

Flanking Sequence Effects within the Pyrimidine Triple-Helix Motif Characterized by Affinity Cleaving[†]

Laura L. Kiessling, Linda C. Griffin, and Peter B. Dervan*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received September 3, 1991; Revised Manuscript Received December 30, 1991

ABSTRACT: Nearest neighbor interactions affect the stabilities of triple-helical complexes. Within a pyrimidine triple-helical motif, the relative stabilities of natural base triplets T·AT, C+GC, and G·TA, as well as triplets, D₃·TA and D₃·CG, containing the nonnatural deoxyribonucleoside 1-(2-deoxy-β-D-ribofuranosyl)-4-(3-benzamido)phenylimidazole (D₃) were characterized by the affinity cleaving method in the context of different flanking triplets (T·AT, T·AT; T·AT, C+GC; C+GC, T·AT; G+GC, C+GC). The T·AT triplet was shown to be insensitive to substitutions in either the 3' or 5' directions, while the relative stabilities of triple helices containing C+GC triplets decreased as the number of adjacent C+GC triplets increased. Triple helices incorporating a G·TA interaction were most stable when this triplet was flanked by two T·AT triplets and were adversely affected when a C+GC triplet was placed in the adjacent 5' direction. Similarly, complexes containing D₃·TA or D₃·CG triplets were most stable when the triplet was flanked by two T·AT triplets but were destabilized when the adjacent 3' neighbor position was occupied with a C+GC triplet. This information regarding sequence composition effects in triple-helix formation establishes a set of guidelines for targeting sequences of double-helical DNA by the pyrimidine triple-helix motif.

Triple-stranded polynucleotide RNA complexes were identified shortly after the discovery of the DNA double helix (Felsenfeld et al., 1957). These structures consist of one purine and two pyrimidine strands, and it was postulated that this interaction occurs through the formation of a triple helix in which a pyrimidine strand binds in the major groove of double-helical DNA (or RNA) with specific hydrogen bonds formed to the Watson-Crick purine strand. To date, no high-resolution X-ray structure for a triple helix has been reported; however, chemical and spectroscopic studies support triple-helical structures derived from the formation of T·AT and C+GC base triplets (Felsenfeld et al., 1957; Lipsett, 1963,

1964; Michelson et al., 1967; Morgan & Wells, 1968; Lee et al., 1979; Moser & Dervan, 1987; Praseuth et al., 1988; Rajagopal & Feigon, 1989a,b; de los Santos et al., 1989; Sklenar & Feigon, 1990; Pilch et al., 1990; Radhakrishnan et al., 1991a) (Figure 1).

In exploring the possibility that the triple-helix structure could function as a recognition motif for single sites in large double-helical DNA, it was demonstrated that a pyrimidine oligonucleotide binds specific purine tracts within large duplex DNA (Moser & Dervan, 1987; Le Doan et al., 1987). The pyrimidine third strand is parallel to the purine Watson-Crick strand (Moser & Dervan, 1987). Since these initial findings, the pyrimidine triple-helix motif has been utilized to target single sites within megabase DNA (Strobel & Dervan, 1990, 1991); the recognition code has been broadened to include G·TA triplets (Griffin & Dervan, 1989); and a nonnatural

[†] We are grateful to the Office of Naval Research for generous support.

* To whom correspondence should be addressed.

