J|A|C|S
COMMUNICATIONS

Published on Web 09/11/2008

A Janus-Wedge DNA Triplex with A-W_1-T and G-W_2-C Base Triplets

Han Chen, Meena, and Larry W. McLaughlin*

Department of Chemistry, Merkert Chemistry Center, Boston College, 2609 Beacon Street, Chestnut Hill, Massachusetts 02467

Received June 17, 2008; E-mail: mclaughl@be.edu

We describe here the development of a Janus-Wedge (J-W) triple helix involving target A-T and G-C base pairs. Each base triplet is formed from a target base pair and a third residue (a wedge residue) capable of hydrogen bonding with the Watson−Crick faces of the base pairing partners.

Triplexes as described by Dervan and Helene are formed from a (third) DNA strand capable of binding to the Hoogsteen face of purine residues of purine−pyrimidine base pair.[1,2] With that design the third strand is bound in the major groove of duplex DNA. The Dervan/Helene approach can be very useful but is generally limited to the targeting of specific polypurine sequences[1,2] (with the formation of T-A-T and C^+-G-C base triplets). To generalize duplex targeting, a mode of recognition must be developed whereby all four possible base pairs can be targeted. Many improvements over the Dervan/Helen targeting design have been reported using derivatives that employ only a single hydrogen bond,[3] recognition of each base pair as a unit,[4] or being able to bind purines in either strand,[5] but improvements in targeting methods[6−8] have achieved only moderate successes.

The Janus-Wedge (J-W) triple helix is based upon a recognition motif first suggested by Lehn[9] in his work with heterocycles; it involves the ability of the incoming third strand to hydrogen bond with the Watson−Crick (W−C) faces of the two target strands (see Figure 1). In the present study, we have used W_1 to target A-T (or by simple rotation of the heterocycle, T-A) base pairs as we described earlier,[10] and W2 for G-C (or C-G) base pairs. We have used a neutral PNA backbone for the heterocycles but have incorporated one lysine (K) residue at the C-terminus to aid in aqueous solubility and to provide some complementary charge−charge interactions. The choice of a PNA backbone follows the observation that certain PNA sequences are known to bind DNA duplexes to form D-loop structures, a form of strand invasion that could be an important intermediate in the pathway to a J-W triplex.[11]

We previously described[10] the synthesis of the protected W_1 monomer from the readily available 2,4-diaminopyridin-6(5*H*)-one. A similar approach to the W_2 monomer (Figure 1b) from the readily available 6-amino-1,3,5-triazine-2,4(1*H*,3*H*)-dione failed, largely owing to the very poor solubility of the starting heterocycle. The solution to this synthesis was to use a cyclization reaction previously reported for this class of compounds.[12] The pathway involves the preparation of an acyclic amidinourea,[13] but that product is unstable and cyclizes to form the aminotriazenedione with the amino group protected as the CBz carbamate. We could use this same reaction using the amine-containing linker (that bridging the triazene and PNA backbone), and after cyclization couple the resulting product heterocycle to the PNA backbone to generate the desired protected monomer (Scheme 1).

In principle this derivative should be ideal for the synthesis of PNA sequences containing W_2 residues, and in fact an efficient coupling by the CBz-protected W_2 monomer to a solid-phase bound W_1 residue was observed. However, upon piperidine deprotection of the bound W_2 residue we could not detect the presence of the expected primary amine necessary to continue elongation. The initial cyclization reaction

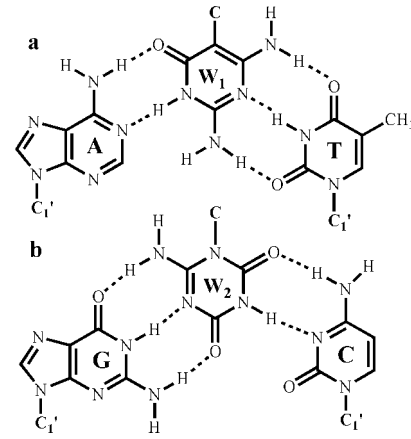
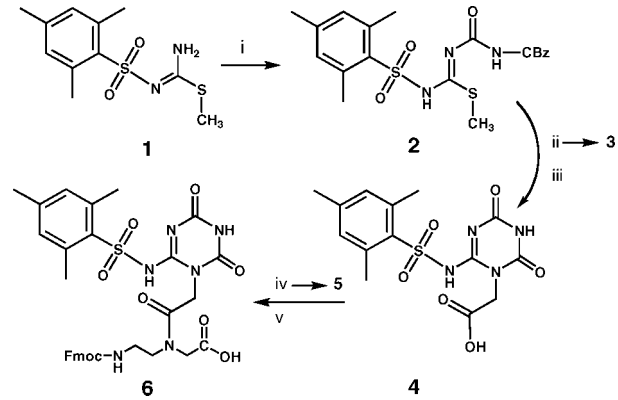


