

Variability of DNA structure and protein-nucleic acid recongnition

O. P. Boryskina, M. Yu. Tkachenko, A. V. Shestopalova

A. Usikov Institute of Radio Physics and Electronics, NAS of Ukraine
12, Akademika Proskury Str., Kharkov, Ukraine, 61085

shestop@ire.kharkov.ua; ypmelezhik@yahoo.com

Revealing molecular mechanisms of sequence-specific recognition of DNA by proteins is one of the key tasks of biology. The current review presents the results of statistical analysis of the structural databases obtained by different scientific groups studying the conformational features of free and protein-bound DNA fragments that could be used for clarifying the mechanisms of protein-nucleic acid recognition. The analysis of the published data allowed us to make the following generalizations. The ability of DNA double helix to adopt alternative conformations, including the ones of sugar-phosphate backbone, is an intrinsic characteristic of certain DNA sequences. Such conformational transitions are the potential sources of formation of unique geometry of the dinucleotide steps and/or individual nucleotides and lead to alteration of base stacking and/or changes of the assessable surface area of atoms, and can be the criteria of recognition of DNA by protein as well. Changes in the physical properties that depend on the DNA structure, i. e. the polar/unpolar profile and electrostatic potential of the grooves, can also be used by protein for DNA readout.

Keywords: protein-nucleic acid recognition, variability of DNA structure, sugar-phosphate DNA backbone, alternative conformations.

Introduction. Uncovering the molecular mechanisms of sequence-specific recognition of DNA by proteins is one of the important biological tasks. It is necessary to understand the mechanisms of regulation of biological processes underlying storage, readout and transfer of genetic information as well as to control these processes [1-3]. The data on atomic structure of DNA fragments, proteins and their complexes obtained by X-ray structure analysis are widely used to study the principles of protein-nucleic acid recognition [4-6].

Rapid development of crystallographic analysis of DNA structure started in 1979, when a detailed structure of hexamer d(CGCGCG)₂ was obtained [7]. The results were quite unexpected because the double helix of the hexamer differed radically from either canonic A- or B-forms of DNA, described earlier by the diffraction analysis of DNA fibers. The structure of d(CGCGCG)₂ oligomer was a zigzag left-handed helix, named DNA Z-form. Simultaneously, the crystalline oligomers with other nucleotide sequences corresponding to A-DNA (octamer d(GGTATACC)₂)

[8] and B-DNA (dodecamer $d(\text{CGCGAATTCGCG})_2$) [9, 10] were obtained.

Dodecamer $d(\text{CGCGAATTCGCG})_2$, the binding site of *EcoRI* restrictase, thanks to its mixed nucleotide content, appeared to be a convenient object to explore the dependence of DNA helix parameters on nucleotide consequence. During the following decade, 22 variants of $d(\text{CGCGAATTCGCG})_2$ fragment obtained under different conditions of crystallization were examined in various laboratories [11]. The results showed that both the width of the minor groove and the bend of DNA helix axis were sequence-specific.

Significant narrowing of the minor groove (up to 3.5 Å) appeared to be characteristic for AT-sites [12] and not inherent in GC-fragments of DNA. As a result of small width of the minor groove, an ordered structure of hydrated water is formed on the AATT-fragments [13], so called hydration "spine", which in its turn contributes to the abnormally narrow minor groove in these regions. The conclusion about a sequence-specific character of deformation of the DNA double helix axis was made after revealing its pronounced bend (10^0 - 20^0) in the places of GC/AT junctions [10].

Dickerson and some other authors [10-13] suggested that on the basis of a thorough analysis of local helical parameters of various nucleotide sequences along with the data on the contacts between bases and amino acid residues in protein-nucleic acid complexes it would be possible to determine the principles of dependence of a double helix structure on a nucleotide sequence and to formulate the general rules for recognition of certain DNA sequences by proteins. However, up to date these rules have not been established because of high DNA conformational mobility and non-predictability of amino acid-base interactions in binding sites.

In the present decade, the mechanisms of protein-nucleic acid recognition have been studied using the statistical analysis of the data obtained by the X-ray structure analysis of DNA crystals, proteins and their complexes [14-16], which are available thanks to creation of online structural databases (PDB, NDB, Swiss Prot, etc.) [17, 18]. Besides, *in silico* experiments by the methods of numerical simulation (see [19-24] for details) allowed assessing

changes of various physicochemical parameters of protein and nucleic acid after their complexes formation. They are: the changes in surface area, accessible for the solvent [16], varying in conformational parameters and DNA packing under interaction with proteins [2-4], the deformation of relief of protein and DNA surfaces at the complex formation [25]. As a result of application of this complex approach two basic mechanisms of protein-nucleic acid recognition were identified:

1. Direct, or specific, readout implies that proteins interact with DNA sequences forming direct contacts (H-bonds) between amino acid residues and unique atomic groups of bases, (the substituents of pyrimidines in C4 position and purines in C6 and N6 positions), situated in the major groove [1, 6].

2. Indirect readout implicates that a protein binds non-specifically to the base and sugar-phosphate DNA backbone moieties. Recognition is determined by conformational features of a certain DNA fragment or its deformational ability [1, 6, 17, 26-28]. In these cases, the criteria of recognition can be: alterations in local geometry of base pairs or sugar-phosphate backbone; bends or kinks of DNA fragments; peculiarities of both major and minor grooves relief; ordered hydrated shell. Such features can pre-exist in the certain DNA fragment or appear as a result of binding with ligands (other proteins, ions, biologically active small molecules) and dehydration of the binding site, preceding the interaction.

As a result of combined analysis of structural data and physicochemical properties of the complexes, the main types of contacts between proteins and DNA were determined: electrostatic, Van der Waals forces and hydrophobic interactions [29-32]; formation of H-bonds [15, 33], including C-H...O bonds [34]; cation-π-system interactions [35]; interaction mediated by molecules of bound water [19, 36].

The main conclusions made on the basis of these investigations are the following:

1. DNA-protein complexation generally occurs via both direct and indirect mechanisms.

2. Contribution of these mechanisms to the realization of the protein-nucleic acid interaction varies depending on the protein family, DNA sequence in the binding site and other factors.

3. The relative contribution is determined by the specific protein type and the DNA fragment structure. Thus, it is difficult to formulate the general rules of protein-nucleic acid recognition [37].

In the present review, we focus our attention on the results of statistical analysis of structural databases obtained by various research groups, studying the conformation features of both naked DNA and DNA-proteins complexes. Systematization of such information will allow us to formulate more accurately the problems, appearing at description of protein-DNA interaction mechanisms, including more diverse mechanism of indirect recognition.

Variability of the DNA double helix. *DNA variability and the problem of protein-nucleic acid recognition.* DNA polymorphism and structural changeability at the formation of complexes with protein are considered to be an important biological problem. To solve this problem we need: (1) to estimate a conformation mobility of DNA structure, and (2) to compare it with the data on DNA double helix rearrangement induced by interactions with proteins. Up to date, the DNA double helix is known to retain its own structure when the complexation with various biologically significant molecules occurs. Furthermore, a considerable variability revealed for short DNA fragments, produces various conformation states characterized by high sequence specificity. It is still unclear whether the structural flexibility of DNA is a capacity attributed to certain sequences or it is stimulated by the interactions with proteins.

The capability to change the double helix structure of DNA is especially important for proteins for reliable recognition of certain DNA sites during the protein-nucleic acid complex formation [38]. In general, protein binding leads to alterations in the bases alignment and sugar-phosphate backbone structure. Such local conformational DNA rearrangement can result in bends or kinks, like in the sequence-specific CAP-DNA complex [39] or at the nucleosome formation when DNA is winding around the histones [40]. It is obvious that sequence-specific deviations of DNA structure from canonic B-form are more likely a rule than an exception and play a crucial role in the process of protein-nucleic acid recognition [24].

