Sequence-specific transitions of the torsion angle gamma change the polar-hydrophobic profile of the DNA grooves: implication for indirect protein–DNA recognition

Mariia Yu. Zhitnikova, Olena P. Boryskina* and Anna V. Shestopalova

O. Ya. Usikov Institute for Radiophysics and Electronics of the National Academy of Sciences of Ukraine, Acad. Proskura Street, 12, Kharkiv, 61085, Ukraine

Communicated by Ramaswamy H. Sarma

(Received 19 November 2012; accepted 29 July 2013)

Variations of the shape and polarity of the DNA grooves caused by changes of the DNA conformation play an important role in the DNA readout. Despite the fact that non-canonical \( \text{trans} \) and \( \text{gauche} \) conformations of the DNA backbone angle \( \gamma \) (O5'–C5'–C4'–C3') are frequently found in the DNA crystal structures, their possible role in the DNA recognition has not been studied systematically. In order to fill in this gap, we analyze the available high-resolution crystal structures of the naked and complexed DNA. The analysis shows that the non-canonical \( \gamma \) angle conformations are present both in the naked and bound DNA, more often in the bound vs. naked DNA, and in the nucleotides with the A-like vs. the B-like sugar pucker. The alternative angle \( \gamma \) torsions are more frequently observed in the purines with the A-like sugar pucker and in the pyrimidines with the B-like sugar conformation. The minor groove of the nucleotides with non-canonical \( \gamma \) angle conformation is more polar, while the major groove is more hydrophobic than in the nucleotides with the classical \( \gamma \) torsions due to variations in exposure of the polar and hydrophobic groups of the DNA backbone. The propensity of the nucleotides with different \( \gamma \) angle conformations to participate in the protein–nucleic acid contacts in the minor and major grooves is connected with their sugar pucker and sequence-specific. Our findings imply that the angle \( \gamma \) transitions contribute to the process of the protein–DNA recognition due to modification of the polar/hydrophobic profile of the DNA grooves.

**Keywords:** DNA backbone; DNA conformations; structural database; protein–DNA recognition; accessible surface area

**Introduction**

Reliable recognition of DNA by proteins is an indispensable condition in the processes of DNA replication, transcription, and chromatin compaction (Svozil, Kalina, Omelka, & Schneider, 2008). Evidences are accumulating that preferential binding sequences for proteins are determined not only by the sequence-specific chemical contacts between a protein and DNA bases (base or direct readout) (Couloucheria, Pigisa, Papavassiliou, & Papavassiliou, 2007; Rohs et al., 2010; Seeman, Rosenberg, & Rich, 1976), but also by a suitable steric arrangement of the DNA fragment and/or by its propensity to adopt a deformed conformation facilitating the protein binding (shape or indirect readout) (Aeling et al., 2006; Debnath, Roy, Bera, Ghoshal, & Roy, 2013; Djuranovic & Hartmann, 2003; Garten & Laughton, 2013; Gromiha, Siebers, Selvaraj, Kono, & Sarai, 2004; Lankas, Sponer, Langowski, & Cheatham, 2003; Paillard & Lavery, 2004; Rohs et al., 2010; Rohs, West, Sosinsky, et al., 2009; Sarai & Kono, 2005; Schroeder, Roongta, Fu, Jones, & Gorenstein, 1989; Tullius, 2009; Watkins, Mohan, Koudelka, & Williams, 2010; Yamasaki, Terada, Kono, Shimizu, & Sarai, 2012).

It is well-known that the conformation of DNA in complexes with proteins often differs significantly from the canonical B-form of the double helix (Boryskina, Tkachenko, & Shestopalova, 2010, 2011; Kitayner et al., 2010; Olson, Gorin, Lu, Hock, & Zhurkin, 1998). These conformational transitions are either pre-existing (Aeling et al., 2006; Cordeiro et al., 2011; Locasale, Napoli, Chen, Berman, & Lawson, 2009; Rohs et al., 2010; Rohs, West, Liu, & Honig, 2009; Rohs, West, Sosinsky, et al., 2009; Stella, Cascio, & Johnson, 2010; Xi et al., 2011) or induced by protein binding (Little, Babic, & Horton, 2008; Otwinowski et al., 1988; Watkins et al., 2010). In this process, the sugar-phosphate backbone does not act as a passive link that just holds the bases at their positions. Instead, its inherent flexibility contributes

*Corresponding author. Email: ypboryskina@gmail.com

© 2013 Taylor & Francis
to, and limits the base mobility. The local DNA structure forms as an interplay between the optimal positions of bases and conformations of the backbone that have a free energy advantage over the scope of others, this value being sequence-specific (Becker, Wolff, & Everaers, 2006; Svozil et al., 2008; Watkins et al., 2010).

Typical re-arrangements of the DNA sugar-phosphate backbone in the protein–DNA complexes involve rotations around the four torsion angles: α (O3′–P–O5′–C5′), γ (O5′–C5′–C4′–C3′), ε (C4′–C3′–O3′–P) and ζ (C3′–O3′–P–O5′) with the possible conformations gauche+ (g+), trans (t), and gauche- (g–).

In sugars with a south B-like pucker (phase angle of pseudorotation P: 90° < P < 210°), the simultaneous transitions of the ε and ζ torsion angles from their canonical state ε/ζ: t/g– with (ε–ζ) = −90° to a non-canonical one ε/ζ: g–/t with (ε–ζ) = +90° constitute the basis for classification of the B-like nucleotides into the classical BI and alternative BII conformers (Schneider, Neidle, & Berman, 1997). The alternative BII backbone conformation, observed both in the naked (18%) and complexed DNA (11%) in crystals (Castagne, Murphy, Gronenborn, & Delpeire, 2000; Djuranovic, & Hartmann, 2003; Gorenstein, 1994; Heddi, Oguey, Lavelle, Foloppe, & Hartmann, 2010), as well as in DNA oligomers in solution (Heddi, Foloppe, Bouchemal, Hantz, & Hartmann, 2006), and hydrated DNA films (Khesbak, Savchuk, Tsushima, & Fahmy, 2011) is shown to be accompanied by a strong negative roll, a large twist, and a positive slide of the bases resulting in a large and shallow major groove and a deep and widened minor groove (Bertrand, Ha-Duong, Fermandjian, & Hartmann, 1998; Djuranovic & Hartmann, 2004; Djuranovic & Hartmann, 2003; Fratini, Kopka, Drew, & Dickerson, 1982; Hartmann, Piazzola, & Lavery, 1993; Madhumalar & Bansal, 2005; Oguey, Foloppe, & Hartmann, 2010). The importance of the BI conformation in the DNA readout is proteins is shown for several families of proteins, including repressors (trp repressor (Otwinowski et al., 1988)), transcription factors (NF-xB (Tisdé, Hartmann, & Delpeire, 1999), and DNA–DNase I (Heddi, Abi-Ghanem, Lavigne, & Hartmann, 2010). In particular, as is evidenced from both the NMR and molecular mechanics (Tisdé et al., 1999), interaction of the transcription factor NF-xB with its target DNA sequence d(CTTGGGACCTTCCAGG) is governed by the BI → BII transition in the flanking TG and CA steps that induces dynamic curvature of the DNA helix and widens the major groove. These deformations result in the greater accessibility of the hydrogen-bonding atoms in the major groove where all protein–DNA contacts are formed. The sequence-preference of the BII conformation has been studied in a number of crystallographic surveys, NMR experiments, as well as molecular dynamics simulations (Djuranovic & Hartmann, 2003; Heddi et al., 2006; Heddi, Oguey, et al., 2010; Madhumalar, & Bansal, 2005; Oguey et al., 2010; Svozil et al., 2008). These studies show that greater preference for the BII state is seen for the CG, GC, GA, CA, TG, and AA sequences, whereas the pyrimidine–pyrimidine steps disfavor it.

