Molecular Recognition of 5-Amino-dU in the Central Strand of a DNA Triplex: Formation of Triads A*U#:A in Parallel and G*U#:A in Antiparallel Motif[†,‡]

Vipul S. Rana, Dinesh A. Barawkar, and Krishna N. Ganesh*

Division of Organic Synthesis, National Chemical Laboratory, Pune 411008, India

Received February 1, 1996

One of the most exciting developments in DNA molecular recognition with a direct potential for practical therapeutics is the discovery of triple-stranded DNA complexes.[1] These arise from a major groove binding of either pyrimidine (Y) or purine (R) rich oligonucleotides in parallel (*p*) or antiparallel (*ap*) orientation, respectively, to polypurine stretches of Watson−Crick (WC) double-stranded DNA. The specificity in triplex formation is derived from third-strand Hoogsteen (HG) hydrogen bonding by which thymine recognizes the AT base pair (T*A:T triad) and protonated cytosine, a GC base pair (C^{+}*G:C triad) in the "pyrimidine" motif (Y*R:Y). Adenine recognition of AT (A*A:T) and guanine recognition of GC (G*G:C) occurs by reverse HG mode in the "purine" (R*R:Y) motif.[1,2] A common feature to both of these well-established motifs is the necessity of a purine (A or G) in the central position of triplex triads, since only these provide two sets of hydrogen bond donors/acceptors in the major groove of double helix. Pyrimidine bases devoid of this feature are not generally compatible in the middle position and lead to decreased triplex stability from HG mismatches.[3] Among the eight possible triads with T or C in the middle, only G*T:A and T*C:G are accommodated with reasonable stability within the established motifs.[4] This limitation of triplex formation has led to exploration for new triad combinations involving unnatural base components[5] that sterically and electronically complement to recognize T and C of the base pairs TA and CG when located in the third position of the triad. Alternatively, pyrimidines can be engineered to endow dual recognition properties for placement as a central base of triplex triads, as in pseudonucleobases, ΨU, ΨC, and ΨiC that possess extra hydrogen bonding sites in the major groove of the derived WC type duplex.[6]

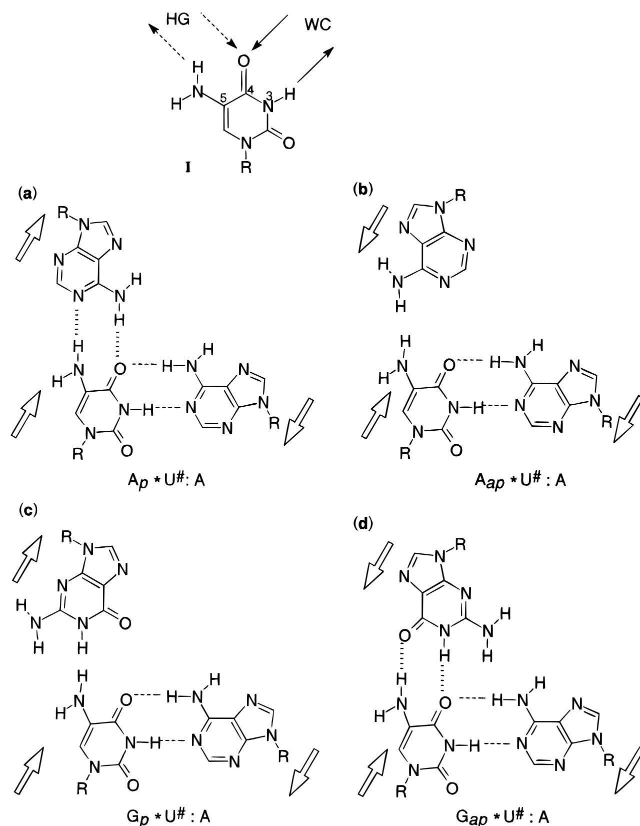

Figure 1. Schematic representation of base triads A*U#:A (a, b) and G*U#:A (c, d). Arrows indicate 5′ → 3′ direction of phosphodiester backbone: (- - -) WC base pairs, (|||) HG-type base pairs.

