THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Theoretical Study on Steric Effects of DNA Phosphorothioation: B-Helical Destabilization in Rp-Phosphorothioated DNA

Yi-Chao Zhang,[†,∥] Juan Liang,[†,∥] Peng Lian,[†] Yiwen Han,[†,⊥] Yifan Chen,[†,⊥] Linquan Bai,[†] Zhijun Wang,[†] Jingdan Liang,[†] Zixin Deng,[†] and Yi-Lei Zhao*[,†,‡,§]

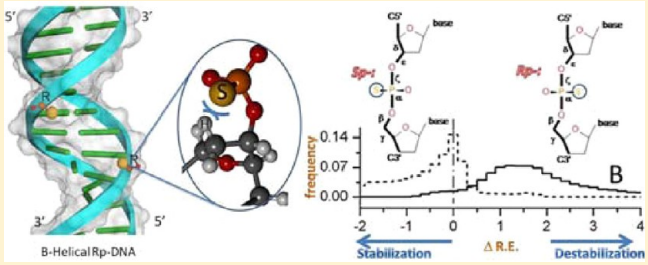
[†]State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, P. R. China

[‡]Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, 200240, P.R. China

[§]MOE Key Laboratory of Scientific and Engineering Computing, Shanghai Jiao Tong University, Shanghai, 200240, P.R. China

Supporting Information

ABSTRACT: Phosphorothioation, with sulfur replacing a nonbridging oxygen of phosphate, has surfaced in bacterial DNA electrophoresis. To understand structural characteristics of the thio-substituted DNA, we have investigated the correlation between the relative energy of phosphate/phosphorothioate linkage and the backbone torsions. The relative energies (R.E.) computed by the quantum mechanical method, the PBE1PBE(CPCM, solvent=water)//PBE1PBE/6-31+G(2df) level of theory, were used to construct energy-scoring functions against backbone torsion variables, resulting in the squared correlation coefficients r^2 of 0.90−0.95. Then, the DNA energy alteration by phosphorothioation is estimated with the relative energy difference (ΔR.E.) between phosphate and phosphorothioate of the phosphate linkages in the DNA crystallographic database (NDB). As a result, Rp-phosphorothioation shifts the relative energy of B-helical structures by 2.7 ± 3.4 kcal/mol, destabilizing about 95% linkages, while Sp-phosphorothioation by −1.4 ± 2.4 kcal/mol, stabilizing over 84% linkages in the data sets. The B-helical destabilization is likely caused by the steric effect between the sulfur atom of Rp-phosphorothioate and the neighboring C−H groups of deoxyribose on the groove wall in B-helix. The unfavorable interaction may be magnified by the increasing rigidness of P−O-involving backbone torsions α and ζ upon the nonbridging phosphorothioations. Since B-helix is the most prevalent DNA double-helical structure and Rp-phosphorothioation is the exclusive configuration in bacteria thio-DNA found to date, the observed stereospecificity-destabilization correlation may reflect a structure−function relationship of biological DNA-phosphorothiation.

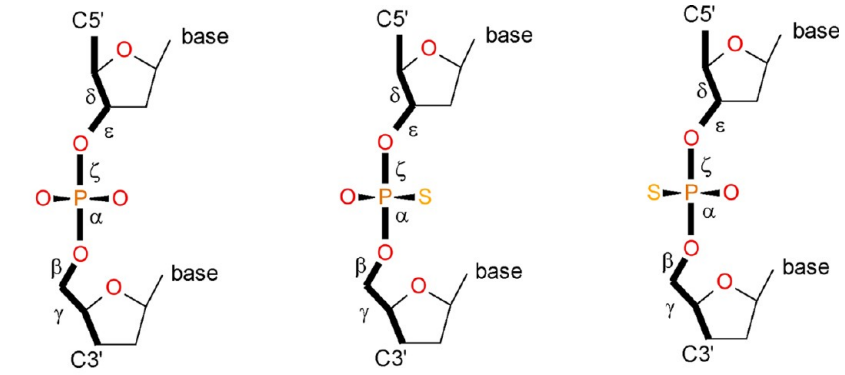


INTRODUCTION

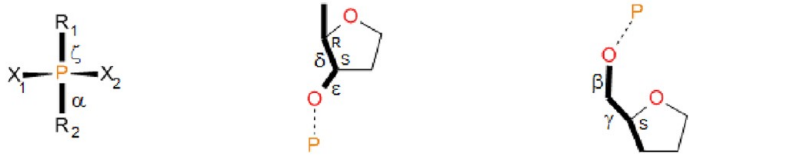
The DNA degradation, namely *Dnd* phenotype, first observed during electrophoresis of genomic DNA from *Streptomyces lividans*, was thought to involve a postreplicative DNA modification.[1] The molecular basis for the degradation is the endogenous replacement of a nonbridging oxygen atom in phosphate linkage on DNA backbone by a sulfur atom.[2,3] Further biochemical experiments have shown that the endogenous modification in bacteria and archarea is sequence-selective and stereospecific.[4,5] Although a diversity of consensus sequences were found nowadays to be related to the endogenous modification on the genome, DNA phosphorothioate is in the Rp-configuration exclusively and no Sp-configuration is observed yet. The stereoselectivity may concern the chiral preference of the enzyme system during the DNA chemical modification or special structural needs for certain biological functions, or both.[6] For instance, it has been found that phosphorothioated DNA functions similarly to methylated DNA under restriction nuclease.[7−9] Methylated DNA can affect the helicity of DNA, implying a certain relationship between DNA helicity and functions.[10,11]

X-ray crystallography suggests three basic types of helical conformations for the known DNA structures, i.e., A-, B-, and Z-form.[12] B-helix is the most prevalent under common conditions, the more compact A-helix appears in the presence of salt or organic solvent, and the alternating purine−pyrimidine sequence of DNA can adopt the Z-helical conformation under certain conditions. Besides the regular antiparallel double-helical structures, DNA occasionally forms a quadruplex, i-motif or cytosine tetraplex, four-way junction, triple helix, adenine zipper, and so on.[13] Typically, the conformations of DNA are a result of a balance of internal and external interactions, including backbone preference, hydrogen bonding between two strands, pi−pi stacking between aromatic groups of base pairs, solvent effect, salt

Received: January 16, 2012
Revised: August 1, 2012
Published: August 3, 2012



Phosphate **Rp-Phosphorothioate** **Sp-Phosphorothioate**



(2*R*, 3*S*)-2-methyl-3-tetrahydrofuranol
MTHF

(*S*)-2-tetrahydrofuranmethanol
THFM

Model	X1	X2	R1	R2
1 DMP	O	O	OCH_3	OCH_3
2 DMPOH	OH	O	OCH_3	OCH_3
3 DMPT	O	S	OCH_3	OCH_3
4 DMPT-OH	OH	S	OCH_3	OCH_3
5 DMPT-SH	O	SH	OCH_3	OCH_3
6 DMTHFP	O	O	MTHF	THFM
7 *R*-DMTHFPT	O	S	MTHF	THFM
8 S-DMTHFPT	S	O	MTHF	THFM

Figure 1. Backbone torsion pattern and calculated diester models for normal and phosphorothioate DNA.

effects, etc. Computational chemists have developed a few force field methods to investigate these interactions in DNAs and RNAs.[14−16] In this work, we try to unveil phosphorothioate effects on DNA helicity, via studying the backbone preference.