Much less is known about the role of transition of the α and γ angles in the DNA readout process. The transitions of these two torsion angles occur in correlation both in RNA and DNA (Pearlman & Kim, 1986; Schneider et al., 1997). The conformation of the α/γ pair is used for classification of the nucleotides with an A-like north (−30° < P < 90°) sugar pucker into the classical AI ones with α/γ: g–/t and alternative AI ones with α/γ: t/t (Haran, Joachimiak, & Sigler, 1992). The BI nucleotides are divided into the classical ones with α/γ: g+/t state and «switched BI» ones with the α/γ pair in the t/t state (Svozil et al., 2008). Recent crystallographic studies and molecular dynamic simulations show that there are also other allowed combinations of the α/γ torsion angle conformations that are present in the B-like nucleotides: g+/g– and g+/t (Djuranovic & Hartmann, 2003; Svozil et al., 2008). Conformation of the α/γ pair correlates with the inter-base helical parameters. Thus, the alternative α/γ: t/t steps usually have low twist and positive roll values (Djuranovic & Hartmann, 2003). The other major alternative α/γ: g+/g– combination does not show such effects, since it does not cause any changes in either twist or roll average values (Djuranovic & Hartmann, 2003). The unusual α/γ conformations also show a pronounced effect on the global characteristics of the DNA oligomers. Thus, the high ratio of the α/γ alternative conformations is usually associated with under-twisted (global twist ~31.2°) and moderately bent oligomers (~28° for the curvature) (Djuranovic & Hartmann, 2003).

The analysis of the crystallographic data shows that the unusual α/γ backbones are rarely encountered in free B-DNA (~2%), while they appear more frequently in the DNA complexed to proteins (~5%) (Djuranovic & Hartmann, 2003; Varnai, Djuranovic, Lavery, & Hartmann, 2002). This signals that such type of the DNA backbone re-arrangements is of importance for the proper functioning of nucleoprotein complexes. Moreover, it was shown that the backbone of a nucleosome (in particular PDB:1KX5, resolution 1.94 Å (Davey, Sargent, Luger, Maeder, & Richmond, 2002)) in addition to a fairly regular periodic alteration of the BI and BII conformers also exhibits a small amount of flipped α/γ torsions, that are present at points of direct protein/DNA contacts.

It is estimated that for the angles α and γ, the energy barriers of transition between the g and the t domains are within the range of 5–7 kcal/mol, while the g+ state for the α angle and the g– state for the γ angle
are disfavored by 5 kcal/mol (Varnai et al., 2002). It is known that the energy of deformation of the DNA sugar-phosphate backbone depends on the sequence (Fujii, Kono, Takenaka, Go, & Sarai, 2007; Lavery, 2005; Olson et al., 1998; Olson & Zhurkin, 2011). Therefore, conformational variability in the $\alpha/\gamma$ angles may constitute a molecular mechanism of indirect sequence readout in protein–DNA complexes. However, the sequence-specificity of the $\alpha/\gamma$ transitions remains the subject of the continuing research.

It was shown recently that variations of the solvent accessibility of the DNA helix, and the polar-hydrophobic profile of the DNA grooves are strong discrimination factors influencing the protein–DNA readout (Bishop et al., 2011; Greenbaum, Pang, & Tullius, 2007; Lindemose, Nielsen, Hansen, & Mollegaard, 2011; Mrinal, Tomar, & Nagaraju, 2011; Xi et al., 2011). For example, the electrostatic profile and the width of the grooves are shown to be important for the readout of A-tracts in their preferable B′-conformation (Bishop et al., 2011; Greenbaum, Pang, & Tullius, 2007; Lindemose, Nielsen, Hansen, & Mollegaard, 2011; Mrinal, Tomar, & Nagaraju, 2011; Xi et al., 2011). For electrostatic potential of the DNA. Such unusual shape and electrostatic properties of the minor groove are recognized by arginines of the proteins that make non-sequence specific hydrogen bonds with phosphates there (Rohs, West, Sosinsky, et al., 2009; Wang, Ulyanov, & Zhurkin, 2010). This phenomena is particularly important for nucleosome positioning, where the minor-groove narrowing is often associated with the presence of A-tracts, and positions of arginines of the histone core closely correspond to the narrowest minor groove (Wang, Ulyanov, & Zhurkin, 2010). The features of the DNA hydrophobic accessible surface have also been shown to be important for the protein–DNA recognition. In particular, the B → A sugar switching in the DNA-protein complexes enhances the readout mechanisms in the minor groove, where hydrogen bonds per se are insufficient for the rigorous selection of the DNA sequence. This effect is achieved by widening of the DNA hydrophobic surface accessible for contacts with proteins in the minor groove (Tolstorukov, Jernigan, & Zhurkin, 2004).

To the best of our knowledge, there has been no explicit analysis reported on the effects of $\alpha/\gamma$ backbone transitions on the polar-hydrophobic profile of the grooves and its sequence specificity as well as their role in the protein–DNA indirect recognition. For example, in a recent analysis of accessible surfaces of individual nucleotides made by Ahmad (2009) using a database of nearly 85 non-redundant DNA fragments (both naked and complexed DNA) it is shown that the accessible surface area (ASA) values of nucleotides are very much influenced by the entity of the immediate neighbors, however the nucleotides in this work have not been classified according to their sugar pucker and backbone angle values.

In the present study, we analyze a set of crystal structures of oligonucleotides and protein–DNA complexes with the aim to investigate the role of the variation in the torsion angle $\gamma$ in interactions of DNA with proteins. Specifically, we focus on the question of how the transition of the $\gamma$ angle in the DNA sugar-phosphate backbone affects the local DNA interface accessible for interaction with proteins in the minor and major grooves and evaluate the sequence dependence of such transitions.

**Materials and methods**

**Data set of crystallographic structures**

The data-set of the crystallographic coordinates of the DNA oligonucleotides (naked, i.e. non-complexed A-DNA, naked B-DNA, and protein–DNA complexes) is extracted from the Nucleic Acid Data Bank (June 2012) (Berman et al., 1992). The complete list of 226 structures selected for further analysis is shown in Supplementary Table 1.

The cutoff criteria used for the structure selection are as follows:

- The resolution of structures solved by the X-ray crystallography is better than 1.9 Å. This resolution has been previously identified as the one that ensures accurate determination of the sugar pucker and backbone torsion angles as well as statistical analysis of their distribution (Schneider et al., 1997).
- The DNA oligonucleotides are double-stranded and contain 4 and more nucleotides in each chain.
- The bases are unmodified.
- The set of structures is non-redundant, i.e. only one structure is selected from the group of complexes containing an identical protein or its mutants. Preference is given to the complexes with wild-type proteins and complexes with better resolution.

The pre-calculation treatment of the structural data includes erasing of the terminal unpaired nucleotides and splitting the oligonucleotides containing ruptures or modified bases so that the modified bases or the base pairs on both sides of the break are eliminated. In order to avoid the end effects, the terminal base pairs on both ends of each DNA structure are not included into analysis. Details of the selection process are shown in Figure 1 (a) and (b).
As a result, we have obtained three data-sets of oligonucleotides (Naked A-DNA, naked B-DNA, and complexed DNA). The composition of the final data-sets and their nucleotide content is shown in Table 1.

Calculation of the DNA structural parameters
The values of the dihedral angle $\gamma$ of the DNA backbone $(O5'-C5'-C4'-C3')$ and the phase angle of pseudorotation ($P$) of the sugar ring for each nucleotide are calculated by means of the 3DNA/compDNA analyzer (Lu & Olson, 2003).

