THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

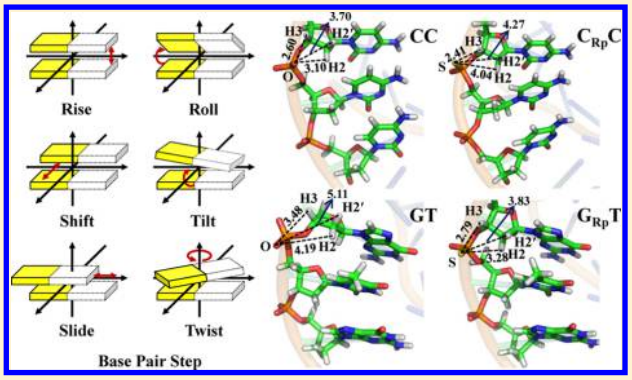
Theoretical Study on the Relationship between Rp-Phosphorothioation and Base-Step in S-DNA: Based on Energetic and Structural Analysis

Limeng Chen, Xiao-Lei Wang, Ting Shi, Tingting Wu, Zixin Deng, and Yi-Lei Zhao*

State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

Supporting Information

ABSTRACT: Phosphorothioation (PT), previously used in synthetic antisense drugs to arrest the transcription or translation process, is also a novel physiological modification in bacteria DNAs. In the previous study, we reported that Rp-phosphorothioation (Rp-PT) destabilizes B-type helix significantly, using a quantum-mechanics-based energy scoring function developed with a dinucleotide model (Zhang et al. *J. Phys. Chem. B*, **2012**, *116*, 10639−10648). A consequent question surfaces in the field of the phosphorothioated DNA (S-DNA) research: does the endogenous chemical modification interact with the base sequence in the bacterial genomes, e.g., in terms of the most common structure of the B-type helix? In this work, we carried out further energetic analysis on the backbone relative energies calculated with the scoring function according to 16 groups of base-step classifications. Moreover, we conducted molecular dynamics simulations of the B-helical structure with the different base-pair steps, to investigate the detailed structural changes upon the O-/S-substitution. As a result, the Rp-PT modification definitively enhances the stiffness of the backbone and differentiates backbone stability as an interaction with base-steps. Furthermore, certain exceptional sequences such as GT and CC were highlighted in the structural analysis of the sulfur local contacts and relative orientation of double strands, indicating that Rp-PT can cross-talk with particular base-steps. The special effects between the phosphorothioation and base-step may be related to the conservative consensus observed highly frequently in bacterial genomes.



INTRODUCTION

DNA phosphorothioation (PT) modification refers to the replacement of one nonbridge oxygen atom of phosphodiester linkage by sulfur atom. This modification, found in *Streptomyces lividans*,[1] confers the centered phosphorus atom with Rp or Sp chirality. The chemical modification changes physicochemical aspects of DNA, such as its chemical stability,[2] conformational stability,[3] and biological activity.[4]

The initial discovery of the physiological Rp-PT originated from a DNA degradation phenomenon in 1988. After that, the degradation was found to be related to DNA S-modification and [5] then confirmed experimentally as the Rp chirality with respect to the phosphorus center in the DNA backbone by Wang et al.[1] Nine types of the Rp-PT-modified base-steps were reported in the biological studies, i.e., GpsG, GpsA, GpsT, CpsC, CpsA, ApsA, ApsC, TpsA, and TpsC.[6] With the deep sequencing technology, Rp-PT could be quickly determined in genomic scope, leading to the additional discovery of GpsAAC/GpsTTC double-strand motifs in *Escherichia coli* B7A and the CpsCA single-strand motif in *Vibrio cyclitrophicus* FF75.[7]

It is well-recognized that sequence context exerts a significant influence on the structure and function of DNA molecules.[8] Many intensive studies on sequence effects of normal B-DNA have been reported. For example, by analyzing geometries of phosphodiester backbone of crystal structures in the Nucleic acid Data Base (NDB),[9] Svozil et al. assessed the sequence preferences of A- and B-helix with statistical goodness-of-fit tests.[10] Besides, the Ascona B-DNA consortium showed nearest-neighbor effects on base pairs.[11−13] Using the sequence-dependent molecular dynamics simulation of normal duplex DNA, Okonogi et al. suggested that the dominant characteristic can be best interpreted in the aspect of purine- and pyrimidine-type steps.[14] Moreover, Saad et al. investigated the origins of distortions in the base pair step adjacent to platinum anticancer drug−DNA adducts based on NMR data.[15] Thus, the base-step conformation map has already been constructed systematically and comprehensively, using base pair step parameters on normal DNA, but not in the case of S-DNA.[16]

Although crystallography, and, to a lesser extent, NMR spectroscopy, provide a growing structural data of phosphorothioated nucleotides oligomers, it is still necessary to make

Received: November 12, 2014
Revised: December 17, 2014
Published: December 18, 2014

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474−481

reliable explanation for the Rp-PT-modified DNA. One of the most famous and relevant investigations about the PT-modified DNA were conducted by Hartmann et al. in 1999.[17] With the molecular dynamic simulations of 10 double-stranded dinucleotide steps, the authors indicated that (1) the destabilization of the PT-modified oligomers was mainly due to an alteration in the electrostatic term, (2) the Sp-PT-modified oligomers possess a larger chirality effect than Rp-PT, (3) the base preceding the PT-modified site has a major effect on the stability, (4) and the Rp-PT is more stable than Sp-PT. Other literature showed that PT modification significantly changes the conformation of RNA molecules,[18] but only has small effects on the DNA double helix.[19] However, the quantum-mechanics-based (QM-based) scoring shows upside-down relative energy stability between Rp and Sp configurations, and the high-level QM calculation indicated that the steric effect is dominant in double-helical structures, rather than the electrostatic interaction in the force field calculation.[3] Therefore, sequence preference of S-DNA needs to be revisited so as to solve the confliction between the force field study and quantum calculations.

