Published in final edited form as:

J Biomol Struct Dyn. 2008 December ; 26(3): 283–292.

A Computational Study of Expanded Heterocyclic Nucleosides in DNA

Peter I. O'Daniel, Malcolm Jefferson, Olaf Wiest¹, and Katherine L. Seley-Radtke^{*}

¹Department of Chemistry and Biochemistry University of Notre Dame Notre Dame, Indiana 46556-5670, USA

Abstract

The first molecular dynamics study of a series of heterospacer-expanded tricyclic bases in DNA using modified force field parameters in AMBER is detailed. The expanded purine nucleoside monomers have been designed to probe the effects of a heteroaromatic spacer ring on the structure, function, and dynamics of the DNA helix. The heterobase scaffold has been expanded with a furan, pyrrole, or thiophene spacer ring. This structural modification increases the polarizability of the bases and provides an additional hydrogen bond donor with the amine hydrogen of the pyrrole ring or hydrogen bond acceptor with the furan or thiophene ring free electron pairs. The polarizability of the expanded bases were determined by AM1 calculations and the results of the MD simulations of 20-mers predict that the modified curvature of the expanded base leads to a much larger major groove, while the effect on the minor groove is negligible. Overall, the structure resembles A-DNA. MD simulations of 10-mers suggest that the balance between base pairing vs. base stacking and intercalation can be shifted towards the latter due to the increased surface area and polarizability of the expanded bases.

Keywords

Molecular dynamics; Expanded DNA; Tricyclic bases; Expanded bases

Introduction

The relationship of structure, properties, and function in the DNA double helix is a central topic of biochemistry, organic chemistry, and materials science (1-5). Studies from several groups have focused on testing this relationship by changing a variety of structural characteristics. Hydrogen bonding and hydrophobic interactions, stacking effects and shape have all been explored, providing new insights into recognition and function (6-20). The foundation for many of those studies was provided by Nelson Leonard's elegant work (21-23) with the expanded purines, which allowed investigation of the effects on enzyme recognition as well as base pairing.

To expand upon that premise, a series of expanded purine nucleosides have been designed (Figure 1) that offers several advantages over previously studied modifications. In contrast to

^{*}Phone: 410-455-8684 Fax: 410-455-2608 Email: kseley@umbc.edu

Department of Chemistry and Biochemistry University of Maryland Baltimore County 1000 Hilltop Circle Baltimore, Maryland 21250, USA

Supplemental Information Standard Deviations and a complete listing of expanded DNA values for Table I, plots of total energy, volume, density, and backbone RMS over the course of the MD simulation, as well as force field parameters, xleap off files, and PDB files for the modified bases. Supplemental materials can be acquired, free of charge, from the contact author or from Adenine Press for \$50.

Leonard's linear benzene-separated (*lin*-benzo) analogues, utilization of a thiophene, furan, or pyrrole spacer ring results in an "arced" nucleoside that only moderately repositions the hydrogen bonding elements required for recognition and base pairing, while lessening the impact on the width of the helix inherent with the *lin*-benzo analogues. More importantly, the heteroaromatic spacer rings provide a unique opportunity to investigate the implications of additional hydrogen bonding capabilities, as well as to increase stacking interactions as a result of the increase in polarizability of the nucleobases that occurs.

One important aspect of helix stability involves base stacking. Stacking effects have been shown to increase helix stability by approximately 0.4 to 3.6 kcal/mol (24). Two structural conditions required for enhancing intra- and inter-strand stacking effects include increased surface area (25) and increased polarizability of the nucleobase (25,26). In that regard, incorporation of the heteroaromatic spacer rings fulfills both requirements. As a first step towards more fully exploring the intrinsic characteristics of the DNA helix, a series of modified DNA strands containing the expanded purine nucleobases were modeled. Using 4 ns molecular dynamics simulations with a modified AMBER force field, helix stability, base stacking and pairing, as well as the effect of base curvature and expansion of the purine scaffold on the DNA helix were investigated.

Materials and Methods

Electrostatic potentials and polarizabilities were calculated for optimized geometries using Spartan04 (27) at the AM1 level of theory on a Dell Dimension 8200 2-GHz Pentium 4 running Windows XP. The electrostatic potentials are shown with a density of 0.002 electrons/Å³ and a relative range from -77.21 electrons/Å³ (red) to 52.21 electrons/Å³ (blue) with red being more electronegative and blue more electropositive.

