4644–4649 *Nucleic Acids Research, 1997, Vol. 25, No. 22*

Relative stability of triplexes containing different numbers of T·AT and C$^+$·GC triplets

Melanie D. Keppler and Keith R. Fox*

Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, UK

Received July 7, 1997; Revised and Accepted September 22, 1997

ABSTRACT

We have used DNase I footprinting to compare the stability of parallel triple helices containing different numbers of T·AT and C$^+$·GC triplets. We have targeted a fragment containing the 17mer sequence 5′-AGGAA-GAGAAAAAAGAA with the 9mer oligonucleotides 5′-TCCTTCTCT, 5′-TTCTCTTTT and 5′-TTTTTTCTT, which form triplexes at the 5′-end, centre and 3′-end of the target site respectively. Quantitative DNase I foot-printing has shown that at pH 5.0 the dissociation constants of these oligonucleotides are 0.13, 4.7 and >30 μM respectively, revealing that increasing the proportion of C$^+$·GC triplets increases triplex stability. The results suggest that the positive charge on the protonated cytosine contributes to triplex stability, either by a favourable interaction with the stacked π system or by screening the charge on the phosphate groups. In the presence of a naphthylquinoline triplex binding ligand all three oligonucleotides bind with similar affinities. At pH 6.0 these triplexes only form in the presence of the triplex binding ligand, while at pH 7.5 footprints are only seen with the oligonucleotide which generates the fewest number of C$^+$·GC triplets (TTTTTTCTT) in the presence of the ligand.

INTRODUCTION

Triple-stranded DNA is formed when an oligonucleotide binds in the major groove of duplex DNA (1,2). Binding of the third strand to the duplex is sequence specific, enabling a very versatile approach to sequence-specific recognition of duplex DNA. Triplexes are stabilized by formation of hydrogen bonds between the third strand bases and substituents on purine bases of the duplex. Two types of triple helix have been described which vary according to the composition and orientation of the third strand. Pyrimidine-rich third strands bind parallel to the purine strand of the duplex and are characterized by formation of T·AT and C$^+$·GC triplets (3–5). In contrast, purine-rich third strands bind in an antiparallel orientation, generating G·GC, A·AT and T·AT triplets (6–8). Both triplex motifs are stabilized by divalent metal ions, particularly magnesium and manganese.

Formation of the C$^+$·GC triplet requires conditions of low pH (<6.0), necessary for protonation of the third strand cytosine. The free base has a pK of ~4.5, though this may be elevated on triplex formation. Blocks of contiguous C$^+$·GC triplets are particularly unstable (9), presumably as a result of repulsion between the adjacent positive charges. In an attempt to overcome this limitation several cytosine analogues have been synthesized. 5-Methylcytosine has a slightly higher pK value than cytosine (10–13) and triplexes containing this base are more stable at higher pH, but are still not formed under physiological conditions. Indeed, the increased stability of ^{5Me}C$^+$·GC may result from the extra spine of methyl groups within the DNA major groove (13). 6-Oxocytosine (14,15) and pseudoisocytosine (16) also have significantly higher pK values. Another promising cytosine analogue is 2-aminopyridine, which has a pK closer to physiological pH (17). A different strategy uses purine analogues such as 8-oxoadenine (18,19), N7-G (20,21) or P1 (22) for recognition of GC. Some of these analogues retain the positive charge on the base, including 2-aminopyridine and 5-methylcytosine, while others are uncharged species, such as pseudoisocytosine and 6-oxocytosine.

Since T·AT and C$^+$·GC triplets are isohelical (23) and in both cases the third strand base is held in place by two hydrogen bonds, differences in the relative stability of these two triplets at pH 5.0 must reflect the effect of the positive charge on the C$^+$·GC triplet. It has been shown that for intramolecular triplexes the T_m of the triplex to duplex transition increases with increasing C$^+$·GC content, suggesting that the free energy change for formation of C$^+$·GC is greater than for T·AT (24). In this paper we have examined the relative contributions of C$^+$·GC and T·AT to triplex stability by comparing the binding of three 9mer oligonucleotides of different base composition to different regions of a 17 base oligopurine tract, as shown in Figure 1b.