DNA helicity can be characterized by a set of backbone torsions: α (O3′−P−O5′−C5′), β (P−O5′−C5′−C4′), γ (O5′−C5′−C4′−C3′), δ (C5′−C4′−C3′−O3′), ε (C4′−C3′−O3′−P), and ζ (C3′−O3′−P−O5′), as shown in Figure 1. It should be noted that torsion around glycosidic bonds (such as χ) is also related to the DNA helicity, via coupling with the backbone torsions. Berman, Schneider, and co-workers have studied the overall torsion distributions of DNA structures in the Nucleic Acid Database.[12,17] On the basis of analysis of 96 crystal structures of oligodeoxynucleotides, Berman et al. reported a series of torsion mean values of (α, β, γ, δ, ε, ζ) for high-frequent conformations of A-, B-, and Z-helices, i.e., (293°, 174°, 56°, 81°, 203°, 289°), (298°, 176°/146°, 48°, 128°/144°, 184°/246°, 265°/174°), and (71°/201°/168°, 183°/225°/166°, 54°/179°, 95°/141°/148°, 189°/240°/267°, 52°/75°/301°), respectively.[12b] The multiple modes of B- and Z-helices reflect subtype conformations such as BI/BII, ZI/ZII clustering. Schneider and co-workers then identified more subclasses, such as AI/AII, A/B and B/A mix, RestA, and RestB conformation by analyzing 447 crystal structures, clarifying (295°, 173°, 54°, 82°, 206°, 285°), (146°, 192°, 183°, 85°, 197°, 289°), (299°, 179°, 48°, 133°, 182°, 263°), and (293°, 143°, 46°, 143°, 251°, 168°) as variants of AI, AII, BI, and BII, repectively.[12c] In Schneider's papers, the BI-form with the torsion set of (299°, 179°, 48°, 133°, 182°, 263°) is the most popular and stable DNA form in the crystal data.

Because it is impossible to distinguish the origin of backbone preference based on limited experimental data, quantum mechanical (QM), semiempirical, and molecular mechanical methods provide a promising means to understand the geometrical and energetic properties.[18−30] High level *ab initio* QM calculations, including density functional theory and Møller−Plesset perturbation theory, have been employed for decades to study the electron correlation energy in base-pair conformational properties of DNA, providing meaningful energetics for base pairing and stacking.[23−26] The resulting QM energetic of the torsional contribution has been incorporated into the molecular mechanical programs such as AMBER[27,28] and CHARMM.[29,30] For example, a set of in-house parameters was developed to simulate the double-helical structure of a specific sequence, $d(GC)_5d(GC)_5$.[31] It should be noted that the relationship of helicity and chirality is not yet clear in the cases of phosphorothioated DNA.

In this study, a series of models (as shown in Figure 1), including dimethyl phosphate anion (DMP, model **1**, methoxy-phosphate-methoxy), dimethyl phosphoric acid (DMPOH, model **2**, methoxy-phosphate_OH-methoxy), *O*,*O*-dimethyl phosphorothioate anion (DMPT, model **3**, methoxyl-phosphorothioate-methoxy), *O*,*O*-dimethyl *O*-hydrogen phosphorothioic acid (DMPT−OH, model **4**, methoxy-phosphorothioate_OH-methoxy), *O*,*O*-dimethyl *S*-hydrogen phosphorothioic acid (DMPT-SH, model **5**, methoxy-phosphorothioate_SH-

dx.doi.org/10.1021/jp302494b | *J. Phys. Chem. B* 2012, 116, 10639−10648

methoxy), (2*R*,3*S*)-2-methyl-3-tetrahydrofuranol (*S*)-2-tetrahydrofuranmethanol phosphate anion (DMTHFP, model **6**, sugar−phosphate-sugar), (2*R*,3*S*)-2-methyl-3-tetrahydrofuranol (*S*)-2-tetrahydrofuranmethanol (*R*)-phosphorothioate anion (R-DMTHFPT, model **7**, sugar-R_phosphorothioate-sugar), and (2*R*,3*S*)-2-methyl-3-tetrahydrofuranol (*S*)-2-tetrahydrofuranmethanol (*S*)-phosphorothioate anion (S-DMTHFPT, model **8**, sugar-S_phosphorothioate-sugar), were computed at the levels of density functional theory (DFT) and complete basis set method (CBS) as well. The CBS-QB3 calculation on the pK_a difference between phosphate and phosphorothioate verifies the reliability of the PBE1PBE (CPCM, solvent=water) energy calculation with the geometry optimized at the PBE1PBE/6-31+G(2df) level of theory, i.e., the PBE1PBE-(CPCM, solvent=water)//PBE1PBE/6-31+G(2df) method used in this work. Both of the theoretical methods confirm the anionic state for the phosphorothioate diester under physiological conditions, similar to phosphate diester in normal DNA. The PBE1PBE method was then chosen to compute models **6**−**8** to unmask the interaction between the deoxyribose (pentose) and phosphorothioate. The resulting relative energies of DMTHFP, R-DMTHFPT, and S-DMTHFPT (models **6**−**8**) were used to build up energy-scoring functions to estimate the Rp and Sp thio-substituent effect on backbone preference. Our calculations show that steric hindrance of neighboring pentose could lead to significant repulsion to the sulfur atom in the case of B-helical Rp-DNA, while no such effect was observed in the case of Sp-configuration.[32]

METHODS

Conformational spaces of models **1**−**5** were computed by systematic conformation search (SCS), using an interval of 5° for the rotatable torsions, and all space-allowed structures were preliminarily minimized by the MMFF force field[33] and MMFF partial charge in Discovery Studio 3.0.[34] The resulting nonredundant conformations were further optimized at the level of the hybrid B3LYP functional and the basis set of 6-311++G** in the gas phase. The low-energy conformations were used as a training set to search for an efficient method for the larger systems with sugar. In the functional-basis-set-testing calculations, the relative electronic energies were used as criteria in the assessment of combinations of functional and basis set. The extensive combinations of 12 exchange-correlation functionals and 308 basis sets in the Gaussian 09 package[35] were tested. Compared to the time-demanding calculations with the largest basis set of aug-cc-pVQZ and the corresponding functional, the combination of the PBE1PBE hybrid functional and the 6-31+G(2df) basis set provides values with a small deviation (0.04 kcal/mol) with respect to the references. In addition, the pK_a calculation of the diesters (models **1** and **3**) suggests that the CPCM solvent model works well with the combination. Therefore, the PBE1PBE(CPCM, solvent=water)//PBE1PBE/6-31+G(2df) method was chosen to compute relative energies in models **6**−**8**.

The low-energy conformations of models **1**−**5** were verified as energy minima by harmonic vibrational frequency analysis. Geometries and energies were also computed with the CBS-QB3 method and continuum CPCM model in the aqueous solution. The solvation energies are calculated with the PBE1PBE(CPCM)//PBE1PBE/6-31+G(2df) method with Pauling (Merz−Kollman) atomic radii. Since phosphorothioate diester is more acidic than phosphate diester, the anionic form was used in the following calculations of models **6**−**8**.

The conformational space of models **6**−**8** was prepared with the following procedure: the MMFF-resulting structures from conformation search were reoptimized at the PBE1PBE/6-31+G(2df) level of theory, and electronic energies of the nonredundant 649 structures were calculated at the PBE1PBE-(CPCM, solvent=water) level of theory (Tables S1−S3, Supporting Information). Backbone torsions of all the calculated conformations were measured, and scoring functions of relative energies were then formulated against the torsions according to the linear regression algorithm.

