PCCP

Accepted Manuscript

Nucleotides containing variously modified sugars: energetics, structure, and mechanical properties

Yevgen P. Yurenko,[1*] Jan Novotný,[1] Tymofii Y. Nikolaienko,[2*] and Radek Marek[1,3,4*]

[1] CEITEC – Central European Institute of Technology, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic

[2] Molecular Physics Department, Faculty of Physics, Taras Shevchenko National University of Kyiv, 4 Hlushkova ave., 03022 Kyiv, Ukraine

[3] National Center for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic

[4] Department of Chemistry, Faculty of Science, Masaryk University, Kamenice 5, CZ-62500 Brno, Czech Republic

* Corresponding Authors

Yevgen P. Yurenko: <u>yevgen.yurenko@ceitec.muni.cz</u>, Tymofii Y. Nikolaienko: <u>tim_mail@ukr.net</u>, Radek Marek: <u>rmarek@chemi.muni.cz</u>

Graphical abstract
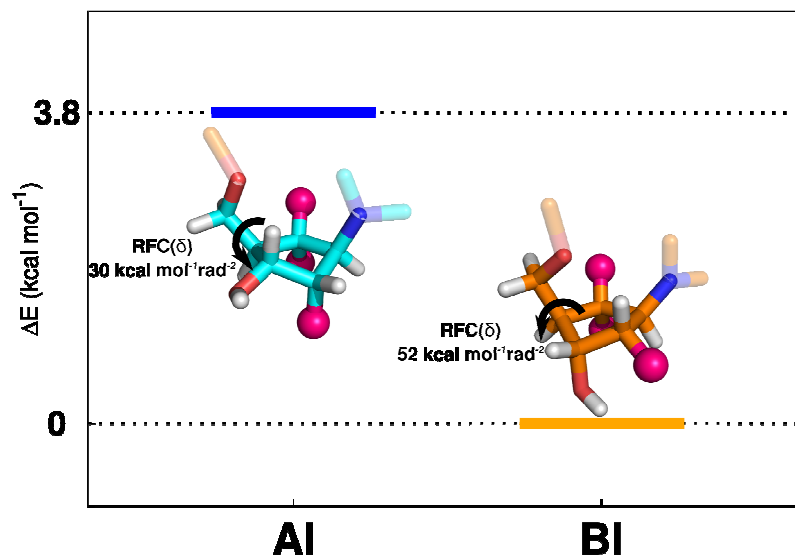
Conformational flexibility and energetic stability of nucleotides are tightly interconnected: more energetically stable conformers are characterized by higher values of relaxed force constants (RFC) for the δ torsion angle.

Abstract

The influence of various sugar-residue modifications on intrinsic energetic, conformational, and mechanical properties of 2′-deoxyribonucleotide-5′-monophosphates (dNs) was comprehensively investigated using modern quantum chemical approaches. In total, fourteen sugar modifications, including double bonds and heteroatoms (S, N) inside the sugar ring, as well as fluorination in various positions, were analyzed. Among hundreds of possible conformational states of dNs, only two - AI and BI, corresponding to the most biologically significant forms of a double-helical DNA, were considered for each dN. It was established that the most of the studied modifications tend to strongly stabilize either AI or BI conformation of dNs both in the gas phase and in aqueous solution (modeled by implicit solvent models). Therefore, some of these modifications can be used as a tool for reducing structural polymorphism of nucleic acids in solution as well as for designing oligonucleotides with specific structural features. The evaluation of relaxed force constants (RFC) for glycosidic bonds suggests that the majority of the studied modifications of the sugar residue yield increased strengths of glycosidic bonds in dNs, and can therefore be used for designing modified nucleic acids with an increased resistance to abasic lesions. The most significant reinforcement of the glycosidic bond occurs in dNs containing the CF_2 group instead of the O4' oxygen and the fluorine atom at the 2'-α-position. The calculation of the RFC and vibrational root-mean-square (VRMS) deviations for conformational degrees of freedom revealed a strong dependence between mechanical properties of dNs and their energetic characteristics. In particular, electronic energies of AI and BI conformers of dNs calculated *in vacuo* are closely connected with the values of relaxed force constants (RFC) for the δ angle: the higher RFC (δ) values correspond to more energetically favorable conformers.

1 INTRODUCTION

Nucleic acids are key molecules in biology playing a variety of biological roles in coding, decoding, regulation, and expression of genes.[1] The diversity of biological functions performed by nucleic acids requires a high degree of their structural adaptability. In particular, the DNA molecule is able to adopt different double-helical (A, B, Z) forms as well as various triple-stranded and quadruhelical structural variants.[2,3] Due to this fact, DNA is commonly believed to be a flexible molecule. Its flexibility is apparent, for instance, in sequence-dependent structural variations,[4] the wrapping around the histone core,[5] the formation of tri- and tetra-loop hairpins,[6] bending and kinking of DNA upon binding to a variety of proteins,[7] and DNA deformation upon drug binding.[8]

The DNA conformational heterogeneity and flexibility are directly connected to its mechanical properties, i.e. ability to be bent or stretched upon external deformations. The investigation of the DNA mechanical properties and deformability can trail a way to understanding its compaction in the limited space of cells and viruses.[9] The DNA compaction process was first associated with some disruption of the base-pair stacking which gives rise to formation of free hinges or 'kinks'.[10] However, later it was suggested[11] that strong bends in the double-helical DNA may occur without de-stacking *via* low-energy fluctuations of the bond and torsion angles. In recent decades, several models describing the DNA mechanical properties on the large scale and their dependence on the molecule length and base sequence were developed (see, e.g. refs. 12, 13 and refs. therein). These models provided useful insights into our understanding of the DNA flexibility and characterized the main bending and stretching modes. However, the prediction of mechanical properties of the whole DNA molecule on different length scales has not been achieved because its mechanics is governed by various entropic macroscopic factors. To gain a firmer understanding of the DNA structural adaptability and strengthen the predictive potential of its behavior on the larger scale, the mechanical properties of its building blocks – nucleotides should be studied.

The first step toward understanding the DNA mechanical properties at the intermediate scale, i.e. at the level of nucleic acid constituents – *nucleosides*, was done in the work by Nikolaienko et al.[14] This study was based on the relaxed force constants (RFC) (somewhat similar to adiabatic force constants[15]) and vibrational-root-mean-square (VRMS) formalisms for nucleic acids, which were developed for the first time in the preceding work.[16] In contrast to traditional or rigid force constants, the RFC are characterized by a unique value corresponding to each structural parameter (bond length, valence angle, and dihedral angle). Moreover, the RFC are applicable to such 'collective' conformational parameters as sugar pseudorotation angle[17] determined by positions of more than four atoms. The VMRS deviation is another useful characteristic that describes fluctuation 'noise' of conformational parameters caused by

molecules' nuclei thermal and zero-point vibrations. It was shown[14] that in canonical nucleosides the β torsional angle and the sugar pseudorotational phase P are rather soft conformational parameters characterized by smaller values of RFC whereas the γ angle is rather rigid. The investigation of VRMS deviations[16] provides evidence that quantum zero-point motion of nuclei makes a perceptible contribution to fluctuation of torsion angles in nucleosides due to the presence of the low-frequency vibrations. At temperature of 298 K, the magnitude of these fluctuations becomes comparable with the half-width of classical *gauche*$^+$, *gauche*$^-$, and *trans* conformational sectors (30°) and can increase even more at higher temperatures. However, the presence of phosphate group in DNA can have an important influence on the rigidity and deviation of individual torsion angles.[18] Therefore, for predicting and understanding some DNA properties on the macro scale, the monomeric units with a phosphate group – *nucleotides*, should be studied.

In recent decades a great deal of effort has been devoted to development and investigation of chemically modified nucleosides and nucleotides that could be used as alternative nucleic acid building blocks.[19,20] These modified biomolecules can serve as efficient tools for studying the evolution and function of living organisms. For instance, Kool and co-workers have developed size-expanded DNA (including 'extended' DNA and 'wide' DNA abbreviated and xDNA and yDNA, respectively) by adding a benzene ring to the canonical nucleobases (A, T, C, G) in 2'-deoxyribonucleosides.[21-23] Natural DNA, which is 20 Å wide, and benzene, with a girth of 2.4 Å, produced a wholly new wider double helix.[21] It was proposed[24] that the xDNA and yDNA systems may have most of features of natural DNA, which suggests the future possibility of a functioning, replicable genetic system using size-expanded DNA as the genetic material. Computational studies on xDNA and yDNA nucleotides/nucleosides provided useful information about their hydrogen-bonding and stacking abilities[25] and energetics of rotation about the glycosidic bond,[26] explained the origins of their optical absorption spectra,[27] analyzed the local aromaticity of extended bases,[28] and cast some light on possibility of further extension of purine and pyrimidine rings.[29,30] In addition to xDNA and yDNA, other alternative genetic platforms including, among others, locked nucleic acids (LNA),[31] peptide nucleic acids (PNA),[32] and threose nucleic acids (TNA)[33] were proposed and investigated using both experimental[34-36] and computational[37-45] approaches. In addition to their role in investigating the origins of life, modified nucleotides and oligonucleotides have various applications in bio- and nanotechnology,[46] molecular medicine,[20] and diagnostics.[47] In particular, fluorinated nucleotides have been used in designing the nucleic acid aptamers,[48] antisense agents,[49] anti-micro RNA oligomers[50] as well as powerful tools for modulating the conformational preferences of nucleic acids.[51] Moreover, 2'-fluoro-substituted nucleic acids are known to induce RNA interference[52] and therefore, can be used to control the expression of fundamentally interesting or therapeutically relevant genes.

In this work we apply modern methods of quantum chemistry to analyze intrinsic structural, energetic, and mechanical properties of the DNA building blocks - 2'-deoxyribonucleotide-5'-monophosphates (dNs). In addition to canonical nucleotides, we performed a comprehensive investigation of dNs units with fourteen types of the sugar residue modifications, including double bonds and heteroatoms (N and S) within the deoxyribose ring, extra bridge connecting O2' and C4´ atoms that 'locks' the ribose in the North (C3'-endo) conformation (so-called locked nucleic acids or LNA),[53] and fluorination in various positions (e.g., replacement of O4' atom by CF_2 group). These sugar modifications were shown to be useful, e.g., in design and investigation of drugs based on nucleoside and nucleotide analogues,[54-57] nucleic acid aptamers,[58] and artificial DNA structures with improved emergent properties[59-63] such as an ability to adopt some particular conformation, reduced structural polymorphism, higher stability coming from non-covalent interactions. We aimed to answer the following questions: i) To what extent the flexibility of individual conformational degrees of freedom determines the mechanical properties of canonical dNs and their energetic preferences; ii) How various modifications within the deoxyribose ring affect the strengths of glycosidic bonds and the rigidity of the main conformational parameters. Since dNs are able to adopt multiple (hundreds) conformations,[64-67] we analyzed only two conformers of each nucleotide - AI and BI corresponding to well-known forms of double-helical DNA.[1] The A and B forms are considered to be the most biologically significant conformations of the double-helical DNA. Whereas the B form prevails under physiological conditions[68] and participates in most of biochemical processes *in vivo* due to its extensive structural heterogeneity,[69] the A-DNA is known to be more conformationally homogenous,[70] which points to its role in preserving the genetic material. For instance, in active sites of DNA polymerases DNA is known to adopt A form to improve accuracy of DNA biosynthesis.[71]

The results of this work are expected to contribute to understanding the intrinsic energetic, conformational, and mechanical features of DNA building blocks and explanation of properties of modified nucleic acid structures containing these blocks.