The fundamentals for sequence specific protein-nucleic acid recognition, analogous to the principles of complementary governing the formation of DNA double helix, have not yet been formulated [41]. Therefore, thorough investigation of the mechanisms of DNA conformational variability at the formation of complexes with proteins is an actual scientific challenge. However, the idea of a “recognition code” between amino acids and nucleotides has not been confirmed so far [42]. Impossibility to establish such a code is explained, in particular, by the existing of the vast number of degrees of freedom in the protein-DNA contact sites during formation of their complexes [29].

Forms of DNA double helix and their parameter. As stated, protein-nucleic acid recognition is directly associated with the ability of DNA molecule to change its conformation. One of the most important biological issues is how DNA conformational variability on the local level (alignment of nucleotide pairs, configuration of distinct double helix “steps” – dinucleotides, sugar-phosphate backbone conformation) influences both the global structure of DNA molecule and its interaction with proteins [3]. Therefore, the criteria are required to establish whether there are the conformational rearrangements in question and, if yes, to identify their type.

The structural differences between A- and B-DNA double helices obtained by X-ray structure analysis of the fibers, are characterized by a number of conformational parameters [43-49], namely helix parameters (an angle of helical twist between the two neighboring residues – *twist*; a distance between nucleotides along the helical axis – *rise*; a helical step – *axial rise*; an angle of pair slope – *tilt*), displacement of the base pairs relative to the helical axis (*x*-displacement), width and depth of the grooves, virtual interchain distances (d_{pp} , $d_{C1'C1'}$) (Table 1); a phase angle of sugar pseudorotation *P*, and sugar-phosphate backbone conformation.

The sugar-phosphate backbone conformation is described by preferential configurations of the sugar residues and torsion angles of the polynucleotide chains. Analysis of the nucleotide structure showed that sugar residues accept one of the two most probable conformations: C3'-endo with $0^\circ < P < 36^\circ$ (B-DNA) or

Table 1
Structural parameters of A- and B-DNA for the fibers and crystal DNA structures (X-ray structure analysis).

Parameter	A-DNA		B-DNA	
	Fibers	Crystal	Fibers	Crystal
Helical parameters				
angle of helical twist (degree)	32,7 1,2	32,5	36,0	36,5
displacement of base pairs along the axis (axial rise, z, A)	2,55	2,8	3,4	3,3
height per coil (rise A) (helical step)	28,8	–	35,6 ± 4,4	–
displacement from helical axis (x, A)	4,4 4,9	4,2	–0,14	–0,2
angle of pair slope (tilt, degree)	20,0	–	–5,9	–
Groove size, A				
major groove	8,0 ± 2,7 ^a	12,9± 2,6	17,2± 0,1 ^a	17,4
minor groove (width)	16,7± 0,1 ^a	15,8± 0,5	12,1± 0,4 ^a	10,8
major groove	13,0 ± 0,5 ^b	–	8,5 ^b	–
minor groove (depth)	2,6 ± 0,2 ^b	–	7,5 ^b	–
Alignment of base pairs				
twist (degree)	30,3	31,1	36,0	36,0
roll, (degree)	12,4	8,0	0,6	1,7
slide, (A)	-1,4	–1,53	0,45	0,23
tilt (base-pair tilt, degree)	22,6	14,6	2,8	2,1
Virtual interchain distances, A				
d _{pp}	5,5	6,0	6,6	6,7
d _{C1'C1'}	5,4	5,5	4,9	4,9
Phosphate position, A				
X _p	–0,9	–1,7	–3,0	–3,0
Y _p	8,4	8,5	8,9	8,9
Z _p	2,5	2,2	0,6	–0,4

^a[51]; ^b[45]; ^c[43]; ^d[48].

C2'-endo with $144^\circ < \text{P} < 190^\circ$ (A-DNA) [43]. Sugar switching from C3'- to C2'-endo-conformation is accompanied by alteration of nucleotide position relative to desoxyribose (Fig. 1). As a result, availability of hydrophobic atoms for the formation of contacts with proteins is different in A- and B-DNA [52].

It has been shown by the X-ray structural analysis of DNA fibers that sugar conformation [53, 54], helical parameters *rise* and *twist* [43], width and depth of the

grooves [55] are the important criteria for referring the DNA structure to either A- or B-type of double helix. Obviously, the values of those parameters are interrelated.

Different conformations of desoxyribose determine the variations in distances between phosphorus atoms of neighbouring phosphate residues in the same polynucleotide chain of A- and B-DNA double helices (Table 1, d_{pp}). Consequently, the A- and

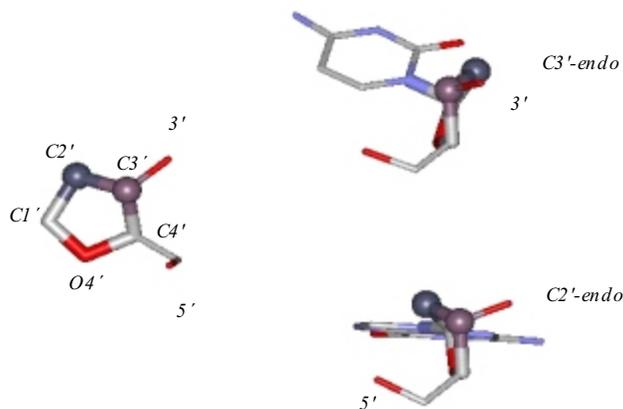


Fig. 1. Switching of desoxyribose from C3'- to C2'-endo-conformation

B-DNA helixes are twisted in a different manner that corresponds to various angles of helical twist. The *tilt* value also significantly differs for A- and B-types and, thus, correlates with the sugar conformation (C3'-endo in B-DNA and C2'-endo or equivalent C3'-exo in A-form) [43].

In addition, A- and B-DNAs differ by the values of displacement of base pairs relative to the helical axis (*x*-displacement). In B-DNA the helical axis goes through the nucleotide pairs (displacement ~ 0.2 E). Due to a larger value of displacement of base pairs from the helical axis (-4.4 μ - 4.9 E) in A-DNA, polynucleotide chains wind around the axis and form a hollow cylinder inside. Nucleotide pairs are situated on the helix periphery thus leading to the formation of a very deep and narrow major groove together with shallow and wide minor groove. In B-DNA, the grooves are less marked, their depths are approximately identical, however, they differ in width, the major groove being wider than the minor one (Table 1).

More subtle features of A- and B-DNA double helixes have been determined by the high resolution X-ray structure analysis [56]. On the level of discrete dinucleotide "steps", the DNA structure has appeared to be irregular, sufficiently depending on the nucleotide sequence, having on each "step" the characteristics different from the mean values for canonical A- and B-forms.

