Biochemistry (Moscow), Vol. 70, No. 11, 2005, pp. 1212-1222. Translated from Biokhimiya, Vol. 70, No. 11, 2005, pp. 1474-1487.

Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM04-338, March 27, 2005.

Analysis of DNA–Protein Interactions in Complexes of Transcription Factor NF-κB with DNA

A. S. Romanenkov[1], A. A. Ustyugov[2], T. S. Zatsepin[1], A. A. Nikulova[2], I. V. Kolesnikov[3], V. G. Metelev[1], T. S. Oretskaya[1], and E. A. Kubareva[4]*

[1]Faculty of Chemistry and [2]Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, 119992 Moscow, Russia; fax: (7-095) 939-3181
[3]Bioline GmbH, Im Biotechnologiepark TGZ 2, Luckenwalde 14943, Germany; fax: +49 (0) 337-168-1244
[4]Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, 119992 Moscow, Russia; fax: (7-095) 939-3181; E-mail: kubareva@belozersky.msu.ru

Received November 10, 2004
Revision received December 9, 2004

Abstract—We have applied bioinformatic analysis of X-ray 3D structures of complexes of transcription factor NF-κB with DNAs. We determined the number of possible Van der Waals contacts and hydrogen bonds between amino acid residues and nucleotides. Conservative contacts in the NF-κB dimer–DNA complex composed of p50 and/or p65 NF-κB subunit and DNA sequences like 5′-G**GG**AMWTT**CC**-3′ were revealed. Based on these results, we propose a novel scheme for interactions between NF-κB p50 homodimer and the κB region of the immunoglobulin light chain gene enhancer (**Ig-κB**). We applied a chemical cross-linking technique to study the proximity of some Lys and Cys residues of NF-κB p50 subunit with certain reactive nucleotides into its recognition site. In all cases, the experimentally determined protein–DNA contacts were in good agreement with the predicted ones.

Key words: transcription factor NF-κB, DNA–protein interaction, hydrogen bond, Van der Waals contact, modified DNA, affinity modification

Recently great attention has been given to transcription factor NF-κB, which is regarded as a protein system elucidating peculiarities in selective activation of gene transcription in response to different pathogenic signals [1, 2]. A significant consequence of the involvement of NF-κB in regulation of activity of multiple genes is strengthening of protective functions of the organism. For instance, in viral infection NF-κB migrates from the cytoplasm into the nucleus, in which it activates transcription of the gene *IFN*-β due to the binding with DNA element **PRDII** of its promoter [3].

Transcription factor NF-κB is a homo- or heterodimer composed of two subunits belonging to the Rel protein family, whose members have a high-homology region called RHR (*Rel* Homology Region) with the product of oncogene *rel* [4]. This family comprises proteins p50, p65 (RelA), p52, RelB, and c-Rel (in mammals), and NF-κB *in vivo* exists mostly in the form of heterodimer composed of subunits p50 and p65 [5].

RHR is localized at the N-termini of Rel proteins and is responsible for dimerization, DNA binding, and transport of active proteins from cytoplasm into the nucleus [6]. Interestingly, the entire high-homology site RHR is necessary for binding of Rel family proteins with DNA, whereas in most transcription factors a small polypeptide fragment can be found possessing virtually the same specificity and affinity to DNA as is the initial protein. This fact indicates that the DNA-binding mode of NF-κB proteins differs from that in common DNA–protein complexes. Luscombe et al. [7] proposed a classification in which transcription factor NF-κB falls into a specific group of proteins characterized by absence of distinctive DNA-binding structural modules.

The DNA site recognized by the transcription factor NF-κB (κB site) is a double-stranded consensus nucleotide sequence of ten base pairs, 5′-G**GG**PuNNPyPy**CC**-3′, in which two exterior conserva-

Abbreviations: IVBP) interior variable base pairs; PDG(**pSS**)) phosphoryldisulfide group; PDB) protein structure databank; Pu) A or G; Py) T or C; N) any of nucleotides; M) A or C; W) A or T; RHR) *Rel* homology region, the high-homology region of *Rel* family proteins; **U*** (**C***)) 2′-*O*-(2-oxoethyl)uridine (2′-*O*-(2-oxoethyl)cytidine); **U**) 2′-deoxy-2′-iodoacetamidouridine; ***i*U**) 5-iodo-2′-deoxyuridine. Prefix "d" (deoxy) is omitted in notations for nucleoside residues, oligodeoxyribonucleotides, and DNA duplexes.

* To whom correspondence should be addressed.

0006-2979/05/7011-1212

tive **G·C** pairs (**GG**-cores) providing specificity of the recognition are divided by interior variable base pairs (IVBP) [8]. The NF-κB proteins have an important ability to modulate expression of various genes depending on nucleotide sequence of the κB site. In particular, NF-κB cannot activate transcription of the gene *IFN*-β if the κB-sequence **PRDII** (5′-GGGAAATTCC-3′) is replaced by the nucleotide sequence **Ig-κB** (5′-GGGACTTTCC-3′) from immunoglobulin light chain gene enhancer [8]. In this case, the loss of transcription activity cannot result from decrease in efficiency of NF-κB binding with other sequence of κB site, because, as in experiments *in vitro*, **Ig-κB** regions bind NF-κB even better than does **PRDII** [8, 9]. Escalante and coworkers [8] hypothesized that the nucleotide sequence of the κB site predetermines the resulting conformation of transcription factor in DNA–protein complexes, thus influencing interaction of NF-κB with other DNA-bound proteins also involved in activation of distinct gene transcription. This suggestion is also supported by the bulk of data from X-ray analysis of various NF-κB complexes, which demonstrates that variations in composition and amount of IVBP in κB sites result in changed spatial orientation of the bound protein about DNA. This conformational mobility of DNA–protein complexes often leads to shifted or disturbed contacts of some amino acid residues of NF-κB protein with κB sites [8, 10, 11].

In this study based on known data from X-ray analysis, we have compared DNA–protein contacts in complexes composed of NF-κB dimers formed by p50 and/or p65 subunits as main components of this protein *in vivo* and various DNAs distinctive by the structure of κB sites [4, 11-15].

We have compared these results with our data on covalent interaction of human transcription factor NF-κB homodimer formed by p50 subunits with reactive DNA derivatives containing a distinct modification type varied in location within the oligonucleotide sequence.

MATERIALS AND METHODS

Computer simulation. Analysis of possible Van der Waals contacts and hydrogen bonds between transcription factor NF-κB and different κB sites of DNA was performed with RasMol software using protein structure databank (PDB) files of DNA–protein complexes (Table 1). We supposed that the hydrogen bond can be established when either hydrogen donors or acceptors in the DNA molecule and corresponding atoms of the protein are within 3.5 Å apart. Similarly, a contact between nonpolar atoms was supposed to be possible when either the C2′-carbon of the sugar DNA moiety or C5-atom of cytosine residue (or the methyl carbon at C5 position of thymine) is no more than 4.5 Å away from a carbon atom of the protein [16, 17].

It is worth noting that RHR sequences of human and murine NF-κB subunits p50 and p65 are homologous with only difference in index numbers of amino acid residues comprising this region. These numbers in murine NF-κB subunits p50 and p65 are three numbers less than that in human p50 and p65 subunits. For example, Cys59 in p50 subunit of murine NF-κB corresponds to the Cys62 in human NF-κB subunit p50. So, to simplify visualization of the data of computer analysis and affinity modification, we used common enumeration of amino acid residues for both proteins, which corresponds to that proposed for murine NF-κB subunits p50 and p65.

