesting Pyr·Pur·Pyr triplex formation. Although separate confirming evidence is not yet available, a possible explanation for this observation in these two cases is that the lowered pH drives disproportionation of the duplex to triplex + single strand. The triplex would in those two cases involve third-strand binding with only nine bases, since it is not fully complementary. The fact that none of the other complexes is pH dependent argues strongly for duplex, and not triplex, structure in those cases.

Correlation with Prior Findings. These results are useful in explaining data in recent published RNA/DNA comparisons (Martin & Tinoco, 1980; Hall & McLaughlin, 1991; Roberts & Crothers, 1992; Han & Dervan, 1993; Escudé *et al.*, 1993; Ratmeyer *et al.*, 1994; Riley *et al.*, 1966). Three recent studies concerning the relative stabilities of RNA (R), DNA (D), and mixed duplexes (Martin & Tinoco, 1986; Roberts & Crothers, 1992; Ratmeyer *et al.*, 1994) have observed that in homopyrimidine−homopurine sequences the relative order of stability is RR > DR > DD > RD. The present data (also using homopyrimidine−homopurine sequences) are in complete agreement with those studies: with naturally substituted strands we find the order RR > DR > DD > RD. Separation of the methyl and hydroxyl effects allows a structural correlation with this relative ordering. It can be concluded that there are two reasons why natural DR Pyr·Pur duplexes are more stable than reversed (RD) cases: first, the methyl effect favors pyrimidine DNA strands. Second, the hydroxyl effect favors purine RNA strands. Duplexes of the RR type are more stable than the DD type because the preference of a hydroxylated strand for another overshadows the methyl effect.

Two studies using mixed pyrimidine−purine sequences (Hall & McLaughlin, 1991; Ratmeyer *et al.*, 1994) have found a different stability order, namely: RR > DD > DR, RD. Our interpretation of this difference is as follows: in mixed sequences the hydroxyl-purine preference will tend to average out, since purines will be placed on both strands. In those cases the dominant trend will be the 0.43−0.47 kcal per base pair preference of a hydroxylated strand for another (discussed above), thus explaining the observed RR > DD ordering in mixed sequences. The greater stability of the DD case over the DR, RD cases may be explained by the methyl effect: in the DD case all possible methyls will be present, while in the DR, RD cases only half (on average) will be. For sequence-averaged duplexes where a T(U) occurs in half the base pairs, this effect might be considerable; based on our measured ~0.3 kcal per methyl effect,

A

purine strand	5'-d(AAAAAAAAAAAA)		
pyrimidine strand	T_m (°C)	$-\Delta G°_{37}$ (kcal)	$-\Delta\Delta G°$ per methyl
5'-d(UUUUUUUUUUUU)	20.4	5.1 (0.5)	--
5'-d(UTUTUTUTUTUT)	26.1	6.1 (0.6)	0.17 (0.13)
5'-d(UUUUUUTTTTTT)	28.6	6.8 (0.7)	0.28 (0.14)
5'-d(TTTTTTTTTTTT)	34.4	8.1 (0.8)	0.25 (0.08)

B

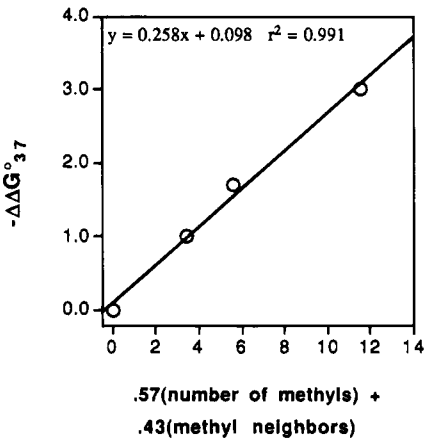

FIGURE 4: The effects of substituting C-5 methyl groups on the dU_{12}·dA_{12} duplex. (A) Melting temperatures (T_m) and free energies ($-\Delta G°_{37}$) for duplexes containing zero, six, or twelve methyl groups, and the average free energy increment ($-\Delta\Delta G°_{37}$) per methyl group. (B) Linear correlation of two factors—the number of methyls and the number of methyl−methyl nearest neighbors—with the corresponding total change in free energy for the duplex. The best fit gives coefficients of 0.57 and 0.43, respectively, for the relative contributions of the two factors.

the difference between DD and (DR, RD) is then predicted to be an average of 0.3 kcal every four base pairs. Indeed, early studies using ΦX174 sequences also found the relative ordering RR > DD > DR, RD (Chamberlin & Berg, 1964).

