enzyme in this organism, which catalyzes the oxidation of aldehydes to carboxylic acids.[22]

It remains to be established, however, exactly why certain acetogenic eubacteria and hyperthermophilic marine archaebacteria use tungsten rather than molybdenum as an essential catalytic component of specific oxidoreductase enzymes, in apparent contrast to the remainder of the microbe world.

(22) White, H.; Strobl, G.; Feicht, R.; Simon, H. *Eur. J. Biochem.* **1989**, *184*, 89-96.

Acknowledgment. We thank Drs. E. I. Stiefel and C. L. Coyle for helpful discussions and the Staff of SSRL for their assistance. SSRL is funded by the DOE (Contract DE-AC03-82ER-13000, Office of Basic Energy Sciences, Division of Chemical Sciences) and NIH (Biotechnology Resource Program, Division of Research Resources). This research was supported in part by a DOE Grant (F609-88ER13901 to M.W.W.A.).

Structural Effects in the Recognition of DNA by Circular Oligonucleotides

Gautam Prakash and Eric T. Kool*

Contribution from the Department of Chemistry, University of Rochester, Rochester, New York 14627. *Received August 8, 1991*

Abstract: It was recently reported that certain pyrimidine-rich circular DNA oligomers can bind strongly and specifically to purine-rich DNA or RNA strands by forming bimolecular triple helical complexes.[1-3] In this study are investigated the effects of structural variations on the strength of binding for this new class of nucleotide-binding ligand. The number of loop nucleotides (nt) which is optimum for bridging the two binding domains of a circle is examined. Comparing loop sizes of 3, 4, 5, 6, and 10 nt, the optimum number of nucleotides in a loop is found to be five for the sequences studied. In order to test the method of construction and the ability of these compounds to bind sites of varied length, we attempted to synthesize circles of varied size. Circles over the size range 24-46 nt were successfully constructed. Varying the target site length shows that oligomers of four, eight, twelve, and eighteen nucleotides can be complexed strongly by circles, with melting temperatures (T_m) 17° to >33 °C higher at pH 7.0 than the corresponding Watson-Crick duplexes of the same length. Also studied is the effect of the covalently closed circular structure in comparison to linear oligomers having the same sequence; it is shown that a covalently closed circle has considerably higher binding affinity than do three different "nicked" circles (linear oligomers) which contain the same bases. The high binding affinities of these circles are thus attributed to the entropic benefit of preorganization. Finally, the ability of such circles to bind to complementary sites within longer oligomers, the ends of which must pass beyond the loops of a circle, is confirmed by melting studies with synthetic target strands 36 bases in length.

Introduction

A common theme in the field of molecular recognition is the idea that macrocyclic molecules are efficient at recognizing and binding specific substrates. Researchers have become skilled at constructing cyclic synthetic hosts for noncovalent binding of complementary-shaped guest molecules and ions.[4-7] Among the earliest examples of cyclic synthetic hosts were the crown ethers, which chelate metal ions with high affinity and selectivity.[8] Macrocyclic structures can give host molecules strong advantages over their linear analogues. One such advantage is strength of binding, due in large part to the entropic benefit of preorganization.[9] A second advantage is specificity of recognition, which is the result of defined cavity size and shape.[10] Although it is certainly true that some elegant noncyclic structures can possess these properties as well,[11] the use of macrocyclic structure remains an important strategy in this field.

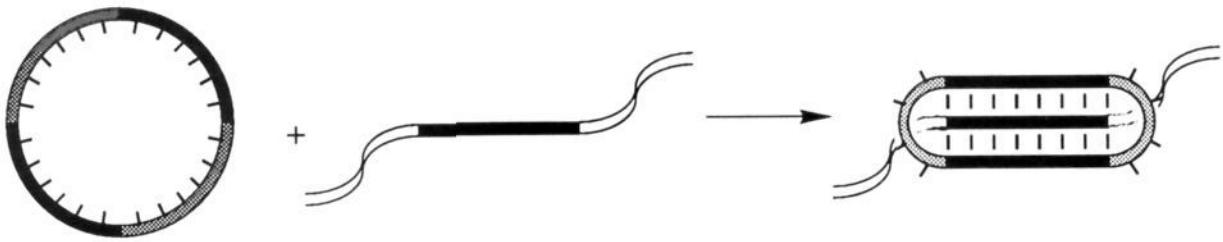
There are also many naturally occurring examples of circular molecules which function as ligands for guest species. For example, the cyclic siderophores act to bind and transport iron,[12] and many other natural ionophores, such as valinomycin[13] and nonactin,[14] are macrocyclic as well. The cyclodextrins, which are circular oligosaccharides, have been studied intensely for their efficient binding of hydrophobic guest molecules inside their central cavity.[15] Several cyclic oligopeptides have also been characterized; examples are the antibiotics echinomycin and triostin A, which bind to specific sequences of duplex DNA.[16] Interestingly, of the three common natural biopolymers—saccharides, peptides, and nucleotides—only the first two have been represented in the known family of cyclic host molecules.