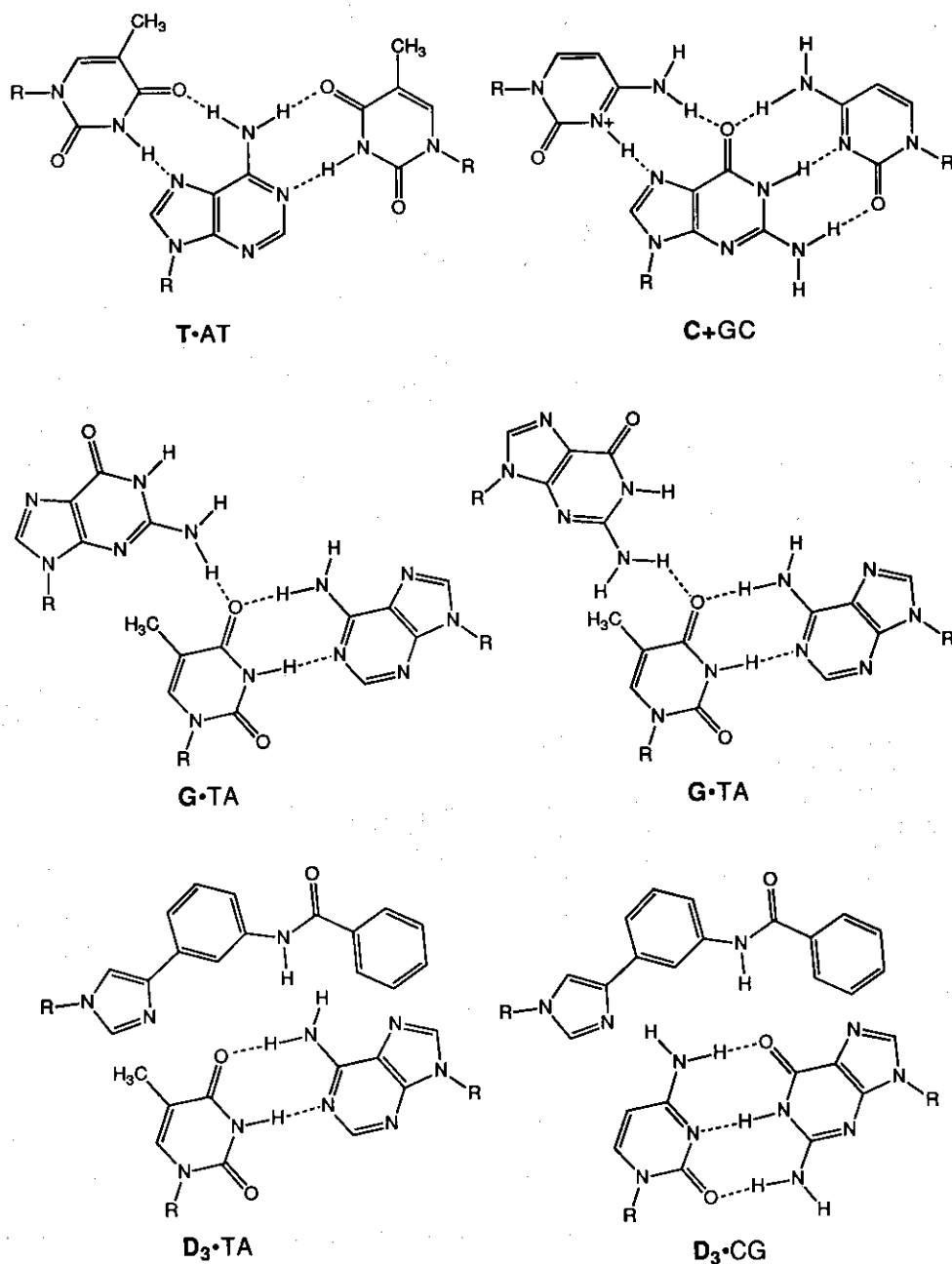


FIGURE 1: Base triplets T-AT, C+GC, and G-TA (Griffin & Dervan, 1989; Radhakrishnan et al., 1991a) and D_3 -TA and D_3 -CG (Griffin, 1990).

deoxyribonucleoside, 1-(2-deoxy- β -D-ribofuranosyl)-4-(3-benzamido)phenylimidazole (D_3) has been designed to bind selectively TA and CG base pairs (Griffin, 1990) (Figure 1).

More recently, an alternate triple helix has been characterized containing a purine third strand that binds the Watson-Crick purine strand (Lipsett, 1964; Mark & Thiele, 1978; Broitman et al., 1987; Letai et al., 1988; Kohwi & Kohwi-Shigematsu, 1988; Cooney et al., 1988; Beal & Dervan, 1991; Radhakrishnan et al., 1991b). Specificity is derived from G recognition of GC base pairs (G-GC triplets) and A or T recognition of AT base pairs (A-AT or T-AT triplets). The orientation of the purine-rich third strand has been shown to be antiparallel to the purine Watson-Crick strand (Beal & Dervan, 1991; Pilch et al., 1991; Radhakrishnan et al., 1991b).