The angle $\gamma$ values are classified as $g^+$, $t$, and $g-$ according to the classical threefold staggered pattern of the dihedrals: $g^+$ (60° ± 30°), $t$ (180° ± 30°), $g-$ (300° ± 30°). The sugar pucker is sorted into the categories of «south» B-like ($120° < P < 220°$) and «north» A-like sugars ($-60° < P < 60°$), the ranges being selected in agreement with the results of the work (Djurandjovic & Hartmann, 2003). The nucleotides with stereochemically unfavorable $P$ values outside of these domains are not included into analysis.

Revealing the atom–atom contacts in the protein–DNA binding sites
The contacting pairs of atoms in the protein–DNA binding sites are identified by means of the modified software package proposed in (Tolstorukov et al., 2004) that distinguishes contacts in the major and minor grooves. In
the case of the backbone atoms, the boundary between the grooves is defined as a plane whose normal is the vector sum of the four normal vectors to the planes defined by each three out of four atoms O4′(i), P(i), P(i+1), and O4′(i+1). See (Tolstorukov et al., 2004) for details. The atoms of bases are also divided into two groups exposed in the major (N4, N6, N7, O4, O6, C5, C6, C7, C8, and C4 in Thymine and Cytosine) and minor groove (N2, N3, O2, C2, and C4 in Adenine and Guanine). Two atoms of the protein and DNA are considered to be in contact if the distance between their centers is less than 4.5 Å (Lejeune, Delsaux, Charloateaux, Thomas, & Brasseur, 2005; Luscombe, Lascovski, & Thornton, 2001). In this way, we find atoms interacting according to all different mechanisms (hydrogen bonds, electrostatic interactions, and van der Waals interactions) in major and minor grooves, without identifying of the exact mechanisms of interaction in each contacting pair. The nucleotides which contain the atoms contacting with protein are classified as interacting nucleotides, with further subdivision into those interacting in the minor and major grooves. Those that have no contacting atoms are assigned as the non-interacting nucleotides. We have also calculated the mean numbers of contacts per interacting nucleotide for each of the four types of nucleotides in all conformations.

### Calculation of the accessible surface of DNA

The ASA of the DNA is calculated using the modified software described in Tolstorukov et al., 2004, built on the basis of the Higo and Go algorithm (Higo & Go, 1989). The principle of the algorithm is as follows: (1) the volume under study is filled with cubes with the edge size 1.5625 × 10⁻² Å; (2) for each cube it is calculated whether it is located inside, outside, or on a surface of the macromolecule; and (3) the surface of cubes located on the surface of the macromolecule gives the value of ASA. The resulting values are insensitive to orientation of the molecule in space. The following atom radii are chosen for calculations: 1.85 Å for C, 1.5 Å for N, 1.4 Å for O, and 2.0 Å for P. Hydrogen atoms are not included into analysis. The values of the atom ASA exposed in the major or minor grooves are calculated separately. Distinguishing the ASA in the grooves is done according to the same procedure as for the atom–atom contacts. The total polar and hydrophobic ASA of each nucleotide in either of groove are calculated as the sum of the ASA of the atoms O3′, O4′, O5′, O1P, O2P, N2, O3, N3, N4, N6, N7, O4, O6, and the atoms C1′, C2′, C3′, C4′, C5′, C4, C5, C6, C8, C7, and C2, correspondingly. The ratio of the total polar to the total hydrophobic ASA gives the polar/hydrophobic ratio.

### Access to the database

A complete set of structure of the protein–DNA complexes studied in this work as well as calculated DNA base pair parameters and ASA of the atoms in the minor and major grooves are freely available in our online ProtNA-ASA database (http://www.protna.bio-page.org/) (Tkachenko, Boryskina, Shestopalova, & Tolstorukov, 2010b).

### Results

The analysis of the DNA crystal structures that we have performed enabled us to determine the percentages of the γ angle conformations among the nucleotides with the B-like and A-like sugar puckers, identify the changes of the polar/hydrophobic profile of the DNA grooves cause by these backbone transitions, and assess the propensity of the nucleotides with different γ angle conformations to participate in the protein–nucleic acid contacts.

### Percentages of the angle γ conformations and the sugar puckers

The percentages of the torsion angle γ conformations in the naked and bound DNA structures are shown in Table 2. Both in the naked and bound DNA structures under study the γ angles are predominantly in the canonical state; however, the alternative states are also present. Higher content of the nucleotides with the alternative γ angle conformations is observed in the bound vs. naked DNA and in the nucleotides with the A-like sugar pucker compared to the B-like ones. The highest percentage of the angle γ flips (31.8%) is seen in the A-like nucleotides in the bound DNA structures. According to our data, the B-like nucleotides populate both the g– and t alternative conformations, flipping to the g– one more often. The A-like nucleotides adopt only the t alternative state.

In order to make sure that the calculated percentages of the γ angle conformations are not the function of the cutoff resolution of the structures used in the analysis, we have further restricted our data-set to the structures

### Table 2. Percentages of the torsion angle γ conformations and sugar pucker in the crystal structures of the naked DNA and the protein–DNA complexes.

| Sugar pucker | Naked DNA | Bound DNA |
|--------------|-----------|-----------|
|              | γ⁺ | γ⁻ | t | γ⁺ | γ⁻ | t |
| A-like       |  87.2  |  .0  | 12.8 |  68.2  |  .0  | 31.8 |
| B-like       |  99.0  |  .8  |  .2  |  93.7  |  3.7  |  2.6  |
whose resolution was better than 1.5 Å. According to Neidle (2008), this resolution guarantees a satisfactory observation of individual atoms in a map. The percentages we obtain (See Supplemental Table 2) are very close to the ones seen in the 1.9 Å data-set, however, the small number of available structures with the 1.5 Å resolution makes any statistical analysis of this data-set inconsistent. Hence, all further calculations are performed using the full data-set of structures with the 1.9 Å resolution.

We have analyzed the number of nucleotides with the flipped γ angle conformation in each structure (the data not shown for brevity). It is remarkable that in the bound DNA, there are 6 oligonucleotides where transition of the sugar to the A-form is accompanied by a 100% transition of the angle γ to the t conformation, and 19 structures whose ~50% of A-like nucleotides have their γ angles in the t conformation. Among those 6 structures, there are two nucleoproteins that contain two A-like nucleotides both in the t state. The other four oligonucleotides have a single nucleotide in A-like form and its γ angle is in the t conformation. These data suggest that a simultaneous transition to the A-like sugar pucker and the t conformation of the γ dihedral can be favorable for the protein binding. In the naked A-DNA, there are also structures containing multiple A-like t nucleotides in one DNA fragment. On the contrary, both in the naked and in bound B-DNA, there are just a few cases when several B-like nucleotides in the same DNA fragment have their γ angle flipped into the g− conformation. The B-like nucleotides with the γ angle t conformation are in the majority of cases isolated.

The percentages of various angle γ conformations for each class of base and sugar pucker are presented in Table 3. The data show that the angle γ flips from the g+ to the t conformation are favored for the A-like purines, Adenine and Guanine (with 29 and 23% in the naked DNA and 33 and 53% in the bound DNA). In B-like nucleotides, the alternative g− and t states are more popular in the pyrimidines, Cytosine and Thymine, although the overall percentages of the alternative conformations in B-like nucleotides are low. We also observe the preference for the angle γ g− state over the t one for the pyrimidine nucleotides and lack of such preference for purines.

**ASA of nucleotides exposed in the minor and major grooves**

With the aid of our algorithm, we have calculated the surface areas of each DNA atom exposed in the minor or major grooves. With this data, we have estimated the total polar and hydrophobic surface areas of the bases and sugar-phosphate backbone accessible in the minor and major grooves. The average values of ASA for the nucleotides with each of the four types of base and the angle γ in different conformations are shown in Table 4, the standard deviation of the data is given in Supplementary Table 2. Figure 2 presents the ratios of the polar and hydrophobic ASA (polar/hydrophobic ratio).