2 COMPUTATIONAL METHODOLOGY

2.1 Studied systems. The molecular models investigated in this work were 2´-deoxyribonucleotide-5´-monophosphates (dNs, Fig. 1). The negative net charges on the phosphate groups were neutralized by Na^+ ions. The hydrogen atom in the terminal 5´-OH hydroxyl group was replaced by the methyl moiety. The base units in dNs (Fig. 1) were adenine (A), thymine (T), guanine (G), and cytosine (C). Since dNs molecules are characterized by extremely high conformational flexibility resulting in hundreds of local minima on their potential energy surfaces,[64-67] we restricted conformational search to two conformers – AI and BI corresponding to the most biologically significant structural variants of a double-helical DNA.

In AI- and BI-form-like conformers of dNs, the conformational angles should fall into ranges which are typical for AI and BI forms of DNA. In particular, AI conformers should satisfy the following requirements: P [17] (Fig. S1 in ESI) $\in$ C3′-endo, $\alpha \in$ gauche⁻, $\beta \in$ trans, $\gamma \in$ gauche⁺, $\varepsilon \in$ trans, $\chi \in$ anti, whereas the corresponding ranges for BI-form DNA-like conformers were: $P \in$ C2′-endo, $\alpha \in$ gauche⁻, $\beta \in$ trans, $\gamma \in$ gauche⁺, $\varepsilon \in$ trans, $\chi \in$ anti. For definition of torsion angles (α, β, γ, ε, and χ), see Fig. 1.
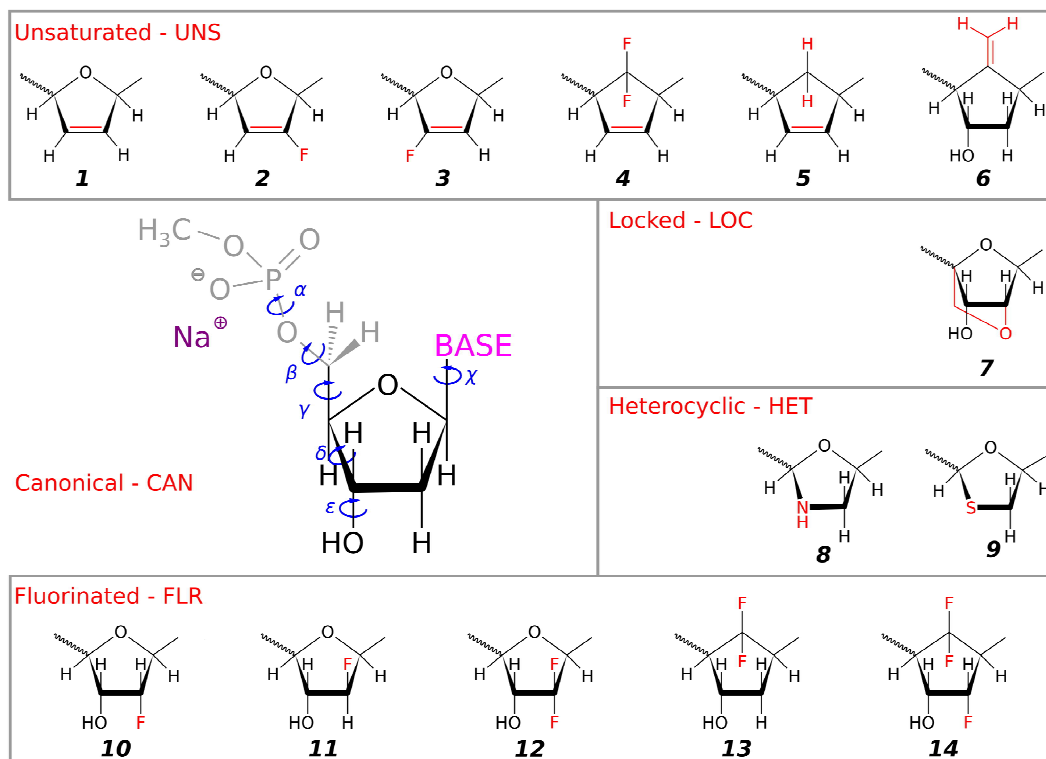
Fig. 1. The deoxyribose modifications studied in this work. The modifications are divided into the following groups: Unsaturated – UNS (modifications from **1** to **6**), Locked – LOC (**7**), Heterocyclic – HET (**8** and **9**), Fluorinated – FLR (from **10** to **14**). The canonical (natural) nucleotides – CAN were also analyzed and served as reference systems to investigate the effects of each modification. The designation of torsion angles correspond to a standard nucleic acid nomenclature.[1]

Besides canonical (Can) nucleotides, dNs incorporating fourteen types of the sugar modifications were investigated. These modifications were classified according to their chemical nature (Fig. 1): unsaturated (UNS) dNs containing a double bond inside the sugar ring (modifications **1-5**) or an exocyclic double bond (modification **6**), locked (LOC),[53] heterocyclic (HET), and fluorinated (FLR) dNs. Overall, the starting structures for geometry optimizations were formed taking into account all possible

combinations of the fifteen types of the sugar residue (Fig. 1) and four nucleobases (A = adenine, G = guanine, T = thymine and C = cytosine) as well as two conformational states (AI and BI).

2.2 Calculation methods. A conformational analysis of dNs was performed by density functional theory (DFT) calculations using the B3LYP[72,73] nonlocal hybrid exchange-correlation functional in combination with the recent D3 empirical dispersion correction[74] and the 6-311++G(d,p) basis set. In the course of geometry optimizations the *tight* convergence criteria and *ultrafine* integration grid options were applied. Vibrational spectra have been calculated at the same level of theory using standard harmonic approximation. The absence of imaginary frequencies in each case confirmed that all optimized structures correspond to local minima on potential energy surface (PES). The starting geometries of canonical dNs and some fluorinated dNs (modifications **10**, **11**, **12** and **13**) were taken from our previous work[75] and re-optimized using the above-mentioned options. The geometry optimization was performed *in vacuo* since we are focusing mainly on *intrinsic* conformational and energetic properties of dNs. The knowledge of intrinsic features of nucleic acid building blocks serves as an important step towards understanding their behavior in different environments (from aqueous solution to active sites of enzymes which are characterized by low dielectric response[76]). All the calculations have been done by means of the Gaussian 09 suite of programs.[77] For dNs which are able to adopt both AI and BI conformers, the relative electronic (ΔE) and vibration-corrected energies (ΔG, denoted as Gibbs energies in Gaussian 09 package) of these conformers were calculated. The ΔG values were sums of electronic energies and the zero-point energies, thermal corrections, and entropy contributions calculated at the B3LYP-D3/6-311++G(d,p) level of theory.

To take into account a portion of possible environmental effects, the relative electronic energies (ΔE) of AI and BI conformers were also estimated in water by means of Conductor-like Polarizable Continuum model (CPCM)[78] as implemented in the Gaussian 09 package.[77] For this purpose, CPCM DFT B3LYP/6-311++G(d,p) single-point calculations were performed for geometries optimized *in vacuo*. In earlier work by Bulavin, Hovorun, and Nikolaienko,[64] the conformational parameters of the most energetically favorable conformers of each of four canonical dNs (with A, T, C, and G) optimized both in the gas phase (at the B3LYP/6-31G(d,p) level of theory) and in aqueous solution (at the same level of theory using IEFPCM model for water with $\varepsilon = 78.4$ as implemented[79] in Gaussian 09 package) were compared. The comparison revealed no significant differences in values of individual conformational parameters in both environments (see Fig. S2 in ESI).

2.3 Estimation of relaxed force constants (RFCs) for glycosidic C-N bonds. The stability of glycosidic bond is one of the essential characteristics of modified DNA systems, determining their resistance to

abasic lesions,[80] i.e., formation of apurinic/apyrimidinic sites. The formation of abasic sites can significantly alter the DNA structure (e.g., changing the G-quadruplex topology[81]) or generate the DNA-strand breaks.[82] Grunenberg and co-workers suggested[83] that compliance constants can be considered as relevant descriptors of intrinsic strengths of covalent and non-covalent[84] bonds due to their independence on the masses of vibrating atoms and the coordinate selection. Moreover, compliance constants allow a direct determination of strengths of covalent and non-covalent bonds without any reference to other states which can be defined arbitrarily and therefore influence the results.[83] The compliance constants represent second derivatives of the potential energy due to an external force:

$$C_{ij} = \frac{\partial^2 E}{\partial f_i \partial f_j} \tag{1}$$

The glycosidic (C-N) bond lengths were used as internal coordinates for the calculation of C_{ij}. The compliance constants measure the displacement of an internal coordinate resulting from a unit force acting on it and can, therefore, be considered as descriptors of intrinsic bond strength. The calculation of compliance constants was performed in Compliance 3.0.2 program.[83,85] Relaxed force constants (RFC) for glycosidic bonds are inverse values of compliance constants calculated by equation (1). For convenience, in the further discussion we will use RFC (in N/m) instead of compliance constants C_{ij}.

2.4 RFCs and vibrational root-mean-square (VRMS) deviations for conformational parameters. For evaluation of flexibility of conformational parameters, we resort to methodology developed by Nikolaienko and co-workers.[16] This methodology was earlier applied to analyze mechanical properties of deoxyribose moiety and canonical nucleosides.[14] Although Compliance 3.0.2 program[83,85] is also able to calculate RFCs for angles by means of Grunenberg's method[83], this feature is considered by its developers as highly experimental and has not been properly tested yet. In contrast, the methodology of Nikolaienko *et al*[16] has been tested[14,16] for a representative set of DNA-related molecules. Moreover, this approach[14] makes possible the evaluation of RFCs not only for torsion angles, but also for more complex conformational parameters, such as phase pseudorotation angle P,[18] depending on coordinates of more than four atoms. The P angle is a crucial conformational characteristic used to distinguish between AI and BI conformations of dNs.

In general, RFC for a conformational parameter τ being a known function f($\mathbf{R_A}$, $\mathbf{R_B}$, $\mathbf{R_C}$, **...**) of nuclei Cartesian coordinates in defined as:[14,86]

$$K_\tau = \left(\frac{\partial^2 \Delta E}{\partial \tau^2} \right)_{\Delta E = \min} \tag{2}$$

where ΔE stands for energy difference $E(\mathbf{R_i})$-$E(\mathbf{R_i^0})$, with $\mathbf{R_i^0}$ representing nuclei equilibrium coordinates and $\mathbf{R_i}$ being such nuclei coordinates, that

$$f(\mathbf{R_i}) = \tau \tag{3}$$

The key idea behind the definition of RFCs in Eq. (2) is indicated by a subscript '$\Delta E = \min$'. To clarify its meaning, we note that Eq. (3) imposes only one constraint on 3N Cartesian coordinates of N atoms in the molecule, so that there remain 3N–1 independent scalar variables ('free parameters') q_i, $i = 1\ldots3N-1$, on which the energy depends in addition to being a function of τ. Thus, the energy E (or its increment ΔE) itself is, generally speaking, not a single-valued function of τ unless some additional condition controls the way in which the parameters q_i change 'in response to' the change of τ. The condition suggested by the definition of RFCs is such that for any given value of τ the parameters q_i should be chosen in a way that allows to minimize (or 'relax') the energy, and this is the condition described by the subscript '$\Delta E = \min$' in Eq. (2). Under this condition the energy becomes a single-valued function of τ and its first, second, and higher-order derivatives become well-defined quantities. It can be easily deduced from this description that the force constant obtained in the suggested way is the lowest possible force constant compared to arbitrary force constants obtained with any other additional constraint (i.e., with the values of q_i others than defined above). For more detailed description of the methodology and its implementation, see previous studies.[14,16]

Whereas RFCs provide information about 'static' properties of molecules, vibrational root-mean-square (VRMS) deviations of conformational parameters allow for understanding some 'dynamical' features. It is known that even at the temperature of 0K the nuclei of molecules are involved in quantum zero-point motions. Therefore, values of conformational parameters are essentially non-stationary. At higher temperatures the motion of nuclei gets even greater. The VRMS deviations for conformational parameters have been defined as $\sigma^\tau = \sqrt{\langle \tau^2 \rangle - \langle \tau \rangle^2}$. For methodology of calculation of VRMS deviations, see refs. 14 and 87.