To understand the structural characteristics of Rp-PT, in the previous work we proposed QM-based scoring with respect to static structures of DNA in the structural databases using the six backbone dihedrals before and after phosphodiester site.[3] In this work we focused on the sequence preference in the Rp-PT-modified B-double helix of DNA. First, the QM-based scoring energies were reanalyzed according to the classification of $4 \times 4 = 16$ types of base-steps (Figure 1). By a proper data-fitting

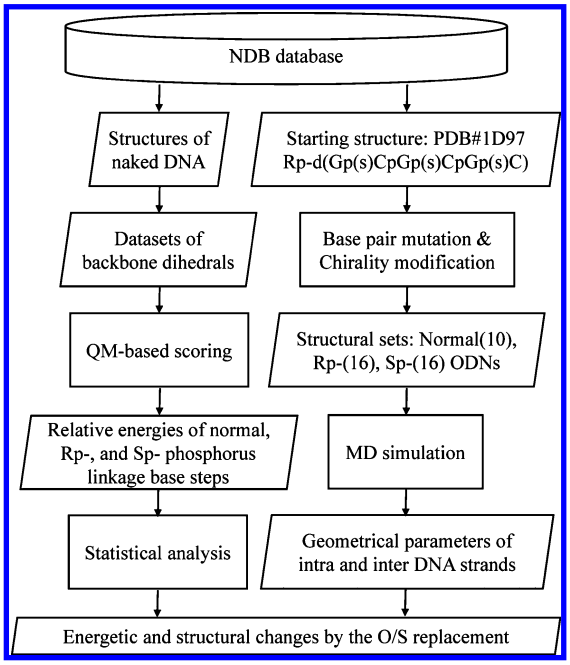


Figure 1. Flowchart of the theoretical investigation.

method, the imbalance inside the DNA structural data set was carefully corrected. Then MD simulations with the AMBER force field were performed on 42 double-stranded B-DNA oligodeoxynucleotides (ODNs), each containing six base pair steps, generated by mutation of a typical crystal structure (PDB code: 1D97).[20] With the in-depth conformation analysis on the local structure and double-helical orientation, we found that Rp-PT changed the local conformation of B-helix via increasing the backbone instability and was partially dependent on the base-steps. Since the B-type double-helical structure is the most basic form of DNA molecules, the current results indicated that the PT modification can couple with the neighboring bases, in particular for the immediately preceding base in certain short consensus. By the mutual corroboration of energetic and structural factors, we show how the Rp-PT modification affects the sequence preference in the form of B-helical structure. Hence, the present theoretical investigations make up the previous MD simulation with the force field by mixing results from quantum calculations and are connected with the recent experimental observations of the S-DNA sequence in bacteria as well.

MATERIALS AND METHODS

Six dihedral angles (backbone torsions) were utilized to investigate the correlation between the relative energy of phosphate/phosphorothioate linkages and base pair steps. These backbone dihedrals were δ (C5′-C4′-C3′-O3′), ε (C4′-C3′-O3′-P), and ζ (C3′-O3′-P-O5′) of 5′-terminal before the modified site and α (O3′-P-O5′-C5′), β (P-O5′-C5′-C4′), and γ (O5′-C5′-C4′-C3′) of 3′-terminal after the modified site. In the structural analysis on the MD trajectory, we studied the inter-base-pair parameters of the three models (phosphate, Rp-PT, and Sp-RT), that is, three translationals (i.e., shift, slide, and rise) and three rotationals (i.e., tilts, roll, and twist) parameters[21] (Figure 2).

395 naked (no-protein) double-stranded B-DNA crystal structures (see Supporting Information Tables S1 and S2), including 6702 base-steps, were collected in this work from the NDB database. Then the relative backbone stability of the collected base-steps were evaluated using the energy scoring function that we constructed according to quantum mechanical calculations of dinucleotide models in the previous work.[3] The QM-based scoring function has a classical cosine form of $x = b_0 + \Sigma b_i \cos(\omega_i - \mu_i)$, where x is the relative energy, ω_i denotes the six backbone torsions, μ_i are the mean values of each torsion, and b_i are the parameters related to the potential energy surfaces for the above torsions. Data sizes of different base-steps changed dramatically, ranging from 124 (CT) to 1693 (CG) as shown in Figure 3a. To cancel the data imbalance, a fitting method was used based on an equation derived from the Maxwell−Boltzmann (M-B) formula,[22] as follows:

$$f(x) = 2\alpha/\sqrt{\pi}(1/\beta)^{3/2}\sqrt{x}\,\exp(-x/\beta) \quad (1)$$

where x and β denote E and kT in the original M-B formula, and α is a normalized parameter depending on the size of the bin used in the fitting. The typical bin finally used in this work is the smallest value that gives the correlation coefficient r^2 above 0.90. So the bins of phosphate, Rp-PT, and Sp-PT models were set as 2.8, 3.8, and 1.8 kcal/mol, respectively (see Supporting Information Figures S1−S3). Parameter β is an intrinsic property of the distribution as a standard to evaluate the relative stability for the different base-steps of the three models, in good agreement with median values of the original data sets in calculation.

Molecular dynamics simulations were performed on the normal, Rp-PT, and Sp-PT double-stranded B-DNA oligomers. The structure (PDB code: 1D97) in NDB database was used as the initial structure, which is a naked B-double helix of DNA with the real Rp-PT modifications (see Supporting Information

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474−481

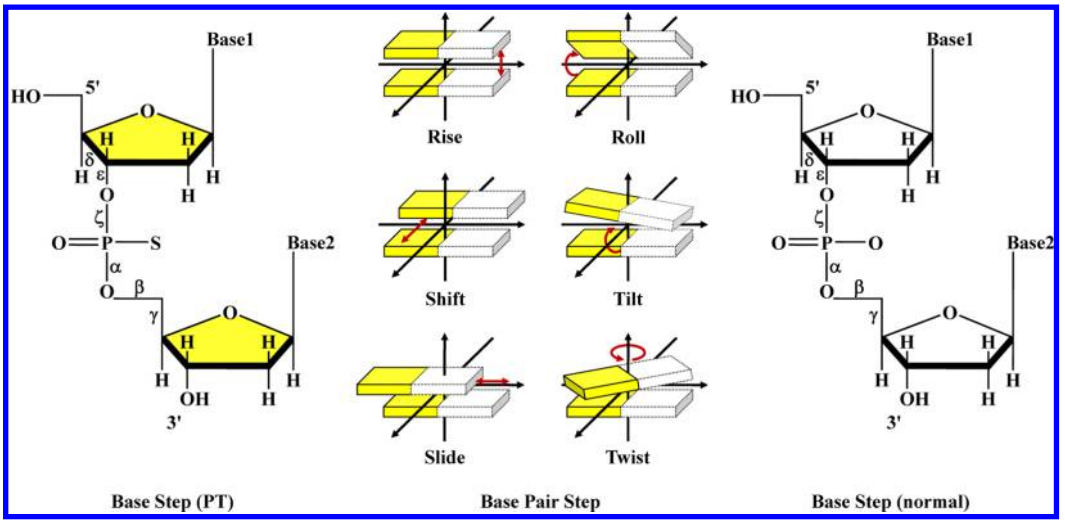


Figure 2. Six backbone dihedral angles and base pair step parameters discussed in this work.

