Published online 24 April 2021

Nucleic Acids Research, 2021, Vol. 49, No. 9 ***4919–4933***
doi: 10.1093/nar/gkab285

A unified computational view of DNA duplex, triplex, quadruplex and their donor–acceptor interactions

Gyuri Park[1], Byunghwa Kang[1], Soyeon V. Park[1], Donghwa Lee[1,2,3,*] and Seung Soo Oh[1,3,4,*]

[1]Department of Materials Science and Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, South Korea, [2]Division of Advanced Materials Science, Pohang University of Science and Technology (POSTECH), Pohang 37673, South Korea, [3]Institute for Convergence Research and Education in Advanced Technology (I-CREATE), Yonsei University, Incheon 21983, South Korea and [4]School of Interdisciplinary Bioscience and Bioengineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, South Korea

Received October 15, 2020; Revised April 07, 2021; Editorial Decision April 07, 2021; Accepted April 14, 2021

ABSTRACT

DNA can assume various structures as a result of interactions at atomic and molecular levels (e.g., hydrogen bonds, π–π stacking interactions, and electrostatic potentials), so understanding of the consequences of these interactions could guide development of ways to produce elaborate programmable DNA for applications in bio- and nanotechnology. We conducted advanced *ab initio* calculations to investigate nucleobase model structures by componentizing their donor-acceptor interactions. By unifying computational conditions, we compared the independent interactions of DNA duplexes, triplexes, and quadruplexes, which led us to evaluate a stability trend among Watson–Crick and Hoogsteen base pairing, stacking, and even ion binding. For a realistic solution-like environment, the influence of water molecules was carefully considered, and the potassium-ion preference of G-quadruplex was first analyzed at an *ab initio* level by considering both base-base and ion-water interactions. We devised new structure factors including hydrogen bond length, glycosidic vector angle, and twist angle, which were highly effective for comparison between computationally-predicted and experimentally-determined structures; we clarified the function of phosphate backbone during nucleobase ordering. The simulated tendency of net interaction energies agreed well with that of real world, and this agreement validates the potential of *ab initio* study to guide programming of complicated DNA constructs.

INTRODUCTION

Deoxyribonucleic acid (DNA) is a unique material as composed of nitrogenous bases (adenine (A), thymine (T), guanine (G) or cytosine (C)), sugar rings, and phosphate groups. DNA is programmable, so it can be rationally designed into molecular structures ranging from simple Watson–Crick base-pairing primers to DNA origami-based complex 3D constructs (1,2). Even spatial and temporal control of DNA nanostructures is achievable; sophisticated DNA molecular machines can perform a series of nanomechanical motions in a controllable manner, so this ability provides unprecedented applications in bio- and nanotechnology (3–6). Such active use of DNA requires fundamental understanding of DNA folding and its stabilization. In particular, DNA hybridization programming extensively exploits understanding of various donor-acceptor interactions of DNA, including hydrogen bonds, π–π stacking interactions, and electrostatic potentials (7).

Hydrogen bonds and π–π stacking are among the most important intra- and inter-molecular interactions in DNA (8–11). In general, nucleobases create specific hydrogen bonds between a purine (A or G) and a pyrimidine (C and T), which yield planar Watson–Crick base pairs (A–T; G–C). Between adjacent base pairs, π–π stacking interactions result in sequential stacking of base pairs. Therefore, hydrogen bonds and π–π stacking drive the formation of a duplex as a basic structure of DNA. However, the nucleobases can be influenced by different interactions due to pH and metal ions, so triplexes and quadruplexes are sometimes produced (12,13). For instance, C is protonated at slightly acidic pH to become a hydrogen-bond acceptor C^+, which binds to the guanine of a G–C pair as a bond donor to yield a C^+•G–C triad in a DNA triplex (14). The C^+ can be also paired with a non-protonated form of C, and the resulting C–C^+ pairs involve in formation of C-quadruplex, i.e., i-motif (15). Moreover, guanines can interact strongly

*To whom correspondence should be addressed. Tel: +82 54 279 2144; Email: seungsoo@postech.ac.kr
Correspondence may also be addressed to Donghwa Lee. Tel: +82 54 279 2160; Email: donghwa96@postech.ac.kr

with metal ions, and monovalent cations can influence a guanine-rich DNA strand to form a G-quadruplex (16,17). DNA should build thermodynamically-favored structures, which is a basic principle behind the spatiotemporal control of DNA nanoconstructs by modifying their environmental conditions.

Rational creation of advanced DNA structures requires in-depth understanding of DNA interactions under a variety of conditions; *ab initio* calculations can provide valuable insights in the behavior of proposed designs (18). This method provides reasonable values of dipole moments, charge distributions, and vibrational frequencies, so it can be useful to describe non-covalent molecular interactions (19,20). Importantly, all donor-acceptor interactions can be easily itemized, so that their independent functions in structure stabilization can be readily analyzed, and the relevant net energy is computed precisely. Therefore, the *ab initio* simulation has been exploited to interpret nucleic acid interactions. However, most previous studies have considered only interactions in gas phase (21–24), whereas actual DNA structures are in aqueous solutions. Some researchers have conducted phase-dependent simulation (25–27), but the calculation works used different computational conditions, making itemized DNA interactions not comparable to each other. Furthermore, the difference between computational and experimental results could not be well interpreted due to lack of effective analytical factors (26,28); in some calculations, the Watson–Crick base pair was less stable than a mismatch (MM) pair, i.e., a non-Watson–Crick base pair (26), and stacking of bases did not yield planarity (28). Therefore, there is a strong need for more realistic calculations of itemized DNA interactions under the unified conditions and analyses, making the computational results comparable with experimental ones.

In this work, we conducted advanced *ab initio* calculations to simulate DNA duplexes, triplexes, and quadruplexes by using identical computation conditions and rationally itemized donor-acceptor interactions (Scheme 1a). To provide the fundamental information for designing and manipulating complicated DNA structures along with predictions of their potential stability and transient motions, we systematically analyzed stability orders of the itemized interactions and relevant structural changes. Specifically, we used the water continuum model to mimic the real environment in nature, and nucleobase interactions were specifically focused on; all related 3D structures were intensively investigated with important parameters such as numbers of bases and layers, and environmental effects (e.g., protonation, water solvation, and metal–ion interactions). Moreover, key interactions such as Watson–Crick and Hoogsteen base pairs, metal ionic bonds, and π–π stacking were independently componentized to facilitate identification of correlations among interactions and stabilizing structures. To fully understand G-quadruplex behaviors, we considered both base-base and ion-water interactions; to our best knowledge, we first provided the clue to the preference of potassium ion (K^+) in G-quadruplex formation at an *ab initio* level. We also devised effective analytical factors, including hydrogen bond length (HBL), glycosidic vector angle (GVA), and twist angle (TA) (Scheme 1b). Using these structural factors, we thoroughly analyzed the relaxed structures by *ab initio* calculations and experimentally-determined structures by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, and those analyses were highly useful to identify how phosphate backbones influence nucleobase interactions. Owing to the unified computational view, various DNA constructs could be compared to each other to provide useful information, such as a stability trend among Watson–Crick base pairing, Hoogsteen base pairing, stacking, and metal-ion binding.

MATERIALS AND METHODS

For our *ab initio* calculations, we used the Gaussian 09 (G09) program. To effectively investigate nucleobase interactions, DNA backbones including sugar rings and phosphate groups were replaced with hydrogen atoms. Based on the ideal structure of B-DNA, an initial model of two-layer structure was constructed with a parallel arrangement of base-pair layers, maintaining 3.4-Å stacking distance and 36° angle under its center of mass. Previous calculation and experimental determination reports have validated the feasibility of this model (24,29). The double-layer model was guided to initially have a fixed dihedral angle (Opt = ModRed), minimizing possible non-planarity caused by the absence of a backbone, and its optimization process was then followed. Every structure optimization and relaxation was conducted under M05–2X/6–31G(d,p)//M05–2X/6–31G(d,p) condition, and the basis set superposition error (BSSE) calculations were further corrected by the counterpoise (CP) method (30). The water solvent effect was realized by using a conductor-like screening model (COSMO) (31).

Shapes of DNA structures were determined by considering the numbers of bases and layers. The number of bases affects the hydrogen bonding among the nucleobases, whereas the number of layers represents π–π stacking interactions among the base-pair layers. Metal ions were added to construct G-tetrad or G-quadruplex structures in which the cations interact electrochemically with guanines.

All interactions are expressible in simple formulae. The binding energy ΔE_{bind} is determined as the difference between the energy $E_{complex}$ of a complex system and the energy E_{sub} of an isolated subsystem:

$$\Delta E_{bind} = E_{complex} - \Sigma E_{sub}$$

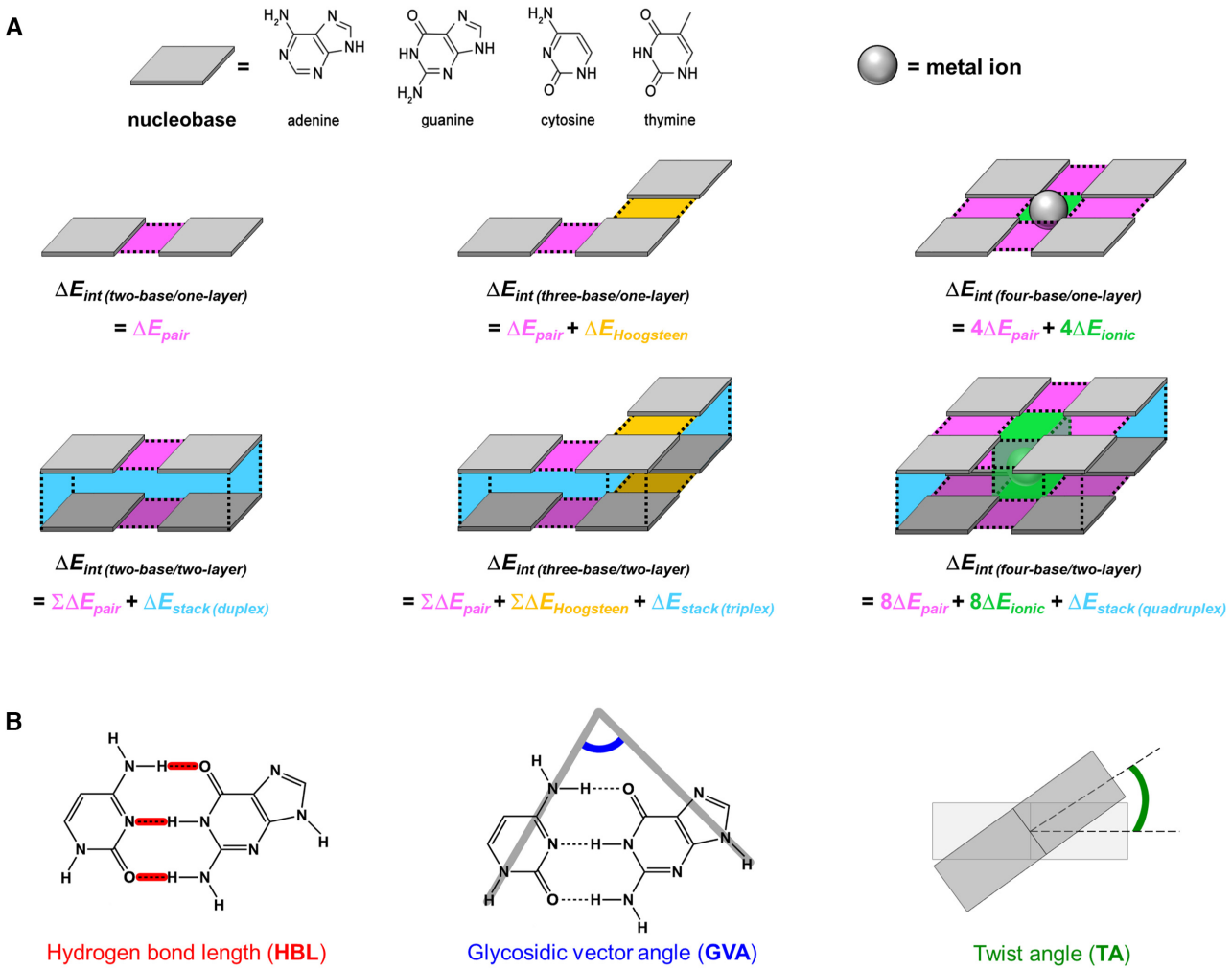
The ΔE_{bind} can be rewritten for a case of typical AB dimer, which is formed by binding monomer A and B, and ΔE_{bind} can be defined:

$$\begin{aligned}\Delta E_{bind}(AB) &= E_{AB}^{AB}(AB) - E_{A}^{A}(A) - E_{B}^{B}(B) \\ &= E(AB) - E(A) - E(B)\end{aligned}$$

In this formula, superscripts, subscripts, and symbols in parenthesis denote basis, geometry, and chemical system, respectively.