A theoretical model for the complex of NF-κB subunit p50 homodimer with **Ig-κB** sequence of DNA was constructed using the program SwissPDBViewer from PDB files of the complexes (p50·p50)–(**Id-κB**) and (p50·p65)–(**Ig-κB**) by way of spatial superposition of their 3D structures at certain protein sites.

Approximate distances between the electrophilic center of the 2′-aldehyde group of DNA and spatially drawn together nucleophilic protein groups were calculated using the program DS ViewerPro Trial 5.0.

Oligonucleotide synthesis. Oligonucleotides used in this study were synthesized by the phosphoroamidite method on an Applied Biosystems 392A (USA) automated synthesizer according to a standard protocol. Oligonucleotides containing 2′-*O*-(2,3-dihydroxypropyl)uridine units were synthesized as described in [18, 19], and oligonucleotides containing phosphoryldisulfide group as described in [20-22].

Enzymes and proteins. T4 polynucleotide kinase (5 U/μl) and terminal deoxynucleotidyl transferase (20 U/μl) were the commercial enzymes purchased from SibEnzyme (Russia) and New England Biolabs (USA). A plasmid carrying the gene encoding human NF-κB p50 subunit with glutathione S-transferase (GST) on its N-terminus was kindly provided by Prof. A. Israel (France). The substitution C59S in the amino acid sequence of the transcription factor was introduced by site-directed mutagenesis in NF-κB p50 coding sequence of the plasmid p50-GST-NF-κB. Site-directed mutagenesis was conducted using a Quikchange (Stratagene, USA) kit containing Pfu DNA polymerase. Polymerase chain reaction was performed with oligodeoxyribonucleotide primers 5′-CGTTTCCGTTATGTAAGTGAAGGCCCATCCCATG-3′ (Nf59f) and 5′-CATGGGATGGGCCTTCACTTACATAACGGAAACG-3′ (Nf59r) (MWG Biotech AG, Germany). Point mutation C59S with absence of mutations in other part of coding sequence of NF-κB p50 subunit was confirmed by primary sequence analysis of the mutant gene performed at MWG Biotech AG. Both chimeric wild-type protein p50-GST NF-κB (wt p50-GST) and mutant p50-GST (C59S) NF-κB were isolated from *Escherichia coli* cell cultures by affinity chromatography on glutathione-agarose [23]. It is known that GST (26 kD) does not have an influence on chimeric

p50 protein binding with DNA [24]. The wt p50-GST and p50-GST (C59S) preparations (both 68.5 kD, 1.0 mg/ml) were stored in buffer A (7.5 mM HEPES, pH 7.5, 34 mM NaCl, 1 mM $MgCl_2$, 0.05 mM EDTA, and 0.5 mM dithiothreitol (DTT)).

Synthesis of DNA duplexes containing 2′-*O*-(2-oxoethyl)uridine residue (2′-aldehyde group). Oligonucleotides containing aldehyde group in 2′-position of the ribose ring were obtained from synthetic oligonucleotides with single 2′-*O*-(2,3-dihydroxypropyl)uridine residues. 5′-^{32}P-labeled oligonucleotide (10 pmol) containing 2′-*O*-(2,3-dihydroxypropyl)uridine residue was dissolved in 15 μl of 30 mM sodium acetate (pH 4-5). Then 15 μl of 0.23 M $NaIO_4$ was added, and the reaction mixture was incubated for 1 h at 25°C. Then 150 μl of 2 M $LiClO_4$ was added followed by precipitation of oligonucleotide with five volumes of acetone, washing with acetone, and air-drying of the pellet. A complementary template (10 pmol) was added to the modified oligonucleotide dissolved in 20 μl of buffer A and annealed at 70°C.

Interaction of DNA duplexes containing 2′-aldehyde group with NF-κB p50-GST homodimer. *Gel retardation method.* 5′-^{32}P-labeled oligonucleotide duplexes **I-XI** and duplex **K** (0.5 pmol) were incubated with 10 pmol of NF-κB p50-GST in 10 μl of the buffer A containing 10% glycerol and 0.25 μg of poly(dI·dC) for 20 min at 25°C. After the incubation, reaction mixtures were analyzed by PAGE in non-denaturing 6% polyacrylamide gel in TBE buffer (50 mM Tris-borate, pH 8.3, and 1 mM EDTA).

Cross-linking. DNA duplexes containing 2′-aldehyde group were reacted with NF-κB p50-GST homodimer under the previously optimized conditions [25]. 5′-^{32}P-labeled oligonucleotide duplexes **I-XI** (0.5 pmol) were incubated with 10 pmol of NF-κB p50-GST in 10 μl of buffer A containing 10% glycerol and 0.25 μg of poly(dI·dC) for 30 min at 25°C. Subsequently, 1 μl of 275 mM $NaBH_3CN$ (final concentration 25 mM) was added, and the reaction mixture was incubated for 1 h at 25°C. Reaction products were separated by SDS-PAGE in 8% polyacrylamide gel with 4% concentrating gel strip. Before application on the gel, the reaction mixtures were maintained in denaturing buffer B (60 mM Tris-HCl, pH 6.8, 1% glycerol, 2% SDS, and 0.05% 2-mercaptoethanol) for 3 min at 95°C.

Cross-linking of DNA duplexes containing phosphoryldisulfide group with NF-κB p50-GST homodimer. 3′-^{32}P-labeled with terminal transferase and [α-^{32}P]UTP

Table 1. Complexes of the transcription factor NF-κB dimers with synthetic DNA duplexes, which were investigated by X-ray analysis and presented in PDB

PDB code of the complex	Structure of the DNA duplex (κB site is underlined)	Nature of κB site	Composition of NF-κB dimer (origin)	Resolution of X-ray analysis (Å)	Reference
1SVC	5'-AGAT**GGGGAATCCC**CTAGA 3'-AGAT**CCCCTAAGGG**GTAGA	κB site of the **MHC** (major histocompatibility complex) group I enhancer	p50 · p50 (human)	2.6	[12]
1NFK	5'-T**GGGAATTCCC** 3'- **CCCTTAAGGG**T	"ideal" κB site (**Id-κB**)	p50 · p50 (mouse)	2.3	[13]
1RAM, 2RAM	5'-CGGCT**GGAAATTTCC**AGCCG 3'-GCCGA**CCTTTAAAGG**TCGGC	P-sequence of the interleukin-4 gene promoter (**κB-33**)	p65 · p65 (mouse)	2.4	[14]
1VKX	5'-TG**GGGACTTTCC** 3'- C**CCCTGAAAGG**A	κB site of the immunoglobulin light chain gene enhancer (**Ig-κB**)	p65 · p50 (mouse)	2.9	[4]
*1LEI**	5'-CTCA**GGG-A-A-AGT-AC**AGA 3'-GAGT**CCC*i*U*i*U*i*UCA*i*UG**TCT	κB site of the urokinase gene promoter (**uPA-κB**)	p65 · p50 (mouse)	2.7	[11]
1LE5 (two complexes)	5'-T**GGGAAATTCC**T 3'- **CCCTTTAAGG**AA	κB site of the β-interferon gene promoter (**PRDII**)	p65 · p50 (mouse)	2.75	[15]
1LE9 (two complexes)	5'-T**GGGACTTTCC**T 3'- **CCCTGAAAGG**AA	κB site of the immunoglobulin light chain gene enhancer (**Ig-κB**)	p65 · p50 (mouse)	3.0	[15]

* ***i*U**, 5-iodo-2′-deoxyuridine.

oligonucleotide duplexes **XII-XIV** (0.5 pmol) were incubated with 10 pmol of either p50-GST or p50(C59S)-GST in 10 μl of DTT-free buffer A containing 10% glycerol and 0.25 μg of poly(dI·dC) for 30 min at 25°C. Reaction products were separated by SDS-PAGE in 8% polyacrylamide gel with 4% concentrating gel strip. Before application on the gel, the reaction mixtures were dissolved in denaturing buffer B without 2-mercaptoethanol.