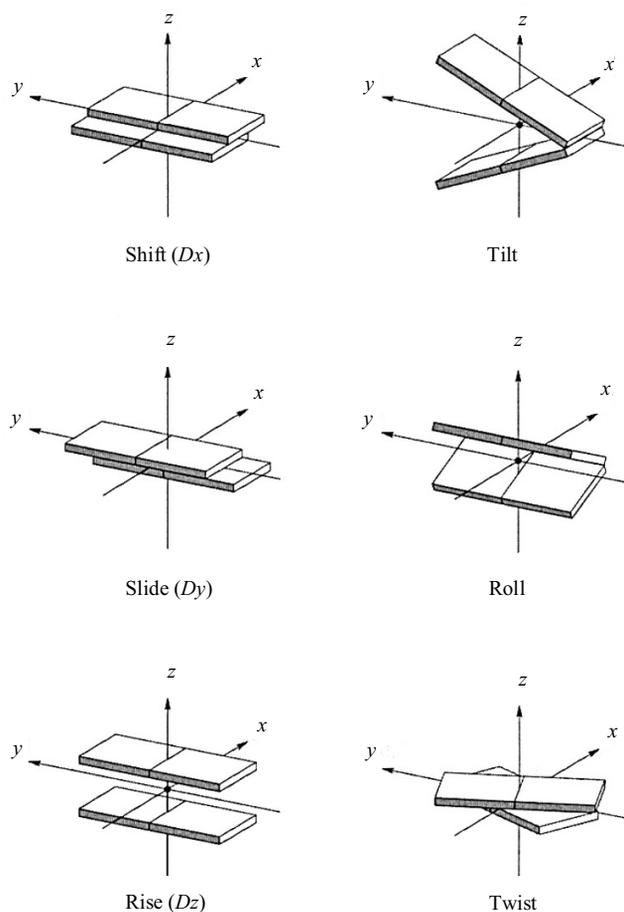


Fig. 2. Local dimer step parameters [58]

The primary data on crystal structures of A- and B-DNA revealed significant differences in three of six local dimer step parameters – twist, roll, slide (Fig. 2)[57].

The most prominent distinctions of *slide* and *roll* parameters were revealed by statistical analysis of crystal structures of A- and B-DNAs. A-DNA's *slide* is equal to $-1.57 (\pm 0.38)$ Å, *roll* to $-7.9 (\pm 5.6)^\circ$; whereas B-DNA's *slide* is equal to $-0.21 (\pm 0.074)$ Å and *roll* to $-0.2 (\pm 5.7)^\circ$ [59], i.e. the B \rightarrow A transition is accompanied by base pairs untwisting, *roll* increase and *slide* decrease (Fig. 3).

An additional discriminatory factor to identify the dinucleotide type may be a mean distance between phosphorus atoms of neighbour nucleotides along *z* axis Z_p (Table 1): for A-DNA $Z_p > 1.5$ Å, for B-DNA $Z_p < 0.5$ Å [50, 60]. Furthermore, a Z_p value correlates

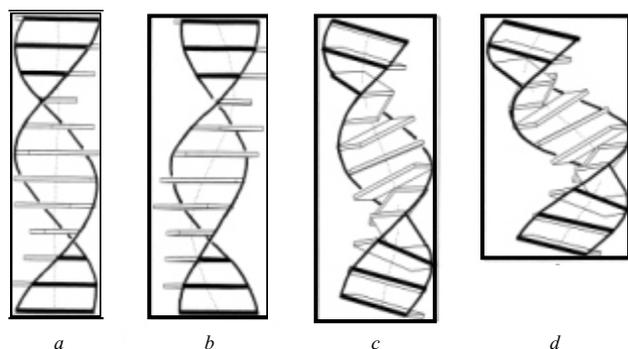


Fig. 3. Two-stage transition of B-DNA into A-form under the changes of helical parameters (slide and roll): *a* – idealized B-DNA; *b* – transition into intermediate form as a result of mutual shifting of base pairs on 1.5 Å (slide alteration); *c* – alternative intermediate form obtained by base pair rotation by 12° (roll alteration); *d* – A-type obtained by simultaneous alteration of slide and roll (base pairs displacement and rotation) [56]

with a value of angle and the coefficient of correlation is approximately equal to -1.0 .

Additionally, “A-phylic” (or “A-like”) dimer steps (GG-CC, GT-AC), including those which are found in B-DNA, can have helical parameters specific for A-DNA, i.e. negative *slide*, positive *Z_p*, large value of *roll* and lower – of *twist*. “A-phobic” (or “B-like”) dimer steps (AA-TT and GA-TC) have the same values of these parameters as those in B-form [3].

Stacking in A- and B-DNA is also different. In the double helix of B-type, stacking is generally limited to interactions between bases of the same polynucleotide chain (intrachain stacking). In A-helix bases from different chains participate in stacking as well, i.e. there are both inter- and intrachain types of stacking. This is determined by two reasons [43]. On the one hand, the angle between neighbouring nucleotides (*twist*) in A-DNA helix is smaller ($30^\circ - 32,7^\circ$) than in B-type helix ($36^\circ - 45^\circ$) that facilitates the formation of both stacking types. On the other hand, the angle of pair tilt, that is positive in A-type helix and negative in B-type, causes an increase in pair overlapping in the first case and a decrease - in the second.

DNA in complex with proteins is known to contain more A-“like” nucleotides, arranged as small clusters in one of DNA chains in the sites of contacts between DNA and proteins. It remains obscure whether such

B A transition is induced by interaction with protein or occurs independently playing a role of recognition factor for a protein [52].

Local variations of sugar-phosphate backbone also can be the “signals” at realization of indirect mechanism of protein-nucleic acid recognition and participate in structural “matching” of DNA binding site [62, 63].

It is known that conformation of sugar-phosphate backbone of double helix in direction of P O5' C5' C4' is described by torsion angles $\tau_1, \tau_2, \tau_3, \tau_4$, and τ_5 , five endocyclic torsion angles of sugar $\phi_0, \phi_1, \phi_2, \phi_3$ and angle χ , which determines the arrangement of base in relation to the sugar ring (Fig. 4).

The torsion angle values varying in the range of $180^\circ \pm 60^\circ$ ($120^\circ - 240^\circ$) are usually denoted as *trans* (*t*)-conformation, and those in the ranges of $60^\circ \pm 60^\circ$ ($0^\circ - 120^\circ$) and $300^\circ \pm 60^\circ$ ($240^\circ - 360^\circ$) - as *gauche*⁺ (*g*⁺)- and *gauche*⁻ (*g*⁻)-conformations correspondently.

Mean values of all torsion angles were determined for A- and B-DNAs (Table 2). Comparison of the results of numerous studies shows that any angle alone can not be considered as criterion of A- or B-form.

Based on the analysis of significant deviations of the torsion angle values observed in dimer steps of dodecamer $d(\text{CGCGAATTCGCG})_2$ [10], it was assumed that the prime cause of high conformational flexibility of B-DNA can be the rotations around O3'P-O5'C5' (angle τ_1) and C3'O3' – PO5' (angle τ_2). However, it was found that angle τ_1 significantly differs for dinucleotides steps ($311^\circ - 170^\circ$) while angle τ_2 changes within the range of *g*⁻ conformation ($319^\circ - 280^\circ$) (Table 2) [64].

Various authors proposed the combinations of two angles, hypothetically specific for one of the double helix forms. Significant correlations were found for pairs $\tau_1 - \tau_2, \tau_1 - \tau_3, \tau_2 - \tau_3, \tau_1 - \tau_4$ and $\tau_2 - \tau_4$ in B-DNA structure (X-ray structural analysis of dodecamer $d(\text{CGCGAATTCGCG})_2$ [64], and for pairs $\tau_1 - \tau_2, \tau_1 - \tau_3, \tau_2 - \tau_3, \tau_1 - \tau_4$ and $\tau_2 - \tau_4$ in A-DNA structure (X-ray structure analysis of tetramer $d(\text{CCGG})_2$ [53].

The values of τ_1 and τ_2 were found to correlate with sugar conformation and thus could be recommended as a criterion of A- or B-DNA [65]. This was confirmed by the analysis of numerous crystal structures and the results of molecular dynamics simulation [66, 67].