This binding energy formula can be used to derive the hydrogen bonding energy ΔE_{pair} of Watson–Crick or MM base pairs:

$$\Delta E_{pair} = E_{complex} - \Sigma E_{base}$$



Scheme 1. (**A**) Simulated nucleobase model structures and their itemized interactions. Planar base pair, triad, and tetrad and their double-layered structures are constructed for *ab initio* simulation. Gray tiles: nucleobases; spheres: monovalent metal ions. One-layer structures can have three different interactions: Watson–Crick or mismatch (MM) base-pairing (magenta), Hoogsteen base-pairing (yellow), and ionic bonding (green), which have interaction energies ΔE_{pair}, $\Delta E_{Hoogsteen}$, and ΔE_{ionic}, respectively. We note that one metal cation forms an ionic bond per nucleobase in the base tetrad. Two-layer structures additionally include a stacking interaction (cyan) which has energy ΔE_{stack}. Overall interaction energy of a simulated structure is described as ΔE_{int}. (**B**) Structural factors for conformational analysis. Hydrogen bond length (HBL, red) is defined as the length of hydrogen bonds of paired nucleobases. Glycosidic vector angle (GVA, blue) is an angle between two glycosidic vectors of nucleobases. Twist angle (TA, olive) is a rotation angle of one layer with respect to another. GVA and TA are obtained by vector dot-product calculations.

E_{sub} can be specifically defined; for instance, within a triad, E_{sub} can be divided into the interaction energy ΔE_{pair} of a base pair and the energy of all bases ΣE_{base}, so the Hoogsteen base-pairing energy $\Delta E_{Hoogsteen}$ can be obtained as

$$\Delta E_{Hoogsteen} = E_{complex} - \left(\Delta E_{pair} + \Sigma E_{base}\right)$$

Similarly, the interaction energy of G-tetrad ($\Delta E_{G\text{-}tetrad}$) is given by

$$\begin{aligned}\Delta E_{G-tetrad} &= E_{complex} - \left(4\Delta E_{pair} + 4E_{base} + E_{ion}\right) \\ &= 4\Delta E_{ionic}\end{aligned}$$

$\Delta E_{G\text{-}tetrad}$ is four times the interaction energy ΔE_{ionic} between a metal ion and a guanine because the monovalent cation forms one ionic bond per guanine.

The stacking interaction energy of a duplex is

$$\Delta E_{stack} = E_{complex} - \left(\Sigma\Delta E_{pair} + \Sigma E_{base}\right)$$

We can describe those of triplex and quadruplex in a similar way.

To obtain the CP-corrected ΔE_{bind} ($\Delta E^{CP}_{bind}(AB)$), we additionally performed BSSE calculations (30). The basis of isolated monomer A and B is overlapped when the monomer A and B bind to each other to form a dimer AB. The AB can be further stabilized because of the basis overlap. To correct the additional binding interactions by BSSE, an extra basis overlap to each monomer should be removed, and the BSSE for monomer A ($E_{BSSE}(A)$) and B ($E_{BSSE}(B)$) can be described as:

$$E_{BSSE}(A) = E^{AB}_{AB}(\mathrm{A}) - E^{A}_{AB}(\mathrm{A})$$

$$E_{BSSE}(B) = E_{AB}^{AB}(\mathrm{B}) - E_{AB}^{B}(\mathrm{B})$$

$E_{AB}^{AB}(\mathrm{A})$ and $E_{AB}^{AB}(B)$ are obtained from an optimized dimer geometry by assigning a ghost fragment to each monomer; $E_{AB}^{A}(\mathrm{A})$ and $E_{AB}^{B}(B)$ are the energies of individual bases from the optimized complex geometry.

Then, the $\Delta E_{bind}^{CP}(AB)$ can be obtained by subtracting $E_{BSSE}(A)$ and $E_{BSSE}(B)$ from ΔE_{bind}:

$$\Delta E_{bind}^{CP}(AB) = \Delta E_{bind}(AB) - E_{BSSE}(A) - E_{BSSE}(B)$$

By this formula, we corrected the ΔE_{bind} and acquired the $\Delta E_{bind}^{CP}(AB)$ values for all structures in our *ab initio* calculation study.

Finally, the overall interaction energy ΔE_{int} of intermolecular interactions can be shown as

$$\Delta E_{int} = \Sigma\Delta E_{pair} + \Sigma\Delta E_{Hoogsteen} + \Sigma\Delta E_{ionic} + \Delta E_{stack}$$

In this formula, each interaction energy term, which is corrected by the CP method, can be zero depending on the type of structure; for example, in a two-base/two-layer structure, only $\Sigma\Delta E_{pair}$ and ΔE_{stack} are non-zero.

Without phosphate backbones, DNA structures in our calculations could be structurally distorted to find the most stable conformations. To identify the structural changes by the presence of the phosphate backbones, we compared the computationally-predicted DNA structures and the experimentally-determined structures. Among a number of DNA structures that have been determined by X-ray crystallography and solution NMR spectroscopy, we only considered the typical DNA helical structures; the DNA-containing complex structures (e.g., peptide nucleic acid–DNA, RNA–DNA and protein–DNA) and the DNA structures that possess nucleobase analogs were excluded. For G-quadruplexes, we analyzed both parallel and anti-parallel ones originated from *Oxytricha nova*, *Tetrahymena* and human telomeres. The PDB IDs of chosen DNA structures are 1ZF7, 1ZFB, 1ZFC, 1ZFH, 1ZFM (32), 1BWT (33), 1CS2 (34), 1D68 (35) for duplexes, 149D (36), 1BWG (37), 1D3X (38) for triplexes, and 1JB7 (39), 2AQY (40), 143D (41), 186D (42), 201D (43), 230D (44), 352D (45), 1JPQ (46), 1JRN (46), 1K8P (47), 1KF1 (47), 2GKU (48), 2JPZ (49) for G-quadruplexes.

RESULTS AND DISCUSSION

Overall *ab initio* calculation conditions and analytical factors

Using G09, we conducted every optimization and relaxation step under the M05–2X/6–31G(d,p)//M05–2X/6–31G(d,p) condition. The M05–2X method (50) is a specially-designed tool that describes non-covalent interactions of various biomolecular structures, making it possible to obtain calculation results of hydrogen bonding and base stacking interactions. Previously, this method has proven its suitability for nucleic acid calculations (51–53); specifically, hydrogen bonding and stacking patterns of DNA by M05–2X were comparable with those by second-order Møller–Plesset perturbation theory (MP2) (53,54), a relatively expensive method. The 6–31G(d,p) basis set (55,56) belongs to the double-zeta polarization group that represents reliable intra- and inter-molecular geometries (57), so DNA nucleobase structures can be described sufficiently well. The solvent effect of water phase was obtained using COSMO, which is useful to describe the Gibbs free energy, cavitation, internal energy, and entropy effect of the solvent (31). Computational results obtained using the COSMO solvation model agree well with experimental data when small molecules, such as nucleobases, are simulated (58,59). By applying the COSMO, we provided the appropriate solvation environment to DNA nucleobases during *ab initio* calculations. Our computational conditions yielded the calculation results remarkably close to those of triple-zeta basis set 6–311G(d,p) (Supplementary Figure S1a, b), and our COSMO model resulted in the interaction energy order that matches the general trend of nature better than that of integral equation formalism polarizable continuum model (IEFPCM) (Supplementary Figure S1c).

After the simulation, we systematically analyzed the relaxed structures by comparing structurally-important factors including HBL, GVA and TA (Scheme 1b). HBL is the length of hydrogen bonds between two adjacent nucleobases; it is involved directly in a pairing-interaction energy. As GVA is an angle between two glycosidic vectors of nucleobases, it is strongly correlated with structural integrity. If GVA deviates from the ideal value (72° for DNA duplexes), the DNA structure that contains the paired bases is assumed to be mechanically distorted. TA is a rotation angle of one layer with respect to another layer, and when twists occur within a helical DNA, torsional stress is expected. Using these structural factors, we systematically compared the relaxed DNA structures by *ab initio* calculations and the experimentally-determined ones by X-ray crystallography and NMR spectroscopy.

Two-base/one-layer structures: Watson–Crick and MM base pairs

A duplex is the most basic structure among 3D DNA constructs, and base pairing is formed by interactions between nucleobases. We calculated the interaction energy ΔE_{pair} of all possible two-base structures including the complementary Watson–Crick and non-complementary MM base pairs both in vacuum and in water (Supplementary Figures S2 and S3, Tables S1 and S2).

The ΔE_{pair} order in the water phase suggests that the number of hydrogen bonds between nucleobases influences the stability of base pairs (Table 1, left). The G–C pair, which has the lowest ΔE_{pair}, is highly stable due to its three interbase hydrogen bonds. In contrast, the C–C pair has a single hydrogen bond, so it is the least stable. Thus, we reconfirmed that the number of hydrogen bonds is a dominant factor to determine the degree of stability.

The simulated HBL and GVA of Watson–Crick and wobble pairs are quite close to the experimentally-obtained values (PDB ID: 1ZF7, 1ZFB, 1ZFC, 1ZFH, 1ZFM, 1BWT, 1CS2, and 1D68) (Figure 1 and Table S2). Analysis of experimentally-determined structures indicates that the Watson–Crick base pairs have 1.90 Å mean HBL and ∼72° GVA (71.8°) (Supplementary Tables S3 and S4), and only the three most stable pairs (i.e., G–C, A–T and G–T) satisfy both ranges; importantly, their stability order is G–C > A–T > G–T in water (Table 1, left), which is consis-

Table 1. Calculated interaction energy (kcal/mol) of base pair ΔE_{pair}, Hoogsteen pair (in triad) $\Delta E_{Hoogsteen}$, and ionic bond (in tetrad) ΔE_{ionic} in water phase. Hydrogen bonds are indicated by hyphens (-) for Watson–Crick or MM base pairs, and by bullets (•) for Hoogsteen base pairs. Ionic bonds between guanines and metal ions are represented by ellipsis (···). ΔE values are arranged in order of decreasing stability. However, $\Delta E_{Hoogsteen}$ values are grouped whether the Hoogsteen base pair includes a flipped third base or not. The values in parenthesis indicate energy per hydrogen bond.

