ISSN 0006-2979, Biochemistry (Moscow), 2007, Vol. 72, No. 9, pp. 913-919. © Pleiades Publishing, Ltd., 2007.
Original Russian Text © G. B. Zavilgelsky, S. M. Rastorguev, 2007, published in Biokhimiya, 2007, Vol. 72, No. 9, pp. 1125-1132.

REVIEW

DNA Mimicry by Proteins as Effective Mechanism for Regulation of Activity of DNA-Dependent Enzymes

G. B. Zavilgelsky* and S. M. Rastorguev

State Research Institute of Genetics and Selection of Industrial Microorganisms, 1-yi Dorozhnyi Proezd 1, 117545 Moscow, Russia; fax: (495) 315-0501; E-mail: zavilgel@genetika.ru; rsergey@molgen.ru

Received March 27, 2007
Revision received May 21, 2007

Abstract—Modern concepts on mechanisms of DNA-dependent enzyme regulation involving specific DNA-mimicking proteins are considered. There are proteins that share structural resemblance with DNA duplexes. These include inhibitors of type I restriction–modification enzymes (Ocr and ArdA), inhibitors of DNA gyrase MfpA and QnrABS, etc. We describe here structural features of these proteins and mechanisms responsible for their interaction with DNA-dependent enzymes and then discuss perspectives of use of DNA-mimicking proteins in analysis of replication, repair, recombination, mechanisms underlying resistance to antibiotics, and also fields of applied biotechnology.

DOI: 10.1134/S0006297907090015

Key words: DNA-mimicry, antirestriction, drug resistance, quinolone, ArdA, MfpA, QnrABS, Ocr

Protein mimicry of DNA is a recently discovered direct mechanism of regulation of DNA-dependent enzyme activity by means of proteins that mimic DNA structure and interact with a target enzyme and inhibit (or modulate) its activity. The phenomenon of DNA mimicry was originally discovered during studies of Ugi protein, an inhibitor of uracil-DNA glycosylase (UDG), an enzyme involved in DNA repair [1-3]. A similar type of DNA mimicry has been found in various groups of proteins: in ribosomal elongation factor EF-G (tRNA-like motif) and in the component $dTAF_{II}230$ of eukaryotic transcription factor TFIID (DNA-like domain) [4]; in DinI, a negative regulator of SOS response in *Escherichia coli* [5]; in *Haemophilus influenzae* HI1450 involved in nucleosome formation [6]. There are DNA-mimicking proteins of the family of DNA gyrase inhibitors (MfpA from *Mycobacterium tuberculosis* and QnrABS, which determine resistance of bacterial cells to quinolone antibiotics [7]) and also proteins of the family of inhibitors of type I restriction–modification enzymes (the antirestriction proteins Ard, Ocr, etc.) encoded by conjugated plasmids and bacteriophages [8, 9].

In this review, we consider some examples of DNA-mimicking proteins, their structural features and action mechanisms, and then discuss possible variants of use of the DNA-mimicking proteins in applied biotechnology.

Antirestriction proteins. Antirestriction proteins, also known as antirestrictases, are an impressive example of DNA-mimicking proteins. (It should be noted that even though the antirestriction proteins inhibit activity of the restriction–modification system, they lack their own intrinsic catalytic activity, and therefore the term "antirestrictase" may be used arbitrarily.) During evolution, special systems responsible for overcoming intercellular restriction barriers appeared in conjugated plasmids and some bacteriophages. This phenomenon was named "antirestriction" [10, 11]. Studies of mechanisms involved in antirestriction in conjugated plasmids resulted in identification of a special protein, an antirestrictase, the product of the *ard* gene (alleviation of restriction of DNA). Originally (1984-1985), *ard* genes were found in the N-incompatibility group plasmid [12, 13] and then in other plasmids [14, 15]. In 1991-1995, *ard* genes were sequenced and primary structures of Ard proteins were determined [16, 17]. The *ard* genes (three variants are now known: *ard*A, *ard*B, and *ard*C) encode small extremely acidic proteins (of 140-170 residues) that have characteristic negative charge (–10...–30) and are inhibitors of cell type I restrictases-methyltransferases. It has been shown that *ard* genes play an important role in the life cycle of conjugated plasmids; they are responsible for overcoming the restriction barriers during DNA trans-

Abbreviations: PRP) pentapeptide repeat proteins; Rfr) repeated five-residues; UDG) uracil-DNA glycosylases.

* To whom correspondence should be addressed.

fer from one bacterial species into others (so-called horizontal gene transfer), thus promoting evolution processes. Some bacteriophages (T1, P1, etc.) also have genes encoding antirestrictase proteins (the gene 0.3 (*ocr*) of phage T7 and gene *dar*A of phage P1 [18-20]). Plasmids (Ard) or phage (Ocr and DarA) antirestriction proteins inhibit only type I restriction–modification systems; their genes (*hsd*RMS) are usually located on a bacterial chromosome; these proteins do not inhibit type II restriction endonucleases encoded by genes localized in plasmids.

The antirestriction proteins inhibit both endonuclease (restriction) and methylase (modification) activities of the type I enzyme complex. Type I restriction–modification enzymes consist of five subunits, R_2M_2S, where R is endonuclease, M is methylase, and S recognizes and contacts with a specific site of double-stranded DNA [21]. It should be noted that the antirestriction proteins exhibit equal inhibitory activity with respect to all groups of type I restriction–modification enzymes (IA, IB, IC, ID) and their inhibitory activity is site-independent. The latter property is quite explainable because conjugated plasmid transfer and cell infection with phage are random processes under natural conditions; they may occur during contacts with bacteria containing various specificities of the restriction–modification system.