Another method for enhancing triplex stability is to use ligands which selectively bind to triplex, but not duplex, DNA (25–33). In general these compounds bind selectively to T·AT rather than C$^+$·GC triplets (28,32) and, although they have pronounced effects on triplex stability, they do not remove the requirement for conditions of low pH. We have therefore examined the stability of the triplexes shown in Figure 1b in the presence of a naphthylquinoline triplex binding ligand (29–32).

*To whom correspondence should be addressed. Tel: +44 1703 594374; Fax: +44 1703 594459; Email: krfi@soton.ac.uk

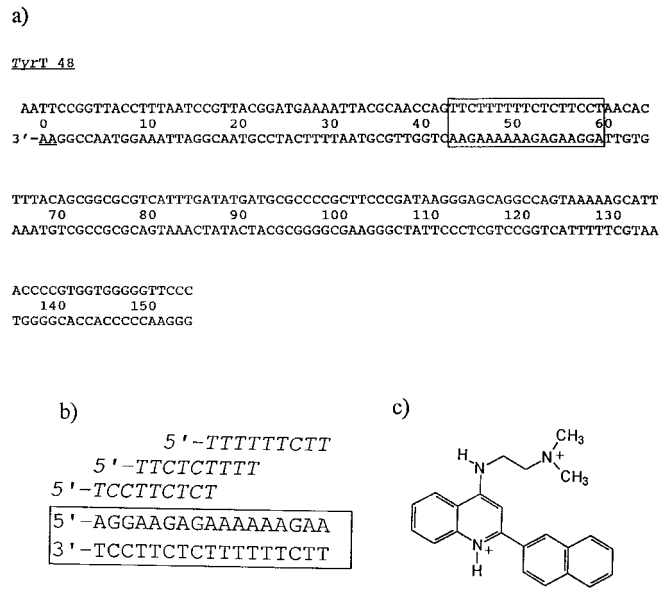

Figure 1. (**a**) Sequence of *tyr*T48. The numbers correspond to those used in the original *tyr*T sequence (34). The box indicates the 17 base oligopurine tract. The bases bearing the radiolabel are underlined. (**b**) Sequence of the 17 base purine tract together with the three 9mer oligonucleotides designed to form triplexes with different regions of the target site. (**c**) Structure of the naphthylquinoline triplex binding ligand.

MATERIALS AND METHODS

Chemicals and enzymes

Oligonucleotides were purchased from Oswel DNA Service and were stored at –20°C in water. DNase I was purchased from Sigma and stored at –20°C at a concentration of 7200 U/ml. Reverse transcriptase was purchased from Promega; restriction enzymes were purchased from Promega, Pharmacia or New England Biolabs. The naphthylquinoline triplex binding ligand (Fig. 1c) was a gift from Dr L.Strekowski (Department of Chemistry, Georgia State University, Atlanta, GA). This was stored as a 20 mM stock solution in dimethylsulphoxide at –20°C and diluted to working concentrations in the appropriate buffer prior to use.

DNA fragments

The sequence of the 160 bp *tyr*T fragment (34) was modified by PCR-mediated site-directed mutagenesis, introducing a 17 base homopurine tract between positions 43 and 59. The mutated fragment was cloned between the *Eco*RI and *Ava*I sites of pUC18. The full details of preparation of this plasmid, which was a gift from Dr Philip Brown, will be published elsewhere. The 160 bp *tyr*T48 fragment (Fig. 1a) was obtained by digesting the plasmid with *Eco*RI and *Sma*I and was labelled at the 3′-end of the *Eco*RI site with [α-^{32}P]dATP using reverse transcriptase. The labelled fragment of interest was separated from the remainder of the plasmid DNA on a 6% non-denaturing polyacrylamide gel, eluted and dissolved in 10 mM Tris–HCl, pH 7.5, containing 0.1 mM EDTA at a concentration of ~10 c.p.s/µl (~10 nM). This procedure labelled the purine-containing strand of the triplex target site.