3 RESULTS AND DISCUSSION

3.1 Energetic preferences and structural features of modified 2′-deoxyribonucleotide-5′-monophosphates (dNs). The study of energetic preferences of dNs is essential for understanding the effects of the sugar modifications. Moreover, the knowledge of intrinsic energetic properties of dNs as building blocks of the DNA molecule is of paramount importance for a rational design of novel nucleic acid structures containing modified units. These artificial DNA structures can have beneficial potential

applications in various areas ranging from analytical chemistry to molecular therapeutics.[54-63] To characterize intrinsic energetic properties of dNs we used two main characteristics – relative electronic (ΔE) and vibration-corrected energies (ΔG, T = 298 K, hereafter called Gibbs energies according to specification in the Gaussian package) of AI and BI conformers obtained for the structures optimized *in vacuo*. For evaluation of a possible impact of the aqueous environment on equilibrium between AI and BI conformational states, the relative electronic energies of these conformers, ΔE_{CPCM}, were estimated (without geometry re-optimization, *vide infra*) by using the CPCM[78] model for water. Although we are aware of the limitations of continuum solvent models for analyzing biomolecules in solution,[88] this approach can serve as an initial step towards understanding a possible influence of the environment on intrinsic energetic properties of dNs. It should be noted, however, that energetic analysis was applied only to the systems which are able to adopt both AI and BI conformers. Therefore, in case of modifications from **1** to **5** (Fig. 1, Table S1) the relative energies were not estimated since the presence of a double bond inside the deoxyribose ring makes it almost planar and thus, it is not possible to distinguish between AI and BI conformers for these types of the sugar modification.

The first studied group of deoxyribose modifications, *UNS* (*unsaturated*, modifications from **1** to **6**, Fig. 1, Table S1) consists of dNs containing either the endocyclic C2'=C3' double bond inside the deoxyribose residue (modifications **1-5**, Fig. 1) or an exocyclic double bond at the position C4' (unsaturated dNs with the sugar modification **6** which can be abbreviated as *UNS*-**6**-dNs, Fig. 1). Nucleosides with the sugar residue **1** are known to possess potent antiviral and antitumoral properties[89] and have been studied previously by various theoretical methods.[55,56,90,91] The modifications **2-4** differ from **1** by the presence of exocyclic fluorine atoms at the positions 2' (*UNS*-**2**-dNs), 3' (*UNS*-**3**-dNs), or 4' (*UNS*-**4**-dNs where the O4' atom is replaced by CF_2 group[75]). In contrast, in *UNS*-**5**-dNs the O4' atom is substituted by CH_2 group. The conformational analysis of dNs with the sugar residue **1-5** leads to the following conclusions: 1) The modifications **2** and **3** have a marginal effect on the conformational parameters (Fig. 1, Table S1) relative to *UNS*-**1**-dNs. However, the substitution of a carbon atom for the O4' oxygen in *UNS*-**4**-**5**-dNs leads to a simultaneous rise of the β angle which is getting closer to 180º and the glycosidic torsion angle χ that approaches *high-anti* conformational region (~ -90º).[1] In the crystal structures of the BI DNA the mean value of the β angle, 176º,[92] is similar to the values in *UNS*-**4**-**5**-dNs). The sugar ring in *UNS*-**1**-**5**-dNs is almost totally planar, consequently, the pseudorotation phase angle *P* (Fig. S2) can, in principle, take any value and is not considered as a distinct conformational parameter.

As regards *UNS*-**6**-dNs with an exocyclic double bond, they are able, like canonical (*CAN*) nucleotides, to adopt both AI and BI conformations (Table S1). From the energetic point of view, the modification **6** results in an additional stabilization of BI conformers relative AI ones (Table S1). This can be seen from the comparison of electronic energies and Gibbs energies *in vacuo* (ΔE_{VAC} and ΔG) of AI

and BI conformers and electronic energies in aqueous solution (ΔE_{CPCM}) for *UNS*-**6**-dNs and *CAN*-dNs**.** In case of *UNS*-**6**-dNs the BI conformers are more stable than AI structures both *in vacuo* and water environment, thus, the inclusion of an implicit solvation model does not reverse the relative stability of these two conformational states contrary to *CAN*-dNs (Table S1).

The next group of modifications, *LOC* (*locked*), contains only a single type of the sugar residue, **7** (Fig. 1), occurring in locked nucleic acids (LNA).[53] Due to the extra bridge connecting the O2' and C4´ atoms that 'locks' the ribose in the North (C3'-endo) conformation, **7**-dNs are not capable of adopting BI conformations. In this regard, only AI conformers of **7**-dNs were analyzed (Table S1). The χ and δ angles are more uniformly distributed in *LOC*-**7**-dNs compared to *CAN*-dNs, which is not surprising, taking into account structural constraints resulting from the presence of an extra bridge in the sugar fragment. Overall, there is a nice structural compatibility between *LOC*-**7**-dN and AI conformers of *CAN*-dNs.

The third group, *HET* (*heterocyclic*), includes two types of the sugar residue, **8** and **9**, containing at the 3'-position nitrogen and sulfur atoms, respectively (Fig. 1). The analysis of energetic characteristics of *HET*-**8**-dNs shows a surprisingly strong stabilization of AI relative to BI conformers both *in vacuo* and in aqueous medium (Figs. 2 and 3, Table S1). In particular, the respective values of ΔG (Fig. 2) and ΔE_{CPCM} (Fig. 3) for BI conformers are by 4.18÷5.69 kcal/mol and 4.10÷5.02 kcal/mol higher than for AI structures. On the contrary, for *HET*-**9**-dNs the energy gaps between AI and BI conformers are rather small (within 1 kcal/mol). Despite this fact, both ΔG and ΔE_{CPCM} manifest the same trend (for purine nucleotides dA-*HET*-**9**-dN and dG *HET*-**9**-dN the BI conformations are slightly preferred both *in vacuo* and in solution whereas the opposite situation is observed for pyrimidine nucleotides dC-*HET*-**9**-dN and dT-*HET*-**9**-dN). Regarding the structural features, the values of conformational angles are very similar in *HET*-**8**-dNs, *HET*-**9**-dNs, and *CAN*-dNs. However, analogously to compounds **1-5**, these modifications can be used only as 3'-terminal residues.
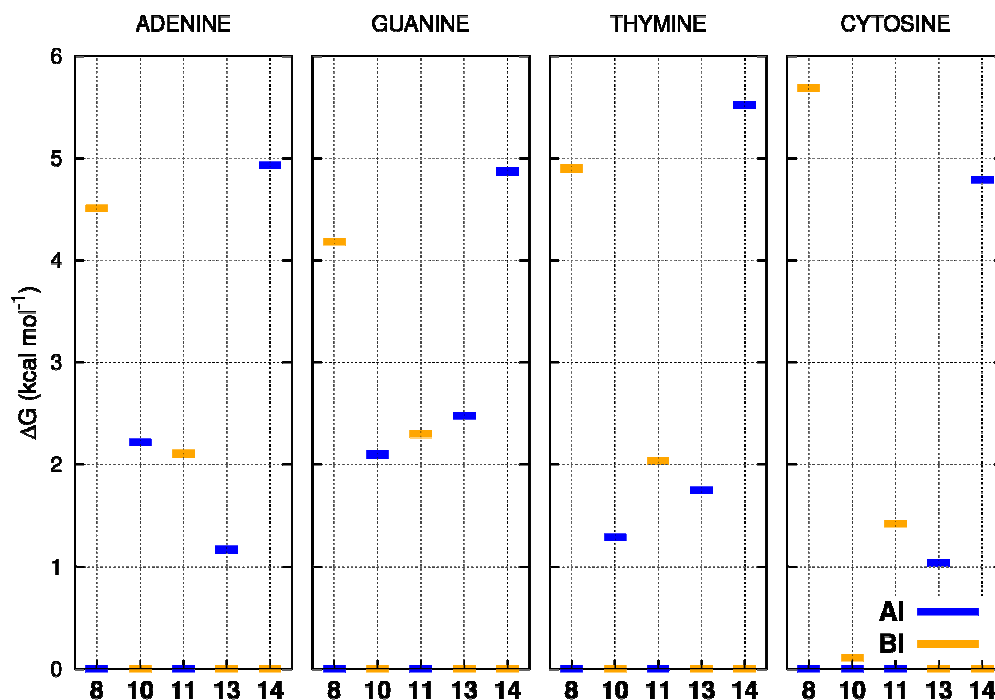
Fig. 2. Relative Gibbs energies (in kcal/mol) for dNs with selected sugar modifications (**8**, **10**, **11**, **13**, and **14**) including four canonical nucleobases (adenine, guanine, thymine and cytosine). Data for modifications **10**, **11** and **13** are taken from ref. 75.
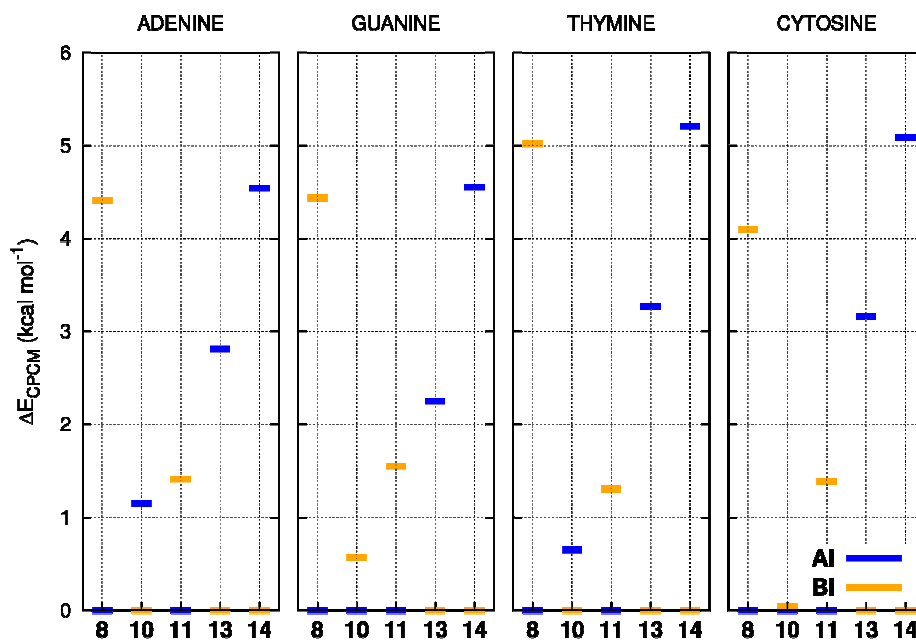

Fig. 3. Relative electronic energies (in kcal/mol) in water estimated by CPCM model[78] for dNs with selected sugar modifications (**8**, **10**, **11**, **13**, and **14**) including four canonical nucleobases.