It has been suggested that the antirestriction proteins of the Ard family and also Ocr and DarA proteins belong to a group of protein modulators which mimic B-form DNA. These proteins have characteristic distribution of negatively charged D and E residues (aspartate and glutamate), imitating distribution of negatively charged phosphate groups along DNA duplex [10]. In other words, antirestrictases mimic DNA structure. The goal of this mimicry is to replace a DNA strand in its complex with a restriction–modification enzyme. This helps conjugated plasmid and bacteriophage DNA to avoid "strict immigration control" by the type I restriction–modification cell system. In 2002, Walkinshaw et al. [8] demonstrated the spatial structure of phage T7 Ocr, the minimal protein of the antirestrictase family (116 residues, total charge –25). According to X-ray structure analysis by scientists from Edinburgh and Cambridge Universities guided by Professor T. Dryden, the structure of Ocr protein is similar to B-form DNA (Fig. 1, taken from [22]; see color insert). The main core of Ocr monomer consists of three α-helices: A (including residues 7-24), B (residues 34-44), and the long, slightly bent D helix (residues 73-106). They form a tightly packed bundle with strictly ordered distribution of negatively charged carboxyl groups of D and E residues along the rod-like structure, which basically coincides with the distribution of negatively charged phosphate groups along DNA duplex. The short α-helix C (residues 49-57) determines the contact between two monomers of Ocr and formation of a stable dimer. The contact between two monomers is determined by van der Waals forces due to interaction of clusters of hydrophobic residues (A50, F53, S54, M56, A57, and also V77) located in the middle of α-helix C. Protein mimicry of DNA duplex allows competition between antirestrictase and DNA for binding to the restriction–modification enzyme and inhibition of DNA degradation (restriction) and methylation (modification) processes.

The mechanism of antirestriction is an example of competitive inhibition from the viewpoint of classical enzymatic catalysis: an inhibitor molecule competes with a natural substrate due to their structural similarity. There is a characteristic position of monomers in the dimer $(Ocr)_2$: an angle of 34° between the longitudinal axes of the monomer (Fig. 1). This dimer structure basically coincides with the structure of DNA duplex possessing a kink in the recognition site. This kink is formed in the center of the recognition sites during complex formation between type I restriction–modification enzyme and DNA. The recognition sites for type I restriction–modification enzymes belong to the group of "sites with hyphen"; for example, the site recognized by *Eco*KI is defined as 5′-AAC$(N)_6$GTGC-3′. The protein $(Ocr)_2$ contacts with S- and R-subunits, and this antirestrictase does not require additional energy for binding to *Eco*KI. This results in effective displacement of DNA duplex from the complex with enzyme by $(Ocr)_2$ (the constant of complex formation between $(Ocr)_2$ and *Eco*KI is about two orders of magnitude higher than that of enzyme–DNA complex formation) [22].

Multifunctional type I restriction–modification systems have been found in 50% of all prokaryotes studied [23]. GenBank now contains about 1000 nucleotide and amino acid sequences homologous to antirestrictases (antirestriction protein). There are 50 sequences homologous to ArdA in the Pfam database [24].

Inhibitors of DNA gyrase. Quinolones and also fluoroquinolones are synthetic derivatives of nalidixic acid; they belong to a group of antibiotics with wide spectrum of action and high activity and inhibit DNA gyrase. Quinolones bind to the gyrase–DNA complex. This results in stabilization of the covalent enzyme tyrosyl-DNA phosphate ester (a transient reaction intermediate) and causes death of bacteria. Quinolones have been successfully used for inactivation of *Mycobacterium tuberculosis* cells. During the first years of clinical use of quinolones, findings of *M. tuberculosis* strains resistant to quinolones were rather rare events. Studies of the nature of resistance to quinolones in the laboratory strains of *M. tuberculosis* and the related strain *M. smegmotis* have shown that this effect is determined by missense mutations (amino acid substitutions) in A-chain of DNA gyrase, or it represents the result of regulatory mutation potentiating expression of a protein pump responsible for the extracellular efflux of toxic compounds. However, wide use of quinolones in medical practice resulted in discovery of a new type of quinolone resistance. It was shown that the gene determining such type of resistance

in *M. smegmotis* and *M. tuberculosis* encodes the MfpA protein, a specific inhibitor of DNA gyrase [7, 25]. The MfpA proteins of *M. tuberculosis* and *M. smegmotis* consist of 183 and 192 residues; they share 67% identity.

In 1998, the resistance to quinolones found in *Klebsiella pneumoniae* was shown to be encoded by the *qnr*A gene and transferred by the conjugated plasmid [26]. Subsequent studies resulted in identification of QnrA protein of 218 amino acid residues. The QnrA protein from *K. pneumoniae* (as well as MfpA) is an inhibitor of DNA gyrase [27]. Later genes of the *qnr* group encoding protein inhibitors of gyrases and determining resistance of cells to quinolones were found in genomes (in chromosomes, plasmids, transposones) of various bacterial species isolated in clinics of many counties [28-30]. There are three variants of the *qnr* gene: *qnr*A, *qnr*B, and *qnr*S. The proteins encoded by these genes QnrA, QnrB, and QnrS, respectively, exhibit the same function of DNA gyrase inhibition. The highly homologous QnrA and QnrS proteins share >60% amino-acid identity. The QnrB proteins (usually of 226 amino acid residues) exhibit low homology with QnrA and QnrS proteins. For example, QnrB1 has 40% amino acid identity with *M. tuberculosis* MfpA [29]. There are differences by five amino acids between QnrB1 and other proteins of the QnrB group; genes encoded these proteins have been found in plasmids isolated from strains of *E. coli*, *Enterobacter cloacae*, *Citrobacter koseri*, and *C. freundi*. All these proteins (QnrA, QnrB, and QnrS) carry similar total negative charge (−6...−10) and contain a significant number of tandem pentapeptide repeats.

Figure 2 (taken from [31]; see color insert) shows the spatial structure of the MfpA protein; it consists of a right-handed β-helix, which corresponds to B-form DNA in size, shape, and electrostatics [7]. MfpA consists of 183 amino acid residues and belongs to the family of pentapeptide repeat proteins (PRP); in these pentapeptides each second amino acid is D or N and each third amino acid is L or F. In solutions, MfpA forms a dimer due to hydrophobic contact of several amino acids located at the C-end of an α-helical site. The monomeric MfpA consists of eight coils, and four repeated pentapeptides form four sides of a quadrant (1-4) with slightly differed characteristics (Table 1). Such spatial structure was named RHQBH (right-handed quadrilateral beta-helix) or "Rfr" (Repeated five-residues).