RESULTS AND DISCUSSION

We have picked eight PDB files of the X-ray-analyzed complexes between various dimers of the transcription factor NF-κB and synthetic DNA duplexes (Table 1). The main criterion for this selection was the presence of p50 and/or p65 subunits of the transcription factor NF-κB in the complex with DNA. Data processing revealed that the complexes ***1LE5*** and ***1LE9*** have two double-strands of DNA and, correspondingly, four protein subunits (two p50 and two p65) in the asymmetric unit of their crystal lattice cell. To simplify the analysis, each complex was divided to have any file containing only one DNA helix bound with proteins. DNA–protein contacts found in various complexes were determined using the RasMol software. Both hydrogen bonds and nonpolar carbon–carbon interactions of amino acid residues of the protein with heterocyclic bases and sugar-phosphate backbone of the DNA molecule were taken in account during the analysis.

The analysis of complexes ***1SVC*** [12], ***1NFK*** [13], and ***1RAM*** and ***2RAM*** [14] with p50 and p65 homodimers revealed several changes in position of some individual DNA–protein contacts that are not characteristic of other studied complexes. This phenomenon is apparently associated with unusual structure of κB sites in DNA duplexes that were crystallized (Table 1).

In the complex ***1SVC***, the NF-κB p50 homodimer was crystallized with DNA duplex containing non-complementary pair A/A in the center of a pseudo-symmetric κB site [12]. In complex ***1NFK***, unlike the consensus structure of κB site with five IVBP between **GG** cores, the exterior **GG** cores of "ideal" κB site of the DNA duplex (**Id-κB**) composed of palindrome oligonucleotide are divided by six interior base pairs [13].

In complexes ***1RAM*** and ***2RAM***, the homodimer of NF-κB subunit p65 was crystallized with DNA duplex containing a pseudo-symmetric (**κB-33**) κB site, whose structure (5′-GAAAATTTCC-3′) differs from that of the consensus one. This κB sequence of DNA only binds specifically with the homodimer p65 [14]. Chen et al. [26] also described the crystal structures of two other complexes of p65 homodimer with unusually structured κB sites, namely pseudo-symmetric κB site from the gene encoding interleukin 8 (**IL-8**: 5′-GGAATTTCC-3′) and asymmetric κB site from enhancer of the gene encoding collagen type VII (**κB-55**: 5′-GGAATTCCC-3′), but did not present them as PDB files. The data of the study [26] are indicative of unique ability of p65 homodimer (which differentiates it from other transcription factor NF-κB dimers) to bind specifically κB sites of various length and composition wherein the segment Arg41-Ser42-Ala43 provides conformational mobility of DNA-binding sites of the protein.

It is worth noting that in natural systems transcription factor NF-κB heterodimer composed of subunits p50 and p65 (p50·p65) is most prevalent. It is the heterodimer that plays a pivotal role in regulation of gene expression. All of the analyzed complexes of NF-κB heterodimer with DNA (***1VKX***, ***1LEI***, ***1LE5***, and ***1LE9***) are characterized by a common feature (Table 1): despite the difference in nucleotide sequences, their κB sites are of similar structure in which the exterior **GG** cores are separated by five IVBP [4, 11, 15]. Complex ***1LEI*** is noticeable for its modified κB site in which thymidines are substituted by 5-iodo-2′-deoxyuridines [11]. The structure of the complex ***1LE5*** between the transcription factor NF-κB heterodimer p50·p65 and DNA duplex containing κB site of the β-interferon gene promoter (**PRDII**) [15] was also studied by Escalante et al. using X-ray analysis [8]. Despite the absence of a corresponding PDB file in the database, a scheme of DNA–protein contacts described in [8] follows the pattern of these contacts in complex ***1LE5***.

After comparison of X-ray analysis data on interaction of NF-κB subunits p50 and p65 with the common-structure DNA sites in which exterior **GG** cores are separated by five interior variable base pairs, we determined conservative DNA–protein contacts formed in all of the analyzed complexes ***1VKX***, ***1LEI***, ***1LE5***, and ***1LE9***. Moreover, we compared the nucleotide sequences of κB sites in these complexes and aligned amino acid sequences of murine NF-κB subunits p50 and p65 using the GeneDoc software (Fig. 1).

Relying on these data, we made a scheme of conservative contacts between a transcription factor NF-κB dimer composed of p50 and/or p65 subunits and κB site 5′-G**GG**AMWTT**CC**-3′ (Fig. 2). Aligned amino acid sequences corresponding to each other and forming identical contacts has the same alphabet code (for example, Lys123 (in p65) = Lys144 (in p50) = Lys A, Fig. 1). The order of code appointment to amino acid residues corresponds to their sequence in the protein chain.

One can see from the generalized scheme (Fig. 2) that a number of conservative contacts between NF-κB dimer and the recognizing site of DNA are significant. The numbers of hydrogen bonds are 13 with sugar-phosphate backbone and 21 with heterocyclic DNA bases. Carbon–carbon contacts with the sugar-phosphate backbone of DNA are only formed by residues GlnB, TyrA, and CysA, and those with heterocyclic bases by TyrA and GluA. It is important that one of the subunits (p50) is in