The torsion angle sets of double-helical structures were collected from the Nucleic Acid Dababase (http://ndbserver.rutgers.edu).[12d] 2847 crystal DNA structures with double-helical features were filtered for energy-scoring analysis. 103686 sets of torsion angles belong to 296 A-form PDBs (6333 sets), 2223 B-form PDBs (78586 sets), and 86 Z-form PDBs (939 sets), respectively. 18428 sets of torsion angles from 519 PDB structures were classified as "other" due to multiple annotations in the PDB structure such as A/B mix, general right-handed helix, loop, quadruple helix, triple helix, single strands. Thus, there is no overlap among the A-, B-, and Z-data sets used in this work; the data sets from the Rutgers server were used to estimate the relative energy shift upon virtual thio-substitution, without any further treatment.

It should be noted that the advanced quantum-mechanical methods such as PBE-D, M06, MP2, and CBS may describe dispersion energy better, as Svozil and Sponer and co-workers pointed out.[19a] For the verification of the conclusions, we have carried out the M06 and MP2 calculations with the CPCM model[35] and the PBE-D calculation[44] with the COSMO model[44c] (Tables S7−S10, Supporting Information).

RESULTS AND DISCUSSION

High Acidity of Phosphorothioate Diester. We computed models **4** and **5** at the high level of the quantum mechanistic method (CBS-QB3). Model **4** is lower in energy than model **5** by 2.3 kcal/mol in the aqueous solution. The selected density functional theory method, PBE1PBE(CPCM solvent=water)//PBE1PBE/6-31+G(2df), results in energy differences of 1.8 kcal/mol in the gas phase and 2.1 kcal/mol in aqueous solution, respectively. It indicates that the acid would be dominant (98%) in the form of model **4**, with sulfur double bond and hydroxyl group [>P(=S)OH]. The results are consistent with the previous studies by Karplus and co-workers.[36] It should be noted that the negative charge favors S atom in the anionic form [>P(=O)S$^-$]. Such electronic structure variation has been reported by Frey and Sammons in 1985.[37] The long distance between phosphorus and sulfur may attenuate negative charge in both the anion and acid forms.

To estimate the pK_a of phosphorothiate diester, we set up a hypothetic equilibrium (eq 1) to cancel computational error in proton hydration and ion solvation. The pK_a difference of phosphorothioate and phosphate diester was estimated by eq 2.

$$\mathrm{DMPT}[(CH_3O)_2P(=O)S^-] + \mathrm{DMPOH}[(CH_3O)_2P(=O)OH] \rightarrow \mathrm{DMPT{-}OH}[(CH_3O)_2P(=S)OH] + \mathrm{DMP}[(CH_3O)_2P(=O)O^-] \quad (1)$$

dx.doi.org/10.1021/jp302494b | J. Phys. Chem. B 2012, 116, 10639−10648

Table 1. The Calculated Free Energies in Aqueous Solution (CBS-QB3 SCRF=(CPCM), in Atomic Units), Gas Phase (PBE1PBE/6-31+G(2df), in Atomic Units), and Aqueous Solvation Energy in Parentheses (PBE1PBE SCRF=(CPCM), in kcal/mol)

	DMP (a.u.)	DMPT−OH (a.u.)	DMPT (a.u.)	DMPOH (a.u.)	ΔG_{rxn} (kcal/mol)	ΔpK_a
CBS-QB3	−721.4705	−1044.4879	−1044.0567	−721.9064	3.0	−2.20
PBE1PBE	−721.6697	−1045.0687	−1044.5540	−722.1914	3.1	−2.27
(solv.)	(−61.4)	(−5.3)	(−58.4)	(−7.0)		

$$\Delta pK_a = pK_a^{DMPT} - pK_a^{DMP} = -\frac{\Delta G_{rxn}}{2.303RT} = -\frac{1}{2.303RT}(G_{ca.}^{DMP} + G_{ca.}^{DMPT-OH} - G_{ca.}^{DMPT} - G_{ca.}^{DMPOH}) \quad (2)$$

As shown in Table 1, the pK_a of the phosphorothioate diester is lower than that of the phosphate diester by about 2.2, indicating that phosphothioate linkage in DNA will be in anionic form [>P(═O)S$^-$] under physiological conditions. The PBE1PBE (CPCM, solvent=water)//PBE1PBE/6-31+G-(2df) calculation gave the corresponding ΔpK_a value of 2.3, consistent with the CBS-QB3 method.

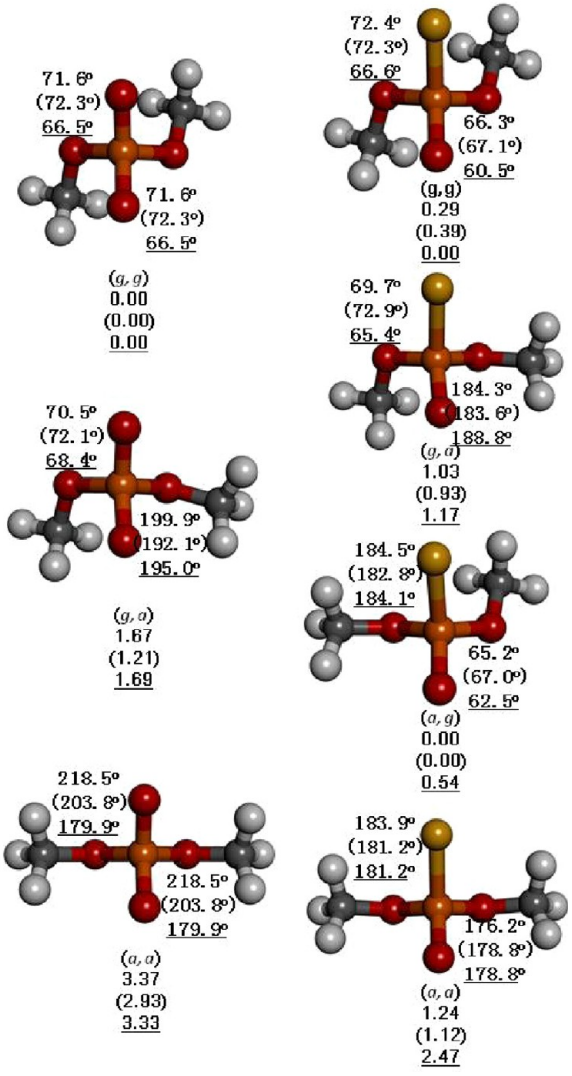


Figure 2. The minima of models **1** and **3**, optimized with the CBS-QB3 (gas phase, plain; water, underlined) and PBE1PBE/6-31+G-(2df) (gas phase, in parentheses) methods, where the relative energies are in kcal/mol and the two torsion angles are marked for C−O−P−O and O−P−O−C.