The dimer $(MfpA)_2$ has a rod-like shape 100 Å in length and 27 Å in diameter. The total charge of the dimer is −10, but the negative charges are distributed non-randomly. This results in $(MfpA)_2$ dimer, which mimicks a 30-bp segment of B-form duplex DNA. Docking analysis revealed the existence of tight contact between $(MfpA)_2$ dimer and A_2 dimer of the DNA gyrase A subunit (Fig. 2b) due to electrostatic complementation between strongly cationic "seat" of the A_2 dimer interface and a strongly anionic surface of the $(MfpA)_2$ dimer. Structural similarity of MfpA protein with DNA duplex of the gyrase substrate determines effectiveness of competitive inhibition of the gyrase; this represents the molecular basis of bacterial resistance to quinolone antibiotics. It should be noted that in contrast to gyrase inhibition by quinolones, the inhibition of gyrase by MfpA protein is not accompanied by cell chromosome degradation. Consequently, the presence of the genes *mfp*A or *qnr*ABS in the bacterial genome is very important because the "fee" for rescue from the inactivating effect of antibiotics is delayed development of the cell. It is possible that the main function of DNA-mimicking inhibitors of gyrase consists in modulation of DNA supercoiling, which may potentiate supercoiling at the stage of DNA replication and decrease the rate of supercoiling when the level of chromosome compactness becomes optimal in a particular cell.

Pentapeptide repeat family proteins. MfpA protein and also QnrA, QnrB, and QnrS proteins belong to the

Table 1. Position of pentapeptides along the axis of the MfpA protein molecule

Coil	Quadrant sides				Amino acid position in the protein chain
	1	2	3	4	
1	QQWVD	CEFTG	RDFRD	EDLSR	21
2	LHTER	AMFSE	CDFSG	VNLAE	41
3	SQHRG	SAFRN	CTFER	TTLWH	61
4	STFAQ	CSMLG	SVFVA	CRLRP	81
5	LTLDD	VDFTL	AVLGG	NDLRG	101
6	LNLTG	CRLRE	TSLVD	TDLRK	121
7	CVLRG	ADLSG	ARTTG	ARLDD	141
8	ADLRG	ATVDP	*VLWRT*	ASLVG	161
	ARVDV	*DQAVA*	*FAAR*	*GLCLA*	181

PRP family. Amino acid sequences of these proteins contain a repeated pentapeptide with the consensus [S, T, A, V][D, N][L, F][S, T, R][G]. Table 1 shows that MfpA protein consists of 30 pentapeptides, which determine characteristic features of its spatial structure.

The first protein of the PRP family was originally found in *Anabaena* cyanobacteria [32]. The HglK protein (encoded by the *hgl*K gene and consisting of 727 residues) contains a series of 36 tandem pentapeptides with the consensus sequence ADLSG. Using methods of bioinformatics, a group of proteins belonging to PRP family has been identified in *Synechocystis* cyanobacteria; there are 15 proteins with series of tandem pentapeptide repeats varying from 13 to 44 [33]. Now proteins of the PRP family have been found in almost all living organisms excluding yeasts. According to data analysis [31], 525 proteins (484 prokaryotic and 41 eukaryotic) with the pentapeptide motif have been identified.

Sequencing of the genome of the cyanobacterium *Cyanothece* sp. PCC 51142 revealed 35 pentapeptide-containing proteins. Buchko et al. [34] determined the spatial structure of the Rfr32 protein, which consists of 167 residues. The authors demonstrated that 21 tandem pentapeptide repeats (with the consensus motif A(N/D)LXX) fold into a right-handed quadrilateral β-helix, or Rfr-fold (as in the case of the MfpA protein); this structure imitates the rod-like structure of B-form DNA. The Rfr structure is also typical for another protein, Rfr23, encoded by a gene that has also been found in the genome of *Cyanothece* sp. PCC 51142 [35]. The real functions of the pentapeptide-containing proteins found in cyanobacteria remain unknown.

Some proteins determining immunity of bacteria to their own synthesized antibiotics also belong to the PRP family. These include the McbG protein (encoded by a *mcb*G gene located in the operon responsible for biosynthesis of microcin B17 [36]) and the OxrA protein determining resistance of *Bacillus megatherium* to oxetanocin A [37]. In contrast to quinolones, microcin B17 interacts with B-subunit of DNA gyrase, but even though the mechanism underlying resistance to antibiotics obviously involves interaction of a DNA-mimicking protein molecule with DNA gyrase.

A significant group of pentapeptide repeat family proteins has complex structure and contains several domains, including those with catalytic functions. However, the functional role of the pentapeptide repeats in this group remains unknown. But if the putative catalytic function of such protein consists of posttranslational modification of some DNA-binding protein (e.g. histone acetylation), one can suggest that binding of the target protein to the pentapeptide domain would significantly increase selectivity of such a modification reaction.

Ugi, an inhibitor of uracil DNA glycosylase (UDG), and other DNA-mimicking proteins. The phenomenon of DNA-mimicry was discovered during studies of the Ugi protein, which consists of 84 amino acid residues (total charge is −12) and is encoded by PBS2 bacteriophage (*B. subtilis* is the host bacterium). Ugi is a protein inhibitor of the repair enzyme UDG [1, 2]. From the evolutionary viewpoint, this represents a unique and interesting example. The DNA of bacteriophages PBS1 and PBS2 contains uracil instead of thymine. It is suggested that this gives the phage DNA the opportunity to escape degradation by cell restrictases. However, bacteria contain the UDG enzyme, which is mainly involved in DNA repair: UDG excises uracils that appear in DNA due to cytosine deamination or replication errors, and the Ugi protein is an effective inhibitor of UDG "designed" for inhibition of cell UDG activity.

Figure 3 (taken from [3]; see color insert) shows the spatial structure of the Ugi–UDG complex, which is compared with the DNA–UDG structure. Since UDG recognizes in DNA just uracil, the contact of this enzyme with DNA is limited by several nucleotides. Consequently, the Ugi inhibitor contains a rather small fragment imitating the structure of several base pairs. The central part of spatial structure of the Ugi protein contains five anti-parallel β-strands surrounded by two rather small α-helices (Fig. 3a). The central hydrophobic core forms a cavity for the hydrophobic side chain of UDG, which binds (fixes) uracil during contact with DNA and moves it into the active site of UDG (Fig. 3, c and d). Negatively charged aspartate and glutamate residues imitate phosphate groups of DNA and play an important role in solution of the geometric task by forming hydrogen bonds with corresponding groups in UDG. It is important to emphasize that in contrast to the UDG–DNA complex, neither UDG nor Ugi change their spatial structures during complex formation. Consequently, as in the case of interaction of the antirestrictase Ocr with *Eco*KI, the UDG–Ugi complex is more energetically favorable than the UDG–DNA complex, and this determines the effectiveness of inhibition of UDG by the Ugi protein.

