JOC Article
pubs.acs.org/joc

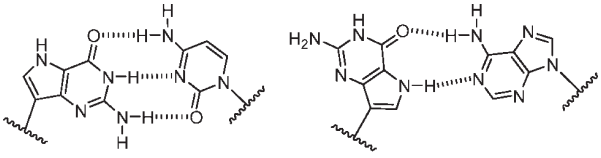
Oligonucleotide Incorporation and Base Pair Stability of 9-Deaza-2′-deoxyguanosine, an Analogue of 8-Oxo-2′-deoxyguanosine

Michelle L. Hamm,*,† Anna J. Parker, Tyler W. E. Steele, Jennifer L. Carman, and Carol A. Parish*,†

Department of Chemistry, University of Richmond, Gottwald Science Center, Richmond, Virginia 23173. †*Contact M.L.H. for synthesis and melting studies and C.A.P. for quantum studies.*

mhamm@richmond.edu; cparish@richmond.edu

Received June 2, 2010



9-Deaza-2′-deoxyguanosine (CdG) is a C-nucleoside and an analogue of the abundant promutagen 8-oxo-2′-deoxyguanosine (OdG). Like 2′-deoxyguanosine (dG), CdG should form a stable base pair with dC, but similar to OdG, CdG contains an N7-hydrogen that should allow it to also form a relatively stable base pair with dA. In order to further investigate the base pairing of CdG, it was incorporated into DNA and paired with either dC or dA. Melting studies revealed CdG:dC base pairs are less stable than dG:dC base pairs, while CdG:dA base pairs are less stable than OdG:dA base pairs. In order to gain a deeper understanding of these results, quantum studies on model structures of nucleoside monomers and base pairs were performed, the results of which indicate that (i) CdG:dC base pairs are likely destabilized relative to dG:dC as a result of structural constraints imposed by the C-nucleotide character of CdG, and (ii) CdG:dA base pairs may be less stable than OdG:dA base pairs, at least in part, because of a third long-range interaction that is possible in OdG:dA but not in CdG:dA.

Introduction

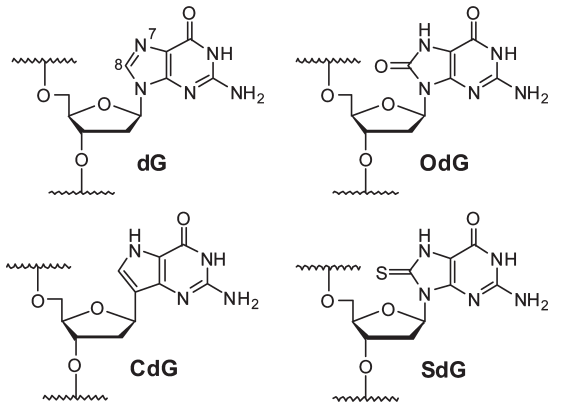
8-Oxo-2′-deoxyguanosine (OdG; Scheme 1) is one of the most prominent damaged nucleotides in mammalian cells.[1] It arises through reaction of 2′-deoxyguanosine (dG) with reactive oxygen species that can be produced by radiation[2] or chemical carcinogens[3] and during normal metabolic respiration.[4] Much research has been focused on OdG, not only because of its prominence, but also because it is a known promutagen; OdG can form stable base pairs with both 2′-deoxycytosine (dC) and 2′-deoxyadenosine (dA; Scheme 2), thus allowing for the production of dG→T transversions during replication.

Crystal structure data from DNA duplexes containing OdG:dC and OdG:dA base pairs have revealed that the glycosidic bond conformation of OdG varies depending on the opposing base. When paired to dC, OdG is in the *anti* conformation and the Watson−Crick hydrogen bonding faces interact.[5,6] When base paired to dA however, OdG is in the *syn* conformation and uses its Hoogsteen edge to hydrogen bond with the Watson−Crick face of dA.[7,8] Consistent with recent work that indicates the number of hydrogen bonds is the best predictor of base pair stability,[9] OdG:dC base pairs, which contain three hydrogen bonds, are more

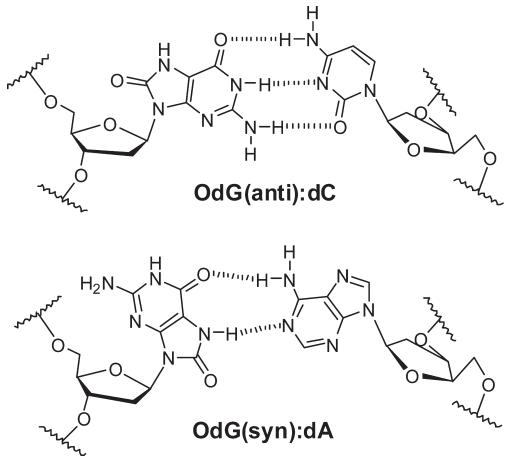
(1) Kunkel, T. *Trends Genet*. **1999**, *15*, 93–94.
(2) Hutchinson, F. *Prog. Nucleic Acid Res. Mol. Biol*. **1985**, *32*, 115–154.
(3) Floyd, R. A. *Carcinogenesis* **1990**, *11*, 1447–1450.
(4) Pryor, W. A. *Annu. Rev. Physiol*. **1986**, *48*, 657–667.
(5) Lipscomb, L.; Peek, M.; Morningstar, M.; Verghis, S.; Miller, E.; Rich, A.; Essigmann, J.; Williams, L. *Proc. Natl. Acad. Sci. U.S.A*. **1995**, *92*, 719–723.
(6) Oda, Y.; Uesugi, S.; Ikehara, M.; Nishimura, S.; Kawase, Y.; Ishikawa, H.; Inoue, H.; Ohtsuka, E. *Nucleic Acids Res*. **1991**, *19*, 1407–1412.
(7) McAuley-Hecht, K. E.; Leonard, G. A.; Gibson, N. J.; Thomson, J. B.; Watson, W. P.; Hunter, W. N.; Brown, T. *Biochemistry* **1994**, *33*, 10266–10270.
(8) Kouchakdjian, M.; Bodepudi, V.; Shibutani, S.; Eisenberg, M.; Johnson, F.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1991**, *30*, 1403–1412.
(9) Geyer, C. R.; Battersby, T. R.; Benner, S. A. *Structure* **2003**, *11*, 1485–1498.