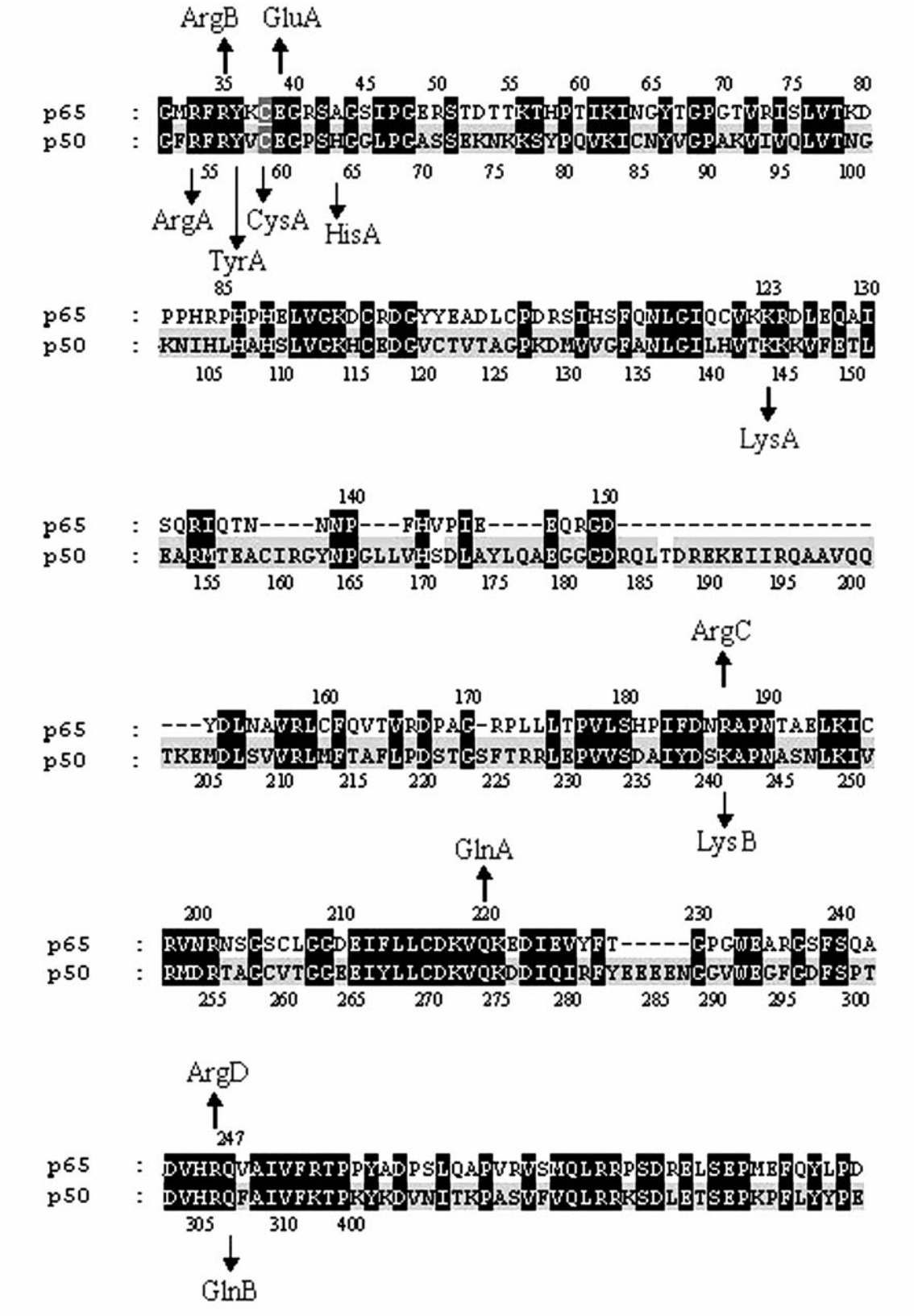

Fig. 1. Comparative analysis of amino acid sequences of proteins p50 and p65 performed using the GeneDoc software. Numeration for murine NF-κB p65 subunit is given above the amino acid sequences, and the numeration for murine NF-κB p50 subunit is given below. The amino acid residues corresponding to each other according to alignment and involved in conservative contacts with DNA in DNA–protein complexes are shown separately.

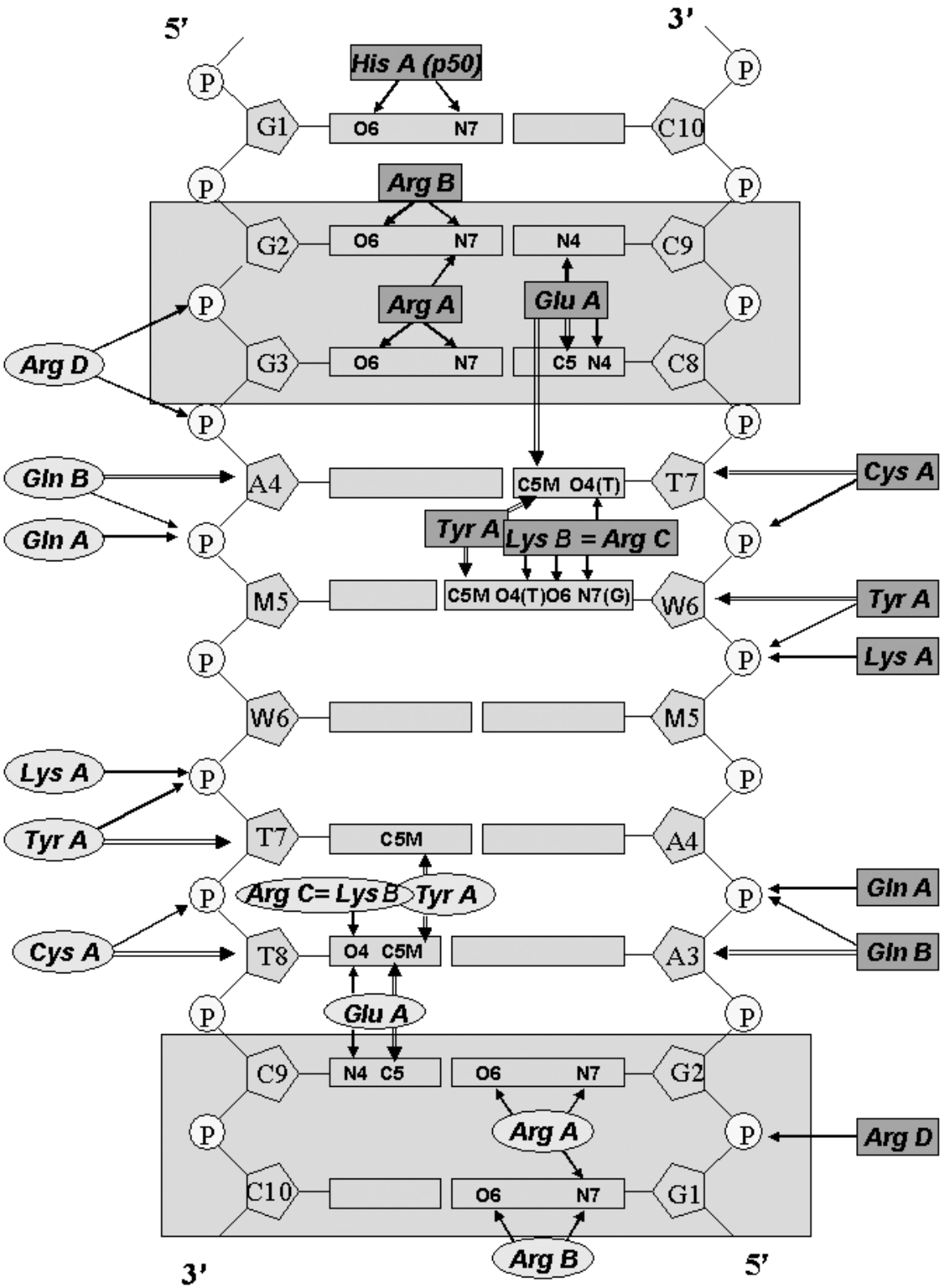

Fig. 2. Generalized scheme of conservative contacts between transcription factor NF-κB dimer composed of p50 and/or p65 subunits and κB site 5′-G**GG**AMWT**CC**-3′ (exterior **GG** cores are separated by five interior base pairs). Double arrow (⇒) indicates nonpolar carbon–carbon contacts; single arrow (→) indicates hydrogen bonds; rectangles encase amino acids belonging to one of the subunits and ellipses to other subunit; **GG** cores are in frames.

preferable contact with heterocyclic bases of the "half-site" 5′-G**GG**AM, whereas the other (p65) with heterocyclic "half-site" TT**CC**-3′. The determining factor of this preference is the establishment of hydrogen bond between the NH-group (proton donor) of the imidazole ring of His64 in the p50 subunit and atoms N7 or O6 (proton acceptors) of a guanine residue localized just before the **GG** core of sequence 5′-G**GG**AM. In subunit p65, the residue corresponding to His64 is Ala43, which cannot form a hydrogen bond with heterocyclic base and, correspondingly, cannot participate in DNA recognition. Otherwise the subunits p50 and p65 interact with heterocyclic bases of both "half-sites" in a similar way, due to formation of hydrogen bonds and Van der Waals contacts with amino acid residues of the recognizing loop L1: ArgA, ArgB, TyrA, and GluA.

X-Ray analysis apparently provides the most reliable information concerning the structure of DNA–protein complex, but despite considerable achievements in studies of DNA-binding proteins, X-ray analysis in some cases does not reveal conformational changes of proteins upon their interaction with DNA. So, we were to determine how much the proposed model of conservative contacts between NF-κB protein and DNA were in agreement with experimental data on affinity modification of transcription factor NF-κB with various reactive DNA analogs selectively reacting with distinct amino acid residues involved in DNA–protein complex formation.