We envisaged that the pyrimidine derivative 5-amino-dU (U#) (**I**), would be suitable as the middle base of a triplex triad, since it has electronic requirements for simultaneous recognition of complementary bases of triad. The availability of one acceptor (O4 carbonyl) and a donor (5-amino) in U# provides orientation specificity for recognition of the U#:A duplet by a third base in the major groove. In this paper, we report the recruitment of U# in the generation of inverted triplex triads R*U#:A (R = A or G) and demonstrate a novel specificity in recognition of A or G that is dependent on the orientation of the unmodified HG strand. The rationale is derived from Figure 1, which reveals that accommodation of U# in the middle strand of the established pyrimidine and purine motifs is possible only when the HG strand containing A is parallel (A*U#:A) (Figure 1a) and that with G is antiparallel (G*U#:A) (Figure 1d) to U# in central strand. In order to probe such a novel molecular recognition of U# in triplexes, the oligodeoxynucleotide (ODN) sequences **1**−**7** were designed to combinatorially generate four triplexes that differ with respect to third-strand orientation and a base within the single triad site X*Y:Z. The oligonucleotide **6** having T in place of U# in **5** was used to constitute relevant control triplexes.

The ODNs **1**−**7** were prepared by standard procedures as reported before[7] and their homogeneity established by HPLC. UV melting experiments indicated formation of duplex **7:5** (T_m = 61 °C) with U# containing ODN with slight destabilization (ΔT_m = 1.5 °C) relative to the duplex

* To whom correspondence should be addressed. Fax: 091 (212) 335153. E-mail: kng@ncl.ernet.in.

† Contribution No. 6343 from NCL.

‡ Abbreviations: X*Y:Z indicates the three strands of triplex in which * and : represent Hoogsteen and Watson−Crick hydrogen bonding patterns, respectively.

(1) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *15*, 7749. (b) Thoung, N. T.; Helen,C. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 666. (c) Crooke, S. T.; Lebleu, B. *Antisense Research and Application*; CRC: Boca Raton, Ann Arbor, London, Tokyo, 1993.

(2) (a) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360. (b) Maher, L. J., III. *BioEssays* **1992**, *14*, 807. (c) Radhakrishnan, I.; Patel, D. J. *Biochemistry* **1994**, *33*, 11405.

(3) (a) Roberts, R. W.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9397. (b) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *32*, 2999. (c) Best, G. C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 1187. (d) Greenberg, W. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 5016.

(4) (a) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967. (b) Yoon, K.; Hobbs, C. A.; Koch, J.; Sardaro, M.; Kutny, R.; Wies, A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3840. (c) Wang, E.; Malek, S.; Feigon, J. *Biochemistry* **1992**, *31*, 4838. (d) Dittrich, K.; Gu, J.; Tinder, R.; Hogan, M.; Gao, X. *Biochemistry* **1994**, *33*, 4111.

(5) (a) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976. (b) Krawczyk, S. H.; Milligan, J. F.; Wadwani, S.; Moulds, C.; Froehler, B. C.; Matteucci, M. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3761. (c) Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Fiegon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908. (d) Huang, C.-Y.; Cushman, C. D.; Miller, P. S. *J. Org. Chem.* **1993**, *58*, 5048. (e) Stilz, H. U.; Dervan, P. B. *Biochemistry* **1993**, *32*, 2177. (f) Durland, R. H.; Rao, T. S.; Bodepudi, V.; Seth, D. M.; Jayaraman, K.; Revanker, G. R. *Nucl. Acids Res.* **1995**, *23*, 647.

(6) (a) Trapane, T. L.; Christopherson, M. S.; Roby, C. D.; Ts'o, P. O. P.; Wang, D. J. *Am. Chem. Soc.* **1994**, *116*, 8412. (b) Trapane, T. L.; Ts'o, P. O. P. *J. Am. Chem. Soc.* **1994**, *116*, 10437. (c) Bandaru, R.; Hashimoto, H.; Switzer, C. *J. Org. Chem.* **1995**, *60*, 786.

