ISSN 0006-2979, Biochemistry (Moscow), 2011, Vol. 76, No. 1, pp. 94-117. © Pleiades Publishing, Ltd., 2011.
Original Russian Text © G. A. Nevinsky, 2011, published in Biokhimiya, 2011, Vol. 76, No. 1, pp. 115-141.

REVIEW

Main Factors Providing Specificity of Repair Enzymes

G. A. Nevinsky

Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, pr. Akademika Lavrentieva 8, 630090 Novosibirsk, Russia; fax: (3832) 363-5153; E-mail: nevinsky@niboch.nsc.ru

Received June 23, 2010
Revision received October 11, 2010

Abstract—Specific and nonspecific DNA complex formation with human uracil-DNA glycosylase, 8-oxoguanine-DNA glycosylase, and apurine/apyrimidine endonuclease, as well as with *E. coli* 8-oxoguanine-DNA glycosylase and RecA protein was analyzed using the method of stepwise increase in DNA-ligand complexity. It is shown that high affinity of these enzymes to any DNA (10^{-4}-10^{-8} M) is provided by a large number of weak additive contacts mainly with DNA internucleoside phosphate groups and in a less degree with bases of nucleotide links "covered" by protein globules. Enzyme interactions with specific DNA links are comparable in efficiency with weak unspecific contacts and provide only for one-two orders of affinity (10^{-1}-10^{-2} M), but these contacts are extremely important at stages of DNA and enzyme structural adaptation and catalysis proper. Only in the case of specific DNA individual for each enzyme alterations in DNA structure provide for efficient adjustment of reacting enzyme atoms and DNA orbitals with accuracy up to 10-15° and, as a result, for high reaction rate. Upon transition from nonspecific to specific DNA, reaction rate (k_{cat}) increases by 4-8 orders of magnitude. Thus, stages of DNA and enzyme structural adaptation as well as catalysis proper are the basis of specificity of repair enzymes.

DOI: 10.1134/S0006297911010111

Key words: mechanism of action, repair enzymes, uracil-DNA glycosylase, apurine-apyrimidine endonuclease, 8-oxoguanine-DNA glycosylase, RecA protein

Base modifications in DNA in response to endogenous and exogenous stress are an important source of mutations and play the key role in human mutagenesis, carcinogenesis, and senescence [1, 2].

The appearance of uracil in DNA results from the deamination of cytosine, either spontaneous or caused by chemical agents (nitrous acid, sodium bisulfite) as well as during substitution of dUMP for dTMP upon incorrect DNA synthesis [1, 3]. Cytosine modification (deamination) results in appearance of the U:G pair, which can result in GC→AT transition during the subsequent replication cycle; the formed a U:A pair may prevent specific protein binding [4, 5].

Abbreviations: AP, apurine/apyrimidine; APE, apurine/apyrimidine endonuclease; BER, base excision repair; Fpg, *E. coli* 8-oxoguanine-DNA glycosylase; hOGG1, human 8-oxoguanine-DNA glycosylase; I_{50}, the compound concentration at which 50% reaction inhibition is observed; IN, HIV integrase; NA, nucleic acids; ON, oligonucleotide; 8-oxoG, 8-oxoguanine; PN, protein–nucleic; SILC, stepwise increase in DNA-ligand complexity; ss and ds, single- and double-stranded, respectively; scDNA, supercoiled DNA; Topo, DNA topoisomerase I; UDG, uracil-DNA glycosylase.

Uracil-DNA glycosylase (UDG) has been found in all studied pro- and eukaryotic organisms as well as in pox and herpes virus families [6]. It is a highly conservative enzyme that hydrolyzes the *N*-glycoside bond between sugar-phosphate skeleton and uracil in DNA [1, 3, 6]. Unlike other glycosylases, UDG is able to remove uracil residues both from ss- and dsDNA with similar efficiency [6].

Type I herpes simplex virus UDG (complexed with free uracil and $d(pT)_3$ [7]) as well as of recombinant human UDG in complex with 6-aminouracil [8] and with specific enzyme inhibitor from pBS2 phage [9] have been studied by X-ray analysis. The results revealed some general features of these enzymes and identified conservative amino acid residues involved in uracil binding and in catalysis of the UDG-dependent reaction. Thermodynamic and kinetic analyses of the interaction of UDG purified from human placenta with ss- and dsON, as well as analysis of the specificity of this enzyme have been described [10-14]. Some reviews deal with pathways of appearance of uracil in DNA, main physical and biochemical enzyme properties, chemical mechanisms of damaged base removal, and X-ray analysis of the enzyme [1-3, 15-19].

It has been shown that oxidative stress is one of the main factors leading to DNA damage and mutation [2, 20, 21]. Oxygen-containing radicals (superoxide, hydroxyl, etc.) formed during oxidative stress induce various types of DNA damage (8-oxoguanine, 8-oxoadenine, formamidopyrimidines, thymine glycols, etc.). The most frequent is the formation of 8-oxoguanine (8-oxoG), which is removed from DNA by 8-oxoguanine DNA glycosylases. Human 8-oxoguanine DNA glycosylase (hOGG1) exhibits two catalytic activities: DNA glycosylase (removal of 8-oxoG) and AP lyase (cleavage of the phosphodiester bond at the formed apurine site by β-elimination) [22, 23]. The substrate specificity of this enzyme is described [23]. hOGG1 is not homologous to prokaryotic Fpg enzyme with analogous function and belongs to the endonuclease III superfamily [24, 25]. hOGG1 and its complexes with DNA have been studied by X-ray analysis [26-28]; complete dissimilarity of hOGG1 and Fpg tertiary structures was confirmed [29].

Unlike hOOG1, Fpg also catalyzes the δ-elimination reaction when further cleavage of 3′-terminal aldehyde takes place with formation of 4-oxopent-2-enal or 4-hydroxypent-2,4-dienal [30, 31]. Fpg substrate specificity was studied using ss- and dsDNA containing one modified unit [32-36]. X-Ray data for this enzyme have been reported [27].

AP sites can also emerge in DNA due to spontaneous apurinization, effects of ionizing radiation and active oxygen forms, etc. [1, 37, 38]. In this case modified (in particular, alkylated) bases often exhibit higher ability for spontaneous apurinization [39]. AP sites also emerge upon removal from DNA of various bases with single damages using base excision repair (BER) enzymes—DNA-*N*-glycosylases [20, 38]. AP endonucleases remove from DNA the AP sites that have emerged in different ways, as well as by chipping out 3′-terminal unsaturated aldehydes formed by bifunctional DNA glycosylases. Later the gap formed at the place of excised nucleotide is completed by various DNA polymerases, and then DNA ligase synthesizes the last phosphodiester bond, thus finishing the BER process. BER is described in more detail in reviews [2, 17, 20, 38].

Pro- and eukaryotic AP endonucleases are divided to two main families according to primary protein sequence [39, 40]. One of these families includes *E. coli* protein Xth and human protein APE1. *Escherichia coli* endonuclease IV (Nfo) is the prototype of the second family of AP endonucleases. No human Nfo-like AP endonucleases are yet found, though in other eukaryotes they may be the main enzymes carrying out this function, in particular, the main AP endonuclease of yeast (Apn1) belongs to the Nfo family. APE1 and APE2 of higher organisms are homologous to exonuclease III (ExoIII) of *E. coli* [41, 42]. APE1 is the most important eukaryotic endonuclease; it exhibits strong AP endonuclease and weak 3′-phosphodiesterase and 3′-5′ exonuclease activities [41, 42]. APE2 is less studied; this enzyme exhibits weak AP endonuclease but strong 3′-phosphodiesterase and 3′-5′ exonuclease activities and mainly exhibits these activities towards DNA substrates containing incorrect base pairs [42].

Despite structural differences, AP endonucleases of both families are rather similar in their specificity to single damages [21, 40]. Substrate specificity of a number of AP endonucleases has been intensely studied [40, 43-47]; X-ray analysis data are available for Xth, APE1, and Nfo [40, 48].

Homologous recombination is one of most important processes necessary for maintenance of genetic variability and damaged DNA repair in all living organisms. Protein RecA plays the main role in homologous recombination in *E. coli*, providing for search of homology between two DNA molecules and exchange of homologous strands [49]. RecA is an ATP-dependent DNA-binding protein (37.8 kDa) [49]. The binding of RecA to DNA is a three-step process: *presynapsis*, during which RecA is polymerized on ssDNA forming a right-helical nucleoprotein filament; *synapsis*, when dsDNA binds to presynaptic complex and homology with ssDNA is searched; and *strand exchange*, during which a new DNA duplex is formed and ssDNA is displaced. At the first stage RecA is polymerized on DNA, and this process is more efficient on ssDNA. RecA binding to ssDNA is unspecific, but formation of a protein filament on poly(dT) and on GT-rich sequences is preferable [50-53]. In the presence of ATP or its nonhydrolyzable thio-analog (ATPγS), RecA forms with DNA a right-helical filament with pitch of 95 Å and diameter of 100 Å [54]. The assembly of the filament is cooperative in the 5′→3′ direction of ssDNA [55, 56]. In its complex with protein, DNA is extended by approximately 50%, internucleotide distance increases to 5.1 Å, and three RecA monomers and nine nucleotide residues of DNA fall within one pitch of the filament helix [54]. In the case of binding to dsDNA filament parameters are the same as for ssDNA, while duplex in the complex is untwisted compared to DNA B form [57, 58]. In the absence of ATP a more compact inactive RecA filament with a pitch of 64 Å and internucleotide distance of 2.1 Å is formed [59].

After filament formation, RecA can bind dsDNA, which interacts with the second DNA-recognizing protein site. The second RecA site is also able to bind ssDNA, interacting with the filament even more efficiently compared to dsDNA. The role of second site of RecA protein after homologous strand exchange is leaving-strand binding after new duplex formation [60].

After dsDNA binding to the filament, the homology between corresponding strands is searched and subsequent strand exchange occurs. Several models have been proposed for search of homology and strand exchange [61-64], but they have not been experimentally confirmed (see below). Since RecA protein is a DNA-dependent ATPase and catalyzes the search for interstrand homolo-

gy and strand exchange, it should be considered as a multifunctional repair enzyme (especially in the case of SOS repair).

