J|A|C|S
COMMUNICATIONS

Published on Web 03/05/2002

# Changing DNA Grooves − A 1,4,5,8-Naphthalene Tetracarboxylic Diimide Bis-Intercalator with the Linker ($\beta$-Ala)$_3$-Lys in the Minor Groove

Vladimir Guelev, Steven Sorey, David W. Hoffman, and Brent L. Iverson*

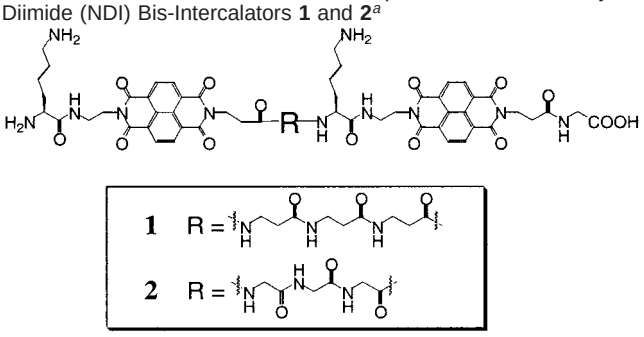
*Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78722*

Received August 13, 2001

Threading polyintercalators that occupy both DNA grooves are expected to display high DNA binding affinity and/or specificity, slow rates of DNA dissociation, and potent inhibition of protein−DNA interactions.[1] 1,4,5,8-Naphthalene diimide (NDI) monointercalators have been known to bind DNA via threading intercalation,[1b] and a series of NDI polyintercalating peptides was synthesized previously,[2] including the first known tetrakis-[2a] and octakis-[2b] intercalators (Scheme 1). A combinatorial library screen identified bis-intercalators **1** and **2** with similar linkers but different DNA binding specificities.[2d] NMR and biochemical experiments on **2** previously revealed threading intercalation with sequence-specific peptide contacts in the DNA major groove.[2e] Here we present NMR/molecular modeling data for the specific complex of **1**, in which the tris-$\beta$-alanyl linker resides in the DNA minor groove. Given the facile synthesis[2a,c] and modular nature[2a−d] of these compounds, the results reported here may enable the construction of threading polyintercalators with sequence-specific interactions in both DNA grooves.

The oligonucleotide d(CGATAAGC)·d(GCTTATCG) used in the NMR studies was selected on the basis of DNAse I footprinting experiments[2d] and the DNA alkylation pattern of an *N*-bromoacetyl derivative of **1** (Supporting Information).[3] Addition of **1** produced well-resolved $^1$H NMR spectra, consistent with a predominant 1:1 **1**-DNA complex. All internal imino protons are observed in the exchangeable 1D spectrum of the complex, suggesting intact base pairing. T3B, T6B, G2A, and G7A imino proton resonances are shifted upfield, consistent with intercalation.[4] In the NOESY spectra the H1′/H2′/H2″−H6/H8 NOE connectivities[5a,b] are weak or interrupted for the Pu−Pu (Py−Py) steps G2A−A3A (T6B−C7B) and A6A−G7A (C2B−T3B), and several DNA-NDI NOEs are observed in their place (Figure 1a), suggesting intercalation at these steps. Intermolecular NOEs between the peptide and DNA protons H1′, H4′, and Ade H2 place the linker in the minor groove. Each of the four AH2 protons has a set of strong NOEs to one distinct linker residue. The DQF-COSY spectrum of the complex suggests a trans conformation for two of the $\beta$-alanines, and a gauche conformation for the internal NDI side chains. Lys 6 displays NOEs with residues A3A and T4A. The N- and C-peptide termini are poorly defined. Only one NOE was identified, between the N-terminal ethylene linker and H8 of A6A in the major groove. The ammonium and H$\epsilon$ protons of Lys 1 are coupled in the TOCSY ($H_2O$) spectrum, suggesting hydrogen bonding, most likely to backbone phosphates. Line broadening in the COSY spectra of the complex, and the high number of overlapping DNA backbone resonances, precluded detailed analysis of DNA conformation. However, a qualitative analysis of the spectra is consistent with a B-like conformation of the internal six base pairs, with the exception of the intercalation bases C2B and T6B, that appear to be in a C2′-endo/C3′-endo equilibrium.[5b,c] B-like DNA structure was previously concluded for the complex of **2**.[2e] In the NOESY and ROESY spectra (Supporting Information), there is evidence for slow (on the NMR time scale) NDI ring flipping inside the intercalation sites, as first observed for the structurally related [mono]imide bis-intercalator LU 79553.[6]

***Scheme 1.*** Structures of the 1,4,5,8-Naphthalene Tetracarboxylic Diimide (NDI) Bis-Intercalators **1** and **2**[a]




[a] An NMR-derived model of the DNA complex of **2** (Figure 2b) was recently reported.[2e]