We recently found, however, that cyclic oligonucleotides can also act as efficient hosts for molecular recognition. It was demonstrated that circular DNA oligomers can bind strongly and sequence specifically to single strands of DNA and RNA by forming triple helical complexes.[1,2] Circular oligomers were shown to bind their complements six orders of magnitude more strongly than standard Watson-Crick complementary oligomers at pH 7.0 and are completely resistant to degradation by exonucleases.[1] It was subsequently found that these circular oligomers are less tolerant of mismatches in the target sequence than are standard linear oligomers, and thus show higher sequence selectivity in binding.[2] Because of the similarities in structure, binding strength, and selectivity to the crown family of compounds, these circular ligands were referred to as "crown nucleotides".[1]

(1) Prakash, G.; Kool, E. T. *J. Chem. Soc., Chem. Commun.* **1991**, 1161.
(2) Kool, E. T. *J. Am. Chem. Soc.* **1991**, *113*, 6265.
(3) D'Souza, D. J.; Kool, E. T., submitted for publication.
(4) Lehn, J. M. *Struct. Bond* **1973**, *16*, 1.
(5) Hamilton, A. D.; Pant, N.; Muehldorf, A. *Pure Appl. Chem.* **1988**, *60*, 533.
(6) Gutsche, C. D. *Acc. Chem. Res.* **1983**, *16*, 161.
(7) Diederich, F. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 362.
(8) Pedersen, C. J. *J. Am. Chem. Soc.* **1967**, *89*, 7017.
(9) Cram, D. J. *Chem. Tech.* **1987**, 120.
(10) Cram, D. J.; Ho, S. P. *J. Am. Chem. Soc.* **1986**, *108*, 2998.
(11) Rebek, J. *Science* **1987**, *235*, 1478.
(12) (a) Pollack, J. R.; Neilands, J. B. *Biochem. Biophys. Res. Commun.* **1970**, *38*, 989. (b) O'Brien, I. G.; Gibson, F. *Biochim. Biophys. Acta* **1970**, *215*, 393.
(13) (a) Dobler, M. *Ionophores and Their Structures*; Wiley: New York, 1981. (b) Neupert-Laves, K.; Dobler, M. *Helv. Chim. Acta* **1971**, *54*, 1709.
(14) Kilbourn, B. T.; Dunitz, J. D.; Pioda, L. A. R.; Simon, W. *J. Mol. Biol.* **1967**, *30*, 559.
(15) (a) Bender, M. L.; Komiyama, M. *Cyclodextrin Chemistry*; Springer-Verlag: New York, 1978. (b) Trainor, G.; Breslow, R. *J. Am. Chem. Soc.* **1981**, *103*, 154.
(16) Ughetto, G.; Wang, A. H.-J.; Quigley, G. J.; Van der Marel, G. A.; Van Boom, J. H.; Rich, A. *Nucleic Acids Res.* **1985**, *13*, 2305.

0002-7863/92/1514-3523$03.00/0

Chart I



Previous studies have focused on the construction, strength of binding, and sequence specificity as well as on the effects of factors such as cations, solvent, and pH on the complexes.[1-3] In the present study we explore the generality of the circle·single strand binding phenomenon by varying the structure of circular oligomers, examining both the effects of circle size and structure, and of binding site size, and we examine the origin of the strength of binding by comparing linear and circular oligomers which can form the same noncovalent bonds.

Experimental Section

Oligonucleotide Synthesis and Purification. Oligodeoxynucleotides were synthesized on a Pharmacia LKB instrument using standard phosphoramidite chemistry.[17] β-Cyanoethyl phosphoramidite monomers were obtained from Cruachem. Phosphorylation on the 5′ ends of oligomers was carried out with a commercially available reagent, also purchased from Cruachem. Oligomers were purified by electrophoresis on 20% acrylamide gels with 8 M urea and Tris·borate·EDTA buffer.[18] The DNA was isolated from the gels by excision followed by elution and dialysis and was quantitated by absorbance at 260 nm using extinction coefficients which were calculated by the nearest neighbor method.[19]