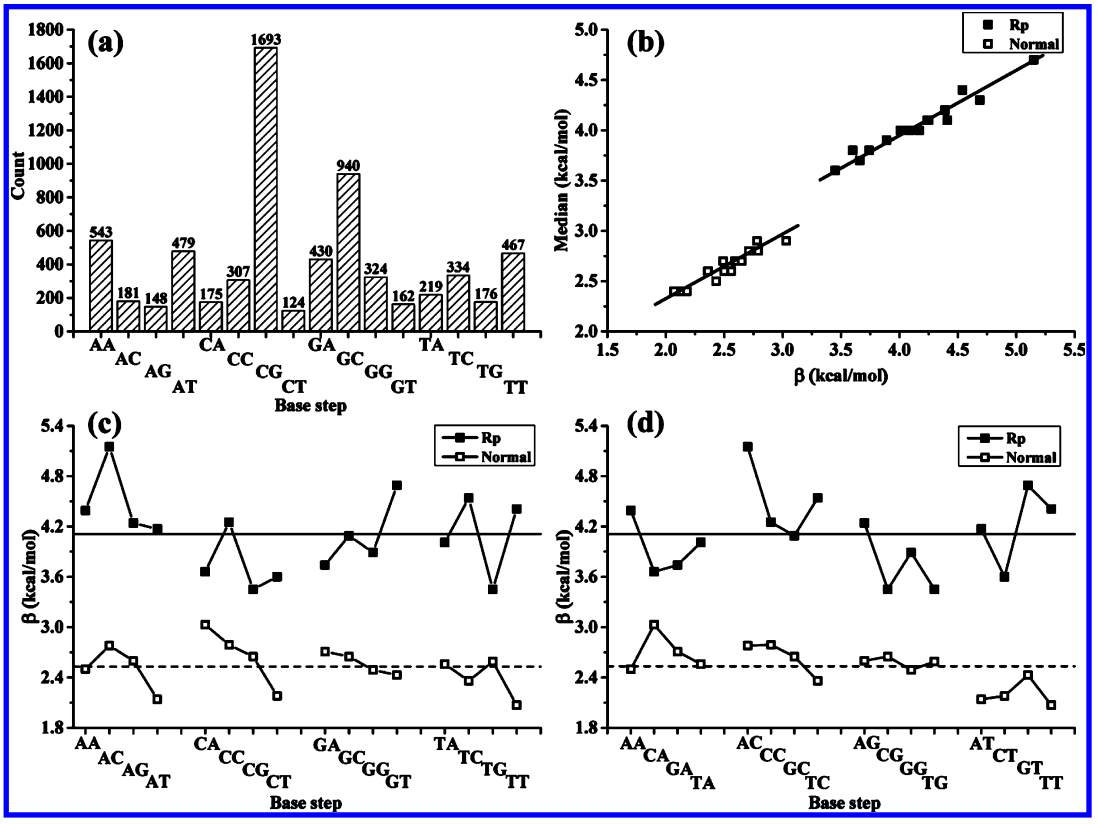


Figure 3. (a) Data sets size of the 16 base pair steps in the statistical analysis. (b) Median value and parameter β of the calculated backbone energies. The scatter diagram of the calculated β values of normal and Rp-PT-modified DNA grouped as 5′- (c) and 3′- terminal (d). The solid squares stand for the β value of Rp-PT-modified base pair step and the hollow squares for those of normal DNA. The horizontal line stands for the average value of all of the 16 β for Rp-phosphorothioated (solid) and normal (hash) DNA. The statistical difference between the two groups is significant with 95% confidence interval (p = 2.919e^{-11}).

Table S3). To investigate the different influences caused by PT modification and different base-steps, we rebuilt all possible types of B-DNA double-stranded oligomers structures in a "XYXYXY" pattern, each containing the target base-step (XY) in the middle with two caps at the two terminals; therefore, 10 normal, 16 Rp-PT, and 16 Sp-PT models were generated. Discovery Studio 3.5 software[23] was used to conduct all of the necessary substitutions of sulfur atom and base pairs.

All of the molecular dynamics simulations were carried out with AMBER12 software.[24] The FF99SB force field was applied to the normal nucleotides, and the force field for the phosphorothioated models was built based on parameters developed by Mukherjee and Bhattacharyya.[25] Sodium ions were added into the systems as counterions to neutralize the charged oligonucleotides, and then the systems were solvated with TIP3P water in a 10 Å octahedron box. The systems were first energy-minimized by 1000 steps of steepest descent,