DNase I footprinting

Radiolabelled DNA (1.5 µl) was mixed with 1.5 µl oligonucleotide, dissolved in an appropriate buffer and 1.5 µl triplex binding ligand or buffer. This mixture was left to equilibrate at 20°C for at least 1 h. Experiments at pH 5.0 and 6.0 were performed in 50 mM sodium acetate containing 10 mM $MgCl_2$; for pH 7.5 the buffer used was 10 mM Tris–HCl containing 50 mM NaCl and 10 mM $MgCl_2$. The mixture was then digested by adding 2 µl DNase I, diluted in 20 mM NaCl containing 2 mM $MgCl_2$ and 2 mM $MnCl_2$ and stopped after 1 min by adding 3.5 µl of a solution containing 80% formamide, 10 mM EDTA and 0.1% (w/v) bromophenol blue. Samples were boiled for 3 min before electrophoresis.

Gel electrophoresis

The products of DNase I digestion were resolved on 40 cm long, 0.3 mm thick, 9% polyacrylamide gels containing 8 M urea, which were run at 1500 V for ~2 h. The gels were then fixed in 10% (v/v) acetic acid, dried under vacuum at 80°C and exposed for autoradiography at –70°C using an intensifying screen. Bands in the digest were assigned by comparison with Maxam–Gilbert markers for guanine and adenine.

Quantitative analysis

Autoradiographs of DNase I digestion patterns were scanned using a Hoefer GS365 microdensitometer. For the analysis we chose a band in each site which was well resolved and cut well in the control (ApG, position 53). The intensity of each band was estimated using the manufacturers software and normalized with respect to the intensity of two bands outside the target site (positions 38 and 61 or 69). It should be noted that under all conditions the concentration of the third strand oligonucleotide is much greater than that of the DNA target. As a consequence, the amount of bound oligonucleotide will be determined by the equilibrium dissociation constant, rather than the stoichiometric ratio of third strand to target. Footprinting plots (35) were constructed from these data and C_{50} values, indicating the oligonucleotide concentration which reduced the band intensity by 50%, were derived by fitting a simple binding curve to plots of band intensity against oligonucleotide concentration using FigP for Windows (Biosoft). These were fitted to the equation $I_c = I_0[C_{50}/(L + C_{50})]$ where I_c is the band intensity in the presence of the ligand, I_0 is the band intensity in the control and L is the oligonucleotide concentration.

RESULTS

*tyr*T48 contains the 17 base oligopurine tract 5′-AGGAAGAGAA-AAAAGAA between positions 43 and 59 (Fig. 1a). We have targeted different regions of this sequence with the oligonucleotides 5′-TCCTTCTCT, 5′-TTCTCTTTT and 5′-TTTTTTCTT, which interact with the 5′-end, centre and 3′-end of this tract. These generate 9 base triplets containing four, two and one C^+·GC triplets respectively, as shown in Figure 1b. The results of DNase I footprinting experiments performed at pH 5.0 with these oligonucleotides are presented in Figure 2. In Figure 2 the left hand portion of each panel shows interaction with the oligonucleotide

