The Energetic Contribution of a Bifurcated Hydrogen Bond to the Binding of DAPI to dA-dT Rich Sequences of DNA

Tao Lan and Larry W. McLaughlin*

Department of Chemistry, Merkert Chemistry Center
Boston College, Chestnut Hill, Massachusetts 02167

Received September 21, 2000

Bifurcated or three-centered hydrogen bonds, in which the hydrogen atom is within bonding distance of two heteroatoms have been suggested to be present in selected small-molecule structures,[1] both DNA[2] and RNA[3], proteins,[4] enzyme transition states[5] and complexes of ligands and nucleic acids, like those with minor groove binding ligands.[6−9] A number of heterocyclic compounds such as distamycin,[9] netropsin,[7] Hoechst 33258,[8] and 4′,6-diamidine-2-phenyl indole (DAPI)[6] bind within the minor groove of dA-dT-rich sequences. These agents make hydrogen bonds with the functional groups on the edges of the dA-dT base pairs. For all of the ligands noted above, some hydrogen-bonding interactions are bifurcated. For the DAPI−DNA complex, the two terminal amidines form conventional hydrogen bonds while the central indole N−H forms a bifurcated hydrogen bond with the O^2-carbonyls of the central two dT residues (Figure 1).[6] To probe the contribution of such a three-centered hydrogen bond to the stability of this complex we prepared a series of analogue sequences in which either the O^2-carbonyls of selected dT residues, or the N^3-nitrogen of a selected dA residue were deleted and replaced simply by hydrogen or a C−H, respectively. These deletions were accomplished by the incorporation of either a dm^3-2P or a dc^3A base analogue[10] for dT or dA, respectively (Figure 2). A recent crystallographic analysis indicates that $dm^3$2P forms a Watson−Crick base pair with dA.[11]

As a target sequence we chose a $d(GAATTC)_2$-containing sequence very similar to the Dickerson dodecamer used in the crystallization of the DAPI−DNA complex,[6] but we altered the nature of the dG-dC sequences such that the dodecamer was no longer self-complementary (Table 1). This approach resulted in a duplex in which we could make single analogue substitutions in either strand. We employed two assays to assess the effects resulting from specific eliminations of minor groove functional groups on the stability of the ensuing DAPI−DNA complex. In one case we employed a thermal denaturation assay since the binding to the minor groove of ligands such as DAPI tends to overstabilize the B-form helix and increase the observed T_M values.[12] A second assay involved the measurement of dissociation constants (K_D) for both native and analogue complexes.

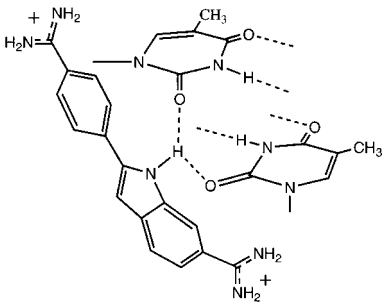

Figure 1. Illustration of the bifurcated hydrogen bonding interactions[6] between the indole N−H of DAPI and the O^2-carbonyls of the two centrally positioned dT residues (bold) of the duplex containing the sequence: 5′ d(...GAA**T**TC...)−3′ d(...CT**T**AAG...). Additional hydrogen bonding interactions for the terminal amidines are not shown.

In the presence of one equivalent of DAPI in a 3 μM solution of the native DNA duplex, the T_M value was observed to increase by 7.9 °C relative to that measured for the DNA duplex alone (Table 1). Replacement of either one of the central dT residues with the $dm^3$2P analogue (and removing the corresponding O^2-carbonyl) resulted in an increased T_M of only 4.5 °C. With both of the central dT residues replaced no significant enhancement in T_M value was observed (<1 °C). By comparison, replacement of one of the dA residues by the analogue dc^3A resulted in an enhanced T_M value of 8.5 °C, slightly larger than that observed for the native sequence in the presence of DAPI. The final sequence containing one dc^3A residue and one $dm^3$2P residue exhibited an intermediate T_M enhancement of 6.4 °C. These results suggested that the absence of either or both O^2-carbonyls from the central dT residues reduced binding by the ligand, while the corresponding deletion of the N^3-nitrogen from a central dA residue actually enhanced binding.

