Biophysical Chemistry 235 (2018) 19–28

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: www.elsevier.com/locate/biophyschem

Non-canonical DNA structures: Comparative quantum mechanical study

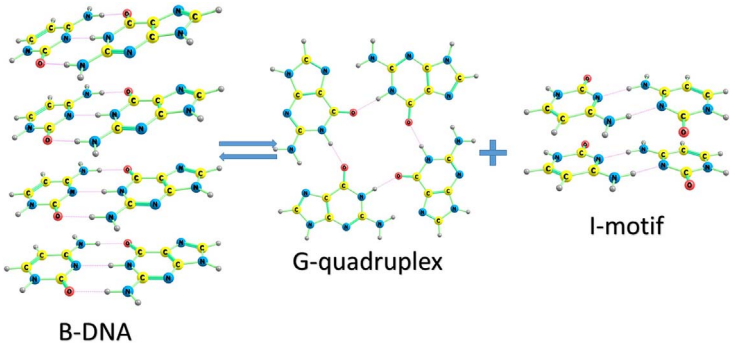
Stanislav S. Bachurin[a,*], Mikhail E. Kletskii[b], Oleg N. Burov[b], Sergey V. Kurbatov[b]

[a] FSBEI HE Rostov State Medical University of Ministry of Health of the Russian Federation, 29 Nakhichevansky st., Rostov-on-Don 344022, Russia
[b] Chemical Faculty of Southern Federal University, 7 Zorge St., Rostov-on-Don 344090, Russia

HIGHLIGHTS

- Combinations of non-canonical structures, which can compete with Watson-Crick pairs form, were identified.
- The thermodynamic stability of non-canonical DNA structures were studied by methods of quantum chemistry.
- The mechanism of the non-canonical DNA structures formation was proposed.
- The model of DNA pharmacological target was built.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Non-canonical DNA structure
G-quadruplex
Triplex
i-Motif
DFT calculations

ABSTRACT

A study of relative thermodynamic stability of non-canonical DNA structures (triplexes, G-quadruplexes, i-motifs) for the first time was conducted on the basis of quantum chemical DFT/B3LYP/6-31 + + G (d) calculations. Results of the calculations completely reproduce the experimental data on stability of G-quadruplexes comparatively Watson-Crick B-DNA. It was discovered that combinations of non-canonical DNA structures were energetically more favorable than separated nitrogenous bases. Supramolecular complexes of the non-canonical DNA structures (NSs) can be considered as a biological drug targets in gene regulation (for example in tumor therapy), in contrast to previous works, where NSs were studied independently.

1. Introduction

DNA unwinding (as a result of the replication and transcription processes) can lead to formation of structures different from the B-form founded by Watson and Crick [1]. Secondary, and respectively tertiary DNA structures stability in solutions depend on many factors, including nucleotide sequence, solvent, metal ions, low-molecular ligands, specific proteins, and especially the dynamic balance between reciprocal antagonistic enzymes - DNA gyrase (responsible for superstructure) and DNA topoisomerase I (eliminating superstructure). This complex dependence leads to structural polymorphism [2]. Properties of the nucleic-acid bases in free and hydrogen-bonded states are very different [3]. The allocation of the electric density on nitrogenous bases atoms allows forming hydrogen bonds between a large atoms range. Quite a rare non-canonical DNA structures (DNA NSs), such as i-motif [4], G-quadruplex [5], cross, hairpin [6] and triplex [7] which have a significant biological role [8], was discovered. In these structures, H-bonds form between different atoms than B-forms, and such H-bonds are called Hoogsteen H-bonds [1,4–8].

At first time, i-motif (protonated – Fig. 1 Structure 1,[1] non-

Abbreviations: NS(s), non-canonical structure(s)

* Corresponding author.
E-mail address: bachurin_ss@rostgmu.ru (S.S. Bachurin).