Through the formation of specific base triplets, oligonucleotide-directed triple-helix formation appears to provide a generalizable chemical method for recognition of double-helical DNA. However, oligonucleotide-directed sequence-specific recognition of double-helical DNA is sensitive to temperature,

salt, pH, length, and sequence composition (Moser & Dervan, 1987). Ultimately, successful utilization of the triple-helix motif for targeting endogenous sequences within megabase DNA will be more fully realized by understanding those factors influencing the stabilities of local triple-helical complexes. Clearly, the match between bases involved in primary hydrogen-bonding interactions is important (Moser & Dervan, 1987; Praseuth et al., 1988; Griffin & Dervan, 1989; Belotserkovskii et al., 1990). On the basis of current understanding of sequence composition effects on the stability of double-helical nucleic acids (Tinoco et al., 1971, 1973; Uhlenbeck et al., 1973; Nelson et al., 1981; Freier et al., 1984; Breslauer et al., 1986), one would anticipate that interactions between nearest neighbor pairs of base triplets would also influence the stability of triple-helical complexes. To address this issue, we characterize here the influence of flanking triplets on the stability of the natural triplets T-AT, C+GC, and G-TA within the pyrimidine triple-helix motif. In addition, we examine the stabilities of triplets containing nonnatural bases (D_3 -TA and

D_3 -CG) with regard to nearest neighbor interactions.

EXPERIMENTAL PROCEDURES

Materials. Automated syntheses of all oligonucleotides were performed on an ABI 380B DNA synthesizer using β -cyanoethyl phosphoramidite chemistry (Gait, 1984). The phosphoramidite derivatives of the nucleoside analogues, T* and D_3 , were prepared according to published procedures (Dreyer & Dervan, 1985; Griffin, 1990). Oligodeoxyribonucleotides (9a-16a, 9b-16b, 25a-28a, 25b-28b, 33a-36a, 33b-36b) were deprotected under standard conditions using ammonium hydroxide. Oligodeoxyribonucleotides containing D_3 or T*, 1-8, 17-24, 29-32 were deprotected with 0.1 N NaOH at 55 °C, for 24 h (Dreyer & Dervan, 1985, Griffin et al., 1991). All oligonucleotides were purified by electrophoresis on 15% or 20% denaturing polyacrylamide gels and then dialyzed exhaustively against water. After dialysis, the oligonucleotide solutions were passed through a 0.45- μ m Centrex filters (Schleicher and Schuell). Oligonucleotide concentrations were determined from absorbance at 260 nm (A_{260}) using extinction coefficients ($M^{-1} \text{ cm}^{-1}$) of 8700 (T and T*), 7400 (C), 15 400 (A), 11 700 (G), and 20 300 (D_3). For the preparation of the duplex targets, oligonucleotides 9b-16b, 25b-28b, and 33b-36b were labeled at the 5' end with T4 polynucleotide kinase (Boehringer-Mannheim) and [γ - 32 P]ATP (Amersham) and then hybridized to their Watson-Crick complement. The labeled duplexes were then purified on nondenaturing 15% polyacrylamide gels. The DNA was eluted from the crushed gel with water and then subjected to ethanol precipitation and wash. The 32 P-end-labeled duplexes were then resuspended in water and used in affinity cleaving studies.

Affinity Cleaving Reactions. The cleaving reactions were executed in a total volume of 20 μ L with final concentrations of each species as indicated. A mixture of the oligonucleotide-EDTA (1 μ M) and ferrous ammonium sulfate (25 μ M) was added to the 32 P-end-labeled duplex in a solution of Tris-acetate (25 mM), pH 7.4, NaCl (10 mM), sonicated calf thymus DNA (Pharmacia) (100 μ M bp), and 35% ethanol. To this mixture was added a solution of spermine tetrahydrochloride (Aldrich) (1 mM, pH 7.4), and the oligonucleotide was allowed to equilibrate with the DNA duplex target at 37 °C for 1 h. The reactions were initiated by the addition of dithiothreitol (3 mM) and allowed to proceed at 37 °C for 6 h. Termination of the reactions was accomplished by freezing followed by lyophilization. The residue was resuspended in 5 μ L of Tris-borate-EDTA (TBE) buffer and 80% formamide solution. The 32 P-labeled products were separated by 20% denaturing polyacrylamide gel electrophoresis.

Quantitation. Relative cleavage efficiencies were determined by quantitation on a Molecular Dynamics 400S PhosphorImager. Gels were exposed to phosphor-storage screens for 24 h. Cleavage bands were quantitated by drawing a rectangle around the band at high magnification (2 \times or 4 \times) and then integrating by volume using the ImageQuant program. The same size rectangle was used for each cleavage band on a particular gel. The percent cleavage was obtained by dividing the value of the cleavage band by the integrated volume of the entire lane.