Dissociation constants (K_D values) for the DAPI−DNA complex were obtained by the titration of a fixed concentration of the DAPI ligand by an increasing concentration of the oligonucleotide.[13] The DNA−ligand complex exhibits a dramatically increased quantum yield relative to the free ligand, and this fluorescence signal can be used to quantify binding.[14] The method of continuous variation in ligand concentration (Job plot analysis[15,16]) was used to determine a DNA:ligand stoichiometry of 1:1. The dissociation constant obtained for the native duplex (K_D = 8.8 nM) at 25 °C (Table 1) is consistent with similar values obtained at 5.5 °C.[14] With a target containing one centrally located $dm^3$2P residue at either position within the target site, the measured dissociation constant increases between 25- and 30-fold. With both of the central dT residues replaced by analogues, we could not detect any significant binding at oligonucleotide concentrations as high as 3 μM. The loss of the related N^3-nitrogen of an adjacent dA residue resulted in a reverse effect. Introduction of a single residue of dc^3A into the binding site resulted in a

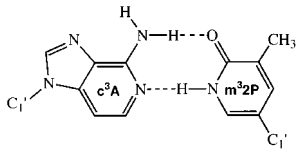

Figure 2. Illustration of Watson−Crick like base pairing[10,11] between dc^3A and $dm^3$2P base analogues.

* To whom correspondence should be addressed. Telephone: (617) 552-3622. Fax: (617) 552-2705. E-mail: larry.mclaughlin@bc.edu.

(1) Tari, L. W.; Secco, A. S. *Acta Crystallogr., Sect. C* **1992**, *48*, 905−907.
(2) Skelly, J. V.; Edwards, K. J.; Jenkins, T. C.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 804−808.
(3) Auffinger, P.; Westof, E. *J. Mol. Biol.* **1999**, *292*, 467−483.
(4) Keeffe, L. J.; Quikr, S.; Gittis, A.; Sondek, J.; Lattman, E. E. *Protein Sci.* **1994**, *3*, 391−401.
(5) Torres, R. A.; Schiott, B.; Bruice, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 8164−8173.
(6) Larsen, T. A.; Goodsell, D. S.; Cascio, D.; Grzeskowiak, K.; Dickerson, R. E. *J. Biomol. Struct. Dyn.* **1989**, *7*, 477−491.
(7) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376−1380.
(8) Pjura, P.; Grzeskowiak, K.; Dickerson, R. E. *J. Mol. Biol.* **1987**, *197*, 257−271.
(9) Coll, M.; Frederick, C. A.; Wang, A. H.-J.; Rich, A. *Proc. Natl. Acad. Sci U.S.A.* **1987**, *28*, 310−320.
(10) Lan, T.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2000**, *122*, 6512−6513.
(11) Wiliams, L. D. Private communication, 2001.
(12) Cory, M.; Tidwell, R. R.; Fairley, T. A. *J. Med. Chem.* **1992**, *35*, 431−438.
(13) Lohman, T. M.; Mascotti, D. P. *Methods Enzymol.* **1992**, *212*, 424−458.
(14) Lootiens, F. G.; McLaughlin, L. W.; Diekmann, S.; Clegg, R. M. *Biochemistry* **1991**, *30*, 182−189.
(15) Job, P. *Ann. Chim. (Paris)* **1928**, *9*, 113−203.
(16) Bostock-Smith, C. E.; Searle, M. S. *Nucleic Acids Res.* **1999**, *27*, 1619−1624.