For understanding the fine mechanisms of any enzyme interactions with specific and nonspecific DNA, quantitative data concerning the efficiency of interactions of the enzyme globule with each of the covered nucleotide units of extended DNA is very important. Therefore, some data on repair enzymes have been investigated using stepwise increase in DNA-ligand complexity (SILC) and X-ray analysis.

METHOD OF STEPWISE INCREASE IN DNA-LIGAND COMPLEXITY

X-Ray analysis, providing data on protein–nucleic acid (PN) interactions at molecular level, is now one of most informative methods of enzyme analysis [1-23]. However, this approach does not allow quantitative estimation of efficiency of formation of the revealed contacts, which in turn excludes the possibility of estimating the relative contribution of individual specific and nonspecific contacts to total nucleic acid (NA) affinity to the enzymes [1-3]. The recognition of small ligands by an enzyme is usually due to formation of several strong specific contacts like hydrogen bonds, electrostatic contacts, or stacking interactions [65]. In this case all enzyme contacts with low molecular weight substrate are realized within a single molecule, which is a unique structural unit upon transition to polymers. Upon transition to extended molecules, the interval of K_d values characterizing their affinity to enzyme (compared to small ligands) is essentially unchanged (to 10^{-7}-10^{-10} M). The available X-ray data on NA–enzyme complexes (in the absence of quantitative data concerning relative contribution made by different structural elements to total affinity) led to rather incorrect concepts that, as in the case of low molecular weight ligands, just the specific contacts of the enzyme with a separate structural unit (or units), such as hydrogen bonds of ssDNA atoms and base groups forming Watson–Crick bonds in dsDNA with different amino acid residues of enzyme active sites, can provide for both high affinity to extended DNA and specificity of enzyme action [1-23]. In a polymeric molecule specific recognition by enzymes may not involve only contacts within a single structural unit, as in the case of low molecular weight ligand recognition. If it were so, the enzymes could not discriminate between polymeric molecules containing a specific unit and a low molecular weight ligand corresponding to this unit. Besides, X-ray analysis gives information on crystalline enzyme complexes with NA. However, the enzyme–NA complexes in a crystal and in solution may significantly differ because such complexes in a crystal are to a certain extent "frozen", while in solution they are more dynamic.

Enzyme-catalyzed reactions proceed in least in three stages [1-3, 27, 65]:

$$E + NA \overset{K_d(1)}{\longleftrightarrow} E{\cdot}NA \overset{K_d(2)}{\longleftrightarrow} E{\cdot}NA^* \overset{k_{cat}(3)}{\longrightarrow} E + \text{product}.$$

After primary E·NA complex formation ($K_d(1)$), both components undergo conformational changes ($K_d(2)$). The third stage is direct catalysis of the reaction characterized by rate constant ($k_{cat}(3)$). The interaction of DNA with replication, topoisomerization, integration, restriction, and repair enzymes in the first stage of primary complex formation was analyzed using the SILC technique, while other stages were analyzed using different methods. The capabilities and features of this method as well as some previously obtained data have been described earlier in reviews [66-68].

It is known that interactions of enzyme with ligands accompanied by formation of several bonds can be cooperative or non-cooperative [65]. In the case of formation of several bonds and absence of cooperativity the Gibbs' free energy value ($\Delta G^\circ = -RT\ln K_d$), characterizing efficiency of primary complex formation, is usually equal to the sum of ΔG° values related to separate types of interactions: $\Delta G^\circ = \Delta G^\circ_1 + \Delta G^\circ_2 + \ldots + \Delta G^\circ_n = -RT[\ln K_d(1)\cdot K_d(2)\cdot \ldots \cdot K_d(n)]$ [65-69]. It follows from additivity of ΔG° for individual ligand contacts with the enzyme that K_d value indicating total substrate affinity to a polymer is equal to the product of K_d values related to the separate contacts ($K_d = K_d(1)\cdot K_d(2)\cdot \ldots \cdot K_d(n)$).

It was shown that the mechanism of PN interactions can be analyzed at the molecular level using methods of synthesis and analysis — by stepwise increase or decrease in the complexity of the studied ligand (stepwise increase in DNA-ligand complexity, SILC) in accordance with the following scheme: *ortho*-phosphate or mononucleotide (as minimal enzyme ligand) → single-stranded unspecific homo-d(pN)$_n$ → single-stranded unspecific hetero-d(pN)$_n$ → single-stranded specific hetero-d(pN)$_n$ → double-stranded unspecific homo-d(pN)$_n$ → double-stranded specific hetero-d(pN)$_n$ → extended specific DNA.

This approach was used for analysis of sequence-independent enzymes (various DNA polymerases of prokaryotes, eukaryotes, viruses, and Archaebacteria [66-72], human DNA ligase and RNA gyrase [68]), as well as of sequence-specific enzymes (restriction endonuclease *Eco*RI [73], human topoisomerase I [74-78], and HIV integrase [79]). It was shown that any DNA ligand is a competitive inhibitor relative to specific DNA substrates, while minimal ligands are *ortho*-phosphate (P_i), deoxyribophosphate (dRMP), and dNMP, modeling separate structural elements of DNA [70, 74, 75, 79, 80]. To detect additivity in enzyme interactions with separate DNA units, the data are presented as logarithmic dependences of K_d values for d(pN)$_n$ on the number of mononucleotide units (n). Sharp changes in logK_d values could be expect-

ed at any *n* value in the case of cooperative interactions or synergism in binding of separate DNA units or fragments covered by the enzymes. However, such dependences were linear for all studied enzymes at $n \leq 7$-20, which was indicative of additivity of Gibbs' free energies characterizing their interactions with separate $d(pN)_n$ units [66-79]. For example, Fig. 1 shows data for the Klenow fragment of *E. coli* DNA polymerase I. The *n* values at which affinity changes correlated with relative size of the enzyme globules. Thus, in the case of *E. coli* DNA polymerase I and human DNA polymerase α, log dependences are linear for DNA template at $n = 0$-20, where 0 corresponds to *ortho*-phosphate [72, 81]. The existence of contacts with the 20-nucleotide-long template units was suggested by Kornberg on the basis of estimation of the diameter of this enzyme molecule (65 Å) and the size of a single complementary pair of DNA duplex [82]. For some enzymes the number of covered DNA units was estimated by the length of DNA fragments protected against hydrolysis by DNases [83, 84] and on the basis of X-ray analysis of the studied enzymes [7-9, 26-29, 40, 41]. Note that all studied repair enzymes have comparable molecular masses (~30-40 kDa), and for them linear increase in $\log K_d$ is observed for $n = 0$-10-11.

The *f* factor characterizing the increase in enzyme affinities to various $d(pN)_n$ upon their elongation by one unit (Tables 1 and 2) were estimated from slopes of linear parts of log dependences (Fig. 1). A 1.2-2.7-fold change in affinity of all enzymes to ssON upon their elongation by one unit (factor *f*) corresponds to $K_d = 1/f = 0.37$-0.83 M) and to change in Gibbs' free energy (from –0.12 to –0.45 kcal/mol), which can be indicative of efficiency of interaction of the enzyme with a single unit of ssDNA.

The increase in *f* factors in the case DNA polymerases [66-72], Topo [74-78], and RNA helicase [68] proceeds in the same order ($d(pC)_n < d(pT)_n < d(pG)_n < d(pA)_n$), as relative hydrophobicity of bases (C < T < G < A) (Tables 1 and 2). However, the affinity of all homo-$d(pN)_n$ to *Eco*RI is identical [73], and it is independent of relative hydrophobicity of their bases (Table 2). This means that *Eco*RI does not contact DNA bases but interacts with the sugar-phosphate skeleton of nucleic acids. A more complicated situation is observed for HIV integrase, the efficiency of its interaction with sugar-phosphate skeleton increasing along with increase in the flexibility of the DNA molecule [79].

The increase in enzyme affinity to ssDNA due to its interaction with a single internucleoside phosphate group is estimated from slopes of curves of $\log K_d$ dependence on *n* for $d(pR)_n$ of oligomers deprived of all bases (R is a chemically stable deoxyribose analog – tetrahydrofuran) (Fig. 1a), as well as from log*f* dependences for homo-$d(pN)_n$ on the relative hydrophobicity of the base estimated using isocratic reverse-phase chromatography of appropriate nucleosides (Fig. 1b) [85]. Extrapolation of these dependences to zero hydrophobicity (*ortho*-phosphate) made it possible to estimate the *e* factor (Fig. 1b) [12, 36, 45, 77, 78]. Since this factor is mainly indicative of increased enzyme affinities to sugar-phosphate skeleton due to their electrostatic interactions with a single internucleoside phosphate group (although there can also be weak hydrogen bonds), it was called electrostatic factor (*e* factor) [66-68]. The *e* values are given in Tables 1 and 2. With account for additive contribution of the sugar-phosphate skeleton structural elements and ON bases to their affinity to enzymes, coefficient of their affinity increase due to hydrophobic and/or van der Waals interactions with one of the bases (hydrophobic *h* factor) was estimated from the ratio $h = f/e$ ([66-68] and references therein). The *h* values are given in Tables 1 and 2. It is shown that interaction of all enzymes with ssON is a superposition of weak unspecific electrostatic and hydrogen bonds as well as of hydrophobic and/or van der Waals interactions with separate structural elements and is described by a single decreasing geometric progression [66-68]:

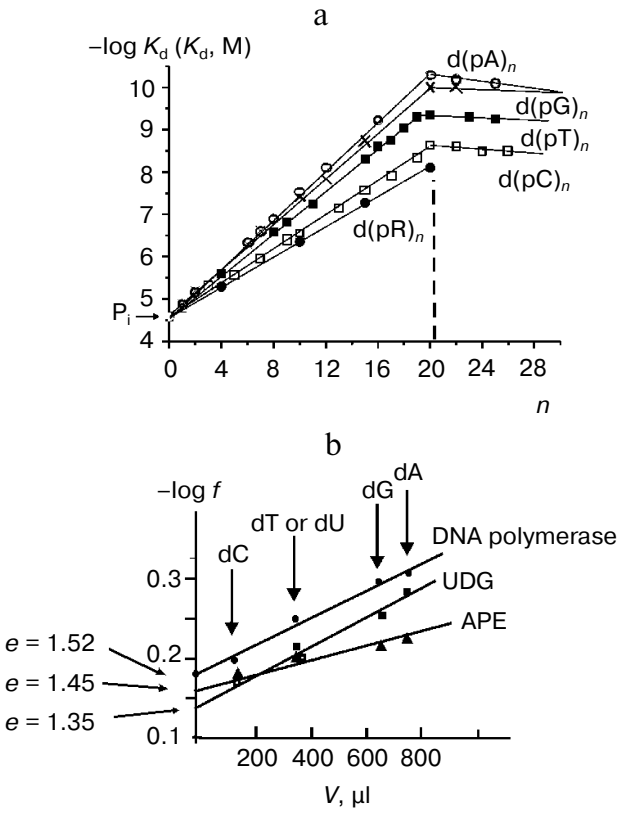

Fig. 1. a) Logarithmic dependences of K_d of complexes of *E. coli* DNA polymerase I Klenow fragment template region with different $d(pN)_n$ oligonucleotides on ligand length (*n*). b) Estimation of electrostatic factor *e* indicating the increase in efficiency of different enzyme interactions with extended DNA due to one internucleoside phosphate group using the factor *f* logarithm dependence on relative base hydrophobicity. Relative hydrophobicity of different bases was determined using isocratic reverse phase chromatography of the corresponding nucleosides.