Table 2. Linear Regression of the Calculated Relative Energies against the Backbone Torsions, with a Functional Form of $f = b_0 + \Sigma b_i \cos(x - \mu_i)$, where μ_i is the Mean Value of Torsion for Each Mode[a]

	model 6		model 7		model 8	
	μ_i	b_i	μ_i	b_i	μ_i	b_i
δ	82.4	−4.79	81.8	−6.42	80.6	−4.62
	139.3	−5.02	138.4	−6.76	137.8	−4.93
	65.5	−1.17	66.6	−0.35	64.3	0.52
ε	203.1	−4.46	209.0	−4.11	207.2	−3.44
	266.5	−4.27	261.7	−3.83	262.1	−3.35
	76.0	−0.64	76.0	−3.71	60.7	−2.88
ζ	185.0	0.62	175.4	−3.30	188.9	−2.45
	287.1	−0.76	296.2	−3.90	284.6	−2.55
	70.8	−1.48	64.2	−5.01	73.1	−2.90
α	181.2	0.04	187.5	−4.39	176.9	−2.66
	286.8	−1.22	289.6	−4.60	298.6	−3.64
	93.1	−1.57	91.9	−1.88	96.6	−0.57
β	181.9	−2.45	182.1	−2.44	180.6	−1.36
	265.6	−1.75	253.6	−2.01	257.0	−0.67
	52.0	−49.38	52.7	−86.76	52.4	−20.98
γ	183.8	−49.75	184.4	−86.99	183.9	−21.26
	293.3	−49.23	292.5	−86.70	292.2	−20.81
b_0		63.75		109.12		37.21
r^2		0.9382		0.9141		0.9513

[a]For convenience, we classify all modes into three groups, [0.0−120.0], [120.0−240.0], and [240.0−360.0]. Torsion δ is bimodal, for lack of data in [240°, 360°].

The energy-minimum structures of models **1** and **3** were plotted in Figure 2. In terms of the P−O-involving torsion angles (C−O−P−O and O−P−O−C) on the backbone, the global minima are the (*gauche, gauche*) conformation in model **1** and the (*anti, gauche*) and (*gauche, gauche*) conformations in model **3** in the gas phase and aqueous solution, respectively. The energy differences between (*gauche, gauche*) and (*gauche, anti*) or (*gauche, gauche*) and (*anti, gauche*) conformations decrease by 0.5−1.5 kcal/mol after phosphorothiation. The extended conformation (*anti, anti*) is in relatively high energy in the case of either normal or phosphorothioated diesters.

Energy-Scoring Functions against Backbone Torsions. The conformation searching with the MMFF force fields resulted in hundreds of energy-minimum structures for models **6**−**8**. After the PBE1PBE/6-31+G(2df) optimization, 294, 165, and 190 nonredundant conformations were obtained for models **6**−**8**, respectively. Some representative structures of models **6**−**8** are drawn in Figure 3, marked with the P−O bond

dx.doi.org/10.1021/jp302494b | J. Phys. Chem. B 2012, 116, 10639−10648

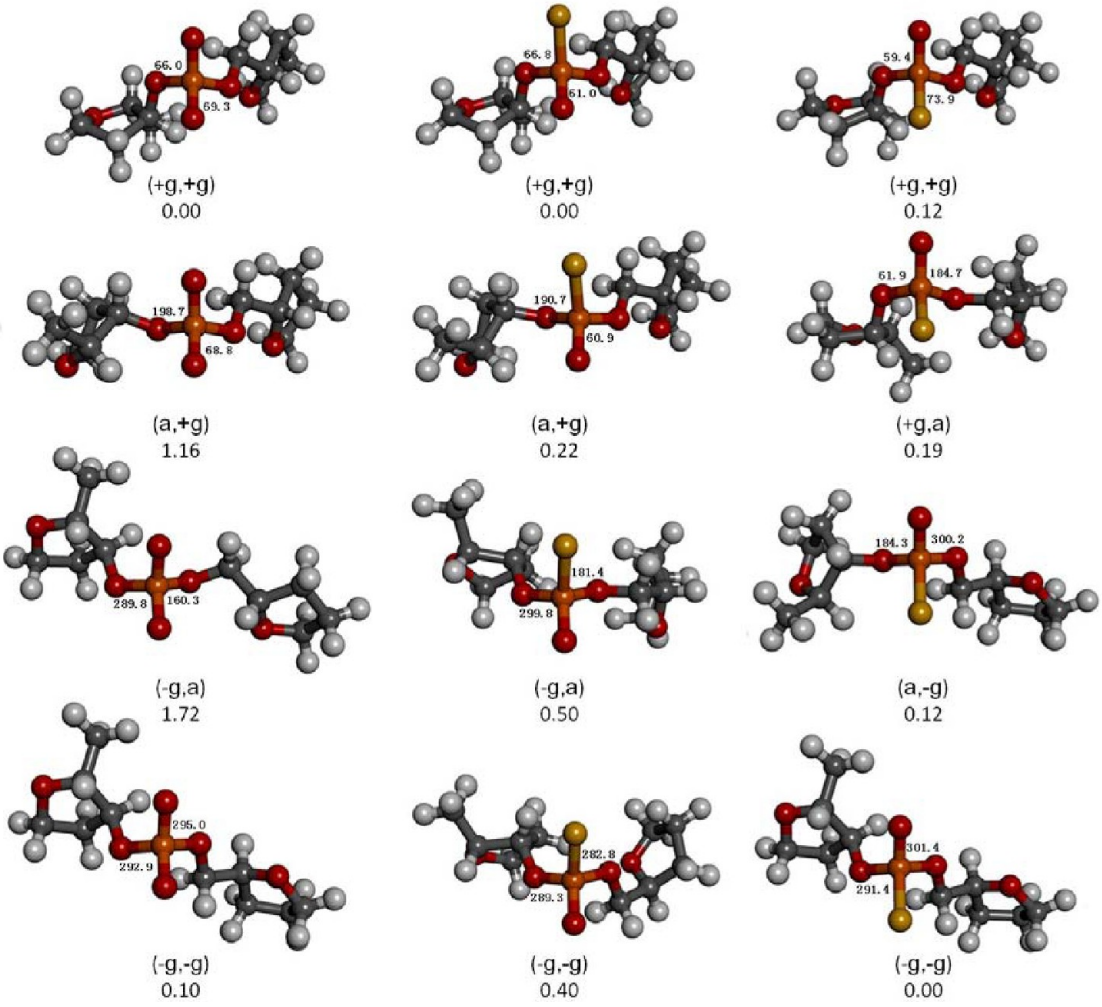


Figure 3. The minima of models **6**−**8**, optimized at the PBE1PBE/6-31+G(2df) level of theory, where the relative energies are in kcal/mol and the two torsion angles are marked for C−O−P−O and O−P−O−C.

conformation (ζ and α). The backbone torsion distribution is shown in Figure 4.

Relative energies (R.E.), with reference to the global minimum (e.g., the most stable structures as shown in Figures 2 and 3), are derived from the single-point electronic energy calculations at the PBE1PBE (CPCM, solvent=water)/6-31+G(2df) level of theory on the geometries fully optimized with the PBE1PBE/6-31+G(2df) method. As shown in Figure 3, the lowest-energy structures of models **6**−**8** have a (*gauche, gauche*) conformation, in which Rp-phosphorothioation and normal nucleotide in a (g+, g+) conformation and Sp-phosphorothioation prefers a mirror (g−, g−) conformation in the calculations, consistent with the results in models **1** and **3**. The relative energies are in a range of 8.7, 6.7, and 7.3 kcal/mol for models **6**, **7**, and **8**, respectively. Most of the R.E. variation is related to the subtle changes of backbone torsions. As shown in Figure 4, torsions α, β, γ, ζ (and probably ε) are trimodal, and δ is bimodal.

At the first glance, the distributional patterns of models **6**−**8** are highly similar. In order to quantify the thio-substituent effect, we used a traditional cosine function to fit the calculated R.E. values.[38] Table 2 lists the parameters optimized with the linear regression, and Figure 5 shows the correlation between the regression-estimated and original QM-calculated R.E. values. In the regression, the overall deviation is estimated to be about 0.45 kcal/mol and the squared correlation coefficient is 0.9382, 0.9141, and 0.9513 for models **6**−**8**, respectively. It means that the energy-scoring function is effective to predict relative energy, at least for the cases of models **6**−**8**.