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474−481

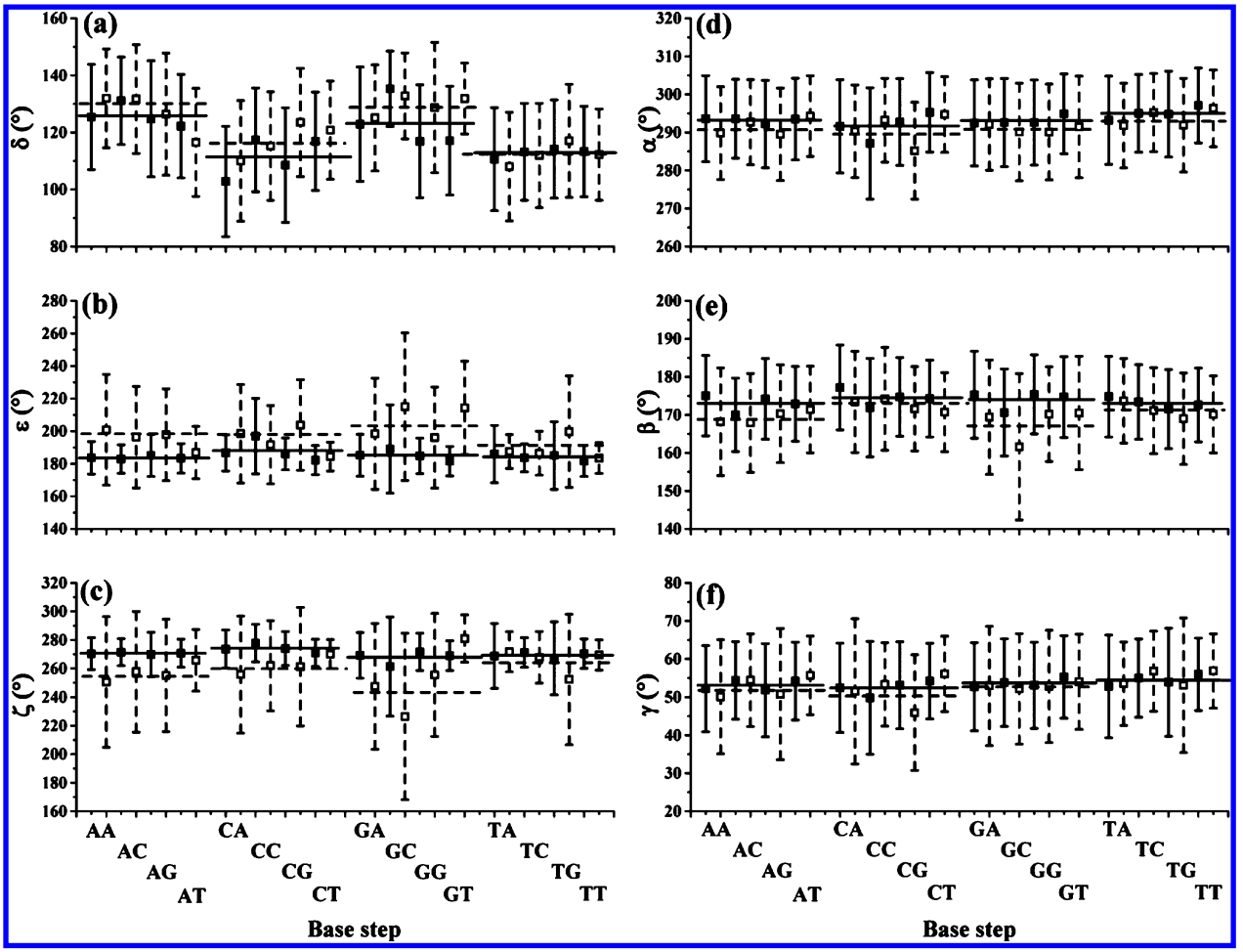


Figure 4. Average values (squares) and standard deviations (vertical bars) of the six backbone dihedrals. The solid core and hollow squares stand for dihedrals of Rp-phosphorothioate and normal base-steps calculated by MD. The horizontal lines are the average values of the corresponding dihedral angles, with the solid and dashed lines for the Rp-PT and normal ones.

followed by 1000 steps of conjugate gradient, with a 500 kcal/(mol·Å^2) restraint force on the DNA molecules. Then the systems were further energy-minimized by 5000 steps of steepest descent followed by 15000 steps of conjugate without restraints. After that, the system was heated to 300 K in 50 ps with 10 kcal/(mol·Å^2) restraints on the DNA. The systems were then equilibrated for 100 ps without restraints. Such an equilibration procedure was repeated 7 times, which allowed the DNA to be equilibrated adequately. Finally, molecular dynamics simulations were carried out for 20 ns, with conformational snapshots saved every 0.1 ps. The mass weighted root-mean-square deviations (RMSDs) from the first structures of the trajectory during the preceding half 10 ns were calculated to check the stability of the DNA structures, and all of the 42 systems converged quickly to a reasonably stable structure (see Supporting Information Figures S10–S12). The remaining 10 ns trajectories were used for further analysis, which contains totally 4.2 million snapshots of structures for the 42 models. The six dihedral angles (δ, ε, ζ, α, β, γ,) in each snapshot were calculated using the *ptraj* program[26] in AmberTools13, and the six inter-base-pair parameters (rise, slide, shift, twist, roll, and tilt) were calculated with the curves+ and canal programs.[27] The frequency distributions of the six dihedrals were calculated and analyzed (see Supporting Information Figures S13–S19).

RESULTS AND DISCUSSION

Direct QM-Based Scoring of the B-DNA Backbone in NDB. Since the fitting parameter β can reduce the perturbance of the data-size imbalance, it is employed to reflect backbone stability and compare the subtle influence of different base pair steps. According to the direct QM-based energy scoring, the Rp-PTs are in higher energy than normal DNA, suggesting the obvious backbone instabilities of the modified B-double helix (Figure 3b, Supporting Information Tables S4 and S5 and Figures S4–S9). The fitting results are in agreement with our previous theoretical work.[3] Then parameter β is used to exhibit the characteristics of 16 base pair steps (Figure 3c and 3d). The data are divided into eight groups to exhibit the 5′- and 3′-terminal base effects. The resulting sequence effects are visible but relatively minor in the figures. First, backbone energies of normal and Rp-PT-modified DNA fluctuate in a range of 2.1 and 4.1 kcal/mol, respectively. This result is consistent with the previous conclusion that the Rp-PT destabilizes the B-type helical structure as well.[3] Second, combined with the information on base-step classification in Figure 3c,d, the backbone energies are interrelated with the preceding 5′-terminal and succeeding 3′-terminal bases. For instance, the backbone stability of Rp-PT-modified DNA seems to fluctuate more intensively, compared to that of the normal DNA. In particular, the backbone energies of 5′-A- and 3′-C-terminal

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474–481

Table 1. Average Values of Inter-base-pair Parameters with Their Standard Deviations: Shift, Slide, Rise, Tilt, Roll, and Twist

	rise (Å)	shift (Å)	slide (Å)	roll (deg)	tilt (deg)	twist (deg)
{A_{Rp}A,TT}	3.35 ± 0.31	0.05 ± 0.48	−0.59 ± 0.56	2.2 ± 5.0	−0.9 ± 4.0	32.6 ± 4.7
{AA,TT}	3.43 ± 0.32	−0.38 ± 0.66	−0.14 ± 0.56	0.5 ± 6.9	−3.8 ± 4.3	35.3 ± 6.0
{AA,T_{Rp}T }	3.34 ± 0.33	0.25 ± 0.62	−0.36 ± 0.61	3.3 ± 6.4	1.7 ± 4.2	33.4 ± 5.0
{A_{Rp}C,GT}	3.36 ± 0.30	0.28 ± 0.67	−0.84 ± 0.51	0.7 ± 4.9	−1.1 ± 3.9	30.6 ± 4.2
{AC,GT}	3.36 ± 0.31	−0.43 ± 0.69	−0.31 ± 0.49	−3.0 ± 6.8	0.6 ± 4.0	31.6 ± 5.7
{AC,G_{Rp}T}	3.40 ± 0.31	0.05 ± 0.73	−0.78 ± 0.50	−0.2 ± 5.3	0.9 ± 3.9	30.8 ± 4.8
{A_{Rp}G,CT}	3.56 ± 0.36	−0.78 ± 0.57	0.02 ± 0.76	1.1 ± 5.4	1.0 ± 4.8	35.6 ± 6.7
{AG,CT}	3.38 ± 0.37	−0.29 ± 0.73	−0.61 ± 0.76	4.8 ± 5.8	−0.8 ± 4.3	30.4 ± 5.9
{AG,C_{Rp}T}	3.42 ± 0.39	0.03 ± 0.63	−0.93 ± 0.62	2.9 ± 5.1	0.2 ± 4.0	30.5 ± 5.0
{A_{Rp}T,AT}	3.28 ± 0.28	0.01 ± 0.66	−0.88 ± 0.43	−0.1 ± 4.6	0.1 ± 3.9	29.9 ± 3.6
{AT,AT}	3.19 ± 0.29	0.03 ± 0.69	−0.81 ± 0.43	1.1 ± 5.1	0.0 ± 3.9	29.2 ± 4.1
{C_{Rp}A,TG}	3.32 ± 0.38	−0.28 ± 0.68	−0.46 ± 0.60	10.1 ± 5.9	−0.4 ± 4.7	33.3 ± 6.2
{CA,TG}	3.19 ± 0.33	−0.07 ± 0.72	0.06 ± 0.70	12.0 ± 6.5	0.0 ± 4.8	30.5 ± 6.8
{CA,T_{Rp}G }	3.47 ± 0.35	0.46 ± 0.77	−0.43 ± 0.61	12.6 ± 7.0	2.0 ± 5.1	33.7 ± 5.4
{C_{Rp}C,GG}	3.05 ± 0.49	0.41 ± 0.55	0.17 ± 1.46	5.4 ± 5.6	0.6 ± 4.4	27.4 ± 7.7
{CC,GG}	3.57 ± 0.31	0.40 ± 0.56	−1.13 ± 0.64	2.8 ± 5.6	2.1 ± 4.2	31.4 ± 4.5
{ CC,G_{Rp}G}	3.60 ± 0.35	−0.03 ± 0.65	−1.18 ± 0.82	5.2 ± 5.3	0.3 ± 4.3	31.6 ± 5.1
{C_{Rp}G,CG}	3.31 ± 0.39	0.26 ± 0.81	0.06 ± 0.59	10.3 ± 6.2	1.9 ± 5.2	33.7 ± 7.6
{CG,CG}	3.13 ± 0.35	−0.21 ± 0.71	0.04 ± 0.54	10.0 ± 5.7	−0.7 ± 4.8	30.3 ± 7.1
{G_{Rp}A,TC}	3.35 ± 0.32	0.10 ± 0.64	−0.71 ± 0.68	2.1 ± 5.7	−0.2 ± 4.5	32.9 ± 5.1
{GA,TC}	3.41 ± 0.32	0.13 ± 0.68	−0.55 ± 0.67	1.4 ± 6.0	−0.4 ± 4.5	35.1 ± 5.9
{ GA,T_{Rp}C}	3.37 ± 0.31	0.08 ± 0.70	−0.72 ± 0.72	1.8 ± 5.7	0.8 ± 4.4	33.6 ± 5.5
{G_{Rp}C,GC}	3.47 ± 0.28	0.44 ± 0.75	−0.41 ± 0.62	−2.5 ± 5.6	1.6 ± 4.3	33.7 ± 5.1
{GC,GC}	3.39 ± 0.34	−0.06 ± 0.79	−0.32 ± 0.55	−0.6 ± 6.1	0.3 ± 4.4	31.7 ± 7.1
{T_{Rp}A,TA}	3.40 ± 0.37	0.23 ± 0.77	−0.46 ± 0.75	9.3 ± 8.3	0.7 ± 5.3	34.1 ± 6.2
{TA,TA}	3.48 ± 0.38	0.07 ± 0.72	−0.68 ± 0.64	11.6 ± 7.5	−0.7 ± 5.1	32.8 ± 5.7