The nucleotides with the classical g+ angle γ conformation and both sugar puckers have the polar/hydrophobic ratio >1 in the major groove and ≤1 in the minor groove. The purine nucleotides, especially the ones with the A-like sugar pucker, show a higher polar/hydrophobic ratio in the major groove, due to a higher exposure of the polar surface of the bases in the Adenine and Guanine nucleotides and the hydrophobic ASA of the bases in the Cytosine and Thymine ones. In both grooves, the GC pair is more polar than the AT one in nucleotides with both sugar puckers.

The transition to the alternative γ angle conformations causes remarkable changes of the polar-hydrophobic profiles of both grooves.

In the minor groove, the angle γ flips to the t and g− alternative conformations cause an augmentation of the polar/hydrophobic ratio and its reversal to the values >1 due to a simultaneous increase of the absolute values of the polar and a decrease of the hydrophobic ASA of the DNA backbone. On the average, the polar/hydrophobic ratio is augmented by 2.4 times in the g− nucleotides

| Base       | Naked A-DNA, A-like nucleotides, % | Naked B-DNA, B-like nucleotides, % | Bound A-DNA, A-like nucleotides, % | Bound B-DNA, B-like nucleotides, % |
|------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
|            | g+  | g−  | t    | g+  | g−  | t    | g+  | g−  | t    | g+  | g−  | t    | g+  | g−  | t    |
| **Mean**   | 87  | –   | 13   | 99  | 1   | 0    | 68  | –   | 32   | 94  | 4   | 3    | 76  | –   | 16   |
| Adenine    | 71  | –   | 29   | 100 | –   | –    | 67  | –   | 33   | 96  | 2   | 2    | 67  | –   | 23   |
| Cytosine   | 98.5| –   | 1.5  | 98  | 2   | –    | 76  | –   | 24   | 90  | 6   | 4    | 76  | –   | 23   |
| Guanine    | 77  | –   | 23   | 100 | –   | –    | 47  | –   | 53   | 96  | 2   | 2    | 80  | –   | 20   |
| Thymine    | 100 | –   | –    | 97.6| 1.6 | 8    | 80  | –   | 20   | 92  | 5   | 3    | 80  | –   | 20   |
Table 4. The absolute values of the polar and hydrophobic ASA exposed in the minor and major grooves of the A-like and B-like nucleotides with different $\gamma$ angle conformations in the DNA bound to proteins ($\text{Å}^2$). The upper half of the table shows the ASA values for the sugar-phosphate backbone, and the lower half presents the ones for the bases.

| Base          | A-like nucleotides | B-like nucleotides |
|---------------|--------------------|--------------------|
|               | Major groove       | Minor groove       | Angle $\gamma$ state | Major groove | Minor groove |
|               | $g^+$  | $t$   | $g^+$  | $t$   | $g^+$  | $g^-$  | $t$   | $g^+$  | $g^-$  | $t$   |
| Sugar-phosphate backbone |    |    |    |    |    |    |    |    |    |    |
| Polar ASA,    | Mean | 45.2 | 37.9 | 36.1 | 51.5 | 48.7 | 48.2 | 44.4 | 33.8 | 50.6 | 36.7 |
| Adenine       | 46.3 | 39.0 | 36.4 | 51.7 | 49.6 | 47.8 | 40.7 | 33.4 | 52.0 | 27.1 |
| Cytosine      | 46.0 | 37.3 | 36.6 | 53.7 | 48.8 | 48.1 | 41.2 | 33.8 | 51.2 | 43.6 |
| Guanine       | 48.0 | 37.3 | 34.7 | 50.8 | 48.1 | 46.9 | 48.1 | 34.6 | 47.0 | 39.1 |
| Thymine       | 41.5 | 38.7 | 35.8 | 49.2 | 48.3 | 48.8 | 48.2 | 33.2 | 50.7 | 35.7 |
| Hydrophobic ASA, | Mean | 1.1  | 14.1 | 55.5 | 36.0 | 11.4 | 11.2 | 27.8 | 45.2 | 25.7 | 22.7 |
| Adenine       | 1.5  | 12.8 | 57.7 | 38.9 | 11.3 | 11.2 | 33.5 | 46.5 | 26.7 | 22.5 |
| Cytosine      | 1.1  | 12.7 | 54.8 | 37.9 | 10.9 | 10.6 | 28.2 | 43.9 | 24.7 | 22.7 |
| Guanine       | 1.0  | 16.0 | 57.2 | 34.5 | 12.3 | 15.0 | 22.3 | 47.0 | 27.0 | 25.0 |
| Thymine       | 0.8  | 13.3 | 53.8 | 32.3 | 10.9 | 10.4 | 26.9 | 43.0 | 25.5 | 20.9 |

| Bases         | Mean | 12.2 | 18.1 | 8.1  | 5.7  | 11.6 | 11.9 | 11.8 | 7.0  | 5.6  | 6.8  |
|---------------|------|------|------|------|------|------|------|------|------|------|------|
| Adenine       | 15.5 | 19.1 | 6.4  | 3.3  | 15.1 | 17.3 | 15.5 | 3.6  | 2.7  | 4.0  |
| Cytosine      | 10.6 | 14.7 | 6.1  | 3.7  | 10.3 | 12.2 | 8.8  | 4.3  | 3.5  | 4.1  |
| Guanine       | 20.2 | 23.5 | 12.5 | 8.5  | 14.1 | 18.6 | 17.9 | 14.1 | 16.7 | 15.5 |
| Thymine       | 7.0  | 8.8  | 9.8  | 5.5  | 5.9  | 7.0  | 6.7  | 5.5  | 4.7  | 5.4  |
| Hydrophobic ASA, | Mean | 15.2 | 18.1 | 2.2  | 1.2  | 18.4 | 26.1 | 21.2 | 1.6  | 1.0  | 2.3  |
| Adenine       | 5.5  | 8.1  | 8.4  | 4.6  | 7.9  | 8.7  | 11.8 | 4.9  | 2.7  | 6.6  |
| Cytosine      | 14.5 | 19.4 | 5.5  | 2.2  | 19.0 | 21.0 | 18.6 | 0.0  | 0.0  | 1.2  |
| Guanine       | 4.4  | 11.2 | 1.2  | 5.5  | 8.9  | 17.5 | 10.7 | 6.6  | 3.9  | 1.6  |
| Thymine       | 30.5 | 49.4 | .5   | .0   | 40.2 | 41.3 | 41.0 | .1   | .0   | .0   |

Figure 2. Average ratio of the polar/hydrophobic ASA exposed in the major (upper blue columns) and minor grooves (lower orange columns) in nucleotides with the A-like and B-like sugar puckers and the $g^+$, $g^-$ and $t$ conformations of the torsion angle $\gamma$ in the DNA bound to proteins.
with the B-like sugar pucker and by 2 times in the \textit{t} nucleotides with the B-like and A-like sugar puckers.

In the major groove, the \textit{g+} to \textit{t} flips cause an opposite effect, i.e. simultaneous oppositely directed changes of the absolute values of the polar and hydrophobic surface of the sugar-phosphate backbone that lead to the decrease of the polar/hydrophobic ratio by a factor of 1.8 and 2 in the B-like and A-like nucleotides, respectively. The most dramatic changes are observed in the case of Guanine with the A-like sugar pucker, where ASA is reduced by a factor of 5.7. The \textit{g+} to \textit{g}– flips in the B-like nucleotides do not considerably change either ASA or the polar/hydrophobic ratio in the major groove.

Same as the classical \textit{g+} nucleotides, the flipped purine ones show a higher polar/hydrophobic ratio in the major groove than the flipped pyrimidines due to the differences of the polar and hydrophobic surfaces of the bases. The GC pair in all conformations is found to be more polar in both grooves than the AT one (except for the B-like sugar pucker and \textit{g}– angle \textit{γ} conformation in the major groove).