The last group of the sugar modifications, *FLR* (*fluorinated*) contains dNs incorporating fluorine atoms at various positions of the sugar residue (Fig. 1) such as 2'-α (modification **10**), 2'-β (**11**) and simultaneous fluorination at both 2'-α and 2'-β positions (**12**). The modification **13** (Fig. 1) consists in the replacement of O4' by the CF_2 group that was suggested to be isopolar and isosteric to oxygen atom.[93] It should be remarked that modifications **10-13** were studied and discussed in our previous work[75] and therefore will be mentioned here just briefly. Whereas the modifications **10-12** were studied earlier in various experimental and theoretical works (see, e.g., refs. 60-62 and refs. therein), the O4'→CF_2 substitution (modification **13**) in nucleic-acid constituents was suggested and studied for the first time in our research group.[75] In particular, it was shown[75] that *FLR*-**13**-dNs (Fig. 1) can be considered as promising building blocks for designing artificial nucleic acids due to their intrinsic propensity to favor BI conformation as well as due to the substantial increase in the glycosidic bond strengths compared to those for natural nucleotides. In this work, we suggest a new modification, **14** (Fig. 1), combining both fluorination at the 2'-α position and the O4'→CF_2 substitution. Since both these modifications were shown[75] to yield a noticeable stabilization of BI conformations relative to AI ones and significant increase in glycosidic bond strength, it is reasonable to expect that their combination will magnify these two effects.

As shown in ref. 75, the modifications **10** and **13** give rise to an additional energetic stabilization of BI conformers relative to AI conformations compared to *CAN-dNs* as evidenced by analysis of relative electronic energy in vacuum and Gibbs energy values (Table S1, Fig. 2). In this regard, it would be interesting to know whether the solvent effects can shift the equilibrium between BI and AI forms. Indeed, in case of *FLR*-**10**-dNs the inclusion of the implicit solvent model reduces the energy gaps between BI and AI conformational states as suggested by ΔE_{CPCM} values (Figs. 2 and 3). In Gua-*FLR*-**10**-dN the energetic order between AI and BI is reversed, i.e. the AI structure becomes more stable. On the contrary, the analysis of the ΔE_{CPCM} values for *FLR*-**13**-dNs suggests that like in the gas phase[75], the BI conformers in solution are expected to be much more energetically favorable than AI structures.

For dNs with the sugar residues **11** and **12**, it was demonstrated[75] that AI conformations are more stable compared to their BI counterparts. One of the possible reasons for energetic destabilization of BI conformers is associated with the repulsive O...F intra-sugar interactions between the O5' oxygen and fluorine atom at the 2'-β position. As evidenced by the electron density topological analysis based on Quantum Theory of Atoms in Molecules (QTAIM)[94], these steric interactions arise only in BI (C2'-endo) conformers.[75] The energetic trend is preserved in solution as follows from the comparison of ΔE_{CPCM} energies with the ΔG and ΔE_{VAC} values calculated *in vacuo* (Figs. 2 and 3, Table S1).

Finally, we investigate here a new modification **14** combining fluorination at the 2'-β position and the O4'→CF_2 substitution. Since both these modifications lead to stabilization of BI-form conformers, it

can be assumed that in *FLR*-**14**-dNs this stabilization will be stronger, i.e. the energy gaps between BI and AI structures will rise. Indeed, both *in vacuo* and in-solvent energies (Figs. 2 and 3, Table S1) confirm this assumption.

From a conformational point of view, the sugar-ring pucker in AI and BI conformers of *FLR*-**14**-dNs is slightly shifted from C3'-endo and C2'-endo conformations to C2'-exo and C3'-exo, respectively. This is also true for *FLR*-**13**-dNs as described previously.[75] However, since the pucker still remains in South (AI) and North (BI) regions, this shift should not affect the compatibility of modified nucleotides with the natural ones and should not induce any steric penalties upon incorporation of *FLR*-**13**-dNs or *FLR*-**14**-dNs units into AI and BI DNA structures.

Finally, for some selected modifications (**8**, **10**, **11**, **13**, and **14**) we applied another continuum solvent model – SMD, which has been recently developed by Marenich, Cramer, and Truhlar.[95] This allowed comparing two alternative approaches – CPCM and SMD in evaluating possible energetic consequences of the sugar residue modification (Fig. 3 and Fig. S3 in ESI). As a result, two models provide the same energetic trends for selected modifications. The most substantial difference between CPCM and SMD approaches is observed for *FLR*-**10**-dNs: the AI and BI conformers of Ade-*FLR*-**10**-dN and Thy-*FLR*-**10**-dN become energetically quasi-degenerate, when the SMD model is applied (Fig. S3 in ESI).

3.2 The investigation of intrinsic strengths of glycosidic bonds in dNs using Relaxed Force Constants (RFC) methodology. We start analyzing mechanical properties of dNs by applying the Grunenberg's relaxed force constants (RFC) formalism[83-85] to evaluation of intrinsic strengths of glycosidic bonds. This methodology proved to be efficient in establishing strengths of both covalent[96,97] and non-covalent[55,84,98] bonds.

The strength of glycosidic bond represents one of the crucial characteristics that determine the general stability not only of single dNs units but also of the whole DNA/RNA structures containing these units. For instance, the likelihood of abasic lesions in nucleic acids resulting in appearance of apurinic/apyrimidinic (AP) sites is, together with other factors, related to intrinsic strength of glycosidic bonds. In our previous work,[75] we applied RFC methodology for estimating the kinetic stability of glycosidic bonds in fluorinated dNs with the sugar modifications **10**, **11**, **12**, and **13** (Fig. 1). Note that all these modifications yield substantially stronger glycosidic bonds[75] compared to *CAN*-dNs (Fig. 4). Among these modifications, the greatest increase in ΔRFC was observed for *FLR*-**13**-dNs with the O4'→CF$_2$ substitution. Thus, it can be predicted that *FLR*-**14**-dNs combining both O4'→CF$_2$ replacement (modification **13**) and fluorination at the 2'-α-position (modification **10**), will lead to the strongest glycosidic bonds. It turned out that dNs with the sugar residue **14** (Fig. 1) are characterized by the greatest

values of RFC (463-485 N·m^{-1}) not only in the *FLR* group, but also among all compounds studied in this work. Fig. 4 displays the relative changes in RFC (ΔRFC) in selected modified dNs with respect to their canonical counterparts. Clearly, *FLR*-**14**-dNs (Fig. 4) show the greatest ΔRFC values ranging from 43.4 N·m^{-1} (BI conformation of the Cyt-*FLR*-**14**-dN molecule) to 87.7 N·m^{-1} (AI conformer of Thy-*FLR*-**14**-dN).
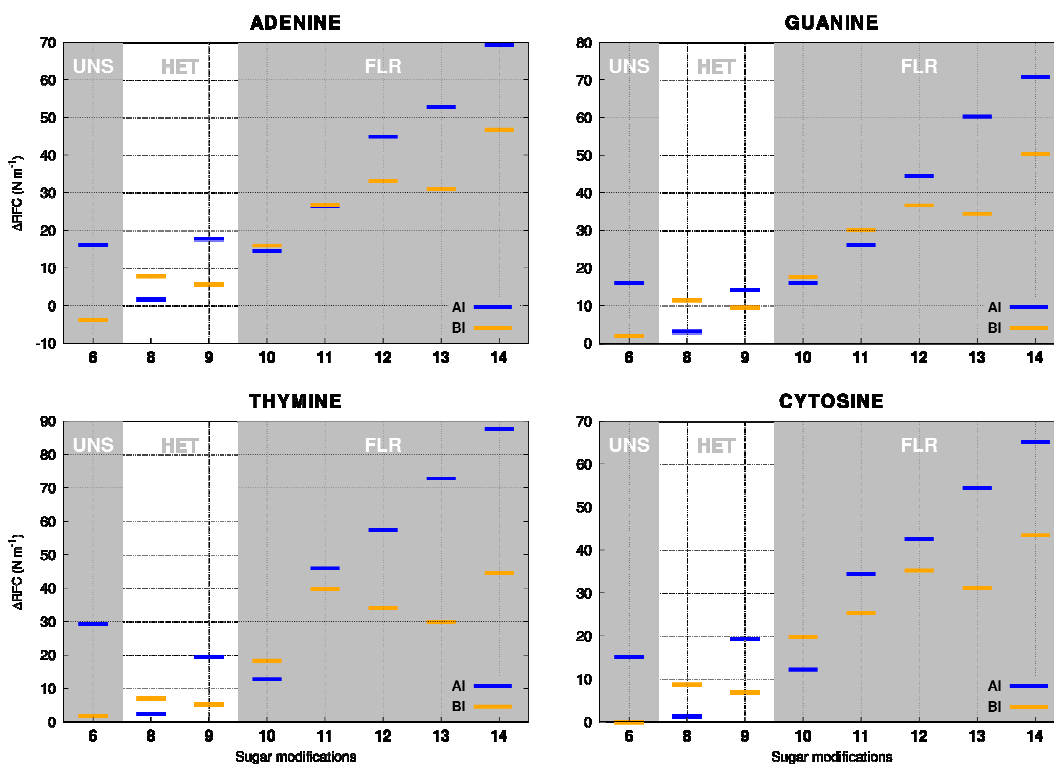

Fig. 4. Relative changes in Relaxed Force Constant (RFC) values, ΔRFC (in N·m^{-1}), for glycosidic bond in AI and BI conformers of modified dNs compared to their canonical (unmodified) counterparts. The data for modifications **10**, **11**, **12**, and **13** are taken from ref. 75. The positive ΔRFC values suggest that the studied sugar modifications (Fig. 1) result in glycosidic bond strengthening.

In the *UNS* group, the dNs with the sugar **1** containing the C2'=C3' double bond (Fig. 1) have RFC mean value of 387.5 N·m^{-1} that lies between the mean values of AI (374.8 N/m) and BI (425.8 N/m) conformers of *CAN*-dNs. As expected, fluorination at different positions (UNS-**2-4**-dNs), and especially the O4'→CF$_2$ substitution (modification **4**) substantially enhances the glycosidic bond compared to UNS-**1-**dNs (Table S1). On the contrary, the modification **6** including an external C=C double bond at the 4'-position enhances glycosidic bonds only in AI conformers (relative to Can-dNs), but has no effect on BI structures.

In the locked nucleotides (LOC-dNs, modification **7**, Fig. 1), which are able to adopt only the AI conformation, the average value of RFC for glycosidic bonds (391.1 N/m) exceeds the respective value for AI conformers of CAN-dNs (374.8 N/m). Therefore, this modification is efficient in strengthening these bonds in A-DNA.

Finally, both modifications **8** and **9** (*HET* subgroup) (Fig. 1) give rise to reinforcement of glycosidic bonds (Fig. 4, Table S1). However, this effect is more pronounced for *HET*-**9**-dNs (Fig. 4) containing a sulfur heteroatom at the 3'-position of deoxyribose ring.

To sum up, the analysis of the Grunenberg's force constants leads us to conclude that all the modifications studied in this work, except modification **6** (BI conformers), result in some strengthening of glycosidic bonds relative to canonical nucleotides, and therefore can be efficiently used for designing nucleic-acid-based systems (only for units with available 5'-/3'-OH groups or as 3'-terminal residues) with increased resistance to abasic lesions. The most significant reinforcement of the glycosidic bonds occurs for systems with the sugar **14** (Fig. 1), containing the CF_2 group at the position 4' and the 2'-α-F atom.