Table 1. Kinetic and thermodynamic characteristics of interaction between sequence-independent enzymes and RecA protein with DNA*

Characteristics of ligands and enzymes	DNA ligase [68]	RNA helicase [68]	Klenow fragment, template [72]	RecA protein [94-96]
P_i; K_d, M (minimal ligand)	$1.0 \cdot 10^{-3}$	$1 \cdot 10^{-1}$	$2.6 \cdot 10^{-5}$	$5 \cdot 10^{-1}$**
Various dNMP; K_d, M	$(5\text{-}8) \cdot 10^{-4}$	$\sim 6.3 \cdot 10^{-2}$	$1.0 \cdot 10^{-5}$	$I_{50} = (1.2\text{-}2.0) \cdot 10^{-2}$ (first DNA-binding site)* $I_{50} = (3\text{-}6) \cdot 10^{-2}$ (second DNA-binding site)
n (number of binding subsites)	8-10	10	19-20	>1000
ss-d(pN)$_n$; $K_d(1)$, M	$1.0 \cdot 10^{-8}$	$5.1 \cdot 10^{-6}$	$5.0 \cdot 10^{-10}$	$(2\text{-}5) \cdot 10^{-4}$ ($n = 10$)*** $(1\text{-}240) \cdot 10^{-7}$ ($n = 20$) (first DNA-binding site)
Electrostatic e factor	2.14	1.61	1.52	1.56
f_C (h_C, hydrophobic factor); $f = h_N\ e$	1.1	1.05	1.58 (1.04)	2.04 (1.32; n = 1-20)
f_T (h_T, hydrophobic factor)	1.29	1.26	1.78 (1.17)	2.04 (1.32; n = 1-20)
f_G (h_G, hydrophobic factor)	1.51	1.48	1.95 (1.28)	not determined
f_A (h_A, hydrophobic factor)	1.62	1.59	2.0 (1.32)	2.12 (1.0; n = 1-10) 1.32 (1.0; n = 10-20)
ds-d(pN)$_n$; $K_d(2)$, M	$1.0 \cdot 10^{-9}$	$5.0 \cdot 10^{-6}$	$\sim 1.0 \cdot 10^{-11}$	$(2\text{-}35) \cdot 10^{-4}$ ($n = 10$) $(3.2\text{-}23) \cdot 10^{-6}$ ($n = 20$) (second DNA-binding site)
$K_d(1)/K_d(2)$	~10	~1.0		two separate sites

* RecA protein forms a filament on any length single-stranded d(pN)$_n$ beginning from n = 10-20 (first DNA-binding site) and then the second DNA-binding site, formed during filamentation, is able to interact with ss- and ds-DNA.

** Within different single-stranded d(pN)$_n$ the first terminal internucleoside phosphate group exhibits different affinity to the first RecA monomer of a formed filament depending on the type of the first terminal base.

*** Logarithmic dependence for d(pA)$_n$ has point of inflection at n = 6-8 (Fig. 10a).

$$K_d[d(pN)_n] = K_d[(P_i)]\ (e)^{1-n}\ (h_C)^{-c}\ (h_T)^{-t}\ (h_G)^{-g}\ (h_A)^{-a}, \quad (1)$$

where $K_d[(P_i)]$ is K_d for minimal ligand *ortho*-phosphate; h_C, h_T, h_G, and h_A correspond to hydrophobic factors h_N indicative of increased efficiency of enzyme interaction due to incorporation into d(pN)$_n$ of one of the bases (C, T, G, or A), the number of which in ON is equal to c, t, g, and a, respectively. Factor e is indicative of increased enzyme affinity due to enzyme interaction with a single internucleoside phosphate group [66-68]. This equation describes the interaction of any ss- or ds-DNA with any sequence-independent enzyme as well as interaction of nonspecific DNA with any studied sequence-dependent enzymes. Upon transition from one enzyme to another there is a pronounced change in numerical $K_d[(P_i)]$ values and slight changes in e and h_N factors shown in Tables 1 and 2.

It is known that specificity of action of many sequence-dependent enzymes, including those studied by us, is up to eight and even nine orders of magnitude (one error per 10^5-10^9 catalytic acts) [1, 38, 66-68]. It has been shown that at the stage of complex formation the interactions with specific nucleotide units are not able to provide either for total high affinity to DNA (10^{-7}-10^{-10} M) or for enzyme action specificity observed in cells *in vitro* and *in vivo*. Usually unspecific additive interactions provide 4-8 orders of magnitude, while specific interactions provide only 1-2 orders of magnitude of total DNA affinity to enzymes [66-68]. As for other enzymes [65], despite detected differences in enzyme interactions with low molecular weight and polymeric ligands, the basis of specificity of DNA-dependent enzyme actions are stages of enzyme-dependent conformational DNA adaptation and catalysis proper: the k_{cat} values increase by 5-8 orders of magnitude upon transition from nonspecific to specific DNA (for reviews see [66-68]).

Table 2. Kinetic and thermodynamic characteristics of sequence-dependent enzyme interactions with nonspecific (ns) and specific (s) DNA*

Ligand and enzyme characteristics	Enzyme		
	topoisomerase I [74-78]	HIV-integrase [79]	*Eco*RI endonuclease [73]
Minimal ligands			
P_i; K_d, M	$3.8 \cdot 10^{-1}$	$3.3 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$
dN(C,T,G,A)MP; K_d, M	$(1.7\text{-}2.2) \cdot 10^{-2}$	$(0.77\text{-}15) \cdot 10^{-3}$	$(1.2\text{-}2.3) \cdot 10^{-3}$
Interaction with nonspecific DNA			
n (number of interaction subsites)	20 (10 with strong contacts)**	⩾24***	7-8
ss-d(pN)$_n$; $K_d(1)$, M	$8.5 \cdot 10^{-5}$	$(0.012\text{-}1.0) \cdot 10^{-6}$	$(5\text{-}6) \cdot 10^{-6}$
ds-d(pN)$_n$; $K_d(2)$, M	$1.0 \cdot 10^{-5}$	$4.0 \cdot 10^{-8}$	$(5\text{-}6) \cdot 10^{-6}$
$K_d(1)/K_d(2)$	8.5	23-83	1.0
e	1.67	e depends on n	2.0
f_C (h_C, hydrophobic factor)	1.7 (1.02)	~1.64 (n = 5-24)****	2.0 (1.0)
f_T (h_T, hydrophobic factor)	1.94 (1.16)	~1.26 (n = 5-24)	2.0 (1.0)
f_G (h_G, hydrophobic factor)	2.1 (1.27)	–	2.0 (1.0)
f_A (h_A, hydrophobic factor)	2.3 (1.40)	~1.35 (n = 5-24)	2.0 (1.0)
Interaction with specific DNA			
ss-d(pN)$_n$; $K_d(3)$, M	$2.0 \cdot 10^{-6}$	$1 \cdot 10^{-8}$	–
ds-d(pN)$_n$; $K_d(4)$, M	$7.0 \cdot 10^{-8}$	$1.3 \cdot 10^{-9}$	$1.0 \cdot 10^{-7}$
$K_d(3)/K_d(4)$	29	7.7	–
Comparison of interaction with specific and nonspecific DNA			
$K_d(1)/K_d(3)$	43	~1-100	–
$K_d(2)/K_d(4)$	143 (240)	31	50-60
dsDNA; [k_{cat} (s)/k_{cat} (ns)]	$10^5\text{-}10^6$	$10^6\text{-}10^7$	$10^5\text{-}10^6$

Note: Dash means value is not determined.

* In the case of topoisomerase I, *Eco*RI endonuclease, and HIV-integrase, single-stranded oligonucleotides do not serve as substrates [73-79].

** Topoisomerase covers 20 base pairs of DNA, but only 10 of them efficiently interact with the enzyme.

*** Oligonucleotides up to 24 units in length were used.

**** Hydrophobic factors f for HIV-integrase like e factor depend on oligonucleotide length; f factor is shown at n = 5-24.

INTERACTION OF NUCLEIC ACIDS WITH EXCISION REPAIR ENZYMES

Different length nonspecific mono- and oligonucleotides are competitive inhibitors relative to specific DNA substrates in reactions catalyzed by studied repair enzymes [10, 35, 45, 84] (Table 3). Minimal ligands for excision repair enzymes (UDG, Fpg, hOGG1, and APE), like for other above-described enzymes, were *ortho*-phosphate, specific and nonspecific NMP and dNMP (Tables 3 and 4). Affinity of specific ligands (dUMP, oxo-dGMP, and 5′-deoxyribose phosphate) to these enzymes was 1.8-45 times higher than for nonspecific dNMP [12, 36, 45, 86]. However, these enzymes recognized specific and nonspecific dNMP due to interaction with the base, sugar residue, and phosphate group.