base-steps are likely higher than others. These results are consistent with Hartmann's results about the preceding base effect in 1999.[17] With comparison of normal and the Rp-PT-modified DNA, it can be concluded that Rp-PT modification differentiated the backbone stability in the form of B-helical structure.

Distribution Changes of Backbone Dihedral Angles upon Phosphorothioation. Based on the well-optimized 10 ns molecular dynamics simulation trajectory consisting of 10^5 conformation snapshots for each type of base pair step sequence, the six backbone dihedral angles (δ, ε, ζ, α, β, γ) located near the Rp-PT linkage and corresponding normal linkage, as well as the mean values and standard errors, were calculated to demonstrate the sequence effect change upon the Rp-PT modification, as shown in Figure 4.

It is clearly shown that, for all types of base-steps, the standard errors of ε and ζ shrink dramatically after the Rp-PT modification, and those of α, β, and γ exhibit a smaller but considerable decrease, except for δ of GT; ε, ζ, and γ of TA; and α and γ of CC. The drop down of standard error indicates that the corresponding backbone torsion changes less during the simulation, which implies a more intensive stiffness of the S-DNA backbone if comparing with normal DNA. We have demonstrated in the previous study that based on a QM-scoring function, the potential energy surfaces of δ, ζ, α, and γ become sharper after the Rp-PT modification in a dinucleotide model, which is consistent with the present observation in the molecular dynamics simulations. Sharper potential energy surface surrounding the Rp-PT modification site will make the local deformation more difficult, which may affect its tolerance toward restriction nuclease cleavage and other modifications.[4]

For each group of base-steps starting with a same base at the 5′-terminal, the mean value of each torsion angle was also calculated. On the whole, the mean values of ε and ζ (5′-terminal) change more significantly after Rp-PT modification compared to those of α and γ (3′-terminal), while δ and β also have moderate changes. For the mean value of δ, ε, ζ, and β, base-steps starting with 5′-G exhibit the largest shifts in the four groups, and in contrast base-steps starting with T change the least in all six dihedral angles upon the Rp-PT modification. The mean value of dihedral angle stands for the equilibrium position in the dynamics. Thus, the Rp-PT modification makes the potential energy surface of the DNA backbone not only sharper but also shifted upon the different base-steps. This means that the PT can couple with the base-step in terms of backbone conformation. Among the six backbone torsion angles, δ, ε, and ζ are relatively larger contributors, indicating that the conformation change caused by the Rp-PT modification mainly cross-talks with the preceding base. Our previous work also showed that the sulfur atom sterically conflicts with a hydrogen atom on the preceding ribose ring (see Supporting Information Figure S20).

Specifically, grouping the 16 base sequences according to the 5′-terminal nucleotide, base-steps starting with G (5′-G group) exhibit the largest mean value shifts of δ, ε, ζ, and β after the Rp-PT modification, whereas the 5′-T group has minor mean value shifts of all six dihedral angles, indicating a significant 5′-terminal effect. The different influences of 5′-bases on the Rp-PT caused conformation change reflect different responses to the sulfur atom conflicting contacts. Compared to thymine, guanine has stronger base pair hydrogen bonding and bulkier molecular structure, which may make it less flexible upon Rp-PT modification. The structural difference could further affect the epigenetic function or binding affinity to enzymes and