RESULTS AND DISCUSSION

The affinity cleaving method was employed to study the effect of altering the neighboring base triplets around a central base triplet (Moser & Dervan, 1987; Griffin & Dervan, 1989). Several 30 bp DNA duplexes were synthesized that contain

15 bp triple-helix target sequences that vary in the sequence composition of the central three base pairs. The formation of the different triple helices was studied with oligonucleotides containing the DNA-cleaving moiety T* (thymidine-EDTA) (Dreyer & Dervan, 1985; Moser & Dervan, 1987). The triple-helical complexes were formed with duplexes that had been labeled with 32 P. Cleavage by oligonucleotide-EDTA-Fe was initiated by treatment with the reducing agent dithiothreitol. The products of the cleavage reaction were then separated on high-resolution polyacrylamide gels. Because the EDTA-Fe moiety generates a diffusible oxidant yielding a Gaussian, sequence-independent, cleavage pattern, the amount of cleavage observed is assumed to be proportional to the fractional occupation of the target site by the oligonucleotide-EDTA-Fe. Therefore, the relative stabilities of various triple helices under a defined set of conditions were assessed by comparing the intensities of the cleavage patterns.

Base Triplets T-AT and C+GC. A series of DNA duplexes was designed to study how flanking triplets (C+GC and T-AT) might affect the stabilities of T-AT and C+GC (Figure 2A). There are formally four different pairs of nearest neighbors to the 5' and 3' side of each triplet, (T-AT, T-AT; T-AT, C+GC; C+GC, T-AT; C+GC, C+GC). For this study, directionality to the 5' or 3' side refers to the purine Watson-Crick strand. We find that the T-AT base triplet is relatively insensitive to nearest neighbor triplets (Figure 2B,C). For the C+GC base triplet, there are differences in stability depending on nearest neighbor triplets. A C+GC triplet is less stable as the number of nearest neighbor C+GC triplets increases (Figure 2B,C). This effect might be the result of destabilizing electrostatic interactions for contiguous C+GC triplets if each of the cytosines in the third strand is protonated in the complex. An alternative explanation is that not all cytosines in the third strand are protonated at pH 7.4 and the stability of a C-GC base triplet with one hydrogen bond is less than that of a T-AT interaction.

Base Triplets G-TA, D_3 -TA, and D_3 -CG. It was of particular interest to investigate the G-TA, D_3 -TA, and D_3 -CG triplets, as two-dimensional hydrogen-bonding patterns are likely oversimplified (Figure 1) (Griffin, 1990). Undoubtedly, nearest neighbors contribute importantly to the stabilization of these interactions. For example, a G-TA triplet was found to be stable when flanked by T-AT triplets on both sides (Griffin & Dervan, 1989). In contrast, studies of H-form triple-helical DNA at pH 4.17 reveal that triple-helical complexes containing a G-TA base triplet with a C+GC base triplet in the adjacent 5' position was destabilized relative to alternative mismatched triplets (Belotserkovskii et al., 1990). Recent NMR structural studies of the G-TA triplet (pH 4.85, 5 °C) by Patel and co-workers are consistent with the model of Griffin and Dervan (1991). For a G-TA triplet flanked by T-AT triplets, guanosine is in an anti orientation and pairs through a single hydrogen bond from one of its 2-amino protons to the 4-carbonyl group of thymidine in the Watson-Crick TA base pair (Radhakrishnan et al., 1991a). A set of NOEs between adjacent triplets establishes that the G-TA triplet stacks between the flanking T-AT triplets in the G-TA triplex (Radhakrishnan et al., 1991a). Assignment of which of the 2-amino protons of G pairs to the 4-carbonyl of T will require more detailed direct studies by NMR (Figure 1).

The effects of altering the flanking sequence around a G-TA triplet were examined (Figure 3). The I-TA triplet was included as a reference (Figure 3A) because inosine lacks the NH_2 group of guanine assumed to be important in the G-TA interaction (Griffin & Dervan, 1989). These studies demon-