Table 3. Affinity of APE to minimal ligands, their derivatives, as well as to ss- and dsDNA [45]

Ligand	I_{50}, μM	K_I, μM*	Ligand	I_{50}, μM	K_I, μM*	Ligand	I_{50}, μM	K_I, μM*
NaH_2PO_4	1080	360	D-ribose	>0.5 M	>0.17 M	d(ribose phosphate)	75	25
dAMP	495	165	dTMP	490	**163.3****	dCMP	490	163.3
Internucleoside phosphate	–	100 or 264***	d(ribose phosphate)	177	**59.0**	dGMP	500	**166.6**
Single-stranded oligonucleotides								
$d(pA)_2$	150	50	$d(pT)_2$	350	116.6	$d(pC)_2$	420	140
	–	–	$d(pT)_3$	200	**66.7**	$d(pC)_3$	180	60
$d(pA)_4$	100	33.3	$d(pT)_4$	135	45	–	–	–
	–	–	–	–	–	$d(pC)_5$	65	21.7
$d(pA)_6$	51.6	17.2	$d(pT)_6$	74	24.6	–	–	–
	–	–	–	–	–	$d(pC)_7$	30	10
$d(pA)_8$	7.5	2.5	$d(pT)_8$	25	8.3	–	–	–
	–	–	–	–	–	$d(pC)_9$	14	4.7
$d(pA)_{10}$	5.0	**1.66**	$d(pT)_{10}$	7.5	**2.5**	$d(pC)_{10}$	10	**3.33**
	–	–	$d(pT)_{11}$	7.5	**2.5**	$d(pC)_{11}$	10	3.33
$d(pA)_{12}$	5.4	1.7	$d(pT)_{12}$	7.5	2.5	–	–	–
	–	–	–	–	–	$d(pC)_{13}$	10	3.33
$d(pA)_{14}$	5.3	1.7	$d(pT)_{14}$	7.5	2.5	–	–	–
	–	–	$d(pT)_{15}$	7.7	2.57	–	–	–
$d(pA)_{16}$	5.1	1.66	–	–	–	–	–	–
$d[(pR)_3pT]$	34	**11.7**	$d(pG)_2$	306	**102**	–	–	–
$d[(pR)_5pT]$	15.6	**5.2**	$d(pG)_4$	115	38.3	–	–	–
$d[(pR)_7pT]$	6.9	**2.3**	$d(pG)_6$	43	14.4	–	–	–
$d[(pR)_9pT]$	3.0	1.0	$d(pG)_8$	16.3	**5.4**	–	–	–
Double-stranded oligonucleotides								
$d(pA)_2 \cdot d(pT)_2$	110	36.6	$d(pA)_8 \cdot d(pT)_8$	28.0	9.3	$d(pA)_{14} \cdot d(pT)_{14}$	1.0	0.33
$d(pA)_4 \cdot d(pT)_4$	80	26.6	$d(pA)_{10} \cdot d(pT)_{10}$	1.0	**0.33**	$d(pA)_{16} \cdot d(pT)_{16}$	1.0	0.33
$d(pA)_6 \cdot d(pT)_6$	35.5	**11.8**	$d(pA)_{12} \cdot d(pT)_{12}$	1.1	0.36	$d(pA)_{20} \cdot d(pT)_{20}$	1.0	**0.33**

Note: Dash means not determined.

* Error in determination of K_I and I_{50} values 10-30%; mean values from 3-4 experiments are shown.

** For some ligands competitive inhibition towards standard APE substrate was shown and K_I values were determined (they are shown in bold in the table); in other cases I_{50} values were found and K_I values were then calculated using the equation $I_{50} = 3K_I$ at specific substrate concentration $2K_m$.

*** K_I value for the internucleoside phosphate group was calculated by extrapolation of linear log dependences for $d(pN)_n$ (300 mM) and for $d[(pR)_npT]$ (100 mM) to $n = 0$ (R is the tetrahydrofuran analog of deoxyribose).

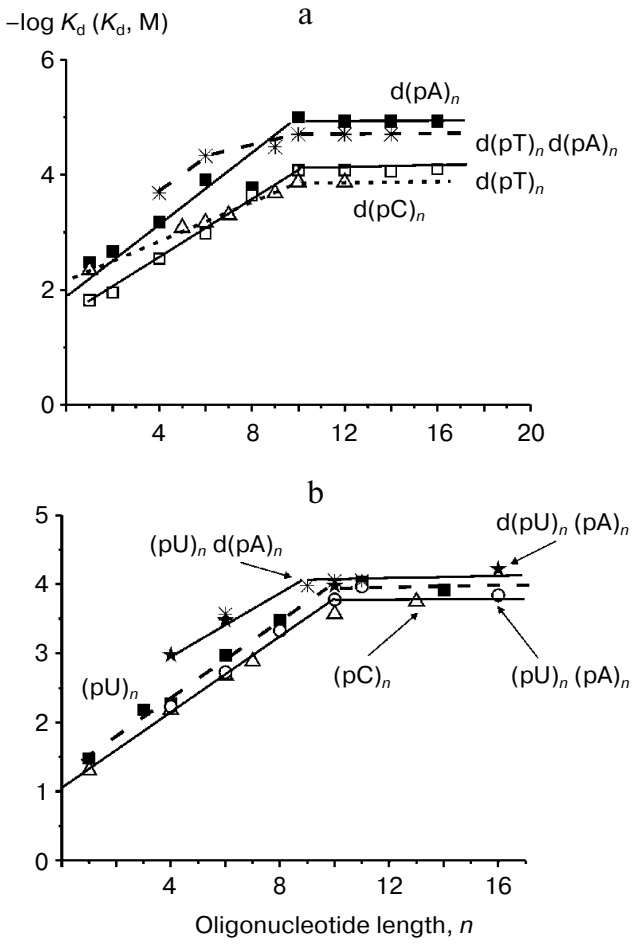

Fig. 2. Logarithmic dependences of K_d values of UDG complexes with different d(pN)$_n$ (a) and ribo(pN)$_n$ (b) oligonucleotides and their duplexes on ligand length (n).

In this case the main contribution to affinity was interactions with the phosphate group, but the contribution of all structural elements of dNMP to the affinity to the enzyme was nearly additive. Thus, Fpg affinity to different dNMP structural elements increased in the order: deoxyribose (K_d = 157 mM) < various bases (18.3-48.3 mM) < *ortho*-phosphate (10 mM) < 5′-deoxyribose phosphate (8.3 mM) = dTMP and dCMP (8.3 mM) < dGMP and dAMP (3.3 mM) < oxo-dGMP (K_I = 0.7 mM) [36]. Data for APE are given in Table 3.

For excision repair enzymes, as for other above-described enzymes, log-dependences of K_d values for ss- and ds-d(pN)$_n$ on n were plotted (Figs. 2-4) [12, 36, 45, 86]. They were linear at $n \leq 10$, which correlated with their relatively small size (30-40 kDa) and was indicative of additivity of free energies characterizing interaction of 10 subsites within the DNA binding site in BER enzymes intended for separate d(pN)$_n$ unit recognition [12, 36, 46, 86]. On the whole, the interaction of repair enzymes with d(pN)$_n$, in accordance with their elongation correlates with the data shown in Fig. 5. The enzyme active site (zero subsite) exhibits increased affinity to free mononucleotide or to one of the units of extended d(pN)$_n$ interacting with 10 enzyme subsites. Dinucleotide forms di- and trinucleotides, three alternative enzyme complexes, etc., until decanucleotide occupies all subsites of the complete DNA binding site. When there is any length-specific unit in the DNA, just this unit interacts with the zero subsite (active site).

It should be emphasized that the repair enzyme affinity to ss- and dsDNA only 2-30-fold exceeds that to corresponding canonical ss- and dsRNA (Figs. 2-4). In the stage of complex formation, these enzymes also poorly discriminate between DNA–DNA, DNA–RNA, and RNA–RNA duplexes (Figs. 2-4). Values of factors f, e, and h_N characterizing the interaction of UDG, Fpg,

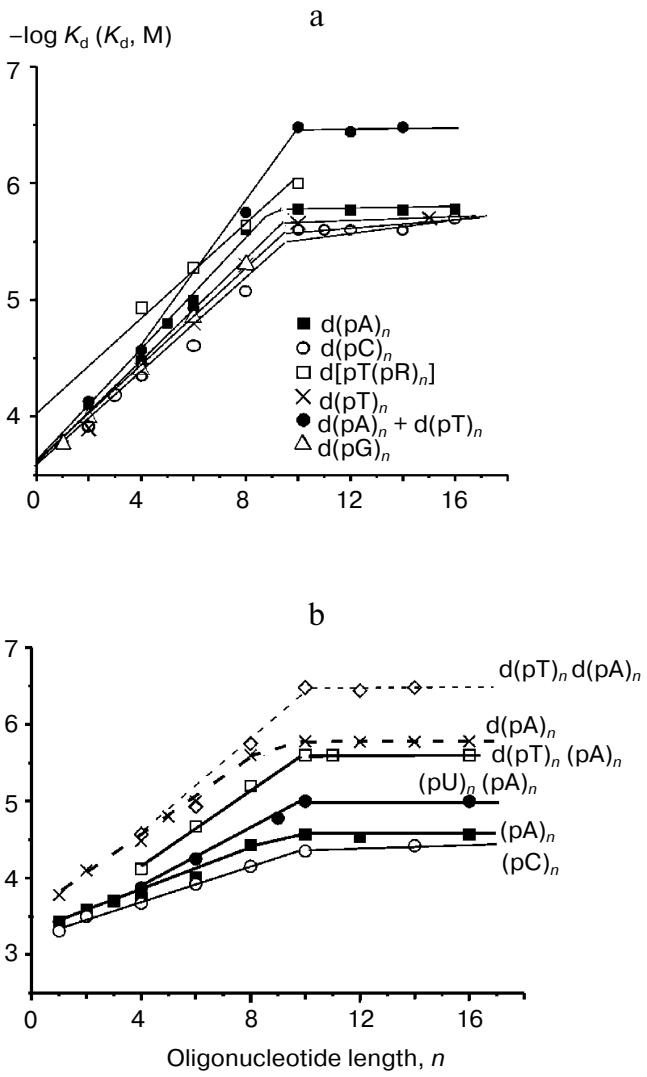

Fig. 3. Logarithmic dependences of K_d values of APE complexes with different d(pN)$_n$ (a) and ribo(pN)$_n$ (b) oligonucleotides and their duplexes on ligand length (n).