The proposed scheme of conservative contacts (Fig. 2) shows that several Lys residues and one Cys, which form a portion of hydrogen bonds between DNA and protein, are involved in interaction of transcription factor NF-κB dimer with κB site 5′-G**GG**AMWTT**CC**-3′. In our laboratory, we synthesized a set of DNA reagents that can selectively interact with amino- and thio-groups of protein amino acid residues Lys and Cys, respectively [18, 19, 21, 22, 25, 27-30]. These reagents are synthetic duplexes containing a reactive group in a specified position of the sugar-phosphate backbone of the κB site. These reactive groups either modified the 2′-position of a nucleotide unit (2′-aldehyde (Fig. 3a) and 2′-iodoacetamide [30] groups) or substituted the natural phosphodiester bond (internucleotide phosphoryldisulfide (Fig. 3b) and modified pyrophosphate groups [28, 29]). To get an exhaustive interpretation of our data on cross-linking of transcription factor NF-κB with reactive DNA analogs and to confirm the expedience of particular modification type for probing of DNA–protein complex structures, we were to generalize the data of X-ray analysis and construct schemes of hydrogen bond and Van der Waals contact formation between those κB sites of DNA and those NF-κB proteins that were studied in our experiments.

The following modified DNA duplexes containing one of two recognition site types were used for affinity modification of transcription factor NF-κB: the first containing the "ideal" κB site (**Id-κB**: 5′-**GG**GAATTC**CC**-3′) and the second containing the κB sequence of the immunoglobulin light chain gene enhancer (**Ig-κB**: 5′-G**GG**ACTTT**CC**-3′). These κB sites differ from each other in nucleotide sequence and number of base pairs between **GG** cores. We used human transcription factor NF-κB p50 homodimer as a model protein structure which, like a p50·p65 heterodimer, effectively binds to κB sites of DNA [31].

X-Ray analysis data are available for the complex of NF-κB p50 homodimer with "ideal" κB sequence of DNA (***1NFK***, Table 1), and we have constructed the scheme of hydrogen bond and carbon–carbon contact formation from these data (Fig. 4a). Similar data are absent for the complex of NF-κB p50 homodimer with **Ig-κB** site of DNA. So, we used the SwissPDBViewer software to make a theoretical model for this complex. With this aim, we took the PDB files of the complexes ***1VKX*** (p50·p65–(**Ig-κB**)) and ***1NFK*** (p50·p50–(**Id-κB**)). Atomic coordinates of these complexes were brought into coincidence by homologous amino acid sequences of the p50 and p65 subunits, which are involved in interaction of these proteins with DNA according to the previously proposed scheme of conservative contacts (Fig. 2). Then subunit p65 from the complex of p50·p65 heterodimer with **Ig-κB** DNA was replaced by the p50 subunit from the complex ***1NFK***. Finally, we examined possible intermolecular hydrogen bonds and Van der Waals contacts between protein and DNA carbons in resulting complex

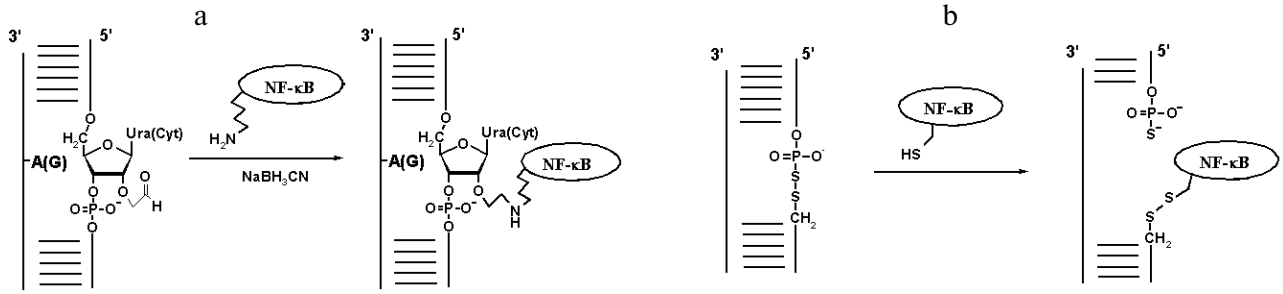

Fig. 3. Affinity modification of lysine and cysteine residues of transcription factor NF-κB with modified DNA duplexes containing a 2′-*O*-(2-oxoethyl)uridine (2′-*O*-(2-oxoethyl)cytidine) residue (a) or a phosphoryldisulfide group (b).