Table 2.
Values of torsion angles for A- and B-DNAs

DNA form	Angle, degree						
A-DNA ^a	308 ± 2	175 ± 3	42 ± 1	79 ± 1	212 ± 2	305 ± 2	203 ± 1
A-DNA ^b	293 ± 17	174 ± 14	56 ± 14	81 ± 7	203 ± 12	289 ± 12	199 ± 8
B-DNA ^c	314	213	36	157	155	264	262
B-DNA ^d	299 ± 8	180 ± 14	57 ± 8	122 ± 21	177 ± 25	269 ± 34	241 ± 14
B-DNA ^e	298 ± 15	176 ± 9	48 ± 11	128 ± 13	184 ± 11	265 ± 10	241 ± 8

^a [45]; ^c [44] – data for DNA fibers; ^{b, d} [64, 65]; ^e [10, 56] – data for DNA crystals.

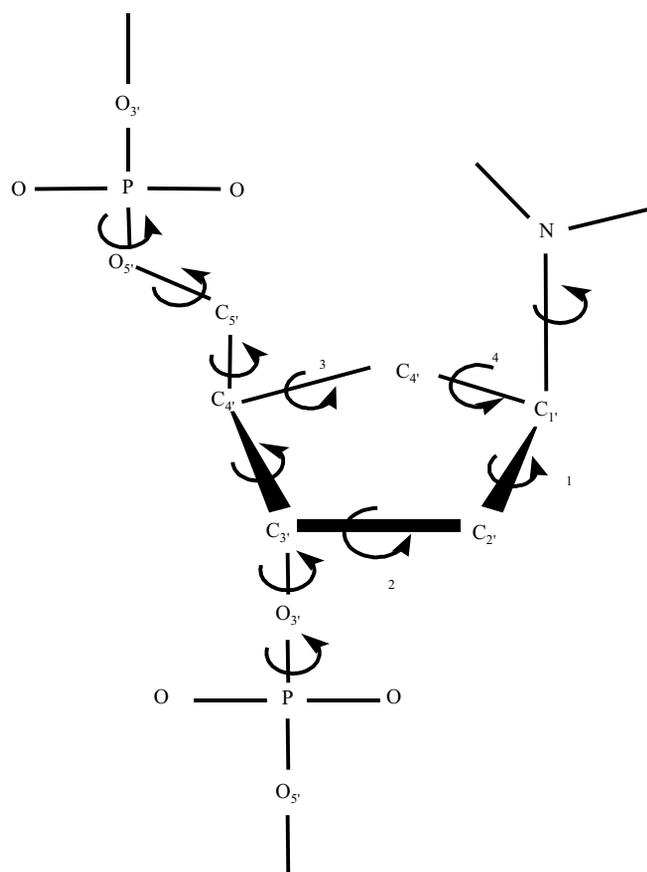


Fig. 4. Atom numbering system and torsion angle determination [43]

Subtypes of B- and A-DNA forms. The results presented in the Table 2 show a slight difference between the values of angles α and β of sugar-phosphate backbone in A- and B-DNA, which remain within t- and

g^+ -conformations respectively. Angle α is more variable ($PO_5^- - C_5' - C_4'$), and its value correlates with values of angles β and γ along the chain. Angle β associated with the sugar conformation, is the most variable, whereas the pair of angles α and β changes within the frame of two discrete states: (t, g^-) or (g^- , t). Due to the presence of two conformations in the pair α and β angles the concept of existing of two B-DNA subtypes, i.e. BI ($\alpha = t, \beta = g^-$) and BII ($\alpha = g^-, \beta = t$) was introduced.

Subtype BI corresponds to the classical B-form of DNA [68] and is characterized by sugar packing C_2' -endo with angle $\alpha = 135^\circ$, g^- conformations of torsion angles β and $\gamma + 1$ (the value of 260° is lower than in A-DNA) as well as by the values of δ typical for B-DNA ($\delta = 260^\circ$) (Table 3).

The conformations of BI and BII are different by the position of phosphate groups relative to the grooves of DNA double helix. In BI subtype, a phosphate is situated almost symmetrically to both grooves, while in subtype BII it is turned towards the minor groove (Fig. 5, a, b).

Transition between the conformations of BI and BII takes place due to simultaneous change of two torsion angles α and β . Therefore, the subtypes of B-DNA can be described by the value of difference $\alpha - \beta$: -90° for BI and $+90^\circ$ for BII [67]. In BII-DNA torsion angles β , $\gamma + 1$ and $\delta + 1$ are altered as well, however these changes usually compensate each other.

Equilibrium BI - BII is sensitive to nucleotide composition and sequence [70, 71]. Conformation BI is more often observed in pyrimidine (R)-pyrimidine (R) and purine (Y)-purine (Y) steps and more seldom in

Table 3
Mean values of torsion angles for the subtypes of A- and B-DNAs [38]

DNA subtype	Angle, degree						
BI	299	179	48	133	187	263	250
BII	293	143	46	143	251	168	278
AI	295	173	54	82	206	285	200
AII	146	192	183	85	197	289	203

combined steps YR and RY. On the other hand, the combined steps are preferable for conformation BII, in particular, for the step YR (Table 4).

Two or more BII conformers rarely follow one another, because BII conformation is accompanied by a local and global distortion of DNA structure, e.g. tetranucleotide, which possesses all four phosphate groups of the BII type, has a bend similar to that observed in complexes with proteins [71]. That is, steps BII-BII require stabilization by external forces, such as crystal packing or interaction with proteins.

The phosphate group conformation, corresponding to BII subtype, affects the helical parameters of bases [71]: the *roll* values are only negative, *twist* has high values.

Mean x-displacement values are sensitive to the number of BII dinucleotide steps in the fragment with classic B-form structure. At BII steps ~ 20% x-displacement is positive, i.e. the bases are shifted towards the major groove, whereas for the B-form the mean value of x-displacement is equal to -1.5 Å (see Table 1). The grooves of BII-DNA and BI-DNA only slightly differ by their width with the apparent tendency of opening major groove [68].

A-DNA also has two subtypes: AI, the classic A-form, and AII. For the first time an alternative conformation AII was observed for the central dinucleotide step C-G in the crystal structure of the duplex d(CCCCGGGG)₂ [72]. The conformations of the angles τ , ϕ , and ψ of the sugar-phosphate backbone of this step correspond to *t*-conformation, local *twist* is much smaller (25°) as compared to its value for the A-form (33.5°).

Mean values of all torsion angles for subtypes AI and AII are presented in Table 3. For different DNA fragments with the A-form double helix, relatively small spread of values is recorded only for three of seven sugar-phosphate backbone torsion angles - τ , ϕ , and to a lesser extent, ψ . Therefore, the criteria for AI and AII subtypes are the values of torsion angles ($O3'P-O5'C5'$) and ($O5'C5'-C4'C3'$). Thus, two conformations of A-DNA differ in the orientation of sugars relative to the bonds P-O5'-C5'-C4' [38].

Subtype AI is characterized by g^-/g^+ -conformation of torsion angles τ/ϕ , while for AII the values of τ and ϕ angles are within the *t*-region. This determines nearly planar arrangement of atoms O3'-P-O5'-C4'-C3' of the 3'-terminal nucleotides, easily accepted by purines and worse - by pyrimidines. Such values of τ and ϕ angles can be formed from classic A-form with $\tau/\phi = 300^\circ/60^\circ$ by crankshaft-like rotation around the respective bonds, which effectively compensates switching of the torsion angles so that the general direction of backbone remains practically unchanged [73, 74]. That is, the AI - AII transition is accompanied by a correlated change in torsion angles in the direction of P - O5' - C5' - C4' that leads to a relative reorientation of desoxyriboses with the bases remaining almost unchanged (Fig. 5, c, d) [61].