The Origin of the Methyl Effect. The reason for stabilization of nucleic acid helices by pyrimidine methylation, previously observed in polymer studies (Barszcz & Shugar, 1968), has been explained variously as an increase in base stacking proficiency due to increased polarizability (Sowers *et al.*, 1987) or as a hydrophobic effect of favorable contacts between adjacent methyl groups (Han & Dervan, 1993; Escudé *et al.*, 1993). We carried out a simple experiment to test these two possibilities: we constructed four deoxyribose dodecamers complementary to the sequence dA_{12} (Figure 4). One oligomer has no methyl groups, and one has twelve; the other two have six methyl groups each, but in one the methyls are adjacent, while in the second they are placed at alternating positions. Comparison of the unmethylated and fully methylated cases again shows a methyl stabilization which averages 0.25 (±0.08) kcal per methyl group. Examination of the two half-methylated cases shows that even when the methyls are nonadjacent, a significant effect is still seen, with a stabilization of 0.17 kcal per methyl. Thus, it is clear that hydrophobic methyl−methyl contacts cannot explain all of the methyl stabilization effect. There does appear to be a small preference for adjacent placement of the six methylated bases. A linear

Table 2: Melting Temperatures (T_m (°C)) and Free Energies ($-\Delta G°_{37}$ (kcal)) for Third-Strand Binding to an Unmodified DNA Duplex at pH 5.5[a]

	pyr•pur duplex		5'-d(AAGAAAGAAAAG) 3'-d(TTCTTTCTTTTC)	
Pyr third strand	T_m (°C)[b,c]	$-\Delta G°_{37}$ (kcal)[c]	methyl effect	2'-OH effect
5'-d(UUCUUUCUUUUC)	18.5	2.7	–	–
			8.9 °C	
5'-d(TTCTTTCTTTTC)	27.4	4.5	-1.8 ± 0.6 kcal	–
5'-r(UUCUUUCUUUUC)	38.3[d]	–[d]	–	19.8 °C
5'-r(TTCTTTCTTTTC)	45.7[d]	–[d]	7.4 °C	18.3 °C

[a] The melting temperature and free energy differences are also separated into methyl and 2'-OH effects. [b] Conditions: 4.5 μM NaCl, 10 mM $MgCl_2$, 10 mM Na•PIPES, pH 5.5. [c] Error limits for individual measurements are estimated at ±0.5 °C in T_m and ±10% in free energy. [d] Third-strand transition coincides with duplex transition.

correlation of free energy versus these two factors suggests a 57:43 (number of methyls:methyl−methyl neighbors) weighting as optimum (Figure 4B).

Thus, the present data provide support for the hypothesis that the methyl group alone increases base stacking proficiency and that this is the major contributing effect. The data do indicate that adjacent interactions between methylated bases may also play a smaller role, although this increased interaction is not likely to be due to hydrophobic effects, since methylation actually *increases* the water solubility of uracil (Sowers *et al.*, 1987). The simplest explanation for both effects is that the methyl increases polarizability, allowing for more favorable van der Waals interactions with neighboring bases (Sowers *et al.*, 1987); when the neighbor is also methylated, the induced dipole−induced dipole attractions would become even larger. A favorable direct methyl−methyl contact may also be used to explain the small adjacent effect; the present data alone cannot distinguish the two explanations.