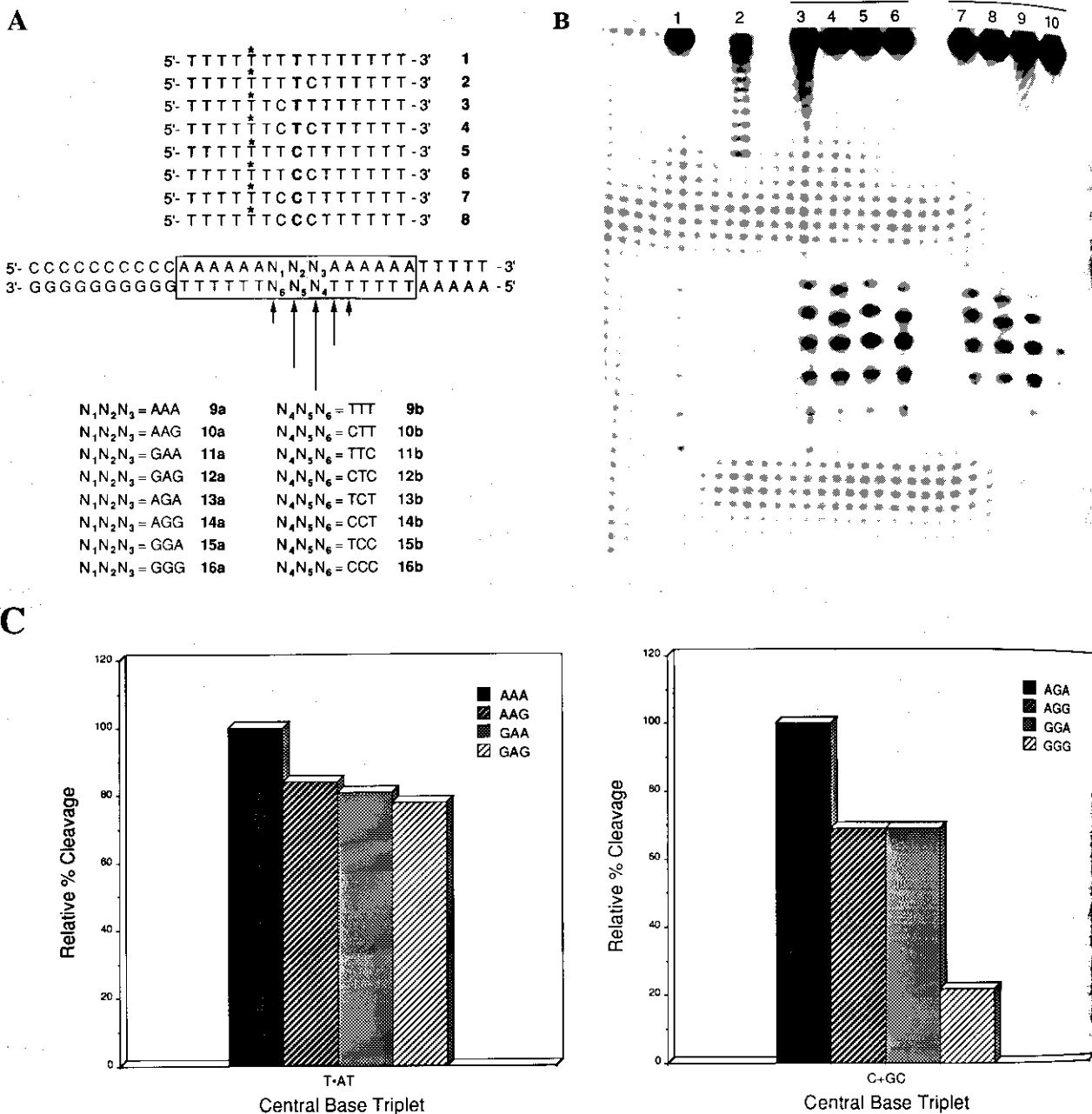


FIGURE 2: (A) (above) Sequences of oligonucleotide-EDTAs 1-8 where T* represents the thymidine-EDTA. The oligonucleotides differ at the central three positions, and the central nucleotide is indicated in bold type. (center) The box indicates the duplex region bound by the oligonucleotide-EDTAs. The arrows depict the general cleavage pattern observed. (below) Sequences of duplex DNA targets 9-16. (B) Autoradiogram of the 20% denaturing polyacrylamide gel obtained from the reactions of oligonucleotides 1-8 with duplexes 9-16. The cleavage reactions were carried out as described. Lane 1, intact 5'-end-labeled DNA after subjection to the reaction conditions in the absence of oligonucleotide-EDTA. Lane 2, products G+A chemical sequencing reactions (Maxam & Gilbert, 1980). Lanes 3-10, cleavage products obtained from reaction of oligonucleotide-EDTA-Fe 1-8 with duplexes 9-16: lane 3, 1 + 9; lane 4, 2 + 10; lane 5, 3 + 11; lane 6, 4 + 12; lane 7, 5 + 13; lane 8, 6 + 14; lane 9, 7 + 15; lane 10, 8 + 16. (C) (left) Bar graphs representing the relative ratios of cleavage obtained from triple-helical complexes containing a central T-AT triplet within the Watson-Crick sequences 5'-AAA-3', 5'-AAG-3', 5'-GAA-3', and 5'-GAG-3' and (right) containing a central C+GC triplet within the Watson-Crick sequences 5'-AGA-3', 5'-AGG-3', 5'-GGA-3', and 5'-GGG-3'. The graphs depict the average of four separate experiments analyzed by PhosphorImaging. The data are reproducible within 11% of the reported values (standard error with a 95% confidence).