DOI: 10.1021/jo1010763 Published on Web 07/29/2010

SCHEME 1



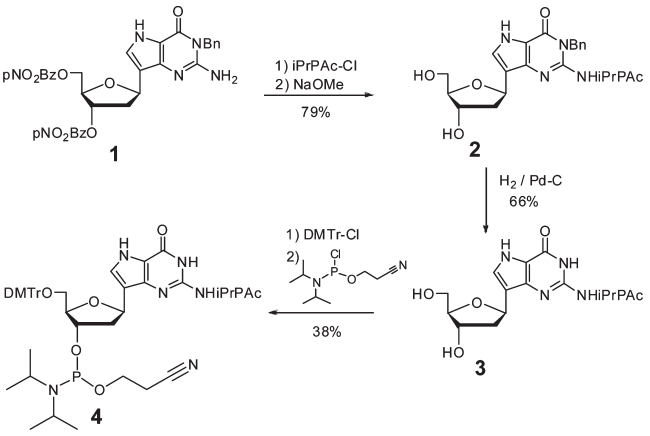
SCHEME 2



stable than OdG:dA mismatches, which contain only two classical hydrogen bonds.[10] Additionally, most likely due to the abundance and promutagenic character of OdG, cells have evolved repair enzymes that specifically remove 8-oxoguanine from OdG:dC base pairs[11,12] and adenine from OdG:dA mismatches.[13]

The base pairing, replication, and repair of OdG described above differs significantly from that of dG despite varying from it at only the N7 and C8 positions. Thus it is possible that one or both of these sites are key to the distinct biochemical activities of OdG. In order to better address the individual importance of these two sites, we have set about creating analogues of OdG that alter either the N7-hydrogen or C8-oxygen while leaving the other site unaffected. Toward this goal, we report herein the oligonucleotide incorporation and base pair stability of the C-nucleotide 9-deaza-2′-deoxyguanosine (CdG; Scheme 1); by replacing the nitrogen at position 9 of dG with carbon, CdG eliminates the C8-oxygen while preserving the N7-hydrogen required for base pairing to dA. Thus, by comparing the base pairing of CdG to that of

SCHEME 3



OdG, further insights into the importance of the C8-oxygen to the promiscuous base pairing of OdG can be obtained.

Results

Oligonucleotide Incorporation of CdG. 9-Deaza-2′-deoxyguanosine has been incorporated into oligonucleotides previously. Revankar et al. deoxygenated 9-deazariboguanosine at the 2′ position before using it to determine that CdG prevents G-tetrad formation.[14] Since publication of that study, a more efficient route to the CdG monomer has been published;[15] using this route, we were able to generate the CdG derivative **1** (Scheme 3) in 7 steps and 7% overall yield starting from commercially available 2-amino-4-hydroxy-6-methylpyrimidine. Additionally, though Revankar et al. used an isobutyryl group to protect the exocyclic amine of the base during DNA synthesis, in our hands the deprotection conditions (29.7% NH_4OH, 55 °C, 15 h) required for removal of the isobutyryl group resulted in roughly 40% α/β epimerization at C1′ of the monomer according to NMR. Thus we instead used an isopropylphenoxyacetyl (iPrPAc) group, which can be removed under much milder conditions, to protect the exocyclic amine. Thus nucleoside **1** was treated with isopropylphenoxyacetyl chloride before selective deprotection of the sugar alcohols with sodium methoxide (Scheme 3). The N1-benzyl group left over from CdG synthesis was then removed by pressurized catalytic hydrogenation to yield nucleoside **3**. To confirm that no epimerization would occur during deprotection of the iPrPAc group, nucleoside **3** was treated with 29.7% NH_4OH for 18 h at room temperature. NMR analysis of the product revealed full deprotection occurred with no α-anomer present. Finally, in order to prepare CdG for solid phase synthesis, the 5′- and 3′-oxygens were protected as a dimethoxytrityl ether and activated as a phosphoramidite, respectively, to yield the solid phase synthesis ready nucleotide **4** as a single diastereomer.

The CdG phosphoramidite **4** was used to synthesize an 11 nucleotide long oligonucleotide with the sequence 5′-dC-CATCXCTACC-3′, where X is CdG (**5**). All standard procedures were used except that the oligonucleotide was deprotected with ammonium hydroxide for only 18 h at room

(10) Plum, G. E.; Grollman, A. P.; Johnson, F.; Breslauer, K. J. *Biochemistry* **1995**, *34*, 16148–16160.
(11) Boiteux, S.; Gajewski, E.; Laval, J.; Dizdaroglu, M. *Biochemistry* **1992**, *31*, 106–110.
(12) Bjoras, M.; Luna, L.; Johnson, B.; Hoff, E.; Haug, T.; Rognes, T.; Seeberg, E. *EMBO J.* **1997**, *16*, 6314–6322.
(13) Michaels, M. L.; Tchou, J.; Grollman, A. P.; Miller, J. H. *Biochemistry* **1992**, *31*, 10964–10968.
(14) Rao, T. S.; Lewis, A. F.; Durland, R. H.; Revankar, G. R. *Tetrahedron Lett.* **1993**, *34*, 6709–6712.
(15) Gibson, E. S.; Lesiak, K.; Watanabe, K. A.; Gudas, L. J.; Pankiewicz, K. W. *Nucleosides Nucleotides* **1999**, *18*, 363–376.

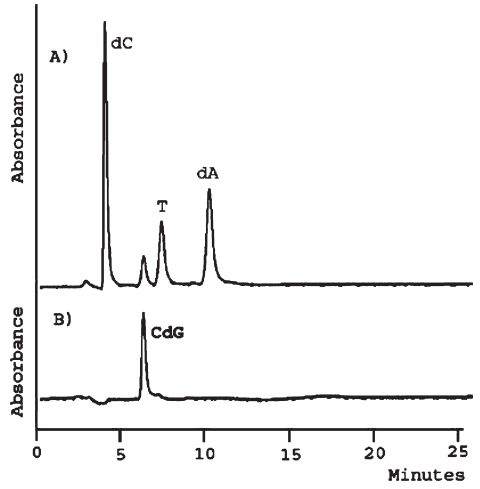


FIGURE 1. Nuclease digest of oligonucleotide **5** (A) and an authentic standard of CdG (B).

TABLE 1. Melting Temperatures (°C) of Duplexes Containing the Various Base Pairs[a]

5′-dCCATC**X**CTACC-3′ 3′-dGGTAG**Y**GATGG-5′				
	X = dG[b]	**X** = CdG	**X** = OdG[b]	**X** = SdG[b]
Y = dC	57.5 ± 0.5	54.1 ± 0.2	52.7 ± 0.5	46.6 ± 0.2
Y = dA	43.2 ± 0.2	45.2 ± 0.4	48.0 ± 0.3	47.7 ± 0.3

[a]Conditions: 1 M NaCl, 0.1 mM EDTA, 5 μM duplex, and 100 mM sodium phosphate, pH 7.0. Average T_m values ± standard deviation were calculated from three or more melts. [b]Taken from ref 16.

temperature. The oligonucleotide was then purified by denaturing PAGE and reverse-phase HPLC.

To further confirm if epimerization occurred during DNA synthesis or purification, oligonucleotide **5** was not only characterized by MALDI-TOF but also digested to its individual nucleosides and analyzed by analytical reverse-phase HPLC. As seen in Figure 1A, the digest produced only four main peaks (corresponding to dC, CdG, T, and dA), suggesting little α-anomer of CdG was present; to further characterize the CdG peak, it was compared to an authentic sample of CdG (Figure1B) as well as the α-anomer of CdG (data not shown).[15] The peaks corresponding to CdG from the digest of **5** and the CdG standard had similar retention times of 6 min and coeluted, while the α-anomer of CdG had a retention time of 9 min.