All potentials for the tricyclic bases for use in AMBER (28) were derived from geometry optimizations of the modified bases capped with a methyl group to represent the sugar moiety at the B3LYP/6-31G** level of theory, performed using Gaussian98 (G98) (29) on a Beowulf cluster with Xenon processors running Linux. All Amber calculations were run on a Beowulf-type cluster with Xenon processors running Linux. Bases were converted in antechamber from the G98 output to a prepi file using the optimized geometries. Three character codes were devised for each tricyclic base and added to the Amber prepi file. Each prepi file was loaded in xleap, off files were created and partial charges from the G98 output for each atom were added for each perspective tricyclic base. Pdb files of the expanded bases were derived from G98 log files using Molden (30) and the pdb atom lists were rearranged to match the atom listing of the Amber off files. The new parameters are shown in the **Supplemental Information**.

Normal DNA strands were built using the Accelrys Biopolymer module in InsightII (31). The guanosine or adenosine bases were deleted from the DNA strand and the pdb file of the guanosine or adenosine like tricyclic base was loaded into InsightII. The tricyclic bases were connected to the DNA stand using the Biopolymer module. The alignment of the tricyclic base to the complementary pyrimidine base was adjusted by modifying the dihedral angle between the base and the sugar. The extreme overlap was resolved by adjusting the dihedral angle of the phosphate backbone starting from the center base pair of the DNA strand and moving to each perspective end. Then adjusting the dihedral angle of the base/sugar and the phosphate backbone again manually aligned the Watson-Crick hydrogen-bonding pairs. The modified strand was minimized through steepest descent optimization with 100 iterations and a 0.001 derivative in InsightII and the structure was saved as a pdb file.

The modified parameter file, tricyclic base .off file and then the modified DNA strands were loaded into xleap. Missing force constants for unknown bond angles and bond lengths were calculated using parmcal and unknown dihedral angles were derived from equivalent dihedrals in the Amber parameter file or from previous density functional theory and *ab initio* calculations (32-34). The force constants were then added to the xleap parameter file. Sodium counter ions were added using xleap to neutralize the phosphodiester backbone charge and the strand was embedded in an 8 Å layer of water in the shape of a truncated octahedron using the TIP3P water file. The structure was refined using a sequence of minimizations of alternating steps of a constrained minimization, followed by a 400 step free minimization until the constrained minimization ran for approximately 2000 steps. Strands were checked for correct base pairing. If bases were misaligned the strands were readjusted in Insight and then loaded back into Amber and minimized.

All MD simulations were run without restraints with a 1 fs time step for a minimum of 4 ns and data was saved every 1.25 ps. Since CURVES (35) does not recognize the tricyclic bases, the data for the appropriate parameters for individual snapshots had to be extracted manually following a published procedure (24). To analyze the strands and calculate deviations, a minimum of three averaged DNA structures for each strand was used per modeled 20-mer. All end base pairs were excluded from data acquisition due to base buckling or pyrimidine base flipping. The three averaged structures were extracted from the data after the initial backbone movement stabilized (~1-1.5 ns) and where three separate time frames in the plot of the total strand backbone movements were within 1 Å deviation. The results for each parameter from the three structures were averaged and standard deviations were calculated using Microsoft Excel®. Normal DNA strands were also simulated using the modified parameter file to check that the additional parameters did not interfere with original simulation parameters. Data tables and plots of backbone root mean square deviation, energy, temperature, density as well as the pdb base files from the Gaussian output, Amber tricyclic base off files, and parameter file force constants can be found in the *Supplemental Information*.

Results and Discussion

Electrostatic and Surface Potentials

Insertion of a furan, pyrrole, or thiophene ring between the imidazole and pyrimidine rings of the purine scaffold extends the overall length of the bases by ~1.4 Å and the bases curve approximately 30°, 33°, and 43°, respectively, towards the minor groove. As previously mentioned, this structural modification should increase the stacking effects in the helix due to (i) increased polarization, (ii) increased surface area, as well as (iii) greater overlap by the bases.

To investigate the increase in polarization of the heterobase surface area, the dipoles, and surface potentials were calculated for normal guanosine and adenosine bases, as well as the benzene-, furan-, pyrrole-, and thiophene-expanded analogues (Figure 2). For clarity, the sugar moieties were removed and replaced with methyl groups, since our initial calculations (data not shown) indicated that the sugars do not affect the polarizability or dipole moments of the bases. The results indicate that the polarizability of the expanded furan, pyrrole, and thiophene nucleosides was increased as compared to the parent adenosine and guanosine and previously reported (36) *lin*-benzo expanded nucleosides (Figure 2 a-e and k-o). This feature should translate into greater dispersion forces and, along with the extended π -system, should result in an increase in the stacking contributions and the overall stability of the duplex. It is well known (11,25,26) that increasing the polarizability of nucleobases in DNA contributes significantly to stacking effects especially when it is an unpaired dangling base (26) and has also been shown to be crucial for recognition and replication by polymerases (11,26).