Figure 1. Janus-Wedge base triplets: (a) the third-strand residue W_1 binds to the W−C faces of the target A and T or (b) W_2 binds similarly to G and C.

to generate the aminotriazenedione ring involved attack of a secondary amine on a CBz protecting group.[13] Fmoc deprotection unmasked a primary amine; we reasoned this amine could attack the remaining CBz protecting group (forming a 10-membered ring, see Supporting Information). Mass spectral analysis confirmed this possibility. Replacing that CBz protecting group with trimethyl-phenylsulfonyl (Scheme 1) eliminated the putative side reaction, and PNA sequences containing W_1 and W_2 could be prepared.

We prepared a mixed 8-mer PNA sequence with a terminal lysine residue: $(W_1)_3W_2(W_1)_4K$, purified it (HPLC), and confirmed its mass. In previous work[10] we had also prepared the $(W_1)_8K$ PNA sequence and showed the binding orientation of the heterocycles and the polarity of the J-W PNA relative to the target strands. We

Scheme 1[a]



[a] Steps: (i) benzyloxycarbonyl isocyanate; (ii) $HCl\cdot H_2NCH_2COOtBu$, Et_3N → **3** (*t*Bu ester of **4**); (iii) TFA (iv) [2-(9*H*-fluoren-9-ylmethoxycarbonylamino)ethyl-amino]-acetic acid *t*Bu ester, HCl/EDCI/DIPEA, → **5** (tBu ester of **6**); (v) TFA.

10.1021/ja804607v CCC: $40.75

Table 1. Thermal Stabilities of J-W Triplexes

X-Y	wedge strand	T_M
T-C	$W_1W_1W_1W_1W_1W_1W_1W_1K$	45.8, 69.2 (±0.5) °C
T-A	$W_1W_1W_1W_1W_1W_1W_1W_1K$	44.9, 69.1 (±0.5)
T-C	$W_1W_1W_1W_1W_2W_1W_1W_1K$	36.4, 70.1 (±0.3)
G-C	$W_1W_1W_1W_1W_2W_1W_1W_1K$	46.9, 69.8 (±0.3)

Table 2. Thermal Stabilities of J-W Triplexes

-X X X X X X X X-	T_M
-C C C C C C C C-	46.9, 69.8 (±0.3) °C
-C A C C C C C C-	44.3, 69.1 (±0.4)
-C A C C C A C C-	34.2, 67.6 (±0.7)
-C A C C C A C A-	19.9, 67.2 (±0.7)
-A A A A C A A A-	− 78.7

also prepared two DNA sequences in which T-C was used for the target for W_1 and G-C was used for the target of W_2 (Table 1).[14]

The observed T_M values for the $(W_1)_8K$ probe and the T-C target was 45.8 and 69.2 °C. The latter transition corresponds to the duplex-random coil equilibrium and the former characterizes the triplex−duplex equilibrium. Introduction of one W_2 residue into the probe strand should represent a mismatch condition for the all T-C target, and in fact the T_M characterizing the triplex−duplex equilibrium in this case decreased by roughly 9 to 36.4 °C (Table 1). When the T-C target at the W_2 site was changed to the canonical G-C base pair, the target for W_2, the T_M value characterizing the triplex−duplex equilibrium rose by over 10 to 46.9 °C (Table 1).

It is noteworthy that placing a single W_2 residue into a sequence that targets T-C residues results in a significant decrease in T_M value and that thermal stability is recovered when the G-C target is place to interact with the W_2 residue. An examination of potential hydrogen bonding interactions for the T-W_2-C base triplet suggests the formation of an essentially isomorphic base triplet (Figure 2), but the existence of such a base triplet does not contradict the data. The measured T_M values simply indicate a loss in complex thermal stability when T-C is the target for W_2. A more destabilizing mismatch situation would likely result in complete dissociation of the triplex. The reduction in thermal stability likely arises from effects including (i) the tridentate interaction with G versus a bidentate interaction with T, (ii) reduced base stacking effects for a pyrimidine versus a purine, and (iii) unfavorable dipole−dipole interactions (Figure 2).