Circularization of linear 5′-phosphorylated precircle oligomers was carried out as previously described.[1] In summary, linear 5′-phosphorylated circle precursors were synthesized and then hybridized with short complementary DNA templates. These precircle complexes bring the reactive 3′-hydroxyl and 5′-phosphate ends adjacent to one another, and these ends are ligated using aqueous BrCN/imidazole/Ni^{2+}. In all cases in this study the precircle ends were joined at the center position of a Watson–Crick binding domain. We have found that this approach gives higher yields than does ligation within the Hoogsteen domain, although the products are identical. Isolation of the circular product (the principal new product from the reaction) was carried out using gel electrophoresis. The circular products typically migrate on a gel at ~0.9 times the rate of their linear precursors. Circularity was confirmed by resistance both to 3′-exonuclease cleavage by T4 DNA polymerase and to 5′-dephosphorylation by T4 polynucleotide kinase.

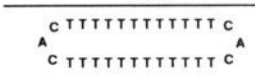
Thermal Denaturation Methods. Melting studies were carried out using a Gilford 250 UV–vis spectrophotometer equipped with a Gilford 2527 thermocontroller, as described previously.[1,2] Solutions contained 3 μM of each oligomer in a buffer of 100 mM NaCl, 10 mM $MgCl_2$, and 10 mM Tris·HCl (pH 7.0). Melting temperatures (T_m) were taken to be the temperature of half-dissociation and were obtained from a plot of the derivative of $1/T$ vs absorbance at 260 nm. Precision in T_m values, estimated from variance in repeated experiments, is ±0.5 °C. Complexes of circles showed a single, sharp melting transition; free energies of complexation were derived by fitting the melting data with a two-state model for denaturation.[2,20]

Results and Discussion

Optimization of Loop Size. The circular ligands are constructed to contain two opposed pyrimidine-rich binding domains, the Watson–Crick and Hoogsteen domains. The Watson–Crick domain is complementary to the target sequence in the conventional antiparallel sense, and the Hoogsteen domain is complementary in the parallel sense.[21] The circular ligand forms a chelate-like complex around two sides of a complementary purine-rich target sequence. Thus, a pyrimidine·purine·pyrimidine triple helix[22] is formed, with nucleotide loops bridging the outer two strands (Figure 1). Previous studies employed circles with 5-nt loops of sequence –CACAC–.[1-3] This sequence was chosen arbitrarily to contain both pyrimidines and purines and to be noncomplementary with other portions of the circle.

To study the effects of varied loop size on these complexes, we synthesized circles with loops of 3, 4, 5, 6, and 10 nt, with sequences 5′-CAC–, –CACA–, –CACAC–, –CACACA–, and –CACACACACA–, respectively. Attempted construction of circles with loops of 2 nt failed; apparently the loops are too small

Table I. The Effects of Varied Loop Size on the Melting Temperature (T_m) of the Complexes of Circular Oligonucleotides with d(A)12 and with a 36mer Containing the $d(A)_{12}$ Sequence[a]

circle	loop size (nt)	T_m of complex (°C) 12mer substrate	T_m of complex (°C) 36mer substrate
	3	49.2	
	4	54.9	49.7
	5	58.6	56.2
	6	55.8	52.3
	10	54.1	

[a] See text for 36mer sequence and for experimental conditions.

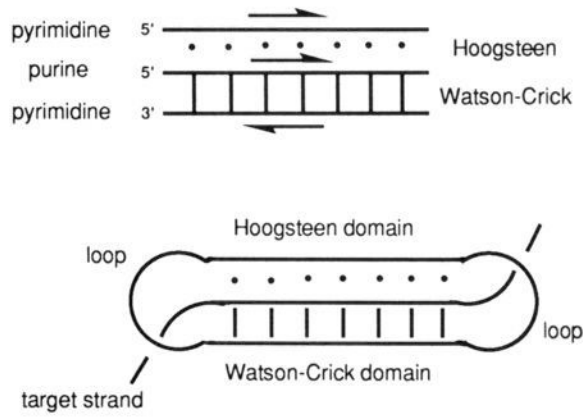


Figure 1. Comparison between a termolecular pyr·pur·pyr triple helix (top) and a bimolecular circle complex (bottom). Short straight lines represent Watson–Crick hydrogen bonds, and Hoogsteen bonds are shown with dots. Arrows denote 5′ to 3′ directionality.