Haemophilus influenzae HI1450 is a very acidic protein of 107 amino acid residues. Its spatial structure was determined for elucidation of the functional role of this protein [6]. The results of structural studies have shown that there is certain similarity between spatial structures of Ugi and HI1450. The distribution of amino acid residues E and D on the protein surface corresponds to the distribution of negatively charged phosphate groups in B-form DNA. This protein was shown to form a complex with the Hu-α protein by competing with DNA [38]. Homologs of HI1450 (the open reading frame homologs) have been found in genomes of *E. coli* and other bacterial species. Since Hu proteins belong to the group of histone-like proteins, are widely distributed, and are involved in compactization of bacterial DNA, it is possible that HI1450 and related proteins are involved in regulation of nucleoid formation in bacterial cells.

The protein DinI plays the role of negative regulator of SOS response in *E. coli* cells. DinI contains a single α-helical domain with ordered position of negatively charged E and D along the protein surface, thus imitating the structure of a single-stranded DNA [5]. DinI interacts with the RecA proteins without displacement of the single-stranded DNA from its complex with RecA. This causes structural changes of the complex accompanied by the decrease of activity of the RecA protease.

DNA-mimetics also exist in eukaryotes. For example, there is the *Drosophila* $dTAF_{II}230$ protein. It is involved in complex formation with TBP (TATA-box binding protein) and repression of transcription [4]. The $dTAF_{II}230$ protein contains N-terminal DNA-mimicking domain, which consists of 77 amino acid residues. In the contact surface of $dTAF_{II}230$ there are two rows of negatively charged amino acid residues for interaction with arginine and lysine residues of the TBP protein. Hydrophobic contacts also play a significant role in the complex formation. The $dTAF_{II}230$ protein mimics grossly distorted DNA with impaired double-helical structure usually formed in the complex of DNA with TBP [4, 39]. This is a good illustration of protein mimicry of a specific site in DNA (TATA-box).

Characteristics of DNA-mimicking proteins. How is it possible to distinguish a DNA-mimetic from an ordinary protein? Since an algorithm for identification of DNA-mimetics by amino acid sequence has not been developed by bioinformaticians, this still represents a significant problem. It should be noted that DNA-mimicking proteins significantly differ in their amino acid sequences. Even in the group of proteins exhibiting almost identical functions (e.g. antirestrictases employing the same mechanism of competition with DNA for complex formation with type I restriction–modification enzymes) it is impossible to find common amino acid sequences in the primary structure of the proteins Ocr and ArdA.

This situation is complicated by the fact that in some proteins DNA mimicry involves just a part (domain) of the macromolecule. For example, X-ray analysis of the Ugi protein has shown the presence of a domain exhibiting similarity with B-form DNA, but the whole Ugi molecule has a globular structure (Fig. 3) [3]. There is principal difference in the case of protein inhibitors of restriction–modification enzymes and inhibitors of DNA gyrase. These proteins have extended rod-like shape, and their molecular structure and electrostatic characteristics are similar to those of B-form DNA (Figs. 1 and 2).

Figure 4 (see color insert) shows structures of Ocr and MfpA monomers and B-form DNA. Their comparison emphasizes the extraordinary capacities of living nature to develop unique forms crucial for adaptation. The most surprising thing is that nature has chosen different ways for design of proteins mimicking the DNA duplex. In one case (e.g. Ocr) these are tightly packed α-helices, in the other it is a right-handed β-helix (MfpA).

Existence of significant negative charge (of the whole macromolecule or particular domain) required for similarity with the DNA polyanion is a common feature of DNA-mimicking proteins. However, this is a necessary but not sufficient precondition. At the moment the only reliable method for detection of DNA-mimicking proteins is X-ray analysis.

Our data on antirestriction activity of the ArsR repressor proteins (inhibitors of transcription of the *ars*-operons) may be used as the example of successful search of the DNA-mimetics with particular function by primary structure of proteins. It should be noted that in the Ugi protein residues E20, E28, and E31, located in the N-domain, determine the key negative charge (Fig. 3c). We have noted that similar motif have also been found in the ArsR proteins, the repressors of the *ars*-operons responsible for resistance of bacterial cells to arsenate ions (Table 2). Cloning of the *ars*R genes from various conjugated plasmids R64 (incI1), R773 (incFI), and pKW301 from *Acidiphilium multivorum* onto a multicopied vector and their administration to *E. coli* K12 caused significant inhibitory effect on the *Eco*KI enzyme [40, 41]. ArsR encoded by the pKW301 plasmid demonstrated the highest inhibitory activity among proteins of this group (alleviation of *Eco*KI restriction on unmodified phage λ by two orders of magnitude) [41]. ArsR pKW301 (as well as Ugi) consists of 84 amino acid residues and total charge of the ArsR pKW301 protein is –4; this suggests the acidic nature of this protein. Consequently, the ability of the ArsR protein repressor to inhibit type I restrictases is determined by N-domain, and total negative charge of the ArsR protein promotes augmentation of its antirestriction activity. It is especially interesting that the *ars*-operons are usually located in the conjugative plasmids, which can carry out "horizontal" gene transfer between bacteria of various species.

Application of DNA-mimicking proteins. Since genes encoding DNA-mimicking proteins (e.g. *ard*A and *qnr*A, *qnr*B, and *qnr*S) are located on transmission elements, transposons, and plasmids, this promotes their wide distribution among bacteria of various species and genera.

Table 2. DNA-mimicking motif in the N-domain of Ugi and ArsR proteins

Protein	Source	Sequence
Ugi	PBS2	20*E*...25LPE*E*VE*E*VI
ArsR	R773	16*E*...24LLR*E*MG*E*LC
ArsR	pKW301	15*E*...23LLR*E*LG*E*LC
ArsR	R64	16*E*...24LLR*E*MG*E*LC

Note: Key amino acids are italicized.