Pair	ΔE_{pair}	Triad	$\Delta E_{Hoogsteen}$	Tetrad	ΔE_{ionic}
G–C	–11.45 (–3.82)	**C^+•G**	–11.04 (–5.52)	**G4···Li^+**	–31.28 (–7.82)
A–T	–6.80 (–3.40)	**G•G**	–7.81 (–3.91)	**G4···Na^+**	–27.68 (–6.92)
G–T	–6.77 (–3.38)	**T•A**	–6.50 (–3.25)	**G4···K^+**	–20.26 (–5.07)
G–G	–6.75 (–3.37)	**A•A**	–1.69 (–0.84)		
C–T	–6.58 (–3.29)	**rC^+•G**	–11.08 (–5.54)		
T–T	–6.43 (–3.21)	**rG•G**	–7.64 (–3.82)		
A–G	–6.41 (–3.20)	**rT•A**	–4.90 (–2.45)		
A–C	–3.39 (–1.70)	**rA•A**	–3.83 (–1.92)		
A–A	–3.25 (–1.62)				
C–C	–3.16				

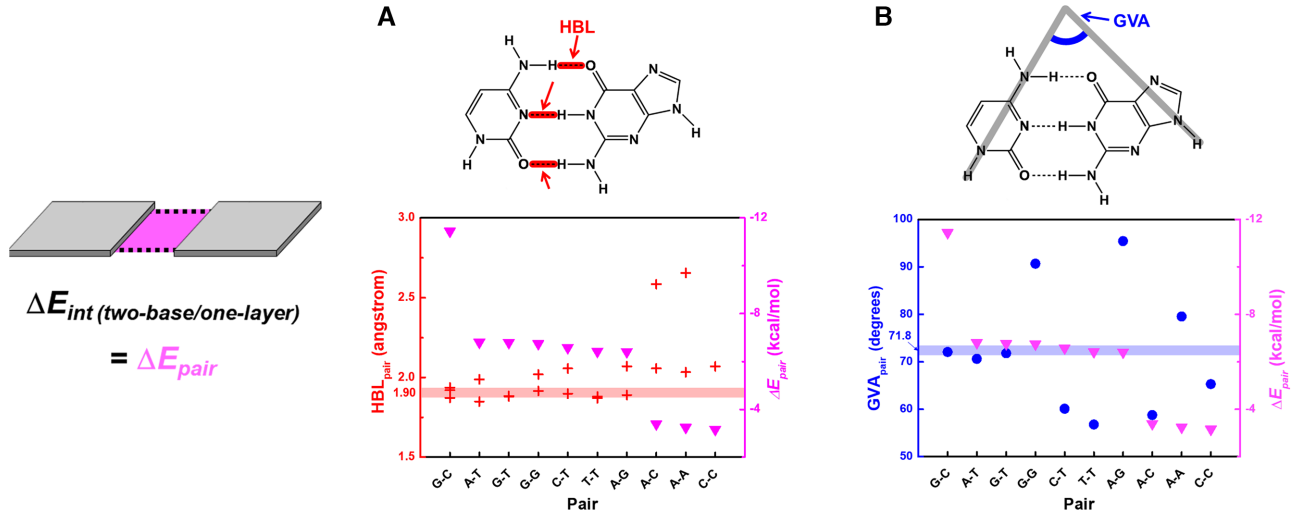


Figure 1. (**A**) HBL (red cross) and (**B**) GVA (blue circle) of Watson–Crick and MM base pairs, along with ΔE_{pair} (magenta inverted triangle). Solid lines indicate the mean HBL and GVA of Watson–Crick base pairs obtained from X-ray crystallography and NMR measurements (PDB ID: 1ZF7, 1ZFB, 1ZFC, 1ZFH, 1ZFM, 1BWT, 1CS2 and 1D68).

tent with the actual character of DNA (60). We note that when we conducted coupled cluster single-double excitation (CCSD) calculations (61–63), one of the most accurate and expensive calculations, the stability order was identical, justifying the reliability of our M05–2X method for optimizations (Supplementary Figure S1b). In our calculations, the G-T wobble pair has similar mean HBL and GVA values to those of the perfectly-matched A-T pair (Supplementary Table S2). Nevertheless, this non-Watson–Crick pair is less stable than the A–T pair due to the inconsistent pattern of hydrogen donor and acceptor. In Watson–Crick base pairs, the keto O4 of thymine and amino N2 of guanine participate in formation of intrabase hydrogen bonds for A–T and G–C pairs, respectively. However, in the G–T pair, they are not involved in those bonding patterns, so this pair has low stability (64).

Compared to the top three stable pairs, other MM pairs (e.g., G–G, C–T, T–T, A–G, A–C, A–A, and C–C) have GVAs that deviate significantly from the ideal value of ∼72° (Figure 1B). Especially, purine-purine pairs have GVA >> 72°, so steric hindrance between the large purines is anticipated. The effect of the steric hindrance in MM pairs also can be observed by long mean HBL values, which are 1.98 Å for A–G, 1.97 Å for G–G, and 2.34 Å for A–A (Supplementary Table S2). However, A–G and G–G have GVA close to 90°, so these purine-purine pairs have potential to seize each other in a square to form a tetrad (65). The T–T MM pair has a similar mean HBL to that of the G–T wobble pair (Supplementary Table S2), but the T–T MM pair is less stable than the G–T wobble pair. Moreover, as the T–T MM pair has the GVA << 72°, so the duplex would be further destabilized within a DNA double helix due to structural distortion. Thus, we conclude that for a duplex to be stable in nature, both HBL and GVA values should be in the appropriate range.

A material's phase has a critical influence on the ΔE_{pair} order. For example, in vacuum, the Watson–Crick A–T pair is less stable than some MM pairs (Supplementary Table S1) as it has a deviated GVA = 67.4° (Supplementary Table S2). However, in an aqueous environment, the A–T pair has appropriate GVA = 70.6°, so the stability order of base pairs becomes the actual G–C > A–T > G–T. This result confirms that the vacuum condition does not represent the actual tendency of natural DNA, and that nucleobase simulation should be performed in an aqueous environment.

The ΔE_{pair} order and the GVAs of base pairs would be highly useful for a duplex design. For example, single nucleotide polymorphism (SNP) detection requires to design DNA probes to selectively bind the SNP site (66), which would be benefited by our ΔE_{pair} order analysis. Additionally, the GVAs involve in the width and depth of grooves in DNA, related to protein–DNA binding interactions, and by incorporating MM pairs, groove manipulation would be achievable for gene regulation, including transcriptional control (67).

Two-base/two-layer structures: Watson–Crick base pair and stacking interactions

π–π stacking interaction has a strong influence on formation of double helix structures, i.e., a stack of base-pair layers. We next considered the stacking interaction for construction of DNA duplex, and two-layer structures of perfect-matched Watson–Crick base pairs were simulated (Supplementary Figures S4 and S5).

In vacuum, ΔE_{stack} is affected by the degree of layer (non)planarity (Supplementary Figure S4). Even though all structures are initially constructed with parallel arrangements of two planar base pairs, the base-pair layers become nonplanar after structural optimization (Supplementary Figure S4, Table S2). The calculations in the gas phase cannot include the neutralization effect of the water continuum; additional hydrogen bonds can form between the layers, so the initial parallel structures are forced to be severely distorted. Compared to the mean GVA of single layers, those of the double layers are also decreased as a result of this structure distortion (Supplementary Table S2).

When the phase condition is changed from gas to water, the base-pair layers become planar (Supplementary Figure S5), and all the GVA are close to 71.8°, as expected from conventional Watson–Crick pairings (Supplementary Table S2). The introduction of the water continuum model successfully excludes the undesired dipole-dipole interactions in the gas by using Van der Waals interactions between solvent and solute. Formation of unwanted hydrogen bonds is interrupted between the base-pair layers, so the planar surfaces can be maintained even after optimization. In water, all calculated structures have relatively similar ΔE_{stack} (Table 2, left); this result may occur because when two layers of purine-pyrimidine pairs are stacked on each other, the degree of layer overlapping would be quite similar under the influence of the π–π stacking.

Consequently, this reduces the influence of ΔE_{stack} on the order of ΔE_{int} in water; the influence of ΔE_{pair} is significant when perfect-matched base pairs are layered (Supplementary Table S1). The ΔE_{int} order mainly depends on the percent of G–C content in the entire structures, meaning that the number of hydrogen bonds dominantly contributes to determining the DNA stability, and that the influence of π–π interactions is relatively less important.

The significance of our calculation results is that the stability order by ΔE_{int} calculations shows an almost identical trend to those of nearest-neighbor thermodynamics results (68): G–C/C–G = C–G/G–C > G–C/G–C > G–C/A–T = C–G/T–A = C–G/A–T = G–C/T–A > A–T/A–T > A–T/T–A > T–A/A–T (Figure 2A). However, the experimental result shows a ΔG difference of A–T/T–A and T–A/A–T because the actual DNA structure has directionality as a result of the orientation of the 3′ and 5′ carbons along the DNA backbone, whereas this effect is not considered in our calculation model. Moreover, we did not consider several factors such as salts (cations and anions), pH, and strand-end effects, because these factors have much less influence on the stability of DNA than do the interactions between bases (8–11). Inclusion of these factors might enable correct prediction of the order of A–T/T–A and T–A/A–T.

The TA calculation results also support the anticipated ΔE_{int} order (Figure 2b, Supplementary Table S5). Unlike actual DNA duplex structures, base-pair layers in our computational model do not have phosphate backbones, so the layers can rotate freely to find the most stable conformations. Compared to the ideal double helix that has TA = 36.0°, underrotated or overrotated helical structures must be less stable, so the TA value can be an indicator of potential torsional stress, showing the influence of the phosphate backbones to base-base and stacking interactions. The real ΔE_{int} values of G–C/T–A and A–T/T–A are supposed to be higher than their calculated ΔE_{int} due to their underrotation; this does not affect the ΔE_{int} order, but their energy differences with G–C/A–T and A–T/A–T increase respectively. The torsional stress effect on G–C/C–G is not clear yet, so further investigation is ongoing. However, these TA calculations would be still useful information for precisely designing complex DNA structures, such as DNA origamis. Given that torsional stress of phosphate backbones is readily considered, the sum of TAs for all layers ideally determines the turns of helix; by calculating half or full turns of helices, the front and back of 2D patterns could be well distinguished (69), and even the curvature (70) or twist (71) of 3D DNA structures would be finely tuned based on the relation between TAs and DNA turns. Furthermore, as the TA values reflect the latent torsion stress, the TA calculations could be valuable for the design of DNA structures performing nanomechanical motions (72,73).

Three-base/one-layer structures: Watson–Crick and Hoogsteen base pairs

The DNA duplex can be extended to construct a triplex when third neighboring bases are introduced to provide additional electron donor-acceptor interactions. The additional interactions can be found at a Hoogsteen edge, termed a Hoogsteen base pair. To analyze the Hoogsteen interaction, we conducted simulation of planar three-base structures by positioning the third bases near the optimized Watson–Crick base pairings. In this study, we chose six natural (C^+•G–C, T•A–T, rT•A–T, rA•A–T, G•G–C and rG•G–C, where r indicates a flipped base) (Table 1, see notation) (12,74,75) and two unnatural (rC^+•G–C and A•A–T) triads (Supplementary Figures S6 and S7). C^+•G–C and T•A–T triads have been widely used to design DNA triple-stranded helices, wherein the C^+•G–C requires a slightly acidic pH environment to protonate the third cytosine base (14). When a third strand is purine-rich, it can bind to a double helix in an antiparallel manner to form reversed Hoogsteen pairs, and three different triads (rT•A–T, rA•A–T and rG•G–C) that contain flipped third bases

Table 2. Calculated stacking interaction energy ΔE_{stack} (kcal/mol) of duplex, triplex, quadruplex in water. Layers are divided by slash (/). ΔE_{stack} is arranged in order of decreasing stability, but ΔE_{stack} values of triplexes are grouped according to whether stacked third bases are parallel or antiparallel. Values in parenthesis: energy per ionic bond.

Duplex		Triplex		Quadruplex	
G–C/C–G	–5.46*	**A·A–T/A·A–T**	–11.87*	**G4/G4**	–13.16
A–T/A–T	–5.22	**C+·G–C/T·A–T**	–8.56*		
G–C/A–T	–5.21	**G·G–C/G·G–C**	–7.78*	**$\Sigma\Delta E_{ionic}$**	
A–T/T–A	–5.05*	**T·A–T/T·A–T**	–7.37	**G4···Na^+···G4**	–39.1 (–4.9)
G–C/G–C	–4.88	**C+·G–C/C+·G–C**	–3.77	**G4···K^+···G4**	–37.0 (–4.6)
G–C/T–A	–4.04*	**rT·A–T/rT·A–T**	–13.07*	**G4···Li^+···G4**	–32.8 (–4.1)
		rA·A–T/rA·A–T	–12.19		
		rG·G–C/rG·G–C	–7.07		
		rC+·G–C/rC+·G–C	–4.10*		

*If model structures further include sugar rings and phosphate backbones, additional torsional stress, *i.e.,* stability loss, is anticipated based on TA comparison between experimental and computational results.