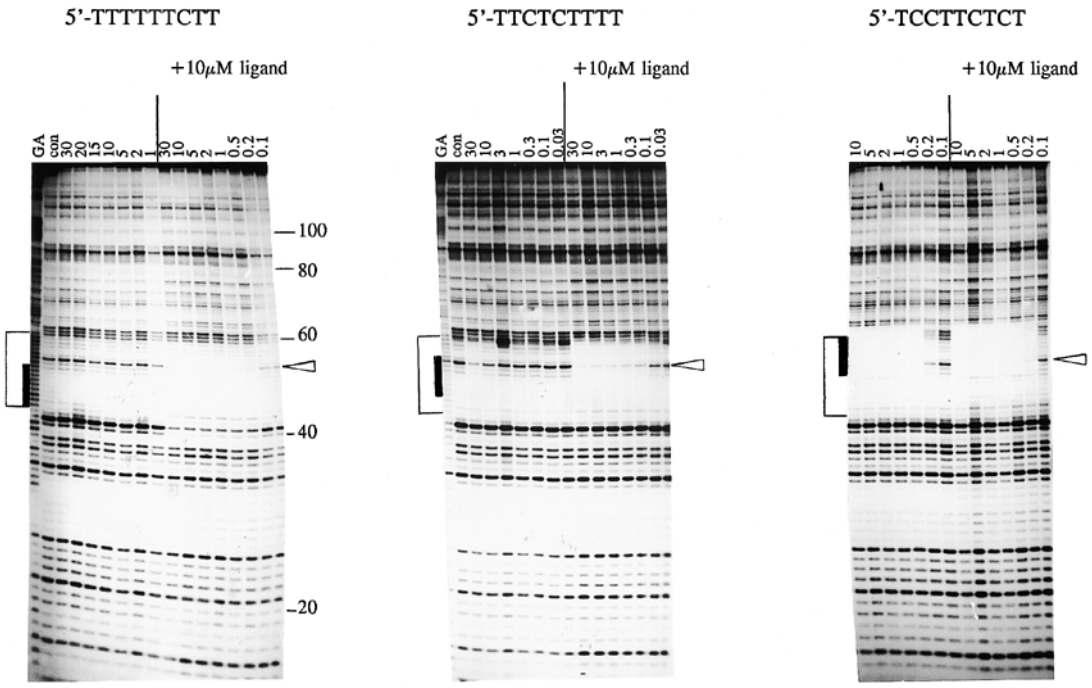

Figure 2. DNase I cleavage patterns of *tyr*T48 in the presence of varying concentrations of the three oligonucleotides at pH 5.0. In each panel the left hand lanes show digestion in the presence of the oligonucleotide alone, while the right hand lanes were performed in the presence of 10 μM triplex binding ligand. The oligonucleotide concentration (μM) is indicated at the top of each lane. Tracks labelled 'con' show digestion of the DNA in the absence of added oligonucleotide or triplex binding ligand. Tracks labelled 'GA' are Maxam–Gilbert markers specific for purines. The brackets indicate the position of the 17mer purine tract, while the filled boxes show the bases targeted by each 9mer third strand. The reactions were performed in 50 mM sodium acetate, pH 5.0, containing 10 mM $MgCl_2$. The arrow indicates the band that was subjected to densitometric analysis.

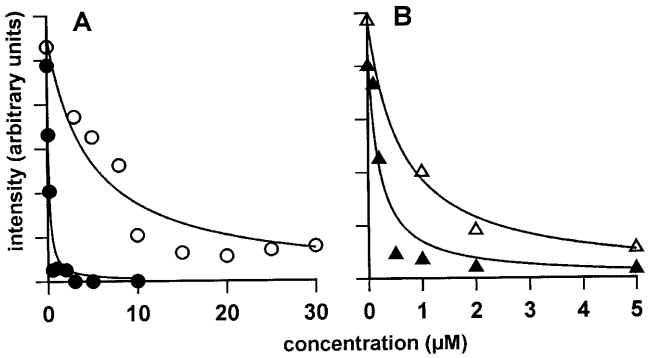

Figure 3. Footprinting plots showing interaction of the 9mer oligonucleotides with the 17 base target site. The intensity of each band was determined from densitometric scans of the autoradiographs shown in Figures 2 and 4. The ordinate shows the oligonucleotide concentration (μM), the abscissa shows the band intensity (arbitrary units). The curves correspond to the binding parameters shown in Table 1. (**A**) No triplex binding ligand, pH 5.0: ○ 5′-TTCTCTTTT; ● 5′-TCCTTCTCT. (**B**) 10 μM triplex binding ligand: Δ 5′-TCCTTCTCT, pH 6.0; ▲ 5′-TTTTTTCTT, pH 5.0.