As can be seen in Figure 2, the expansion of the bases will increase the surface area. Qualitatively, the energy gained by π -stacking is a function of the surface area of a base that is in contact with the adjacent base if appropriately aligned. The addition of the hetero rings adds at least an additional heteroatom to the surface area and can aid in overall stacking of the bases. Increasing the surface area of unnatural bases also has been shown to enhance the overall stability of DNA (25,37). Increasing the length of the bases can also enhance the stability by increasing the overlap of the bases. As will be shown later, the increase in size increases the ability of the base to overlap more with intra-strand and inter-strand bases. This increase in overlapping has been shown to add to the overall stability of DNA, especially if it is a dangling end (10,38,39).

The electrostatic surface potentials shown in Figure 2 f-j and p-t indicate that the normal hydrogen-bonding base functionalities remain unchanged for all of the expanded nucleosides as compared to the natural bases. Leonard's *lin*-benzo-expanded tricyclic nucleosides (Figure 1 g and q) also show very little change to the electrostatic surface potential compared to the natural nucleosides (Figure 1 f and p), thus this indicates a negligible net gain/loss with regards to the polarizability of the heterobase. For the furan-separated tricyclics **1** and **4** (Figure 2 h and r), an additional electronegative/hydrogen bond acceptor (yellow to red) surface is observed on the top center of the bases. With the pyrrolo-spacer tricyclics **2** and **5** (Figure 2 i and s), a more electropositive/hydrogen bond donor (green to blue) surface area is observed. Both the thiazole tricyclic **3** and **6** (Figure 2 j and t) are slightly more electropositive (green) than the furan expanded nucleosides and is likely due to the increased size and dispersive effect of the sulfur atom. Like the furan spacer ring, the unshared pair of electrons of the thiophene sulfur should act as a hydrogen bond acceptor, the hydrogen bond acceptor (or donor) atoms can also be expected to increase hydrogen bonding to water in the spine of hydration (40) in the major groove, which would also theoretically increase the T_m of a DNA strand. The heteroatom spacer rings are also expected to increase the recognition of the modified oligonucleotide in the major groove by enzymes, as well as a greater chance of triplex and tetraplex strand formation (41). This latter effect has been shown to regulate site specific DNA cleavage, protein binding, and gene expression (41) and has been exploited and studied by several groups since the discovery in 1987 (41-49).

Molecular Dynamics Simulations

A critical question for the design of expanded heterobases is the effect on the structure and dynamics of the DNA duplex. It is surprising that, to the best of our knowledge, no unrestrained MD simulations of expanded DNA have been reported to date, although *ab initio* calculations (50-53) and some minimized structures (18) of expanded nucleosides have been reported. These calculations are important but they do not consider the mobility and stability of a modified DNA strand and are likely to describe a local minimum on the complex energy landscape of DNA. Molecular dynamics calculations at the ns timescale using appropriate parameters have been shown (54-56) to accurately describe both regular and modified DNA. As a result, they provide a promising tool for the investigation of expanded DNA.

Beginning with 20-mer modified DNA strands, the base sequences shown in Figure 3 were simulated for comparison to previously reported DNA structures (36,57). The simulations produced DNA strands with 15 base pairs per turn for most of the strands (Table S1). All strands were stable for the entire 4 ns simulation, and as shown in Figure 4, the average structure for the expanded strand displayed a much wider helix structure than B-DNA and thus can be considered "A-type" DNA. The increase in helix width also resulted in a decrease in the overall height of the structure if compared to an unmodified 20-mer of B-DNA (Figure 4).

The parameter values for B-DNA, A-DNA, as well as the overall maximum and minimum measured values for the expanded DNA's in this study are given in Table I. The full listing of

parameter values and standard deviations for each modified DNA strand can be found in **Supplemental Information**, Table S1. As shown in Table S1, the number of bases per turn was found to be a minimum of 13 for the pyrrole extended purines **2** and **5**, while the other four extended bases surprisingly exhibit 15 bases per turn. This is significantly larger than the standard periodicity for regular B-DNA, but is consistent with the values obtained for a helix twist where one rotation is 360 degrees. The helix pitch per base pair rise for the pyrrole expanded bases was found to be lower than the value obtained from the other two measurements due to the fact that the value was not corrected for base inclination, and as a result, was not considered in the bases per turn calculations.