3.3 Mechanical properties (relaxed force constants and vibrational root-mean-square deviations) of conformational angles in dNs and their relation to structural and energetic characteristics. The investigation of relaxed force constants for main conformational parameters in dNs (torsion angles α, β, γ, δ, and χ and pseudorotational angle P) is essential for understanding and predicting the structural behavior of oligonucleotide sequences. Moreover, the information on mechanical properties of dNs, the basic nucleic-acid building blocks, can be useful in various areas ranging from force-field development for modified nucleic acids[100] to design of nucleic-acid-based aptamers and nanodevices with specific properties.

The mechanical properties of some biologically significant conformers of canonical 2'-deoxyribonucleosides were recently reported by Nikolaienko et al.[14] As regards 2´-deoxyribonucleotide-5´-monophosphates (dNs), the DNA basic building blocks, only limited information is available on the flexibility of their conformational degrees of freedom.[18] Moreover, the influence of the sugar modifications on the intrinsic rigidity of dNs molecules as well as the relationship between the energetic preferences and mechanical properties has not been explored yet. For this purpose, we analyze the relaxed force constants (RFCs) and vibrational root-mean-square (VRMS) deviations for the major conformational parameters in dNs calculated according to the methodology developed in ref. 16.
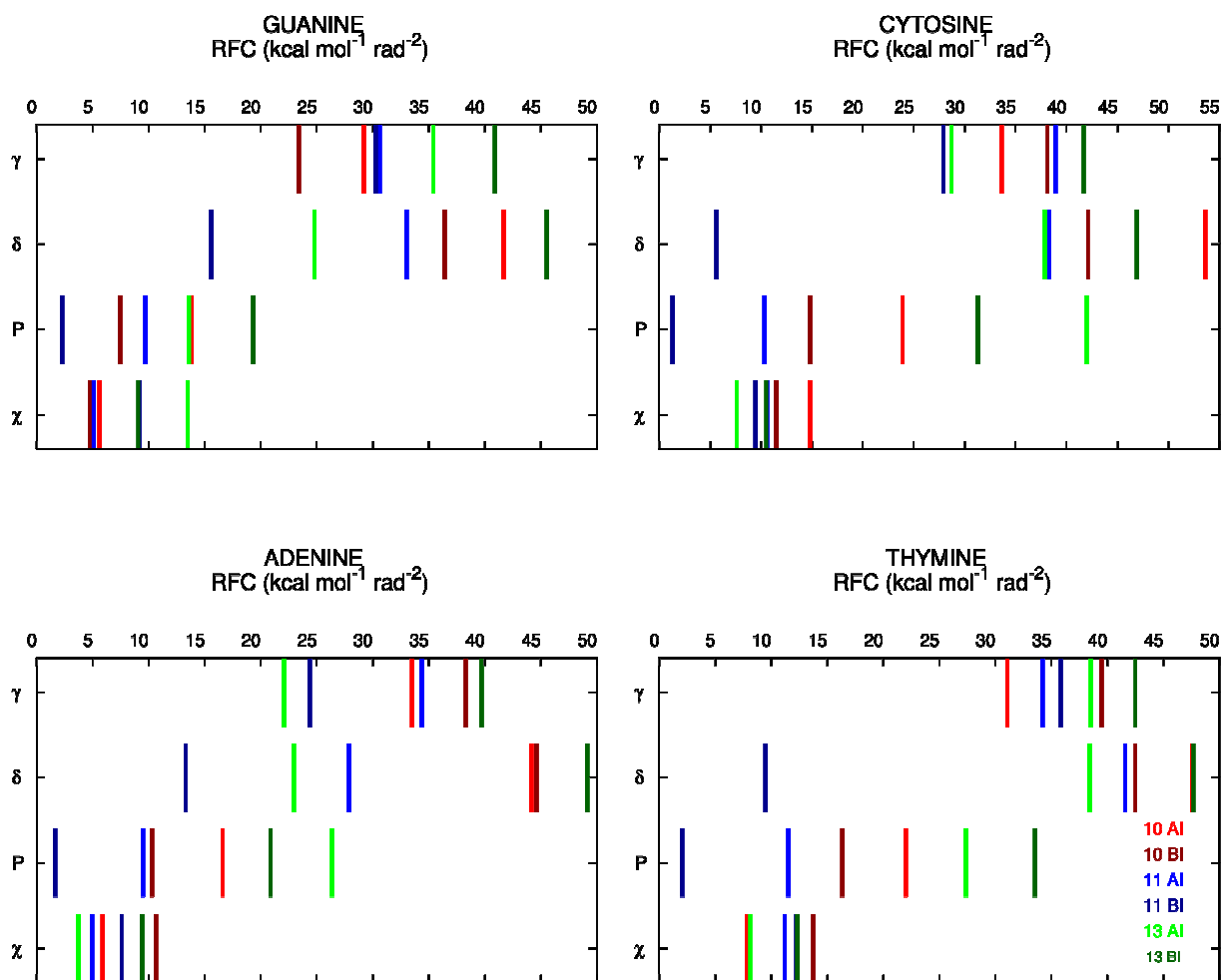
Fig. 5. The scale of relaxed force constants (RFC) for 2′-deoxyribonucleotide-5′-monophosphates (dNs) in their AI and BI conformations with the sugar residues **10**, **11**, and **13** (Fig. 1). The RFC for dNs containing each canonical nucleobase (guanine, cytosine, adenine, thymine) are shown separately.

The calculated RFCs for the main conformational parameters in dNs (the α, β, γ, δ, χ, and P conformational angles) are shown in Table S1. Fig. 5 displays the distribution of RFC for some selected conformational angles (γ, δ, P, and χ) for dNs with the sugar residues **10**, **11**, and **13** (Fig. 1). Before analyzing the influence of chemical modifications on mechanical properties of dNs, we discuss the canonical systems, i.e. *CAN*-dNs. Among various conformational parameters in *CAN*-dNs, the χ angle appeared, as expected, to be the softest one, with RFC (χ) ranging from 3.9 to 12.2 kcal·mol^{-1}·rad^{-2} (Table S1). The three other angles (torsion angles α, β, and pseudorotational angle P) are characterized by similar degrees of rigidity, with RFCs falling into the ~8 to ~20 kcal·mol^{-1}·rad^{-2} interval (the smallest RFCs of ~8 kcal·mol^{-1}·rad^{-2} correspond to BI conformers of *CAN*-dNs with adenine and guanine bases).

Finally, the γ and δ angles are characterized by the highest RFC values of 32.4÷39.0 and 26.3÷43.6 kcal·mol[-1]·rad[-2], respectively (Table S1). In the previous work on 2'-deoxyribonucleosides[14] substantially smaller RFC values were reported for β (~ 1÷5 kcal·mol[-1]·rad[-2]) and γ (~ 17÷ 27 kcal·mol[-1]·rad[-2]) angles. As assumed, our results indicate that the presence of a phosphate group in dNs (this group is absent in 2'-deoxyribonucleosides) imparts an additional rigidity to the β and γ angles.
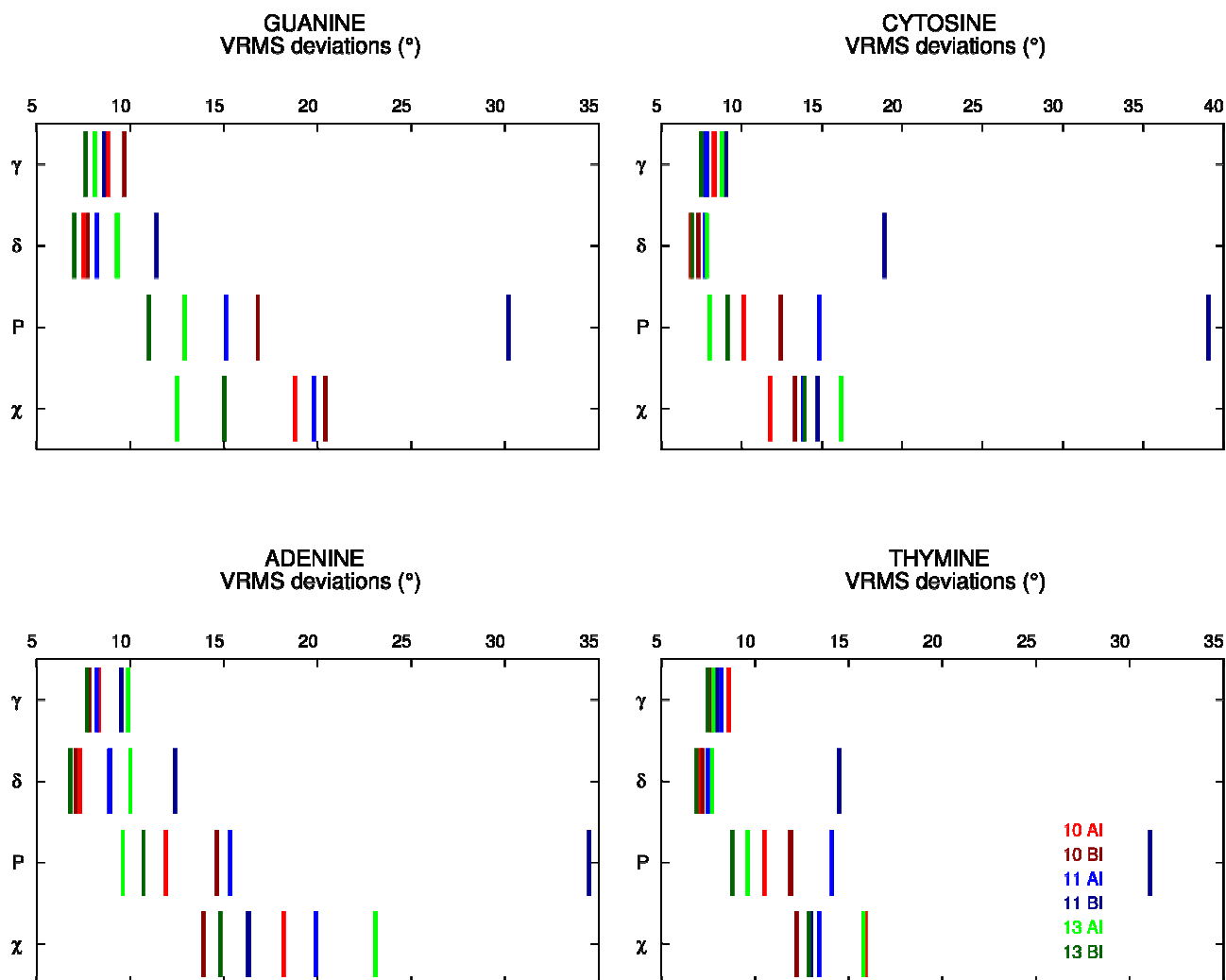

Fig. 6. The scale of vibrational root-mean-square (VRMS) deviations at T = 298K for dNs with the sugar residues **10**, **11**, and **13** (for structures, see Fig. 1). The VRMS deviations for dNs containing each canonical nucleobase (guanine, cytosine, adenine, thymine) are shown separately.

In addition to RFCs, vibrational root-mean-square (VRMS) deviations for the conformational parameters were investigated at two temperatures, T = 0 K and T = 298 K. These deviations occur since

in any real molecule atomic nuclei are not fixed, and participate in quantum zero-point and thermal motions. Clearly, the greatest VRMS deviations are observed for the softest conformational parameters (Figs. 6 and S4, Table S2). For instance, for the pseudorotational phase angle P, the VRMS deviation at T = 298 K (from 11.5° to 15.7°) is comparable with a half-width (18°) of a C2'-endo or C3'-endo conformational range.