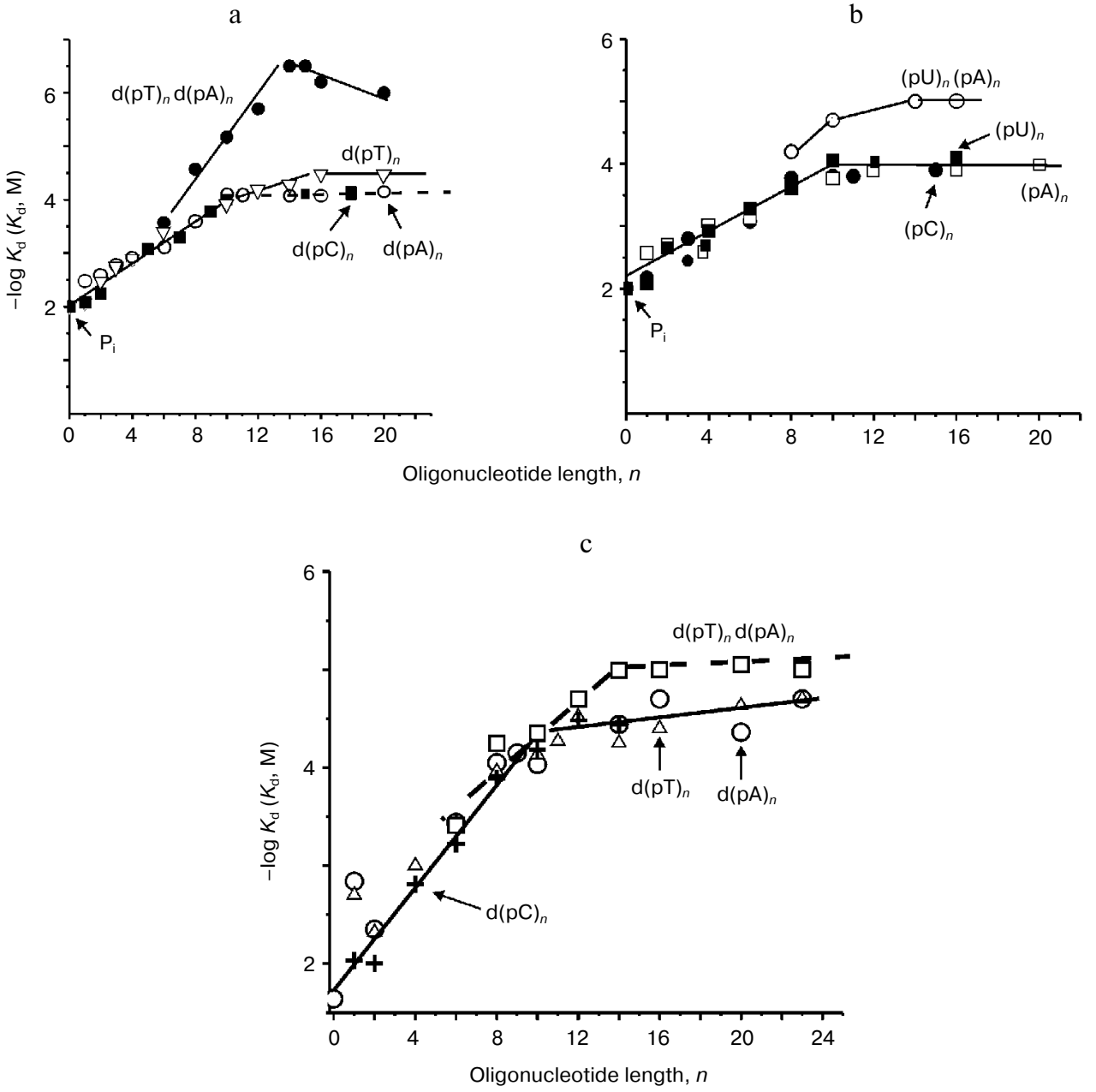

Fig. 4. Logarithmic dependences of K_d values of Fpg (a, b) and hOGG1 (c) complexes with different single-stranded d(pN)$_n$ and ribo(pN)$_n$ oligonucleotides and their duplexes on ligand length (n).

hOGG1, and APE with DNA were calculated as described in the previous section. It is interesting that these factors were similar for DNA and RNA, while the observed difference in affinity of the latter to repair enzymes is defined by the fact that active sites of these enzymes interact with equally lower efficiency with NMP as with dNMP, as well as with one unit (Fig. 5) of nonspecific extended RNA compared to one unit of DNA (Figs. 2-4) [12, 36, 44, 45]. It was shown that the interaction of UDG, Fpg, hOGG1, and APE with nonspecific ss- and dsDNA and RNA having canonical structure is described using the same decreasing geometric progression (Eq. (1)) as the interaction of different enzymes, dependent or independent of specific DNA sequences.

Both ds- and ssDNA are substrates for UDG, Fpg, and APE1, but only dsDNA is a substrate for hOGG1 [11, 12, 32-36, 43-45, 86]. It is interesting that UDG even better removes modified units from ss- than from dsDNA, while Fpg and APE demonstrate higher affinity and conversion rate in the case of dsDNA. The change by 1.46-1.92-fold (f factor; Table 4) of the affinity of UDG,

Table 4. Kinetic and thermodynamic characteristics of repair enzyme interactions with specific (s) and nonspecific (ns) DNA

Ligand and enzyme characteristics	Repair enzymes			
	UDG [12]	Fpg [36]	APE [45]	hOGG1 [86]
Minimal ligands				
P_i; K_d, M	$1.7 \cdot 10^{-2}$	$1.0 \cdot 10^{-2}$	$3.6 \cdot 10^{-4}$	$6.8 \cdot 10^{-2}$
dN(C,T,G,A)MP; K_d (ns), M	$(3.3\text{-}15) \cdot 10^{-3}$	$(3.3\text{-}8.3) \cdot 10^{-3}$	$(1.6\text{-}1.7) \cdot 10^{-4}$	$(1.5\text{-}9.4) \cdot 10^{-3}$
Specific dN(R)MP; K_d (s), M	$1.8 \cdot 10^{-3}$	$7.0 \cdot 10^{-4}$	$2.5 \cdot 10^{-5}$	$1.6 \cdot 10^{-3}$
K_d (ns)/K_d (s)	1.8-8.3	4.7-12.0	6.4-6.8	7.2-45.0
Nonspecific DNA				
n (number of subsites)	9-10	10-13*	9-10	10-13*
ss-d(pN)$_n$; K_d(1), M	$1 \cdot 10^{-5}$**	$(3.3\text{-}10) \cdot 10^{-5}$**	$2.1 \cdot 10^{-6}$	$(2.0\text{-}3.2) \cdot 10^{-5}$
ds-d(pN)$_n$; K_d(2), M	$(1.5\text{-}1.8) \cdot 10^{-5}$	$(0.75\text{-}1.3) \cdot 10^{-6}$	$0.5 \cdot 10^{-6}$	$8.7 \cdot 10^{-6}$
$K_d(1)/K_d(2)$	0.6-0.7	90-233	~4-12	1.5-3.3
e	1.35	1.56	1.51	1.78
f_C (h_C, hydrophobic factor)	1.46 (1.08)	1.56 (1.0)	1.53 (1.01)	1.78 (1.0)
f_T (h_T, hydrophobic factor)	1.65 (1.22)	1.56 (1.0)	1.59 (1.05)	1.78 (1.0)
f_G (h_G, hydrophobic factor)	1.82 (1.35)	1.56 (1.0)	1.62 (1.07)	1.78 (1.0)
f_A (h_A, hydrophobic factor)	1.92 (1.42)	1.56 (1.0)	1.66 (1.10)	1.78 (1.0)
Specific DNA				
ss-d(pN)$_n$; K_d(3), M	$3.3 \cdot 10^{-5}$	$(5\text{-}21) \cdot 10^{-7}$	$(3.5\text{-}3.9) \cdot 10^{-7}$	$8.0 \cdot 10^{-8}$
ds-d(pN)$_n$; K_d(4), M	$1.3 \cdot 10^{-5}$	$(6.7\text{-}10) \cdot 10^{-9}$	$0.13 \cdot 10^{-6}$	$1.1 \cdot 10^{-8}$
$K_d(3)/K_d(4)$	2.5	10-300	~3	~7.3
Difference in affinity to specific and nonspecific DNA				
$K_d(1)/K_d(3)$	2-3	4-30	5-6	250
$K_d(2)/K_d(4)$	10-30	7-70	3.8-11.5	791
dsDNA; [k_{cat} (s)/k_{cat} (ns)]	$(5\text{-}50) \cdot 10^4$	10^6-10^8	10^6-10^7	10^6-10^8

* Fpg and hOGG1 interact with 10 units of ssDNA and 13 base pairs of dsDNA.
** Range of values is given when affinity very strongly depends on DNA sequences; one K_d value is given for DNA exhibiting maximal affinity.

Fpg, hOGG1, and APE to ssDNA upon their elongation by one unit corresponds to $K_d = 1/f = 0.52\text{-}0.68$ M and $\Delta G^o = -0.23 \ldots -0.39$ kcal/mol, indicative of the interaction efficiency of the enzyme with one unit of ssDNA. The ΔG^o values for such interactions are significantly lower than those for strong electrostatic contacts (−1 ... −2 kcal/mol) or hydrogen bonds (−1 ... −5 kcal/mol) between enzymes and ligands and they are comparable with corresponding values for weaker hydrophobic, ion–dipole, and dipole–dipole interactions [65-68]. Since internucleoside phosphate groups make the main contribution to the affinity of UDG, Fpg, hOGG1, and APE, it should be assumed that interaction of these enzymes with the DNA sugar-phosphate skeleton is not due to contacts between oppositely charged groups, but it is rather due to weak interactions between oppositely charged surfaces of the biopolymers [12, 36, 45, 86].