Equilibrium AI - AII as well as BI - BII depends on the nucleotide composition and base sequence, i.e. AII subtype is more common for YR-steps and less frequent for steps RY and YY (Table 4) [38].

Nonclassical conformations of DNA sugar-phosphate backbone in the sites of protein binding. According to the data of numerous up to date studies on DNA fragment and protein-nucleic acid complex crystallography, the number of nucleotides with nonclassical conformations of sugar-phosphate backbone is essentially higher in the DNA-protein complexes than in the naked DNA. For example, the frequency of alter native conformations of the torsion angle τ in DNA interacting with proteins is increased approximately by 5 times as compared to naked B-DNA [75, 76].

Moreover, A- and B-forms as well as BI- and BII-conformations in naked DNA and DNA complexed with proteins differ not only in conformation of sugar-phosphate backbone but also in

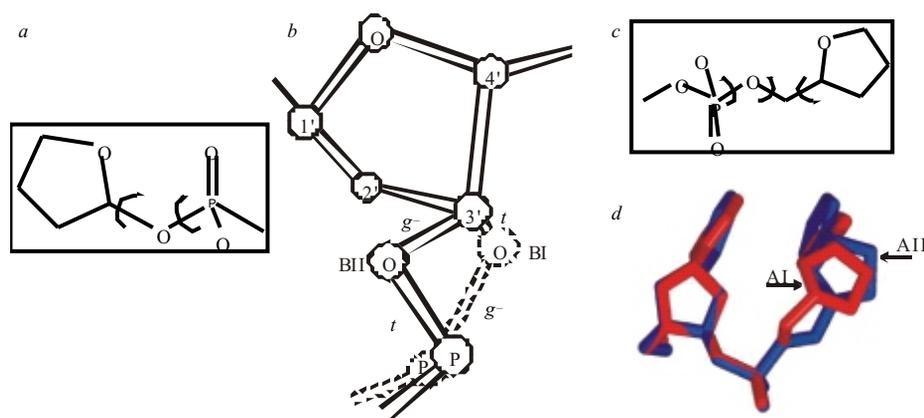


Fig. 5. Diagram of two main conformations of B-DNA: *a* – subtypes BI and BII differ in orientation relative to the C3'-O3'-O5'-P-bond determined by torsion angles τ and ϕ [61]; *b* – in BI conformation (dotted line) angles τ , ϕ have (*t*, *g*⁻)-conformations; in BII subtype angles τ , ϕ have (*g*⁺, *t*)-conformations respectively [64]; *c* – superposition of A-DNA subtypes corresponding to correlated bond restructuring in sugar-phosphate backbone: subtypes AI and AII differ in orientation relative to the PO5'-C5'-C4'-bond determined by the angles τ , ϕ , and ψ (d) [61].

Table 4
Frequency of occurrence of BI, BII, and AII conformations (%) for the steps RR, RY, YR, and YY for the pure DNA [38].

DNA subtype	RR	RY	YR	YY
BI	32,94	16,17	20,29	30,59
BII	20,11	25,44	54,43	0
AI	17,14	30,71	29,64	22,5
AII	32,72	1,8	63,63	1,8

the values of helical parameters *roll*, *twist* and *tilt* (Table 5).

Further we present several typical examples of DNA–protein complexes where nucleotides with nonclassical conformations of sugar-phosphate backbone can be used for DNA readout by proteins.

Transition of the central dinucleotide step C-G of duplex d(CCCCGGGG)₂ to an alternative conformation AII can be used as a signal for recognition by the protein of its “own” binding site, for example, by restrictase *MSPI*, which precisely cleaves the fragment CCGG independently on the nucleotide composition of neighbouring sequences [72].

Protein BPV-1 E2 recognizes DNA sequence (hexamer d(GACGTC)₂) by the bent toward the minor groove and by the BII conformation of the central C-G step [77]. It was found [76] for the complex of insect heterodimer nuclear receptor consisting of ecdysone receptor (EcR) and ultraspiracle (USP) (PDB index 1R00 [78]) having palindrome duplex d(5'-AGGTCAATGACCT-3') as a binding site, that

adenine at the starting point of binding site had an A-like desoxyribose packing and *t*-conformation of torsion angle τ (Fig. 6). For most of the similar complexes (the mammal steroid and non-steroid hormone receptors are bound to DNA just according to this principle), adenine at the start point has B-like desoxyribose packing and *g*⁺-conformation of angle τ typical for the classic DNA B-type [79]. Probably, adenine with alternative conformation can serve as a signal for the proteins involved in formation of this highly specific complex to recognize their binding sites.

Nonclassical conformations of the sugar-phosphate backbone are very common for the nucleosome structure (in particular, nucleosome 1KX5 [36]), where regular interchange of BI- and BII-conformations are observed. In the sites of direct protein-nucleic acid interactions these conformations are transformed to more deformed B-conformers characterized by switching of torsion angles $\tau + 1$ and $\phi + 1$ (“switched” BI-DNA) and by the presence of nucleotides with a wide dispersion of the values of τ and $\phi + 1$. In this case, it can be supposed that nucleotides with alternative conformations of sugar-phosphate backbone in nucleosome are essential not only for ensuring optimal conditions for interactions with histones but also for a possibility of other proteins to recognize indirectly their “own” sites on the nucleosome DNA.

Summarizing, the following conclusion could be made. The ability of DNA double helix, including sugar-phosphate backbone, to acquire alternative

Table 5
Mean values of helical parameters A- and B-DNAs [69] and dinucleotide steps with BI- and/or BII-conformations [71]

DNA type	Twist, degree	Rise, Å	Roll, degree	Tilt, degree
A-DNA	31,8	2,9	1,2	-0,2
Pure B-DNA	34,9	3,4	2,8	0,6
Protein-bound B-DNA	32,1	3,4	-	-
Pure B-DNA				
BI/BI	34,3	3,3	3,9	1,1
BI/BII	37,8	3,4	-3,9	0,4
BII/BII	46,8	3,4	-7,7	-1,9
Protein-bound B-DNA				
BI/BI	31,4	3,4	4,9	0,6
BI/BII	37,2	3,4	-2	0,7
BII/BII	44	3,5	-8,7	-0,8

conformations, can be considered as an “intrinsic” property of certain DNA sequences. Such conformational changes can result in the peculiarities of unique geometry of various dimer steps and/or individual nucleotides, thus leading to disturbances of the base stacking and/or alteration of the accessible surface area of atoms and therefore can be the criteria of protein-DNA recognition. Conformational transitions induce changes in some structure-dependent physical properties, i.e. polar/non-polar profile and electrostatic potential of both minor and major grooves, which can also be used by proteins for specific DNA readout [15, 75]. The latter appeared to be important for the Hox-group proteins functioning [80]. However, further studies are required for understanding general principles of the indirect DNA readout by proteins.