Thus it is important to investigate in detail the structure of such proteins and the mechanisms of their action. The most illustrative example is distribution of gene responsible for resistance to quinolone antibiotics (*qnr*A, *qnr*B, and *qnr*S) among clinical bacterial strains. Only in 2005, the mechanism of inhibitory effect of the DNA-mimicking proteins (QnrA, QnrB, and QnrS) on DNA gyrase was determined [7]. The search for and analysis of genes encoding DNA-mimicking proteins and representing constituents of transmission elements are important tasks.

Below we consider some putative variants of use of DNA-mimicking proteins.

The DNA-mimicking proteins may be successfully used for substitution of DNA during elucidation of spatial structure of the DNA-dependent enzymes by means of X-ray analysis [42]. In some cases, it is difficult to obtain crystals of the complexes of the DNA dependent enzymes and DNA and it is possible that substitution of DNA by the DNA-mimicking proteins may solve this problem. There are examples illustrating successful use of such substitutions: Ugi–UDG [1-3], $dTAF_{II}230$–TBP [4]. It is suggested that substitution of DNA by the Ocr protein might be used for crystallization of Ocr in its complex with S-subunit of *Eco*KI. In this connection it should be noted that spatial structure of S-subunit of two type I restriction–modification enzymes has been determined [43, 44].

The DNA-mimicking proteins can be used in affinity chromatography. Affinity columns with a DNA-mimicking protein can be used with high effectiveness for detection and purification of various types of DNA-dependent enzymes. Use of radioactive or fluorescent labels will increase sensitivity of such method.

Perspectives of *in vitro* construction of new types of DNA-mimicking proteins (i.e. generation of proteins with different "design" and new functions) may be quite wide. These include potential tasks of construction of DNA-mimetics, inhibiting or modulating activity of specific groups of DNA-dependent enzymes and tasks related to site-directed changes in the structure of already known DNA-mimicking proteins. Such works are rather successful. For example, using site-directed mutagenesis we have modified the structure of antirestrictases ArdA and Ocr; the modified proteins selectively inhibit endonuclease (restriction) activity of type I restriction–modification enzymes without any influence on their methylase (modification) activity [45, 46]. Use of such type of antirestrictases in gene engineering works gives a possibility for the development of stable strains with hybrid plasmids because the process of specific modification of chromosome DNA remains unimpaired.

The protein Ocr has already been used as an effective factor promoting significant increase of bacterial transformation by plasmids. Addition of a small amount of the Ocr protein to solution with plasmid DNA causes significant (by several orders of magnitude) increase in effectiveness of cell transformation during electroporation. (In this case unmethylated DNA and host bacteria with active type I restriction–modification system are used.) The DNA-mimetic Ocr (as well as plasmid DNA) easily penetrates inside cells and immediately protects unmethylated DNA against degradation (EPICENTRE Forum 9, 8, http//www.epibio.com/forum.asp).

In general, perspectives of use of the DNA-mimicking proteins might be related diagnostics and therapy of various diseases (e.g. for inhibition of specific enzymes and corresponding biochemical processes in cells).

This work was supported by grants from the Russian Foundation for Basic Research (Nos. 05-04-08033 and 07-04-00726).

REFERENCES

1. Mol, C. D., Arval, A. S., Sandersen, R. J., Slupphaug, G., Kavil, B., Krokan, H. E., Mosbaugh, D. W., and Tainer, J. A. (1995) *Cell*, **82**, 701-708.
2. Savva, R., and Pearl, L. H. (1995) *Nat. Struct. Biol.*, **2**, 1464-1472.
3. Putnam, C. D., and Tainer, J. A. (2005) *DNA Repair*, **4**, 1410-1420.
4. Liu, D., Ishima, R., Tong, K. I., Bagby, S., Kokubo, T., Muhandiram, D. R., Kay, L. E., Nakatani, Y., and Ikura, M. (1998) *Cell*, **94**, 573-583.
5. Ramirez, B. E., Voloshin, O. N., Camerini-Otero, R. D., and Bax, A. (2000) *Protein Sci.*, **9**, 2161-2169.
6. Parsons, L. M., Yeh, D. C., and Orban, J. (2004) *Protein Struct. Func. Bioinform.*, **54**, 375-383.
7. Hegde, S. S., Vetting, M. V., Roderick, S. L., Mitchenall, L. A., Maxwell, A., Takiff, H. E., and Blanchard, J. S. (2005) *Science*, **308**, 1480-1483.
8. Walkinshaw, M. D., Taylor, P., Sturrock, S. S., Atanasiu, C., Berge, T., Henderson, R. M., Edwardson, J. M., and Dryden, D. T. F. (2002) *Mol. Cell*, **9**, 187-194.
9. Dryden, D. T. F., and Tock, M. R. (2006) *Biochem. Soc. Trans.*, **34**, 317-319.
10. Zavilgelsky, G. B. (2000) *Mol. Biol. (Moscow)*, **34**, 854-862.
11. Tock, M. R., and Dryden, D. T. F. (2005) *Curr. Opin. Microbiol.*, **8**, 466-472.
12. Zavilgelsky, G. B., Belogurov, A. A., Mershavka, V. Yu., and Yussifov, T. N. (1984) *Mol. Biol. (Moscow)*, **18**, 1590-1596.
13. Belogurov, A. A., Yussifov, T. N., Kotova, V. Yu., and Zavilgelsky, G. B. (1985) *Mol. Gen. Genet.*, **198**, 509-513.
14. Kotova, V. Yu., Belogurov, A. A., and Zavilgelsky, G. B. (1988) *Mol. Biol. (Moscow)*, **22**, 270-276.
15. Zavilgelsky, G. B., Bakalova, T. L., Duzhiy, D. E., and Kotova, V. Yu. (1994) *Genetika*, **30**, 1582-1586.
16. Chilley, P. M., and Wilkins, B. M. (1995) *Microbiology*, **141**, 2157-2164.
17. Delver, E. P., Kotova, V. Yu., Zavilgelsky, G. B., and Belogurov, A. A. (1991) *J. Bacteriol.*, **173**, 5887-5892.
18. Dunn, J. J., Elzinga, M., Mark, K. K., and Studier, F. W. (1981) *J. Biol. Chem.*, **256**, 2579-2585.