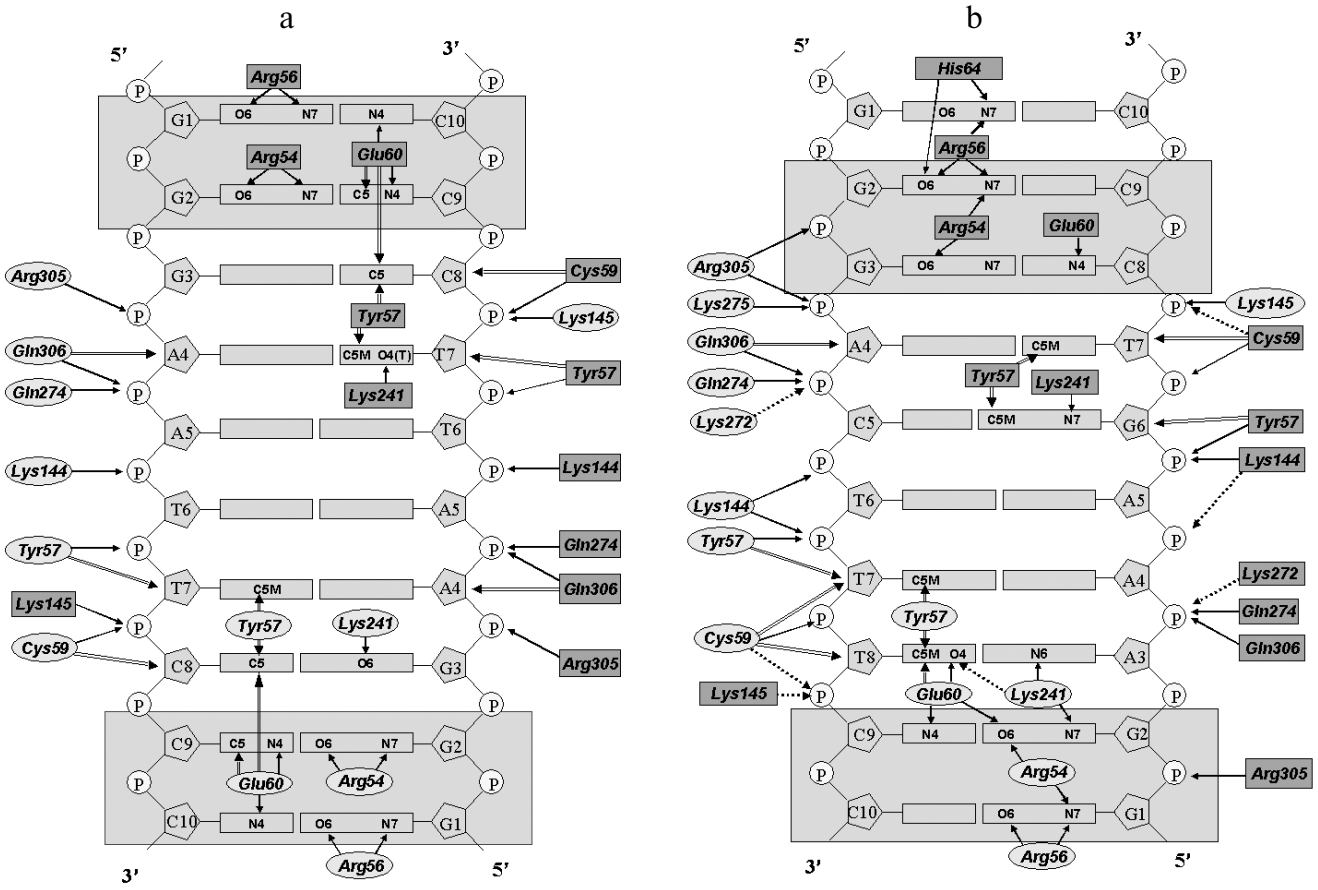

Fig. 4. A scheme for interaction of transcription factor NF-κB p50 subunit homodimer with: "ideal" κB sequence **Id-κB** (exterior **GG** cores are separated by six IVBP) (a) and κB sequence from the immunoglobulin light chain enhancer **Ig-κB** (exterior **GG** cores are separated by five IVBP) (b). Double arrows (⇒) show nonpolar carbon–carbon contacts, single arrows (→) show hydrogen bonds, and dashed arrows (···>) show contacts permitting hydrogen bond formation (3.5-4.0 Å); amino acids belonging to one of subunits are enclosed by rectangles and to another subunit by ellipses; **GG** cores are in frames.

p50·p50–(**Ig-κB**) using the RasMol software (Fig. 4b). It is worth noting that along with conservative DNA–protein contacts (Fig. 2), a great number of additional hydrogen bonds with Lys residues, such as Lys145, Lys272, and Lys275, is established in the complex p50·p50–(**Ig-κB**). Hence, the scheme of DNA–protein contacts in the complex p50·p50–(**Ig-κB**) (Fig. 4b) rather than the generalized scheme of conservative contacts between NF-κB dimer and the κB site 5′-G**GG**AMWTT**CC**-3′ is more applicable for further interpretation of affinity modification data.

Affinity modification of ε-amino groups of Lys residues in a NF-κB p50 homodimer was performed using modified DNAs, namely oligonucleotide duplexes containing aldehyde group in 2′-position of the sugar moiety [18, 19, 27]. The reaction product is a Schiff base which needs to be reduced into the secondary amine with $NaBH_3CN$ (Fig. 3a) [25].

To ascertain peculiarities of complex formation and cross-linking of NF-κB p50 homodimer with 2′-aldehyde-containing DNA derivatives, we constructed a panel of oligonucleotide duplexes **I-XI** containing single 2′-*O*-(2-oxoethyl)uridine (**U***) or 2′-*O*-(2-oxoethyl)cytidine (**C***) residue in varied position of oligonucleotide strand (Table 2). The duplexes **I-VII** contain κB site (underlined) from the immunoglobulin light chain gene enhancer (**Ig-κB**), and the duplexes **VIII-X** – the "ideal" κB site (**Id-κB**). The duplex **XI**, which contains no κB site, served as a control.

We studied, using the gel-retardation method, an ability of modified DNA duplexes containing κB site to form specific complexes with p50 homodimer. The unmodified 5′-^{32}P-labeled DNA duplex **K** containing κB site (underlined) from the immunoglobulin light chain gene enhancer (**Ig-κB**) (Table 2) served as a positive control. Following the incubation under conditions conducive to formation of specific protein–nucleic complex, the reaction mixtures were assayed by PAGE in nondenaturing 6% polyacrylamide gel (Fig. 5a). As it is evident from the picture, the protein effectively binds the

Table 2. Structures of DNA duplexes containing reactive groups*

K	5'-ACCTC**GGAAAGTCCC**CTCT GAG**CCTTTCAGGG**GAGA	**I**	5'-ACCTC**GGU*AAGTCCC**CTCT GAG**CCA TTCAGGG**GAGA
II	5'-ACCTC**GGAU*AGTCCC**CTCT GAG**CCTA TCAGGG**GAGA	**III**	5'-ACCTC**GGAAU*GTCCC**CTCT GAG**CCTTA CAGGG**GAGA
IV	5'-ACCTC**GGAAAU*TCCC**CTCT GAG**CCTTTA AGGG**GAGA	**V**	5'-ACCTC**GGAAAGU*CCC**CTCT GAG**CCTTTCA GGG**GAGA
VI	5'-ACCTC**GGAAAGTC*C**CTCT GAG**CCTTTCAG GG**GAGA	**VII**	5'-ACCTC**GGAAAGTCC*C**CTCT GAG**CCTTTCAGG G**GAGA
VIII	5'-T**GGGAU*TTCCC**CTC A**CCCTA AAGGG**GAG	**IX**	5'-T**GGGAAU*TCCC**CTC A**CCCTTA AGGG**GAG
X	5'-T**GGGAATU*CCC**CTC A**CCCTTAA GGG**GAG	**XI**	5'-CTCCCAGGC**U***CAAAT GAGGGTCCGA GTTTA
XII	5'-ACCTC**GGAAAGpSSTCCC**CTCT GAG**CCTTTC – AGGG**GAGA	**XIII**	5'-ACCTC**GGAAAGTpSSCCC**CTCT GAG**CCTTTCA – GGG**GAGA
XIV	5'-TCGGAAAGT**pSS**TGACTGCACGGT GCCTTTCA – ACTGACGT	**XV**	5'-TC**GGAAAGTCCCC**<u>**U**</u>C AG**CCTTTCAGGG**GAG
XVI	5'-TC**GGAAAG**<u>**U**</u>**CCC**CTC AG**CCTTTCAGGG**GAG	**XVII**	5'-T**GGGAA**<u>**U**</u>**TCCC**CTC A**CCCTTAAGGG**GAG
XVIII	5'-T**GGGAAT**<u>**U**</u>**CCC**CTC A**CCCTTAAGGG**GAG		

* **U*(C*)**, 2′-*O*-(2-oxoethyl)uridine (2′-*O*-(2-oxoethyl)cytidine); **pSS**, phosphoryldisulfide group; <u>**U**</u>, 2′-deoxy-2′-iodoacetamidouridine.

modified duplexes **I-X** and does not interact with duplex **XI**, which lacks the κB site.

Cross-linking of 5′-^{32}P-labeled duplexes **I-XI** with p50 homodimer was conducted as described in "Materials and Methods". Electrophoretic analysis of reaction mixtures demonstrated that the duplexes **I-III** and **VIII** effectively interact with the protein to form a covalent bond. Maximum yield of conjugate was about 16% (Fig. 5b). The duplexes **IV-VII** and **IX-X** demonstrated low efficiency of cross-linking, which was comparable to that of nonspecific duplex **XI**.

In modified duplexes **I-X** with different primary structure of κB site, the aldehyde group was in 2′-position of the sugar moiety of a certain nucleotide that is in close proximity of its 3′-phosphate group. One can suppose that a given DNA type interacted with those lysine residues of p50 homodimer that were in proximity with this area of sugar-phosphate backbone. According to the schemes of contacts between p50 homodimer and recognizing **Id-κB** and **Ig-κB** sites of DNA (Fig. 4), the aldehyde group of duplexes **III** and **VIII** is localized close to Lys144, and aldehyde group of duplex **I** to Lys272. It remained unclear what lysine residue might be involved in formation of a covalent bond between NF-κB p50 and DNA duplex **II**. To reveal this, we used the DS ViewerPro Trial 5.0 software and PDB file of the p50·p50–(**Ig-κB**) complex created by the SwissPDBViewer program to estimate the distances between the electrophilic center of 2′-aldehyde group of duplex **II** and adjacent ε-amino groups of p50 lysines. We found that the 2′-reactive group comprising this DNA duplex can theoretically form a covalent bond with Lys241 of NF-κB p50.