The base inclinations exhibited by the expanded DNAs are much steeper as compared to natural B-DNA and the inclination (12°) is closer to that of standard ADNA (24). The greater inclination can be attributed to the increase in base length and possibly to the curvature of the bases. Interestingly, it has been shown that an “A-DNA” formation can be attributed to a negative slide and a positive roll movement of the bases in the strand (58). There was no indication of negative slide in any of the simulated strands and this increase in inclination has not been seen in previous extended base minimizations (18,36,57).

The extension of the bases as expected also increased the diameter of the helix. The increase in distance from P-P cutting across the strand ranges from 5.4 to 9.8 Å, which is wider than normal B-DNA, and can only partially be attributed to the heteroatom in the spacer ring since the increase is much larger than would be expected from just the extension of the purine bases. The extra length is more likely a result of the increased curvature of the helix axis. The helix axis curvature can also be attributed to the inherent tip towards the major groove by all of the expanded oligomers. While this slight effect varies, the stacking heights between each base were also affected by small differences in rise distances. This is demonstrated by the finding that the base rise distances were greatest nearer the backbone, ~ 4.1 Å, and decreased slightly towards the ends of the bases near the hydrogen bonding moieties closest to the major groove, ~ 3.1 Å. It should also be noted that even though the backbone charge was neutralized with sodium ions, the increased distance at the backbone could be due to the repulsions by the negative charges on the phosphates. Future studies will have to investigate this effect in more detail.

The increased diameter, as well as the curvature of the tricyclic bases, displaces the bases in the x-axis perpendicular to the helix axis from -4.75 to almost -7.00 Å relative to the center of the helix. This is greater than the observed displacement for A-DNA (-4.1 Å) and results in a deeper and more open major groove as depicted in Figure 4. Although the results of the simulations for B-DNA indicate that the neutralization of the backbone charge by the added sodium ions is sufficient, an effect on the observed widening of the duplex cannot be completely dismissed. Nonetheless, these findings provide a tantalizing picture of the DNA structural effects induced by the modified expanded bases **1-6**.

Although base stacking is very important to structure, one must not exclude sugar pucker. The sugar pucker of the ring for the natural B-DNA is the C2'endo envelope configuration. Although other arrangements are observed, they are less important and do not contribute to the overall configuration to any appreciable extent since the other ring conformations are higher energy and are probably pseudo rotation modes (59). For all of the extended DNA strands, several different sugar conformations were noted but the majority proved to be the C2'endo-C1'exo twist conformation. For A-DNA, the C3'-endo conformation is observed and can be considered as the inverse puckering of B-DNA. Since the simulations indicated that the sugar puckering of the heteroatom expanded DNA is close to the puckering of B-DNA, it is expected that there might be slight changes to this puckering. However, this is dependent on the degree of the puckering and is at present considered to be directed by the increased base stacking. This study

indicates that the expanded DNA strands will exhibit the same pattern and conformation as normal B-DNA and also implies that the positioning of the bases in **1-6** is reminiscent of A-DNA type structure, while the sugar conformations are close to B-DNA, highlighting once again the profound structural changes induced by the modified bases.

The results show that the major groove widths and depths are increased. Following the standard protocol for measuring (24), the depth of the major groove only changed by 1.4-1.8 Å, which is reasonable, since the extension of the bases is within this range. As can be seen from Figure 4, a new definition for the depth may need to be defined for these types of strands since visually the major groove depth is much greater than that distance. In contrast, the width of the major groove increased dramatically, ranging from 4.1 Å in **4** to 10.3 Å in **3**. This can be attributed to the curvature of the bases, which repositioned the phosphate backbone away from the major groove and towards each other in the minor groove. Although the phosphates were moved towards the minor groove, the expanded heterocyclic bases increased the distance to where the minor groove was widened by as much as 2.1 Å. The minor groove depth decreased by a minimum of 1.7 up to 3.4 Å, which is surprising, but can be attributed to the curvature of the tricyclic bases towards the minor groove. The distance from the C1' atoms of base pairs across the minor groove was also measured and it was found that the distances averaged between 10.5 and 11.1 Å, well within the range of the B-DNA distance of 10.7 Å. The minor groove widths also are within a reasonable distance of B-DNA widths, and as such, minor groove recognition and binding by enzymes could be possible even with the decrease in depth. In addition, the increase in the width and depth of the major groove may enable major groove binders and enzymes requiring greater access to the major groove to bind more easily since it has been shown that DNA is bent in order to access the bases (60-62). This, along with the additional hydrogen bonding capabilities of the heteroatoms in the spacer rings should result in greater recognition by proteins.