In the naked DNA, we observe similar changes of the polar/hydrophobic profile in flipped nucleotides. However, it appears impossible to accurately determine the exact ASA of the nucleotides, as the small number of nucleotides with the alternative \textit{γ} angle conformations in our data-set give a high dispersion of the ASA values.

### Table 5. Mean values of ASA of the O3′, O5′, and C5′ atoms in the nucleotides with the A-like and B-like sugar puckers and different \textit{γ} angle conformations in the minor and major grooves in the bound DNA (Å²).

| Base | \textbf{A-like nucleotides} | | \textbf{B-like nucleotides} | |
|-------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|       | Major groove | Minor groove | Angle \textit{γ} state | Major groove | Minor groove | |
| \textbf{O3′} | \textbf{Mean} | \textbf{g+} | \textbf{t} | \textbf{g+} | \textbf{g–} | \textbf{t} | \textbf{g+} | \textbf{g–} | \textbf{t} | |
|       | .2 | .2 | 5.0 | 5.3 | .0 | .0 | .0 | 3.3 | 5.5 | 4.8 | |
| Adenine | .1 | .4 | 5.2 | 5.4 | .0 | .0 | .0 | 3.6 | 5.9 | 3.6 | |
| Cytosine | .2 | .3 | 5.0 | 5.2 | .0 | .0 | .1 | 3.1 | 5.3 | 4.8 | |
| Guanine | .3 | .0 | 5.3 | 5.6 | .0 | .0 | .0 | 3.2 | 4.6 | 5.6 | |
| Thymine | .2 | .0 | 4.8 | 4.9 | .0 | .0 | .0 | 3.2 | 5.7 | 5.4 | |
| \textbf{O5′} | \textbf{Mean} | \textbf{g+} | \textbf{t} | \textbf{g+} | \textbf{g–} | \textbf{t} | \textbf{g+} | \textbf{g–} | \textbf{t} | |
|       | .5 | .5 | 0.0 | 5.6 | .4 | 1.7 | 1.0 | .0 | 3.1 | 2.8 | |
| Adenine | .4 | .3 | 1.1 | 6.4 | .5 | 1.4 | 1.5 | .0 | 3.8 | 3.4 | |
| Cytosine | .7 | .0 | 0.0 | 8.0 | .4 | 2.3 | .7 | .0 | 2.6 | 3.1 | |
| Guanine | .6 | 1.1 | 1.1 | 3.5 | .6 | 4.4 | .5 | .1 | 4.9 | 2.2 | |
| Thymine | .3 | .0 | 0.0 | 5.8 | .2 | 1.8 | 1.4 | .0 | 2.6 | 2.4 | |
| \textbf{C5′} | \textbf{Mean} | \textbf{g+} | \textbf{t} | \textbf{g+} | \textbf{g–} | \textbf{t} | \textbf{g+} | \textbf{g–} | \textbf{t} | |
|       | .1 | 13.0 | 23.7 | 6.9 | .1 | 1.7 | 19.0 | 23.8 | 2.9 | 3.9 | |
| Adenine | .1 | 11.8 | 24.3 | 7.8 | .1 | 1.3 | 21.5 | 24.1 | 3.2 | 7.3 | |
| Cytosine | .1 | 11.7 | 23.9 | 7.7 | .1 | 1.5 | 17.9 | 23.5 | 1.9 | 3.4 | |
| Guanine | .1 | 15.0 | 24.0 | 6.2 | .1 | 3.5 | 15.7 | 24.7 | 6.3 | 3.0 | |
| Thymine | .0 | 11.9 | 23.0 | 5.5 | .1 | 1.4 | 20.8 | 22.8 | 2.4 | 2.0 | |
nucleotides stand alone, as they have the lowest values of the O5' ASA in the minor groove in the t conformation in both B-like and A-like nucleotides and the highest ones in the g− state in the B-like nucleotides. The highest values of the O5' ASA are seen in Cytosine in the A-like t conformation.

In the nucleotides with the classical γ angle state and either sugar pucker the atom C5′ is exposed mostly in the minor groove. In the flipped t ones, it traverses to the major groove, where the values of the C5′ ASA in the B-like nucleotides are nearly twice as high as the ones in the A-like nucleotides. On the opposite, the C5′ ASA in the minor groove is higher in the A-like t nucleotides. In the B-like g− nucleotides, the C5′ atom is hidden inside DNA with a minimal exposure into either

of grooves (except Guanine which has the highest values of the C5′ ASA in both grooves).

Protein–nucleic acid contacts

Figure 3 shows the percentages of nucleotides that contact with proteins as well as the average numbers of contacts per interacting nucleotide in each groove. Each two atoms of the protein and the nucleic acid are considered to have a contact if the distance between their centers is smaller than 4.5 Å.

The A-like nucleotides in the classical g+ conformation contact with proteins more actively in the minor groove, while the B-like nucleotides prefer the major groove. In addition, in both grooves the A-like nucleotides surpass the B-like ones both in the total numbers of contacts and in the percentages of contacting nucleotides. This difference is more pronounced in the minor groove.

Transition to the t state appears to be favorable for making contacts with proteins in both grooves for A-like Adenine, Guanine, and Thymine, where both the percentages of interacting nucleotides and their average numbers of contacts are higher in the flipped nucleotides than in the classical ones. It is remarkable that all flipped A-like Thymine nucleotides in our database make protein–nucleic acid contacts both in the minor and the major groove! A-like Cytosine nucleotides stand alone as the flipped nucleotides are less frequently involved in the protein–nucleic acid contacts in both grooves and make fewer contacts per interacting nucleotide in the minor groove than the classical ones.

In the B-like nucleotides, the t transitions augment the number of contacts and percentage of interacting nucleotides in the minor groove in Cytosine and Guanine, do not change these parameters in Thymine and decrease them in Adenine. In the major groove, the percentage of interacting flipped B-like t nucleotides is lower than that of the classical g+ ones. However, the number of contacts per interacting nucleotide in the flipped ones is higher. For example, in the Adenine t nucleotides the number of contacts increases by a factor of ~4.5 compared to the classical g+ ones.

Transitions of the γ angle to the g− conformation in the B-like nucleotides increases the percentage of interacting nucleotides as well as the number of protein–nucleic acid contacts in the Guanine nucleotides in both grooves. In the other g− nucleotides, these parameters noticeably decrease in the minor groove and remain close to the ones seen in the classical g+ nucleotides in the major groove.

The A-like flipped nucleotides generally show a higher percentage of interacting nucleotides as well as a greater number of protein–nucleic acid contacts in the minor groove than the B-like nucleotides with the
same angle $\gamma$ conformation. The only exception is Cytosine with the $t$ angle $\gamma$ state, where these parameters are higher in the B-like nucleotides. In the major groove, the A-like $t$ nucleotides have a higher percentage of interacting nucleotides, while the B-like ones show a greater number of contacts per interacting nucleotide.

**Discussion**

Our results show that transitions of the torsion angle $\gamma$ of the sugar-phosphate DNA backbone from the classical to alternative conformations cause remarkable changes of the polar/hydrophobic profile of the DNA surfaces exposed in the minor and major grooves. The nucleotides with the classical $g^+$ angle $\gamma$ conformation and both sugar pucker are characterized by a polar major groove and a hydrophobic minor groove. In the flipped nucleotides, the minor groove becomes more polar, while the polarity of the major groove decreases due to oppositely directed changes of the polar and hydrophobic ASA of the sugar-phosphate backbone. In particular, we have followed the changes of accessibility of the O3', O5', and C5' atoms that often interact with proteins. In the nucleotides with both alternative angle $\gamma$ conformation the accessibility of the O3' and O5' atoms in the minor groove atoms is increased while the accessible surface of the C5' atom decreases. On the contrary, in the major groove the atom C5' becomes more exposed.