Experimental Base Pair Stabilities. Oligonucleotide **5** was then used to better understand the base pairing of CdG as compared to dG, OdG, and 8-thio-2′-deoxyguanosine (SdG, which also contains a N7-hydrogen but not a C8-oxygen; Scheme 1). It was paired to complementary oligonucleotides (Table 1) and melted under conditions identical to those previously reported for duplexes containing dG, OdG, and SdG.[16] Since the only difference in the various duplexes studied was the one X:Y base pair, any difference in melting temperature is attributable to that particular base pair. Previous studies have shown that OdG:dC and OdG:dA base pairs are much more similar in stability than dG:dC and dG:dA base pairs.[10] It is believed this is due, in part, to a destabilization of OdG:dC base pairs relative to dG:dC, which is likely caused by a steric clash between the C8-oxygen and ribose ring of OdG. Consistent with this theory, other nucleosides with steric bulk off C8, including SdG,[16] show reduced stability in base pairs with dC. Interestingly, though CdG contains little steric bulk off C8, it also shows reduced stability in base pairs with dC as compared to dG.

Also contributing to the relatively similar stability of OdG:dC and OdG:dA base pairs is the increased stability of OdG:dA pairs relative to dG:dA pairs. This is most likely due, at least in part, to the presence of an N7-hydrogen on OdG. This hydrogen allows OdG to form a base pair with dA that mimics a natural dT:dA base pair.[7] dG lacks the required N7-hydrogen and thus can only assume mismatches with dA that do not structurally mimic a natural base pair.[17,18] Though CdG, like OdG and SdG, contains an N7-hydrogen and CdG:dA base pairs likely adopt a structure similar to that of OdG:dA and SdG:dA base pairs (vide infra), they are less stable than either of these pairs.

Computational Study of Monomers and Base Pairs. Since significant changes in structure and electronics can accompany the change from an N- to a C-nucleotide, a B3LYP/6-31G* geometry optimization and frequency analysis using the polarizable continuum model (PCM)[19] for water was undertaken to further investigate the nature of the energetic differences between dG, CdG, OdG, and SdG both as monomers and in base pairs. Previous work[20] has indicated that this level of theory yields heavy atom distances that are approximately 0.1 Å too short but that interaction energies are well reproduced. All calculations were performed on nucleoside structures containing deoxyribose sugars with and without a CH_3-O-CH_2- (methoxymethylene) linkage on the 4′ carbon and a CH_3-O- (methoxy) linkage on the 3′ carbon. These substituents were added to mimic the steric bulk of the phosphodiester chain in DNA. These structural variations are shown in the representative model structure in Figure 2. Base pair interaction energies were determined using $E_{int} = E(\text{base pair}) - E(\text{nucleoside1}) - E(\text{nucleoside2})$ where a negative E_{int} indicates a favorable interaction. Previously, quantum mechanical *gas-phase* calculations have been reported on bases[20−28] and base pairs.[22] In addition, theoretical results incorporating solvent effects have been described for monomers[22,23,28−31] and base pairs lacking a ribose ring.[22] To our knowledge, the results reported here

(16) Hamm, M. L.; Cholera, R.; Hoey, C. L.; Gill, T. J. *Org. Lett.* **2004**, *6*, 3817–3820.

(17) Brown, T.; Hunter, W. N.; Kneale, G.; Kennard, O. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 2402–2406.

(18) Prive, G. G.; Heinemann, U.; Chandrasegaran, S.; Kan, L. S.; Kopka, M. L.; Dickerson, R. E. *Science* **1987**, *238*, 498–504.

(19) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43–54.

(20) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247–3276.

(21) Aida, M.; Nishimura, S. *Mutat. Res.* **1987**, *192*, 83–89.

(22) Cysewski, P. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 3117–3125.

(23) Cysewski, P.; Bira, D.; Bialkowski, K. *J. Mol. Struct. THEOCHEM* **2004**, *678*, 77–81.

(24) Gu, J. D.; Leszczynski, J. *J. Phys. Chem. A* **1999**, *103*, 577–584.

(25) Marian, C. M. *J. Phys. Chem. A* **2007**, *111*, 1545–1553.

(26) McConnell, T. L.; Wheaton, C. A.; Hunter, K. C.; Wetmore, S. D. *J. Phys. Chem. A* **2005**, *109*, 6351–6362.

(27) Reynisson, H.; Steenken, S. *Phys. Chem. Chem. Phys.* **2002**, *4*, 527–532.

(28) Venkateswarlu, D.; Leszczynski, J. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 373–382.

(29) Barbe, S.; Le Bret, M. *J. Comput. Chem.* **2008**, *29*, 1353–1363.

(30) Jang, Y. H.; Goddard, W. A.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. *Chem. Res. Toxicol.* **2002**, *15*, 1023–1035.

(31) Jang, Y. H.; Goddard, W. A.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. *J. Phys. Chem. B* **2003**, *107*, 344–357.

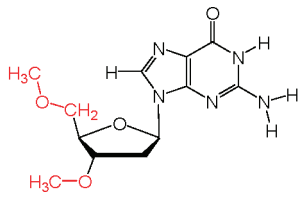


FIGURE 2. Model structure used for B3LYP/6-31G*/PCM(water) geometry optimizations (dG shown as a representative structure; all structures are available in the Supporting Information.) The CH_3-O-CH_2- and CH_3-O- linkages on the 4′ and 3′ carbons, respectively, were included to model the steric bulk of the phosphodiester chain in DNA.

contain the largest oligonucleotide/base pair models investigated to date using quantum mechanical methods and including the effects of solvent.

Monomers. Starting structures for monomer calculations began with bases oriented either *syn* or *anti* relative to sugars. All structures oriented the 4′ axial CH_3-O-CH_2- and 3′ pseudoequatorial CH_3-O- linkages similarly to the phosphodiester arms in the experimental crystal structure of OdG:dC (PDB 183D) or OdG:dA (PDB 178D). Upon geometry optimization, the 4′-methoxymethylene and 3′-methoxy linker arms reoriented slightly with all structures displaying a *trans* rotamer orientation for both sugar substituents. Monomer energies are shown in Table 2 and indicate that for these model structures *syn* and *anti* OdG conformations are relatively isoenergetic, whereas SdG prefers *syn*. Interestingly, for CdG, the preferred base orientation seems to be strongly dependent upon the presence or absence of the 4′-methoxymethylene arm: when the linkage is lacking, the *anti* and *syn* conformations are isoenergetic, but when present, the *anti* conformation is favored by more than 4 kcal/mol. This effect is negated by the addition of the 3′ pseudoequatorial linkage (vide infra).

Base Pairs. To ascertain the importance of base/backbone interactions on the stability of base pairs, we also performed our theoretical study of base pair interactions with and without the 4′ CH_3-O-CH_2- and 3′ CH_3-O- arms that model the phosphodiester backbone. By comparison with the crystal structures of duplexes containing OdG:dC[5] and OdG:dA,[7] we know that our geometry optimized model structures are very similar with respect to the nucleoside-base orientation, but the 4′ linker arm is in a very different position than in the backbone of DNA. This is despite utilizing starting structures with the correct linker orientation. Geometry optimization moves the 4′ linker arm to $\sim 180^o$ orientation, corresponding to a low energy *trans* conformation as would be expected in a steric-free environment (the arm is oriented at 130^o in the crystal structure). This positions the 4′ arm further away from the base than in the corresponding DNA structures. The 3′ CH_3-O- arm in our model compounds maintains a more similar orientation to the crystal structure (152^o crystal vs 168^o model). As shown in Table 3, the energy of interaction is favorable in all cases.