Another aspect is the alignment and stacking of the hydrogen bonding between base pairs. The simulations show that the base pairs of the extended DNA align correctly with respect to one another. There is also complete overlap of the pyrimidine rings of the tricyclic bases to form *inter*-strand stacking when the tricyclic base on the 3' to 5' strand is above the tricyclic base on the 5' to 3' strand, as shown in Figure 5a. This intermolecular stacking is not observed in normal DNA, as shown in Figure 5c. It is likely this overlap will impart greater stability to the helix by adopting a pseudo zipper-like configuration between the strands. However, when the stacking is reversed, the tricyclic bases Figure 5b stack in an equivalent fashion as their normal counterpart bases Figure 5d. Although the stacking in Figure 5b is equivalent to normal base stacking, the interstrand stacking of the tricyclics in Figure 5a may increase the stability of the DNA.

Since simulating 20 mer DNA strands can be computationally intensive and for the purpose of quickly studying different sequences of bases or other types of modified DNA duplexes, it is desirable to have a more efficient computational model in order to study the effects of nonstandard bases on the structure, function, and dynamics of DNA. For that purpose, we next modeled shorter strands in an effort to investigate a more facile model that would adequately represent the structural characteristics obtained for the 20-mer. Figure 6 shows sequences **7** through **12** that were designed to investigate the consistency between the two strand lengths and were calculated in analogy to **1-6**. The parameters for these DNA strands were consistent with those of the initial modified 20 mer DNA strands indicating that the 10 mer strands would be good representatives for the 20 mers.

The next series of strands studied contained the extended bases A_XT or G_XC placed on the same strand (**13** through **18** in Figure 6). Figure 7 shows the structure of **14**, which is a characteristic result from the simulations of dA_{X10}T₁₀. While the bases in the middle of the

strands as well as the overall structure are comparable to the 20-mers, both termini displayed significant structural changes. With the exception of **16**, all of these extended strands shifted by one base pair leaving a dangling base on each end. This feature allowed us to study the stacking ability of an unpaired pyrimidine base and a tricyclic base. Dangling ends have been observed experimentally and are known to increase the stability of DNA (25,38,39) and increasing the polarizability of the dangling end base also increases the stability of the DNA structure (26). Purine dangling ends are known to stabilize DNA structure better than pyrimidine-dangling ends (25). The dangling pyrimidine bases did not remain stacked over any bases and freely rotated in and out of the DNA helix structure indicating no real stacking effects under these simulated conditions. However, the dangling tricyclic bases on these strands either remained stacked above the penultimate tricyclic base or shifted to form interstrand stacking to both the penultimate base on its strand and the terminal base on its complementary strand as shown in Figure 8. When inter strand stacking of the dangling end tricyclic occurred, a shift back to intrastrand stacking was not found. This indicated a greater ability of the tricyclic base to increase the stability of the strand.

Two of the tricyclic bases were also observed to intercalate between their complementary bases in the simulations for **17** and **18**, and involved the 5'-terminal tricyclic base inserting in between the 3'-terminal cytosine and the penultimate cytosine/tricyclic G base pair. The intercalation of **17** is shown in Figure 9, which was initiated by a slide of the base pair pushing the terminal cytosine partially out of the stacking arrangement from the penultimate cytosine base. This enabled the terminal tricyclic G to stack over both penultimate bases. The terminal cytosine then slipped on top to complete the intercalation process. This phenomenon was only observed for dG_{X10}C₁₀ and dA_{X10}T₁₀, indicating that the balance between the correct stacking throughout the strand and the increased stacking interaction at the dangling ends is shifted towards the latter in the shorter sequences of the expanded bases. Thus, intercalation of terminal bases should also increase the stability of DNA by increasing the intramolecular stacking. It is at this time not known whether the base pair slipping is an artifact of the simulation or is due to excess energy in the molecule system. The fact that these results were obtained independently in a number of simulations and are also found in previous experimental observations on other systems (63), increases our confidence in the validity of these results. On the other hand, solution NMR studies of the *lin*-benzo expanded bases (57) do not report intercalation, but rather an increased flexibility of the terminal bases. Ultimately, experimental studies on these new analogues are needed to determine if dangling ends do readily form for these strands.