[1] All mentioned here structures will be discussed later on Figs. 4 and 5.

https://doi.org/10.1016/j.bpc.2018.02.003
Received 4 January 2018; Received in revised form 6 February 2018; Accepted 6 February 2018
Available online 08 February 2018
0301-4622/

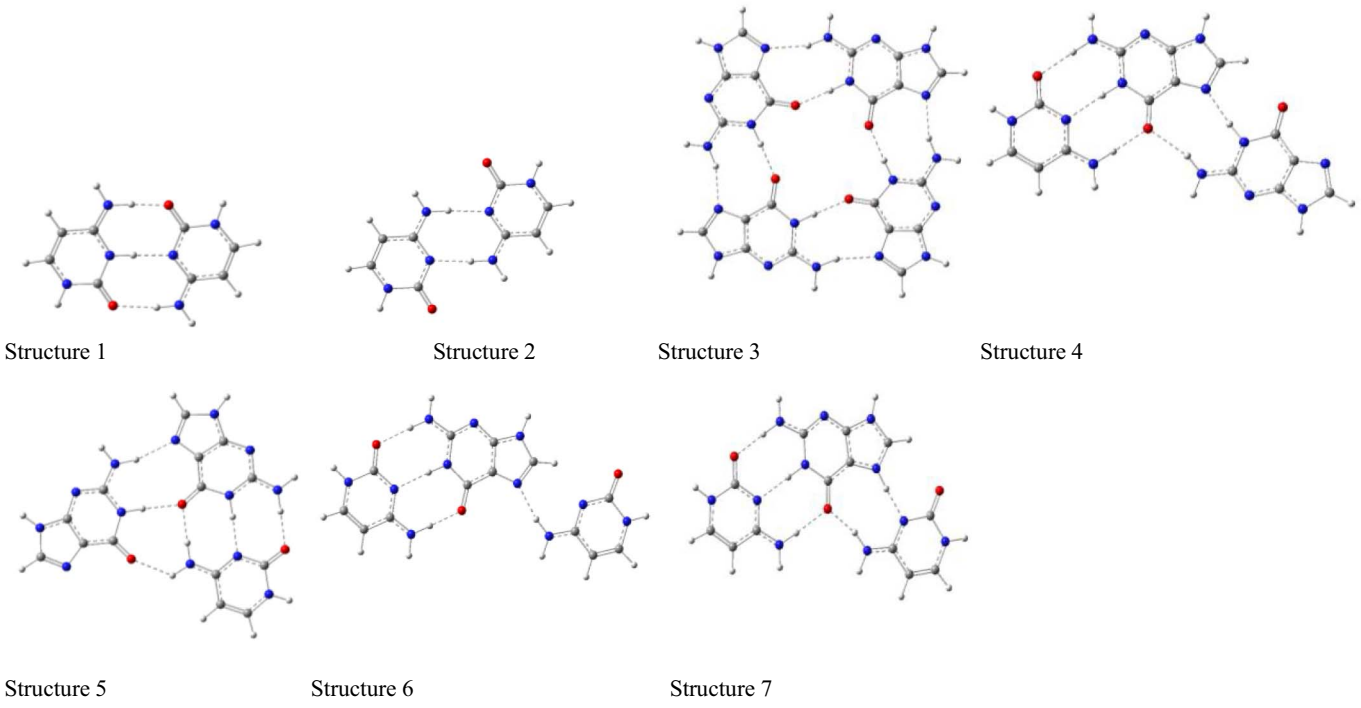


Fig. 1. The representation of all structures considered in present paper.

protonated – Structure 2) was described in [9], and it was further investigated in [10–12]. This DNA NS can be formed in cytosine-rich sequences. i-Motif is the most stable in an acidic environment, and often consists of four DNA strands held by intercalating cytosine-cytosine pairs of bases [10], in which one of the molecules of cytosine is protonated. As already noted, the stability of i-motif is highly dependent on nucleotide sequence of the DNA and the ionic environment [11]. In paper [12] it was shown that such structures as i-motif and G-quadruplex (see below), are mainly formed in solution with 100 mM of K^+ ions, although B-form of DNA is dominated in the presence of the same number of Na^+ ions. The authors of [13] were able to stabilize the i-motif of DNA in the presence of Ag^+ ions at neutral pH.