Taking into account the mechanical properties of *CAN*-dNs, we now start exploring the effects of various sugar modifications on intrinsic flexibility of dNs. In the *UNS* (unsaturated) group of dNs (Fig. 1), the *UNS*-**6**-dNs with a C4'=C double bond deserve a special attention, since as expected this modification noticeably increases the rigidity of the sugar ring. This fact is evident from analyzing the RFC values for δ and P angles. For instance, the RFCs(P) for *UNS*-**6**-dNs lie within the 19.3÷31.6 kcal·mol⁻¹·rad⁻² range, i.e. are perceptibly greater than RFCs(P) for respective *CAN*-dNs (*vide supra*). The enhanced sugar-ring rigidity upon the O4'→C=CH₂ substitution is also evident from analysis of energetic data (Table S1) for *UNS*-**6**-dNs. Compared to *CAN*-dNs, the energy differences between BI and AI conformers augment both *in vacuo* and in water solution (CPCM[78] model) data. Moreover, ΔE_{VAC}, ΔG, and ΔE_{CPCM} manifest the same trend, thus, the solvent effects do not reverse the BI/AI energetic order as in case of *CAN*-dNs. This serves an additional evidence of the elevated rigidity of the deoxyribose cycle.

In the next group, *LOC* (locked) containing only a single sugar type (modification **7**, Fig. 1), we should also point to very large values of RFC(δ) and RFC(P), which is not surprising, since these molecules are covalently 'locked' in the AI conformation. Conversely, RFC(γ) are smaller in *LOC*-**7**-dNs (26.6÷33.1 kcal·mol⁻¹·rad⁻²) than in AI conformations of *CAN*-dNs. Therefore, this modification makes the γ angle more flexible.

In the *HET* (heterocyclic) group, the *HET*-**9**-dNs containing a sulfur atom at the 3'-position of the sugar ring, possess interesting mechanical characteristics. The values of RFC (χ) and RFC (P) in these compounds strongly vary with the type of attached nucleobase (adenine, guanine, cytosine, or thymine) and the sugar conformation (AI or BI). This can be explained by the fact that the sulfur atom is bulkier and more polarizable than carbon, thus, it can be involved in the intramolecular interactions between the base and sugar. In particular, the QTAIM analysis[94] performed on these structures (Table S3 in ESI) indicates that the S atom in the *HET*-**9**-dNs can be involved in pseudo-hydrogen bonds C8H (purine dNs) / C6H (pyrimidine dNs) ... S. These non-covalent interactions affect the degree of flexibility of the sugar residue, and since their strengths are strongly dependent on conformation and the type of the attached nucleobase, the RFC (χ) and RFC (P) noticeably vary from one structure to another.

The last group, *FLR* (fluorinated), includes dNs with the sugar residues from **10** to **14** (Fig. 1). The mechanical properties of fluorinated dNs are tightly connected with their energetic preferences. As mentioned above, modifications **10**, **13**, and **14** stabilize the BI sugar conformation, i.e. decrease the

relative energies of BI structures in AI-BI pairs as compared with *CAN*-dNs. However, the extent of stabilization of BI conformers rises in the order *FLR*-**10**-dNs < *FLR*-**13**-dNs < *FLR*-**14**-dNs (Figs. 2 and 3). Since the AI and BI forms differ only in the type of sugar pucker (C2'-endo in BI and C3'-endo in AI) which is determined by the values of P and δ angles, it is expected that in all three classes of dNs, the RFC(P) and RFC(δ) should augment relative to those for their canonical counterparts, and the changes in RFCs should be in the order ΔRFC_{10} (δ or P) < ΔRFC_{13} (δ or P) < ΔRFC_{14} (δ or P). Indeed, this is true (Table S1, Fig. 5) and the greatest RFC (δ or P) values and the lowest VRMS deviations (Figs. 6 and S4) are observed for BI conformers of dNs with the sugar residues **13** and **14**. In case of *FLR*-**10**-dNs the ΔRFC_{10} (δ or P) values are smaller, therefore, the solvent effects may reverse the BI/AI energetic order as can be seen for Gua-*FLR*-**10**-dN (Figs. 2 and 3, Table S1). However, in *FLR*-**13**-dNs and *FLR*-**14**-dNs, the RFC(δ) and RFC(P) are high enough (Fig. 5) and, consequently, the same energetic trends are observed for modifications **13** and **14** in vacuum and in implicit solvent (Figs. 2 and 3).

Finally, modifications **11** and **12** (Fig. 1) lead to destabilization of BI form, therefore, AI conformers of *FLR*-**11**-dNs and *FLR*-**12**-dNs are more stable in vacuum and in implicit solvent. This destabilization, possibly caused by 2'-β-F...O5' repulsive interactions (*vide supra*), is reflected in very small values of RFC(δ) and RFC(P) for BI conformers of *FLR*-**11**-dNs and *FLR*-**12**-dNs (Fig. 5 and Table S1) and very large VRMS deviations. For BI conformer of the dC-*FLR*-**11**-dN nucleotide (Fig. 6 and Table S1) the VRMS at T = 298 K for P angle is equal to 39.1°, exceeding the width (36°) of the C2'-endo conformational sector. In other words, this conformer is dynamically unstable at T = 298 K, and can convert to another conformer with a different sugar pucker as a result of zero-point and thermal motions. It should be also noted that in *FLR*-**11**-dNs and *FLR*-**12**-dNs the RFC are decreased for α and β angles as compared to *CAN*-dNs. However, the ΔRFC for the above-mentioned angles (P, δ, α, and β) are smaller in case of *FLR*-**12**-dNs (Fig. 5, Table S1) since the presence of a second fluorine in the 2'-α-position of these molecules (Fig. 1) imparts an additional stability to BI conformers despite the 2'-β-F...O5' destabilizing interactions.

In addition, for all the studied systems we noticed a very remarkable feature: there is a relationship between RFC(δ) for AI and BI conformers and their relative electronic energies *in vacuo*, ΔE_{VAC}. In all the cases, higher RFC(δ) values calculated for AI or BI conformers of a particular nucleotide correspond to ΔE_{VAC} = 0 kcal/mol (Table S1). In other words, in a pair of structures (AI and BI) it is possible to predict which conformation is more stable by analyzing RFC(δ), since this conformation is characterized by higher value of RFC(δ). In most of cases (with some exceptions) this is also true for other energetic characteristics, ΔG and ΔE_{CPCM}. However, ΔG is temperature-dependent and includes thermal and entropic effects, whereas ΔE_{CPCM} reflects the possible influence of the solvent.

Therefore, it is logical that the best correlation is observed between RFC(δ) and ΔE_{VAC} since both quantities characterize intrinsic properties of a stationary molecule.

To sum up, for the first time the mechanical properties of dNs were comprehensively described and characterized, and the possible influence of various sugar modifications on the rigidity of these molecules was investigated. As a result, the mechanical characteristics of dNs were found to be tightly related to their conformational and energetic features.

4 CONCLUSIONS

The design of novel building blocks of nucleic acids with some desired properties (e.g., ability to favor a particular conformation and form specific supramolecular structures, modified flexibility of the sugar-phosphate residue) represents one of the most promising and important current trends in NA chemistry, medicine, and biotechnology.[59] This work is a continuation of our previous research[75,101-105] devoted to *in silico* design of novel building blocks for nucleic acid duplexes and quadruplexes with potential applications in various areas ranging from analytical chemistry to bio- and nanotechnology.

Here we investigate for the first time the intrinsic mechanical flexibility of the 2´-deoxyribonucleotide-5´-monophosphates (dNs) and find the relation between mechanical properties of dNs, on the one hand, and their energetic and conformational characteristics, on the other hand. In addition to canonical dNs (*CAN*-dNs) in the biologically significant AI- and BI-form DNA-like conformations, we comprehensively explored the effects of fourteen sugar modifications (Fig. 1) on the energetic and conformational behavior of dNs, as well as on the flexibility of the main conformational parameters.

First, it was established that most of the studied modifications strongly shift the equilibrium between AI-form and BI-form DNA-like conformations either to the former or the later arrangement. This tendency was also preserved in aqueous solution modeled by CPCM solvent model[78] although we are aware of the limitations of this approach. The propensity of modified dNs to strongly favor either BI or AI conformation can have beneficial applications, for instance, in analytical chemistry, since a selective incorporation of these units into nucleic-acid duplexes/quadruplexes should reduce their structural polymorphism in solution.

Second, it was shown that most of the studied sugar modifications benefit from an increased stability of a glycosidic bond as evidenced by relaxed force constants analysis,[83-85] and can therefore be used for designing modified nucleic acids with a higher resistance to abasic lesions. The most significant reinforcement of the glycosidic bonds occurs in case of systems with the sugar **14** (Fig. 1), containing the CF_2 group instead of the O4' oxygen and the fluorine atom at the 2'-α-position.

Third, a strong relation between mechanical characteristics (relaxed force constants and vibrational root-mean-square deviations of the conformational degrees of freedom) of dNs and their conformational and energetic features was found. Most notably, the relative stability (estimated by electronic energy *in vacuo*, ΔE_{VAC}) of AI and BI conformers of a particular nucleotide is closely connected with the values of relaxed force constants (RFC) for the δ angle: the higher RFC(δ) values correspond to more stable conformers.

Finally, the knowledge of force constants for the main conformational parameters can be used in development and refinement of force fields for modeling modified nucleic acids at the larger scale in the presence of an explicit solvent. One of the crucial requirements to the force field parameters obtained for building blocks (such as nucleotides) is their transferability to higher-order structures, i.e. oligonucleotides and nucleic acids structures. We assume that in larger systems RFC for the main conformation parameters should somewhat increase due to a higher structural rigidity imposed by intra- and inter-strand interactions. The solvent molecules interacting with the polar sites (e.g. phosphate and terminal hydroxyl groups) within the sugar residue can also affect the values of RFC for torsion angles. In principle, these effects should not yield significant changes in the RFC values relative to those for single nucleotides, and therefore, these parameters can be transferable to the larger systems. However, this assumption should be tested on larger DNA fragments.

References

1. S. Neidle, Principles of nucleic acid structure, Elsevier, Amsterdam, 2008

2. B. Hartmann, R. Lavery, *Q. Rev. Biophys*., 1996, **29**, 309.

3. G.N. Parkinson, Fundamentals of Quadruplex Structures. In: Quadruplex Nucleic Acids (Eds. S. Neidle, S. Balasubramanian), RSC Publishing, Cambridge, **2006**.