A comparative analysis of the base pair interaction energies indicates that substitution of the sugar significantly affects the energetics of dC containing base pairs but has very little effect on dA containing pairs. A comparison of different base pairs with no substituents on the sugar reveals very similar interaction energies ranging between −6.15 and −7.54 kcal/mol. However, substitution of the sugar with the 3′ and 4′ linkages stabilizes dC containing pairs by ∼5 kcal/mol, while hardly changing the interaction energies of dA containing base pairs. In order to further understand the linkage effects, we also determined the structures and energies for model systems containing only the 3′-methoxy or only the 4′-methoxymethylene arms. In all cases, having only the 3′-methoxy linkage does not change the interaction energies from their values with both arms present. However, having only the 4′-methoxymethylene linkage destabilizes XdG:dC base pairs by 6−10 kcal/mol as compared to when both linkers are present. This effect is not seen in the XdG(*syn*):dA structures.

All optimized base pair structures are planar except SdG:dA and those incorporating CdG, i.e., in each base pair except these, the bases are aligned in the same plane (Table 4). For SdG:dA and CdG containing base pairs, one of the monomers is twisted 3.9−11° relative to the other. For model compounds containing the 4′ CH_3-O-CH_2- linkage, the least stable base pair is CdG:dC (−0.83 kcal/mol). Generally, substitution of the sugar with 3′-methoxy and 4′-methoxymethylene arms does not significantly affect intermolecular base pair geometries. For instance, comparing the average intermolecular distances for base pairs with and without substituents on the sugar, we see that substitution increases the average distances by only 0.001−0.016 Å.

Discussion

Quantum Results Support the Theory That C8 Steric Bulk Can Destabilize XdG:dC Base Pairs. CdG is an interesting nucleoside to study because it lacks significant steric bulk or hydrogen bonding ability at C8 similar to dG but contains an N7-hydrogen similar to OdG. Previous experimental work with dG, OdG, and their analogues[16,32] has suggested that the presence of a large atom off C8 destabilizes base pairs to dC, perhaps due to a steric clash with the backbone sugar that destabilizes the *anti* conformation, leading to a preference for a *syn* conformation. Previous NMR studies support this theory and indicate OdG,[33] and other dG monomers with large atoms off C8, are primarily in the *syn* conformation.[32,34] The quantum results presented here also indicate that large atoms at C8 might cause a preference for the *syn* conformation as evidenced by the *anti* verses *syn* preference for CdG verses SdG when the 4′-methoxymethylene linkage is present (Table 2). It is interesting that the inclusion of the 3′ and 4′ arms stabilize XdG:dC pairs, while the inclusion of only the 4′ CH_3-O-CH_2- linkage destabilizes XdG:dC base pairs by 6.1−10.2 kcal/mol relative to the systems with both the 3′ and 4′ arms present (Table 3). When the two substituents are present, the stabilizing effect of the 3′ arm dominates and even negates the large destabilizing effect of the 4′ linker in CdG:dC. This sensitivity to sugar substitution is not present in the XdG:dA base pairs, further suggesting that an unfavorable interaction with the backbone is playing a role in base pair stability. A careful analysis of the energetic components of base pair E_{int} reveals that the addition of the 4′ arm causes slight destabilization of the

(32) Hamm, M. L.; Rajguru, S.; Downs, A. M.; Cholera, R. *J. Am. Chem. Soc.* **2005**, *127*, 12220–12221.
(33) Culp, S. J.; Cho, B. P.; Kadlubar, F. F.; Evans, F. E. *Chem. Res. Toxicol.* **1989**, *2*, 416–422.
(34) Uesugi, S.; Ikehara, M. *J. Am. Chem. Soc.* **1977**, *99*, 3250–3253.

TABLE 2. B3LYP/6-31G* PCM(water) Nucleoside Energies[a]

	no substituents		4′-methoxymethylene and 3′-methoxy		4′-methoxymethylene		3′-methoxy	
monomer	*E*	Δ*E*	*E*	Δ*E*	*E*	Δ*E*	*E*	Δ*E*
dG(*anti*)	−773.8457475		−1042.197025		−927.6791134		−888.363809	
dG(*syn*)	−773.8456820	−0.04	−1042.196331	−0.43	−927.6784245	−0.43	−888.364011	0.13
CdG(*anti*)	−757.7951492		−1026.148490		−911.6362020		−872.313985	
CdG(*syn*)	−757.7948563	−0.18	−1026.148342	−0.09	−911.6293070	−4.32	−872.313958	−0.02
OdG(*anti*)	−849.0972735		−1117.448744		−1002.930281		−963.615783	
OdG(*syn*)	−849.0969126	−0.23	−1117.448158	−0.37	−1002.929796	−0.30	−963.615390	−0.25
SdG(*anti*)	−1172.056522		−1440.406492		−1325.888170		−1286.575068	
SdG(*syn*)	−1172.059257	1.71	−1440.409658	1.98	−1325.891348	1.99	−1286.577567	1.57

[a]Absolute energies are in atomic units and relative energies are in kcal/mol. Detailed geometries are included in the Supporting Information. Relative energies (Δ*E*) were determined as the difference between the *anti* and *syn* conformers (*anti* − *syn*); a positive difference indicates that the *syn* is favored.

TABLE 3. B3LYP/6-31G* PCM(water) Base Pair Interaction Energies (E_{int}, kcal/mol)[a]

base pair	no substituents	4′-methoxymethylene and 3′-methoxy	4′-methoxymethylene	3′-methoxy
dG:dC	−6.75	−11.61	−5.48	−11.63
CdG:dC	−6.15	−11.01	−0.83	−11.04
OdG:dC	−6.91	−11.76	−5.69	−11.82
SdG:dC	−7.05	−11.82	−5.71	−11.91
CdG:dA	−6.80	−6.79	−6.81	−6.87
OdG:dA	−7.54	−7.27	−7.30	−7.50
SdG:dA	−6.60	−6.37	−6.31	−6.62

[a]Detailed geometries are included in the Supporting Information.