(7) Barawkar, D. A.; Ganesh, K. N. *BioMed. Chem. Lett.* **1993**, *3*, 347.

S0022-3263(96)00207-1 CCC: $12.00 © 1996 American Chemical Society

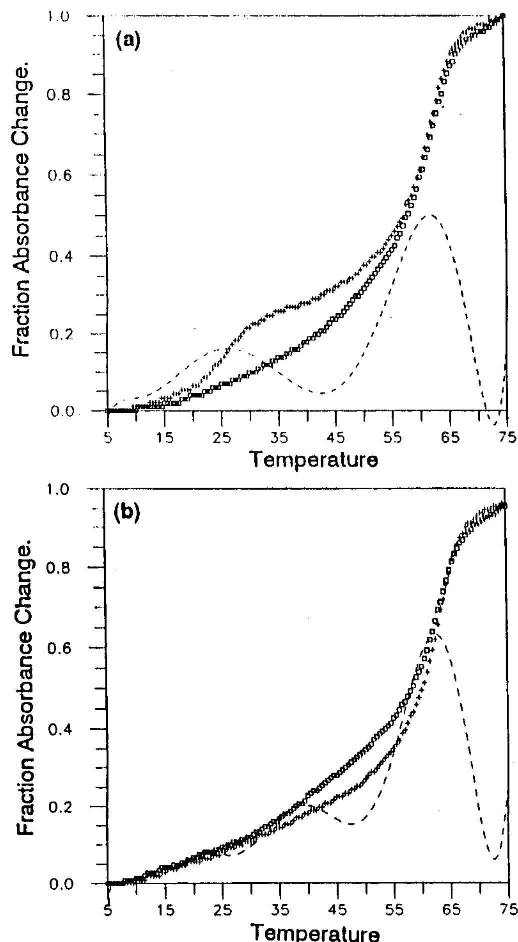

Figure 2. UV absorbance (260 nm) vs temperature profiles for (a) A*U#:A and (b) G*U#:A triads: (+) third strand in parallel orientation, (□) the same in antiparallel orientation, (−) first-derivative plot of parallel strand in (a) and that of antiparallel in (b).

Table 1. UV−T_m of Triplexes

Entry	Triad X*Y:Z	Triplex (Orientation)	Triplex Tm (°C) pH 5.8	Triplex Tm (°C) pH 7.0	Duplex Tm (°C)
1	G*T:A	**2*6:7** (*p*)	25	26	63
2	A*T:A	**1*6:7** (*p*)	24	nd	63
3	A*U#:A	**1*5:7** (*p*)	28	nd	62
4	G*U#:A	**4*5:7** (*ap*)	35	37	62

p: Parallel motif; *ap*: Antiparallel motif; nd: Not detected. No triplex formation was observed with G*T:A in **4*6:7** (*ap*), A*T:A in **3*6:7** (*ap*), A*U#:A **3*5:7** (*ap*), and G*U#:A in **2*5:7** (*p*) under the above conditions.

1. Xp = A; 2. Xp = G 5' -T T C T T T T **Xp** T T C T T C T - 3'
3. Xap = A; 4. Xap = G 3' - A A G A A A A **Xap** A A G A A G A - 5'
5. Y = U#; 6. Y = T 5' -C G A A G A A A A **Y** A A G A A G A C C - 3'
7. Z = A 3' -G C T T C T T T T **Z** T T C T T C T G G - 5'

(**7:6**) with T at the same position. The triplexes containing U# in the central strand and the corresponding control triplexes with T were individually constituted by heating equimolar amounts of the appropriate three strands at 80 °C for 3 min followed by slow cooling in buffer, 100 mM sodium cacodylate containing 20 mM $MgCl_2$ and 1 M NaCl at pH 5.8. The stability of all triplexes was measured by UV melting experiments in which biphasic transitions characteristic of triplex formation were seen (Figure 2) and supported by the mixing curves (not shown), which indicated a 1:1 stoichiometry of duplex and third strand in triplex.