Effects in Termolecular Triple Helices. We also specifically tested the methyl and hydroxyl effects in Pyr•Pur•Pyr triple helical complexes. Two types of triplexes were studied, as models of the two most common triple helical hybridizations reported in the literature. First we studied the effects of methyl and hydroxyl groups in the binding of a pyrimidine third strand to an unmodified purine−pyrimidine DNA duplex (Moser & Dervan, 1987; LeDoan *et al.*, 1987) at pH 5.5. Table 2 lists the results of the four experiments; reported are the T_m and $-\Delta G°_{37}$ values for the triplex−duplex transition, and T_m values for the triplex−single strand transition in cooperative cases.

Results show that both the methyl and hydroxyl group substitutions have a positive effect on binding. The effect of the nine methyl groups is a total 1.8 kcal (or ~7−9 °C in T_m) of stabilization. The effect of the twelve hydroxyl groups is a larger 18−20 °C of thermal stabilization (depending on methyl substitution). Thus, when binding to a purine-rich site in duplex DNA, a methylated pyrimidine strand with ribose sugars can have a large advantage over unmethylated and unhydroxylated versions.

Previous studies on this type of triplex have concluded that RNA third strands have a binding advantage over DNA third strands (Roberts & Crothers, 1992; Han & Dervan, 1993; Escudé *et al.*, 1993; Ratmeyer *et al.*, 1994; Skoog & Maher, 1993). The current experiments allow us to conclude that the hydroxyl groups on the RNA strand greatly favor the RNA binding to a DNA duplex; relative to a DNA third strand binding, the thymine methyl effect, although also stabilizing, is overwhelmed by the ~3-fold larger hydroxyl effect. Methylation of cytosines in third strands has been well-documented to stabilize binding (Lee *et al.*, 1984; Povsic & Dervan, 1990; Xodo *et al.*, 1991); this has been attributed to increased pK_a of this base (Escudé *et al.*, 1993; Xodo *et al.*, 1991). While basicity differences may well be a partial factor, we believe that the cytosine methylation is also likely to have a non-pK_a-related stabilizing effect, possibly due to more favorable base stacking, as seen for uracil in this study. The present results suggest that, in third-strand binding, a general recommendation can be made for use of RNA backbones with C-5-methylated uracils as well as cytosines.

Effects in Bimolecular Triplexes. Pyr•Pur•Pyr triplexes are also involved in a different three-stranded structure, in which purine-rich target strands are bound by a single molecule containing two pyrimidine domains, one a Watson−Crick complement and one a Hoogsteen complement. This mode of binding ("sandwiching" or "clamping" the purine target site between the two binding domains of a probe) is receiving increasing attention because of the high binding affinities and sequence selectivities achieved (Xodo *et al.*, 1990; Kool, 1991; Prakash & Kool, 1991; Giovannangeli et al., 1991, 1993; Prakash & Kool, 1992; D'Souza & Kool, 1992; Rumney & Kool, 1992; Salunkhe *et al.*, 1992; Rubin *et al.*, 1993; Gryaznov & Lloyd, 1993). Another advantage of this system is the fact that binding is often cooperative, with thermal dissociation directly from triplex to unbound components; this allows measurement of methyl and hydroxyl effects simultaneously in all parts of a triplex, without the need to separate Hoogsteen and Watson−Crick interactions. We tested the substituent effects by constructing four circular pyrimidine-rich oligonucleotides of identical sequence but with varying degrees of methylation and hydroxylation (Table 3). These were used to bind to complementary 12-base purine DNA or RNA strands, and experiments were carried out at two pH values (pH 7.0 and 5.5). The loops bridging the binding domains in all four cases are the same, having sequence CACAC, without methyls or 2'-hydroxyls. For all eight complexes studied, single, apparently two-state, transitions were observed (see Figure 1C).