alone, while the right hand lanes contain 10 μM naphthylquinoline triplex binding ligand. As previously noted (34), DNase I cleavage within the oligopurine tract is generally poor. However, oligonucleotide-induced footprints can clearly be seen and are most evident from inhibition of the strong cleavage products at position 53 (ApG) and at the upper (5′) end of the 17 base target site. Looking first at the left hand panel, it can be seen that in the absence of the ligand 5′-TTTTTTCTT, which should bind to the (lower) 3′-end of the target site, does not alter the DNase I digestion pattern even at a concentration as high as 30 μM. In the presence of the ligand a clear footprint can be seen which persists to the lowest oligonucleotide concentration (0.1 μM). In the absence of the ligand TTCTCTTTT, which binds to the centre of the oligopurine tract, alters the DNase I cleavage pattern at concentrations of 30 and 10 μM. In the presence of the ligand this footprint persists to an oligonucleotide concentration of ~0.2 μM. In contrast, 5′-TCCTTCTCT, which binds to the (upper) 5′-end of the oligopurine tract produces a clear footprint in the absence of the ligand, which persists to a concentration of 0.2 μM and which shows little change on adding 10 μM triplex binding ligand. Inspection of these cleavage patterns reveals that although the footprints extend by a few bases on either side of the target regions, the oligonucleotides have bound selectively to their intended target sites, both in the presence and absence of the triplex binding ligand. Since these oligonucleotides interact with different regions of the same target site, it seems reasonable to suppose that the differences in their affinity arise from variations in their sequence composition, i.e. the relative numbers of T·AT and C^+·GC triplets. The oligonucleotide generating the greatest number of C^+·GC triplets binds with the highest affinity, in contrast to that forming only one C^+·GC triplet, which does not produce a DNase I footprint. A quantitative estimate of the relative binding affinities of the three 9mer oligonucleotides was