The UV−T_m data obtained for different triplexes at pHs 5.8 and 7.0 are shown in Table 1. At pH 5.8 stable triplex formation was observed for G*T:A triad in triplex (**2*6:7**) within the parallel pyrimidine motif (Table 1, entry 1) but not in the antiparallel purine motif (**4*6:7**). The triplex (**1*6:7**) with triad A*T:A (Table 1, entry 2) was slightly less stable than the triplex (**2*6:7**) with G*T:A triad in the pyrimidine motif in agreement with the literature.[3,4] In comparison, the parallel triplex (**1*5:7**) having triad A*U#:A with a modified base in the central strand (Table 1, entry 3) had a higher UV−T_m compared to the corresponding control triplex A*T:A (**1*6:7**) in pyrimidine motif; no triplex was detected in the corresponding antiparallel mode (Table 1, footnote). In the case of triad G*U#:A, triplex formation was seen only in the antiparallel orientation (**4*5:7**) (Table 1, entry 4), but not in the parallel form (**2*5:7**) (Table 1, footnote). The G:U#:A antiparallel triplex exhibited a higher stability than the G*T:A triplex (**2*6:7**) seen in the parallel motif (Table 1, entry 1). Thus, triplexes with modified base U# in central strand (**1*5:7** and **4*5:7**) exhibited not only a higher UV−T_m stability compared to corresponding control T analogues, but also displayed a remarkable orientation selectivity in third-strand recognition. Only marginal differences ($\Delta T_m = 1-2$ °C) for duplex T_m in triplexes was noticed among T and U# triplexes and the T_m values, close to that of duplex alone, were not influenced much by the third base of the triad. The fraction absorbance change for triplex melting in the antiparallel motif (Figure 2b) was also much less than that in the parallel motif (Figure 2a) as expected from a poorer base stack[2c] in the former as compared to the latter. The pyrimidine motif triplexes containing a A*U#:A triad and a A*T:A triad were formed only at pH 5.8 and not detected at pH 7.0. In contrast, the antiparallel purine motif triplex (**4*5:7**) devoid of the base C in third strand was observed in both pH ranges, with a slightly higher stability at pH 7.0 (T_m = 37 °C) compared to that at 5.8 (T_m = 35 °C).

The above experimental results are in accordance with the hydrogen-bonding scheme shown in Figure 1 and implies a novel molecular recognition of U#: *U# of WC base pair U#:A recognizes third strand A only in the parallel motif and G recognition occurs in the antiparallel motif.* The antiparallel purine motif, formed only with G*U#:A but not with G*T:A, seems to be extremely sensitive to the presence of U#. It may be pointed out that hydration sites in the Crick−Hoogsteen groove of a triplex are important determinants for stability in antiparallel purine motif.[2c] The replacement of hydrophobic 5-methyl group in T by hydrophilic 5-amino function as in U# may have vital consequences, since the amino group can favorably participate in the hydration network to offer crucial stability for stabilization of the G*U#:A triad in the antiparallel mode. Preliminary molecular modeling points to nonparticipation of N7 of purine in recognition of U# in the central strand since the derived triad geometries are not isomorphous with the established geometries of T*A:T, A*A:T, and G*G:C triplexes,[3c,d] leading to strained backbone.

The use of modified ψ bases in central strand of triplexes allows formation of triple-stranded helices at single strand target sites of unrestricted sequence employing two oligonucleotide probes, one of which contains modified pyrimidines.[6] In this context, successful application of a simple pyrimidine derivative 5-amino dU in second strand for selective triplex formation as reported here adds a new repertoire to nucleic acid recognition. Further work is underway to examine the sequence context effects and other recognition tolerants in U# triads.

Acknowledgment. V.S.R. and D.A.B. acknowledge CSIR (New Delhi) for fellowships. K.N.G. thanks the Jawaharlal Nehru Center for Advanced Scientific Research, Bangalore, of which he is a Senior Honorary Fellow.

Supporting Information Available: UV mixing and melting curves and general comments (5 pages).

JO9602074