(17) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859.
(18) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*, 2nd ed.; Cold Spring Harbor Press: 1989.
(19) Borer, P. N. In *Handbook of Biochemistry and Molecular Biology*; G. D. Fasman, Ed.; CRC Press: Cleveland, 1975; p 589.
(20) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256.
(21) Moser, H. E.; Dervan, P. B. *Science (Washington, D.C.)* **1987**, *238*, 645.
(22) Arnott, S.; Selsing, E. *J. Mol. Biol.* **1974**, *88*, 509.

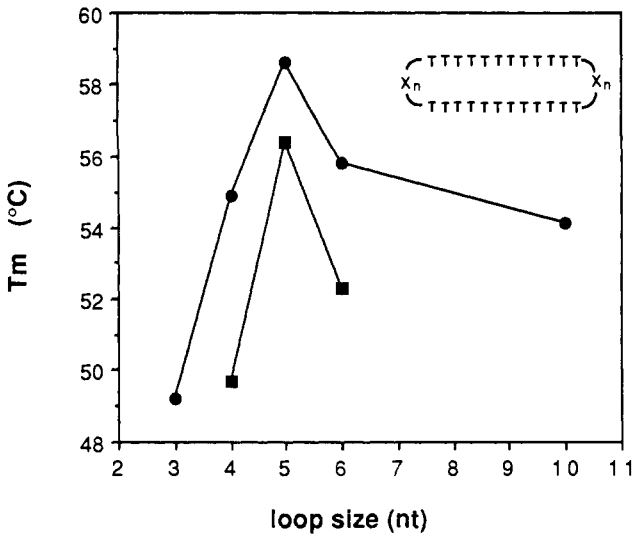


Figure 2. Effects of loop size in nucleotides (nt) on the T_m of a circle complex with $d(A)_{12}$ (circles) and with an A_{12} site in a 36-nt oligomer (squares). Sequence of 36mer is 5′-GGACTCTATCAGAAAAAAAAAAAAGGACTCTATCAG. Loop sequence is given in Table I.

to allow correct alignment of the precircle ends in the ligation reaction. This series of variable-loop circles (Table I) all contained the same $d(T)_{12}$ binding domains, and the circles were hybridized both to $d(A)_{12}$ and to a 36-nt oligomer of sequence 5′-GGACTCTATCAGAAAAAAAAAAAAGGACTCTATCAG, which has an internal $d(A)_{12}$ site, for study of the complex strengths.

Measurement of the melting temperatures for these complexes reveals a clear preference for five-base loops for the sequences studied. Using the $d(A)_{12}$ substrate, the T_m's are 3–4 °C higher with the five-base loop than with four or six, and 5 °C higher than with the 10-base loop. Using the longer substrate strand, the trend is the same, but with greater preference for the five-base loop (Figure 2).

This trend for the loop sizes is reasonable, considering the structures involved. It is likely that loops which are shorter than optimum will distort the ends of the helix in such a complex, resulting in weaker binding by lessening base stacking and hydrogen bonding contributions. Loops which are larger than the optimum may weaken such a complex entropically. Small circles are more conformationally restricted than larger ones, and, as a result, the circles with smaller loops probably lose less entropy upon complexation than do those with larger loops.[23,24]

These findings are consistent with studies done on hairpin loops in duplex RNA[24] and DNA.[25] In those cases, loops of four or five bases were found to be the most stable, with sequence playing only a small role in this preference. Models indicate that in a pyr·pur·pyr triple helix, the distance between the two pyrimidine strands is only a little larger than the distance between two strands in Watson–Crick duplex, and this is consistent with our finding of a preference for five-base loops in the present case. We are currently examining the extent of sequence dependence in the loops on the stability of circle complexes.

Complexes of circles with longer oligomers have some structural similarities to H-DNA, which is a related secondary structure that forms at mirror repeat sequences in duplex DNA.[26] H-DNA is formed when a pyrimidine strand, under superhelical stress or at low pH, dehybridizes from the duplex and folds back on a neighboring purine site to form a triple helical secondary structure. This forms a loop across the purine strand, bridging between the pyrimidine strands of the triplex, the same situation as is found in one-half of a circle complex (Figure 3). A study of loop sizes in H-DNA structures concluded that a loop of four bases gave a more stable triplex than did loop sizes of six, eight, and ten bases.[27] H-DNA loop sizes of three, four, five, and six bases have apparently not been compared experimentally, although molecular modeling studies have been reported.[28] Our data suggest that five-base loops might be preferred over four- and six-loops for bridging the pyrimidine strands in H-DNA.

