

2-Aminobenzene Derivatives as Unnatural Nucleobases and Their DNA Duplex Stabilities

Ju Young Heo and Gil Tae Hwang*

Department of Chemistry, Kyungpook National University, Daegu 702-701, Korea. *E-mail: giltae@knu.ac.kr
Received September 29, 2010, Accepted October 2, 2010

Key Words: DNA, Nucleoside, Nucleotide, Thermal stability, Melting temperature

The thermal stability and faithful replication of a natural DNA duplex is based on the Watson-Crick hydrogen-bonding (H-bonding) patterns of dA:dT and dC:dG base pairs. Additional stable and selective unnatural base pairs would facilitate hybridization and encoding experiments in cases where natural sequences cross hybridize or where increased information storage is desirable.¹ Recent studies have shown that H-bonds between unnatural nucleobases are not absolutely essential for selective base pairing.² On the basis of this concept, a large number of nucleotides bearing large and predominantly hydrophobic nucleobases have been synthesized and characterized according to the isocarboxystyryl, naphthyl, and indole scaffolds, whose pairing is mediated by hydrophobic interactions.² From these studies, several self-pairs (formed by pairing of identical analogues) and heteropairs (formed by pairing of 2 different analogues) that rival natural base pairs in terms of thermal stability and selectivity against mispairing have been identified. A larger aromatic surface area increases the thermal stability and selectivity of unnatural base pairs, and a structural study suggested that analogues intercalate into the opposite strand between the pairing nucleobase and an adjacent nucleobase.³ We therefore became interested in examining nucleobase analogues that lack a large aromatic surface area but are optimized to interact edge-on analogous to natural Watson-Crick base pairing.

Recent works have focused on the systematic examination of the contribution to base pair stability of nucleobase shape, hydrophobicity, dipole moment, and polarizability within the benzene, pyridine, or pyridone scaffolds.⁴ Our current focus is identifying combinations of methyl-group substitutions at the major groove position and heteroatom derivatizations at the *ortho*-position in the unnatural nucleobase scaffold responsible for the improved stability of these smaller unnatural nucleobases. Although 2-aminophenyl derivatives are good candidates as hydrophobic surrogates, relatively little work has been done on their deoxy-C-nucleosides. Some research groups have synthesized these derivatives by inducing a titanium aldol reaction/O-H insertion reaction between diazoketone and aminophenyl derivative and 2'-deoxy-D-ribo-lactone.⁶ However, these processes have the disadvantages of long steps and low stereoselectivity. Further, no previous study has examined the effect of 2-aminophenyl derivatives on the thermal stability of the DNA duplex. To further understand how 2-aminobenzene derivatives contribute to DNA stability, we synthesized and characterized the 2-aminobenzene nucleotide analogues **AM** and **ADM**

(Figure 1, Scheme 1). We also synthesized the *N*-methylated aminobenzene derivatives **MAN**, **MMAN**, and **DMMAN** to evaluate the amino functional group at the 2-position for its role in optimizing DNA stability. A palladium-mediated Heck-type coupling reaction⁷ between the silyl-protected glycol **2**⁸ and the 2-amine-substituted iodobenzenes **1a-1e** yielded the corresponding β -anomers, which had high stereoselectivity. After coupling, the mixture was treated with *n*-TBAF under mild acidic conditions, and the resulting keto group was stereoselectively reduced with NaBH(OAc)₃, yielding the free nucleosides **3a-3e**. The primary amines of free nucleosides **3a** and **3b** were