19. Kruger, D. H., Hansen, S., and Reuter, M. (1983) *J. Virol.*, **45**, 1147-1149.
20. Iida, S., Hiestand-Nauer, R., Sandmeier, H., Lehnherr, H., and Arber, W. (1998) *Virology*, **251**, 49-58.
21. Murray, N. E. (2000) *Microbiol. Mol. Biol. Rev.*, **64**, 412-434.
22. Atanasiu, C., Su, T. J., Sturrock, S. S., and Dryden, D. T. (2002) *Nucleic Acids Res.*, **30**, 3936-3944.
23. Roberts, R. J., et al. (2005) *Nucleic Acids Res.*, **33**, D230-D232.
24. Finn, R. D., Mistry, J., Schuster-Bockler, B., et al. (2006) *Nucleic Acids Res.*, **34**, D247-D251.
25. Montero, C., Mateu, G., Rodrigez, R., and Takiff, H. F. (2001) *Antimicrob. Agents Chemother.*, **45**, 3387-3393.
26. Martinez-Martinez, L., Pascual, A., and Jacoby, G. A. (1998) *Lancet*, **351**, 797-799.
27. Tran, J. H., and Jacoby, J. A. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 5638-5642.
28. Li, X. Z. (2005) *Int. J. Antimicrob. Agents*, **25**, 453-463.
29. Jacoby, G. A., Walsh, K. E., Mills, D. M., Walker, V. I., Oh, H., Robicsek, A., and Hooper, D. C. (2006) *Antimicrob. Agents Chemother.*, **50**, 1178-1182.
30. Bonemann, G., Stiens, M., Puhler, A., and Schluter, A. (2006) *Antimicrob. Agents Chemother.*, **50**, 3075-3080.
31. Vetting, M. V., Hegde, S. S., Fajardo, E., Fiser, A., Roderick, S. L., Takiff, H. E., and Blanchard, J. S. (2006) *Biochemistry*, **45**, 1-10.
32. Black, K., Buikema, W. J., and Haselkorn, R. (1995) *J. Bacteriol.*, **171**, 6440-6448.
33. Bateman, A., Murzin, A. G., and Teichman, S. A. (1998) *Protein Sci.*, **7**, 1477-1480.
34. Buchko, G. W., Ni, S., Robinson, H., Welsh, E. A., Pakrasi, H. B., and Kennedy, M. A. (2006) *Protein Sci.*, **15**, 2579-2595.
35. Buchko, G. W., Robinson, H., Ni, S., Pakrasi, H. B., and Kennedy, M. A. (2006) *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, **62** (Pt. 12), 1251-1254.
36. Pierrat, O. A., and Maxwell, A. (2005) *Biochemistry*, **44**, 4202-4215.
37. Morita, M., Tomita, K., Ishizawa, M., Tagaki, K., Kawamura, F., Takahashi, H., and Morino, T. (1999) *Biosci. Biotechnol. Biochem.*, **63**, 563-566.
38. Parsons, L. M., Liu, F., and Orban, J. (2005) *Protein Sci.*, **14**, 1684-1687.
39. Goodsell, D. S. (2005) *J. Mol. Recognit.*, **18**, 427-430.
40. Rastorguev, S. M., Zavilgelsky, G. B., and Tchurikov, N. A. (1998) *FEBS Lett.*, **426**, 21-23.
41. Rastorguev, S. M., Zavilgelsky, G. B., Suzuki, K., and Sakka, K. (2001) *Mol. Biol. (Moscow)*, **35**, 79-82.
42. Dryden, D. T. F. (2006) *Trends Biotech.*, **24**, 378-382.
43. Kim, J.-S., DeGiovanni, A., Jancarik, J., Adams, P. D., Yokota, H., Kim, R., and Kim, S.-H. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 3248-3253.
44. Calisto, B. M. (2005) *J. Mol. Biol.*, **351**, 749-762.
45. Rastorguev, S. M., and Zavilgelsky, G. B. (2003) *Genetika*, **39**, 218-224.
46. Zavilgelsky, G. B., Letuchaya, E. A., and Rastorguev, S. M. (2006) *Genetika*, **42**, 333-338.

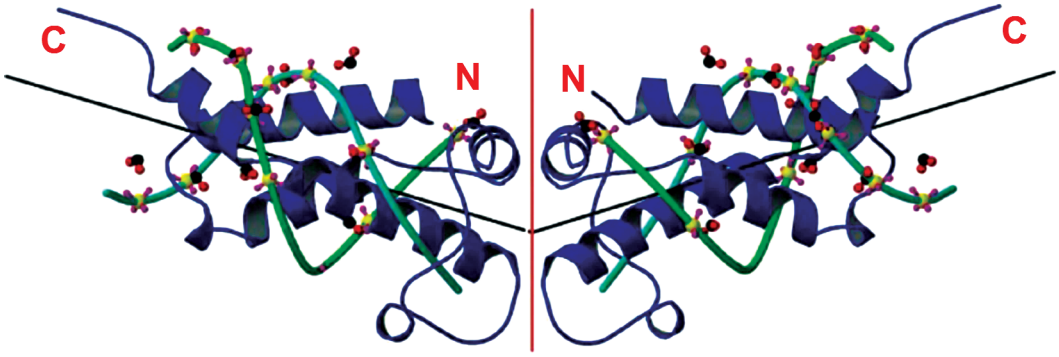

Fig. 1. (G. B. Zavilgelsky and S. M. Rastorguev) Spatial structure of the antirestrictase $(Ocr)_2$ dimer based on X-ray data.