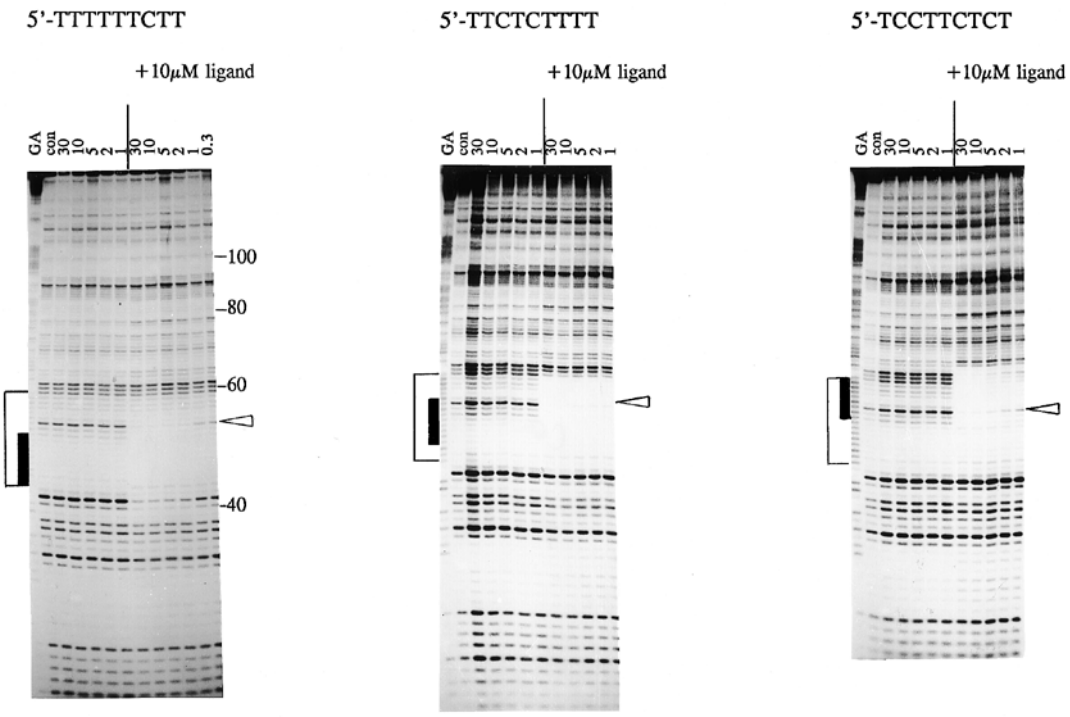

Figure 4. DNase I cleavage patterns of *tyr*T48 in the presence of varying concentrations of the three oligonucleotides at pH 6.0. In each panel the left hand lanes show digestion in the presence of the oligonucleotide alone, while the right hand lanes were performed in the presence of 10 μM triplex binding ligand. The oligonucleotide concentration (μM) is indicated at the top of each lane. Tracks labelled 'con' show digestion of the DNA in the absence of added oligonucleotide or triplex binding ligand. Tracks labelled 'GA' are Maxam–Gilbert markers specific for purines. The brackets indicate the position of the 17mer purine tract, while the filled boxes show the bases targetted by each 9mer third strand. The reactions were performed in 50 mM sodium acetate, pH 6.0, containing 10 mM $MgCl_2$. The arrow indicates the band that was subjected to densitometric analysis.

obtained by densitometric analysis of the bands marked by an arrow, as described in Materials and Methods, producing the footprinting plots (35). C_{50} values, representing the oligonucleotide concentration required to reduce the band intensity by 50%, were calculated from these plots and provide an estimate of the dissociation constant. Representative examples of these plots are presented in Figure 3. The C_{50} values obtained are presented in Table 1 and confirm that the oligonucleotide containing four cytosines binds much better than those containing one or two cytosines.

The results of similar experiments performed at pH 6.0 are presented in Figure 4. It can be seen that under these conditions none of the oligonucleotides produces a footprint in the absence of the ligand, even at concentrations as high as 30 μM. This confirms the requirement for protonation of the third strand cytosine. It is possible that the isolated cytosine in TTTTTTCTT has a higher pK value, but since this oligonucleotide does not produce a footprint at pH 5.0, it is not possible to assess how it is affected by changes in pH. In the presence of 10 μM triplex binding ligand all three oligonucleotides produce clear footprints at their target sites, which persist to similar concentrations of ~1–3 μM. C_{50} values estimated from densitometer scans of these autoradiographs are presented in Table 1. It can be seen that the complex with TTCTCTTTT, binding at the centre of the oligopurine tract, is marginally more stable than the other two complexes. This result may reflect the fine balance between the requirement for C^+·GC triplets to give a strong interaction and the selective stabilization of T·AT triplets by the ligand.

Table 1. C_{50} values (μM) obtained from footprinting plots for the interaction of the three oligonucleotides with their target sequences

	pH 5.0		pH 6.0	pH 7.5
	No ligand	10 μM ligand	10 μM ligand	10 μM ligand
5′-TTTTTTCTT	No footprint	0.19 ± 0.06	0.35 ± 0.10	<2[a]
5′-TTCTCTTTT	4.6 ± 1.1	0.13 ± 0.03	0.15 ± 0.05	≈30[a]
5′-TCCTTCTCT	0.13 ± 0.03	≈0.1[a]	0.62 ± 0.10	No footprint

The value corresponds to the oligonucleotide concentration (μM) required to reduce the intensity of bands in the footprint by 50%.
[a]Accurate values could not be determined from a quantitative analysis; the figures quoted are based on visual inspection of the autoradiographs.

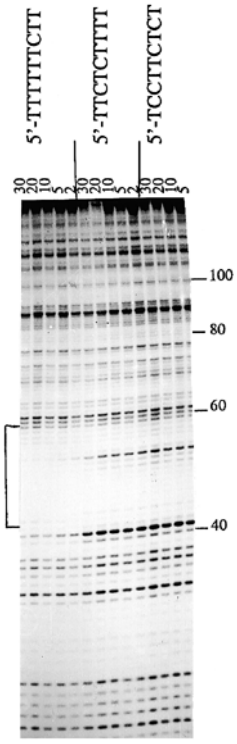

Figure 5. DNase I cleavage patterns of *tyr*T48 in the presence of varying concentrations of the three oligonucleotides at pH 7.5. The oligonucleotides alone do not affect the DNase I cleavage pattern; all reactions contained 10 µM triplex binding ligand. The oligonucleotide concentration (µM) is indicated at the top of each lane. The bracket indicates the position of the 17mer purine tract. The reactions were performed in 10 mM Tris–HCl, pH 7.5, containing 50 mM NaCl and 10 mM $MgCl_2$.