The polar/hydrophobic profiles of nucleotides in the classical, as well as in alternative, angle $\gamma$ conformations have their sequence-specific features. The purine nucleotides in all conformations show a higher polar/hydrophobic ratio in the major groove than the pyrimidine ones due to more exposed polar or hydrophobic surfaces of the bases, correspondingly. Our data confirm the earlier finding of Moravec et al. that the GC pair is more polar than the AT one in both grooves, which has been received on the basis of analysis of the database of protein–DNA complexes with a lower resolution (Moravec, Neidle, & Schneider, 2002). However, we show that this is true for all conformations except for the B-like $g^-$ one, where in the major groove the ASA is more polar in the AT pair.

The analysis of the protein–nucleic acid interactions shows the sequence-specific features of the propensity of the flipped nucleotides to interact with proteins. For A-like nucleotides, we observe a pronounced increase of the number of such contacts in both grooves in all flipped $t$ nucleotides except Cytosine. It is remarkable, that 100% of the A-like $t$ Thymine nucleotides in our database interact with proteins both in the minor and major groove. The B-like $t$ conformation appears to be favorable for establishing the protein–nucleic acid contacts for Cytosine and Guanine in the minor groove, while the B-like $g^-$ one is advantageous only for Guanine for contacting with proteins in both grooves. We also observe a higher propensity of the A-like nucleotides to contact with proteins in the minor groove than the B-like ones having the same conformation, reported earlier in (Elrod-Erickson, Rolund, Nekludova, & Pabo, 1996; Lejeune et al., 2005; Tolstorukov et al., 2004). However, we show that this is not characteristic for Cytosine in the $t$ state of the $\gamma$ angle.

We propose that the nucleotides with the flipped torsion angle $\gamma$ can be indirectly recognized by proteins due to their unusual polar/hydrophobic profile. Other known mechanisms that exploit similar phenomena include the protein–DNA recognition in the minor groove by amino acids with hydrophobic side chains, due to augmentation of the available hydrophobic surface of DNA in the case of the sequence-specific B-to-A sugar repucker that usually occurs in several consequent nucleotides along the DNA helix (Tolstorukov et al., 2004). Also, enhancement of the negative electrostatic potential in the narrow minor groove of the A-tracts in their preferable B-conformation specifically attracts positively charged arginines (Parker & Tullius, 2011; Rohs et al., 2010). This mechanism is shown to be important for instance for specific binding of the Hox proteins to DNA (Joshi et al., 2007; Slattery et al., 2011), specific binding of the DNA-bending protein Fis (Stella et al., 2010), and nucleosome positioning (Olson & Zhurkin, 2011; Richmond & Davey, 2003; West, Rohs, Mann, & Honig, 2010). According to our data, in nucleosomes the number of nucleotides with the alternative $\gamma$ angle conformations is also quite high (Tkachenko, Boryskina, Shestopalova, & Tolstorukov, 2010a).

The specific role of nucleotides with the non-classical $\gamma$ angle in a process of protein–DNA recognition and binding can be different.

Firstly, despite a relatively high energetic barrier of transitions of the angle $\gamma$ to the alternative conformations, which is more than 4 kcal/mol for the GpC step (Varnai et al., 2002), small amounts of non-classical nucleotides are present in naked DNA according to our analysis as well as the data of earlier studies (Djuranovic & Hartmann, 2003; Schneider et al., 1997; Svozil et al., 2008; Varnai et al., 2002). Thus, it is possible that proteins can recognize them during the search of their target sites, just as they recognize the clusters of nucleotides with A-like pucker of the sugar ring, e.g. the GpC steps that tend to spontaneously acquire A-like sugar pucker in the naked DNA (Dickerson & Ng, 2001; Svozil et al., 2008; Vargason, Henderson, & Ho, 2001; Wu & Koudelka, 1993). The analysis of the distribution of the angle $\gamma$ conformation in nucleotides with different sugar pucker in the naked DNA has shown that the B-DNA adopts both $g^-$ and $t$ angle $\gamma$ conformations. The A-DNA, which contains a higher overall amount of alternative nucleotides, does not show the $g^-$ angle $\gamma$ transitions (see also earlier works Djuranovic et al., 2003; Varnai et al., 2002).
According to our data, the propensity of nucleotides to undergo transitions into alternative angle $\gamma$ conformation in naked DNA has some sequence-specific features, i.e. in the A-DNA, angle $\gamma$ flips appeared to be very common in purines (approximately 25% of either Adenine or Guanine in A-DNA have the $t$ conformation of the $\gamma$ angle), and atypical in pyrimidines. In B-DNA, single cases of the angle $\gamma$ flips are seen only in pyrimidines.

Secondly, we observe a considerable increase of the frequency of the angle $\gamma$ flips in the complexes compared to the naked DNA. For instance, the percentage of alternative $\gamma$ angle conformation in the B-like nucleotides in the bound DNA exceeds 20% for the nucleotides containing all four types of base pair, and in Guanine is as high as 53%. The angle $\gamma$ transitions can be induced by the process of protein binding by lowering the energy barrier, and contribute to proper positioning and fitting of the binding compounds. Here, the propensity of the DNA fragment to acquire a non-classical conformation, i.e. the energy costs of this transition can be one of the recognition criteria. A similar mechanism has been shown to play an important role in the process of selective binding of antibiotics to their target sites on rRNA. The high flexibility of the rRNA sugar-phosphate backbone and its capability to form non-standard conformations appeared to be a discriminative criterion for selection of target sites of antibiotics among thousands of similar ones (David-Eden, Mankin, & Mandel-Gutfreund, 2010). Also, the differences in plasticity of the various DNA sequences and sequence-specific energy of the A-DNA, angle $\gamma$ flips appeared to be only in pyrimidines.

To better illustrate the possible roles of the angle $\gamma$ transitions in the process of protein–DNA interaction, let us consider several complexes from our database.

**Complex of restriction endonuclease EcoRV bound to the target DNA site (1SX5)**

EcoRV restriction endonuclease recognizes a dual-symmetric 5′-GATATC-3′ target DNA site and cleaves the phosphodiester backbone at the central TA step (Figure 4 (a) and (b), stereo representation of the complex is shown in Supplementary Figure 1) (Horton & Perona, 2004). The sharp 50° bend into the major groove at this step, needed for a proper juxtaposition of the catalytic side chains and divalent metal ions among the scissile phosphate, prevents the enzyme functional groups from making discriminating hydrogen bonds with the bases. Instead, the center step is recognized primarily by

---

**Figure 4.** Complex of restriction endonuclease EcoRV bound to the target DNA site (1SX5). (a) Fragment of the crystal structure of restriction endonuclease EcoRV–DNA complex. The Adenine-containing nucleotides A4 in the opposite DNA strands are shown in ball and stick representation. The O3′, C5′, O5′ atoms of the sugar-phosphate backbone of the nucleotides A4 are highlighted by different colors and labeled as follows: red (1), orange (2), and green (3), respectively. The black line connects the centers of the base pairs and shows the sharp bend of the DNA axis. The arrows point to the cleaved phosphodiester bonds. The picture shows a fragment of protein, the important residues K119 (LYS), and N120 (ASN) of the protein chains A and B that contact with the A4 nucleotides are shown in ball and stick representation, and the atom N is marked by a blue ball with number 4. (b) Sequence of the DNA duplex; the protein target site is highlighted in green and boxed; the arrows point to the cleaved phosphodiester bonds. (c) Conformations of nucleotides defined by their ribose pseudorotation angle $P$ as A-like (A) or B-like (B), and the configuration of the $\gamma$ angle; values of ASA of the atoms C5′, O3′, O5′ in the minor and major grooves. X and Y are the DNA strands 5′–3′ and 3′–5′, respectively.

|      | Minor groove | Major groove |
|------|--------------|--------------|
| N    | P            | $\gamma$    | C5′ | O3′ | O5′ | C5′ | O3′ | O5′ |
| A4(X)| A            | $t$          | 6.2 | 6.1 | 7.9 | 13.3| 0   | 0   |
| A4(Y)| A            | $t$          | 6.6 | 5.7 | 8.0 | 12.8| 0   | 0   |
indirect readout and stabilized by numerous protein–DNA non-specific contacts (Martin, Sam, Reich, & Peroña, 1999). Adenine 4 on both strands of the binding site have flipped $t$ angle $\gamma$ conformation and the A-like sugar pucker, which makes the polar/hydrophobic profile of these nucleotides recognizable in the minor groove and contributes to the indirect readout of the target site. High surface areas of the O3$'\,$ and O5$'\,$ atoms accessible in the minor groove of the Adenine 4 nucleotides (Figure 4(c)) facilitate formation of electrostatic contacts with the nitrogen atoms of ASN (N120) and LYS (K119) (Figure 4(a)) that stabilize the fully bent state.