Conclusions

Molecular dynamics simulations indicate that all of the expanded DNA strands will have increased structure parameters and will form “A-DNA”-type helices, even though the sugar residues mostly retain their B-DNA conformations. The heteroaromatic spacer rings of the tricyclic bases increase the polarizability and the additional hydrogen bonding moieties will allow for new interactions in the major groove of the helix. The data shows that the increased overlap and polarizability should increase the stacking ability of the bases in the helix structure and increase the overall stability of DNA. The data also indicates that the modified DNA minor groove is comparable to the minor groove in natural B-DNA, while the major groove is greatly expanded. Simulations of shorter strands suggest that the balance of base pairing and stacking is shifted in short fragments towards the latter, leading to base intercalation and dangling ends. The results of this computational study provide some fascinating predictions about the structural effects of the expanded bases on duplex DNA. We are currently in the process of synthesizing and fully characterizing these modified DNA strands and the results of these studies will be reported elsewhere.

Acknowledgements

We gratefully acknowledge financial support of this research by the National Institutes of Health (Grant #GM 073645-01 to K.L. S-R) and the National Science Foundation (Grant DMR-0079647 to O.W.).

References and Footnotes

1. Liu, XD.; Yamada, M.; Matsunaga, M.; Nishi, N. *Functional Materials and Biomaterials: Functional Materials Derived from DNA*. Springer; 2007. p. 149-178.
2. Gardiner EJ, Hunter CA, Lu X-J, Willett P. *J Mol Biol* 2004;343:879–889. [PubMed: 15476807]
3. Berg MA, Coleman RS, Murphy CJ. *Phys Chem Chem Phys* 2008;10:1229–1242. [PubMed: 18292856]
4. Conroy RS, Danilowicz C. *Contemp Phys* 2004;45:277–302.
5. Tinoco I Jr. *J Phys Chem B* 1996;100:13311–13322.
6. Kool ET, Morales JC, Guckian KM. *Angew Chem Int Ed Engl* 2000;39:990–1009. [PubMed: 10760910]
7. Guckian KM, Schweitzer BA, Ren RX-F, Sheils CJ, Paris PL, Tahmassebi DC, Kool ET. *J Am Chem Soc* 1996;118:8182–8183.
8. Matray TJ, Kool ET. *J Am Chem Soc* 1998;120:6191–6192.
9. Guckian KM, Krugh TR, Kool ET. *J Am Chem Soc* 2000;122:6841–6847.
10. Burkard ME, Kierzek R, Turner DH. *J Mol Biol* 1999;290:967–982. [PubMed: 10438596]
11. Henry AA, Yu C, Romesberg FE. *J Am Chem Soc* 2003;125:9638–9646. [PubMed: 12904030]
12. Kool ET. *Ann Rev Biophys Biomol Struc* 2001;30:1–22.
13. Lai JS, Qu J, Kool ET. *Angew Chem Int Ed Engl* 2003;42:5973–5977. [PubMed: 14679546]
14. Lin K-Y, Matteucci MD. *J Am Chem Soc* 1998;120:8531–8532.
15. Matsuda S, Henry AA, Schultz PG, Romesberg FE. *J Am Chem Soc* 2003;125:6134–6139. [PubMed: 12785844]
16. Minakawa N, Kojima N, Hikishima S, Sasaki T, Kiyosue A, Atsumi N, Ueno Y, Matsuda A. *J Am Chem Soc* 2003;125:9970–9982. [PubMed: 12914460]
17. Newcomb LF, Gellman SH. *J Am Chem Soc* 1994;116:4993–4994.
18. Krueger AT, Lu H, Lee AHF, Kool ET. *Acc Chem Res* 2007;40:141–150. [PubMed: 17309194]
19. Heckel A. *Chem Bio Chem* 2004;5:765–767.
20. Liu H, Gao J, Kool ET. *J Org Chem* 2005;70:639–647. [PubMed: 15651812]
21. Leonard NJ. *Biopolymers* 1985;24:9–28. [PubMed: 3886035]
22. Leonard NJ, Hiremath SP. *Tetrahedron* 1986;42:1917–1961.
23. Lessor RA, Gibson KJ, Leonard NJ. *Biochemistry* 1984;23:3868–3873. [PubMed: 6487584]
24. Bloomfield, VA.; Crothers, DM.; Tinoco, JI. *Nucleic Acids: Structure, Properties, and Functions*. University Science Books; 2000. p. 794
25. Guckian KM, Schweitzer BA, Ren RX-F, Sheils CJ, Tahmassebi DC, Kool ET. *J Am Chem Soc* 2000;122:2213–2222.
26. Rosemeyer H, Seela F. *J Chem Soc Perkin Trans* 2002;2:746–750.
27. Spartan '04. Wavefunction, Inc.; Irvine, CA:
28. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J., et al. *AMBER 8*. University of California; San Francisco: 2004.
29. Frisch, MJ.; Trucks, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA., et al. *Gaussian 03, Revision C.02*. Gaussian, Inc.; Wallingford CT: 2004.
30. Schaftenaar G, Noordik JH. *J Comput-Aided Mol Design* 2000;14:123–134.
31. InsightII, Biopolymer. Accelrys, Inc.; San Diego, CA: 2001.
32. Kupka T, Wrzalik R, Pasterna G, Pasterny K. *J Mol Structure* 2002;616:17–32.
33. Lohr S, Yonemura M, Orita A, Imai N, Akashi H, Otera J. *Acta Cryst* 2003;E59:594–595.
34. Simandiras ED, Handy NC, Amos RD. *J Phys Chem* 1988;92:1939–1942.