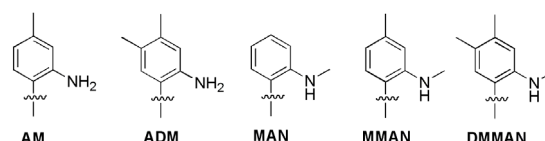
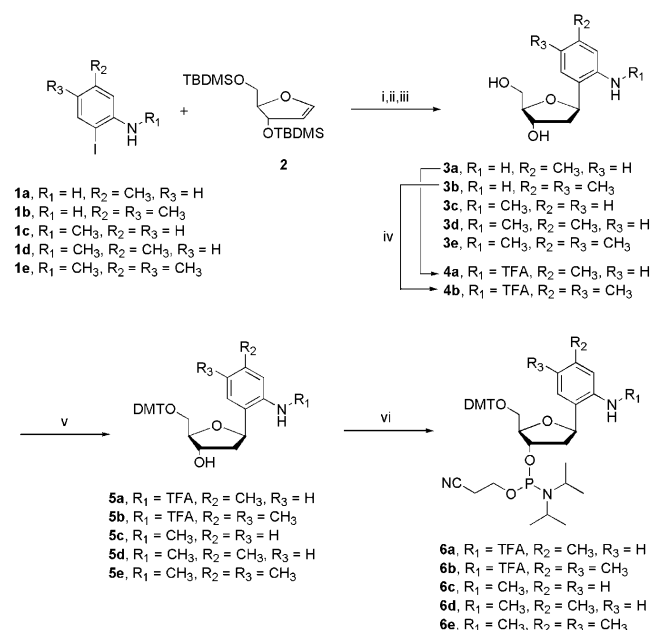


Figure 1. Unnatural nucleobases. Sugars and phosphates have been omitted for clarity.



Reagents: (i) Pd(OAc)₂, Ph₃As, *n*Bu₃N, DMF, 70 °C; (ii) TBAF, 0 °C; (iii) NaB(OAc)₃H, AcOH, CH₃CN, 0 °C; (iv) (CF₃CO)₂O, pyridine, 0 °C; (v) DMT-Cl, DMAP, DIPEA, pyridine, rt; (vi) 2-Cyanoethyl diisopropylchlorophosphoramidite, DIPEA, CH₂Cl₂, rt.

Scheme 1. Synthesis of aminobenzene derivatized nucleotides

Table 1. T_m values for duplexes containing unnatural base pairs^a

5'-d(GCGTACXCATGCG)-3' 3'-d(CGCATGYGTACGC)-5'					
X	Y	T_m (°C)	X	Y	T_m (°C)
AM	AM	55.1	ADM	ADM	54.1
	A	49.3		A	47.6
	C	45.3		C	43.6
	G	49.6		G	49.2
	T	47.8		T	46.0
MAN	MAN	50.5	MMAN	MMAN	51.3
	A	47.5		A	48.6
	C	43.3		C	43.0
	G	47.1		G	48.3
	T	46.1		T	47.1
DMMAN	DMMAN	50.3			
	A	46.8			
	C	42.1			
	G	46.8			
	T	45.8			

^aUncertainty in values is less than 0.1 °C. See Experimental Section for details.

protected with a trifluoroacetyl group, which was eliminated during DNA purification. The phosphoramidite **6** used in automated DNA synthesis was obtained by standard dimethoxytritylation and phosphorylation.

To determine the thermodynamic stability of the unnatural base pairs in duplex DNA, all unnatural and natural nucleosides were incorporated into the complementary oligonucleotides 5'-d(GCGTACXCATGCG)-3' and 3'-d(CGCATGYGTACGC)-5' at the positions labeled X and Y. The melting temperature (T_m) of the resulting duplexes was determined *via* thermal denaturation experiments.⁹ For comparison, the T_m for a duplex containing a natural base pair (X:Y = dT:dA) is 58.7 °C (Table 1).^{2a}

The unnatural self-pairs (formed between two identical analogues) are generally more stable than mismatches of natural nucleotides, with T_m values that range from 50.3 to 55.1 °C. The relative stability of the **AM** self-pairs probably reflects the optimized packing of the 4-position methyl groups in the interbase interface, as compared to the packing of the **ADM** self-pair, which has an additional 5-position methyl group. This may also hold true for the **MMAN** self-pair with only 4-position methyl groups, which is more stable than the **MAN** self-pair, which has no methyl groups at the 4-position, and the **DMMAN** self-pair, which has an additional 5-position methyl group. The stability of the unnatural self-pairs of **MAN**, **MMAN**, and **DMMAN** was lower than that of **AM**, **ADM**, and even simple benzene ($T_m = 52.8$ °C)¹⁰ self-pairs. This result indicates that the methylation at the *ortho*-amine probably disrupts the optimized packing of the 2-position amine group in the interbase interface and hence significantly destabilized the self-pairs. Despite the absence of H-bond donors/acceptors and the significantly reduced aromatic surface area, the T_m of the **AM** self-pair is only 3.6 °C less than that of a natural dT:dA pair in the same sequence context.