The results of similar experiments performed at pH 7.5 are presented in Figure 5. In the absence of the ligand none of the oligonucleotides produce a DNase I footprint at concentrations as high a 30 µM (not shown). In the presence of 10 µM triplex binding ligand TTTTTTCTT produces a clear footprint which persists to the lowest oligonucleotide concentration (2 µM). The intensity of bands in the target is attenuated at the highest concentration of TTCTCTTTT (30 µM), while TCCTTCTCT has no effect on the cleavage pattern.

DISCUSSION

The results presented in this paper show that under conditions in which the third strand cytosines should be protonated (pH 5.0) there is a clear relationship between the number of C^+·GC triplets and the binding affinity of these 9mer oligonucleotides. Since T·AT and C^+·GC triplets are isostructural and in both cases the third strand makes two hydrogen bond contacts to substituents in the duplex major groove, this difference is most likely due to the presence (C^+·GC) or absence (T·AT) of the positive charge on the third strand base. Previous studies have suggested that blocks of C^+·GC triplets are unstable because of repulsion between the adjacent positive charges, possibly lowering the pK value (9). The present results suggest that provided the cytosine residues are protonated, the C^+·GC triplet makes a greater contribution to triplex stability than T·AT. This conclusion is in agreement with the results of melting studies on intramoleuclar triplexes, which suggested that the free energy change for formation of C^+·GC is greater than T·AT (24).

We can offer two explanations for the greater stabilizing effect of the C^+·GC triplet. Firstly, it is possible that the positive charge may shield the electrostatic repulsion between the phosphate groups on the third strand and the duplex and may thereby facilitate oligonucleotide binding. Alternatively, the positive charge on the protonated cytosines may form favourable interactions with the negative π system of the stacked DNA bases. We favour the second explanation, since the three oligonucleotides show the same order of binding at 1 M NaCl (unpublished observations). In either case this observation has clear implications for the design of novel cytosine analogues for generating stable triplexes at physiological pH. If the positive charge is important then derivatives which retain this charge, such as 2-aminopyridine, may be more successful than uncharged analogues, such as 6-oxocytosine.

The observation that at pH 5.0 the three oligonucleotides bind with similar affinities in the presence of the triplex binding ligand confirms that each of the positions in the oligopurine tract is capable of forming triple helical complexes. Since previous studies have shown that the protonated form of this ligand is more efficient at stabilizing triple helices (32), it is possible that the positive charge on the ligand affects triplex stability in a similar way to the presence of a C^+·GC triplet. This compound, along with other triplex binding ligands, selectively stabilizes T·AT triplets (32). As a consequence the binding stoichiometry may not be the same for each of the triplexes investigated and this may explain why it appears to have a greater effect on the less stable complexes, which contain a greater number of T·AT triplets.

ACKNOWLEDGEMENTS

This work was supported by grants from the Cancer Research Campaign and the Medical Research Council.

REFERENCES

1 Soyfer,V.N. and Potaman,V.N. (1996) *Triple-helical Nucleic Acids.* Springer-Verlag, Berlin, Germany.
2 Thuong,N.T. and Hélène,C. (1993) *Angew. Chem.*, **32**, 666–690.
3 Moser,H.E. and Dervan,P.B. (1987) *Science*, **238**, 645–650.
4 Le Doan,T., Perrouault,L., Praseuth,D., Habhoub,N., Decout,J.L., Thuong,N.T., Lhomme,J., and Hélène,C. (1987). *Nucleic Acids Res.*, **15**, 7749–7760.
5 Radhakrishnan,I. and Patel,D.J. (1994) *Structure*, **2**, 17–32.
6 Beal,P.A. and Dervan,P.B. (1991) *Science*, **251**, 1360–1363.
7 Chen,F.-M. (1991) *Biochemistry*, **30**, 4472–4479.
8 Radhakrishnan,I. and Patel,D.J. (1993) *Structure*, **1**, 135–152.
9 Kiessling,L.L., Griffin,L.C. and Dervan,P.B. (1992) *Biochemistry*, **31**, 2829–2834.
10 Lee,J.S., Wordsworth,M.L., Latimer,L.J.P. and Morgan,A.R. (1984) *Nucleic Acids Res.*, **12**, 6603–6614.
11 Povsic,T.J. and Dervan,P.B. (1989) *J. Am. Chem. Soc.*, **111**, 3059–3061.
12 Xodo,L.E., Manzini,G., Quadrifoglio,F., van der Marel,G. and van Boom,J. (1991) *Nucleic Acids Res.*, **19**, 5625–5631.
13 Singleton,S.F. and Dervan,P.B. (1992) *Biochemistry*, **31**, 10995–11003.
14 Xiang,G., Soussou,W. and McLaughlin,L.W. (1994) *J. Am. Chem. Soc.*, **116**, 11155–11156.
15 Xiang,G., Bogacki,R. and McLaughlin,L.W. (1996) *Nucleic Acids Res.*, **24**, 1963–1970.