**Complex of restriction endonuclease ThaI bound to the target DNA site (3NDH)**

The PD-(D/E)XK type II restriction endonuclease ThaI cuts the target DNA sequence CG/CG and produces the blunt-ended fragments (Figure 5(a) and (b), stereo representation of the complex is shown in Supplementary Figure 2). The ThaI–DNA interactions are very tight. They result in major DNA distortions, i.e. doubling of the CG step rise, and drastic unwinding of the DNA helix. In many instances, the enzymes of this type explore the intrinsic, sequence-dependent DNA deformability, and thus add an indirect contribution to direct readout. The bulk of ThaI is located on the minor groove side of the specifically recognized bases, where the protein makes a single direct hydrogen bond with the DNA basis. All other direct interactions between the DNA bases and ThaI endonuclease occur on the major groove side. Among those contacts is the one of the Guanine 7 with the $\varepsilon$-amino group of Lys104 (Firczuk, Wojciechowski, Czapinska, & Bochtler, 2011). Guanine 7 has as alternative B-like $g$–conformation. Its sugar-phosphate backbone also contacts with protein, and the O5$'$ atom, whose ASA in the minor groove is relatively high due to the alternative conformation of the $\gamma$ angle (Figure 5(c)), makes electrostatic contact with the nitrogen atom of the ARG (R53) (Figure 5(a)) which contributes to stabilization of the complex.

**Complex of restriction endonuclease PvuII bound to the target DNA site (3PVI)**

PvuII is the smallest restriction endonuclease that recognizes the sequence 5$'\,$-CAGCTG-3$'$ and cleaves it between the central Guanine and Cytosine bases in both strands to generate the products with blunt ends (Figure 6(a) and (b), stereo representation of the complex is shown in Supplementary Figure 3). Recognition of the binding DNA site occurs in the major groove through two antiparallel $\beta$ strands from the sequence recognition region of the protein. Additional binding site locates in the DNA minor groove, where the side chains of GLN33–ASP34–ASN35 of PvuII form the bottom of the DNA binding cleft that span the minor groove to directly contact the sugar-phosphate backbone (Cheng, Balendir, Schildkraut, & Anderson, 1994). Among the contact that those side chains make are the electrostatic

![Figure 5](image-url)

**Figure 5.** Complex of restriction endonuclease ThaI bound to the target DNA site (3NDH). (a) Fragment of the crystal structure of the PD-(D/E) XK type II restriction endonuclease ThaI–DNA complex. The Guanine-containing nucleotides G7 in the opposite DNA strands are shown in ball and stick representation. The O3$'\,$, C5$'\,$, O5$'\,$ atoms of the sugar-phosphate backbone of the nucleotides G7 are highlighted by different colors and numbered as follows: red (1), orange (2), and green (3), respectively. The arrows point to the cleaved phosphodiester bonds. The picture shows a fragment of protein, the important residues R53 (ARG) of the protein chains A and B that contact with the nucleotides G7 are shown in ball and stick representation, the atom N is marked by a blue ball with number 4. (b) Sequence of the DNA duplex; the protein target site is highlighted in green and boxed; the arrows point to the cleaved phosphodiester bonds. (c) Conformations of nucleotides defined by their ribose pseudorotation angle P as A-like (A) or B-like (B), and the conformation of the $\gamma$ angle; values of ASA of the atoms C5$'\,$, O3$'\,$, O5$'\,$ in the minor and major grooves. X and Y are the DNA strands 5$'$$-3'\,$ and 3$'$$-5'\,$, correspondingly.
contacts of the nitrogen atom of GLN33 with the O5' atom and the nitrogen atom of ASN35 with the O5' and O3' atoms of the Thymine-containing nucleotides T8 (Figure 6(a)), where the ASAs of the O5' and O3' atoms in the minor groove are relatively high, due to the unusual A-like t conformation of these nucleotides (Figure 6(c)). The above contacts contribute to stabilization of the orientation of the catalytic region for efficient cleavage.

Conclusion

The structural variations of the sugar-phosphate backbone of DNA from its canonical B-form characterized by unusual shape, width, and polarity of the DNA grooves facilitate reliable recognition of the target DNA sites by proteins. The present work has focused on transitions of the γ backbone angle conformations and their possible role in the DNA readout process. We have performed a detailed analysis of 226 high-resolution crystal structures of naked (non-complexed) and complexed DNA structures. As a result, we have determined the percentages of the γ angle conformations among the nucleotides with the B-like and A-like sugar pucker, identified the changes of the polar/hydrophobic profile of the DNA grooves caused by these backbone transitions, and assessed the propensity of the nucleotides with different γ angle conformations to participate in the protein–nucleic acid contacts.

The unusual γ backbones are present both in the naked DNA and bound DNA structures. The incidence of the γ angle flips is higher in the bound vs. naked DNA and in the nucleotides with the A-like sugar pucker compared to the B-like ones. In the A-like nucleotides in the bound DNA structures, the percentage of flipped nucleotides is as high as 31.8%. The B-like nucleotides populate both the g− and t alternative conformations, flipping to the g− one more often. The A-like nucleotides adopt only the t alternative state. The alternative angle γ torsions are more often found in the purines with the A-like sugar pucker and in the pyrimidines with the B-like sugar conformation.

Compared to the nucleotides with the classical γ angle conformation, the flipped ones show a more polar minor groove and a more hydrophobic major groove due to variations in exposure of the polar and hydrophobic groups of the DNA backbone. The purine nucleotides with all types of the angle γ conformations show a higher polar/hydrophobic ratio in the major groove than the pyrimidines due to the differences of accessibility of the polar and hydrophobic surfaces of the bases. The GC pair in all conformations is found to be more polar in both grooves than the AT one (except for the combination of the B-like sugar pucker and g− angle γ conformation in the major groove).

The propensity of the nucleotides with different γ angle conformations to participate in the protein–nucleic acid contacts in the minor and major grooves appeared to be connected with their sugar pucker and sequence-specific. The combination of the A-like sugar pucker and the t angle γ state appeared to be favorable for all nucleotides except Cytosine in both grooves. The B-like t combination is advantageous for Cytosine and Guanine in the minor groove, and the B-like g− combination is beneficial for Guanine in the major groove.
Finally, our findings imply that the $\gamma$ angle transitions and changes of the polar/hydrophobic profile of the DNA grooves induced by them should be accounted for as the factors important for the protein–DNA indirect recognition, from the early stage discrimination of the correct target to fine-tuning of the interacting compounds during complexation.

Supplementary materials

The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2013.830579.

List of abbreviations

- DNA: Deoxyribonucleic acid
- DNase: Deoxyribonuclease
- A: Adenine
- G: Guanine
- C: Cytosine
- T: Thymine
- ASA: Accessible surface area

Acknowledgments

This work was supported by NAS of Ukraine under the grant #0107U001079 (Project MODEL, 2007–2011) and Grant #0111U010475 (Project MODEL-2, 2012–2016). The authors thank M. Tolstorukov for his development of the codes used for the structure analysis.