To examine the thermal selectivity of the unnatural base pairs, we determined T_m values with the unnatural bases incorporated

at position X and natural nucleotides incorporated at position Y (Table 1). These mismatches are significantly less stable than even the least stable self-pair. This result demonstrates that these small unnatural bases differentiate strongly between the hydrophobic nucleobases and the natural nucleobases. This is probably, at least in part, because of the forced desolvation of the more hydrophilic natural bases. In a manner similar to natural mispairing, uncompensated desolvation of hydrophilic functional groups of natural nucleobase likely renders the mispairing with hydrophobic analogues energetically unfavorable. Furthermore, the pyrimidine scaffold of dC and dT might be too small to interact across the duplex with unnatural nucleobases. In all, unfavorable desolvation/hydrophobic effects and geometrical constraints of the mismatches between hydrophobic nucleobase and natural nucleobase contribute extensively to the thermal instability of nucleic acid.

It is interesting to consider how nucleobase analogues with such limited surface area and no ability to form H-bonds can pair with reasonable stability and thermal selectivity. In this study, amino and methyl substitutions of the small phenyl ring nucleobase scaffold were seen to have large effects on unnatural base pair and mismatch stability, suggesting that hydrophobicity, dipole moment, and polarizability may all contribute to base pair stability. In addition, modifications at the 2- and 4-positions may generally be useful, at least for the benzene scaffold, for stabilizing DNA duplex.

Experimental Section

General procedure for Heck coupling. Compound **1a-1b**,¹¹ **1c-1e**,¹² and **2**⁸ were prepared according to reported procedures. A mixture of palladium acetate (0.10 equiv) and triphenylarsine (0.20 equiv) in DMF (0.4 M) was stirred under argon atmosphere at room temperature for 20 min. Then **1** (1 equiv), 1,4-anhydro-3,5-bis-*O*-(*tert*-butyldimethylsilyl)-2-deoxy-D-erythro-pent-1-enitol **2** (1.1 equiv), and tri-*n*-butylamine (1.3 equiv) in DMF (0.4 M) were added, and the resulting reaction mixture was stirred under nitrogen at 70 °C for 24 h. The mixture was cooled to 0 °C, and 1 M tetra-*n*-butylammonium fluoride (2.3 equiv) in THF (0.15 M) was added and stirred for 1 h. The reaction mixture was filtered through Celite and extracted with EtOAc/Et₂O (1:1, v/v) and saturated aq. NaHCO₃. The combined organic layers were dried over anhydrous Na₂SO₄. After filtration and evaporation, the product, without further purification or characterization, was dissolved in acetic acid and acetonitrile (0.15 M), the solution was cooled to 0 °C, and sodium triacetoxyhydride (2.0 equiv) was added and stirred for 1 h. The reaction mixture was extracted with EtOAc and saturated aq. NaHCO₃. The combined organic layers were dried over anhydrous Na₂SO₄. After filtration and evaporation, the residue was purified by silica gel column chromatography (0 - 5% MeOH in EtOAc) to afford **3** as a white solid.

General procedure for DMT-protection. To a solution of free nucleoside **3c-e** or **4a-b** (1 equiv) and 4-(dimethylamino)pyridine (0.5 equiv) in anhydrous pyridine (0.25 M) was added *N,N*-diisopropylethylamine (3.0 equiv) and 4,4'-dimethoxytrityl chloride (1.5 equiv) and the mixture was stirred overnight under argon at room temperature. The reaction was quenched by

addition of MeOH. The mixture was evaporated and purified by chromatography on a silica gel column (33% EtOAc in hexane).

General procedure for phosphoramidation. 2-Cyanoethyl-diisopropyl chlorophosphoramidite (1.5 equiv) was added dropwise to a solution of **5** (1 equiv) and *N,N*-diisopropylethylamine (4 equiv) in CH₂Cl₂ at room temperature. After the reaction reached completion (30 min), the mixture was concentrated in vacuo and purified by chromatography through a short column of SiO₂ (15% EtOAc in CH₂Cl₂) to yield **6**.