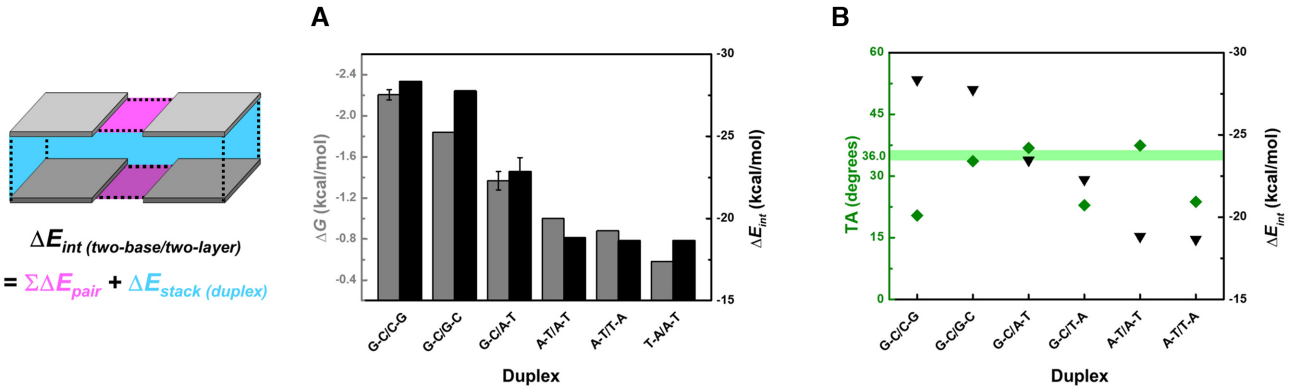


Figure 2. (**A**) Comparing thermodynamic ΔG (gray) (68) with calculated ΔE_{int} (black). Thermodynamic ΔG results of duplexes show the trend, G–C/C–G = C–G/C–G and G–C/A–T = C–G/A–T = C–G/T–A = G–C/T–A. Compared to experimental nearest-neighbor models, our computational ones cannot have the directionality, so ΔE_{int} of A–T/T–A and T–A/A–T are identical in simulation. (**B**) TA (olive diamonds) and ΔE_{int} (black inverted triangles) of two-base/two-layer structures. The two-base/two-layer structures are ordered by stability of duplex, and ΔE_{int} is the sum of $2\Delta E_{pair}$ (top and bottom layers) and $\Delta E_{stack\ (duplex)}$. Solid line: TA of duplex from crystallography and NMR data (PDB ID: 1ZF7, 1ZFB, 1ZFC, 1ZFH, 1ZFM, 1BWT, 1CS2 and 1D68).

are found in the antiparallel triple helix (76). G•G–C triads are not common, but they have been discovered in several high-resolution crystallography studies (77). For comparison with the natural three-base structures, two unnatural triads (rC^+•G–C and A•A–T) were also chosen.

Regardless of phase and base flipping, the stability order determined by $\Delta E_{Hoogsteen}$ is (r)C^+•G > (r)G•G > (r)T•A > (r)A•A (Table 1, middle). When we obtained the mean HBL of Hoogsteen base pairs from available solution NMR data (PDB ID: 149D, 1BWG and 1D3X), (r)C^+•G and (r)A•A show somehow the difference between experimental determination and calculation. The calculated HBLs of (r)C^+•G are shorter than the experimentally-determined ones (Figure 3A); when the (r)C^+•G–C triad is connected to a phosphate backbone, the (r)C^+•G is supposed to be under tensile stress, so the actual $\Delta E_{Hoogsteen}$ is higher than the calculated one. The simulated (r)A•A shows longer HBLs, so compressive stress is expected in presence of the backbone, similarly decreasing the stability. On the contrary, the (r)G•G and (r)T•A show the relatively smaller difference between experimentally-observed and calculated HBLs, indicating that their $\Delta E_{Hoogsteen}$ would be less relevant to the absence of the phosphate backbone. Interestingly, the calculated $\Delta E_{Hoogsteen}$ of a triad is proportional to its mean HBL (Supplementary Table S2), whether the third base is flipped or not. Compared to the Hoogsteen pairs, Watson–Crick base pairs within the simulated triads have similar HBLs and GVAs to each other (Supplementary Tables S6 and S7, Figure S8), and they are quite close to the ideal values. This result suggests that the pre-existing Watson–Crick pairs are not significantly affected by the access of third bases during formation of Hoogsteen base pairs. Hoogsteen interactions do not much depend on Watson–Crick interactions, so $\Delta E_{Hoogsteen}$ would be crucial to stabilize a triple helix.

Among the calculated triads that have unflipped third bases, C^+•G can form the most stable Hoogsteen pair (Table 1), due to the molecular ion-molecular dipole interaction among nucleobases (78). In contrast, the A•A Hoogsteen base pair is energetically disfavored within the A•A–T triad. Interestingly, $\Delta E_{Hoogsteen}$ of G•G–C is slightly lower than that of T•A–T even though the T•A–T is found much more frequently than G•G–C in nature. Although the HBLs of their Hoogsteen pairs are similar to each other, the G•G–C has large GVA > 90° as a result of its purine-purine pairing (Figure 3B). This large GVA can cause a severe structural distortion and therefore cannot be accepted in an

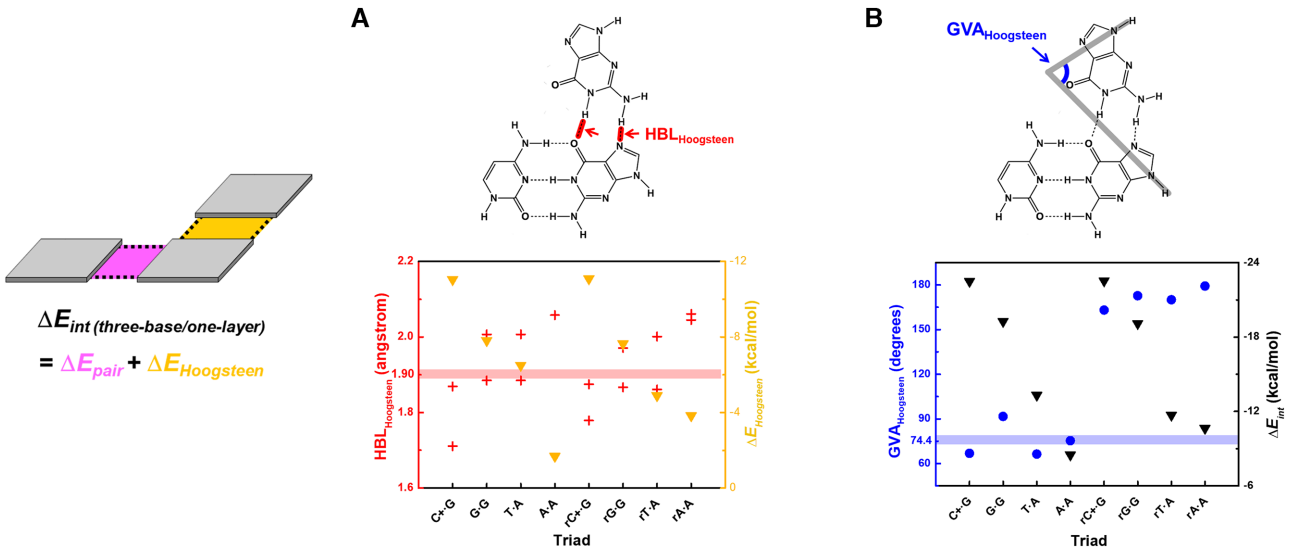


Figure 3. (**A**) HBL (red cross) and $\Delta E_{Hoogsteen}$ (yellow inverted triangle) of Hoogsteen base pairs in simulated triads. Solid line indicates the NMR mean HBL of C$^+$•G and T•A (PDB ID: 149D, 1BWG and 1D3X). (**B**) GVA of Hoogsteen base pairs (blue circle) and ΔE_{int} of triads (black inverted triangle), which is the sum of ΔE_{pair} and $\Delta E_{Hoogsteen}$. Solid line indicates the NMR mean GVA of C$^+$•G and T•A.

actual triple helix. However, our simulated model freely accepts this distortion due to exclusion of the phosphate backbone. According to this structural freedom, $\Delta E_{Hoogsteen}$ of G•G–C is comparable with that of T•A–T only in the calculation; in reality, the G•G Hoogsteen pair would be less stable than the T•A Hoogsteen pair.

Surprisingly, the most stable triad is rC$^+$•G–C, which is an unnatural triad that has not been reported in nature (12,74,75). The stability of rC$^+$•G–C is expected to be caused by the flipped third base; as a result, the oxygen atom of rC$^+$ has a chance to approach the hydrogen atom of C and thereby causes formation of an additional hydrogen bond (Supplementary Figure S9). However, the actual formation of DNA triple helix must consider the backbone directionality of the third strand. The triads that contain flipped third base show calculated GVA closed to 180° (Figure 3b), which indicates that when a third strand is located within the major groove of double helix, its positioning can be disturbed by steric hindrance or helical under/overwinding. In addition, the rC$^+$•G–C may cause electrostatically repulsive interactions with neighboring layers within the triple helix. The effect of third base flipping is strongly relevant to stacking interactions, so it is further evaluated in the next section.

Three-base/two-layer structures: Hoogsteen base pairs and stacking interactions

In nature, a base triad alone is not stable, but its stacking structure, i.e., a triple helix, can be retained. In that configuration, π–π stacking interactions between the triad layers greatly contribute to formation of a stable triple-stranded helical structure. For next three-base/two-layer modeling, we first arranged the well-known triads (C$^+$•G–C and T•A–T) to form C$^+$•G–C/C$^+$•G–C, C$^+$•G–C/T•A–T and T•A–T/T•A–T triplexes. Second, to investigate a stacking effect of purine-purine Hoogsteen base pairs, we stacked G•G–C and A•A–T onto the same triads to construct G•G–C/G•G–C and A•A–T/A•A–T double layer structures. Finally, we simulated homolayers composed of triads that contain flipped third bases (rC$^+$•G–C/rC$^+$•G–C, rT•A–T/rT•A–T, rG•G–C/rG•G–C and rA•A–T/rA•A–T) to observe the effects of antiparallel strands in a triplex (Supplementary Figures S10 and S11).

According to the ΔE_{stack} order in both phases, the C$^+$•G–C/C$^+$•G–C stacking is the least stable among the simulated triplexes (Table 2, middle; Supplementary Table S1). The stack instability is induced by strong charge-charge repulsion between C$^+$ of each layer, and this repulsion distorts the triad layer into a nonplanar structure (Supplementary Figures S10 and S11). Therefore, when triple helices are rationally designed, successive C$^+$•G–C layers should be avoided, and C$^+$•G–C/T•A–T and T•A–T/T•A–T can be favored instead (79). C$^+$•G–C/T•A–T triplexes have lower ΔE_{stack} than T•A–T/T•A–T due to the different π–π stacking energy $\Delta E_{stack(3rd\ base)}$ between the third bases. $\Delta E_{stack(3rd\ base)}$ can be simply obtained by subtracting $\Delta E_{stack(duplex)}$ from $\Delta E_{stack(triplex)}$ (Figure 4A). The calculated $\Delta E_{stack(3rd\ base)}$ suggests that the C$^+$/T stack of C$^+$•G–C/T•A–T is more stable (–3.35 kcal/mol) than the T/T stack of T•A–T/T•A–T (–2.15 kcal/mol). These observations indicate that insertion of C$^+$•G-C/T•A–T would stabilize the triple-stranded helix, in terms of stacking, better than insertion of T•A–T/T•A–T. However, the C$^+$•G–C/T•A–T has smaller TA than the experimentally-determined value (Figure 4B, Supplementary Table S8), inferring that during triplex structure design, repetitive insertion of C$^+$•G–C/T•A–T should be avoided because of undesired torsional stress.