16 Ono,A., Ts'o,P.O.P. and Kan,L. (1991) *J. Am. Chem. Soc*., **113**, 4032–4033.
17 Bates,P.J., Laughton,C.A., Jenkins,T.C., Capaldi,D.C., Roselt,P.D., Reese,C.B. and Neidle,S. (1996) *Nucleic Acids Res*., **24**, 4176–4184.
18 Jetter,M.C. and Hobbs,F.W. (1993) *Biochemistry*, **32**, 3249–3254.
19 Miller,P.S., Bhan,P. Cushman,C.D. and Trapane,T.L. (1992) *Biochemistry*, **31**, 2999–3004.
20 Brunar,H. and Dervan,P.B. (1996) *Nucleic Acids Res*., **34**, 1601–1604.
21 Koshlap,K.M., Schultze,P., Brunar,H., Dervan,P.B. and Feigon,J. (1997) *Biochemistry*, **36**, 2659–2668.
22 Koh,J.S. and Dervan,P.B. (1992) *J. Am. Chem. Soc*., **114**, 1470–1478.
23 Giovannangeli,C., Rougée,M., Garestier,T., Thuong,N.T. and Hélène,C. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 8631–8635.
24 Völker,J. and Klump,H.H. (1994) *Biochemistry*, **33**, 13502–13508.
25 Mergny,J.L., Duval-Valentin,G., Nguyen,C.H., Perrouault,L., Faucon,B., Rougée,M., Montenay-Garestier,T., Nisagni,E. and Hélène,C. (1992) *Science*, **256** 1681–1684.
26 Pilch,D.S., Waring,M.J., Sun,J.-S., Rougée,M., Nguyen,C.-H., Bisagni,E., Garestier,T. and Hélène,C. (1993) *J. Mol. Biol*., **232**, 926–946.
27 Lee,J.S., Latimer,L.J.P. and Hampel,K.J. (1993) *Biochemistry*, **32**, 5591–5597.
28 Moraru-Allen,A.A., Cassidy,S., Alvarez,J.-L.A., Fox,K.R., Brown,T. and Lane,A.N. (1997) *Nucleic Acids Res*., **25**, 1890–1896.
29 Wilson,W.D., Tanious,F.A., Mizan,S., Yao,S., Kiselyov,A.S., Zon,G. and Strekowski,L. (1993) *Biochemistry*, **32**, 10614–10621.
30 Cassidy,S.A., Strekowski,L., Wilson,W.D. and Fox,K.R. (1994) *Biochemistry*, **33**, 15338–15347.
31 Chandler,S.P., Strekowski,L., Wilson,W.D. and Fox,K.R. (1995) *Biochemistry*, **34**, 7234–7242.
32 Cassidy,S.A., Strekowski,L. and Fox,K.R. (1996) *Nucleic Acids Res*., **24**, 4133–4138.
33 Fox,K.R., Polucci,P., Jenkins,T.C. and Neidle,S. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 7887–7891.
34 Drew,H.R. and Travers,A.A. (1984) *Cell*, **37**, 491–502.
35 Dabrowiak,J. C. and Goodisman,J. (1989) In Kallenbach,N.R. (ed.), *Chemistry and Physics of DNA–Ligand Interactions*. Adenine Press, New York, NY, pp. 143–174.