Е. П. Борискина, М. Ю. Ткаченко, А. В. Шестопалова

Полиморфизм ДНК и белково-нуклеиновое узнавание

Резюме

Расшифровка молекулярных механизмов узнавания белками конкретных последовательностей ДНК – одна из основных задач современной биологии. Настоящий обзор посвящен результатам статистического анализа структурных баз данных,

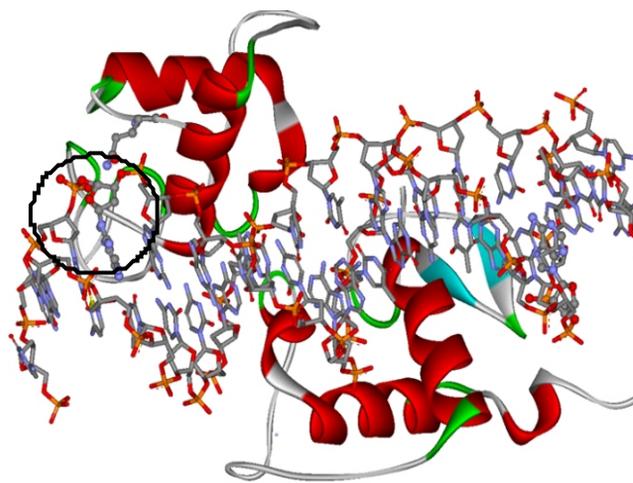


Fig. 6. A-like packing of desoxyribose and *t*-conformation of angle in adenine at the starting point of protein binding site in a complex of heterodimer receptor with duplex d(5' -AGGTCAATGACCT-3') (PBD index 1R0O) [78]

полученных различными исследователями при изучении конформационных особенностей фрагментов свободной ДНК и ДНК в комплексах с белками, которые могут быть полезными при выяснении механизмов белково-нуклеинового узнавания. Анализ литературных данных позволил сделать следующие обобщения. Способность двойной спирали ДНК принимать альтернативные конформации, в том числе сахарофосфатным остовом, является «внутренним» свойством, присутствующим определенным последовательностям ДНК. Такие конформационные перестройки могут служить потенциальными источниками проявления уникальных особенностей геометрии различных динуклеотидных шагов и/или отдельных нуклеотидов и нарушать стэкинг оснований и/или изменять площадь доступной поверхности атомов и быть критериями для узнавания участка ДНК белком. Изменения физических свойств, зависящих от структуры ДНК, таких как полярно-неполярный профиль и электростатический потенциал желобков, также могут использоваться белками для узнавания конкретного сайта ДНК.

Ключевые слова: белково-нуклеиновое узнавание, полиморфизм ДНК, сахарофосфатный остов ДНК, альтернативные конформации.

О. П. Борискина, М. Ю. Ткаченко, Г. В. Шестопалова

Полиморфизм ДНК і білково-нуклеїнове впізнавання

Резюме

Розшифрування молекулярних механізмів упізнавання білками конкретних послідовностей ДНК – одна із основних задач сучасної біології. Огляд присвячено результатам статистичного аналізу структурних баз даних, отриманих різними групами дослідників при вивченні конформаційних особливостей фрагментів вільної ДНК і ДНК у комплексах з білками. Такі дані використовують при дослідженні механізмів

білково-нуклеїнового впізнавання. Аналіз літературних джерел дозволив зробити наступні узагальнення. Здатність подвійної спіралі ДНК набувати альтернативних конформацій, у тому числі сахарофосфатним остовом, є «внут- рішньою» властивістю, яка притаманна визначеним послідовностям ДНК. Подібні конформаційні перебудови можуть слугувати потенційним джерелом прояву унікальних особливостей геометрії різних динуклеотидних фрагментів і/або окремих нуклеотидів та призводити до порушення стекінгу основ і/або змін площ доступної поверхні атомів, а також бути критеріями впізнавання білками специфічних послідовностей ДНК. Зміни фізичних властивостей, які залежать від структури ДНК, серед яких полярно-неполярний профіль і електростатичний потенціал борозенок, також можуть використовуватися білками для впізнавання кон- кретного сайта ДНК.

Ключові слова: білково-нуклеїнове впізнавання, поліморфізм ДНК, сахарофосфатний остов ДНК, альтернативні конформації.

REFERENCES

- Gromiha M. M., Siebers J. G., Selvaraj S., Kono H., Sarai A. Intermolecular and intramolecular readout mechanisms in protein-DNA recognition // *J. Mol. Biol.*—2004.—**337**, N 1.—P. 285–294.
- Steitz T. A. Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding // *Q. Rev. Biophys.*—1990.—**23**, N 3.—P. 205–280.
- Marathe A., Karandur D., Bansal M. Small local variations in B-form DNA lead to a large variety of global geometries which can accommodate most DNA-binding protein motifs // *BMC Struct. Biol.*—2009.—**9**.—P. 24.
- Kono H., Sarai A. Structure-based prediction of DNA target sites by regulatory proteins // *PROTEINS: Structure, Function, and Genetics.*—1999.—**35**, N 1.—P. 114–131.
- Bernstein F. C., Koetzle T. F., Williams G. J., Meyer E. F. Jr., Brice M. D., Rodgers J. R., Kennard O., Shimanouchi T., Tasumi M. The Protein Data Bank: a computer-based archival file for macromolecular structures // *J. Mol. Biol.*—1977.—**112**, N 3.—P. 535–542.
- Sarai A., Kono H. Protein-DNA recognition patterns and predictions // *Annu. Rev. Biophys. Biomol. Struct.*—2005.—**34**.—P. 379–398.
- Wang A. H., Quigley G. J., Kolpak F. J., Crawford J. L., van Boom J. H., van der Marel G. A., Rich A. Molecular structure of a lefthanded double helical DNA fragment at atomic resolution // *Nature.*—1979.—**282**, N 5740.—P. 680–686.
- Shakke Z., Rabinovich D., Kennard O., Cruse W. B., Salisbury S. A., Viswamitra M. A. Sequence-dependent conformation of an A-DNA double helix. The crystal structure of the octamer d(G-G-T-A-T-A-C-C) // *J. Mol. Biol.*—1983.—**166**, N 2.—P. 183–201.
- Wing R., Drew H., Takano T., Broka C., Tanaka S., Itakura K., Dickerson R. E. Crystal structure analysis of a complete turn of B-DNA // *Nature.*—1980.—**287**, N 5784.—P. 755–758.
- Dickerson R. E., Drew H. R. Structure of a B-DNA dodecamer. II. Influence of base sequence on helix structure // *J. Mol. Biol.*—1981.—**149**, N 4.—P. 761–786.
- Dickerson R. E. What do we really know about B-DNA? // *Structure and methods DNA and RNA: Proceedings of the Sixth Conversation in Biomolecular Stereodynamics* / Eds R. H. Sarma, M. H. Sarma.—New York: Adenine press, 1990.—Vol. 3.—P. 1–38.
- Drew H. R., Dickerson R. E. Structure of a B-DNA dodecamer. III. Geometry of hydration // *J. Mol. Biol.*—1981.—**151**, N 3.—P. 535–556.
- Kopka M. L., Fratini A. V., Drew H. R., Dickerson R. E. Ordered water structure of a B-DNA dodecamer. A quantitative study // *J. Mol. Biol.*—1983.—**163**, N 1.—P. 129–146.
- Jones S., van Heyningen P., Berman H. M., Thornton J. M. Protein-DNA interactions: A structural analysis // *J. Mol. Biol.*—1999.—**287**, N 5.—P. 877–896.
- Luscombe N. M., Laskowski R. A., Thornton J. M. Amino acids-base interaction: a three-dimensional analysis of protein-DNA interactions at an atomic level // *Nucl. Acids Res.*—2001.—**29**, N 13.—P. 2860–2874.
- Ahmad S., Gromiha M. M., Sarai A. Analysis and prediction of DNA-binding proteins and their binding residues based on composition, sequence and structural information // *Bioinformatics.*—2004.—**20**, N 4.—P. 477–486.
- Berman H. M. The Protein Data Bank: a historical perspective // *Acta Crystallogr. A.*—2008.—**64**, pt 1.—P. 88–95.
- Berman H. M., Olson W. K., Beveridge D. L., Westbrook J., Gelbin A., Demeny T., Hsieh S. H., Srinivasan A. R., Schneider B. The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids // *Biophys. J.*—1992.—**63**, N 3.—P. 751–759.
- Reddy C. K., Das A., Jayaram B. Do water molecules mediate protein-DNA recognition? // *J. Mol. Biol.*—2001.—**314**, N 3.—P. 619–632.
- Ahmad S., Keskin O., Sarai A., Nussinov R. Protein-DNA interactions: structural, thermodynamic and clustering patterns of conserved residues in DNA-binding proteins // *Nucl. Acids Res.*—2008.—**36**, N 18.—P. 5922–5932.
- Wellenzohn B., Flader W., Winger R. H., Hallbrucker A., Mayer E., Liedl K. R. Complex of B-DNA with polyamides freezes DNA backbone flexibility // *J. Am. Chem. Soc.*—2001.—**123**, N 21.—P. 5044–5049.
- Wellenzohn B., Flader W., Winger R. H., Hallbrucker A., Mayer E., Liedl K. R. Indirect readout of the trp-repressor-operator complex by B-DNA's backbone conformation transitions // *Biochemistry.*—2004.—**41**, N 12.—P. 4088–4095.
- Hartmann B., Sullivan M. R., Harris L. F. Operator recognition by the phage 434 cI repressor: MD simulations of free and bound 50-bp DNA reveal important differences between the OR1 and OR2 sites // *Biopolymers.*—2003.—**68**, N 2.—P. 250–264.
- Rohs R., West S. M., Liu P., Honig B. Nuance in the double-helix and its role in protein-DNA recognition // *Curr. Opin. Struct. Biol.*—2009.—**19**, N 2.—P. 171–177.
- Pabo C. O., Nekludova L. Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition // *J. Mol. Biol.*—2000.—**301**, N 3.—P. 597–624.
- Lankas F., Sponer J., Langowski J., Cheatham T. E. 3rd. DNA basepair step deformability inferred from molecular dynam-