Addition of the second strand differently influences the affinity of the duplex compared to ssDNA (Figs. 2-4). Like in the case of DNA polymerases [66-72], topoiso-

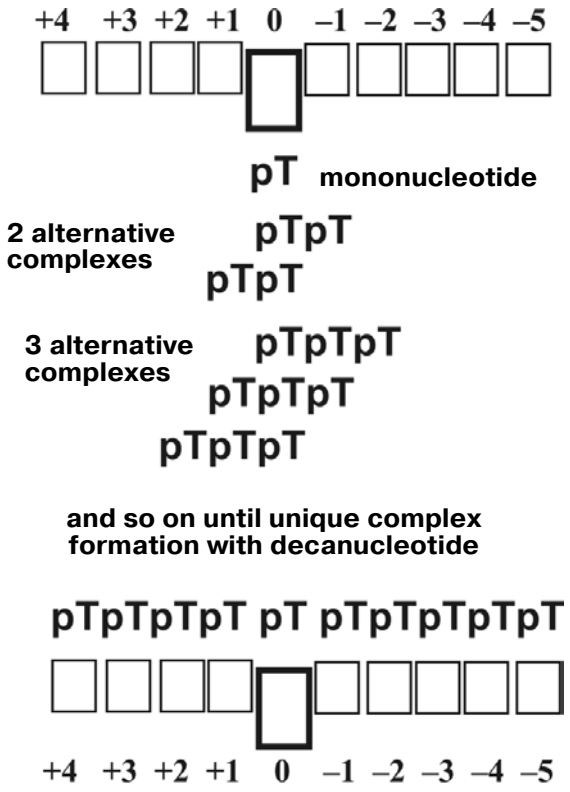

Fig. 5. Scheme of enzyme interaction with different length nonspecific oligonucleotides. The enzyme active site (subsite 0) exhibits increased affinity to dNMP or to one of its units within any length $d(pN)_n$. As ON length increases, the number of alternative enzyme complexes with $d(pN)_n$ increases until monomeric units of oligonucleotide occupy all enzyme subsites intended for their recognition. In the case of specific DNA, a specific modified DNA unit exhibiting increased affinity to the enzyme always binds its active site.

merases I [74-78], and HIV integrase [79] (Tables 1 and 2), the second chain noticeably increases the affinity of the duplex compared to that of ssDNA in the case of APE, Fpg, and hOGG1 [36, 45, 86], but it does not significantly contribute to the affinity of the duplex to *Eco*RI [73] and UDG [12] (Tables 1-4). Whereas the first chain provides 5-8 orders of magnitude of BER enzyme total affinity to dsDNA, the contribution of the second chain usually does not exceed two orders of magnitude (Table 4). This correlates with X-ray data showing that these enzymes establish the most important and strongest contacts mainly with one strand of duplex [7-9, 26, 29, 87-93]. Besides, interaction between DNA duplex strands within its complex with enzymes is weakened due to almost complete or partial dsDNA melting caused by intercalation of hydrophobic amino acid residues in dsDNA and formation of contacts with each DNA strand.

Threshold K_d values of repair enzyme complexes with different specific ss- and ds-$d(pN)_n$ have been estimated. The increase in affinity upon transition from nonspecific to specific ss- and dsON of the same length was somewhat dependent on the sequence adjacent to the specific unit and was 3.8-11.5 (APE), 4-70 (Fpg), 2-30 (UDG), and 250-791 (hOGG1) (Table 4) [12, 36, 45, 86]. It should be noted that for APE, Fpg, and UDG the increase in the complex formation efficiency upon transition from nonspecific to specific ligands was very close at the level of minimal ligands (dNMP) and ss- and dsDNA. In other words, the difference in enzyme affinity for pairs of specific (s) and nonspecific (ns) ligands (s-NMP–ns-NMP, ss-s-$d(pN)_n$–ss-ns-$d(pN)_n$, and ds-s-$d(pN)_n$–ds-ns-$d(pN)_n$) was practically identical. This pointed to approximately additive contribution of each nucleotide unit specific for these enzymes and interacting with the active site and of nine nonspecific nucleotides of ssDNA (or nine base pairs of dsDNA (scheme of Fig. 5) to total affinity of specific dsDNA to these enzymes (Table 4). Only in the case of hOGG1 there was no correlation between efficiency of enzyme interaction with specific 8-oxoG unit at the level of dNMP and ss- and dsDNA [86]. Whereas difference in affinity of specific and nonspecific dNMP for hOGG1 was 7.5-42, it increased to 250 upon transition to ssDNA and to 791 times upon transition to dsDNA (Table 4). This showed that in the case of hOGG1 cooperative interactions between the enzyme active site recognizing the 8-oxoG unit and enzyme regions interacting with nonspecific units of cleavable and uncleavable strands are observed [86].

Thus, it is obvious that at the stage of complex formation the efficiency of interaction of repair enzyme with DNA specific structural elements cannot provide for specificity of their action observed *in vivo*. In fact, the efficiency of formation of specific contacts between these enzymes and DNA does not exceed two orders of magnitude in affinity, while the relative contribution of nonspecific interactions to total enzyme affinity to DNA is 5-7 orders of magnitude (Table 4). It should be stressed that exactly the same situation with affinity increase by 1-2 orders of magnitude for specific DNA is also observed for different sequence-specific enzymes of topoisomerization, restriction, and integration [66-68].

Based on data of thermodynamic and X-ray analysis from the literature concerning complexes of these repair enzymes with DNA, we described their interaction with specific and nonspecific DNA using models of interaction (Figs. 6-9). These models produce ideas concerning relative contribution to the repair enzyme affinities (ΔG^o) to different specific and nonspecific units of extended DNA at fixed temperature.

It was shown [46, 50, 57, 61, 66-68, 70, 86] that UDG, Fpg, APE, and hOGG1 recognize uncharacteristic nonspecific RNA–RNA and RNA–DNA duplexes with affinity only 6-20 times lower than that for DNA–DNA duplexes. However, the enzymes do not catalyze RNA–RNA and RNA–DNA transformation even at sat-

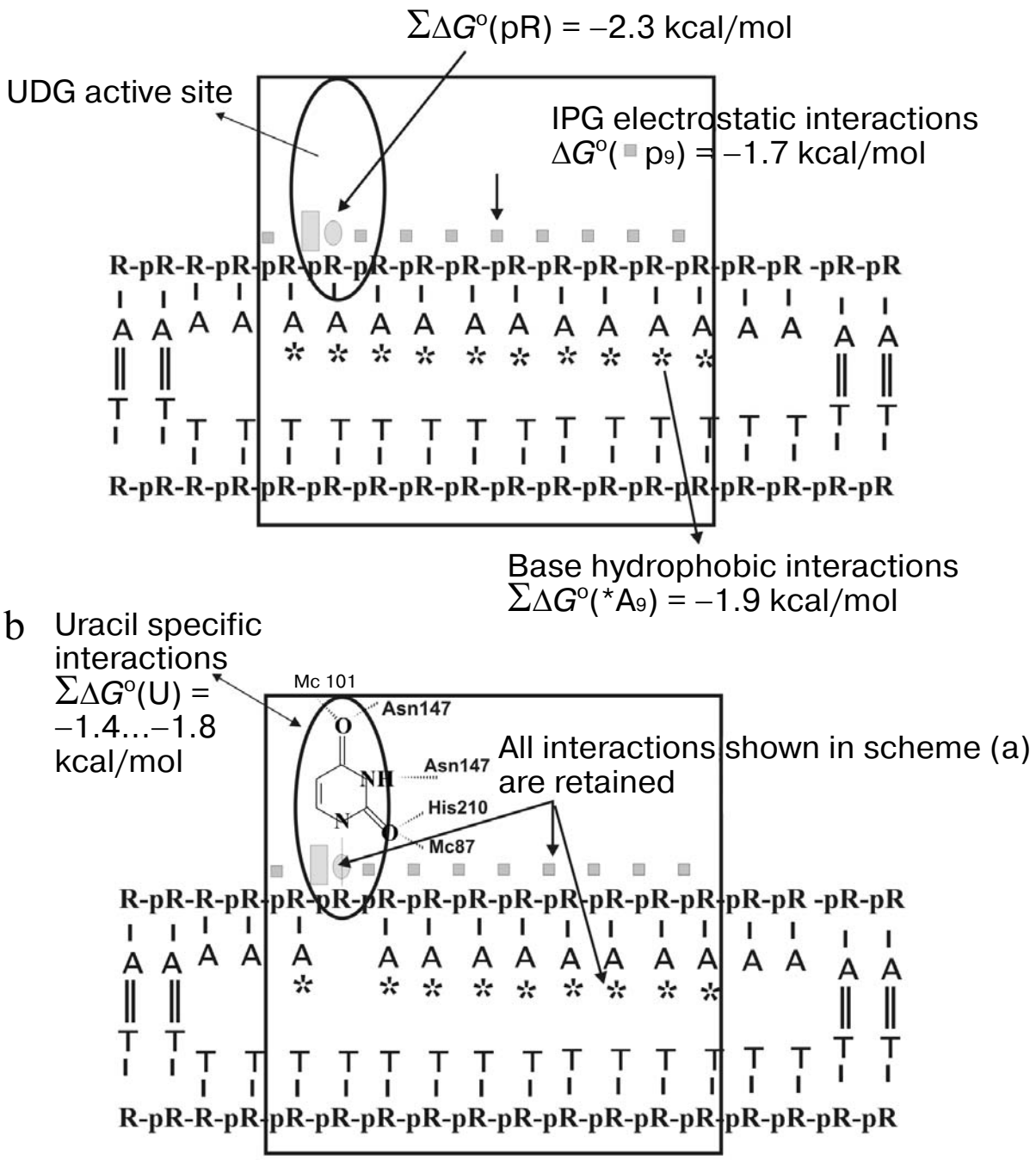

Fig. 6. Models of the interaction of UDG with nonspecific (a) and specific DNA (b) containing a dU unit in one strand: gray squares correspond to weak electrostatic contacts of internucleoside phosphate groups (IPG); asterisks indicate weak hydrophobic or van der Waals contacts of the enzyme with bases; gray rectangles and ellipse show nonspecific contacts of d(pR)-sugar-phosphate residue of nonspecific dA- or specific dU unit with the enzyme active site. b) Specific hydrogen bonds of a uracil residue with amino acid residues of the enzyme revealed by X-ray analysis are shown [7-9]; Mc101 and Mc87 correspond to atoms of N–H peptide bonds of Phe101 and Gln87 belonging to the main chain; Asn147 forms a hydrogen bond with an oxygen atom and an NH group of a uracil residue via NH_2- and C=O groups of the side chain, respectively; His210 interacts with the uracil oxygen atom via the NH group of the imidazole ring. The range of ΔG^o changes for the dU unit depending on the substrate base sequence is shown.