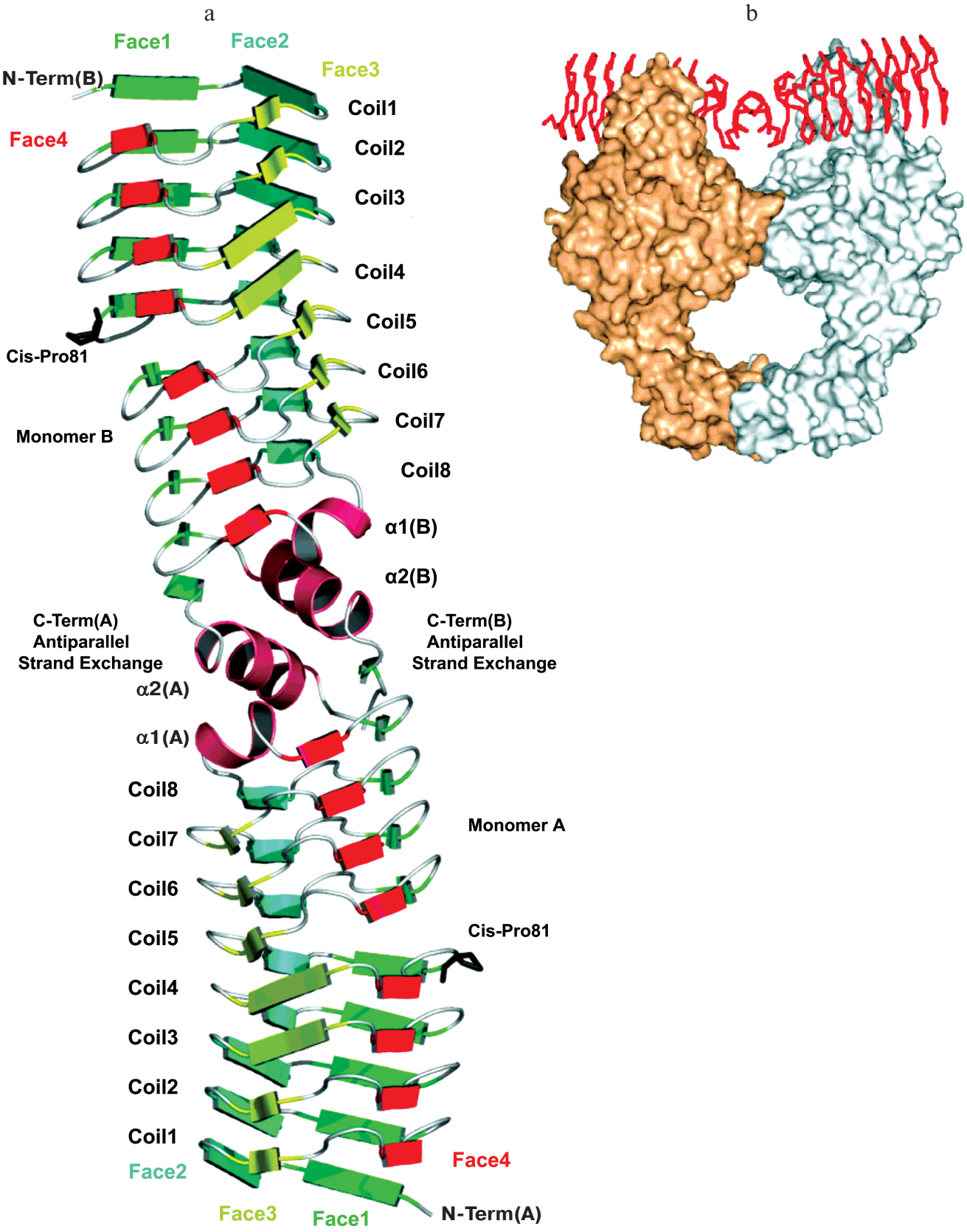

Fig. 2. (G. B. Zavilgelsky and S. M. Rastorguev) Spatial structure of MfpA protein, the inhibitor of DNA gyrase: a) schematic presentation of the structure of $(MfpA)_2$ dimer; b) molecular model of the complex of $(GyrA)_2$ and $(MfpA)_2$ dimers.