Effect of Helix Length. To compare the effects of helix length on stability, we have constructed circles with binding domains of varied length. Examination of the data, compiled from this and a previous study,[1] gives some insight into the scope of possible variations in circle and binding site size for these complexes. Figure 4 shows a plot of the melting temperatures for four circle complexes, with triple helix lengths of 4, 8, 12, and 18 bases. The circles are complementary to simple $d(A)_n$ oligomers, and the stabilities are compared with $d(A)_n$·$d(T)_n$ duplexes for the same lengths and under the same conditions (pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$). Loop sequences for these circles were –CACA– except for the $n = 18$ case, which had loops of –CACAC–. The binding domains were runs of $d(T)_n$. Three circles were actually constructed, for $n = 8$, 12, and 18. In the case of the four-base helix, the $n = 8$ circle was hybridized to $d(A)_4$, because synthesis of the 16-base $n = 4$ circle was unsuccessful.

Thermal denaturation studies show that the circle·single strand complexes are consistently more stable than Watson–Crick duplexes of the same length (see Figure 4 for the comparison). For the 18-base circle complex, the T_m is ~17 °C higher than the T_m interpolated for the duplex $d(A)_{18}$·$d(T)_{18}$. The differences in T_m increase for the shorter helix lengths; for example, for $n = 8$, the circle complex is 31 °C higher in T_m, and for $n = 4$, the circle complex has a T_m of 33 °C, while the $d(A)_4$·$d(T)_4$ duplex has a T_m below 0 °C. It should be noted that this study was carried out before the optimum loop size was determined, and so even larger differences would be expected with the optimized five-base loops. In any case, these complexes are among the strongest known intermolecular DNA complexes at neutral pH. This property is important in the development of potential biologically active agents: several studies have shown that increasing the binding strength of antisense oligomers has a positive effect on antiviral activity in cell culture.[29]

The helix length—and thus binding site size—is significant for reasons of specificity as well as for overall complex strength. Longer binding sites imply higher site-selectivity of binding, since the longer a specific target sequence, the fewer times it will occur in random DNA. It has been estimated that recognition of 15–17 contiguous bases is necessary to achieve single-site specificity in the 3-gigabase human genome.[29] This estimate assumes perfect base specificity in sequence matching, which may not be achieved in reality.[30] We have previously shown that circles are considerably more selective for their matched target sequence than are linear oligomers,[2] implying that for circles, selectivity as high as a standard linear oligomer may be achievable with somewhat shorter target sites. In any case, this recognition site size falls well within the capabilities for circle construction.

Advantages of Covalently Closed Circular Structure. It is possible for linear oligonucleotides to form triple helical complexes with single strands, by looping around to form both Watson–Crick and Hoogsteen hydrogen bonds.[31,32] To investigate the relative

(23) Jencks, W. P. *Adv. Enzymol. Rel. Areas Mol. Biol.* **1975**, *43*, 219.
(24) Groebe, D. R.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1988**, *16*, 11725.
(25) Haasnoot, C. A. G.; Hilbers, C. W.; van der Marel, G. A.; van Boom, J. H.; Singh, U. C.; Pattabiraman, N.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1986**, *3*, 843.
(26) (a) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomolec. Struct. Dyn.* **1986**, *4*, 667. (b) Voloshin, O. N.; Mirkin, S. M.; Lyamichev, V. I.; Beloserkovskii, B. P.; Frank-Kamenetskii, M. D. *Nature* **1988**, *333*, 475. (c) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *J. Biol. Chem.* **1989**, *264*, 5950.
(27) Shimizu, M.; Hanvey, J. C.; Wells, R. D. *J. Biol. Chem.* **1989**, *264*, 5944.
(28) Harvey, S. C.; Luo, J.; Lavery, R. *Nucleic Acids Res.* **1988**, *16*, 11795.
(29) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543.
(30) Li, Y.; Zon, G.; Wilson, W. D. *Biochemistry* **1991**, *30*, 7566.
(31) Xodo, L. E.; Manzini, G.; Quadrifoglio, F. *Nucleic Acids Res.* **1990**, *18*, 3557.
(32) Giovannangeli, C.; Montenay-Garestier, T.; Rougée, M.; Chassignol, M.; Thuong, N. T.; Hélène, C. *J. Am. Chem. Soc.* **1991**, *113*, 7775.

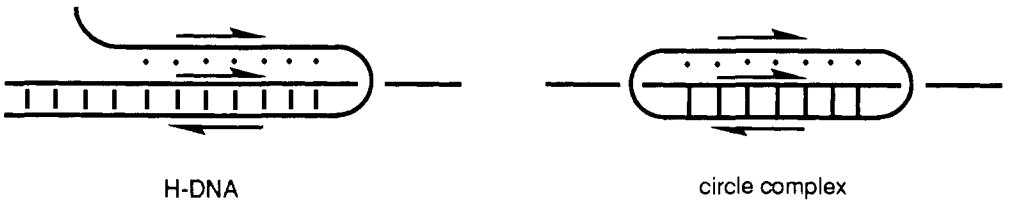


Figure 3. Structural similarities between H-DNA and circle-single strand complexes.