The smaller coefficient b_i means a flat potential energy surface surrounding the minima in the torsion coordinate. Among them, torsion γ is found to be the "hardest", consistent with the correspondingly sharpest distribution in Figure 4. It is intriguing that the phosphorus-neighboring torsions (α and ζ) increase their "hardness" after phosphorothioation, and those in the Rp-configuration (model **7**) are even harder than those in the Sp-configuration (model **8**). However, the mean value (μ_i) of backbone torsion does not show any statistically meaningful change, whereas Rp's appears opposite to Sp's in some cases. In addition, we did not observe that correlation is improved by adding the pseudotorsional angle of the pentose (not shown here).

Thio-Substituent Effect on the Known DNA Structures. More than 5000 crystal structures have been reported in the NDB and PDB databases. The resulting scoring functions provide a fast means to estimate the relative energy of different linkages in the DNA structures. We collected the backbone torsions of the phosphate linkages of DNA released in 2011. The backbone torsion distribution of A-, B-, and Z-helical structures is plotted in Figure 6. The conformational space of

dx.doi.org/10.1021/jp302494b | J. Phys. Chem. B 2012, 116, 10639−10648

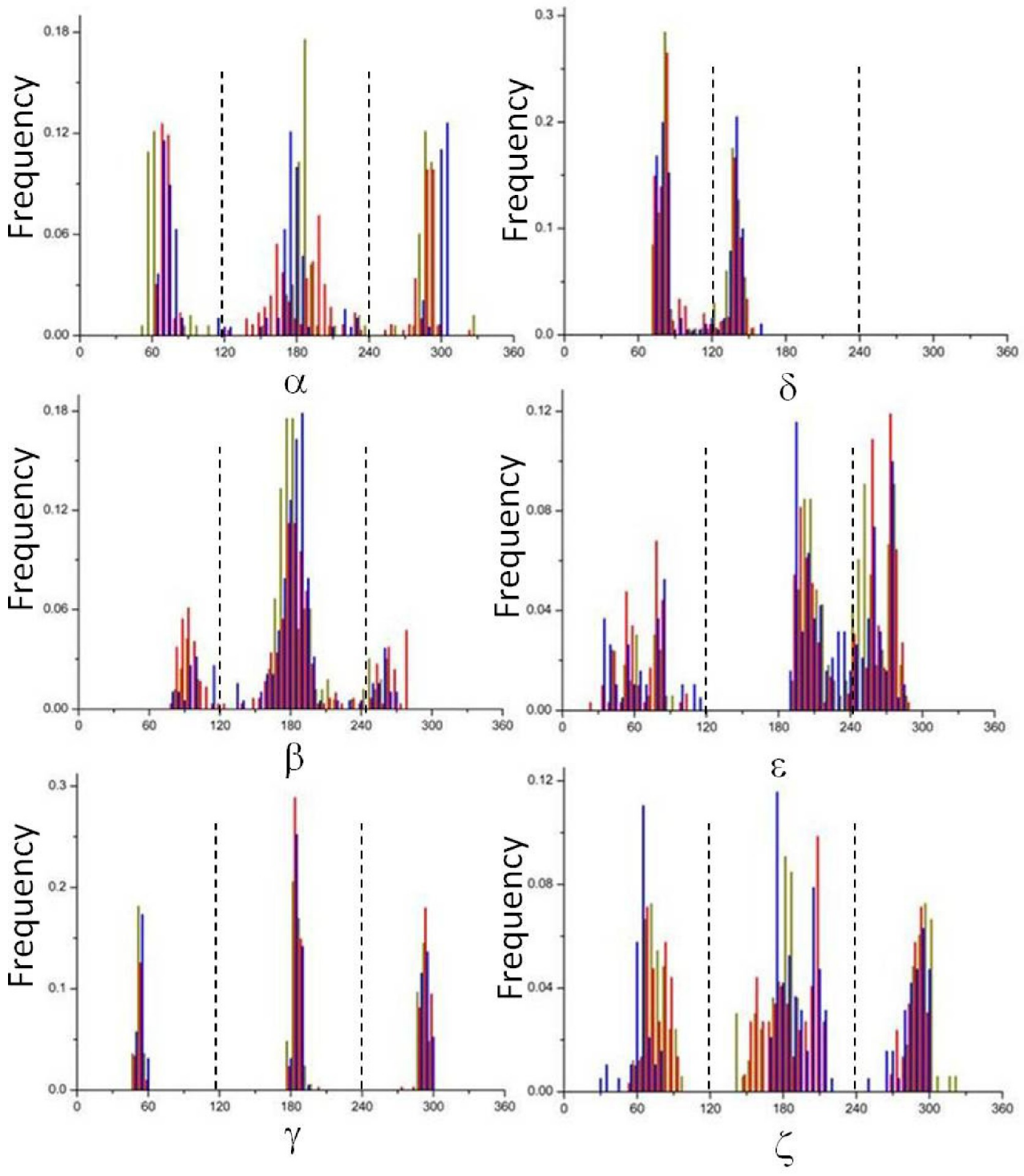


Figure 4. Backbone torsion distribution for calculated structures with the PBE1PBE(CPCM)/6-31+G(2df) method. (Models **6**, **7**, and **8** are labeled with red, green, and blue bars, respectively. The frequencies of 5° bins were normalized by the total numbers of conformations of each model.)

the phosphate linkages in models **6**–**8** (Figure 4) obviously encompasses those from the DNA crystal structures (Figure 6). For example, no observable distribution for the A, B, and Z double-helical structures for torsions β and ε is in a range of [0°, 120°], γ in [240°, 360°], or ζ in [120°, 240°]. It is due to the fact that the QM-derived conformations are constructed in the gas phase calculations, while those in the DNA crystal are under severe compact effect. In other words, the calculations of models **6**–**8** present all possible low-energy conformations in the gas phase, while the real conformations in the DNA crystal structures are restricted in a small range. This guarantees that the double-helical backbone conformations are scored effectively with the QM-based scoring function.

However, due to the conformation disproportion in the data sets, the overall distribution of the backbone torsion is highly similar to those of B-helices (75% abundance) and non-double-helices (18% abundance). In the data sets, the highest-frequent torsion of (α, β, γ, δ, ε, ζ) is (294°, 173°, 53°, 82°, 207°, 288°)

dx.doi.org/10.1021/jp302494b | J. Phys. Chem. B 2012, 116, 10639–10648

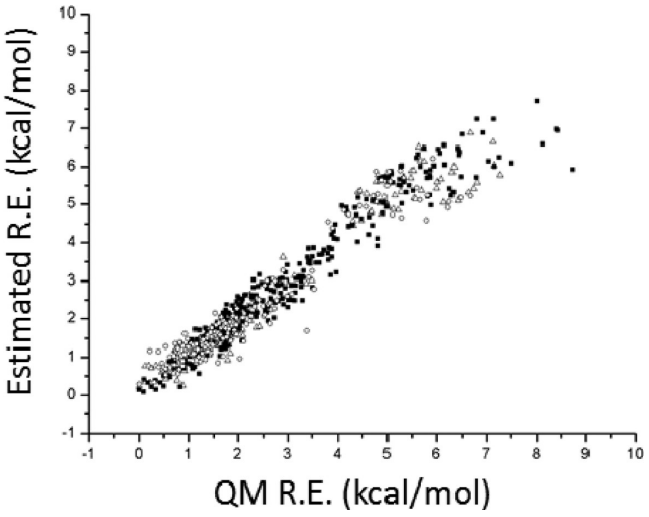


Figure 5. Comparison of relative energies calculated at the PBE1PBE(CPCM)//PBE1PBE/6-31+G(2df) and estimated with the energy scoring functions. (Models **6**–**8** are labeled with solid squares, empty circles, and empty triangles, respectively.)