strate that the G-TA interaction is influenced by its nearest neighbor triplets. A G-TA triplet with a C+GC neighbor in either the 3' or 5' direction causes a decrease in stability (Figure 3B,C). The adjacent 5' C+GC neighbor exerts a greater destabilizing influence. The stability of the I-TA triplet follows a similar trend. Accommodation of a purine base in the third strand of the pyrimidine motif may have several ramifications. For instance, there may be a steric bias that causes the base to be tilted out of the plane of the base pair toward the 5' direction, or the energy derived from base stacking differs.

The same target duplexes were used to study the nonnatural D₃TA triplet (Figure 3). In contrast to the case of G-TA, substitution of a C+GC base triplet in the adjacent 5' direction does not affect binding, but in this case a C+GC neighbor in the 3' direction results in diminished stability of the triple helix (Figure 3B,C). The experiments demonstrate opposite neighbor effects with regard to stabilization of G-TA and D₃TA triplets (Table I). With this information, most sequences containing single TA base pairs within a purine tract can be targeted using either G or D₃. The D₃-CG triplet was also examined in the same manner (Figure 3). As with the

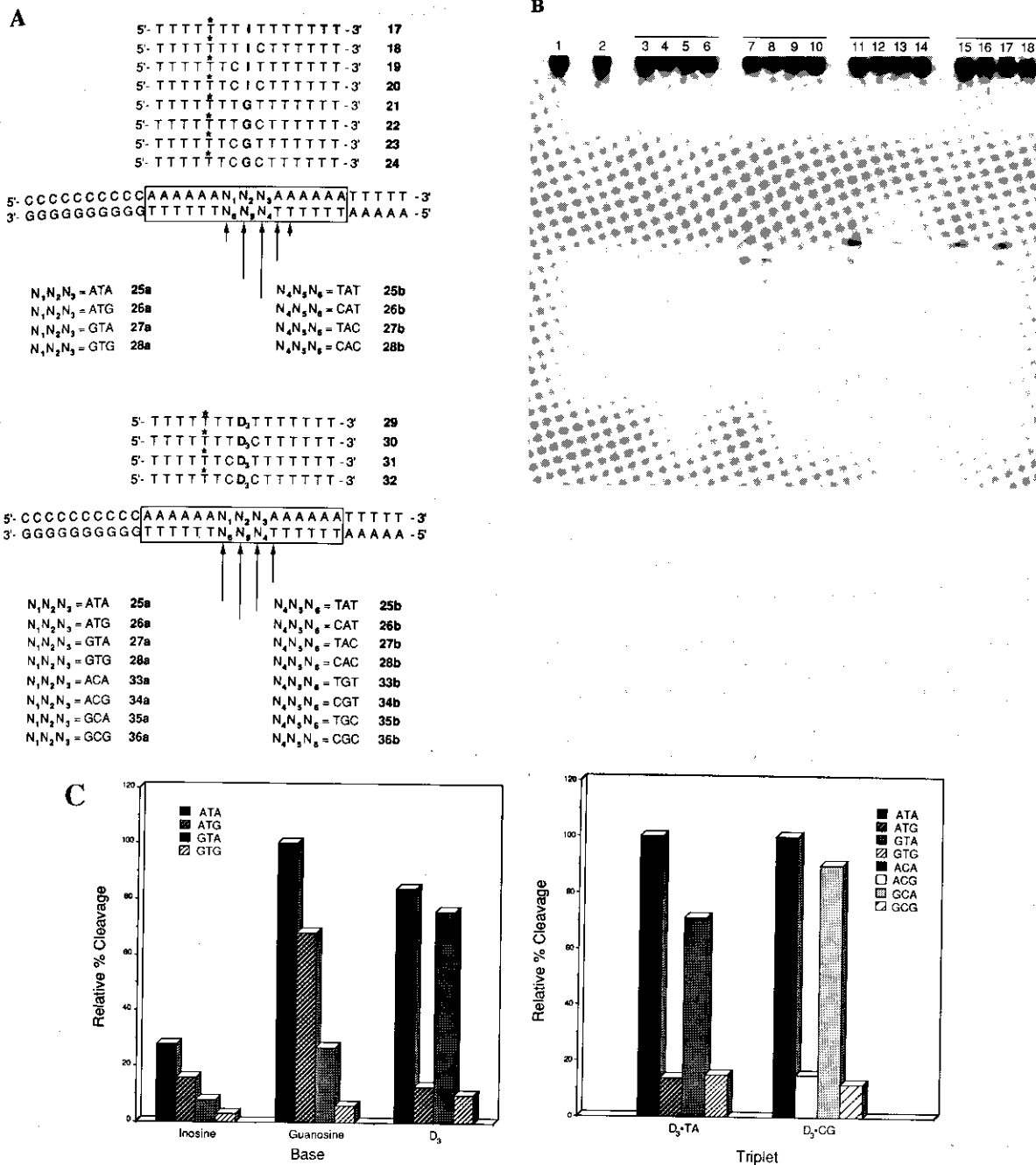


FIGURE 3: (A) (above) Sequences of oligonucleotide-EDTAs 17-24 where I is inosine and T* is thymidine-EDTA. The oligonucleotides differ at the central three positions, and the central nucleotide is indicated in bold type. The box indicates the duplex region bound by the oligonucleotide-EDTAs. The arrows depict the general cleavage pattern observed. Duplex DNA targets are formed from oligonucleotides 25-28. (below) Sequences of oligonucleotide-EDTAs 29-32 where D₃ is a nonnatural base (Figure 1) and T* is thymidine-EDTA. The oligonucleotides differ at the central three positions, and the central nucleotide is indicated in bold type. The box indicates the duplex region bound by the oligonucleotide-EDTAs. The arrows depict the general cleavage pattern observed. Duplex DNA targets are formed from oligonucleotides 25-28 and 33-36. (B) Autoradiogram of the 20% denaturing polyacrylamide gel obtained from the reactions of oligonucleotides 17-24 and 29-32 with duplexes 25-28 and 33-36. The cleavage reactions were carried out as described: lane 1, intact 5'-end-labeled DNA after subjected to the reaction conditions in the absence of oligonucleotide-EDTA; lane 2, products G+A chemical sequencing reactions (Maxam & Gilbert, 1980); lanes 3-18, cleavage products obtained from reaction of oligonucleotides 17-24 and 29-32; lane 3, 17 + 25; lane 4, 18 + 26; lane 5, 19 + 27; lane 6, 20 + 28; lane 7, 21 + 25; lane 8, 22 + 26; lane 9, 23 + 27; lane 10, 24 + 28; lane 11, 29 + 25; lane 12, 30 + 26; lane 13, 31 + 27; lane 14, 32 + 28; lane 15, 29 + 33; lane 16, 30 + 34; lane 17, 31 + 35; lane 18, 32 + 36. (C) (left) Bar graphs representing the relative ratios of cleavage obtained from triple-helical complexes containing central I-TA, G-TA, and D₃-TA triplets within the Watson-Crick sequences 5'-ATA-3', 5'-ATG-3', 5'-GTA-3', and 5'-GTG-3'. The graphs depict the average of four separate experiments analyzed by PhosphorImaging. The data are reproducible within 10% of the reported values (standard error with a 95% confidence). (right) Bar graph representing the relative ratios of cleavage obtained from triple-helical complexes containing a central D₃-TA triplet within the Watson-Crick sequences 5'-ATA-3', 5'-ATG-3', 5'-GTA-3', and 5'-GTG-3' and a central D₃-CG triplet within the Watson-Crick sequences 5'-ACA-3', 5'-ACG-3', 5'-GCA-3', and 5'-GCG-3'. The graphs depict the average of four separate experiments analyzed by PhosphorImaging. The data are reproducible within 10% of the reported values (standard error with a 95% confidence).

D₃-TA interaction, a C+GC triplet neighbor on the 3' side leads to a decrease in triple-helix stability (Figure 3B,C). This suggests that the binding affinity of D₃ for TA and CG base

pairs may be derived from similar interactions. Because the array of hydrogen-bond donor and acceptor groups is distinct for each base pair, it is likely that, in both cases, binding is

Table I: Guidelines for Targeting Purine Tracts Containing TA or CG Base Pairs Utilizing G·TA, D₃·TA, or D₃·CG Triplets Depending on Flanking Sequence

target sequence	G	D ₃
5'-ATA-3'	+++	+++
ATG	++	+
GTA	+	+++
GTG	-	-
5'-ACA-3'	-	+++
ACG	-	+
GCA	-	+++
GCG	-	-

a result of favorable van der Waals interactions with the TA or CG and its neighboring base pairs. Therefore, the non-natural base D₃ likely exhibits a type of sequence-dependent shape selectivity (Griffin, 1990).