Results show (Table 3, Figure 5) that methyl and hydroxyl effects can be of similar magnitude in this triplex type, and the effects are essentially independent of one another. Both types of substitutions are stabilizing; interestingly, the methyl effect is the dominant effect when the target is a DNA purine strand, but when the target is RNA, the hydroxyl effect is larger. For example, at pH 5.5 a normally methylated DNA circle binds a DNA complement with a substantial 8.2 (±2.5) kcal/mol higher affinity than does an unmethylated circle; averaged over 18 methyls, this is 0.45 (±0.14) kcal/mol of stabilization per methyl group (Figure 5). With the circular RNA probes the methyl effect is almost identical, with an −8.0 (±2.9) kcal total stabilization. The hydroxyl effect is much smaller by comparison: for example, the advantage in binding of a DNA target by a hydroxylated (RNA) circle relative to an unhydroxylated (unmethylated DNA) circle is only 0.15 (±0.09) kcal per hydroxyl group; thus the methyl effect in this case is 3-fold larger than the backbone effect.

With an RNA purine-rich target, however, results are different: the hydroxyl effect becomes much larger, while the methyl effect becomes smaller. The effect of adding 23

Table 3: Melting Temperatures (T_m) and Free Energies ($-\Delta G°_{37}$) for Bimolecular Triplexes with Varied Combinations of 2′-Hydroxyl and C-5 Methyl Groups at pH 5.5 and 7.0[a]

		purine strand	5′- d(AAGAAAGAAAAG)		5′- r(AAGAAAGAAAAG)	
pyrimidine strand		pH	T_m (°C)[b,c]	$-\Delta G°_{37}$ (kcal)[c]	T_m (°C)[b,c]	$-\Delta G°_{37}$ (kcal)[c]
A CUUCUUUCUUUUCC A C deoxy- C A CUUCUUUCUUUUCC A		7.0	43.7	10.5 (1.1)	38.7	9.1 (0.9)
		5.5	54.3	13.0 (1.3)	37.0	8.9 (0.9)
A CTTCTTTCTTTTCC A C deoxy- C A CTTCTTTCTTTTCC A		7.0	54.5	14.5 (1.5)	42.8	10.4 (1.0)
		5.5	68.8	21.2 (2.1)	40.5	10.1 (1.0)
A CUUCUUUCUUUUCC A C ribo- C A CUUCUUUCUUUUCC A		7.0	48.5	11.8 (1.2)	51.2	12.6 (1.3)
		5.5	62.3	16.4 (1.6)	62.9	17.7 (1.8)
A CTTCTTTCTTTTCC A C ribo- C A CTTCTTTCTTTTCC A		7.0	59.9	16.0 (1.6)	59.2	16.5 (1.7)
		5.5	73.4	24.4 (2.4)	69.2	19.5 (2.0)

[a] Underlined C residues in ribo-circle binding domains lack a 2′-OH for simplification of synthesis; CACAC loop sequences in all four compounds are composed of 2′-deoxy residues. [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; for free energy, error limits are given in parentheses.

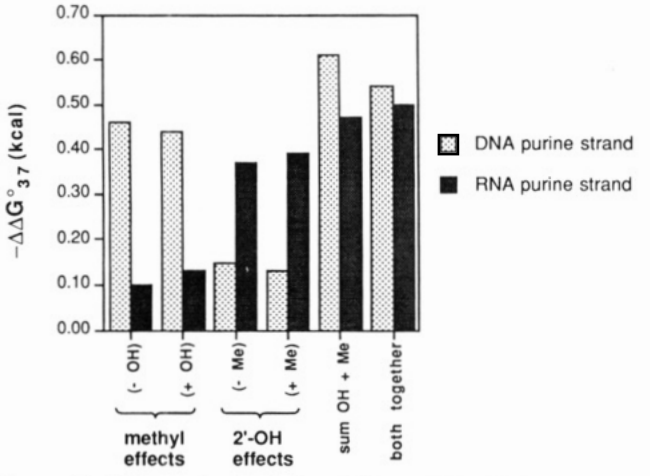

FIGURE 5: Effects of substituting eighteen C-5 methyl or twenty-three 2′-OH groups on a circular pyrimidine-rich probe in the triplex binding of a purine DNA or RNA complement at pH 5.5 and 7.0. The sequences are given in Table 3. Shown are average free energy increments ($-\Delta\Delta G°_{37}$) for adding a single substituent. The last two entries are the mathematical sum of the two individual effects and the experimental effect of making both changes together. Free energies for individual complexes were measured from melting curves carried out at a total strand concentration of 3.0 μM in a buffer containing 100 mM NaCl, 10 mM $MgCl_2$, and 10 mM Na·PIPES (piperazinebis(ethanesulfonate)) buffer.