$\Delta E_{stack(3rd\ base)}$ of the A/A stack in A•A–T/A•A–T (–6.65 kcal/mol) is the lowest among the triplexes that do not have flipped third bases. However, $\Delta E_{Hoogsteen}$ of A•A Hoogsteen base pair (–1.69 kcal/mol) is significantly higher than the others (Table 1, middle), and the small TA = 24.6°

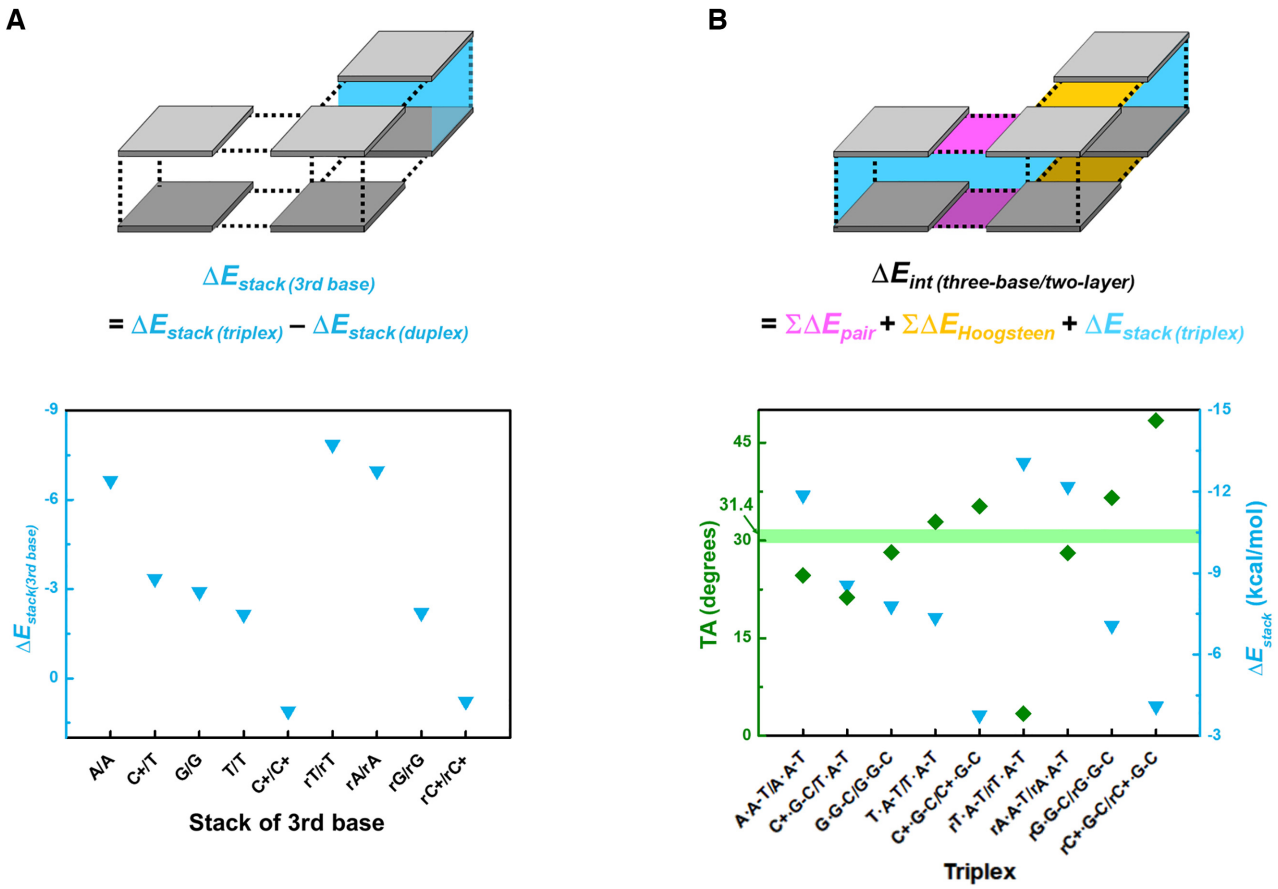


Figure 4. (**A**) Stacking interaction energy of third bases ($\Delta E_{stack\ (3rd\ base)}$), which is obtained by subtracting $\Delta E_{stack\ (duplex)}$ from $\Delta E_{stack\ (triplex)}$. (**B**) TA (olive diamonds) and $\Delta E_{stack\ (triplex)}$ (cyan inverted triangle) of triplexes. The mean TA of experimentally-determined C^+•G–C/T•A–T and T•A–T/T•A–T is represented as solid line (PDB ID: 149D, 1BWG, and 1D3X).

of A•A–T/A•A–T indicates that this underrotated structure would be affected by torsional stress; this may be the reason that tandem repeats of A are not included in the third parallel strand of a natural triple helix, but can be only permitted by specifically-designed DNA structures in the laboratory (80).

Another triad stacking structure that contains purine-purine Hoogsteen pairs, G•G–C/G•G–C, has a value of $\Delta E_{stack(\ 3rd\ base)}$ (–2.91 kcal/mol) that is higher than that of A•A–T/A•A–T, but is lower than that of T•A–T/T•A–T. Again, the calculated $\Delta E_{Hoogsteen}$ value of G•G–C is comparable with that of naturally-occurring T•A–T, and the simulated TA of the G•G–C/G•G–C double layer has a negligible difference from the ideal TA value; it is therefore assumed to be unaffected by torsional stress. These observations suggest that G tandem repeats could be readily used to design a parallel triple-stranded helix if their steric hindrance is systematically compensated for GVA > 90° (Figure 3B).

Although rT•A–T/rT•A–T gives the most stable stacking interaction (Table 2, middle), such triad double layers are rare in naturally-occurring triple helices; the rarity may be caused by the preference for layer sliding as a phosphate backbone is absent in our computational model (81). During structure optimization, the rT•A–T triads slide preferentially to each other (Supplementary Figure S11) rather than to maintain the initial twist condition between layers (TA = 3.4°, Figure 4B). The highly decreased TA in *ab initio* calculations suggests that rT•A–T/rT•A–T double layers without the phosphodiester backbone may be stabilized at global minima. However, this layer sliding can allow the actual triple helix to underwind, making the helical structure destabilized due to torsional stress. Unlike the rT•A–T/rT•A–T, T•A–T/T•A–T shows no TA difference between the relaxed and experimentally-determined triplexes, indicating that T•A–T/T•A–T already has the preferred stacking structure only by base-base interactions. It would be the reason that when a poly-T strand binds to a DNA duplex, formation of parallel triple helix is more favored in nature than that of antiparallel one (82–84). Other double layers that include flipped third purine bases (e.g., rG•G–C/rG•G–C and rA•A–T/rA•A–T) display TA quite close to the ideal value of 31.4°. These observations may provide a clue to why formation of antiparallel triple helices requires a purine-rich third strand. Furthermore, using under- or overwinding information during triad stacking, optimal antiparallel triplexes can be readily designed by compensation of all torsional stress.

Four-base/one-layer structures: G-tetrad and donor-acceptor bonds in metal ion binding

Nucleobases can form a quadruplex when other elements are involved in its formation. For instance, monovalent cations (M^+) such as Na^+ and K^+ support nucleic acids to construct a G-quadruplex structure; briefly, addition of these ions affects interactions between guanine bases and thereby causes change in various factors (e.g., HBLs, GVAs and TAs). To investigate this ion-dependent formation of quadruplexes, we first focused on a guanine tetrad, which requires widely-accessible monovalent cations such as Li^+, Na^+, and K^+ (Supplementary Figures S13 and S14). To clarify the effect of the metal ion, we simulated and compared G-tetrad structures both with and without metal ions (Supplementary Figure S12).

When monovalent metal cations are located in the guanine tetrads, ionic radii are crucial to determine whether the ion positions can be in-plane or out-of-plane with respect to the planar G-tetrad layers. Electrostatically, the G-tetrad tends to catch a monovalent cation in the center of itself due to partially negatively-charged oxygen atoms in the guanines (Supplementary Figure S15). Even when Li^+ and Na^+ are initially placed out of the preformed G-tetrads in *ab initio* calculations, the electrostatic charge-charge attraction draws the metal ions into an in-plane orientation, so the ions move into the center of the G-tetrad. In contrast, K^+ is not in the center, but moves upwards along the central line perpendicular to the surface of the G-tetrad (Supplementary Figures S13 and S14); this observation is consistent with previous computational result (85). Compared to Li^+ and Na^+ that are readily accepted in the center of G-tetrad, the K^+ has a longer ionic radius. Its relatively large ion size causes severe steric hindrance, so it remains above the G-tetrad.

From our *ab initio* calculations of monovalent ion-containing G-tetrads (G4···M^+) in both phases, the stability orders determined by ΔE_{ionic} and ΔE_{int} are G4···Li^+ > G4···Na^+ > G4···K^+ (Table 1, right; Supplementary Table S1), which is the opposite of the natural behavior of G-quadruplexes: G4···K^+ > G4···Na^+ > G4···Li^+ (86,87). Interestingly, this disagreement is not a result of our computational setup that excludes phosphodiester backbones which strongly affect the size of a G-quadruplex. In our calculations, G-tetrad can freely distort their size and conformation to captures target cation, so the G-tetrad structures in simulation may behave differently compared to the actual ones (39–46,48,49) (Figure 5, Supplementary Table S9). To confirm it, we analyzed the computed HBLs that are involved in the stability of a G-tetrad in the presence of metal ion. G4···M^+ features two hydrogen bonds: NH-N and NH-O (Figure 5). The length of NH-N bond affects the size of the G-tetrad, whereas the NH–O bonds provide a space in which the metal ion is held. Mean computed HBLs of NH-N and NH-O are similar to the experimentally-determined in-plane G4···Na^+ (1JB7, 2AQY, 143D, 186D, 201D, 230D and 352D) (Figure 5A). However, the simulated G4···K^+ has longer NH-N mean HBL than crystallography and NMR spectroscopy measurements (1JPQ, 1JRN, 1K8P, 1KF1, 2GKU and 2JPZ), although NH-O bond lengths differ by < 2% from the experimentally-obtained counterparts (Figure 5B). The increased mean length of NH-N bonds in the optimized G4···K^+ indicates that the simulated K^+-containing G-tetrad is larger than a real one, so compressive stress can be exerted in presence of DNA backbones. Considering the energy loss by additional compression, the actual G4···K^+ should be less stable than our calculations suggests, whereas the G4···Na^+ would not be compressed, i.e., no further destabilization.

The length of the bond between a metal ion and a partially-charged oxygen of guanine (M^+···O) was next considered, and the calculated M^+···O lengths are shorter than the experimentally-determined ones (Figure 5). The difference is a result of stacking interactions between G-tetrad layers. An isolated G-tetrad experiences no π–π stacking interaction and importantly, provides only four oxygen atoms for a metal ion to form M^+···O bonds. Therefore, in a double G-tetrad structure, i.e., a G-quadruplex, the metal ion can be located at its correct position; this change would narrow the difference between computed and measured M^+···O lengths.

Four-base/two-layer structures: G-quadruplex and donor–acceptor bonds in metal ion binding

Two or more G-tetrad layers can be stacked by π–π stacking interactions; the result is called a G-quadruplex. We further simulated G-quadruplex structures with monovalent cations, Li^+, Na^+, and K^+, to investigate the stacking effect between the G-tetrads (Supplementary Figures S16 and S17). In constructing an initial simulation model, we positioned a metal ion in the center of the optimized G-quadruplex, so that the cation was out-of-plane with respect to both G-tetrad layers. This structural arrangement resembles the G-quadruplex formation reported in crystallography and NMR data (45,46). Simulated G-quadruplexes with metal ions (G4···M^+···G4) show similar planarity and TA values to each other, so we focused on optimization of metal ion sites and alteration of relevant bond lengths.

Computational results of G4···M^+···G4 reveal that the position of monovalent cation within an optimized G-quadruplex depends on the ion size (Supplementary Figure S17), as was observed in computational results of G4···M^+. In water, Li^+ does not remain in the center of G-quadruplex; in the two stacked G-tetrads, the Li^+ occupies the central space in only one tetrad layer. Construction of out-of-plane G4···Li^+···G4 requires long Li^+···O bond lengths due to the small size of Li^+, so its most stable arrangement is to have four Li^+···O bonds by forming an in-plane G4···Li^+ and an empty G-tetrad. In contrast, the large K^+ is out-of-plane with respect to both G-tetrad layers, and being centered in the G-quadruplex is maintained. As a result, the optimal G4···K^+···G4 structure has eight K^+···O bonds and is further stabilized. The order of radius of metal ion is $Li^+ < Na^+ < K^+$, so Na^+ can be confined either in a G-tetrad or in a G-quadruplex; this inference has been confirmed by crystallography and NMR studies (39). G-quadruplexes that contain in-plane and out-of-plane Na^+ have $\Delta E_{int} = -105.8$ and –106.3 kcal/mol, respectively (Supplementary Figure S18),

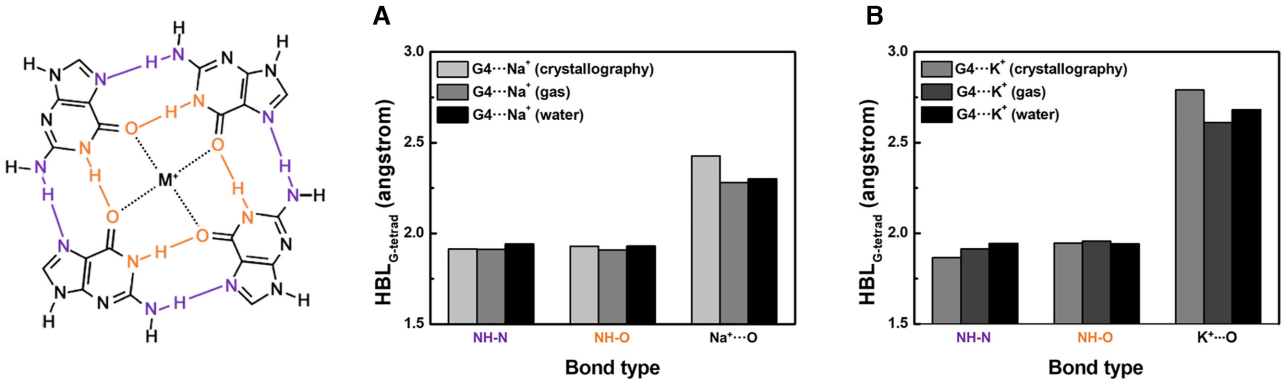


Figure 5. NH-N (purple), NH-O (orange), and M^+···O (black) bonds of G-tetrad in crystallography and NMR data (light gray) and in gas- and water-phase calculations (dark gray and black, respectively). Mean bond length of G4···Na^+ (**A**) and G4···K^+ (**B**). The experimentally-determined bond length of Na^+···O is obtained from in-plane G4···Na^+ tetrads (PDB ID: 1JB7, 2AQY, 143D, 186D, 201D, 230D and 352D).

and because of the similar energetic stability, we believe that both structures can be observable.