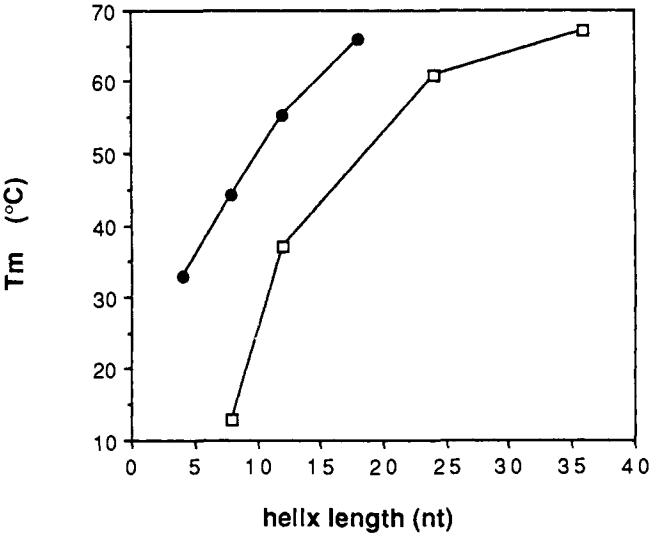


Figure 4. Effects of triple helix length (n, in nucleotides) on T_m of circle complexes with $d(A)_n$ (circles) as compared to duplex DNA having the sequence $d(T)_n$-$d(A)_n$ (squares) at pH 7.0. The loop sequence is 5′-CACA- except where noted in the text.

Table II. Comparison of Binding Affinities of Nicked Circles (Linear Oligomers) with That of a Covalently Closed Circle[a]

entry	complex	T_m (°C)
1	5′ AAGAAAAGAAAG TTCTTTTCTTTC 5′	43.8
2	A CTTCTTTTCTTTC CA C AAGAAAAGAAAG C A CTTCTTT TCTTTC CA P OH	46.4
3	HO P A CTTCTTT TCTTTC CA C AAGAAAAGAAAG C A CTTCTTTTCTTTC C A	46.6
4	P-A CTTCTTTTCTTTC CA AAGAAAAGAAAG C HO-C A CTTCTTTTCTTTC CA	54.9
5	A CTTCTTTTCTTTC CA C AAGAAAAGAAAG C A CTTCTTTTCTTTC CA	62.3

[a] Melting temperatures (T_m) were measured at 260 nm, pH 7.0, for the complexes with the complement 5′-d(AAGAAAAGAAAG). "P" denotes a 5′ phosphate group.

binding strengths for this approach in comparison to the use of fully circular oligomers, we constructed three linear oligomers (Table II), all of which contain the same nucleotides and in the same relative order as a reference circular oligomer. These linear compounds should in principle form the same triple helical complex with the substrate, d(AAGAAAAGAAAG), as is formed by the circle, but with a nick at one of three different positions: in the center of either of the binding domains and at the loop end.

Melting studies of the resulting complexes at pH 7.0 show that all three of these "nicked" complexes (entries 2–4) are somewhat stronger than the reference Watson–Crick duplex (entry 1), and thus probably involve Hoogsteen bonds as well, as would be expected from their sequences. The melting curves for the three complexes show a single transition from bound to dissociated structures. The model duplex melts at 43.8 °C, and the three nicked complexes melt at 46.4 °C for the case nicked in the Watson–Crick domain (entry 2), 46.6 °C for the case nicked in the Hoogsteen domain (entry 3), and 54.9 °C for the case nicked at the loop end (entry 4). By comparison, the covalently closed circle (entry 5) melts at 62.3 °C, for an advantage of 18.5 °C over the Watson–Crick duplex, and 7.4 °C higher than the best linear oligomer in the experiment.

Because they contain the same bases, the nicked oligomers should, in principle, form the same Watson–Crick and Hoogsteen hydrogen bonds with the complement as the circle does. The observation that all three nicked oligomers bind somewhat more tightly to the complement than does the Watson–Crick complementary oligomer suggests that additional interactions, including Hoogsteen bonds, are probably involved in the complexes. The fact that the template-mediated cyclization reaction can be carried out successfully with either of the first two complexes also indicates that both types of hydrogen bonds are involved. The third (entry 4) is similar to an oligomer studied by Xodo et al.,[31] which was shown to form a triple helical complex with a short complement. Interestingly, it binds significantly stronger than the two which are nicked in the binding domains. This difference may reflect some intrinsic advantage of having continuous binding domains of 12 bases rather than short ones of six plus six bases.