hydroxyls to a circular ligand is an increase of 8.8 (±2.0) to 9.4 (±2.2) kcal of stabilization to the complexes at pH 5.5. By contrast, the stabilizing effect of methylation is small, 1.2 (±1.3) to 1.8 (±2.7) kcal altogether (0.10 ± 0.07 to 0.13 ± 0.15 kcal per methyl). The average effect of a hydroxyl group is 0.38 (±0.09) to 0.41 (±0.10) per substitution; this amounts to a ~3-fold larger effect in the opposite direction from the methyl dominance seen with a DNA target strand. It seems likely that at least part of the reason for the lower methyl effect is due to differences in structure and/or conformation in the complexes, limiting the magnitude of base stacking effects. For example, the complex between a DNA circle and an RNA target is likely not three-stranded (see below); since the Hoogsteen domain is not bound, its methyls probably cannot contribute to stabilization.

Overall, the lack of pH dependence for DNA circles binding RNA strands confirms previous observations that nonhydroxylated circles bind RNA strands without triplex-type hybridization (Wang & Kool, 1994); instead, only the Watson−Crick-complementary domain binds, leaving the other domain unbound. This stresses the importance of the 2′-OH in this type of triplex and is also consistent with prior observations that DNA−RNA−DNA triplexes are highly unstable and in fact do not form under most conditions (Roberts & Crothers, 1992; Han & Dervan, 1993; Escudé *et al.*, 1993). The present results confirm that the source of this effect in the Hoogsteen DNA strand is the lack of 2′-hydroxyls, rather than the presence of methyl groups. The results also reiterate the importance of the presence or absence of purine strand 2′-OH groups on helix structure and conformation.

Conclusions. Several general conclusions can be reached from these results with double and triple helical complexes. First, the structural sources of stability differences between complexes involving natural DNA or RNA are 2-fold, with both the C-5 methyl of thymine in DNA and the 2′-OH in RNA playing significant roles. Second, the effects of these two substitutions are largely or completely independent of each other. In purine-rich/pyrimidine-rich duplexes, the effect of methyl group substitution is always stabilizing, while the hydroxyl group can be stabilizing or destabilizing, depending on the case. In triple helices, both methyl and hydroxyl groups are stabilizing; which one of these plays the dominant role depends on whether DNA or RNA is involved and what type of triplex is being formed. In third-strand binding to duplex DNA, the tightest-binding pyrimidine strand will contain the maximum number of both 2′-OH and C-5 methyl groups; in this case, the hydroxyl contributes 2−3 times as much to the stability as the methyl group. In binding of purine single strands by triplex formation, the methyl effect is stronger than the hydroxyl effect when DNA is the target, but the reverse is true when RNA is the target. In sum, the results are helpful in explaining differences seen for naturally substituted DNAs and RNAs and suggest further that unnatural residues such

as 5-methyluridine can be useful in stabilizing complexes with designed nucleic acid probes.

REFERENCES

Barszcz, D., & Shugar, D. (1968) *Eur. J. Biochem. 5*, 91−100.

Beaucage, S. L., & Caruthers, M. H. (1981) *Tetrahedron Lett. 22*, 1859−1862.

Borer, P. N. (1985) in *Handbook of Biochemistry and Molecular Biology* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 589, CRC Press, Cleveland.

Chamberlin, M., & Berg, P. (1964) *J. Mol. Biol. 8*, 297−313.