In conclusion, it has been found that, within the pyrimidine triple-helix motif, the stabilities of C+GC are more sensitive to nearest neighbors than T+AT. The effects observed for triple-helical complexes containing G·TA, D₃·TA, and D₃·CG triplets allow the establishment of a set of guidelines for targeting purine sequences containing pyrimidine base pairs (Table I). In addition, these studies provide some insight into the interactions stabilizing the base triplets comprised of the nonnatural base D₃. The bases G and D₃ likely recognize local structural features of double-helical DNA larger than one base pair. This mode of recognition offers new possibilities for the targeting of mixed sequences by oligonucleotide-directed triple-helix formation.

ACKNOWLEDGMENTS

We are grateful to Scott Singleton for helpful discussions and to the American Cancer Society for a postdoctoral fellowship to L.L.K.

REFERENCES

- Beal, P. A., & Dervan, P. B. (1991) *Science* 251, 1360-1363.
- Belotserkovskii, B. P., Veselkov, A. F., Filippov, S. A., Dobrynin, V. N., Mirkin, S. M., & Frank-Kamenetskii, M. D. (1990) *Nucleic Acids Res.* 18, 6621-6624.
- Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 3746-3750.
- Broitman, S. L., Im, D. D., & Fresco, J. R. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 85, 3781.
- Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., & Hogan, M. E. (1988) *Science* 241, 456.
- de los Santos, C., Rosen, M., & Patel, D. (1989) *Biochemistry* 28, 7282-7289.
- Dreyer, J. B., & Dervan, P. B. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 968.

- Felsenfeld, G., Davies, D. R., & Rich, A. (1957) *J. Am. Chem. Soc.* 79, 2023-2024.
- Freier, S. M., Petersheim, M., Hickey, D. R., & Turner, D. H. (1984) *J. Biomol. Struct. Dyn.* 1, 1229-1242.
- Gait, J. J., Ed., (1984) *Oligonucleotide Synthesis*, IRCC Press, Oxford, and references cited therein.
- Griffin, L. C. (1990) Ph.D. Thesis, California Institute of Technology, Pasadena, CA.
- Griffin, L. C., & Dervan, P. B. (1989) *Science* 245, 967-971.
- Hoogsteen, K. (1959) *Acta Crystallogr.* 12, 822-823.
- Horne, D. A., & Dervan, P. B. (1990) *J. Am. Chem. Soc.* 112, 2435-2437.
- Kohwi, Y., & Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 3781.
- Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V., & Fresco, J. R. (1988) *Biochemistry* 27, 9108.
- Lipsett, M. N. (1963) *Biochem. Biophys. Res. Commun.* 11, 224.
- Lipsett, M. N. (1964) *J. Biol. Chem.* 239, 1256-1260.
- Mark, C., & Thiele, D. (1978) *Nucleic Acids Res.* 5, 1017.
- Michelson, A. M., Massoulie, J., & Guschlbauer, W. (1967) *Prog. Nucleic Acid Res. Mol. Biol.* 6, 83.
- Morgan, A. R., & Wells, R. D. (1968) *J. Mol. Biol.* 37, 63-80.
- Moser, H. E., & Dervan, P. B. (1987) *Science* 238, 645-650.
- Nelson, J. W., Martin, F. J., & Tinoco, I., Jr. (1981) *Bio-polymers* 20, 2509-2531.
- Pilch, D. S., Brousseau, R., & Shafer, R. H. (1990) *Nucleic Acids Res.* 18, 5743-5750.
- Pilch, D. S., Levenson, D., & Shafer, R. H. (1991) *Biochemistry* 30, 6081-6087.
- Radhakrishnan, I., Gao, X., de los Santos, C., Live, D., & Patel, D. J. (1991a) *Biochemistry* 30, 9022-9030.
- Radhakrishnan, I., de los Santos, C., & Palet, D. J. (1991b) *J. Mol. Biol.* 221, 1403-1418.
- Rajagapol, P., & Feigon, J. (1989a) *Nature* 339, 637-640.
- Rajagapol, P., & Feigon, J. (1989b) *Biochemistry* 28, 7859-7870.
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning, a Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Strobel, S. A., & Dervan, P. B. (1990) *Science* 249, 73-75.
- Strobel, S. A., & Dervan, P. B. (1991) *Nature* 350, 172-174.
- Tinoco, I., Jr., Uhlenbeck, O. C., & Levine, M. D. (1971) *Nature* 230, 362-367.
- Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature* 246, 40-41.
- Uhlenbeck, O. C., Borer, P. N., Dengler, B., & Tinoco, I., Jr. (1973) *J. Mol. Biol.* 273, 483-496.