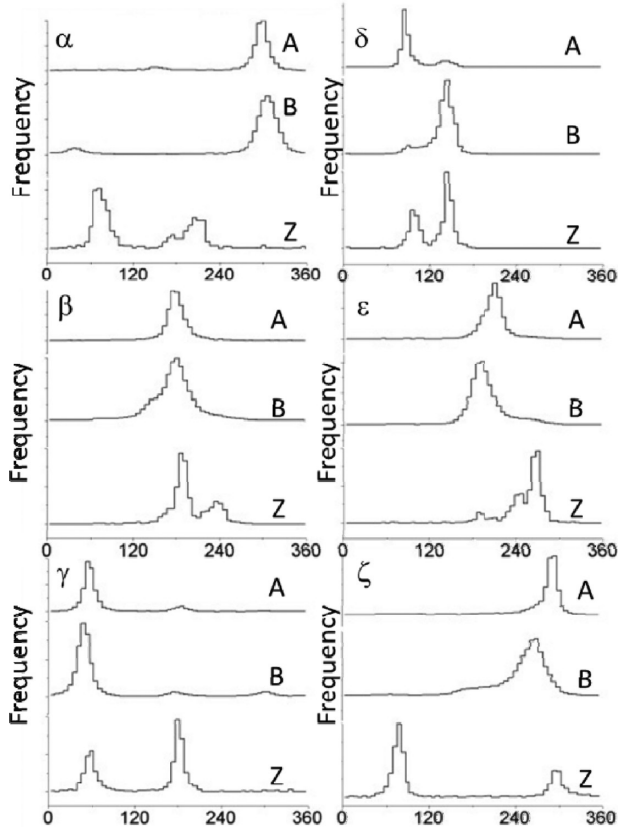


Figure 6. Backbone torsion distributions of helical conformations of the known DNA structures (from top to bottom: A-, B-, and Z-helices).

for A-form, (300°, 174°, 46°, 140°, 186°, 264°) for B-form, and (207°, 229°, 54°, 95°, 241°, 294°) (5′purine-pyrimidine3′) and (66°, 184°, 180°, 143°, 265°, 75°) (5′pyrimidine-purine3′) for Z-form. The highest-frequent torsion angle combinations in the data sets are in excellent agreement with the previously reported subclasses in the literature: AI (295°, 173°, 54°, 82°, 206°, 285°),[12c] BI (299°, 179°, 48°, 133°, 182°, 263°),[12c] ZIR (201°, 225°, 54°, 95°, 240°, 301°),[12b] and ZY (71°, 183°, 179°, 141°, 267°, 75°).[12b]

The relative energy (R.E.) of the linkages of double-helices is estimated with the QM-based energy-scoring functions, respectively; the difference (ΔR.E.) between phosphorothioate (models **7** and **8**) and phosphate (model **6**) is used to show the helical preference related to the chiral thio-substituent effect on the phosphate linkages. The mean values and deviations for the estimated R.E. and ΔR.E. were collected in Table 3, and the detailed distributions of the ΔR.E. for the three helical structures are shown in Figure 7.

The average relative energies of phosphate and phosphorothioate linkages are in a range of 2.2−6.8 kcal/mol for the structures in the data sets. Interestingly, compared to the unmodified phosphate linkage, Rp-phosphorothioate in A- and B-helical structures causes a leverage of 1.6 ± 2.8 and 2.7 ± 3.4 kcal/mol in energy, with odd events (stabilization) less than 2 and 5%, respectively. On the contrary, Sp-phosphorothioate seems to stabilize the double-helical structures slightly. Since the A- and B-helices are the most common secondary structures of DNA, the Rp-destabilization indicates that the Rp-phosphorothioation damages the right-handed double-helical structures.

Figure 8 shows the A- and B-helical structures of Rp-phosphorothioate DNA, in which the charged phosphate groups are fully exposed to the solvent at the edge of the backbone, and the P−S bonds point to the groove inside. The distances between the C′2-hydrogen and sulfur atoms in the Rp-phosphorothioated DNA are similar to those between the C′2-hydrogen and corresponding oxygen atoms in the normal DNA; however, the attractive C−H···O contact changes into the repulsive C−H···S contact in B-helix, since short-range H···S interaction is somewhat unfavorable.[43] The less flexibility of torsions α and ζ would enhance the disadvantageous interaction, leading to the steric effect on the B-helicity.

Our results about Rp-phosphorothiated B-helical destabilization are in agreement with dynamics studies of the phosphorothioated [d(CpGpCpGpApApTpTpCpGpCpG)]$_2$,[39] in which the duplex model containing the homochiral Rp-phosphorothioate strand shows a significant decrease of stability while the model containing the homochiral Sp-phosphorothioate strand has practically the same energy as the unmodified oligodeoxynucleotide duplex. However, for [d(YR)$_8$]$_2$ and d(RR)$_8$-d(YY)$_8$, it has been reported that both Rp- and Sp-phosphorothioations result in a destabilization.[40] Melting point measurement of d(GCpsTpsACG) indicates that the [*Sp, Sp*]-isomer is less stable than the [*Rp, Rp*] isomer in the case of PS-DNA−RNA hybrid duplexes, while the trend is reversed in the case of PS-DNA−DNA duplexes.[32] Thus, the helical stability is also related to the sequential composition,[41] helical type, etc.

CONCLUSIONS

Although many structural and functional features of DNA phosphorothioation have been discovered experimentally,[42] it is the first time that the correlation of helicity and stereospecificity has been systemically illustrated for phosphorothioated DNA using the relative-energy-shift method. In this work, we developed a novel energy-scoring method to valuate the stability of numerous conformations in the DNA databases according to the limited and accurate quantum mechanical calculations: (1) Using the PBE1PBE(CPCM, solvent=water)//PBE1PBE/6-31+G(2df), we can reproduce the

dx.doi.org/10.1021/jp302494b | J. Phys. Chem. B 2012, 116, 10639−10648

Table 3. The Phosphate and Phosphorothioate R.E.$_{(6)}$, R.E.$_{(7)}$, and R.E.$_{(8)}$ Estimated by the above Scoring Functions Built up with the PBE1PBE(CPCM)/6-31+G(2df) Calculations (Energy in kcal/mol)[a]

		estimated R.E.			structural preference			
type	counts	R.E.$_{(6)}$	R.E.$_{(7)}$	R.E.$_{(8)}$	ΔR.E.$_{(R)}$	odds$_{(R)}$	ΔR.E.$_{(S)}$	odds$_{(S)}$
A-helical	6333	2.7 ± 4.0	4.3 ± 6.8	1.9 ± 2.1	1.6 ± 2.8	1.5%	−0.9 ± 2.1	14.2%
B-helical	78586	4.1 ± 4.7	6.8 ± 8.0	2.7 ± 2.4	2.7 ± 3.4	4.8%	−1.4 ± 2.4	16.6%
Z-helical	939	2.9 ± 3.5	4.0 ± 5.7	2.3 ± 1.8	1.1 ± 2.6	23.7%	−0.6 ± 1.8	44.2%
nonhelical	18428	3.3 ± 4.1	5.1 ± 7.0	2.2 ± 2.2	1.8 ± 3.0	6.1%	−1.0 ± 2.1	15.8%

[a]Note: the negative ΔR.E. means stabilization, while positive means destabilization. Odds are the percentages of conformations with ΔR.E. opposite sign to the mean value.