35. Lavery R, Sklenar H. Curves Helical Analysis of Irregular Nucleic Acids 2005;5.1Paris
36. Liu H, Gao J, Maynard L, Saito D, Kool ET. J Am Chem Soc 2004;126:1102–1109. [PubMed: 14746479]
37. Kim TW, Kool ET. J Org Chem 2005;70:2048–2053. [PubMed: 15760186]
38. Ohmichi T, Nakano S-I, Miyoshi D, Sugimoto N. J Am Chem Soc 2002;124:10367–10372. [PubMed: 12197739]
39. Bommarito S, Peyret N, SantaLucia JJ. Nucleic Acids Res 2000;28:1929–1934. [PubMed: 10756193]
40. Lan T, McLaughlin LW. J Am Chem Soc 2000;122:6512–6513.
41. Stilz HU, Dervan PB. Biochemistry 1993;32:2177–2185. [PubMed: 8443159]
42. Wang W, Purwanto MGM, Weisz K. Org Biomol Chem 2004;2:1194–1198. [PubMed: 15064798]
43. Lengeler D, Weisz K. Nucleosides & Nucleotides 1999;18:1657–1658.
44. Dervan PB. Bioorg Med Chem 2001;9:2215–2235. [PubMed: 11553460]
45. Griffin LC, Kiessling LL, Beal PA, Gillespie P, Dervan PB. J Am Chem Soc 1992;114:7976–7982.
46. Horne DA, Dervan PB. J Am Chem Soc 1990;112:2435–2437.
47. Purwanto MGM, Lengeler D, Weisz K. Tetrahedron Lett 2002;43:61–64.
48. Purwanto MGM, Weisz K. J Org Chem 2004;69:195–197. [PubMed: 14703398]
49. Spackova NA, Cubero E, Sponer J, Orozco M. J Am Chem Soc 2004;126:14642–14650. [PubMed: 15521784]
50. McConnell TL, Wetmore SD. J Phys Chem B 2007;111:2999–3009. [PubMed: 17388411]
51. Sharma P, Singh H, Sharma S, Singh H. J Chem Theory Comput 2007;3:2301–2311.
52. Fuentes-Cabrera M, Sumpter BG, Wells JC. J Phys Chem B 2005;109:21135–21139. [PubMed: 16853737]
53. Fuentes-Cabrera M, Lipkowski P, Huertas O, Sumpter BG, Orozco M, Luque FJ, Wells JC, Leszczynski J. Int J Quant Chem 2006;106:2339–2346.
54. Miaskiewicz K, Miller J, Cooney M, Osman R. J Am Chem Soc 1996;118:9156–9163.
55. Spector TI, Cheatham TE III, Kollman PA. J Am Chem Soc 1997;119:7095–7104.
56. Park H, Zhang K, Ren Y, Nadj S, Sinha N, Taylor J-S, Kang C. Proc Natl Acad Sci USA 2002;99:15965–15970. [PubMed: 12456887]
57. Liu H, Lynch SR, Kool ET. J Am Chem Soc 2004;126:6900–6905. [PubMed: 15174859]
58. Lu X-J, Olson WK. Nucleic Acids Res 2003;31:5108–5121. [PubMed: 12930962]
59. Levitt M, Warshel A. J Am Chem Soc 1978;100:2607–2613.
60. Wolfe SA, Ferentz AE, Grantcharova V, Churchill ME, Verdine GL. Chem Biol 1995;2:213–221. [PubMed: 9383423]
61. Dickerson RE. Nucleic Acids Res 1998;26:1906–1926. [PubMed: 9518483]
62. Cloutier TE, Windom J. Mol Cell 2004;14:355–362. [PubMed: 15125838]
63. Brotschi C, Leumann CJ. Angew Chem Int Ed Engl 2003;42:1655–1658. [PubMed: 12698469]