As it was discovered in [15,16], the possibility of i-motifs stabilization with different ligands [14] is considered as a promising direction in oncology due to the presence of rich in cytosine and guanine in the promoters sites of the oncogenes and telomeric DNA sites of cancer cells. Some of the ligands capable to i-motifs stabilization have already been considered as drugs [17–19].

Millimolar concentrations of guanosine and its derivatives in aqueous solutions are able to self-assemble in tetraplexes (NS G-quadruplex - Structure 3), which leads to viscous gel formation [20]. Thus Hoogsteen hydrogen bonds are formed between N_1-O_6 and N_2-N_7 atoms [21]. It is noteworthy that the stability of G-quadruplexes depends on the concentrations of metal ions (K^+, Na^+, Mg^{2+}) in solution. For example, regions with a nucleotide sequence $(TGG)_n$ form antiparallel duplexes or hairpins in low-concentrated solutions of NaCl. In the presence of K^+ at physiological pH 7.36–7.42 these sequences are stabilized in G-quadruplex structures [22,23].

Formation of DNA G-quadruplexes is interesting, because the interaction between poly(G)-oligonucleotide with complementary poly (C)-oligonucleotide is weaker [24] than the formation of the DNA NSs. G-quadruplexes, as well as the i-motifs, are considered as pharmacophore targets in tumor therapy. As mentioned above, the telomeric sections of cancer DNA are rich in cytosine and guanidine bases, and hereby, stabilization of the G-quadruplexes may lead to inhibition of telomerase activity and cancer cell death [21].

Triplex forms of nucleotide sequences are NSs shown on Structures 4–7, where three nitrogenous bases are linked by Watson-Crick and Hoogsteen H-bonds [25–27]. Triplexes are characterized by a large number of isomeric structures [28]. It is assumed that intramolecular triplexes act as molecular switches that regulate gene expression and other processes of DNA metabolism [29]. In [30,31] it was found that short oligonucleotides can be used to induce rupture of DNA strands in a specific binding site due to possibility of triplexes formation. Triplexes are also considered as pharmacological targets due to these properties [32].

The cross and hairpin are composed of similar sections - stem (consisting of a continuous sequence of complementary associated nitrogenous bases) and loop of unpaired nitrogenous bases, covalently connected with the sugar-phosphate backbone. Non-canonical hairpin structure is formed from single-stranded DNA with the ends complementary to each other. The last leads to the formation of Watson-Crick H-bonds. Two sequences of this type can form cross type of NS, when DNA chains are intertwined with each other. At the same time, two hairpins are formed at each of the opposite sides. Cross or hairpin NSs reduce the degree of tension of the DNA helix [33], and affect the interaction of DNA with proteins and even on recognition of nucleotide sequence by the corresponding enzyme [34]. Both structures are not truly self-associated NSs and their topology preferably dictated by stem per loop ratio. Due to this fact this NSs were not considered in present work.

It should be noted that even in the Watson-Crick DNA double helix there is an inequality between the nitrogenous base pairs of adenine-thymine and cytosine-guanine. Theoretical [3] and experimental [35] methods confirmed that different nitrogenous base pairs have different affinity for ligands (metal ions, drug, proteins). Both in the case of isolated DNA bases and Watson-Crick pairs, “electrostatic fingerprints” are found. They depend on the primary DNA structure. These phenomena are involved in the processes of recognition and binding of proteins with the corresponding coding regions of DNA [36]. Presumably, for the nitrogenous bases in the NS DNA conformation, the “electrostatic fingerprints” may exist, determining their biological role.

Obviously, the differences of all NS DNAs depend solely on the nature of the centers bounded by intermolecular Watson-Crick/