10.1021/ja003451s CCC: $20.00
Published on Web 02/13/2001

Table 1. DAPI Binding to Native and Analogue Sequences

5' d(C C G G A A T T C G C C)
3' d(G G C C T T A A G C G G)

sequence[a]	T_M (°C)[b]	ΔT_M (°C)[c]	K_D (nM)[d]	$\Delta G°$ (kcal/mol)	$\Delta\Delta G°$ (kcal/mol)
5′−GAATTC− 3′−CTTAAG−	60.0 ± 0.5	7.9 ± 0.5	8.80 ± 0.85	11.0	-
5′−GAA**t**TC− 3′−CTTAAG−	56.5	4.5	267 ± 24	9.0	2.0
5′−GAATTC− 3′−CT**t**AAG−	57.4	4.6	214 ± 21	9.1	1.9
5′−GAA**t**TC− 3′−CT**t**AAG−	52.4	0.7	>3000	<7	≥4
5′−GAATTC− 3′−CTT**a**AG−	58.3	8.5	2.78 ± 0.62	11.7	−0.7
5′−GAATTC− 3′−CT**ta**AG−	54.9	6.4	51.4 ± 8.9	9.9	1.1

[a] **a** = 3-deazaadenine, **t** = 3-methyl-2-pyridone (see Figure 2). [b] T_M values were obtained in 200 mM NaCl, 20 mM phosphate pH 7.0. [c] ΔT_M values were determined in the presence of 1 equivalent of DAPI. [d] Binding constants were obtained at a DAPI concentration of 1−24 nM and varying olgonucleotide concentrations at 25 °C in 200 mM NaCl, 20 mM phosphate pH 7.0.

moderate but reproducible 3-fold enhancement in binding. With a single dm^{3}2P and dc^{3}A residue present in the binding site, a K_D of 51 nM was obtained, roughly 6-fold greater than that obtained for the native site. This value can be explained in terms of both negative and positive effects on binding; with one dm^3-2P analogue reducing binding affinity by 25-fold and the dc^{3}A residue enhancing that affinity by 3-fold, one could predict a dissociation constant some 8-fold higher than that of the native sequence, consistent with the measured value.

$\Delta G°$ values were calculated for both the native and analogue complexes (Table 1). The loss of either O^2-carbonyl from the center of the binding site results in a loss of 2 kcal/mol in binding energy, while the loss of the related N^3-nitrogen from the adjacent dA residue actually enhances binding by 0.7 kcal/mol. With the loss of both centrally positioned O^2-carbonyls we cannot accurately determine a $\Delta G°$ value, but with a K_D of greater than 3 μM the $\Delta G°$ value for this sequence would be no higher than about 7 kcal/mol ($\Delta\Delta G°$ of ≥4 kcal/mol). The final entry in Table 1 results in an intermediate $\Delta\Delta G°$ value of 1.1 kcal/mol—consistent with a loss of ∼2 kcal/mol in binding energy for the dm^{3}2P substitution and a gain of ∼0.7 kcal/mol for the dc^{3}A substitution.

The crystal structure analysis of DAPI bound to a self-complementary duplex containing the d(GAATTC)$_2$ site has indicated the presence of a bifurcated hydrogen bond involving the indole N−H and the O^2-carbonyls of the central two dT residues[6] (see Figure 1). The introduction of an analogue dm^{3}2P residue for one of these central dT residues disrupts one of the bifurcated hydrogen bonds between the ligand and the target sequence but permits the second interaction to remain. The observation that the loss of either bifurcated hydrogen bond from the complex reduces the binding energy by 2 kcal/mol suggests that each of the bifurcated interactions are at least energetically equivalent to conventional hydrogen bonds between uncharged partners, originally suggested to be between 0.5 and 1.2 kcal/mol for the binding of substrate to a transition state.[17] Similarly, individual hydrogen bonds in base pairs, for example that involving the N^2-amino group of guanine, have been estimated to contribute 0.7 to 1.6 kcal/mol of stabilization energy.[18] Studies with proteins suggest a value of at least 1 kcal/mol and perhaps as high as 2.2 kcal/mol.[19] With both bifurcated bonds absent (substitution of both central dT residues by dm^{3}2P) we cannot effectively determine a binding energy, but the limits of our assay suggest a value of no more than 7 kcal/mol, or a reduction by roughly 4 kcal/mol in binding energy. A linear relationship is present for the values of ΔT_M and $\Delta G°$ reported here (see Supporting Information) and this relationship suggests that the 0.7 °C enhancement in T_M correlates to roughly a $\Delta G°$ value of about −6.5 kcal/mol, consistent with the observation that each interaction of the bifurcated hydrogen bond contributes about 2 kcal/mol in binding energy.