It is clear that the covalent closure of these structures into a true circle results in a much greater binding affinity than is achieved with any of the linear variants. Since it should be possible for the linear variants to form the same hydrogen bonding and base stacking arrangements as the circular version, it is likely that this difference in affinity is due at least in part to an entropic effect. The circular structure is preorganized for complexation relative to the linear compounds, and thus loses less entropy upon forming the triple helical complex.

Binding to Sites within Longer Strands. In order to bridge the two outer strands in a triple-helical circle complex, a nucleotide loop must reach across the central target strand. If compounds such as these are to be targeted to biologically relevant nucleic acids, an important question to be answered is whether this interaction will affect the binding to a complementary site within a long sequence. To address this question, we synthesized two 36-nt oligomers which contain a central 12-base binding site flanked by two noncomplementary 12-base sequences. Table III displays the results of melting studies for two different circle sequences. Each was hybridized to its minimal 12-nt complement as well as to a longer target oligomer, the ends of which must pass beyond the loops in a complex.

First, in order to confirm that the actual binding site in the longer sequence is the central 12-base purine run, we hybridized the A_{12}-binding circle to the mismatched 36mer, in which the purine run has three guanine residues. The T_m for this complex was 33 °C, or 25 °C lower than for the same circle with the correctly matched A_{12} site. This demonstrates that binding occurs at the central 12-base site in the 36mer. Similarly, when the second circle is hybridized to the mismatched 36mer having the A_{12} site, the T_m is once again low (19 °C), as a result of three C–A–C mismatches. These T_m values are quite close to those seen for the same mismatched complexes with the corresponding 12-base oligomers.[1] These experiments confirm that binding

Table III. Comparison of the Affinities (as Measured by T_m) of Two Different Circular Oligomers for a 12mer Substrate or for the Same Sequence Embedded within a 36mer Oligonucleotide[a]

complex[a]	T_m (°C)
TTTTTTTTTTTT / AAAAAAAAAAAA / TTTTTTTTTTTT	57.5
5'-GGACTCTATCA (TTTTTTTTTTTT / GAAAAAAAAAAAAG / TTTTTTTTTTTT) GACTCTATCAG	56.2
TTCTTTTCTTTC / AAGAAAAGAAAG / TTCTTTTCTTTC	59.8
5'-GGACTCTATCA (TTCTTTTCTTTC / GAAGAAAAGAAAGG / TTCTTTTCTTTC) GACTCTATCAG	63.4

[a] Curved lines represent 5-base loops of sequence –CACAC–.

occurs at the expected sites in the center of the 36mer sequences.

Comparisons show (Table III) that the effects of the loop-flanking base interaction are relatively small for the sequences studied. Both complexes with the 36mer targets have T_m's within 4 °C of the complexes with the minimal nonoverhanging 12mers. In one case, the T_m with the short sequence is the same or slightly higher than with the 36mer, and in the second case, the T_m is 3.6 °C higher with the longer sequence. The small difference in the results between the two cases may be due to sequence-dependent conformational differences in the two triplexes. The results do demonstrate that it is possible for circular oligomers to bind strongly to sites in longer sequences as well as to short oligomers. Models suggest that the first base or two of the target sequence flanking the triplex region may be able to form pairs with neighboring loop bases. In the present case these bases are partly complementary to each other: the flanking base is a G on each end of the site, and the loops contain C's at the positions adjacent to the helix. Our preliminary studies have shown that this complementarity may be important, and so we are currently examining further the role of loop bases and binding site flanking bases on binding strength.

On examining the structure of a bound circular oligomer, one question to be answered is whether the central target strand passes *through* the loop in such a complex, or if the circle simply wraps around the central strand without this topological linkage. Models strongly suggest that this threading cannot easily occur (see Figure 5A), and published models of H-DNA loop structure support this assumption.[26,27] Examination of pyrimidine·purine·pyrimidine base triplet structure shows that a loop connecting the two pyrimidines will most easily pass directly from one strand to another, without looping around the central purine strand.