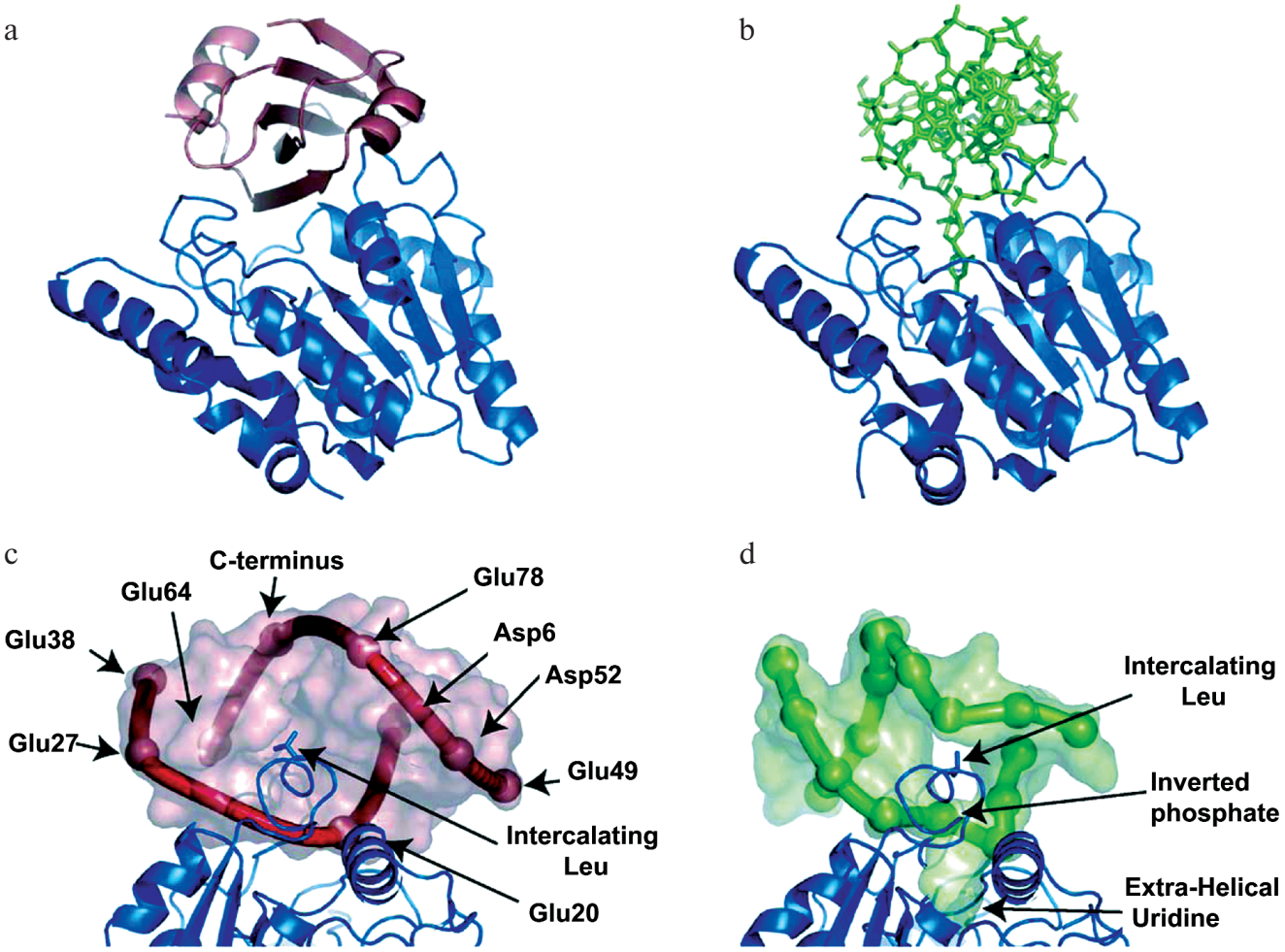

Fig. 3. (G. B. Zavilgelsky and S. M. Rastorguev) Spatial structures of the Ugi–UDG (a, c) and Ugi–DNA (b, d) complexes: a) ribbon model of the Ugi–UDG complex. Five antiparallel β-structures form a core of the Ugi protein; b) diagram of the DNA–UDG complex. In UDG a leucine residue contacting the hydrophobic site of the Ugi protein is marked (c) and in the complex with DNA this residue intercalates into the DNA duplex (d). In the DNA–UDG complex uridine looped out the DNA duplex into the UDG active site and corresponding inverted phosphate of the sugar-phosphate backbone are marked (d).

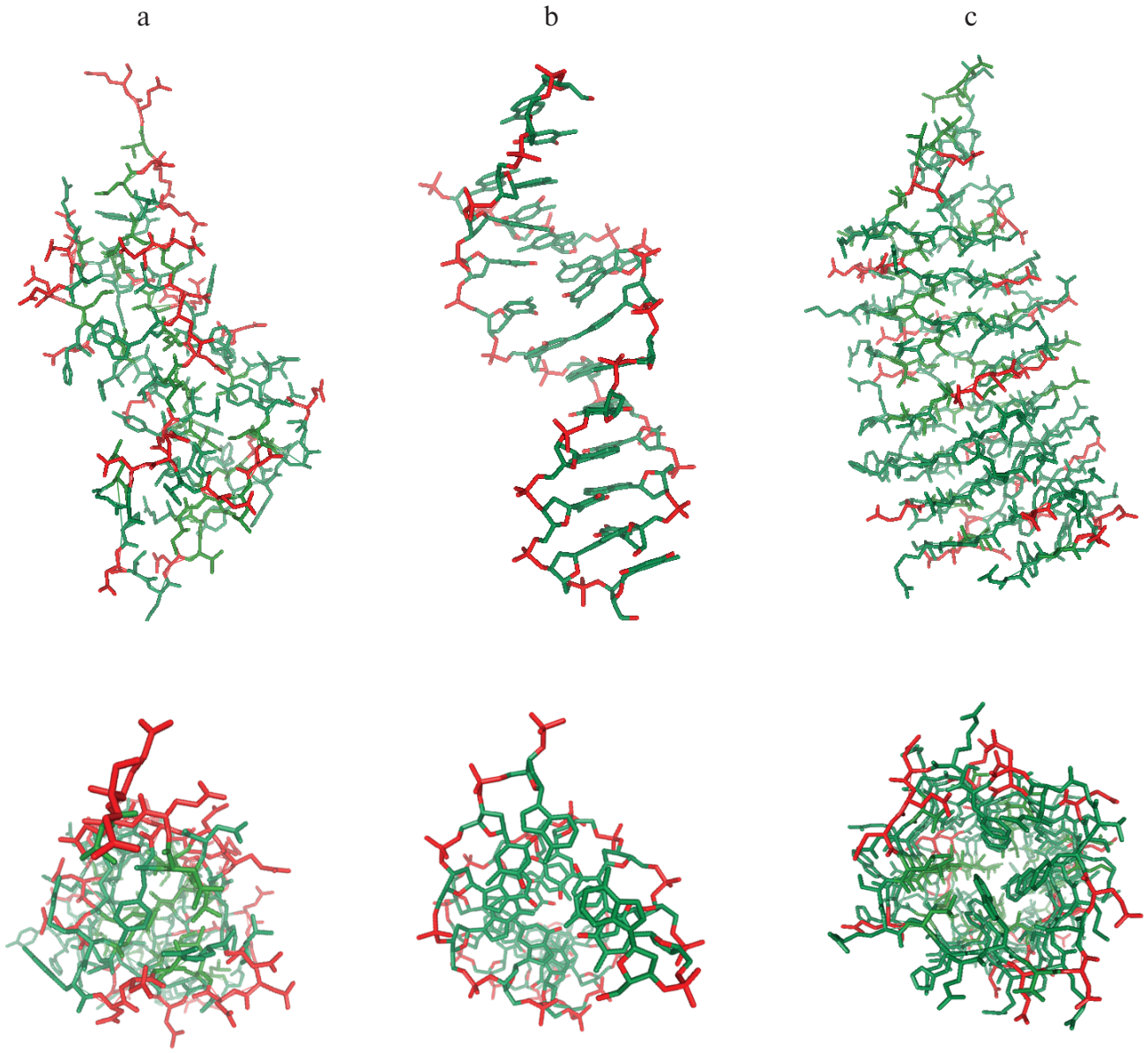

Fig. 4. (G. B. Zavilgelsky and S. M. Rastorguev) Spatial structures of Ocr monomer (a), B-form of DNA duplex (b), and MfpA monomer (c). The upper and lower rows show side and front views, respectively. The diameters of these structures are almost the same.