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474−481

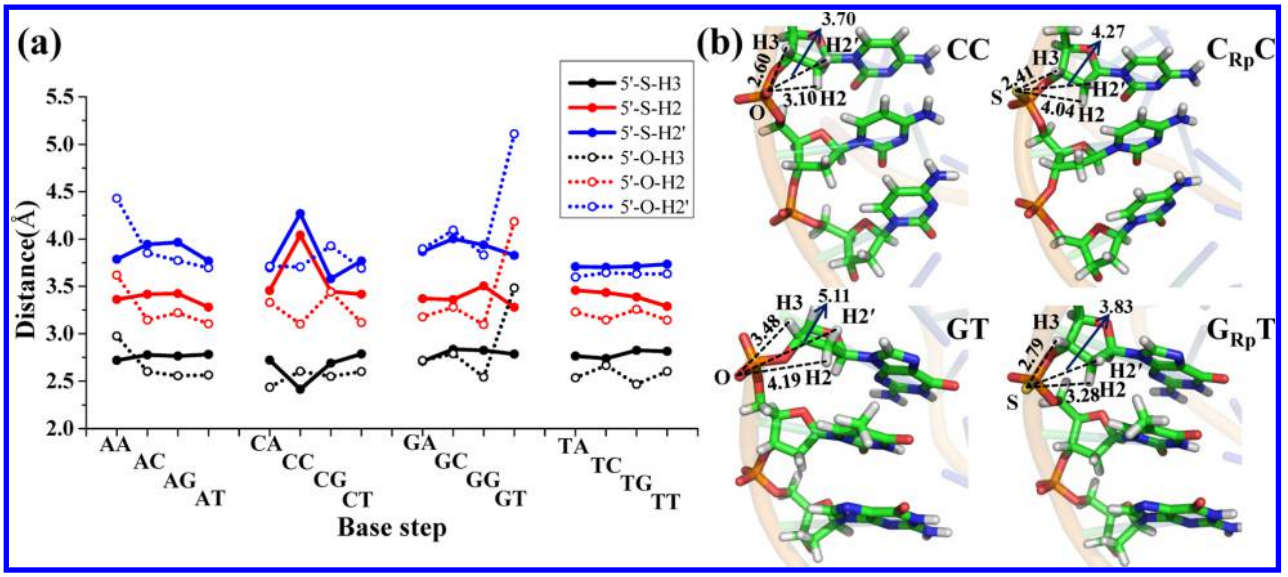


Figure 5. Distances between sulfur and hydrogen atoms (H2, H2′, and H3) of 5′-terminal ribose with structures of normal and Rp-PT-modified CC and GT of the B-helix.

molecules. Since the Rp-PT modification has been found highly sequence-dependent in microorganisms, especially favoring G-preceding sequences,[1] thus the above illustrated 5′-terminal effect may be a reason worth considering.

Orientation of Double Strands upon Phosphorothioation. The base pair step parameters of 16 Rp-PT models and 10 normal models, as well as corresponding standard deviations, are listed in Table 1. Among the three translational parameters, rise values change least upon Rp-PT, and its deviation is also smaller than shift and slide, indicating that Rp-PT has little effect along the helical axis. The axial rise in the case of {$C_{Rp}C$, GG} exhibits exceptional change compared to the others, with the rise value changed −15% compared to {CC, GG}. Horizontal shifts in the cases of {$A_{Rp}A$, TT} (−113%), {$A_{Rp}C$, GT} (−165%), {$A_{Rp}G$, CT} (169%), {$C_{Rp}G$, CG} (−224%), {$G_{Rp}C$, GC} (−833%), {CC, $G_{Rp}G$} (−108%), {AC, $G_{Rp}T$} (−112%), {CA, $T_{Rp}G$} (−757%), and {AA, $T_{Rp}T$} (−166%) alter obviously after Rp-PT. Meanwhile, horizontal slides in the cases of {$A_{Rp}A$, TT} (321%), {$A_{Rp}C$, GT} (171%), {$A_{Rp}G$, CT} (−103%), {$C_{Rp}A$, TG} (−867%), {$C_{Rp}C$, GG} (−115%), {AC, $G_{Rp}T$} (152%), {CA, $T_{Rp}G$} (−817%), and {$A_{Rp}A$, TT} (157%), also show large changes. Especially, the slide of CC changes more than 1 Å. In contrast, AT has the smallest translational change.

As for the three rotational parameters, the standard deviation of tilt angle is smaller than roll and twist for all base-pair steps, possibly due to the structural constraint by helical backbone. Roll values in the cases of {$A_{Rp}A$, TT} (340%), {$A_{Rp}C$, GT} (−123%), {$A_{Rp}G$, CT} (−77%), {$A_{Rp}T$, AT} (−109%), {$C_{Rp}C$, GG} (93%), {$G_{Rp}C$, GC} (317%), {CC, $G_{Rp}G$} (86%), {AC, $G_{Rp}T$} (−93%), and {AA, $T_{Rp}T$} (560%), tilt values of {$A_{Rp}C$, GT} (−283%), {$A_{Rp}G$, CT} (−225%), {AG, $C_{Rp}T$} (−125%), {$C_{Rp}G$, CG} (−371%), {GA, $T_{Rp}C$} (−300%), and {AA, $T_{Rp}T$} (−145%), {$G_{Rp}C$, GC} (433%), {$T_{Rp}A$, TA} (−200%), {$A_{Rp}T$, AT} (−109.09%), and twist values of {$A_{Rp}G$, CT} (17%), {CA, $T_{Rp}G$} (10%), {$C_{Rp}C$, GG} (−13%), and {$C_{Rp}G$, CG} (11%) have significant changes upon Rp-PT modification. Considering the three rotational parameters together, {$A_{Rp}G$, CT}, {$C_{Rp}G$, CG}, {$A_{Rp}A$, TT}, and {AA, $T_{Rp}T$} have larger changes than others, whereas {$A_{Rp}T$, AT} changes least. Combining the translational and rotational parameters, {$A_{Rp}G$, CT}, {$C_{Rp}C$, GG}, {$C_{Rp}G$, CG}, and {$C_{Rp}A$, TG} have significant inter-base-pair conformational change during the molecular dynamics simulations, and {$A_{Rp}T$, AT} hardly changes. It is worth noting that 5′-G base presents significant 5′-terminal effects on intrabackbone torsions, but no distinctive inter-base-pair conformation changes upon the Rp-PT modification. Groove parameters, intra-base-pair parameters, and axis parameters are listed in Supporting Information Tables S6−S8 as well.

Local Spatial Contacts of Rp-Phosphorothioation. In this section, we attempted to understand the local spatial contacts of phosphodiester linkage in B-DNA upon the PT modification. The sulfur atom of Rp-PT orientated to the major groove of B-DNA, whereas it slightly perturbed the groove location. Here we use the average structures obtained by the MD simulations of the last 10 ns trajectory to describe the local structural changes. Distances between the sulfur atom and its closest hydrogen atoms (H2, H2′, and H3) of 5′-terminal ribose were measured with Discovery studio 3.5 software. Figure 5 shows that the distances are basically in the order of H3 < H2 < H2′. Different from other base pair steps, AA and GT have much shorter S---H distances after Rp-PT modifications; CC has the inverse variant tendency in S---H2 and S---H2′, but not S---H3. Collaborated with backbone dihedral parameters, dihedral ζ in CC changes greatly as large as 12.6%, compared with δ (7.4%) and ε (11.2%).