Replacement of various C residues in the target sequences with A residues results in binding to an increased number of A-T base pairs by the W_1 residues (Table 2).[14] The temperature for first transitions for these complexes decreased with increasing A content, likely reflecting the formation of base triplets with a total of five hydrogen bonds (see Figure 1a) in place of the T-W_1-C base triplets containing six hydrogen bonds. When the target sequence contained only A-T and G-C base pairs, the PNA 8-mer was unable to form a stable Janus-Wedge triplex. Presumably the eight-residue probe sequence is simply not long enough to effectively invade the target duplex and form the stable triplex when competing duplex formation is also possible.

The stability of the Dervan/Helene $(T\text{-}A\text{-}T)_{10}$ triplex is reported to be −6.6 kcal/mol under optimal conditions (2 M NaCl, 50 mM $MgCl_2$).[15] We have reported[10] the stability of the $(T\text{-}W_1\text{-}C)_8$ J-W triplex as −15.3 kcal/mol (50 mM NaCl, 10 mM $MgCl_2$). While this value reflects effective binding interactions through the J-W format, it will be moderated when strand invasion is required.

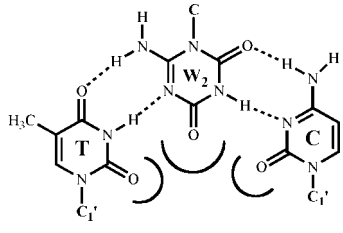


Figure 2. Proposed structure of a T-W_2-C base triplet.

Binding of the PNA sequence containing the W_2 residue to the target DNA containing a G-C base pair (Table 1) was also confirmed by nondenaturing PAGE analysis, which resulted in a K_D of 12 nM (Figure 3).

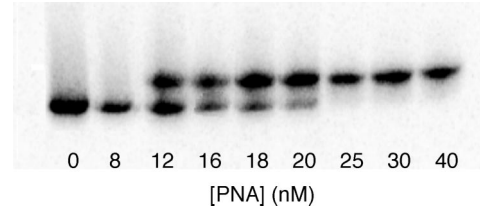


Figure 3. Gel shift assay for $(W_1)_4W_2(W_1)_3K$ and DNA (G-C, Table 1).

Initial studies with a longer $(W_1)_{14}K$ sequence targeting an A_{14}-T_{14} duplex have indicated the formation of a stable complex on the basis of thermal melting analysis and a fluorescence-based assay in which an A-T selective minor groove-binding fluorophore is displaced from the target duplex upon complex formation with the J-W strand (see Supporting Information).

To target all four base pairs we need to design derivatives that can discriminate all four base pairs. We expect that these residues will exhibit both base-pair selectivity and enhanced stability in part as the result of better base-stacking between the J-W residues.

Acknowledgment. This work was supported by the NSF (MCB 0451448).

Supporting Information Available: Synthetic schemes, procedures, and thermal analyses. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645–50.
(2) Francois, J. C.; Saison-Behmoaras, T.; Helene, C. *Nucleic Acids Res.* **1988**, *16*, 11431–40.
(3) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967–71.
(4) Gianolio, D. A.; McLaughlin, L. W. *Nucleosides Nucleotides* **1999**, *18*, 1751–69.
(5) Li, J.-S.; Fan, Y.-H.; Zhang, Y.; Marky, L. A.; Gold, B. *J. Am. Chem. Soc.* **2003**, *125*, 2084–93.
(6) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435–37.
(7) Jayasena, S. D.; Johnston, B. H. *Nucleic Acids Res.* **1992**, *20*, 5279–88.
(8) Gowers, D. M.; Fox, K. R. *Nucleic Acids Res.* **1999**, *27*, 1569–77.
(9) Branda, N.; Kurz, G.; Lehn, J. M. *Chem. Commun.* **1996**, 2443–44.
(10) Chen, D.; Fraley, A. W.; Meena; Sharma, S. K.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2004**, *126*, 70–71.
(11) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–1500.
(12) Yuan, C.; Williams, R. M. *J. Am. Chem. Soc.* **1997**, *119*, 11777–84.
(13) Yuan, C.; Williams, R. M. *Tetrahedron Lett.* **1996**, *37*, 1945–48.
(14) Errors in Tables 1 and 2 are reported for the first transitions.
(15) Pilch, D. S.; Brousseau, R.; Shafer, R. H. *Nucleic Acids Res.* **1990**, *18*, 5743–50.

JA804607V