TABLE 4. B3LYP/6-31G* PCM(water) Base Pair Geometries and Hydrogen Bond Lengths[a]

	with 4′-methoxymethylene linkage					no substituents on sugar				
base pair	base pair geo[b]	base pair H-bonds[c]			average[d]	base pair geo[b]	base pair H-bonds[c]			average[d]
dG:dC	planar	1.883	1.930	1.866	1.893	planar	1.885	1.926	1.854	1.888
CdG:dC	twisted	1.896	1.940	1.913	1.916	twisted	1.884	1.928	1.904	1.905
OdG:dC	planar	1.880	1.910	1.861	1.884	planar	1.880	1.909	1.856	1.882
SdG:dC	planar	1.895	1.901	1.844	1.880	planar	1.896	1.901	1.840	1.879
CdG:dA	twisted	1.867	1.944		1.906	twisted	1.869	1.950		1.910
OdG:dA	planar	1.914	1.856	2.911	1.885	planar	1.907	1.864	2.946	1.885
SdG:dA	twisted	1.875	1.904	2.944	1.890	planar	1.879	1.907	3.001	1.893
	with 4′-methoxymethylene and 3′-methoxy linkages					with 3′-methoxy linkage				
base pair	base pair geo[b]	base pair H-bonds[c]			average[d]	base pair geo[b]	base pair H-bonds[c]			average[d]
dG:dC	planar	1.889	1.925	1.859	1.891	planar	1.885	1.926	1.854	1.888
CdG:dC	twisted	1.882	1.931	1.905	1.906	twisted	1.881	1.928	1.906	1.905
OdG:dC	planar	1.877	1.911	1.861	1.883	planar	1.880	1.909	1.858	1.882
SdG:dC	planar	1.901	1.902	1.845	1.883	planar	1.895	1.902	1.844	1.880
CdG:dA	twisted	1.872	1.939		1.905	twisted	1.871	1.940		1.906
OdG:dA	planar	1.914	1.858	2.919	1.886	planar	1.910	1.862	2.948	1.886
SdG:dA	twisted	1.876	1.908	2.946	1.892	twisted	1.882	1.900	2.996	1.891

[a]Geometries are included in the Supporting Information. [b]Planar = intermolecular torsion angles are ±0.5−3.5; twisted = intermolecular torsion angles are ±3.6−11.0. [c]For dC containing base pairs, the first, second, and third hydrogen bonded values correspond to XdG(C6carbonyl O)−dC(C4amine H), XdG(N1H)−dC(N3), and XdG(C2amine H)−dC(C2carbonyl O), and in dA containing base pairs, the first, second, and third hydrogen bonded values correspond to the XdG(C6carbonyl O)−dA(C6amine H), XdG(N7 H)−dA(N1), and XdG(C8carbonyl O/C8thio S)−dA(C2 H) interactions, respectively. A minimum for dG:dA was not found at this level of theory presumably due to a lack of hydrogen bond donor−acceptor functionality. [d]Average distance computations did not include the third nonclassical H-bonding distance for OdG:dA and SdG:dA base pairs.

XdG(*anti*) monomers and, unexpectedly, a slight stabilization of the XdG(*syn*) monomers, resulting in an overall destabilization of XdG:dC pairs relative to the XdG:dA pairs (monomer data without arms available in the Supporting Information). It should be noted that though all XdG:dC systems were destabilized by inclusion of the 4′ linker, pairs with large atoms off C8 (as in OdG:dC and SdG:dC) were not destabilized to a greater extent than those with small atoms off C8 (as in dG:dC); this is likely due to the differences in linker arm orientation between our models and the crystal structure of DNA. It is possible this difference in configuration limits the destabilizing clash between large atoms off C8 and the backbone.

Stability of CdG:dC Base Pairs. Interestingly, though CdG contains only a hydrogen off C8 like dG, CdG:dC base pairs are less stable than dG:dC base pairs (Tables 1 and 3). Previous experimental studies on the CdG monomer have shown some evidence that, like OdG and SdG, it may have some preference for a *syn* base conformation despite not having a large atom off C8.[15] Such a preference may explain the instability of CdG:dC relative to dG:dC; however, the NMR studies used to determine this finding did not take into account the known change in sugar pucker present in C-nucleosides (to increased 2′-endo character as indicated

(35) Oleary, D. J.; Kishi, Y. *J. Org. Chem.* **1994**, *59*, 6629–6636.

by an increased $J_{1'-2'}$),[35] a property that complicates analysis of the preferred *syn*/*anti* base orientation of the CdG monomer and itself may alter base pair stability.[36] Our quantum studies confirm the increased 2′-*endo* character for CdG relative to the other monomers (as indicated by an increased H2′−C2′−C1′−H1′ torsional angle; data shown in the Supporting Information). However, our theoretical results stand in contrast to the previous NMR studies and indicate that either the CdG *syn* and *anti* conformations are isoenergetic or CdG prefers an *anti* conformation when the 4′-methoxymethylene linkage is present. This finding casts further doubt on the hypothesis that the instability of CdG:dC relative to dG:dC is caused by a preference for the CdG *syn* monomer conformation.

Computational results suggest that on these smaller base pair models CdG:dC is less stable than dG:dC due to a number of relatively subtle factors. First, a comparison of the monomer:monomer dipole interactions with the resultant base pair dipoles indicate that there may be an electrostatic basis for the destabilization: the dipolar orientation is more complementary/favorable for the other dC base pairs than for CdG:dC. This is further complicated by an unusually small dipole moment for the CdG(*anti*) monomer (relative to all other monomers including the CdG(*syn*) monomer; dipole data shown in the Supporting Information).

The second factor is that the hydrogen bonding/electrostatic interactions between monomers in the base pair structures are slightly less favorable for CdG:dC relative to the other dC base pairs as judged by the hydrogen-bonding intermonomer distances and perhaps best summarized by the average of those distances (Table 4). For instance, on average, the CdG:dC monomers are 0.023−0.036 Å further apart than all other dC containing base pairs. Third, all other monomers form optimally oriented, planar base pair structures (Figure 3) except for those with CdG likely due to the C-nucleoside nature of the base (vide infra). The most significant CdG:dC destabilization (of ∼10 kcal/mol relative to the model with both arms) occurs in the presence of the 4′-methoxymethylene arm. An analysis of the orientation of the plane of the base relative to the pseudoplane of the sugar reveals that in all of the 4′-methoxymethylene containing systems except CdG:dC, the free monomer orientation is preserved (to within 1−9°) in the optimized base pair structure. However, the base in the CdG(*anti*) monomer (orientation = 67.7°) must torque 20° in the glycosidic torsional angle (O4′−C1′−C9−C8) in order to base pair with dC (orientation of CdG base relative to sugar in base pair = 47.7°). Thus it appears that whereas all the other monomers are "pre-organized" for bonding with their corresponding base pair, CdG(*anti*) must undergo a significant structural rearrangement in order to pair with dC. This builds an inherent strain into the system, resulting in a much less favorable interaction. In this case, the addition of the 4′ linker arm increases the separation between monomers by 0.011 Å *and* changes significantly the orientation of the base to the sugar.

Finally, in our model systems, the reorientation of the 4′-methoxymethylene arm upon geometry optimization allows a favorable bifurcating nonclassical hydrogen bonding interaction between a proton on the linker arm, a proton on the sugar and the O or S off C8 (XdG(*anti*)) or the N3 (XdG(*syn*)). This interaction is not possible in dG:dC or CdG:dC (Figure 3) and is not present in the experimental crystal structures where instead the O or S off C8 interacts unfavorably with the 4′O in the sugar ring. This indicates that while our model compounds are, to the best of our knowledge, the largest base pair structures to be characterized using quantum mechanical methods and including the effects of solvation, they are structurally too small to fully capture the steric effects present in the oligonucleotides studied experimentally here.