4. R.-J. Sung, M. Zhang, Y. Qi, G.L. Verdine, *J. Biol. Chem*., 2012, **287**, 18044.

5. K. Struhl, E. Segal, *Nat. Struct. Mol. Biol*., 2013, **20**, 267.

6. G.P. Santini, C. Pakleza, J.A. Cognet, *Nucleic Acids Res*., 2003, **31**, 1086.

7. S. Jones, P. van Heyningen, H.M. Berman, J.M. Thornton, *J. Mol. Biol*., 1999, **287**, 877.

8. F. Coste, J.M. Malinge, L. Serre, W. Shepard, M. Roth, M. Leng, C. Zelwer, *Nucleic Acids Res*., 1999, **27**, 1837.

9. A. Gonzalez-Perez, *Curr. Top. Med. Chem*., 2014, **14**, 766.

10. F.H.C. Crick, A. Klug, Nature, 1975, **255**, 530.

11. N.Z. Namoradze, A.N. Goryunov, T.M. Birshtein, *Biophys. Chem*., 1977, **7**, 59.

12. A. Noy, R. Golestanian, *Phys. Rev. Lett*., 2012, **109**, 228101.

13. A.K. Mazur, M. Maaloum, *Nucleic Acids Res*., 2014, **42**, 14006.

14. T.Y. Nikolaienko, L.A. Bulavin, D.M. Hovorun, *Phys. Chem. Chem. Phys*., 2012, **14**, 15554.

15. Z. Konkoli, J.A. Larsson, D. Cremer, *Int. J. Quant. Chem*., 1998, **67**, 41.

16. T.Y. Nikolaienko, L.A. Bulavin, D.M. Hovorun, *J. Biomol. Struct. Dyn*., 2011, **29**, 563.

17. C. Altona, M. Sundaralingam, *J. Am. Chem. Soc*., 1972, **94**, 8205.

18. T. Yu. Nikolaienko, L.A. Bulavin, D.M. Hovorun, *Ukr. Biochem. J*., 2011, **83**, 48 (in Ukrainian).

19. E. Ichikawa, K. Kato, *Curr. Med. Chem*., 2001, **8**, 385.

20. P. Herdewijn (Ed.) Modified nucleosides in biochemistry, biotechnology and medicine, 2008, Weinheim: Wiley.

21. H. Liu, J. Gao, S.R. Lynch, Y.D. Saito, L. Maynard, E.T. Kool, *Science*, 2003, **302**, 868.

22. H. Liu, S.R. Lynch, E.T. Kool, *J. Am. Chem. Soc*., 2004, **126**, 6900.

23. H. Lu, K. He, E.T. Kool, *Angew. Chem. Int. Ed*., 2004, **43**, 5834.

24. M. Winnacker, E.T. Kool, *Angew. Chem. Int. Ed*., 2013, **52**, 12498.

25. T.L. McConnell, S.D. Wetmore, *J. Phys. Chem. B*., 2007, **111**, 2999.

26. L.A. Lait, L.R. Rutledge, A.L. Millen, S.D. Wetmore, *J. Phys. Chem. B*., 2008, **112**, 12526.

27. D. Varsano, A. Garbesi, R. DiFelice, *J. Phys. Chem. B*., 2007, **111**, 14012.

28. O. Huertas, J. Poater, M. Fuentes-Cabrera, M. Orozco, M. Solà, F.J. Luque, *J. Phys. Chem. A*., 2006, **110**, 12249.

29. P. Sharma, L.A. Lait, S.D. Wetmore, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15538.

30. P. Sharma, L.A. Lait, S.D. Wetmore, *Phys. Chem. Chem. Phys.*, 2013, **15**, 2435.

31. S. Obika, D. Nanbu, Y. Hari, K. Morio, Y. In, T. Ishida, T. Imanishi, *Tetrahedron Lett.*, 1997, **38**, 8735.

32. P.E. Nielsen, M. Egholm, R.H. Berg, O. Buchardt, *Science*, 1991, **254**, 1497.

33. K.-U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science*, 2000, **290**, 1347.

34. O. Berger, L. Adler-Abramovich, M. Levy-Sakin, A. Grunwald, Y. Liebes-Peer, M. Bachar, L. Buzhansky, E. Mossou, V.T. Forsyth, T. Schwartz, Y. Ebenstein, F. Frolow, L.J. Shimon, F. Patolsky, E. Gazit, *Nat. Nanotechnol.*, 2015, **10**, 353.

35. Y. You, B.G. Moreira, M.A. Behlke, R. Owczarzy, *Nucleic Acids Res.*, 2006, **34**, e60.

36. M.R. Dunn, A.C. Larsen, W.J. Zahurancik, N.E. Fahmi, M. Meyers, Z. Suo, J.C. Chaput, *J. Am. Chem. Soc.*, 2015, **137**, 4014.

37. D.E. Condon, I. Yildirim, S.D. Kennedy, B.C. Mort, R. Kierzek, D.H. Turner, *J. Phys. Chem. B*, 2014, **118**, 1216.

38. T. Koch, I. Shim, M. Lindow, H. Orum, H.G. Bohr, *Nucleic Acid Ther.*, 2014, **24**, 139.

39. S. Sen, L. Nilsson, *J. Am. Chem. Soc.*, 2001, **123**, 7414.

40. T. Natsume, Y. Ishikawa, K. Dedachi, N. Kurita, *Chem. Phys. Lett.*, 2006, **418**, 239.

41. T. Natsume, Y. Ishikawa, K. Dedachi, T. Tsukamoto, N. Kurita, *Chem. Phys. Lett.*, 2007, **434**, 133.

42. S. Sharma, U.B. Sonavane, R.R. Joshi, *J. Biomol. Struct. Dyn.*, 2010, **27**, 663.

43. T. Rathinavelan, N. Yathindra, *FEBS J.*, 2005, **272**, 4055.

44. J.M. Sanders, M.E. Wampole, C.-P. Chen, D. Sethi, A. Singh, F.-Y. Dupradeau, F. Wang, B.D. Gray, M.L. Thakur, E. Wickstrom, *J. Phys. Chem. B*, 2013, **117**, 11584.

45. R. Soliva, E. Sherer, F.J. Luque, C.A. Laughton, M. Orozco, *J. Am. Chem. Soc.*, 2000, **122**, 5997.

46. C. Fan (Ed.), *DNA Nanotechnology: From Structure to Function*, 2013, Springer-Verlag, Berlin Heidelberg.

47. J.C. Rohloff, A.D. Gelinas, T.C. Jarvis, U.A. Ochsner, D.J. Schneider, L. Gold, N. Janjic, *Mol. Ther. Nucleic Acids*, 2014, **3**, e201.

48. E.W. Ng, D.T. Shima, P. Calias, E.T. Cunningham, Jr., D.R. Guyer, A.P. Adamis, *Nat. Rev. Drug Discov.*, 2006, **5**, 123.

49. B.P. Monia, E.A. Lesnik, C. Gonzalez, W.F. Lima, D. McGee, C.J. Guinosso, A.M. Kawasaki, P.D. Cook, S.M. Freier, *J. Biol. Chem.*, 1993, **268**, 14514.

50. S. Davis, S. Propp, S.M. Freier, L.E. Jones, M.J. Serra, G. Kinberger, B. Bhat, E.E. Swayze, C.F. Bennett, C. Esau, *Nucleic Acids Res*., 2009, **37**, 70.

51. C.J. Lech, Z. Li, B. Heddi, A.T. Phan, *Chem. Commun*., 2012, **48**, 11425.

52. R.A. Blinder, R.P. Hammer, M.J. Lopez, S.O. Robinson, W.T. Monroe, *Chem. Biol. Drug Des*., 2007, **70**, 113.

53. J.S. Jepsen, M.D. Sørensen, J. Wengel, *Oligonucleotides*, 2004, **14**, 130.

54. A.P. Lea, D. Faulds, *Drugs*, 1996, **51**, 846.

55. A.G. Ponomareva, Y.P. Yurenko, R.O. Zhurakivsky, T. van Mourik, D.M. Hovorun, *Phys. Chem. Chem. Phys*., 2012, **14,** 6787.

56. A.G. Ponomareva, Y.P. Yurenko, R.O. Zhurakivsky, T. van Mourik, D.M. Hovorun, *J. Biomol. Struct. Dyn*., 2014, **32**, 730.

57. W.D. Meng, F.L. Qing, *Curr. Top. Med. Chem*., 2006, **6**, 1499.

58. E.W. Ng, D.T. Shima, P. Calias, E.T. Cunningham, Jr., D.R. Guyer, A.P. Adamis, *Nat. Rev. Drug Discovery*, 2006, **5**, 123.

59. O. Doluca, J.M. Withers, V.V. Filichev, *Chem. Rev*., 2013, **113**, 3044.

60. N. Martin-Pintado, G.F. Deleavey, G. Portella, R. Campos-Olivas, M. Orozco, M.J. Damha, C. González, *Angew. Chem. Int. Ed*. 2013, **52***,* 12065.

61. C.J. Lech, Z. Li, B. Heddi, A.T. Phan, *Chem. Commun*. 2012, **48**, 11425.

62. H. Ikeda, R. Fernandez, A. Wilk, J.J. Barchi, Jr., X. Huang, V.E. Marquez, *Nucleic Acids Res*., 1998, **26**, 2237.

63. A. Patra, M. Paolillo, K. Charisse, M. Manoharan, E. Rozners, M. Egli, *Angew. Chem. Int. Ed*. 2012, **51**, 11863.

64. L.A. Bulavin, D.M. Hovorun, T.Y. Nikolaienko, Structure of DNA monomers, Naukova Dumka, Kiev, 2014 (in Russian).

65. T. Yu. Nikolaienko, L.A. Bulavin, D.M. Hovorun, *Phys. Chem. Chem. Phys*., 2012, **14**, 7441.

66. O.V. Shishkin, L. Gorb, O.A. Zhikol, J. Leszczynski, *J. Biomol. Struct.Dyn*., 2004, **22**, 227.

67. O.V. Shishkin, L. Gorb, O.A. Zhikol, J. Leszczynski, *J. Biomol. Struct.Dyn*., 2004, **21**, 537.

68. A.G.W. Leslie, S. Arnott, R. Chandrasekaran, R.L. Ratliff, *J. Mol. Biol*., 1980, **143**, 49.

69. T. Maehigashi, C. Hsiao, K.K. Woods, T. Moulaei, N.V. Hud, L.D. Williams, *Nucleic Acids Res*. 2012, **40**, 3714

70. V.B. Zhurkin, Y.P. Lysov, V.I. Ivanov, *FEBS Lett*. 1975, **59**, 44.

71. Y. Timsit, *J. Mol. Biol*., 1999, **293**, 835.

72. A. D. Becke, *Phys. Rev*., 1988, **38**, 3098

73. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys*., 1988, **37**, 785

74. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, J. Chem. Phys., 2010, 132, 154104

75. Y.P. Yurenko, J. Novotný, V. Sklenář, R. Marek, *Chem. Eur. J.*, 2015, **21**, DOI: 10.1002/chem.201502977

76. E.L. Mertz, L.I. Krishtalik, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 2081.

77. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford CT, 2010.

78. V. Barone, M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995.

79. G. Scalmani, M.J. Frisch, *J. Chem. Phys.*, 2010, **132**, 114110.

80. S. Boiteux, M. Guillet, *DNA Repair (Amst.)*, 2004, **3**, 1.

81. M. Babinský, R. Fiala, I. Kejnovská, K. Bednářová, R. Marek, J. Sagi, V. Sklenář, M. Vorlíčková, *Nucleic Acids Res.*, 2014, **42**, 14031.

82. L. Harrison, K.L. Brame, L.E. Geltz, A.M. Landry, *DNA Repair (Amst*), 2006, **5**, 324.

83. K. Brandhorst, J. Grunenberg, *Chem. Soc. Rev.*, 2008, **37**, 1558.

84. J. Grunenberg, G. Barone, *RSC Adv.*, 2013, **3**, 4757.

85. K. Brandhorst and J. Grunenberg, *J. Chem. Phys.*, 2010, **132**, 184101

86. T. Y. Nikolaienko, I.S. Voiteshenko, L.A. Bulavin, D.M. Hovorun, *Bull. Univ. Kyiv Series: Phys. Math.*, 2010, **N2**, 279 (in Ukrainian).