Substitution of a single dc^{3}A analogue into of the binding site did not result in a loss of binding energy although a minor groove functional group had been deleted. This is consistent with the crystal structure analysis[6] and suggests that the position and targets for formation of the bifurcated hydrogen bond in this complex are solely the carbonyls of the two dT residues in the center of the binding site, although other modes of binding can be induced.[20,21] It was unexpected that the dc^{3}A-containing sequence would exhibit tighter binding with the ligand than did the native sequence. It has previously been suggested, most notably for the structure of netropsin bound to the minor groove of DNA,[7] that the C^2−H sites of the dA residues provide significant hydrophobic contacts with the ligand. In the present complex such an effect is also possible, and when the adjacent N^3-nitrogen is deleted, the minor groove edge of the dA residue should offer a more hydrophobic surface. Hydrophobic interactions have been suggested to contribute significant effects to minor groove binding by other ligands.[22] The enhanced binding by DAPI to the dc^{3}A-containing sequence could then be explained by enhanced hydrophobic interactions, without any loss in interresidue hydrogen bonds. Consistent with this suggestion are the results for the sequence containing both dc^{3}A and dm^{3}2P residues.

Although experimental studies suggest energetic values of 1 to 2 kcal/mol for hydrogen bonding interactions in nucleic acids[18] and proteins,[19] corresponding theoretical studies, at least for proteins,[23,24] suggest that polar functional groups do not contribute significantly to structural stability. An earlier study with distamycin and sequences containing dc^{3}A suggested that the absence of the N^3-nitrogens had little effect on ligand binding.[25] Although we cannot compare that work directly with the current study, our results suggest that the O^2-carbonyls may be the more critical ligand binding sites within the minor groove. An earlier study[26] with peptides in chloroform suggests the energetic superiority of two-centered over three-centered hydrogen bonds, but in systems composed of multiple interactions, more favorable complexes are predicted[27,28] when both acceptors are in the same molecule and optimally positioned. There is some experimental evidence for this effect[29] including the present work.

Acknowledgment. This work was supported by a grant from the NSF (MCB 0077667).

Supporting Information Available: Details of the binding assays and sample plots, a sample Job plot, and the relationship between ΔT_M and $-\Delta G$ (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA003451S

(17) Fersht, A. R.; Shi, J.-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M. H.; Winter, G. *Nature* **1985**, *314*, 235−238.

(18) Turner, D. H.; Sugimoto, N.; Kierzek, R.; Dreiker, S. D. *J. Am. Chem. Soc.* **1987**, *109*, 3783−3785.

(19) Myers; Pace *Biophys. J.* **1996**, *71*, 2033−2039.

(20) Vlieghe, D.; Sponer, J.; Van Meervelt, L. *Biochemistry* **1999**, *38*, 16443−16451.

(21) Trotta, E.; Paci, M. *Nucleic Acids Res.* **1998**, *26*, 4706−4713.

(22) Mazur, S.; Tanious, F. A.; Ding, D.; Kumar, A.; Boykin, D. W.; Simpson, I. J.; Neidle, S.; Wilson, W. D. *J. Mol. Biol.* **2000**, *300*, 321−337.

(23) Harris, B.; Wang, A. S. *Adv. Protein Chem.* **1994**, *46*, 27−58.

(24) Lazaridis, T.; Archontis, G.; Karplus, M. *Adv. Protein Chem.* **1995**, *47*, 231−306.

(25) Lever, C.; Li, X.; Cosstick, R.; Ebel, S.; Brown, T. *Nucleic Acids Res.* **1993**, *21*, 1743−1746.

(26) Yang, J.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 9090−9091.

(27) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008−1010.

(28) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810−2819.

(29) Zimmerman, S. C.; Murray, T. J. *Tetrahedron Lett.* **1994**, *35*, 4077−4080.