Synthesis of oligonucleotides. Oligonucleotides were prepared by the β-cyanoethylphosphoramidite method on controlled pore glass supports (1 mmol) using an Applied Biosystems Inc. 392 DNA/RNA synthesizer as standard method. After automated synthesis, the oligonucleotides were cleaved from the support by concentrated aqueous ammonia for 1 h at room temperature, deprotected by heating at 56 °C for 10 h, and purified by denaturing polyacrylamide gel electrophoresis (12 - 15%, 8 M urea). The oligonucleotides were purified by PAGE, visualized by UV shadowing, and recovered by electroelution. After ethanol precipitation, the concentration of oligonucleotides was determined by UV-vis absorption.

Thermal stability. The unnatural and natural nucleosides were incorporated into the complementary oligonucleotides 5'-d(GC GTACXCATGCG)-3' and 5'-d(CG CATGYGTACGC)-3' at the positions labeled **X** and **Y**. For UV melting experiments the sample (3 μM oligonucleotide, 10 mM PIPES buffer, pH 7.0, 100 mM NaCl and 10 mM MgCl₂) absorption was monitored at 260 nm from 20 °C to 80 °C at a heating rate of 0.5 °C per min, using a Cary 300 Bio UV-vis spectrophotometer. Melting temperatures were determined from the derivative method using the Cary Win UV thermal application software.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0007605).

Supporting Information Available. Characterization data of nucleotides **3-6**. The supporting materials are available via the Internet <http://journal.kcsnet.or.kr>.

References

- (a) Hoheisel, J. D. *Nat. Rev. Genet.* **2006**, *7*, 200-210. (b) Krützfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K. G.; Tuschl, T.; Manoharan, M.; Stoffel, M. *Nature* **2005**, *438*, 685-689. (c) Schwarz, D. S.; Ding, H.; Kennington, L.; Moore, J. T.; Schelter, J.; Burchard, J.; Linsley, P. S.; Aronin, N.; Xu, Z.; Zamore, P. D. *PLoS Genet.* **2006**, *2*, e140.
- (a) Berger, M.; Ogawa, A. K.; McMinn, D. L.; Wu, Y.; Schultz, P. G.; Romesberg, F. E. *Angew. Chem. Int. Ed.* **2000**, *39*, 2940-2942. (b) Wu, Y.; Ogawa, A. K.; Berger, M.; McMinn, D. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 7621-7632.
- Matsuda, S.; Fillo, J. D.; Henry, A. A.; Rai, P.; Wilkens, S. J.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Schultz, P. G.; Spraggon, G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 10466-10473.
- (a) Hwang, G. T.; Romesberg, F. E. *Nucleic Acids Res.* **2006**, *34*, 2037-2045. (b) Hwang, G. T.; Leconte, A. M.; Romesberg, F. E. *ChemBioChem* **2007**, *8*, 1606-1611. (c) Hari, Y.; Hwang, G. T.; Leconte, A. M.; Joubert, N.; Hocek, M.; Romesberg, F. E. *ChemBioChem* **2008**, *9*, 2796-2799.
- Calter, M. A.; Zhu, C. *J. Org. Chem.* **1999**, *64*, 1415-1419.
- Stoop, M.; Zahn, A.; Leumann, C. J. *Tetrahedron* **2007**, *63*, 3440-3449.
- (a) Farr, R. N.; Outten, R. A.; Cheng, J. C. Y.; Daves, G. D. *Organometallics* **1990**, *9*, 3151-3156. (b) Hsieh, H. P.; McLaughlin, L. W. *J. Org. Chem.* **1995**, *60*, 5356-5359.
- (a) Larsen, E.; Jørgensen, P. T.; Sofan, M. A.; Pedersen, E. B. *Synthesis* **1994**, 1037-1038. (b) Cameron, M. A.; Cush, S. B.; Hammer, R. P. *J. Org. Chem.* **1997**, *62*, 9065-9069.
- Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515-537.
- Matsuda, S.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 14419-14427.
- Kajigaeshi, S.; Kakinami, T.; Yamasaki, H.; Fujisaki, S.; Okamoto, T. *Bull. Chem. Soc. Jpn.* **1988**, *61*, 600-602.
- Mouzin, G.; Cousse, H.; Autin, J.-M. *Synthesis* **1981**, 448-449.