D'Souza, D. J., & Kool, E. T. (1992) *J. Biomol. Struct. Dyn. 10*, 141−152.

Escudé, C., François, J.-C., Sun, J., Ott, G., Sprinzl, M., Garestier, T., & Hélène, C. (1993) *Nucleic Acids Res. 21*, 5547−5553.

Giovannangeli, C., Montenay-Garestier, T., Rougée, M., Chassignol, M., Thuong, N. T., & Hélène, C. (1991) *J. Am. Chem. Soc. 113*, 7775−7776.

Giovannangeli, C., Thuong, N. T., & Hélène, C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10013−10017.

Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S., & Singh, R. M. M. (1966) *Biochemistry 5*, 467−477.

Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.

Gryaznov, S. M., & Lloyd, D. H. (1993) *Nucleic Acids Res. 21*, 5909−5915.

Hakimelahi, G. H., Proba, Z. A., & Ogilvie, K. K. (1982) *Can. J. Chem. 60*, 1106−1113.

Hall, K. B., & McLaughlin, L. W. (1991) *Biochemistry 30*, 10606−10613.

Han, H., & Dervan, P. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3806−3810.

Hogrefe, R. I., McCaffrey, A. P., Borozdina, L. U., McCampbell, E. S., & Vaghefi, M. M. (1993) *Nucleic Acids Res. 21*, 4739−4741.

Horn, T., & Urdea, M. (1986) *Tetrahedron Lett. 27*, 4705−4708.

Kool, E. T. (1991) *J. Am. Chem. Soc. 113*, 6265−6266.

LeDoan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J. L., Thuong, N. T., Lhomme, J., & Hélène, C. (1987) *Nucleic Acids Res. 15*, 7749−7760.

Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., & Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603−6614.

Maggio, A., Boyer, V., Aubertin, A., Obert, G., Kirn, A., & Imbach, J.-L. (1991) *Nucleosides Nucleotides 10*, 1431−1449.

Martin, F. H., & Tinoco, I. Jr., (1980) *Nucleic Acids Res. 8*, 2295−2299.

Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645−650.

Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 256−263.

Povsic, T. J., & Dervan, P. B. (1990) *J. Am. Chem. Soc. 111*, 3059−3061.

Prakash, G., & Kool, E. T. (1991) *J. Chem. Soc., Chem. Commun.*, 1161−1162; *ibid.* (1994) 676.

Prakash, G., & Kool, E. T. (1992) *J. Am. Chem. Soc. 114*, 3523−3528.

Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G., & Wilson, W. D. (1994) *Biochemistry 33*, 5298−5304.

Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol. 20*, 359−389.

Roberts, R. W., & Crothers, D. M. (1992) *Science 258*, 1463−1466.

Rubin, E., McKee, T. L., & Kool, E. T. (1993) *J. Am. Chem. Soc. 115*, 360−361.

Rumney, S., & Kool, E. T. (1992) *Angew. Chem. 104*, 1686−1689; (1992) *Angew. Chem., Int. Ed. Engl. 31*, 1617−1619.

Salunkhe, M., Wu, T., & Letsinger, R. L. (1992) *J. Am. Chem. Soc. 114*, 8768−8772.

Scaringe, S. A., Francklyn, C., & Usman, N. (1990) *Nucleic Acids Res. 18*, 5433-5441.

Skoog, J. U., & Maher, L. J., III, (1993) *Nucleic Acids Res. 21*, 2131−2138.

Sowers, L. C., Shaw, B. S., & Sedwick, W. D. (1987) *Biochem. Biophys. Res. Commun. 148*, 790−794.

Thuong, N. T., & Hélène, C. (1991) *J. Am. Chem. Soc. 113*, 7775−7776.

Wang, S., & Kool, E. T. (1994) *Nucleic Acids Res. 22*, 2326−2333.

Xodo, L. E., Manzini, G., & Quadrifoglio, F. (1990) *Nucleic Acids Res. 18*, 3557−3564.

Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., & van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625−5631.

BI942437N