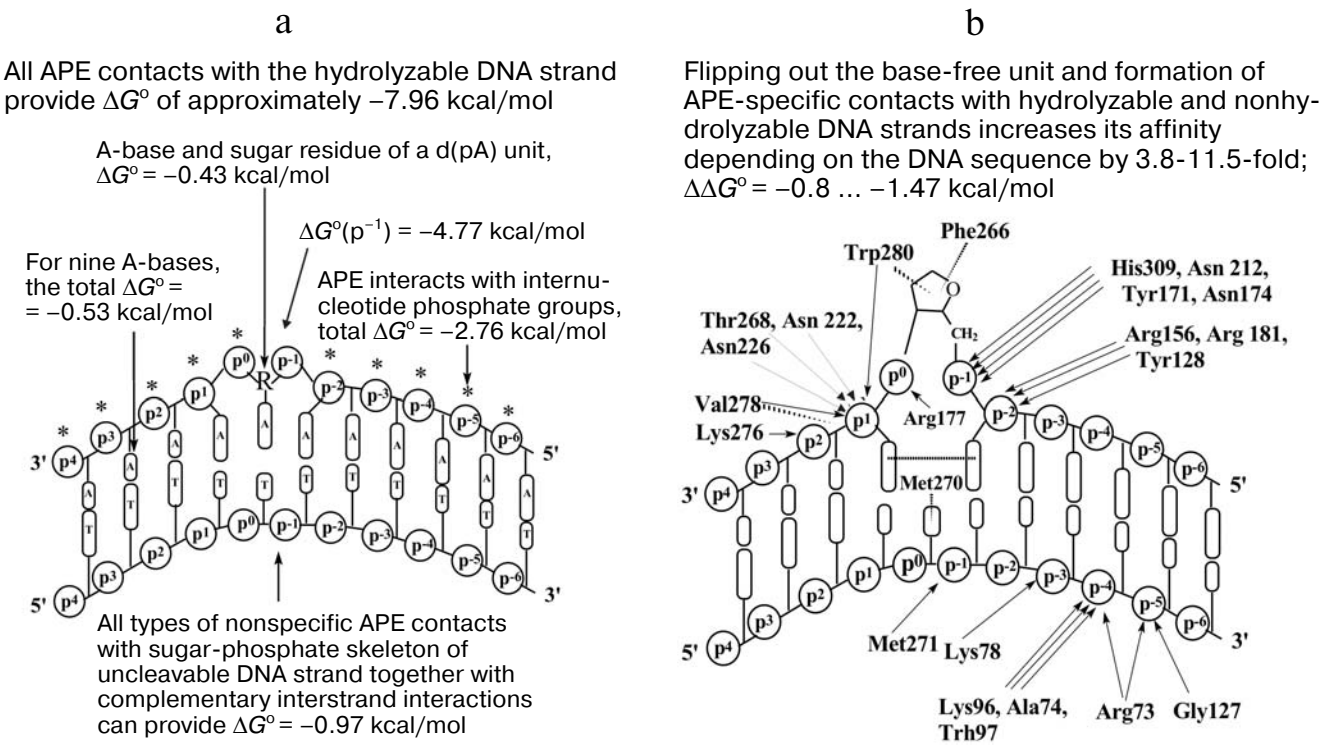

Fig. 7. Models of interaction of human AP endonuclease (APE1) with nonspecific (a) and specific (b) DNA. b) Amino acid residues of APE1 that, according to X-ray analysis, interact with DNA are shown [83] and the difference in structural features and thermodynamic characteristics of interaction of the enzyme with specific compared to nonspecific DNA are indicated.

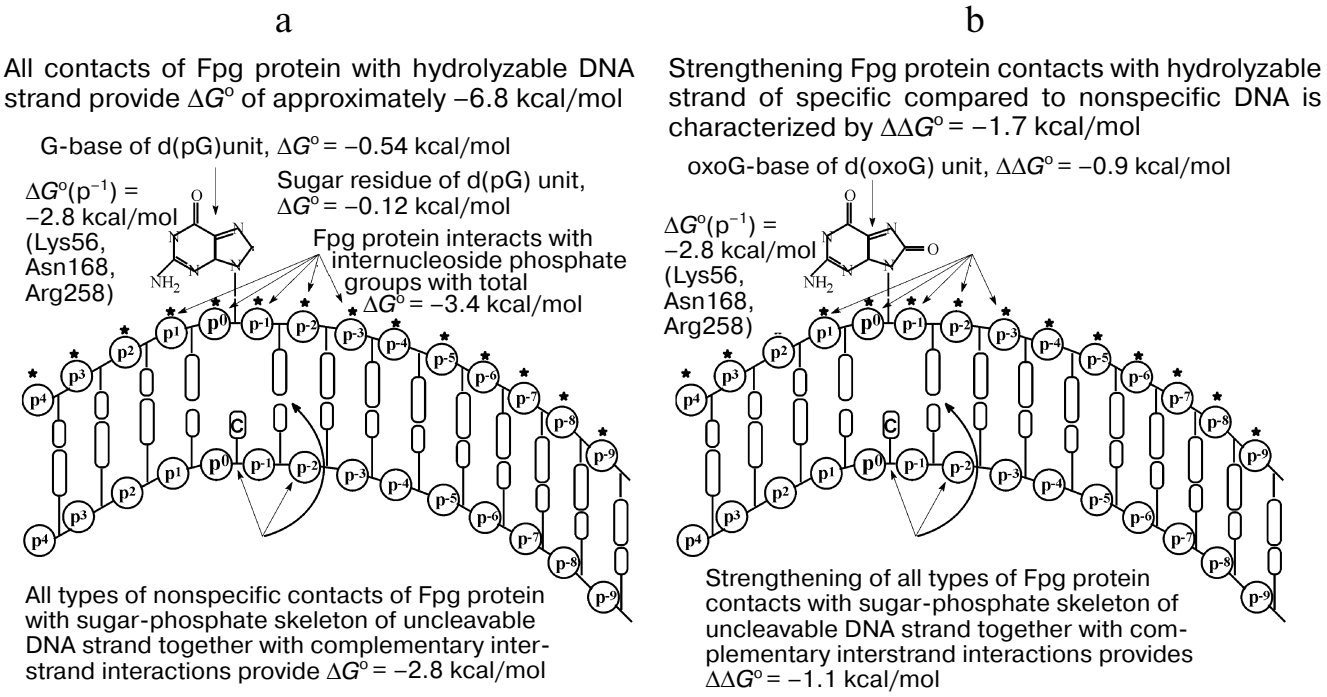

Fig. 8. Models of interaction of *E. coli* 8-oxoguanine DNA glycosylase (Fpg protein) with nonspecific (a) and specific (b) DNA. b) Difference in thermodynamic characteristics of the interaction of the enzyme with specific compared to nonspecific DNA is shown.

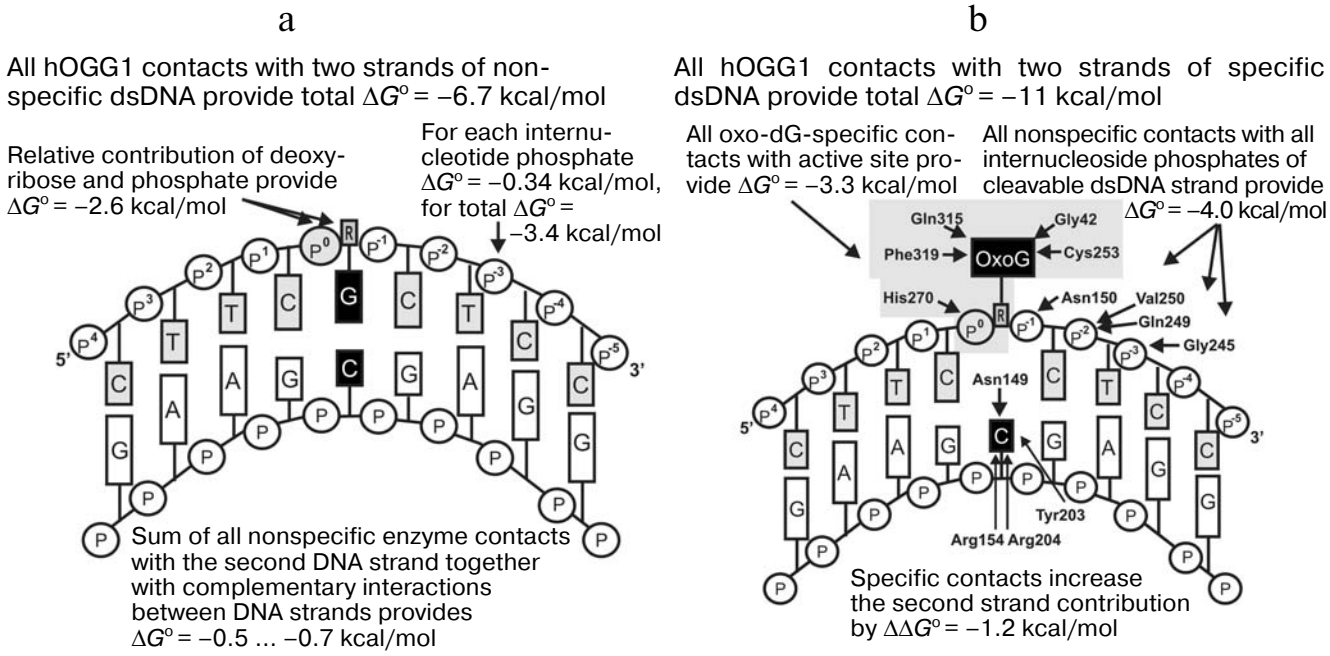

Fig. 9. Models of interaction of human 8-oxoguanine DNA glycosylase (hOGG1) with nonspecific (a) and specific (b) DNA. b) Amino acid residues of hOGG1 that according to X-ray analysis interact with DNA [26-28] are shown along with difference in thermodynamic characteristics of interaction of the enzyme with specific compared to nonspecific DNA.

urating concentrations of these ligands. Thus, at the stage of complex formation the repair enzymes are not able to differentiate with high efficiency between specific and nonspecific DNA–DNA, RNA–RNA, and DNA–RNA duplexes. High affinity of these enzymes is provided due to totality of a large number of weak nonspecific enzyme interactions with the large number of NA nucleotide units. The following questions can be formulated on the basis of data concerning relatively low difference in enzyme affinity to specific and nonspecific DNA: how do DNA-dependent enzymes rapidly search for specific sequences, and what factors provide for their specificity *in vivo*?