Even though metal ions are located in optimal positions of G-quadruplexes, the stability orders determined by the overall interaction energy ΔE_{int} of G4···M^+···G4 in both phases are G4···Na^+···G4 > G4···K^+···G4 > G4···Li^+···G4 (Supplementary Table S1), which does not match the general trend of nature. The formation energies of G-tetrads and their stacking are similar among three different monovalent cations (Supplementary Table S1), so the ΔE_{int} values heavily rely on the sum of M^+···O bond energies. The computed M^+···O mean lengths of G4···M^+···G4 structures are close to the experimentally-measured ones (39–46,48,49) (Figure 6A, B and Supplementary Table S9); this result indicates that the G-tetrad stacking yields the calculated ionic bonds that agree well with an actual G-quadruplex. In our calculations, both Na^+- and K^+-containing G-quadruplexes have eight M^+···O bonds, but ΔE_{ionic} of the Na^+···O bond (-4.88 kcal/mol) is lower than that of K^+···O bond (–4.63 kcal/mol). The in-plane Li^+-containing G-quadruplex, which has only four Li^+···O bonds, also has relatively higher ΔE_{ionic} (–4.10 kcal/mol). In presence of K^+, the calculated NH-N length of the G4···K^+···G4 is even longer than the experimentally-determined length (Figure 6B). This means that if phosphate backbones are further included, the simulated G-quadruplex can be compressively stressed, as was observed for the G-tetrad.

To mimic the interaction of G-quadruplex accurately, metal-ion hydration becomes the key interaction to be considered. The COSMO water continuum was successfully applied to calculate a variety of base-base interactions, but the computation of interactions in a G-quadruplex demands an additional interaction among bases and metal ions. Metal cations are always coordinated with aqua ligands in water and therefore must be dehydrated before entering the G-quadruplex (88). According to the hydration property of metal ion in solution, we adopted a chemical equation wherein a hydrated metal ion reacts with a G-quadruplex to form a metal ion-containing G-quadruplex by the release of water molecules, and the net energy $\Delta E_{M+binding}$ of this reaction determines whether the metal ion enters the G-quadruplex.

$$M^+ + nH_2O \xrightarrow{\Delta E_{hydration}} M[H_2O]_n^+ ,$$

$$8G + M^+ \xrightarrow{\Delta E_{ionic}} M[G]_8^+ ,$$

$$M[H_2O]_n^+ + 8G \xrightarrow{\Delta E_{M^+binding}} M[G]_8^+ + nH_2O,$$

$$\therefore \Delta E_{M^+binding} = \Delta E_{ionic} - \Delta E_{hydration}$$

$\Delta E_{M+binding}$ is the difference between ΔE_{ionic} and the metal ion hydration energy $\Delta E_{hydration}$, so we next calculated $\Delta E_{hydration}$ values. First, the coordination number of the monovalent metal ions should be identified because the $\Delta E_{hydration}$ depends on the number of water molecules surrounding the metal ions. From previous computational and experimental studies (89–93), the coordination numbers by water are widely accepted to be 4 for Li^+, 5 for Na^+, and 6 for K^+. We applied those numbers in our model of hydrated metal ion structures to calculate $\Delta E_{hydration}$ values (Supplementary Figure S19; Table S10) and then obtained $\Delta E_{M+binding}$ values of different Li^+, Na^+, and K^+ within the G-quadruplex (Figure 7, inset). The calculated stability order of hydrated ions by $\Delta E_{hydration}$ was Li^+ > Na^+ > K^+, which is consistent with the observation that the affinity for water molecules decreases as ion size increases (85). Importantly, $\Delta E_{hydration}$ values of Na^+ and K^+ differ by 2.22 kcal/mol, whereas their ΔE_{ionic} values differ by only 0.25 kcal/mol. From these results, we can conclude that the stability order of G-quadruplexes by $\Delta E_{M+binding}$ is G4···K^+···G4 > G4···Na^+···G4 > G4···Li^+···G4. This result (with hydration) corrects the $\Delta Eint$ order to G4···K^+···G4 > G4···Na^+···G4 > G4···Li^+···G4 (Figure 7; Supplementary Table S11), which agrees well with the actual stability tendency of G-quadruplexes in nature. This successful correction indicates that the reason why K^+ binding to G-quadruplex is the strongest is that K^+ is the most easily dehydrated.

Our calculation results reveal the importance of metal ion hydration in determining the G-quadruplex stability,

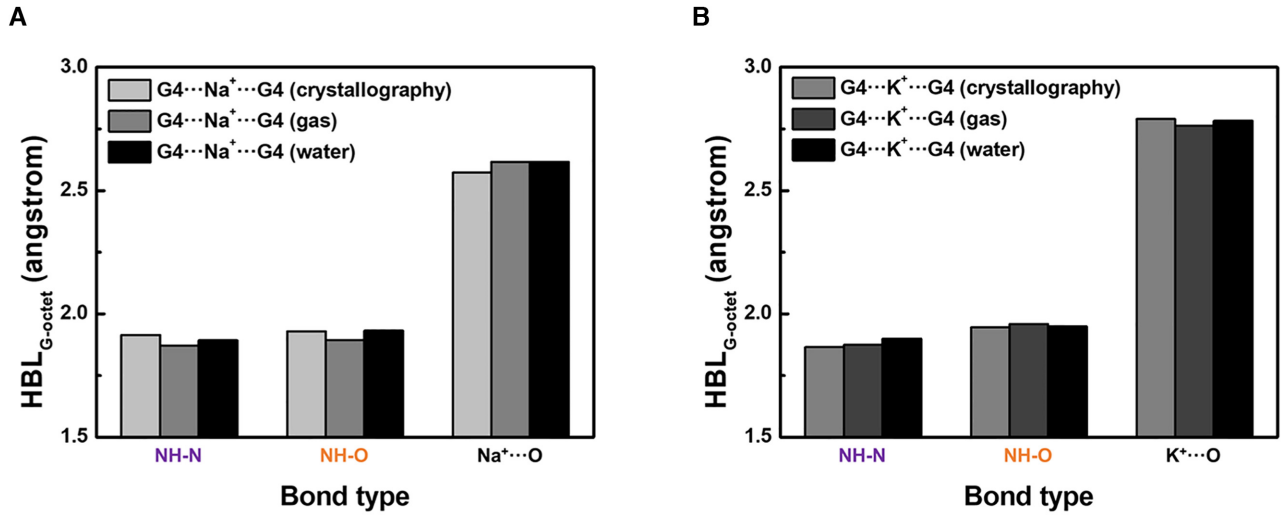


Figure 6. NH-N (purple), NH-O (orange), and M^+···O (black) bonds of G-quadruplex. Mean bond length of G4···Na^+···G4 (**A**) and G4···K^+···G4 (**B**). The experimentally-determined bond length of Na^+···O is originated from out-of-plane G4···Na^+ quadruplexes (PDB ID: 1JB7, 2AQY, 143D, 186D, 201D, 230D and 352D).

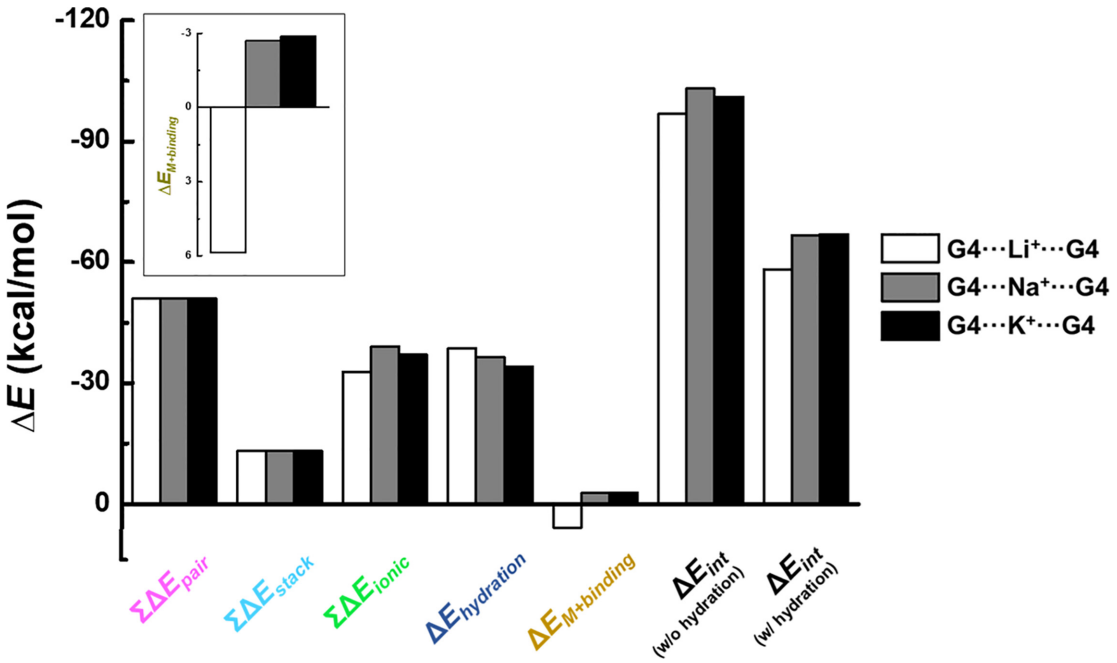


Figure 7. Energy components of G4···Li^+···G4 (white columns), G4···Na^+···G4 (gray columns), and G4···K^+···G4 (black columns). ΔE_{int} (w/o hydration) is summation of $\Sigma\Delta E_{pair}$, $\Sigma\Delta E_{stack}$ and $\Sigma\Delta E_{ionic}$, and ΔE_{int} (w/ hydration) is obtained by subtracting $\Delta E_{hydration}$ from ΔE_{int} (w/o hydration). Inset shows detail $\Delta E_{M+binding}$; specifically, Li^+ shows a positive value because its $\Sigma\Delta E_{ionic}$ value is higher than $\Delta E_{hydration}$ value.

which has been highlighted by previous experimental studies (94). Moreover, from this hydration-dependent stability, we infer that if different solvents or metal-binding ligands are introduced, G-quadruplex-favoring metal ions would be changed.

CONCLUSION

In this study, we explored fundamental interactions of DNA duplexes, triplexes, and quadruplexes by calculating key interactions (*e.g.*, hydrogen bonding, π-π stacking, ion binding, and metal-ion hydration) that strongly correlate with the actual DNA characteristics. To clearly identify the effect of each interaction, the involved interactions in molecular structure formation were componentized as independent terms in our *ab initio* calculations. In general, our simulated DNA structures and calculated energies agreed well with the actual behaviors of DNA; for example, the order in the interaction energy of simulated duplexes matches well with that obtained previously using nearest-neighbor thermodynamics (68). The agreement between computational and experimental results suggests that our improved *ab ini-*

tio calculations based on itemized interactions have potential to mimic natural DNA well.

The advantage of identical computation conditions in duplex, triplex and quadruplex calculations is striking. Regardless of the number of bases and layers, all calculated interaction energies are comparable among the simulated structures. Accordingly, we successfully obtained meaningful information:

First, ΔE_{pair} of pairs and $\Delta E_{Hoogsteen}$ of triads can be compared, and the resulting stability order is G–C (ΔE_{pair} = –11.45 kcal/mol) > C^+•G ($\Delta E_{Hoogsteen}$ = –11.04 kcal/mol) > A–T (ΔE_{pair} = -6.80 kcal/mol) > T•A ($\Delta E_{Hoogsteen}$ = –6.50 kcal/mol). The energy difference between G–C and C^+•G is attributed to the number of hydrogen bonds: the G–C Watson–Crick pair has three hydrogen bonds, but C^+•G Hoogsteen pair has only two. In addition, the canonical A–T Watson–Crick pair shows the greater stability than the non-canonical T•A Hoogsteen pair.