Earlier, affinity modification of lysine residues of proteins, such as transcription factor NF-κB, was successfully performed with specific DNA duplexes containing another modification type of sugar-phosphate backbone, namely internucleotide substituted pyrophosphate (SP) groups [28, 29]. Substituted pyrophosphate groups were inserted in varied position of DNA duplex **K** (Table 2) containing **Ig-κB** site (figures denote the numbers of 5′-nucleoside phosphates in the recognition area):

5′ A-C-C-T-C-**^{1}G-^{2}G-^{3}A-^{4}A-^{5}A-^{6}G-^{7}T-^{8}C-^{9}C-^{10}C**-C-T-C-T-3′

3′ G-A-G–**C–C–T–T–T–C–A–G–G—G**-G-A-G-A-5′.

Cross-linking of NF-κB p50 homodimer with DNA duplexes containing the internucleotide SP group in positions 5, 6, and 8 occurred with high efficiency (conjugate yields were 50, 30, and 35%, respectively). The data for positions 5 and 6 coincide with those obtained from affinity modification of NF-κB p50 homodimer with 2′-aldehyde-containing DNAs. These are precisely the modifi-

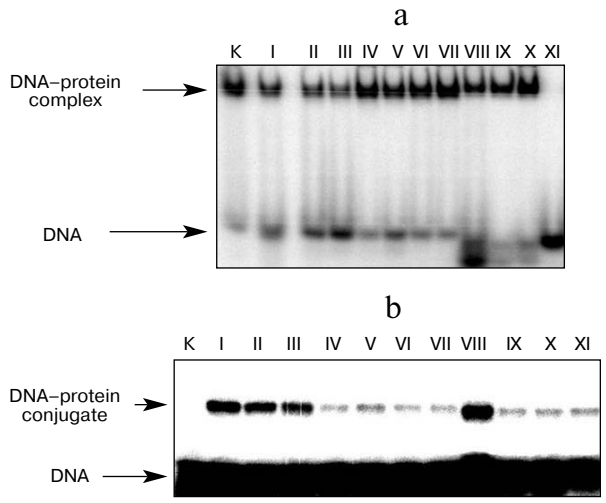

Fig. 5. Interaction of NF-κB p50-GST homodimer with DNA duplexes containing 2′-*O*-(2-oxoethyl)uridine or 2′-*O*-(2-oxoethyl)cytidine residues **U*** or **C*** (**I-X**) in their κB sites. a) Radioautograph of 6% polyacrylamide gel after non-denaturing electrophoresis of reaction mixtures. b) Radioautograph of 8% polyacrylamide gel after denaturing SDS-PAGE of reaction mixtures. The numbers of DNA duplexes are given above the lanes. The unmodified κB-containing duplex **K** was used as a positive control, and the modified κB site-free duplex **XI** as a negative control.

cation positions of κB sites of duplexes **II** and **III** that form covalent bonds with the protein, presumably with its Lys241 and Lys144 residues (Fig. 4b). It is worth noting that DNA duplex containing SP group in position 4, unlike the 2′-aldehyde-containing duplex **I**, whose modification is localized in the same position of sugar-phosphate backbone of DNA, does not form a covalent bond with p50 homodimer. Insertion of a substituted pyrophosphate group most likely leads to local distortion of the structure of DNA duplex, thus counteracting its interaction with the Lys272 residue of NF-κB p50. One can explain formation of a conjugate between modified DNA duplex containing SP group in position 8 and NF-κB p50 homodimer from the interaction between the corresponding area of sugar-phosphate backbone of DNA ligand and the Lys145 residue, according to the proposed scheme of contacts (Fig. 4b). The results presented will be further updated in experiments with mutant p50 proteins containing substitutions of corresponding lysine residues by non-reactive residues.

Affinity modification of the Cys residue involved in formation of contacts between the transcription factor NF-κB and DNA was performed using oligonucleotide duplexes carrying two modifications: internucleotide phosphoryldisulfide and 2′-iodoacetamide groups. These groups interact with sulfhydryl group of cysteine: the first one due to thiol-disulfide exchange (Fig. 3b), and the second due to alkylation [30].

Oligonucleotides containing phosphoryldisulfide group (PDG) in a specified position of the sugar-phosphate backbone were prepared by chemical ligation as previously described [21, 22]. We synthesized DNA duplexes **XII** and **XIII** containing PDG(**pSS**) in the **Ig-κB** site (Table 2). The modified duplex **XIV** lacking a κB site served as a control in investigation of cross-linking specificity.

It was shown earlier that an attack of a nucleophilic group of a cysteine residue is directed to electrophilic sulfur atom of PDG that is bound to a methylene group (Fig. 3b) [21, 22]. Cross-linking of 3′-^{32}P-labeled duplexes **XII-XIV** with NF-κB p50 homodimer was performed in a buffer, pH 7.5, for 30 min in the absence of reducing agents. Electrophoretic analysis of reaction mixtures showed that only duplex **XIII** effectively interacts with the protein (the yield of conjugate was 18%). The yield of conjugate formed by duplex **XII** with p50 homodimer was far less (about 5%), which is comparable with that of conjugate between NF-κB p50 and nonspecific duplex **XIV** (Fig. 6). Using a mutant protein NF-κB p50(C59S)-GST, we confirmed involvement of Cys59 in formation of the covalent bond with DNA. The substitution of Cys59 in p50 subunit by Ser had virtually no effect on formation of specific DNA–protein complex, but counteracted cross-linking with duplexes **XII-XIV** (Fig. 6).

In accordance with the proposed scheme of DNA–protein contacts p50·p50–(**Ig-κB**) (Fig. 4b), the Cys59 residue forms a hydrogen bond with that fragment of the sugar-phosphate backbone of the κB site that corresponds to the modification position in duplex **XII**. However, p50 homodimer covalently binds DNA duplex **XIII** (Fig. 6). One can suppose that a given cysteine residue contacts not only with 5′-phosphate of the seventh nucleotide residue of the **Ig-κB** site, but also is close enough to its 3′-phosphate group (Fig. 4b). Modification can distort the DNA structure, thus influencing the pattern of contacts with protein amino acid residues.

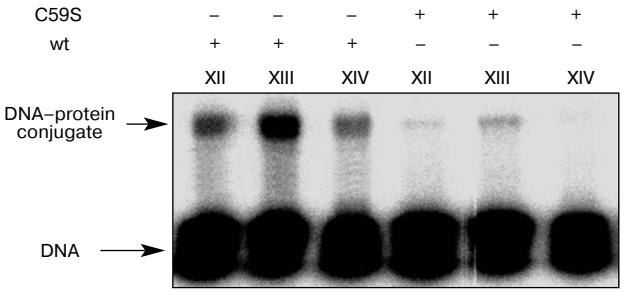

Fig. 6. Products of cross-linking of transcription factor NF-κB p50-GST homodimer with DNA duplexes containing phosphoryldisulfide group (**XII-XIV**). Radioautograph of 8% polyacrylamide gel after SDS-PAGE of reaction mixtures of modified duplexes containing (**XII** and **XIII**) or not containing (**XIV**) κB site with wild type NF-κB p50-GST (wt) or its mutant p50(C59S)-GST. DNA duplex numbers are given above the lanes.