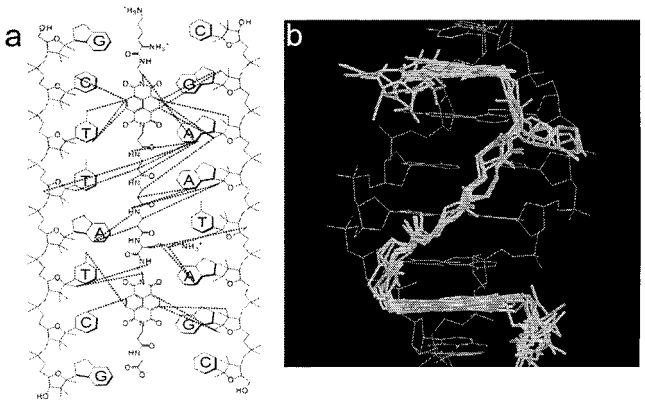



***Figure 1.*** Model of the **1**-DNA complex derived from the NMR data: (a) Observed intermolecular DNA-**1** NOEs (60 ms mixing time, 27 °C); (b) Eight superposed ligand structures satisfying the experimental restraints, calculated over a "fixed" DNA conformation (see Supporting Information).

A model of the **1**-DNA complex, obtained from restrained Molecular Dynamics (rMD) simulations in XPLOR 3.1,[7] is shown in Figures 1b and 2a. The NDI rings are intercalated at the ApG and GpA steps, in a diagonal threading geometry. The linker spans the four base pairs ATAA (TTAT) in the minor groove, following a helical path between the intercalated steps. The Lys 6 side chain is "tucked" toward strand A, in VDW contact with the hydrophobic surface of the T4A ribose.

Comparison of the complexes of **1** and **2** (Figure 2) supports the notion of an "active" linker that significantly contributes to the energetics of binding. While the tris-glycyl linker of **2** directs



**10.1021/ja016834e CCC: $22.00** 

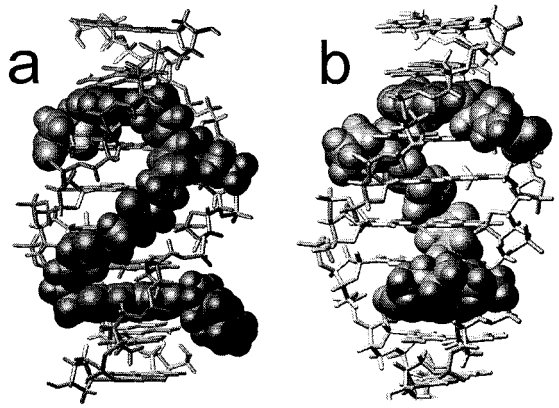



***Figure 2.*** View from the minor groove of (a) the **1**-d**(**CGATAAGC)·d(GCTTATCG) complex (reported here); (b) the **2**-d(CGGTACCG)$_2$ complex (ref 2e).

intercalation at the G−G steps of G|GTAC|C,[2e] the tris-$\beta$-alanyl compound **1** intercalates at the G−A and A−G steps, flanking an A/T-tract. The similarity in the stacking geometry for the three different Pu−Pu steps suggests that optimal stacking geometry may influence linker orientation in the DNA grooves.

Previous footprinting experiments have shown that **2** does not recognize the 5′-CGATAAGC-3′ site preferred by **1**.[2d] The path across four base pairs along the same strand of the minor groove is longer than a diagonal path via the major groove, suggesting that the greater length of the linker in **1** is partially responsible for recognition in the minor groove.

In addition, the linker of **1** is relatively hydrophobic, containing 10 methylene groups between the intercalating units. It is thus well-suited for binding in the narrow and hydrophobic minor groove and suggests that the $\beta$-alanyl linker directs **1** to the ATAA·TTAT site via "classical" minor groove binding,[8a] analogous to the $\beta$-alanine ("$\beta$") moiety used in the hairpin polyamides of Dervan and co-workers.[8b] Minor groove binders are generally known to form hydrogen bonds in the minor groove, between the NH groups and Thy O2/Ade N3.[8] Several such H-bonds can also be postulated for compound **1** (Supporting Information) but could not be unequivocally assigned on the basis of the data. We previously proposed a hydrogen bonding scheme for compound **2** between the backbone N−H groups flanking the NDI rings and two O6 atoms of the guanine residues at the internal positions of the intercalated steps.[2e,6] Analogous in the complex of **1**, the backbone NH groups adjacent to the NDI rings are within H-bonding distance with the O2 atoms of residues T3 and T6. Thus, it appears that adenine at the internal positions of the intercalated steps (3′- or 5′-end) favors minor groove binding, while guanine in those positions favors major groove binding. Such considerations may prove useful for directing linker segments of future threading polyintercalators into a desired DNA groove.

Compound **1** binds preferentially in one of the two possible N-C orientations on the non self-complementary DNA sequence used here. This is perhaps surprising, considering that the minor groove of A/T tracts is virtually C2 symmetric, due to the symmetric disposition of the Thy O2 and Ade N3 groups.[8,9] However, as shown in Figure 3, non-C2 symmetric binding (i.e., with both intercalators placed toward the same DNA strand) of a chiral ligand results in two distinct ligand orientations relative to the binding site.[10] rMD simulations (Supporting Information) suggest that inversion of the binding orientation might cause the Lys 6 side chain to be pointed "out", away from the DNA into the solvent. The different Lys 6-DNA contacts in the two possible ligand orientations either directly contribute to the binding energy difference or indirectly affect H-bonding in the groove by affecting the peptide backbone conformation.