A methodology, developed by McGaughey et al.[28] and embedded in Discovery studio 3.5, was employed to show the $\pi-\pi$ stacking of base pair steps. For most of the base pair steps, the $\pi-\pi$ stacking was found to be reduced after Rp-PT modification. What's more, the decrease only happens on the Rp-PT-modified strand. {$G_{Rp}A$, TC}, {$G_{Rp}G$, CC}, {$G_{Rp}T$, AC}, and {$C_{Rp}C$, GG}, which are reported as the top frequent types, were discussed here as examples. Compared to their normal counterparts, the $\pi-\pi$ stacking of {$G_{Rp}A$, TC} and {$G_{Rp}T$, AC} reduces a lot, since the sulfur atom is bigger than the oxygen atom and P−S has a longer bond, which increases the steric hindrance and electrostatic repulsion between the phosphorothioate moiety and the base moiety as well (see Supporting Information Figure S21). To avoid the extra steric

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474−481

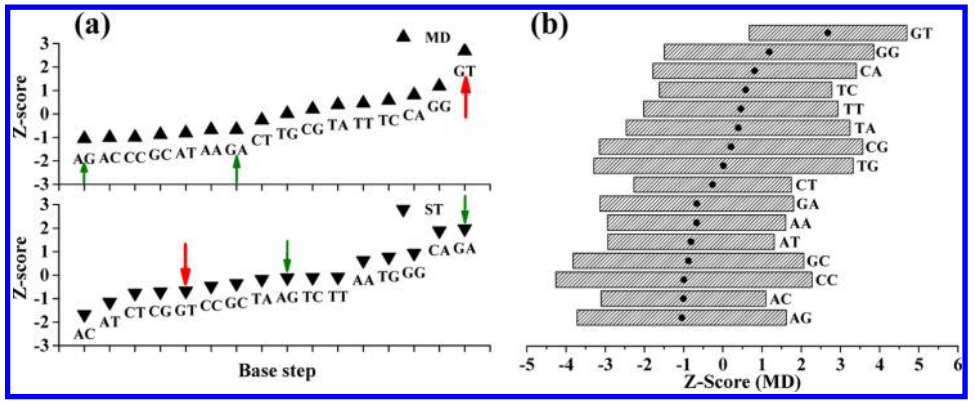


Figure 6. Base-step preference ranks of Rp-phosphorothioation. (a) Triangles stand for base-steps preference rank calculated with six torsion angles gotten from MD simulations (MD) through the last 10 ns trajectory; inverted triangles stand for the preference rank obtained from statistical analysis (ST), which are calculated with dihedrals gotten from the NDB database and the Rp-phosphorothioation fitting function obtained in QM-based scoring with static information. The ranks of base-steps denoted with arrows are affected significantly after PT modification. (b) The horizontal stripe stands for the standard deviation; the solid core circle stands for the average value.

hindrance, neighbor bases tend to separate from each other, resulting in the decrease in distance between the cancroids of each base pair of π rings and weakening $\pi-\pi$ stacking, while base pair steps on the normal strand are too far away to be affected by the sulfur atom, with their $\pi-\pi$ stacking interaction remains intact.

To compare the base pair step preference rank obtained from MD simulations and statistical analysis, median values of the six backbone torsions were calculated, followed by the relative energy calculation with a QM-based scoring function on the backbone stability (Figure 6a). The inconsistency between the statistical and MD-resulting preferences is mainly caused by the dynamic changes in structure after real phosphorothioation, instead of static structures from nonmodified DNA used previously. The maximum-correlation minimum-motion analysis indicates that the GT base-step contributes significantly for the rank correlation coefficient between the two preferences. The motion of the GT position in the sequence preference can change the correlation coefficient from 0.37 to 0.61. The next two motions of the AG and GA base-steps would have Spearman's rank correlation coefficient reach 0.80 with respect to the MD-resulting preference (see Supporting Information Table S9). Moreover, it is likely that the Rp-PT modification in the GT base-step is highly disfavored in the simulation, while those in the AG and GA base-steps are partially relaxed, in terms of the back-and-forth directions of the three motions. It is noteworthy that the local spatial contacts of the sulfur atom are sort of crowded in the GT base-step. For the two bases in a base-step motif, the effect of the preceding base is dominant in determining the structure fluctuation of the Rp-PT. The preceding base is closer to sulfur, leading to a bigger steric hindrance than that of the succeeding base. In particular, Figure 6b shows that the energy deviation in the sequence preference is overwhelming in the molecular dynamic simulation, compared to the subtle difference among the base-steps (except for the GT step).

CONCLUSION

The statistical results with the QM-based scoring function uncovered the potential preference of S-DNA, in terms of backbone stability. With comparison of normal and Rp-PT-modified DNA, the Rp-PT modification obviously destabilized the B-type helical structure and differentiated the backbone stability of B-helix. Furthermore, the structural characteristics were discussed with the molecular dynamic simulations of all types of B-double-stranded oligomers DNA structures in a "XYXYXY" pattern. Besides, special steric effect in the local site and base pair interaction were observed in the study. The reinvestigation is complementary to those of Hartmann et al., which were concluded from the force field calculation. The structural effect may be related to the sequence preference observed in the bacterial genome, such as GT and CC steps, which may change the original conformations of B-type helix and playing important roles in biological processes.

ASSOCIATED CONTENT

Supporting Information

Data preparation details for statistical analysis and MD simulations, statistical analysis results, RMSD of MD trajectories, frequency distribution of backbone dihedrals, comparison of backbone dihedrals between Rp-PT-modified DNA from MD simulations and normal DNA from NDB database, groove parameters, intrabase pair parameters and base pair axis parameters of Rp-PT-modified and normal DNA from MD simulations, comparison between Sp-PT-modified and normal DNA, and evolutionary process from ST preference rank to MD preference rank. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

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

Author Contributions

Y.-L.Z. and Z.D. conceived and designed the investigation. L.C., X.-L.W., and T.W. performed the energetic analysis and MD simulations. T.S., X.-L.W., L.C., and Y.-L.Z. wrote up the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the National High-tech R&D Program of China "863" (Grant 2012AA020403) and the National Basic Research Program of China "973" (Grants 2012CB721005 and 2013CB966802), the National Science Foundation of China (Grants 21377085, 21303101, 31121064, 21102090, 91230105,

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474−481

and J1210047), the China Postdoctoral Science Foundation (Grants 2014T70413 and 2014M561463), the Program for New Century Excellent Talents in University (Grant NCET-12-0354), the Shanghai Municipal Council of Science and Technology (Grants 10PJ1405200 and 13YZ032), and the SJTU-HPC computing facility award for financial support and computational hours. L.C. acknowledges Drs. Chaochun Wei and Jing Li for their helpful discussions.

REFERENCES

(1) Wang, L.; Chen, S.; Xu, T.; Taghizadeh, K.; Wishnok, J. S.; Zhou, X.; You, D.; Deng, Z.; Dedon, P. C. Phosphorothioation of DNA in Bacteria by dnd Genes. *Nat. Chem. Biol.* **2007**, *3*, 709−710.

(2) Boczkowska, M.; Guga, P.; Stec, W. J. Stereodefined Phosphorothioate Analogues of DNA: Relative Thermodynamic Stability of the Model PS-DNA/DNA and PS-DNA/RNA Complexes. *Biochemistry* **2002**, *41*, 12483−12487.