Earlier, we used 2′-iodoacetamide-containing DNA duplexes for affinity modification of the transcription factor NF-κB p50 homodimer [30]. These modified DNAs interacted with Cys residues of the protein in buffer systems with pH 8.3 for no more than 15 min.

To study the interaction of a given modified DNA type with the transcription factor NF-κB p50 homodimer, we constructed oligonucleotide duplexes **XV-XVIII** [30] containing 2′-deoxy-2′-iodoacetamidouridine (**U**) residues in **Ig-κB** and **Id-κB** sites (Table 2). Only DNA duplexes **XVI** and **XVIII** can effectively form conjugates with p50 (the yields of reaction products were 25 and 21%, respectively) [30]. Formation of conjugate between the (**Id-κB**)-containing duplex **XVIII** and p50 homodimer can be explained from a close proximity of Cys59 to the iodoacetamide group (Fig. 4a). The covalent interaction between the (**Ig-κB**)-containing duplex **XVI** and the cysteine residue of the protein is also consistent with the data obtained for DNA duplex **XIII** containing phosphoryldisulfide group (Figs. 4b and 6).

Our experimental data on affinity modification of NF-κB p50 homodimer with different modified DNA duplexes correspond to proposed schemes of DNA–protein contacts. These schemes include contacts between NF-κB p50 homodimer with **Ig-κB** and **Id-κB** sequences, as well as the generalized scheme of conservative contacts between the transcription factor NF-κB dimer composed of its p50 and/or p65 subunits and the κB site 5′-G**GG**AMWTT**CC**-3′. Thus, on one hand, we have demonstrated that the modified DNA used can be effective in examination of DNA–protein contacts in various complexes with some DNA-recognizing proteins. On the other hand, we have proposed schemes of DNA–protein contacts that provide a basis for creation of novel DNA ligands possessing ability to irreversibly bind with transcription factor NF-κB due to the interaction of reactive group inserted in a specified position of a κB site of DNA and definite amino acid residue of the protein. We think that this kind of DNA duplex will compete with the cell DNA for NF-κB, thus regulating activity of this transcription factor.

We are grateful to A. V. Alekseevskii and S. A. Spirin, collaborators of Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, for advice on use of the software.

This study was supported by the State Program "Universities of Russia" (project No. 07.02.567) and the Russian Foundation for Basic Research (projects No. 03-04-48957 and 04-04-48714).

REFERENCES

1. Ghosh, S., May, M. J., and Kopp, E. B. (1998) *Annu. Rev. Immunol.*, **16**, 225-260.
2. Baldwin, A. S., Jr. (1996) *Annu. Rev. Immunol.*, **14**, 649-681.
3. Thanos, D., and Maniatis, T. (1992) *Cell*, **71**, 777-789.
4. Chen, F. E., Huang, D.-B., Chen, Y.-Q., and Ghosh, G. (1998) *Nature*, **391**, 410-413.
5. Sen, R., and Baltimore, D. (1986) *Cell*, **47**, 921-928.
6. Grimm, S., and Baeuerle, P. A. (1993) *Biochem. J.*, **290**, 297-308.
7. Luscombe, N. M., Laskowski, R. A., and Thornton, J. M. (2001) *Nucleic Acids Res.*, **29**, 2860-2874.
8. Escalante, C. R., Shen, L., Thanos, D., and Aggarwal, A. K. (2002) *Structure*, **10**, 383-391.
9. Menetski, J. P. (2000) *J. Biol. Chem.*, **275**, 7619-7625.
10. Mueller, C. W., Rey, F. A., and Harrison, S. C. (1996) *Nature Struct. Biol.*, **3**, 224-227.
11. Chen-Park, F. E., Huang, D.-B., Noro, B., Thanos, D., and Ghosh, G. (2002) *J. Biol. Chem.*, **277**, 24701-24708.
12. Mueller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L., and Harrison, S. C. (1995) *Nature*, **373**, 311-317.
13. Ghosh, G., van Duyne, G., Ghosh, S., and Sigler, P. (1995) *Nature*, **373**, 303-310.
14. Chen, Y. Q., Ghosh, S., and Ghosh, G. (1998) *Nature Struct. Biol.*, **5**, 67-73.
15. Berkowitz, B., Huang, D.-B., Chen-Park, F. E., Sigler, P. B., and Ghosh, G. (2002) *J. Biol. Chem.*, **277**, 24694-24700.
16. Cramer, P., Larson, C. J., Verdine, G. L., and Mueller, C. W. (1997) *EMBO J.*, **16**, 7078-7090.
17. Alexeevski, A., Spirin, S., Alexeevski, D., Klychnikov, O., Ershova, A., Titov, M., and Karyagina, A. (2003) *Biophysics (Moscow)*, **48** (Suppl. 1), 45-55.
18. Kachalova, A. V., Zatsepin, T. S., Romanova, E. A., Stetsenko, D. A., Gait, M. J., and Oretskaya, T. S. (2000) *Nucleosides, Nucleotides, Nucleic Acids*, **19**, 1693-1707.
19. Zatsepin, T. S., Stetsenko, D. A., Gait, M. J., and Oretskaya, T. S. (2005) *Tetrahedron Lett.*, **46**, 3191-3195.
20. Metelev, V. G., Borisova, O. A., Volkov, E. M., Oretskaya, T. S., and Dolinnaya, N. G. (2001) *Nucleic Acids Res.*, **29**, 4062-4069.
21. Metelev, V. G., Kubareva, E. A., Vorob'eva, O. V., Romanenkov, A. S., and Oretskaya, T. S. (2003) *FEBS Lett.*, **538**, 48-52.
22. Romanenkov, A. S., Borisova, O. A., Kubareva, E. A., Oretskaya, T. S., and Metelev, V. G. (2003) *Vestn. Mosk. Univ. Ser. 2: Khim.*, **44**, 208-213.
23. Tanaka, H., Vickart, P., Bertrand, J. R., Rayner, B., Morvan, F., Imbach, J. L., Paulin, D., and Malvy, C. (1994) *Nucleic Acids Res.*, **22**, 3069-3074.
24. Smith, D. B., and Johnson, K. S. (1988) *Gene*, **67**, 31-40.
25. Turutin, D. V., Zatsepin, T. S., Timchenko, M. A., Kubareva, E. A., and Oretskaya, T. S. (2002) *Mol. Biol. (Moscow)*, **36**, 877-879.
26. Chen, Y. Q., Sengchanthalangsy, L. L., Hackett, A., and Ghosh, G. (2000) *Structure Fold Des.*, **8**, 419-428.
27. Zatsepin, T. S., Stetsenko, D. A., Arzumanov, A. A., Romanova, E. A., Gait, M. J., and Oretskaya, T. S. (2002) *Bioconj. Chem.*, **13**, 822-830.
28. Kozlov, I. A., Kubareva, E. A., Ivanovskaya, M. G., and Shabarova, Z. A. (1997) *Antisense Nucleic Acid Drug Dev.*, **7**, 279-289.
29. Kozlov, I. A., Kubareva, E. A., and Shabarova, Z. A. (1998) *Mol. Biol. (Moscow)*, **32**, 389-400.
30. Borisova, O. A., Romanenkov, A. S., Metelev, V. G., Kubareva, E. A., Zubin, E. M., Timchemko, M. A., and Oretskaya, T. S. (2003) *Mol. Biol. (Moscow)*, **37**, 534-543.
31. Kieran, M., Blank, V., Logeat, F., Vandkerckhove, J., Lottspeich, F., LeBail, O., Kourilsky, P., Baeuerle, P. A., and Israel, A. (1990) *Cell*, **62**, 1007-1018.