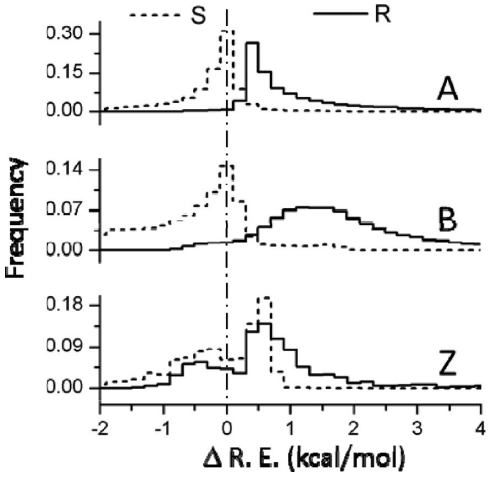


Figure 7. Distribution of the R.E. difference between phosphorothioated and normal DNA linkages. (From top to bottom: A-, B-, and Z-helices, respectively. Rp-phosphorothioate in solid line, and Sp-phosphorothioate in dashed line.)

computer-demanding CBS-QB3 results for models **1**−**5**, identifying the high acidity of phosphorothioate. (2) Hundreds of conformations were optimized by the relatively high-quality quantum mechanical method, and the reliable energy data achieved. (3) A classical cosine form of torsional potential was used to fit the QM data, resulting in high r^2 of above 0.90. (4) The energy-scoring functions were then used to estimate the phosphorothioate effect on the phosphate linkages of the known DNA structures. As a result, we found that Rp-phosphorothioate disfavors the traditionally B-form helices statistically. For confirmation, the above scoring calculation was repeated with the PBE-D(COSMO, solvent=water)/6-31+G-(2df),[44] M06(CPCM, solvent=water)/6-31+G(2df), and MP2-(CPCM, solvent=water)/6-31+G(2df) methods, consistently leading to a relative energy shift of 3.6 ± 5.2, 2.1 ± 3.2, and 1.4 ± 1.6 kcal/mol for the B-helical linkages, respectively (see Table S7, Supporting Information). It is likely that all popular DFT methods perform well in estimating the relative stability of the linkages of double-helical structures. In the methodology assessment work by Mladek et al.,[19a] for the "canonical" geometries (280°−320°, 170°−190°, 40°−55°, 116°−150°, 160°−200°, and 260°−270°), the weighted residual sum of squares (RSoS/n) between the reference CBS(T) and DFT energy with MP2/6-31+G(d) optimized geometries are 0.082 (PBE-D), 0.039 (TPSS-D), 0.145 (TPSS), 0.029 (mPW2-PLYP), 0.189 (B3LYP), and 0.034 (M06) kcal/mol,[19a] and the PBE1PBE/6-31+G(2df) method used in this work gave a RSoS/n value of 0.128 kcal/mol for the canonical structures, slightly larger than the PBE-D method. The hybrid functional does not perform well for the noncanonical conformations, resulting in a high RSoS/n value of 4.086 kcal/mol. Unexpectedly, the PBE-D calculation[44] results in the similar preference of 3.6 ± 5.2 and −2.2 ± 4.4 kcal/mol for Rp- and Sp-phosphorothioations, respectively (Tables S7−10, Supporting Information). This indicates that the B-helical destabiliza-

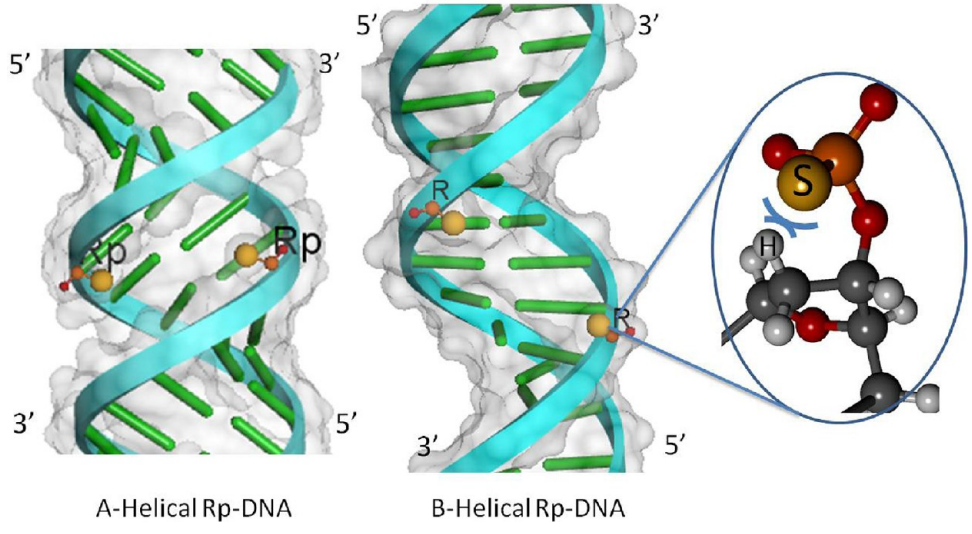


Figure 8. Diagrams of A- and B-helical structures of Rp-phosphorothioated DNA, in which the P−S bonds point to the groove inside. A short-range interaction presents between the sulfur and pro-R hydrogen atom on the sugar C2 in the virtual thio-substitution.

tion caused by Rp-phosphorothioation is of high robustness, insensitive to computational methods.

The origin for the disadvantage of phosphorothioate can be rationalized by the steric effect caused by the unfavorable S···H contact, in particular for the cases of B-helixes of Rp-phosphorothioated DNA. Moreover, the hardness of torsions α and ζ ($dE/d(\text{torsion}) = -b_i \sin(x_i - \mu_i)$) rises significantly for models **7** and **8** (see Table 2), indicating that the backbone rigidness increases in the phosphorothioated DNA. The structural understandings of phosphorothioated DNA may help in finding out functions of the endogenous DNA chemical modification.

■ ASSOCIATED CONTENT

Supporting Information

The (α, ζ) contour PES maps, local energy minima of models **1**−**3**, geometrical parameters of calculated structures of models **6**−**8**, and detailed calculated results. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: yileizhao@sjtu.edu.cn. Phone/Fax: +86-21-34207190.

Author Contributions

‖These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

⊥These authors are students of High School Affiliated to Shanghai Jiao Tong University.

■ ACKNOWLEDGMENTS

The authors wish to thank the National High-tech R&D Program of China "863" (2012AA020403) and the National Basic Research Program of China "973" (2012CB72100), Shanghai Municipal Council of Science and Technology (10PJ1405200), the Specialized Research Fund for the Doctoral Program of Higher Education (Z1025507), Shanghai Municipal Education Commission (Oriental Professorial Scholarship 0900000171), and National Science Foundation of China (30821005, 31121064) for financial support. We also thank the High-School Students Innovation Program in High School Affiliated to Shanghai Jiao Tong University.

■ REFERENCES

(1) Zhou, X.; Deng, Z.; Firmin, J. L.; Hopwood, D. A.; Kieser, T. *Nucleic Acids Res.* **1988**, *16*, 4341−4352.