- mics simulations // *Biophys. J.*—2003.—**85**, N 5.—P. 2872–2883.
27. *Dickerson R. E.* DNA bending: the prevalence of kinkiness and the virtues of normality // *Nucl. Acids Res.*—1998.—**26**, N 8.—P. 1906–1926.
 28. *Aeling K. A., Opel M. L., Steffen N. R., Tretyachenko-Ladokhina V., Hatfield G. W., Lathrop R. H., Senear D. F.* Indirect recognition in sequence-specific DNA binding by *E. coli* integration host factor: the role of DNA deformation energy // *J. Biol. Chem.*—2006.—**281**, N 51.—P. 39236–39248.
 29. *Ahmad S., Sarai A.* Moment-based prediction of DNA-binding proteins // *J. Mol. Biol.*—2004.—**341**, N 1.—P. 65–71.
 30. *Nadassy K., Wodak S. J., Janin J.* Structural features of protein-nucleic acid recognition sites // *Biochemistry.*—1999.—**38**, N 7.—P. 1999–2017.
 31. *Matthew J. B., Ohlendorf D. H.* Electrostatic deformation of DNA by a DNA-binding protein // *J. Biol. Chem.*—1985.—**260**, N 10.—P. 5860–5862.
 32. *Lejeune D., Delsaux N., Charlotiaux B., Thomas A., Bressan R.* Protein-nucleic acid recognition: statistical analysis of atomic interactions and influence of DNA structure // *Proteins.*—2005.—**61**, N 2.—P. 258–271.
 33. *Mandel-Gutfreund Y., Schueler O., Margalit H.* Comprehensive analysis of hydrogen bonds in regulatory protein–DNA complexes: in search of common principles // *J. Mol. Biol.*—1995.—**253**, N 2.—P. 370–382.
 34. *Mandel-Gutfreund Y., Margalit H., Jernigan R. L., Zhurkin V. B.* A role for CH...O interactions in protein–DNA recognition // *J. Mol. Biol.*—1998.—**277**, N 5.—P. 1129–1140.
 35. *Gromiha M. M., Santhosh C., Ahmad S.* Structural analysis of cation- π interactions in DNA binding proteins // *Int. J. Biol. Macromol.*—2004.—**34**, N 3.—P. 203–211.
 36. *Davey C. A., Sargent D. F., Luger K., Maeder A. W., Richmond T. J.* Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution // *J. Mol. Biol.*—2002.—**319**, N 5.—P. 1097–1113.
 37. *Prabakaran P., Siebers J. G., Ahmad S., Gromiha M. M., Singarayan M. G., Sarai A.* Classification of protein–DNA complexes based on structural descriptors // *Structure.*—2006.—**14**, N 9.—P. 1355–1367.
 38. *Svozil D., Kalina J., Omelka M., Schneider B.* DNA conformations and their sequence preferences // *Nucl. Acids Res.*—2008.—**36**, N 11.—P. 3690–3706.
 39. *Schultz S. C., Shields G. C., Steitz T. A.* Crystal structure of a CAP–DNA complex: the DNA is bent by 90 degrees // *Science.*—1991.—**253**, N 5023.—P. 1001–1007.
 40. *Luger K., Mader A. W., Richmond R. K., Sargent D. F., Richmond T. J.* Crystal structure of the nucleosome core particle at 2.8 Å resolution // *Nature.*—1997.—**389**, N 6648.—P. 251–260.
 41. *Matthews B. W.* No code for recognition // *Nature.*—1988.—**335**, N 6188.—P. 294–295.
 42. *Seeman N. C., Rosenberg J. M., Rich A.* Sequence specific recognition of double helical nucleic acids by proteins // *Proc. Nat. Acad. Sci. USA.*—1976.—**73**, N 3.—P. 804–808.
 43. *Saenger W.* Principles of nucleic acids structure.—M.: Mir, 1987.—583 p.
 44. *Arnott S., Hukins D. W.* Refinement of the structure of B-DNA and implications for the analysis of X-ray diffraction data from fibers of biopolymers // *J. Mol. Biol.*—1973.—**81**, N 2.—P. 93–105.
 45. *Chandrasekaran R., Wang M., He R.-G., Puigianer L. G., Byler M. A., Millane R. P., Arnott S.* A re-examination of the crystal structure of A-DNA using fiber diffraction data // *J. Biomol. Struct. Dyn.*—1986.—**6**, N 6.—P. 1189–1202.
 46. *Arnott S., Hukins D. W.* Optimised parameters for A-DNA and B-DNA // *Biochem. Biophys. Res. Commun.*—1972.—**47**, N 6.—P. 1504–1509.
 47. *Dickerson R. E.* Definitions and nomenclature of nucleic acid structure components // *Nucl. Acids Res.*—1989.—**17**, N 5.—P. 1797–1803.
 48. *Sinden R. R.* DNA structure and function.—New York: Acad. press, 1994.—398 p.
 49. *Rich A., Norheim A., Wang A. H. J.* The chemistry and biology of left-handed Z-DNA // *Ann. Rev. Biochem.*—1984.—**53**.—P. 791–846.
 50. *Lu X.-J., Shakked Z., Olson W. K.* A-form conformational motifs in ligand-bound DNA structures // *J. Mol. Biol.*—2000.—**300**, N 4.—P. 819–840.
 51. *van Dam L., Levitt M. H.* BI Nucleotides in the B and C forms of natural-sequence polymeric DNA: a new model for the C form of DNA // *J. Mol. Biol.*—2000.—**304**, N 4.—P. 541–561.
 52. *Tolstorukov M. Y., Jernigan R. L., Zhurkin V. B.* Protein–DNA hydrophobic recognition in the minor groove is facilitated by sugar switching // *J. Mol. Biol.*—2004.—**337**, N 1.—P. 65–76.
 53. *Conner B. N., Yun C., Dickerson J. L., Dickerson R. E.* Helix geometry and hydration in an A-DNA tetramer: IC-C-G-G // *J. Mol. Biol.*—1984.—**174**, N 4.—P. 663–695.
 54. *Dickerson R. E.* Base sequence and helix structure variation in B and A DNA // *J. Mol. Biol.*—1983.—**166**, N 3.—P. 419–441.
 55. *Bhattacharyya D., Bansal M.* Groove width and depth of B-DNA structures depend on local variation in slide // *J. Biomol. Struct. Dyn.*—1992.—**10**, N 1.—P. 213–226.
 56. *Dickerson R. E., Ng H. L.* DNA structure from A to B // *Proc. Nat. Acad. Sci. USA.*—2001.—**98**, N 13.—P. 6986–6988.
 57. *Bingman C., Li X., Zon G., Sundaralingam M.* Crystal and molecular structure of d(GTGCG CAC): investigation of the effects of base sequence on the conformation of octamer duplexes // *Biochemistry.*—1992.—**31**, N 51.—P. 12803–12812.
 58. *Olson W. K., Bansal M., Burley S. K., Dickerson R. E., Gerstein M., Harvey S. C., Heinemann U., Lu X. J., Neidle S., Shakked Z., Sklenar H., Suzuki M., Tung C. S., Westhof E., Wolberger C., Berman H. M.* A standard reference frame for the description of nucleic acid base-pair geometry // *J. Mol. Biol.*—2001.—**313**, N 1.—P. 229–233.
 59. *Gorin A. A., Zhurkin V. B., Olson W. K.* B-DNA Twisting correlates with base-pair morphology // *J. Mol. Biol.*—1995.—**247**, N 1.—P. 34–48.
 60. *Lu X., Olson W. K.* 3DNA: software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structure // *Nucl. Acids Res.*—2003.—**31**, N 17.—P. 5108–5121.
 61. *Sims G. E., Kim S.-H.* Global mapping of nucleic acid conformational space: dinucleoside monophosphate conformations and transition pathways among conformational classes // *Nucl. Acids Res.*—2003.—**31**, N 19.—P. 5607–5616.
 62. *Wenz C., Jeltsch A., Pingoud A.* Probing the indirect readout of the restriction enzyme EcoRV mutational analysis of contacts to the DNA backbone // *J. Biol. Chem.*—1996.—**271**, N 10.—P. 5565–5573.
 63. *Zhang Y., Zhiqun X., Hegde R. S., Shakked Z., Crothers D. M.* Predicting indirect readout effects in protein–DNA inter-