(3) Zhang, Y.-C.; Liang, J.; Lian, P.; Han, Y.; Chen, Y.; Bai, L.; Wang, Z.; Liang, J.; Deng, Z.; Zhao, Y.-L. Theoretical Study on Steric Effects of DNA Phosphorothioation: B-Helical Destabilization in Rp-Phosphorothioated DNA. *J. Phys. Chem. B* **2012**, *116*, 10639−10648.

(4) Roche, C. J.; Tse-Dinh, Y. C. Effect of Phosphorothioate Substitutions on DNA Cleavage by Escherichia coli DNA Topo-isomerase I. *Int. J. Biol. Macromol.* **2001**, *29*, 175−180.

(5) Zhou, X. F.; He, X. Y.; Liang, J. D.; Li, A. Y.; Xu, T. G.; Kieser, T.; Helmann, J. D.; Deng, Z. X. A Novel DNA Modification by Sulphur. *Mol. Microbiol.* **2005**, *57*, 1428−1438.

(6) 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. DNA Phosphorothioation is Widespread and Quantized in Bacterial Genomes. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 2963−2968.

(7) Cao, B.; Chen, C.; DeMott, M. S.; Cheng, Q.; Clark, T. A.; Xiong, X.; Zheng, X.; Butty, V.; Levine, S. S.; Yuan, G.; et al. Genomic Mapping of Phosphorothioates Reveals Partial Modification of Short Consensus Sequences. *Nat. Commun.* **2014**, *5*, 3951−3951.

(8) Kilpinen, H.; Waszak, S. M.; Gschwind, A. R.; Raghav, S. K.; Witwicki, R. M.; Orioli, A.; Migliavacca, E.; Wiederkehr, M.; Gutierrez-Arcelus, M.; Panousis, N. I.; et al. Coordinated Effects of Sequence Variation on DNA Binding, Chromatin Structure, and Transcription. *Science* **2013**, *342*, 744−747.

(9) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S. H.; Srinivasan, A. R.; Schneider, B. The Nucleic Acid Database. A Comprehensive Relational Database of Three-Dimensional Structures of Nucleic Acids. *Biophys. J.* **1992**, *63*, 751−759.

(10) Svozil, D.; Kalina, J.; Omelka, M.; Schneider, B. DNA Conformations and Their Sequence Preferences. *Nucleic Acids Res.* **2008**, *36*, 3690−3706.

(11) Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T., III; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.; et al. A Systematic Molecular Dynamics Study of Nearest-Neighbor Effects on Base Pair and Base Pair Step Conformations and Fluctuations in B-DNA. *Nucleic Acids Res.* **2010**, *38*, 299−313.

(12) Dixit, S. B.; Beveridge, D. L.; Case, D. A.; Cheatham, T. E.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Sklenar, H.; et al. Molecular Dynamics Simulations of the 136 Unique Tetranucleotide Sequences of DNA Oligonucleotides. II: Sequence Context Effects on the Dynamical Structures of the 10 Unique Dinucleotide Steps. *Biophys. J.* **2005**, *89*, 3721−3740.

(13) Beveridge, D. L.; Barreiro, G.; Suzie Byun, K.; Case, D. A.; Cheatham, T. E.; Dixit, S. B.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; et al. Molecular Dynamics Simulations of the 136 Unique Tetranucleotide Sequences of DNA Oligonucleotides. I. Research Design and Results on d(CpG) Steps. *Biophys. J.* **2004**, *87*, 3799−3813.

(14) Okonogi, T. M.; Alley, S. C.; Reese, A. W.; Hopkins, P. B.; Robinson, B. H. Sequence-Dependent Dynamics of Duplex DNA: The Applicability of a Dinucleotide Model. *Biophys. J.* . **2002**, *83*, 3446−3459.

(15) Saad, J. S.; Natile, G.; Marzilli, L. G. Origins of the Distortions in the Base Pair Step Adjacent to Platinum Anticancer Drug-DNA Adducts. Fundamental NMR Solution Studies Utilizing Right-Handed Cross-Link Models Having 5′- and 3′-Flanking Residues. *J. Am. Chem. Soc.* **2009**, *131*, 12314−12324.

(16) Packer, M. J.; Dauncey, M. P.; Hunter, C. A. Sequence-Dependent DNA Structure: Dinucleotide Conformational Maps. *J. Mol. Biol.* **2000**, *295*, 71−83.

(17) Hartmann, B.; Bertrand, H. O.; Fermandjian, S. Sequence Effects on Energetic and Structural Properties of Phosphorothioate DNA: A Molecular Modelling Study. *Nucleic Acids Res.* **1999**, *27*, 3342−3347.

(18) Smith, J. S.; Nikonowicz, E. P. Phosphorothioate Substitution Can Substantially Alter RNA Conformation. *Biochemistry* **2000**, *39*, 5642−5652.

(19) Kanaori, K.; Tamura, Y.; Wada, T.; Nishi, M.; Kanehara, H.; Morii, T.; Tajima, K.; Makino, K. Structure and Stability of the Consecutive Stereoregulated Chiral Phosphorothioate DNA Duplex. *Biochemistry* **1999**, *38*, 16058−16066.

(20) Cruse, W. B.; Salisbury, S. A.; Brown, T.; Cosstick, R.; Eckstein, F.; Kennard, O. Chiral Phosphorothioate Analogues of B-DNA. the Crystal Structure of Rp-d Gp(S)CpGp(S)CpGp(S)C. *J. Mol. Biol.* **1986**, *192*, 891−905.

(21) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X. J.; Neidle, S.; Shakked, Z.; et al. Standard Reference Frame for the Description of Nucleic Acid Base-Pair Geometry. *J. Mol. Biol.* **2001**, *313*, 229−237.

(22) Skrdla, P. J. Crystallizations, Solid-State Phase Transformations and Dissolution Behavior Explained by Dispersive Kinetic Models Based on a Maxwell-Boltzmann Distribution of Activation Energies: Theory, Applications, and Practical Limitations. *J. Phys. Chem. A* **2009**, *113*, 9329−9336.

(23) *Discovery Studio*, version 3.5; Accelrys Software: San Diego, 2010.

(24) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 198−210.

(25) Mukherjee, S.; Bhattacharyya, D. Effect of Phosphorothioate Chirality on the Grooves of DNA Double Helices: A Molecular Dynamics Study. *Biopolymers* **2004**, *73*, 269−282.

(26) Roe, D. R.; Cheatham, T. E., III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.

(27) Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. Conformational Analysis of Nucleic Acids Revisited: Curves. *Nucleic Acids Res.* **2009**, *37*, 5917−5929.

(28) McGaughey, G. B.; Gagne, M.; Rappe, A. K. π-Stacking Interactions. Alive and Well in Proteins. *J. Biol. Chem.* **1998**, *273*, 15458−15463.

DOI: 10.1021/jp511359e
J. Phys. Chem. B 2015, 119, 474−481