It is also noteworthy that our model results suggest that a C-nucleoside is much less flexible and less able to adjust to the local environment than an N-nucleoside. The twisted, nonplanar nature of CdG containing base pairs is one evidence of this. Structural analysis indicates there is a pyramidalization at N9 of up to 8° that is not possible in C-nucleosides where the sp^2 hybridized C9 is locked into a planar geometry (torsional data shown in Supporting Information). This theoretical observation is difficult to verify experimentally; to our knowledge no crystal structure is available containing CdG. However, a comparison with a crystal structure of formycin (PDB 1T8S),[37−39] a C-nucleoside analogue of adenosine, revealed an average C9 pyramidalization of only 0.78°, confirming our findings. Structural data also suggests that the glycosidic bond length in CdG is less able to adjust to the local environment; for instance upon addition of the 4′ linker arm in N-nucleosides the base moves 0.010−0.014 Å away from the sugar, but is only able to move 0.002 Å away in C-nucleosides.

Thus, the quantum studies indicate the lower stability of CdG:dC base pairs as compared to dG:dC base pairs is likely due to the C-nucleoside character of CdG and the various structural and electronic changes that result. These findings are also consistent with previous work that shows oligonucleotide duplexes containing a C-glycoside usually have melting temperatures that are a few degrees lower than an analogous duplex with an N-glycoside.[9]

Stability of CdG:dA Base Pairs. Though CdG contains an N7-hydrogen like OdG and SdG, which should allow it to form relatively stable base pairs to dA, duplexes containing a CdG:dA base pair are less stable than duplexes containing an OdG:dA or SdG:dA base pair (Table 1). This experimentally observed difference in base pair stability may be due to a number of different factors. First, the pK_a at the N7 of CdG is likely much higher in value than that at the N7 of OdG or SdG. Pyrrole (a model compound for the five-membered ring in the base of CdG) has a pK_a value of 17.5,[40] much higher than that of 2-hydroxyimidazole or 2-mercaptoimidazole (comparable analogues for OdG and SdG, respectively), which have pK_a values of 12.8 and 11.3, respectively (see Supporting Information). It is known that the strength of a hydrogen bond increases as the pK_a values for the acceptor

(36) Williams, A. A.; Darwanto, A.; Theruvathu, J. A.; Burdzy, A.; Neidigh, J. W.; Sowers, L. C. *Biochemistry* **2009**, *48*, 11994–12004.

(37) Giranda, V. L.; Berman, H. M.; Schramm, V. L. *Biochemistry* **1988**, *27*, 5813–5818.

(38) Prusiner, P.; Brennan, T.; Sundaral., M *Biochemistry* **1973**, *12*, 1196–1202.

(39) Zhang, Y.; Cottet, S. E.; Ealick, S. E. *Structure* **2004**, *12*, 1383–1394.

(40) Archeson, R. M. *An Introduction to the Chemistry of Heterocyclic Compounds*, 3rd ed.; John Wiley & Sons, Inc.: New York, 1976.

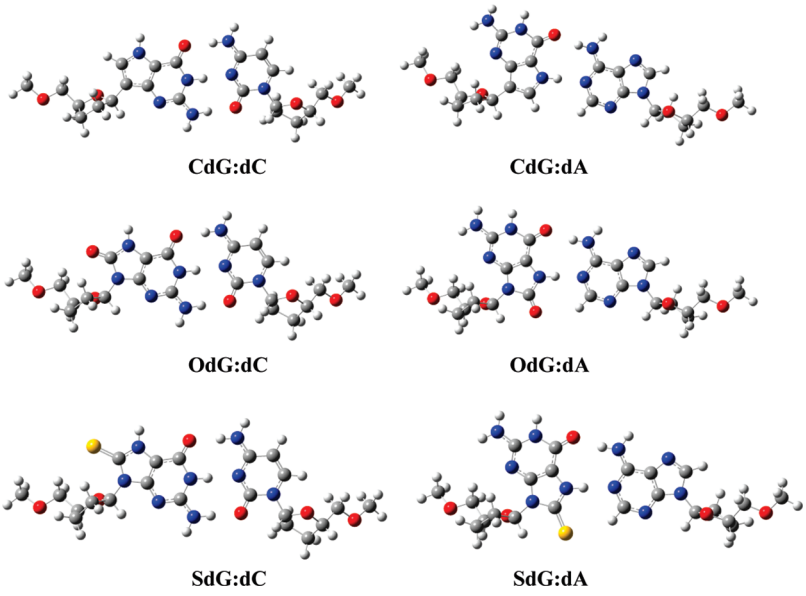


FIGURE 3. B3LYP/6-31G*/PCM(water) optimized base pair structures. For clarity, only model compounds are shown containing 4′ methoxymethylene arms. All other structures may be found in the Supporting Information.

and donor atoms become more similar.[41,42] Since the N1 position of dA (the acceptor) has a pK_a value of 3.5, the N7-hydrogen (the donor) pK_a values for OdG and SdG would be closer than that for CdG, possibly allowing for the formation of a stronger hydrogen bond and stabilizing OdG:dA and SdG:dA base pairs overall.

The instability of CdG:dA base pairs as compared to OdG:dA and SdG:dA base pairs may also be due to a lack of steric bulk at the C8 position of CdG. As explained above, while dG prefers an *anti* conformation about the glycosidic bond, OdG, SdG, and other dG analogues with large atoms at C8 prefer a *syn* base conformation. While the preferred base orientation of CdG is controversial, it is possible that, as compared to OdG and SdG, CdG may be less likely to adopt the *syn* conformation needed for base pairing to dA.

Finally, though computational results suggest that for these model structures, CdG:dA, OdG:dA, and SdG:dA are isoenergetic with interaction energies ranging from −6.20 to −7.54 kcal/mol (Table 3), we do note that the average intermonomer distances in OdG:dA and SdG:dA are similar and ∼0.02 Å shorter than in CdG:dA. We also see the potential for a third, albeit relatively long (∼2.9 Å) electrostatic/nonclassical hydrogen bonding interaction[43,44] present in OdG:dA and SdG:dA that is not possible in CdG:dA (Figure 3). Natural bond orbital analysis and Wiberg Bond Indices (Supporting Information) determined for OdG:dA agree with the structural data and suggest the possibility of a third, albeit weak, long-range interaction. The similarities for OdG:dA and SdG:dA in base pair interaction energy and structure is in accord with previous results suggesting that sulfur is close to oxygen in its hydrogen bond acceptor behavior.[45] The strong correlation between the number of hydrogen bonds formed and base pair stability confirms the results reported previously by Geyer et al.[9]

Monomer Conformational Preferences Are a Strong Predictor of Base Pair Stabilities. When combining the experimental dC and dA base pairing data, we find a trend where CdG forms much more stable base pairs to dC, OdG forms somewhat more stable base pairs to dC, and SdG forms slightly more stable base pairs to dA (Table 1). Though we realize many factors can lead to the observed stability differences in the corresponding duplexes, it is interesting to note that the monomeric conformational preferences of CdG, OdG, and SdG determined computationally are consistent with these findings. With the 4′-methoxymethylene arm present, the CdG monomer has a relatively large energetic preference for the *anti* conformation (Table 2), and experimentally we see a relatively large preference for base pairing to dC where the *anti* conformation is required. The *syn* conformation is more energetically stable than the *anti* conformation in the SdG monomer, and experimentally SdG:dA containing duplexes, where the *syn* conformation of SdG is required, are more stable than SdG:dC containing duplexes. As a note, dG was not addressed in this analysis since it cannot form similar base pairs to dA due to a lack of a N7-hydrogen (a geometry optimization starting with a planar intermonomer arrangement led to a relatively high energy structure containing one hydrogen bond and a twisted (50°) orientation between nucleosides).