The building of models shows that these macrocyclic ligands share some similarities with two general types of host molecule: the crowns and the molecular clefts. Since the uncomplexed circles are not especially rigid, they loosely resemble the crowns in uncomplexed forms, and like crowns, they contain regularly-spaced multiple binding groups. However, in this case the substrate for binding is a long linear one, and, in the complex, the circle presents a long linear (helical) cleft with convergent hydrogen-bonding groups pointing inward toward the guest molecule (see Figure 5B). This convergence is an important functional feature of cleft-shaped host molecules.[33] In the present case, this spatial arrangement allows for very strong binding and high shape-selectivity as well.

(33) Rebek, J.; Marshall, L.; Wolak, R.; Parris, K.; Killoran, M.; Askew, B.; Nemeth, D.; Islam, N. *J. Am. Chem. Soc.* **1985**, *107*, 7476.

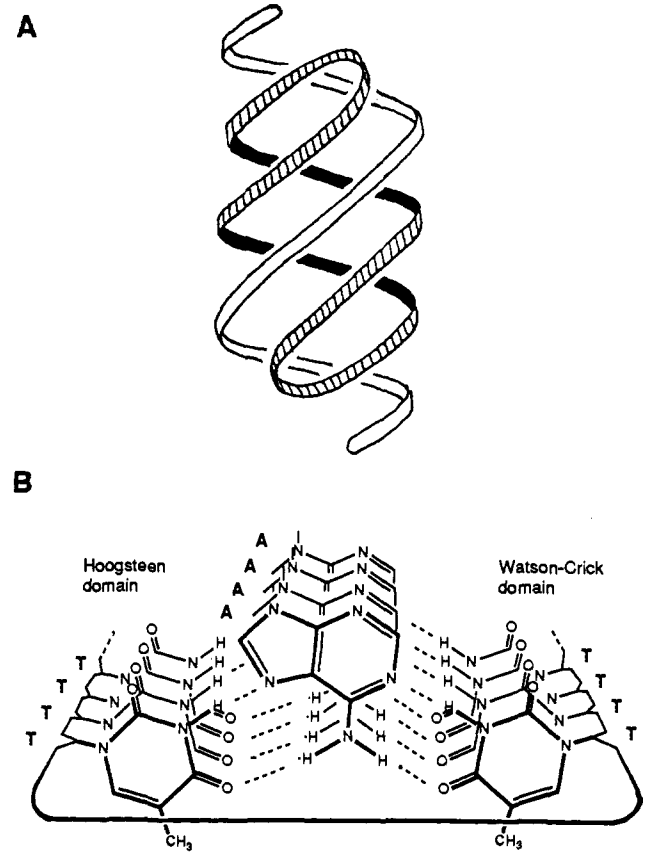


Figure 5. Illustrations of complexes formed between oligonucleotide circles and single strands. (A) Helical representation showing bimolecular triplex structure. (B) End-on view of complex showing convergence of hydrogen bonds from outer circle bases to the central purine strand. For clarity, the helix is unwound and the phosphate–deoxyribose backbone is simplified.

Conclusions

We have delineated a range of variations in loop size and helix length which are tolerated for oligonucleotide circle·single strand complexes and at the same time demonstrate optimum structures which will allow the strongest binding of a substrate. It is concluded that loops of five bases are preferred for such circles. Binding site size for these complexes can be varied at least over the range of 4–18 bases, and the melting temperatures are ~20–30 °C higher than the analogous Watson–Crick duplexes at pH 7.0. Strong-binding circular oligomers have been demonstrated which span the size range of 24–46 nucleotides. It is further shown that covalent closure of triple helix-forming oligomers into circles results in considerably higher binding affinity than can be achieved with the linear analogues, thus demonstrating that the circular structure is the major contributing factor in the binding strength. Finally, it is shown that purine sites within longer DNA sequences can be targeted with no loss of binding strength.

Acknowledgment. We thank the University of Rochester and the National Institutes of Health (RO1-GM46625) for support and Professor D. H. Turner for generously allowing use of the spectrophotometer.

Registry No. $d(A)_{12}$, 125620-91-9; 3 nt cyclic, 139523-50-5; 4 nt cyclic, 139523-51-6; 5 nt cyclic, 139523-55-0; 6 nt cyclic, 139523-58-3; 10 nt cyclic, 139523-59-4; entry 2, 139523-52-7; entry 3, 139523-53-8; entry 4, 139523-54-9; entry 5, 135105-43-0; 5'-AAGAAAAGAAAG, 134881-26-8; 5'-CTTTCTTTTCTT, 139463-03-9; 5'-GGACTCTATCAGAAAAAAAAAAAAGGACTCTATCAG, 139523-56-1; 5'-GGACTCTATCAGAAGAAAAGAAAGGGACTCTATCAG, 139523-57-2.