References

Aeling, K. A., Opel, M. L., Steffen, N. R., Tretyachenko-Ladokhina, V., Hatfield, G. W., Lathrop, R. H., & Senear, D. F. (2006). Indirect recognition in sequence-specific DNA binding by Escherichia coli integration host factor: The role of DNA deformation energy. Journal of Biological Chemistry, 281, 39236–39248.

Ahmad, S. (2009). Sequence-dependence and prediction of nucleotide solvent accessibility in double stranded DNA. Gene, 428, 25–30.

Alexeev, D. G., Lipanov, A. A., & Skuratovskii, I. Ya. (1987). Poly(dA)poly(dT) is a B-type double helix with a distinc- crown monogroove. Nature, 325, 821–823.

Becker, N. B., Wolff, L., & Everaers, R. (2006). Indirect read-out: Detection of optimized subsequences and calculation of relative binding affinities using different DNA elastic potentials. Nucleic Acids Research, 34, 5638–5649.

Berman, H. M., Olson, W. K., Beveridge, D. L., Gelbin, A., Demeny, T., Hsieh, S. H., … Schneider, B. (1992). The nucleic acid database: A comprehensive relational database of three-dimensional structures of nucleic acids. Biophysical Journal, 63, 751–759.

Bertrand, H., Ha-Duong, T., Ferrandjian, S., & Hartmann, B. (1998). Flexibility of the B-DNA backbone: Effects of local and neighbouring sequences on pyrimidine–purine steps. Nucleic Acids Research, 26, 1261–1267.

Bishop, E. P., Rohs, R., Parker, S. C., West, S. M., Liu, P., Mann, R. S., … Tullius, T. D. (2011). A map of minor groove shape and electrostatic potential from hydroxyl radical cleavage patterns of DNA. ACS Chemical Biology, 12, 1314–1320.

Borysikina, O. P., Tkachenko, M. Yu., & Shestopalova, A. V. (2010). Variability of DNA structure and protein–nucleic acid recognition. Biopolymers and Cell, 26, 360–372.

Borysikina, O. P., Tkachenko, M. Yu., & Shestopalova, A. V. (2011). Specificity of protein–nucleic acid complexes and DNA readout mechanisms. Biopolymers and Cell, 27, 3–16.

Burkhoff, A. M., & Tullius, T. D. (1987). The unusual conformation adopted by the adenine tracts in kinetoplast DNA. Cell, 48, 935–943.

Castagne, C., Murphy, E. C., Gronenborn, A. M., & Delepiere, M. (2000). 31P NMR analysis of the DNA conformation induced by protein binding SRY/DNA complexes. European Journal of Biochemistry, 267, 1223–1229.

Cheng, X., Balendiran, K., Schildkraut, I., & Anderson, J. E. (1994). Structure of PvuII endonuclease with cognate DNA. EMBO Journal, 13, 3927–3935.

Cordeiro, T. N., Schmidt, H., Madrid, C., Juárez, A., Bernadó, P., Griesinger, C., … Pons, M. (2011). Indirect DNA read-out by an H–NS related protein: Structure of the DNA complex of the C-terminal domain of Lcr. PLoS Pathogens, 7, e1002380. doi:10.1371/journal.ppat.1002380

Coulouheria, S. A., Pigisa, D. G., Papavassilioua, K. A., & Papavassiliou, A. G. (2007). Hydrogen bonds in protein–DNA complexes: Where geometry meets plasticity. Biochimie, 89, 1291–1303.

Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., & Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. Journal of Molecular Biology, 319, 1097–1113.

David-Eden, H., Mankin, A. S., & Mandel-Gutfreund, Y. (2010). Structural signatures of antibiotic binding sites on the ribosome. Nucleic Acids Research, 38, 5982–5994.

Debnath, S., Roy, N. S., Bera, I., Ghoshal, N., & Roy, S. (2013). Indirect read-out of the promoter DNA by RNA polymerase in the closed complex. Nucleic Acids Research, 41, 366–377.

Dickerson, R. E., & Ng, H. (2001). DNA structure from A to B. Proceedings of the National Academy of Sciences of the USA, 98, 6986–6988.

Djuruvicov, D., & Hartmann, B. (2003). Conformational characteristics and correlations in crystal structures of nucleic acid oligonucleotides: Evidence for sub-states. Journal of Biomolecular Structural Dynamics, 20, 771–788.

Djuruvicov, D., & Hartmann, B. (2004). DNA fine structure and dynamics in crystals and in solution: The impact of BI/ BII backbone conformations. Biopolymers, 73, 356–368.

Drew, H. R., & Dickerson, R. E. (1981). Structure of a B-DNA dodecamer. III. Geometry of hydration. Journal of Molecular Biology, 151, 535–556.

Elrod-Erickson, M., Roud, M. A., Nekludova, L., & Pabo, C. (1996). Zif268 protein–DNA complex re- fined at 1.6 Å: A model system for understanding zinc finger–DNA interactions. Structure, 4, 1171–1180.

Fireczek, M., Wojciechowski, M., Czapinska, H., & Bochtler, M. (2011). DNA intercalation without flipping in the specific Thal–DNA complex. Nucleic Acids Research, 39, 744–754.

Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982). Reversible bending and helix geometry in a B-DNA dodecamer: CCGGAATTBrCGCG. Journal of Biological Chemistry, 257, 14686–14707.
Richmond, T. J., & Davey, C. A. (2003). The structure of DNA in the nucleosome core. *Nature*, 423, 145–150.

Rohs, R., Jin, X., West, S., Joshi, R., Honig, B., & Mann, R. (2010). Origins of specificity in protein–DNA recognition. *Annual Review of Biochemistry*, 79, 233–269.

Rohs, R., West, S. M., Liu, P., & Honig, B. (2009). Nuance in the double-helix and its role in protein–DNA recognition. *Current Opinion in Structural Biology*, 19, 171–177.

Rohs, R. R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S., & Honig, B. (2009). The role of DNA shape in protein–DNA recognition. *Nature*, 461, 1248–1253.

Samanta, S., Chakrabarti, J., & Bhattacharyya, D. (2010). Changes in thermodynamic properties of DNA base pairs in protein–DNA recognition. *Journal of Biomolecular Structure & Dynamics*, 27, 429–442.

Sarai, A., & Kono, H. (2005). Protein–DNA recognition patterns and predictions. *Annual Review of Biophysics and Biomolecular Structure*, 34, 379–398.

Schneider, B., Neidle, S., & Berman, H. M. (1997). Conformations of the sugar–phosphate backbone in helical DNA crystal structures. *Biopolymers*, 42, 113–124.

Schroeder, S. A., Roongta, V., Fu, J. M., Jones, C. R., & Gorenstein, D. G. (1989). Sequence-dependent variations in the 31P NMR spectra and backbone torsional angles of wild-type and mutant Lac operator fragments. *Biochemistry*, 28, 8292–8303.

Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proceedings of the National Academy of Sciences of the USA*, 3, 804–808.

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., ... Mann, R. S. (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell*, 147, 1270–1282.

Stella, S., Cascio, D., & Johnson, R. C. (2010). The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. *Genes & Development*, 24, 814–826.

Svozil, D., Kalina, J., Onelka, M., & Schneider, B. (2008). DNA conformations and their sequence preferences. *Nucleic Acids Research*, 36, 3690–3706.

Tisné, C., Hartmann, B., & Delepierre, M. (1999). NF-kappa B binding mechanism: A nuclear magnetic resonance and modeling study of a GGG→CTC mutation. *Biochemistry*, 38, 3883–3894.

Tkachenko, M. Y., Boryskina, O. P., Shestopalova, A. V., & Tolstorukov, M. Y. (2010a). Gamma angle transitions in nucleosomal DNA backbone: Significance for protein–DNA recognition. *New Biotechnology*, 27(Suppl.), S24.