Final Note. Differences between the theoretically and experimentally determined stabilities of the various base pairs may be due to very small differences in individual base pair interactions that become amplified by interaction with flanking nucleotides when incorporated into the oligomers, an effect too large for current quantum mechanical methods. The model structures employed here also appear to be too small and flexible to capture the true nature of the duplexes containing damaged DNA base pairs. For instance, contributions from base-stacking are not included, and thus additional effects from structural differences in the base pairs

(41) Chen, J. G.; McAllister, M. A.; Lee, J. K.; Houk, K. N. *J. Org. Chem.* **1998**, *63*, 4611–4619.
(42) Shan, S. O.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14474–14479.
(43) Costa, M. T. C. M. *J. Mol. Struct. THEOCHEM* **2005**, *729*, 47–52.
(44) Khan, A. *J. Phys. Chem. B* **2000**, *104*, 11268–11274.
(45) Wierzejewska, M.; Saldyka, M. *Chem. Phys. Lett.* **2004**, *391*, 143–147.

may not be fully accounted for. For example, the twist between the two bases observed computationally for CdG:dC, CdG:dA, and SdG:dA (Table 4) could weaken stacking interactions with the adjacent base pairs, thus helping to account for the relatively lower melting temperatures observed experimentally with these base pairs (Table 1). Also, the overall structure of DNA positions the backbone arm more rigidly near the base, and this likely prevents small changes in the base−sugar orientation that are possible in our model systems. Hybrid (QM/MM) and molecular dynamics (MD) analyses are currently underway to probe these effects and will be reported elsewhere. It is also possible that the experimental effects are not captured with the approach taken here but necessitate a determination of the free energy including the entropy. It is not currently possible to perform a multiconformer entropic analysis with quantum mechanical methods on model systems of this size while including the effects of solvent.

Conclusion

CdG was incorporated into oligonucleotides, and its base pairing, as compared to dG, OdG, and SdG, was studied through both experimental and computational analysis. These findings shed additional light on the reasons for the promiscuous base pairing, and thus promutegenic character, of OdG and are consistent with current theories that suggest the steric bulk of the C8-oxygen of OdG contributes to the destabilization of OdG:dC base pairs. Additionally, the results suggest OdG:dA base pairs are stabilized not only by the presence of an N7-hydrogen but possibly also by the presence of a third, albeit long-range, electrostatic interaction between the C8-oxygen of OdG and the C2-hydrogen of dA.

Experimental Section

General Methods. Methylene chloride used in chromatography was passed through alumina (Active Basic, Activity I), and Dowex (50 × 4-400) was washed with methanol prior to use. MALDI-TOF and HR-ESI analyses were performed at the University of California-Riverside Mass Spectrometry Facility. Preparative and analytical HPLC were performed using a semiprep C18 column (10 × 250 mm) run at 3 mL/min and an analytical C18 column (4.6 × 250 mm) run at 1 mL/min, respectively. HPLC solvents A and B were 0.1 M triethylammonium acetate (TEAAC) pH 7 and acetonitrile, respectively. Merck silica gel, 200−400 mesh, 60 Å was used for column chromatography.

1-*N*-Benzyl-2-*N*-isopropylphenoxyacetal-9-deaza-2′-deoxyguanosine (2). A 340 mg (0.52 mmol) portion of **1**[15] was coevaporated three times with pyridine to remove any associated water before being covered with argon and put on ice. Dry pyridine (5 mL) was added before 108 μL (0.57 mmol) of isopropylphenoxyacetyl chloride was added dropwise over 5 min while stirring. The reaction was then removed from the ice and stirred for 1.5 h before an additional 50 μL of isopropylphenoxyacetyl chloride was added, and the reaction stirred for 30 min longer at room temperature. The solution was concentration in vacuo and used directly in the next reaction without purification. To the resulting oil, 12 mL of 65/35 pyridine/methanol and 440 μL of 25% sodium methoxide in methanol were added. The reaction was stirred for 1.5 h, quenched with Dowex, filtered, concentrated, and purified by silica gel chromatography using 3−5% methanol in chloroform to yield 220 mg (0.41 mmol; 79%) of **2** as a yellow foam. ^{1}H NMR (DMSO-d_6) δ: 11.50 (s, 1H), 11.04 (s, 1H), 8.40 (s, 1H), 7.21 (s, 1H), 5.22 (m, 1H), 5.17 (dd, 1H), 4.92 (m, 1H), 4.25 (m, 1H), 3.79 (m, 1H), 3.54 (m, 1H), 3.47 (m, 1H), 3.09 (s, 3H), 2.99 (m, 3H), 2.32 (dt, 1H), 1.90 (dd, 1H). ^{13}C NMR (DMSO-d_6) δ:157.4, 155.3, 154.6, 143.6, 126.0, 116.3, 116.1, 87.9, 73.8, 73.7, 63.6, 42.5, 40.8, 34.8. HR-ESI (M + H$^+$) for $C_{29}H_{33}N_4O_6$: expected 533.2395, found 533.2389.

2-*N*-Isopropylphenoxyacetal-9-deaza-2′-deoxyguanosine (3). Compound **2** (220 mg, 0.41 mmol) in 150 mL of methanol was added to 75 mg of palladium hydroxide. The solution was shaken for 15 h under 40 psi of hydrogen at room temperature, filtered over Celite, concentrated, and purified by silica gel chromatography using 4% methanol in chloroform to yield 120 mg (0.27 mmol; 66%) of **3** as a white powder. ^{1}H NMR (DMSO-d_6) δ: 11.96 (b, 1H), 11.6 (b, 2H), 7.37 (s, 1H), 7.17 (d, 2H), 6.88 (d, 2H), 5.20 (dd, 1H), 4.95 (m, 1H), 4.79 (s, 2H), 4.20 (m, 1H), 3.71 (m, 1H), 3.44 (m, 2H), 2.83 (m, 1H), 2.14 (m, 1H), 2.00 (m, 1H), 1.17 (d, 6H). ^{13}C NMR (DMSO-d_6) δ: 171.5, 156.4, 153.0, 145.1, 142.5, 141.6, 127.7, 126.4, 117.4, 115.5, 114.9, 73.0, 87.8, 72.3, 67.2, 63.1, 41.9, 33.1, 24.5. HR-ESI (M + H$^+$) for $C_{22}H_{27}N_4O_6$: expected 443.1925, found 443.1940.