Figure 1.
Expanded nucleoside monomers of adenosine (A) and guanosine (G).

Figure 2.

Base polarization and electrostatic surface potentials. Columns left to right are (1) adenosine (A) and adenosine derivative (A_X) polarizabilities (2) A and A_X electrostatic surface potentials (ESP) (3) guanosine (G) and guanosine derivative (G_X) polarizabilities and (4) G and G_X ESP. Rows from the top down are (1) natural bases, (2) lin-benzo bases, (3) furan expanded bases, (4) pyrrole expanded bases, and (5) thiophene expanded bases. Bar units are in electrons/Å³.

$$\text{dG}_x\text{CG}_x\text{G}_x\text{CG}_x\text{CCG}_x\text{CG}_x\text{CG}_x\text{G}_x\text{CG}_x\text{CCG}_x\text{C}$$

$$\text{G}_x = \text{G}_O, \text{G}_{\text{NH}}, \text{G}_S$$

$$1 \quad 2 \quad 3$$

$$\text{dA}_x\text{TA}_x\text{A}_x\text{TA}_x\text{TTA}_x\text{TA}_x\text{TA}_x\text{A}_x\text{TA}_x\text{TTA}_x\text{T}$$

$$\text{A}_x = \text{A}_O, \text{A}_{\text{NH}}, \text{A}_S$$

$$4 \quad 5 \quad 6$$

Figure 3.

Base sequences for the modeled 20-mer DNA strands. Where G_O , G_{NH} , G_S are the Guanosine furan, pyrrole, and thiophene expanded bases and A_O , A_{NH} , A_S are the adenosine furan, pyrrole, and thiophene expanded bases, respectively.

Figure 4.
Representative averaged structure for the tricyclic-modified DNA vs. B-DNA for the same relative base pair sequence.

Figure 5. Interstrand base pair stacking in the thiophene expanded guanosine (**a**) and (**b**) and normal interstrand DNA stacking in B-DNA (**c**) and (**d**).

Figure 6.
Modeled 10-mer DNA strands.

Figure 7.
Tricyclic dangling ends in the case of **14**.

Figure 8.
Tricyclic base intermolecular stacking shown with the thiophene expanded adenosine.

Figure 9.
Base intercalation by dGNHC.

Table I

MD results for Modified DNA[†] of the overall minimum and maximum parameter values from the strands in this computational study compared to the standard B- and A-DNA parameter values. Full parameter value results can be found in the Supplemental information.

Parameters/Strand	B-DNA(24)	A-DNA(24)	Minimum	Maximum
Helix handedness	Right	Right	Right	Right
bp/repeating unit	1	1	1	1
bp/turn	10	11	13	15
Helix twist, (°)	36	32.7	24.5	30.1
Rise/bp, (Å)	3.4	2.9	3.5	3.6
Helix pitch, (Å)	34	32	38.1	45.4
Base pair inclination, (°)	2.4	12	10.1	16.5
P distance from helix axis, (Å)	9.4	9.5	11.8	13.8
X displacement from bp to helix axis, (Å)	0.8	-4.1	-4.8	-6.9
Glycosidic bond orientation	anti	anti	anti	anti
Sugar conformation*	C2'-endo	C2'-endo		
Major groove depth	8.5	13.5	9.9	10.3
width (Å)	11.7	2.7	15.81	22.0
Minor groove depth	7.5	2.8	4.1	5.8
width (Å)	5.7	11	6.2	7.8
C1'-C1' distance (Å)	10.7		10.5	11.1
Diameter P-P (Å)	18.4	26	23.8	28.2

[†] Standard deviation and full data can be found in the Supporting Information.

* Range of conformations for calculation, see Supporting Information.