- actions // Proc. Nat. Acad. Sci. USA.–2004.–**101**, N 22.–P. 8337–8341.
64. *Schneider B., Neidle S., Berman H. M.* Conformations of the sugar–phosphate backbone in helical DNA crystal structures // *Biopolymers.*–1997.–**42**, N 1.–P. 113–124.
65. *Fratini A. V., Kopka M. L., Drew H. R., Dickerson R. E.* Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG* // *J. Biol. Chem.*–1982.–**257**, N 24.–P. 14686–14707.
66. *Varnai P., Djuranovic D., Lavery R., Hartmann B.* - transitions in the B-DNA backbone // *Nucl. Acids Res.*–2002.–**30**, N 24.–P. 5398–5406.
67. *Dixit S. B., Beveridge D. L., Case D. A., Cheatham T. E. 3rd, Giudice E., Lankas F., Lavery R., Maddocks J. H., Osman R., Sklenar H., Thayer K. M., Varnai P.* Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II: sequence context effects on the dynamical structures of the 10 unique dinucleotide steps // *Biophys. J.*–2005.–**89**, N 6.–P. 3721–3740.
68. *Grzeskowiak K., Yanagi K., Prive G. G., Dickerson R. E.* The structure of B-helical C-G-A-T-C-G-A-T-C-G and comparison with C-C-A-A-C-G-T-T-G-G. The effect of base pair reversals // *J. Biol. Chem.*–1991.–**266**, N 14.–P. 8861–8883.
69. *Hartmann B., Piazzola D., Lavery R.* BI-BII transitions in B-DNA // *Nucl. Acids Res.*–1993.–**21**, N 3.–P. 561–568.
70. *Bertrand H., Ha-Duong T., Femandjian S., Hartmann B.* Flexibility of the B-DNA backbone: effects of local and neighbouring sequences on pyrimidine-purine steps // *Nucl. Acids Res.*–1998.–**26**, N 5.–P. 1261–1267.
71. *Djuranovic D., Hartmann B.* Conformational characteristics and correlations in crystal structures of nucleic acid oligonucleotides: evidence for sub-states // *J. Biomol. Struct. Dyn.*–2003.–**20**, N 6.–P. 771–788.
72. *Haran T. E., Shakked Z., Wang A. H., Rich A.* The crystal structure of d(CCCCGGGG): a new A-form variant with an extended backbone conformation // *J. Biomol. Struct. Dyn.*–1987.–**2**, N 2.–P. 199–217.
73. *Wang A. H.-J., Fujii S., van Boom J. H., Rich A.* Molecular structure of the octamer d(G-G-C-C-G-G-C-C): modified A-DNA // *Proc. Nat. Acad. Sci. USA.*–1982.–**79**, N 13.–P. 3968–3972.
74. *Shakked Z., Rabinovich D., Kennard O., Cruse W. B. T., Salsbury S. A., Viswamitra M. A.* Sequence-dependent conformation of an A-DNA double helix. The crystal structure of the octamer d(G-G-T-A-T-A-C-C) // *J. Mol. Biol.*–1983.–**166**, N 2.–P. 183–201.
75. *Tkachenko M. Y., Boryskina O. P., Shestopalova A. V., Tolstorukov M. Y.* ProtNA-ASA: Protein-nucleic acid structural database with information on accessible surface area // *Int. J. Quant. Chem.*–2010.–**110**, N 1.–P. 230–232.
76. *Boryskina O. P., Tkachenko M. Y., Shestopalova A. V., Tolstorukov M. Y.* DNA accessible surface area and indirect protein-DNA recognition: study by bioinformatical approach: Proc. 7th EBSA Eur. Biophys. Congr. (Genova, Italy, July 11–15 2009) // *Eur. Biophys. J.*–2009.–**38** (Suppl. 1).–P. S62.
77. *Rozenberg H., Rabinovich D., Frolow F., Hegde R. S., Shakked Z.* Structural code for DNA recognition revealed in crystal structures of papillomavirus E2-DNA targets // *Proc. Nat. Acad. Sci. USA.*–1998.–**95**, N 26.–P. 5194–5199.
78. *Devarakonda S., Harp J. M., Kim Y., Ozyhar A., Rastinejad F.* Structure of the heterodimeric ecdysone receptor DNA-binding complex // *EMBO J.*–2003.–**22**, N 21.–P. 5827–5840.
79. *Sierk M. L., Zhao Q., Rastinejad F.* DNA deformability as a recognition feature in the reverb response element // *Biochemistry.*–2001.–**40**, N 43.–P. 12833–12843.
80. *Joshi R., Passner J. M., Rohs R., Jain R., Sosinsky A., Crickmore M. A., Jacob V., Aggarwal A. K., Honig B., Mann R. S.* Functional specificity of a Hox protein mediated by the recognition of minor groove structure // *Cell.*–2007.–**131**, N 3.–P. 530–543.

UDC 577.323.5

Received 26.01.10