Second, the simulated triplexes have a wider range of ΔE_{stack} values (–3.77 to –13.07 kcal/mol) than do the duplexes (–4.04 to –5.46 kcal/mol). This observation suggests that in triple helices, the nucleobase pairs display strong sequence-dependence in structural designs, whereas in double-stranded helices, the sequence of bases is unconstrained.

Third, in G-tetrad and G-quadruplex structures that capture identical cations, the stability order by ΔE_{ionic} is G4···Na^+ (–6.92 kcal/mol) > G4···Na^+···G4 (–4.88 kcal/mol), and G4···K^+ (–5.07 kcal/mol) > G4···K^+···G4 (–4.88 kcal/mol). ΔE_{ionic} is affected by M^+···O length, so the relatively large K^+ that cannot be in-plane on a G-tetrad layer shows similar K^+···O length in both the G-tetrad and the G-quadruplex; this result explains the prevalence of K^+-containing G-quadruplex over G-tetrad in nature.

Fourth, the most encouraging result is the identification of metal ion hydration as an explanation of the K^+ preference in G-quadruplex formation. When considering only the water continuum, the stability order by the calculated ΔE_{int} is G4···Na^+···G4 > G4···K^+···G4 > G4···Li^+···G4, while G4···K^+···G4 > G4···Na^+···G4 > G4···Li^+···G4 is observed in nature. Addition of the dehydration effect successfully corrects the calculated order to the natural order. To our best knowledge, this consideration of both base-base and ion-water interactions has yielded the first successful correction of G-quadruplex ion preference by *ab initio* calculations.

Most donor-acceptor interactions in DNA construction are among nucleobases, but to be accurate, *ab initio* calculations must include structural components such as sugar rings and phosphate backbones, and also consider environmental conditions including surrounding ions and water molecules. Several analytical factors such as HBL, GVA, and TA were effective to indirectly examine the effect of DNA backbones, but an improvement of the simulation to consider the backbones is inevitably necessary. It is well known that the metal ions such as Na^+ and Mg^{2+} neutralize negative charges of phosphodiester backbones (95,96), and the DNA backbones determine syn- and anti-type nucleoside, and strand directionality, resulting in formation of 3D complex DNA structures. Moreover, G-quadruplexes can show parallel, anti-parallel, or even hybrid topologies according to the length and the sequence of DNA strand, and the different G-quadruplex topologies can be also affected by syn- and anti-conformations of guanines (97). Furthermore, different shapes of DNA duplexes such as A-, B- or Z-form DNA are influenced by the syn- and anti-conformations, and even sugar puckers (98). Thus, only the calculations of backbone-including DNA structures will enable quantitative analysis of the effects of metal ions, backbone directionality, syn- and anti-nucleosides, and sugar puckers.

Additionally, studies of modified nucleobases have observed that hydration of hydrogen-bonding acceptors in minor grooves can influence the stability of helix (99,100). The relatively small system such as ours tend to be affected by this hydration effect, so investigation of nucleobase interactions in presence of water molecules can give more accurate and meaningful results. Continuing advances in understanding of DNA structures and environment effects will enable calculations, designs, and programming of complicated DNA constructs, for use in DNA nanotechnology, in the near future.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Computational Nano-Materials Design Laboratory for supporting Nucleus and Nature server.

FUNDING

National Research Foundation of Korea (NRF) grant [NRF-2017R1C1B3012050]. Funding for open access charge: Brain Korea 21 FOUR project for Education and research center for future materials [F21YY7105002].
Conflict of interest statement. None declared.

REFERENCES

1. Porchetta,A., Idili,A., Vallée-Bélisle,A. and Ricci,F. (2015) General strategy to introduce pH-induced allostery in DNA-based receptors to achieve controlled release of ligands. *Nano Lett.*, **15**, 4467–4471.
2. Rothemund,P.W.K. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature*, **440**, 297–302.
3. Oh,S.S., Plakos,K., Lou,X., Xiao,Y. and Soh,H.T. (2010) In vitro selection of structure-switching, self-reporting aptamers. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 14053.
4. Ramezani,H. and Dietz,H. (2020) Building machines with DNA molecules. *Nat. Rev. Genet.*, **21**, 5–26.
5. Kang,B., Park,S.V., Soh,H.T. and Oh,S.S. (2019) A dual-sensing DNA nanostructure with an ultrabroad detection range. *ACS Sens.*, **4**, 2802–2808.
6. Yoo,H., Jo,H. and Oh,S.S. (2020) Detection and beyond: challenges and advances in aptamer-based biosensors. *Mater. Adv.*, **1**, 2663–2687.
7. Bloomfield,V. and Crothers,D.M. (2000) In: *Nucleic Acids: Structures, Properties and Functions*. University Science Books, Melville, California.
8. Marmur,J. and Doty,P. (1962) Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.*, **5**, 109–118.
9. Wartell,R.M. and Benight,A.S. (1985) Thermal denaturation of DNA molecules: a comparison of theory with experiment. *Phys. Rep.*, **126**, 67–107.

10. John SantaLucia,J. and Hicks,D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–440.
11. Yakovchuk,P., Protozanova,E. and Frank-Kamenetskii,M.D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic. Acids. Res.*, **34**, 564–574.
12. Frank-Kamenetskii,M.D. and Mirkin,S.M. (1995) Triplex DNA structures. *Annu. Rev. Biochem*, **64**, 65–95.
13. Vorlíčková,M., Chládková,J., Kejnovská,I., Fialová,M. and Kypr,J. (2005) Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic. Acids. Res.*, **33**, 5851–5860.
14. Idili,A., Vallée-Bélisle,A. and Ricci,F. (2014) Programmable pH-triggered DNA nanoswitches. *J. Am. Chem. Soc.*, **136**, 5836–5839.
15. Gehring,K., Leroy,J.-L. and Guéron,M. (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature*, **363**, 561–565.
16. You,J., Li,H., Lu,X.-M., Li,W., Wang,P.-Y., Dou,S.-X. and Xi,X.-G. (2017) Effects of monovalent cations on folding kinetics of G-quadruplexes. *Biosci. Rep.*, **37**, BSR20170771
17. Sigel,A., Sigel,H. and Sigel,R.K. (2016) In: *The Alkali Metal Ions: Their Role for Life*. Springer International Publishing, NY.
18. Šponer,J. and Hobza,P. (2003) Molecular interactions of nucleic acid bases. A review of quantum-chemical studies. *Collect. Czech. Chem. Commun.*, **68**. 2331–2282.
19. Cheng,C., Zhang,M. and Sheng,L. (2013) *Ab initio* and DFT study of non-covalent interactions between rare gas atoms and aromatic rings. *J. Theor. Comput. Chem.*, **12**, 1350012.
20. Raju,R.K., Bloom,J.W.G., An,Y. and Wheeler,S.E. (2011) Substituent effects on non-covalent interactions with aromatic rings: insights from computational chemistry. *Chem. Phys. Chem.*, **12**, 3116–3130.
21. Hobza,P. and Šponer,J. (2002) Toward True DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations. *J. Am. Chem. Soc.*, **124**, 11802–11808.
22. McDonald,A.R., Denning,E.J. and MacKerell,A.D. (2013) Impact of geometry optimization on base–base stacking interaction energies in the canonical A- and B-forms of DNA. *J. Phys. Chem. A.*, **117**, 1560–1568.
23. Swart,M., van der Wijst,T., Fonseca Guerra,C. and Bickelhaupt,F.M. (2007) π-π stacking tackled with density functional theory. *J. Mol. Model.*, **13**, 1245–1257.
24. Cooper,V.R., Thonhauser,T., Puzder,A., Schröder,E., Lundqvist,B.I. and Langreth,D.C. (2008) Stacking Interactions and the twist of DNA. *J. Am. Chem. Soc.*, **130**, 1304–1308.
25. Barone,G., Fonseca Guerra,C. and Bickelhaupt,F.M. (2013) B-DNA structure and stability as function of nucleic acid composition: dispersion-corrected DFT study of dinucleoside monophosphate single and double strands. *ChemistryOpen*, **2**, 186–193.
26. Pacureanu,L. and Simon,Z. (2010) DFT plus PCM calculation for pairing specificity of Watson–Crick-type bases in aqueous solutions. *Int. J. Quantum Chem*, **110**, 1295–1305.
27. Šponer,J., Jurečka,P., Marchan,I., Luque,F.J., Orozco,M. and Hobza,P. (2006) Nature of Base stacking: reference quantum-hemical stacking energies in ten unique B-DNA base-pair steps. *Chem. Eur. J.*, **12**, 2854–2865.
28. Fonseca Guerra,C., van der Wijst,T., Poater,J., Swart,M. and Bickelhaupt,F.M. (2010) Adenine versus guanine quartets in aqueous solution: dispersion-corrected DFT study on the differences in π-stacking and hydrogen-bonding behavior. *Theor. Chem. Acc.*, **125**, 245–252.
29. Wood,B.R. (2016) The importance of hydration and DNA conformation in interpreting infrared spectra of cells and tissues. *Chem. Soc. Rev.*, **45**, 1980–1998.
30. Boys,S.F. and Bernardi,F. (1970) The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.*, **19**, 553–566.
31. Klamt,A. and Schüürmann,G. (1993) COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Transactions*, **2**, 799–805.
32. Hays,F.A., Teegarden,A., Jones,Z.J.R., Harms,M., Raup,D., Watson,J., Cavaliere,E. and Ho,P.S. (2005) How sequence defines structure: a crystallographic map of DNA structure and conformation. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 7157–7162.
33. Aramini,J.M., Germann,M.W. and Mujeeb,A. (1998) NMR solution structures of [d(GCGAAT-3′-3′-αT-5′-5′-CGC)$_2$] and its unmodified control. *Nucleic Acids Res.*, **26**, 5644–5654.
34. Leporc,S., Mauffret,O., Tevanian,G., Lescot,E., Monnot,M. and Fermandjian,S. (1999) An NMR and molecular modelling analysis of d(CTACTGCTTTAG)·d(CTAAAGCAGTAG) reveals that the particular behaviour of TpA steps is related to edge-to-edge contacts of their base-pairs in the major groove. *Nucleic. Acids. Res.*, **27**, 4759–4767.
35. Cheng,J.-W., Chou,S.-H., Salazar,M. and Reid,B.R. (1992) Solution structure of [d(GCGTATACGC)]$_2$. *J. Mol. Biol.*, **228**, 118–137.
36. Radhakrishnan,I. and Patel,D.J. (1994) Solution structure of a pyrimidine·purine·pyrimidine DNA triplex containing T·AT, C^+·GC and G·TA triples. *Structure*, **2**, 17–32.
37. Asensio,J.L., Brown,T. and Lane,A.N. (1999) Solution conformation of a parallel DNA triple helix with 5′ and 3′ triplex–duplex junctions. *Structure*, **7**, 1–11.
38. Tarköy,M., Phipps,A.K., Schultze,P. and Feigon,J. (1998) Solution Structure of an Intramolecular DNA triplex linked by Hexakis(ethylene glycol) units: d(AGAGAGAA-(EG)$_6$-TTCTCTCT-(EG)$_6$-TCTCTCTT). *Biochemistry*, **37**, 5810–5819.
39. Horvath,M.P. and Schultz,S.C. (2001) DNA G-quartets in a 1.86 Å resolution structure of an Oxytricha nova telomeric protein-DNA complex11Edited by I. Tinoco. *J. Mol. Biol.*, **310**, 367–377.
40. Zhang,N., Phan,A.T. and Patel,D.J. (2005) (3 + 1) Assembly of Three Human Telomeric Repeats into an Asymmetric Dimeric G-Quadruplex. *J. Am. Chem. Soc.*, **127**, 17277–17285.
41. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat d[AG$_3$(T$_2$AG$_3$)$_3$] G-tetraplex. *Structure*, **1**, 263–282.
42. Wang,Y. and Patel,D.J. (1994) Solution structure of the Tetrahymena telomeric repeat d(T$_2$G$_4$)$_4$ G-tetraplex. *Structure*, **2**, 1141–1156.
43. Wang,Y. and Patel,D.J. (1995) Solution structure of the Oxytricha telomeric repeat d[G$_4$(T$_4$G$_4$)$_3$] G-tetraplex. *J. Mol. Biol.*, **251**, 76–94.
44. Smith,F.W., Schultze,P. and Feigon,J. (1995) Solution structures of unimolecular quadruplexes formed by oligonucleotides containing *Oxytricha* telomere repeats. *Structure*, **3**, 997–1008.
45. Phillips,K., Dauter,Z., Murchie,A.I.H., Lilley,D.M.J. and Luisi,B. (1997) The crystal structure of a parallel-stranded guanine tetraplex at 0.95Å resolution. *J. Mol. Biol.*, **273**, 171–182.
46. Haider,S., Parkinson,G.N. and Neidle,S. (2002) Crystal structure of the potassium form of an Oxytricha nova G-quadruplex. *J. Mol. Biol*, **320**, 189–200.
47. Parkinson,G.N., Lee,M.P.H. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.
48. Luu,K.N., Phan,A.T., Kuryavyi,V., Lacroix,L. and Patel,D.J. (2006) Structure of the human telomere in K^+ solution: An intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.*, **128**, 9963–9970.
49. Dai,J., Carver,M., Punchihewa,C., Jones,R.A. and Yang,D. (2007) Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.*, **35**, 4927–4940.
50. Zhao,Y., Schultz,N.E. and Truhlar,D.G. (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory. Comput.*, **2**, 364–382.
51. Jissy,A.K. and Datta,A. (2014) Design and applications of noncanonical DNA base pairs. *J. Phys. Chem. Lett.*, **5**, 154–166.
52. Shukla,M.K., Dubey,M., Zakar,E., Namburu,R., Czyznikowska,Z. and Leszczynski,J. (2009) Interaction of nucleic acid bases with single-walled carbon nanotube. *Chem. Phys. Lett.*, **480**, 269–272.
53. Gu,J., Wang,J. and Leszczynski,J. (2011) Stacking and H-bonding patterns of dGpdC and dGpdCpdG: performance of the M05-2X and M06-2X Minnesota density functionals for the single strand DNA. *Chem. Phys. Lett.*, **512**, 108–112.
54. Kabelác,M., Valdes,H., Sherer,E.C., Cramer,C.J. and Hobza,P. (2007) Benchmark RI-MP2 database of nucleic acid base trimers: performance of different density functional models for prediction of structures and binding energies. *Phys. Chem. Chem. Phys.*, **9**, 5000–5008.