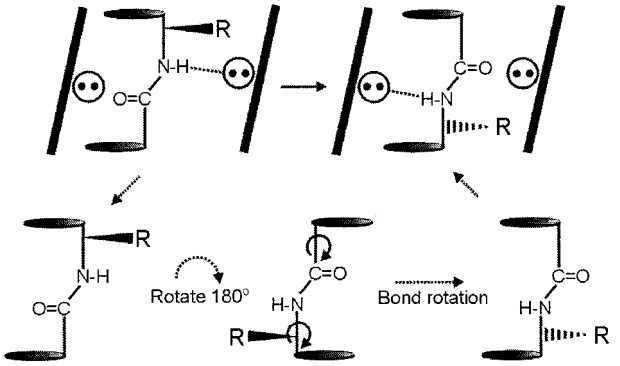



***Figure 3.*** Binding of a hypothetical chiral dipeptide bis-intercalator to the minor groove of an A−T pair. Asymmetric binding toward the same DNA strand results in two nondegenerate orientations, where the chiral residue R points toward or away from the DNA.

The structural analysis presented here for **1**, combined with the data previously reported for **2**, has verified that the polyintercalation strategy based on NDI threading intercalalators is capable of placing linkers in well-defined locations in both the minor and the major grooves. Current work is focused on increasing sequence selectivity of bis-intercalating modules such as **1** and **2**, then linking these minor and major groove-specific modules in hybrid molecules capable of binding specifically longer sequences via linkers bound in alternating grooves.

**Acknowledgment.** This work was supported by the American Cancer Society (RPG-97-085-01-CDD), the Welch Foundation (F-118), and the National Institutes of Health (RO1 GM55646).

**Supporting Information Available:** NMR and modeling data, DNA alkylation by an *N*-bromoacetyl-**1** derivative, experimental procedures (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

## References

(1) (a) Takenaka, S.; Nishira, S.; Tahara, K.; Kondo, H.; Takagi, M. *Supramol. Chem.* **1993**, *2*, 41. Lamberson, C. Ph.D. Thesis, University of Illinois, 23−26, 1991. Veal, J. M.; Li, Y.; Zimmerman, S. C.; Lamberson, C. R.; Cory, M.; Zon, G.; Wilson, W. D. *Biochemistry* **1990**, *29*, 10918−27. Spielmann, H. P.; Wemmer, D. E.; Jacobsen, J. P. *Biochemistry* **1995**, *34*, 8542−53. Jourdan, M.; Garcia, J.; Lhomme, J.; Teulade-Fichou, M. P.; Vigneron, J. P.; Lehn, J. M. *Biochemistry* **1999**, *38*, 14205−13. (b) Tanious, F.; Yen, S.; Wilson, W. D. *Biochemistry* **1991**, *30*, 1813.
(2) (a) Lokey, R. S.; Kwok, Y.; Guelev, V.; Pursell, C.; Hurley, L.; Iverson, B. L. *J. Am. Chem. Soc.* **1997**, *19*, 77202. (b) Murr, M.; Harting, M.; Guelev, V.; Ren, J.; Chaires, J. B.; Iverson, B. L. *Bioorg. Med. Chem.* **2001**, *9*, 1141−1148. (c) Guelev, V.; Cubberley, M.; Murr, M.; Lokey, R. S.; Iverson, B. L. *Methods Enzymol.* **2001**, *340*, 556−72, (d) Guelev, V.; Harting, M.; Lokey, R. S.; Iverson, B. L. *Chem. Biol.* **2000**, *7*, 1−9. (e) Guelev, V.; Lee, J.; Sorey, S.; Ward, J.; Hofmann, D.; Iverson, B. *Chem. Biol.* **2001**, *8*, 415−425.
(3) Baker, B.; Dervan, P. *J. Am. Chem. Soc.* **1989**, *111*, 2700−12.
(4) Feigon, J.; Denny, W. A.; Leupin, W.; Kearns, D. R. *J. Med. Chem.* **1984**, *27*, 450−65.
(5) (a) Sheek, R.; Russo, N.; Boelens, R.; Kaptein, R.; van Boom, J. *J. Am. Chem. Soc.* **1983**, *105*, 2914−2916. Feigon, J.; Wright, J.; Leupin, W.; Denny, W.; Kearns, D. *J. Am. Chem. Soc.* **1982**, *104*, 5540−5541. (b) Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley and Sons: New York, 1986. (c) van Wijk, J.; Huckriede, B. D.; Ippel, J. H.; Altona, C. *Methods Enzymol.* **1992**, *211*, 286−306.
(6) Gallego, J.; Reid, B. *Biochemistry* **1999**, *38*, 15104.
(7) Brünger, A. XPLOR 3.1; Yale University Press: New Haven, CT, 1993.
(8) (a) Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513−538. (b) For a comprehensive review, see: Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 439−461.
(9) Seeman, N. C.; Rosenberg, J. M.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 804−808.
(10) Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 1382.

JA016834E