(2) Zhou, X.; He, X.; Liang, J.; Li, A.; Xu, T.; Kieser, T.; Helmann, J. D.; Deng, Z. *Mol. Microbiol.* **2005**, *57*, 1428−1438.

(3) Wang, L.; Chen, S.; Xu, T.; Taghizadeh, K.; Wishnok, J. S.; Zhou, X.; You, D.; Deng, Z.; Dedon, P. C. *Nat. Chem. Biol.* **2007**, *3*, 709−710.

(4) Wang, L.; Chen, S.; Vergin, K. L.; Giovannoni, S. J.; Chan, S. W.; DeMott, M. S.; Taghizadeh, K.; Cordero, O. X.; Cutler, M.; Timberlake, S.; Alm, E. J.; Polz, M. F.; Pinhassi, J.; Deng, Z.; Dedon, P. C. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2963 −2968.

(5) Liang, J.; Wang, Z.; He, X.; Li, J.; Zhou, X.; Deng, Z. *Nucleic Acids Res.* **2007**, *35*, 2944−2954.

(6) Guga, P. *Curr. Top. Med. Chem.* **2007**, *7*, 695−713.

(7) Liu, G.; Ou, H.-Y.; Wang, T.; Li, L.; Tan, H.; Zhou, X.; Rajakumar, K.; Deng, Z.; He, X. *PLoS Genet.* **2010**, *6*, e1001253.

(8) Dickey, J. S.; Van Etten, J. L.; Osheroff, N. *Biochemistry* **2005**, *44*, 15378−15386.

(9) Eckstein, F.; Gish, G. *Trends Biochem. Sci.* **1989**, *14*, 97−100.

(10) Vargason, J. M.; Henderson, K.; Ho, P. S. A crystallographic map of the transition from B-DNA to A-DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7265 −7270.

(11) Vargason, J. M.; Eichman, B. F.; Ho, P. S. *Nat. Struct. Biol.* **2000**, *7*, 758−761.

(12) (a) Saenger, W.; Hunter, W. N.; Kennard, O. *Nature* **1986**, *324*, 385−388. (b) Schneider, B.; Neidle, S.; Berman, H. M. *Biopolymers* **1997**, *42*, 113−124. (c) Svozil, D.; Kalina, J.; Omelka, M.; Schneider, B. *Nucleic Acids Res.* **2008**, *36*, 3690−3706. (d) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751−759.

(13) *Computational Studies of RNA and DNA*; Šponer, J., Lankaš, F., Eds.; Springer: The Netherlands, 2006.

(14) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(15) MacKerell, A. D.; Wiorkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

(16) Langley, D. R. *J. Biomol. Struct. Dyn.* **1998**, *16*, 487−509.

(17) Gelbin, A.; Schneider, B.; Clowney, L.; Hsieh, S.-H.; Olson, W. K.; Berman, H. M. *J. Am. Chem. Soc.* **1996**, *118*, 519−529.

(18) Poltev, V. I.; Anisimov, V. M.; Danilov, V. I.; van Mourik, T.; Deriabina, A.; González, E.; Padua, M.; Garcia, D.; Rivas, F.; Polteva, N. *Int. J. Quantum Chem.* **2010**, *110*, 2548−2559.

(19) (a) Mladek, A.; Sponer, J. E.; Jurecka, P.; Banas, P.; Otyepka, M.; Svozil, D.; Sponer, J. *J. Chem. Theory Comput.* **2010**, *6*, 3817−3835. (b) Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E., III; Jurecka, P. *J. Chem. Theory Comput.* **2011**, *7*, 2886−2902.

(20) Westerhoff, L. M.; Merz, K. M., Jr. *J. Mol. Graphics Modell.* **2006**, *24*, 440−455.

(21) Shen, X.; Gu, B.; Che, S. A.; Zhang, F. S. *J. Chem. Phys.* **2011**, *135*, 034509.

(22) Svozil, D.; Šponer, J. E.; Marchan, I.; Pérez, A.; Cheatham, T. E.; Forti, F.; Luque, F. J.; Orozco, M.; Šponer, J. *J. Phys. Chem. B* **2008**, *112*, 8188−8197.

(23) (a) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493−2499. (b) Sponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *267*, 263−270.

(24) (a) Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *Chem.—Eur. J.* **1999**, *5*, 3581−3594. (b) Kawahara, S.; Wada, T.; Kawauchi, S.; Uchimaru, T.; Sekine, M. *J. Phys. Chem. A* **1999**, *103*, 8516−8523.

(25) (a) Danilov, V. I.; Anisimov, V. M.; Kurita, C.; Hovorun, D. *Chem. Phys. Lett.* **2005**, *412*, 285−293. (b) Fiethen, A.; Jansen, G.; Hesselmann, A.; Schutz, M. *J. Am. Chem. Soc.* **2008**, *130*, 1802−1803.

(26) (a) Szatyłowicz, H.; Sadlej-Sosnowska, N. *J. Chem. Inf. Model.* **2010**, *50*, 2151−2161. (b) Pacureanu, L.; Simon, Z. *Int. J. Quantum Chem.* **2010**, *110*, 1295−1305. (c) Close, D. M. *J. Phys. Chem. A* **2010**, *114*, 1860−1867. (d) Kozak, C. R.; Kistler, K. A.; Lu, Z.; Matsika, S. *J. Phys. Chem. B* **2010**, *114*, 1674−1683 and references therein..

(27) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287−303.

(28) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.

(29) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(30) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D., Jr. *J. Chem. Theory Comput.* **2012**, *8*, 348−362.

(31) Mukherjee, S.; Bhattacharyya, D. *Biopolymers* **2004**, *73*, 269−282.

(32) Kanaori, K.; Tamura, Y.; Wada, T.; Nishi, M.; Kanehara, H.; Morii, T.; Tajima, K.; Makino, K. *Biochemistry* **1999**, *38*, 16058−16066.

(33) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−519.

(34) *Discovery Studio*, version 3.0; Accelrys Software, Inc.: San Diego, CA, 2010.

(35) Frisch, M. J.; et al. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(36) Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 5010−5018.
(37) Frey, P.; Sammons, R. *Science* **1985**, *228*, 541 −545.
(38) Radom, L.; Hehre, W. J.; Pople, J. A. *J. Am. Chem. Soc.* **1972**, *94*, 2371−2381.
(39) Jaroszewski, J. W.; Syi, J. L.; Maizel, J.; Cohen, J. S. *Anti-Cancer Drug Des.* **1992**, *7*, 253−262.
(40) Hartmann, B.; Bertrand, H.-O.; Fermandjian, S. *Nucleic Acids Res.* **1999**, *27*, 3342 −3347.
(41) Boczkowska, M.; Guga, P.; Stec, W. J. *Biochemistry* **2002**, *41*, 12483−12487.
(42) Guga, P.; Koziołkiewicz, M. *Chem. Biodiversity* **2011**, *8*, 1642−1681.
(43) Domagala, M.; Grabowski, S. J. *J. Phys. Chem. A* **2005**, *109*, 5683−5688.
(44) The PBE-D calculation was carried out with ORCA software: (a) *ORCA*, version 2.9; Max-Planck-Institute for Bioinorganic Chemistry: Muelheim, Germany. The Gimme's dispersion correction and COSMO solvent model: (b) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463−1476. (c) Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F. *J. Phys. Chem. A* **2006**, *110*, 2235−2245.