55. Petersson,G.A., Bennett,A., Tensfeldt,T.G., Al-Laham,M.A., Shirley,W.A. and Mantzaris,J. (1988) A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. *J. Chem. Phys.*, **89**, 2193–2218.
56. Petersson,G.A. and Al-Laham,M.A. (1991) A complete basis set model chemistry. II. Open-shell systems and the total energies of the first-row atoms. *J. Chem. Phys.*, **94**, 6081–6090.
57. Glendening,E.D. and Streitwieser,A. (1994) Natural energy decomposition analysis: An energy partitioning procedure for molecular interactions with application to weak hydrogen bonding, strong ionic, and moderate donor–acceptor interactions. *J. Chem. Phys.*, **100**, 2900–2909.
58. Amovilli,C., Barone,V., Cammi,R., Cancès,E., Cossi,M., Mennucci,B., Pomelli,C.S. and Tomasi,J. (1998), Recent advances in the description of solvent effects with the polarizable continuum model. In: Löwdin,P-.O. (ed). *Adv. Quantum. Chem*. Academic Press, Vol. **32**, pp. 227–261.
59. Jissy,A.K. and Datta,A. (2010) Designing molecular switches based on DNA-base mispairing. *J. Phys. Chem. B*, **114**, 15311–15318.
60. Werntges,H., Steger,G., Riesner,D. and Fritz,H.-J. (1986) Mismatches in DNA double strands: thermodynamic parameters and their correlation to repair efficiencies. *Nucleic. Acids. Res.*, **14**, 3773–3790.
61. Purvis,G.D. and Bartlett,R.J. (1982) A full coupled-cluster singles and doubles model: the inclusion of disconnected triples. *J. Chem. Phys.*, **76**, 1910–1918.
62. Scuseria,G.E., Janssen,C.L. and Schaefer,H.F. (1988) An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations. *J. Chem. Phys.*, **89**, 7382–7387.
63. Scuseria,G.E. and Schaefer,H.F. (1989) Is coupled cluster singles and doubles (CCSD) more computationally intensive than quadratic configuration interaction (QCISD)? *J. Chem. Phys.*, **90**, 3700–3703.
64. Modrich,P. (1987) DNA, mismatch correction. *Annu. Rev. Biochem*, **56**, 435–466.
65. Megger,D.A., Lax,P.M., Paauwe,J., Fonseca Guerra,C. and Lippert,B. (2018) Mixed guanine, adenine base quartets: possible roles of protons and metal ions in their stabilization. *J. Biol. Inorg. Chem.*, **23**, 41–49.
66. LaFramboise,T. (2009) Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic. Acids. Res.*, **37**, 4181–4193.
67. Rathnam,C., Chueng,S.D., Yang,L. and Lee,K.B. (2017) Advanced Gene Manipulation Methods for Stem Cell Theranostics. *Theranostics*, **7**, 2775–2793.
68. SantaLucia,J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 1460–1465.
69. Akbari,E., Mollica,M.Y., Lucas,C.R., Bushman,S.M., Patton,R.A., Shahhosseini,M., Song,J.W. and Castro,C.E. (2017) Engineering cell surface function with DNA origami. *Adv. Mater.*, **29**, 1703632.
70. Dietz,H., Douglas,S.M. and Shih,W.M. (2009) Folding DNA into Twisted and Curved Nanoscale Shapes. *Science*, **325**, 725–730.
71. Zadegan,R.M., Lindau,E.G., Klein,W.P., Green,C., Graugnard,E., Yurke,B., Kuang,W. and Hughes,W.L. (2017) Twisting of DNA Origami from Intercalators. *Sci. Rep.*, **7**, 7382.
72. Lavelle,C. (2008) DNA torsional stress propagates through chromatin fiber and participates in transcriptional regulation. *Nat. Struct. Mol. Biol.*, **15**, 123–125.
73. Reymer,A., Zakrzewska,K. and Lavery,R. (2018) Sequence-dependent response of DNA to torsional stress: a potential biological regulation mechanism. *Nucleic Acids Res.*, **46**, 1684–1694.
74. Beal,P. and Dervan,P. (1991) Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. *Science*, **251**, 1360–1363.
75. Goñi,J.R., de la Cruz,X. and Orozco,M. (2004) Triplex-forming oligonucleotide target sequences in the human genome. *Nucleic Acids Res.*, **32**, 354–360.
76. Aviñó,A., Cubero,E., González,C., Eritja,R. and Orozco,M. (2003) Antiparallel triple helices. Structural characteristics and stabilization by 8-amino derivatives. *J. Am. Chem. Soc.*, **125**, 16127–16138.
77. Van Meervelt,L., Vlieghe,D., Dautant,A., Gallois,B., Précigoux,G. and Kennard,O. (1995) High-resolution structure of a DNA helix forming (C.G)*G base triplets. *Nature*, **374**, 742–744.
78. Sponer,J., Burda,J.V., Mejzlík,P., Leszczynski,J. and Hobza,P. (1997) Hydrogen-bonded trimers of DNA bases and their interaction with metal cations: *ab initio* quantum-chemical and empirical potential study. *J. Biomol. Struct. Dyn.*, **14**, 613–628.
79. Soto,A.M., Loo,J. and Marky,L.A. (2002) Energetic contributions for the formation of TAT/TAT, TAT/CGC^+, and CGC^+/CGC^+ base triplet stacks. *J. Am. Chem. Soc.*, **124**, 14355–14363.
80. Dagneaux,C., Gousset,H., Shchyolkina,A.K., Ouali,M., Letellier,R., Liquier,J., Florentiev,V.L. and Taillandier,E. (1996) Parallel and Antiparallel A*A-T Intramolecular Triple Helices. *Nucleic. Acids. Res.*, **24**, 4506–4512.
81. Ghorbani,M. and Mohammad-Rafiee,F. (2010) Geometrical correlations in the nucleosomal DNA conformation and the role of the covalent bonds rigidity. *Nucleic Acids Res.*, **39**, 1220–1230.
82. Brown,P.M., Madden,C.A. and Fox,K.R. (1998) Triple-helix formation at different positions on nucleosomal DNA. *Biochemistry*, **37**, 16139–16151.
83. Chandler,S.P. and Fox,K.R. (1996) Specificity of Antiparallel DNA Triple Helix Formation. *Biochemistry*, **35**, 15038–15048.
84. Keppler,M.D., Neidle,S. and Fox,K.R. (2001) Stabilisation of TG- and AG-containing antiparallel DNA triplexes by triplex-binding ligands. *Nucleic Acids Res.*, **29**, 1935–1942.
85. Gu,J. and Leszczynski,J. (2000) A Remarkable alteration in the bonding pattern: an HF and DFT study of the interactions between the Metal Cations and the Hoogsteen Hydrogen-Bonded G-Tetrad. *J. Phys. Chem. A.*, **104**, 6308–6313.
86. Kankia,B.I. and Marky,L.A. (2001) Folding of the thrombin aptamer into a G-quadruplex with Sr^{2+}: stability, heat, and hydration. *J. Am. Chem. Soc.*, **123**, 10799–10804.
87. Włodarczyk,A., Grzybowski,P., Patkowski,A. and Dobek,A. (2005) Effect of Ions on the polymorphism, effective charge, and stability of human telomeric DNA. Photon correlation spectroscopy and circular dichroism studies. *J. Phys. Chem. B*, **109**, 3594–3605.
88. Miyoshi,D., Karimata,H. and Sugimoto,N. (2006) Hydration regulates thermodynamics of G-Quadruplex formation under molecular crowding conditions. *J. Am. Chem. Soc.*, **128**, 7957–7963.
89. Norio,O. and Kiyoshi,A. (1979) Neutron diffraction study of aqueous Ionic solutions. I. Aqueous solutions of lithium chloride and caesium chloride. *Bull. Chem. Soc. Jpn.*, **52**, 2755–2759.
90. Pálinkás,G., Radnai,T. and Hajdu,F. (1980) Ion-Solvent and Solvent-Solvent Interactions. X-ray Study of Aqueous Alkali Chloride Solutions. *Z. Naturforsch. A*, **35**, 107–114.
91. Mancinelli,R., Botti,A., Bruni,F., Ricci,M.A. and Soper,A.K. (2007) Hydration of sodium, potassium, and chloride ions in solution and the concept of structure Maker/Breaker. *J. Phys. Chem. B*, **111**, 13570–13577.
92. Azam,S.S., Hofer,T.S., Randolf,B.R. and Rode,B.M. (2009) Hydration of sodium(I) and potassium(I) revisited: a comparative QM/MM and QMCF MD simulation study of Weakly Hydrated Ions. *J. Phys. Chem. A.*, **113**, 1827–1834.
93. Soper,A.K. and Weckström,K. (2006) Ion solvation and water structure in potassium halide aqueous solutions. *Biophys. Chem.*, **124**, 180–191.
94. Hud,N.V., Smith,F.W., Anet,F.A.L. and Feigon,J. (1996) The Selectivity for K^+ versus Na^+ in DNA quadruplexes is dominated by relative free energies of hydration: A thermodynamic Analysis by 1H NMR. *Biochemistry*, **35**, 15383–15390.
95. Tan,Z.J. and Chen,S.J. (2006) Nucleic acid helix stability: effects of salt concentration, cation valence and size, and chain length. *Biophys. J.*, **90**, 1175–1190.
96. Zhou,W., Saran,R. and Liu,J. (2017) Metal sensing by DNA. *Chem. Rev.*, **117**, 8272–8325.
97. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic. Acids. Res.*, **34**, 5402–5415.
98. Dickerson,R.E., Drew,H.R., Conner,B.N., Wing,R.M., Fratini,A.V. and Kopka,M.L. (1982) The anatomy of A-, B-, and Z-DNA. *Science*, **216**, 475–485.
99. Siegfried,N.A., Kierzek,R. and Bevilacqua,P.C. (2010) Role of unsatisfied Hydrogen bond Acceptors in RNA energetics and specificity. *J. Am. Chem. Soc.*, **132**, 5342–5344.
100. Salandria,K.J., Arico,J.W., Calhoun,A.K. and McLaughlin,L.W. (2011) Stability of DNA containing a structural water mimic in an A-T rich sequence. *J. Am. Chem. Soc.*, **133**, 1766–1768.