87. L.A. Bulavin, T.Y. Nikolaienko, D.M. Hovorun, *Rep. Natl. Acad. Sci. Ukraine*, 2010, **N7**, 61.

88. R.E. Skyner, J.L. McDonagh, C.R. Groom, T. van Mourik, J.B. Mitchell, *Phys. Chem. Chem. Phys.*, 2015, **17**, 6174.

89. M. Nasr, S.R. Turk, In: Nucleosides and Nucleotides as Antitumor and Antiviral Agents (Eds. C.K. Chu, D.C. Baker), Springer Science + Business Media, 1993, New York

90. A.G. Ponomareva, Y.P. Yurenko, R.O. Zhurakivsky, T. van Mourik, D.M. Hovorun, *Curr. Phys. Chem.*, 2013, **3**, 83.

91. M. Alcolea Palafox, N. Iza, *J. Mol. Struct*., 2012, **1028**, 181.

92. H.M. Berman, *Biopolymers*, 1997, **44**, 23.

93. G.M. Blackburn, D.A. England, F. Kolkmann, *J. Chem. Soc., Chem. Commun*., 1981, 930-932.

94. R. F. W. Bader, Atoms in Molecules: A Quantum theory, Oxford University Press, New York, US, 1990.

95. A.V. Marenich, C.J. Cramer, D.G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378.

96. G. Markopoulos, J. Grunenberg, *Angew. Chem. Int. Ed*., 2013, **52**, 10648.

97. J. Grunenberg, *Nature Chem*., 2012, **4**, 154.

98. O.O. Brovarets', Y.P. Yurenko, D.M. Hovorun, *J. Biomol. Struct. Dyn*., 2015, **33**, 1624.

99. B.E. Tropp, Molecular Biology: Genes to Proteins, Jones and Bartlett Publishers, Sudbury, US, 2012.

100.   R. Aduri, B.T. Psciuk, P. Saro, H. Taniga, H.B. Schlegel, J. SantaLucia, Jr., *J. Chem. Theor. Comput*., 2007, **3**, 1464.

101.   J. Novotný, P. Kulhánek, R. Marek, *J. Phys. Chem. Lett*., 2012, **3**, 1788.

102.   Y.P. Yurenko, J. Novotný, V. Sklenář, R. Marek, *Phys. Chem. Chem. Phys*., 2014, **16**, 2072.

103.   J. Novotný, Y.P. Yurenko, P. Kulhánek, R. Marek, *Phys. Chem. Chem. Phys*., 2014, **16**, 15241.

104.   Y.P. Yurenko, J. Novotný, M.P. Mitoraj, V. Sklenář, A. Michalak, R. Marek, *J. Chem. Theor. Comput*., 2014, **10**, 5353.

105.   S. Bazzi, J. Novotný, Y.P. Yurenko, R. Marek, *Chem. Eur. J*., 2015, **21**, 9414.
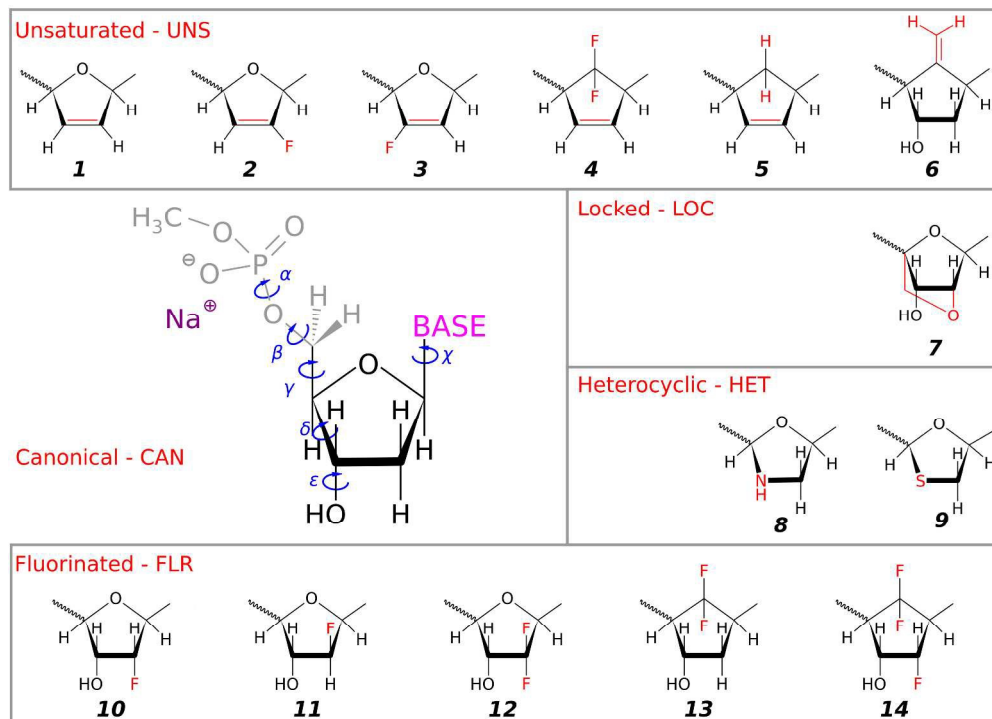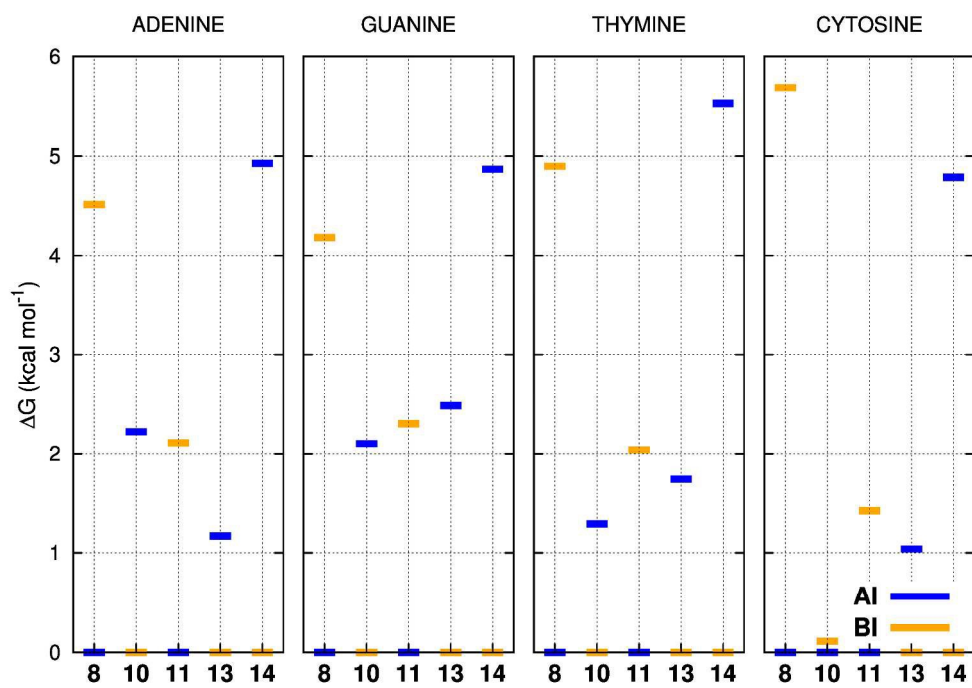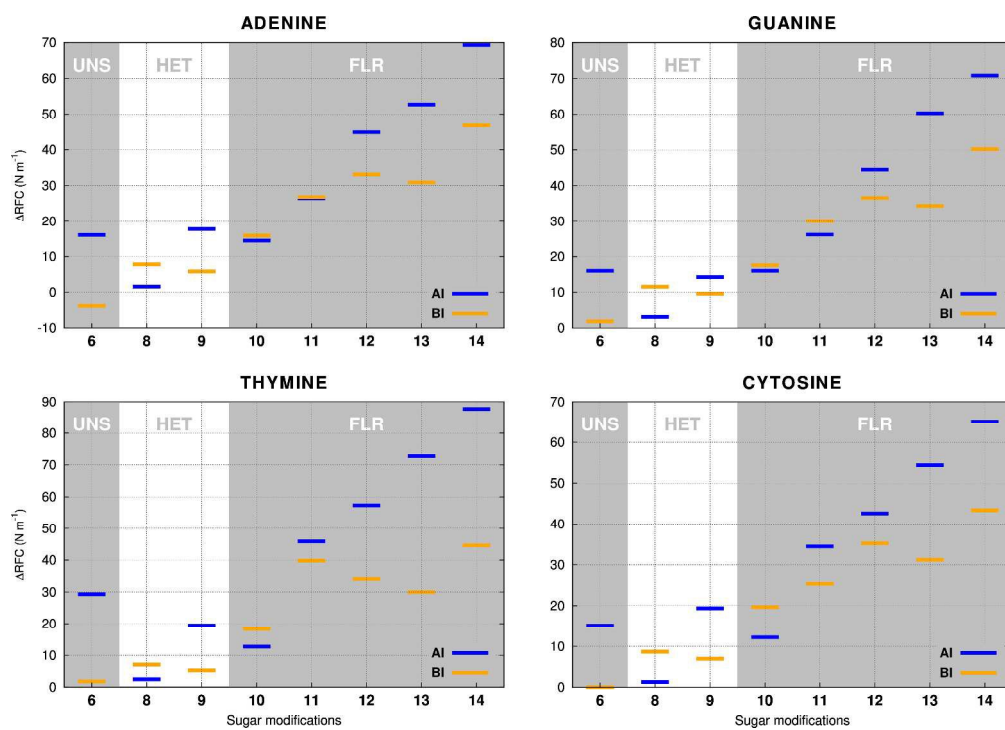
254x177mm (300 x 300 DPI)
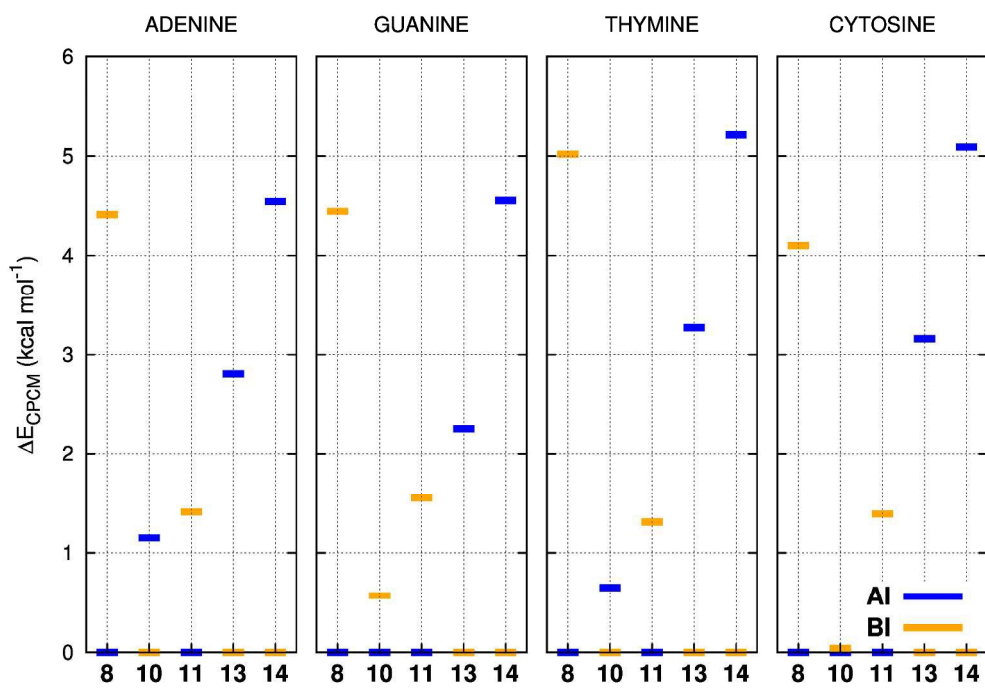
254x177mm (300 x 300 DPI)

485x344mm (300 x 300 DPI)