2-*N*-Isopropylphenoxyacetal-5′-*O*-dimethoxytrityl-9-deaza-2′-deoxyguanosine. Compound **3** (120 mg, 0.27 mmol) was coevaporated three times with pyridine to remove any associated water before addition of 137 mg (0.40 mmol) of dimethoxytrityl chloride and 2 mg (0.016 mmol) of 2-(dimethylamino)pyridine. The flask was covered with argon, and 4 mL of anhydrous pyridine and was added. After 1.5 h of stirring at room temperature, the reaction mixture was concentrated, and the resulting oil was purified by silica gel chromatography using 1−2% methanol in chloroform to yield 160 mg of 2-*N*-isopropylphenoxyacetal-5′-*O*-dimethoxytrityl-9-deaza-2′-deoxyguanosine (0.21 mmol; 80%) as a white foam. ^{1}H NMR ($CDCl_3$) δ: 11.83 (b, 1H), 10.89 (b, 1H), 9.15 (b, 1H) 7.4−7.1 (14H), 6.88 (d, 2H), 6.78 (d, 2H), 5.45 (dd, 1H), 4.60 (s, 2H), 4.50 (m, 1H), 4.11 (m, 1H), 3.75 (s, 6H), 3.36 (m, 1H), 3.18 (m, 1H), 2.90 (m, 1H), 2.31 (m, 1H), 1.89 (m, 1H), 1.24 (d, 6H). ^{13}C NMR ($CDCl_3$) δ: 169.7, 158.5, 154.8, 153.2, 149.8, 144.9, 143.4, 142.6, 136.11, 136.08, 130.1, 128.2, 127.80, 127.75, 126.8, 126.4, 123.8, 117.5, 116.0, 114.8, 113.1, 86.2, 86.0, 74.8, 72.1, 67.2, 64.8, 55.2, 41.0, 33.3, 29.7, 24.1. HR-ESI (M + H$^+$) for $C_{43}H_{45}N_4O_8$: expected 745.3232, found 745.3263.

2-*N*-Isopropylphenoxyacetal-5′-*O*-dimethoxytrityl-9-deaza-2′-deoxyguanosin-3′-yl β-Cyanoethyl-*N*,*N*-diisopropylphosphoramidite (4). A 180 mg portion of 2-*N*-isopropylphenoxyacetal-5′-*O*-dimethoxytrityl-9-deaza-2′-deoxyguanosine (0.24 mmol) was coevaporated three times with pyridine to remove any associated water before addition of 41 mg (0.35 mmol) of 4,5-dicyanoimidazole (Chemgenes). The flask was covered with argon, and 5 mL of anhydrous methylene chloride and 113 μL (0.35 mmol) of 2-cyanoethyl-*N*,*N*,*N*,*N*-tetraisopropyl-phosphane (Chemgenes) were added. The reaction was stirred for 40 min at room temperature before addition of 15 mL of ethylacetate. The solution was washed twice with 10 mL of saturated $NaHCO_3$, dried with Na_2SO_4, and concentrated. The resulting foam was purified by silica gel chromatography using 15−25% acetone and 0.1% triethylamine in methylene chloride to yield 110 mg of **4** (0.114 mmol; 47%) as a white foam. ^{31}P NMR ($CDCl_3$) δ: 148.0. HR-MS (M + Na$^+$) for $C_{52}H_{61}N_6O_9P$-Na: expected 967.4129, found 967.4125.

Oligonucleotide Synthesis. All standard DNA was purchased from IDT Technologies. Synthesis of **5** was performed at the University of Virginia Biomolecular Research Facility using all standard procedures. It was then deprotected and cleaved from the column using 29.7% ammonium hydroxide incubated at room temperature for 18 h. MALDI-TOF for **5** (MH$^+$): expected 3236, found 3236.

DNA Purification. All oligonucleotides were purified by 20% denaturing PAGE before UV visualization. The slowest running

band was excised and soaked twice in water, in the dark, for 24 h. The resulting solutions were then concentrated, combined by resuspension in 1 mL of water, and filtered. The oligonucleotides were then further purified by preparative HPLC using a linear gradient of 5−20% solvent B in A over 30 min.

Nuclease Digestion. Compound **5** (0.2 OD_{260}) was incubated for 16 h at 37 °C with 12 μg units snake venom phosphodiesterase (*Crotalus adamanteus*), 2 units bacterial alkaline phosphatase, 32 mM Tris pH 7.5, and 15 mM $MgCl_2$ in a final volume of 80 μL. When the reaction was complete, 10 μL of 3 M sodium acetate pH 7 and 250 μL of ethanol were added, and the solution incubated for 30 min at −78 °C before centrifugation for 20 min at 12,000 rpm. The supernatant was removed, and the remaining solid was dried in vacuo and resuspended in 150 μL of water before analysis by analytical HPLC using a linear gradient of 5−6.5% B in A over 20 min.

Melting Studies. Each oligonucleotide at 5 μM, 1 M NaCl, 0.1 mM EDTA, and 100 mM phosphate buffer pH 7 in a total volume of 1 mL was heated for 5 min at 90 °C. The solution was then allowed to cool at room temperature for at least 30 min and at 4 °C for at least 30 min. The absorbance at 260 nm was monitored from 20−80 °C with the temperature increased at a rate of 0.5 °C/min. The resulting melting curves were analyzed by least-squares fitting to $\Delta H°$ and T_m for a two state stable transition.

Computational Details. Geometry optimizations of base pairs and the corresponding monomers were performed using the B3LYP density functional with the 6-31G* basis set as implemented in the Gaussian03[46] program. Solvent effects were included at every step of the geometry optimization using the polarizable continuum model (PCM)[19] for water. This approach has been shown to adequately describe similar systems.[29] Harmonic frequency analysis was used to confirm all structures as minima on the B3LYP/6-31G* (PCM) potential energy surfaces. All calculations were performed on nucleoside structures containing deoxyribose sugars with and without 4′ CH_3−O−CH_2− and 3′ CH_3−O− linkages intended to mimic the steric bulk of the phosphodiester chain as shown in the representative dG structure in Figure 2. Base pair interaction energies were determined using the formula $E_{int} = E$(base pair) − E(nucleoside 1) − E(nucleoside 2) where a negative E_{int} indicates a favorable interaction. Interaction energies were not corrected for zero point vibration.

Acknowledgment. The authors thank Bojan Dragulev and René Kanters for assistance with DNA synthesis and pK_a determination, respectively. The work was supported by the NSF CAREER (Grant CHE-0239664), NSF RUI (Grant CHE-0809462), and Henry Dreyfus Teacher Scholar Award (M.L.H. and C.A.P.) programs, as well as the HHMI Foundation and the Floyd D. and Elisabeth S. Gottwald Endowment. Support is also acknowledged from the Donors of the American Chemical Society Petroleum Research Fund.

Supporting Information Available: pK_a determination; ^{1}H and ^{13}C NMR; computational results (geometries, energies, dipoles). This material is available free of charge via the Internet at http://pubs.acs.org.

(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian03, rev. B.04*; Gaussian, Inc.: Pittsburgh, PA, 2003.