MECHANISM OF SEARCHING FOR DNA HOMOLOGY BY RecA PROTEIN

The mechanism of interaction of DNA with RecA filament and search for homology between DNA strands and following strand exchange is of special interest. The interactions of ssDNA with the first DNA-recognizing site of RecA protein were studied using SILC [94, 95]. It was shown that RecA, like other DNA-recognizing enzymes, is able to interact not only with $d(pN)_n$ but also with dNMP and *ortho*-phosphate (Fig. 10a and Table 1). It is seen that the initial step of filamentation is equally efficient for ssDNA of any composition and is practically independent of $d(pN)_n$ ($2 \le n \le 7$-10) sequence. The subsequent filamentation process is more specific and results in formation of firmer RecA contacts with $d(pT)_n$ and $d(pC)_n$ compared to $d(pA)_n$. This may show that necessary filament structure is formed after binding of DNA to more than three RecA molecules, which corresponds to formation of more than one turn of the filament helix.

Some DNA-recognizing enzymes can nonspecifically interact with DNA bases of hydrophobic type [66-68]. Then interaction efficiency increases along with increase in the relative hydrophobicity of the base. In the case with RecA, the inverse dependence is observed: more hydrophobic $d(pA)_n$ ligands less efficiently interact with protein than less hydrophobic $d(pT)_n$ and $d(pC)_n$ oligonucleotides [94, 95]. Comparison of affinity of RecA to $d(pA)_n$, to oligonucleotides ethylated at internucleoside phosphate groups, and to $d(pR)_n$ oligomers lacking bases indicated that RecA within extended filament does not establish efficient contacts with A-bases, while the $d(pA)_n$ affinity is mainly provided by their interaction with the sugar-phosphate skeleton [94, 95].

Since the extent of RecA-dependent ATP hydrolysis induced by different $d(pN)_{20}$ correlates well with the efficiency of protein filamentation on different ssDNA, the influence of nucleotide sequence on efficiency of interaction of RecA with different ss-poly(dN) was estimated by efficiency of ATP hydrolysis (Table 5) [94, 95]. Although guanosine has the same acceptor (O6) and donor (HN1) groups as inosine, poly(dG) like poly(dA) poorly interacted with RecA. However, deamination of poly(dG) or poly(dA) increased the efficiency of ATP hydrolysis and,

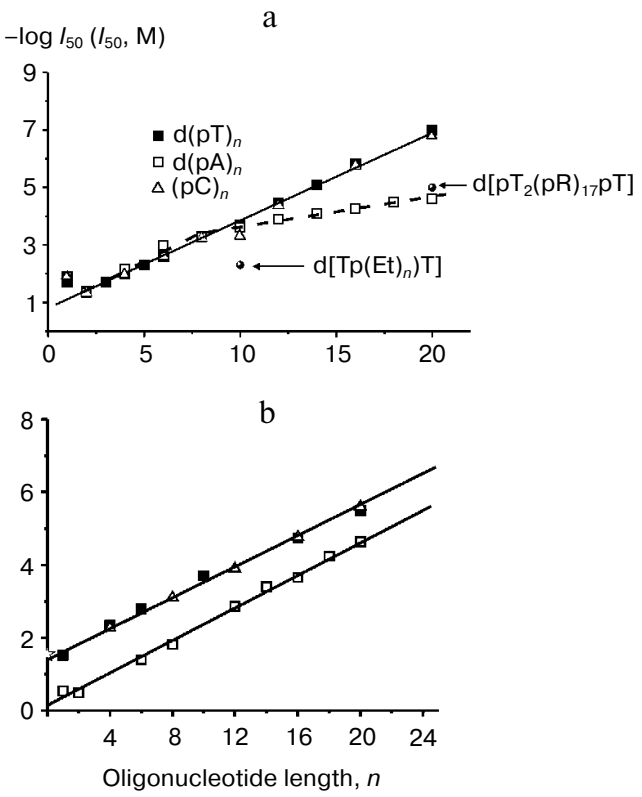

Fig. 10. Logarithmic dependences of RecA filament first (a) and second (b) site affinity (I_{50}) to different d(pN)$_n$ homooligonucleotides depending on their length (n). The I_{50} values correspond to d(pN)$_n$ concentrations at which 50% inhibition of the RecA protein monomer filamentation on d(pN)$_n$ (a) and ON binding to the second site (b) are achieved [87-89]. Arrows point to (a) positions of logI_{50} for some oligonucleotides ethylated at internucleoside phosphate groups (d[Tp(Et)]$_9$T) and for oligomers lacking (d[(pT)$_2$(pR)$_{17}$(pT)]) bases.

respectively, increased efficiency of filamentation. The same increase was noted upon transition from poly(dAG) to poly(dIX). DNA of mixed purine-pyrimidine composition interacted differently with RecA. Thus, poly(dAC) and poly(dTG) bound well to RecA, poly(dAT) exhibited intermediate efficiency between poly(dA) and poly(dT) ligands, while poly(dCG) showed practically no stimulation of RecA-dependent ATP hydrolysis (Table 5). It is necessary to note that in the presence of polyribonucleotides whose sugar-phosphate skeleton is in 3′-endo- but not in 2′-endo-conformation as in DNA, RecA does not hydrolyze ATP [94-96]. For poly(dAT) in regions with alternating purine and pyrimidine nucleotides the ribose conformation changes from 2′-endo- in the case of thymidines to 3′-endo-conformation in the case of adenosines [97], which can prevent binding of RecA to such filaments. Poly(dCG) in the presence of Mg^{2+} is in general characterized by 3′-endo-form and transition to Z-conformation [97]; therefore, such DNA interacts poorly with RecA. Deamination of mixed polynucleotides resulting in dA→dI, dG→dX, and dC→dU transitions increased the efficiency of interaction with RecA, especially in the case of poly(dCG)→poly(dUX), which is not characterized by Z-DNA conformation, and poly(dAG)→poly(dIX) (Table 5). It is interesting that replacement of adenine H_3N6 group by a halogen, which is also an acceptor group due to free electron pairs, also results in more efficient purine ligand interaction with DNA [94-96].

It is known that the A form of RNA is characterized by ribose 3′-endo-conformation [97] that should be the main indication of absence of efficient interactions between RecA and RNA. Since during filamentation from RecA DNA undergoes significant unfolding and extension by 50%, 2′-endo-conformation of sugar-phosphate skeleton should be preferable for this process because in this case the helix is less compact and the distance between nucleotides along helix the axis is 3.03-3.37 Å, unlike that in RNA where it is 2.56-3.29 Å [97]. Thus, preference of one ribose conformation allows RecA to distinguish DNA and RNA and thus to prevent homologous exchange between them [84-96].

Extension of the sugar-phosphate skeleton of DNA during RecA filamentation on it may play the most important role both for preparation of the filament specific complex with ssDNA capable of searching for homologous DNA sequence and for the strand exchange reaction. The filament pitch (95 Å) and diameter (100 Å) are average for various DNA [54]. At the same time, the filament pitch can change depending on the sequence bound to the DNA. Conformation of the DNA sugar-phosphate skeleton can vary within broad limits. For example, such B-DNA parameters as deflection angle and helix pitch is 31.12-41.41 and 3.16-4.08 Å, respectively [97]. Thus, changes induced by DNA in the RecA filament will depend on the initial conformation of the particular ssDNA sequence, while filament parameters should change depending on helix parameters. This means that primary selection of homologous sequence can take place at conformational level when different dsDNA binding efficiency will depend on conformation and structure of the filament formed on ssDNA.

Interaction of different ssDNA with the second DNA-recognizing site of RecA protein has been studied [96]. In the case of recognition of ssDNA by the second site, the protein interacts with each nucleotide unit of the ligand (Fig. 10b). Pyrimidine d(pC)$_n$- and d(pT)$_n$-oligonucleotides interact with the second site of the RecA filament more efficiently than d(pA)$_n$-oligonucleotides. This is the result of tighter interaction of the RecA filament with the 5′-terminal unit of pyrimidine DNA and by the difference in specific conformational transformations of the nucleoprotein filament caused by pyrimidine and purine DNA [96].

The binding of dsDNA to the RecA–ssDNA filament is followed by search for homology between corre-

Table 5. Maximal level of RecA-dependent ATP hydrolysis in the presence of different poly(dN) [94-96]

DNA	Maximal ATP hydrolysis, %*	DNA	Maximal ATP hydrolysis, %*
poly(dA)	1.6	poly(dG)	3.5
poly(dAT)	33.2	poly(dIT)	63.4
poly(dAC)	59.4	poly(dIX)	43.4
poly(dAG)	1.3	poly(dTX)	63.6
poly(dC)	63.1	poly(dXU)	62.1
poly(dGC)	1.5	poly(dU)	63.0
poly(dT)	61.0	poly(dI)	24.3
poly(dTG)	58.5	poly(dX)	28.4

* Error did not exceed 10% of the given value.

sponding strands and subsequent exchange. The mechanism of this process is still not clear. Several possible mechanisms for searching homology are described in the literature. It is supposed that ssDNA may replace one strand through the large or small groove of the duplex [49]. It is assumed that in the first case the exchange reaction is preceded by formation of DNA triplex called the DNA R-form [61, 62]. In the second case the search for homology is carried out only due to formation of standard Watson–Crick hydrogen bonds after base flipping out [63]. It is supposed that since bound ss- and dsDNA are significantly extended and untwisted, the bases can easily leave the bodies of these DNA and be checked for homology.

Recent analysis of data concerning processes of RecA protein filamentation on different ssDNA allowed us to propose a new model describing formation of weak specific additive hydrogen bonds between RecA protein and T-, C-, and G-bases, the existence of which is impossible in the case of A-bases [94-96]. A supposition concerning the existence near acceptor C=O groups of pyrimidines (as well as NH_2- and C=O groups of G-base) of amino acid residues of RecA filament (OH, COOH, etc.) that are proton donors suggested a model of the DNA base sequence recognition in the second site connected with the first site of the RecA filament [95, 96]. In the presence of an appropriate proton donor amino acid residues of the filament for any homology combinations within strands (A·A·T, T·T·A, C·C·G, and G·G·C), initially keto-enol and then amino-imine, shifts are possible which, after rearrangement of hydrogen bonds, result in formation of a new duplex between two DNA strands bound to the first and second filament sites (Figs. 11 and 12) [95, 96]. After strand exchange the second site binds the DNA strand displaced from homologous duplex, stabilizing in this way the newly formed duplex. It seems that the interaction of ssDNA with the second DNA-recognizing site might not be strictly specific, because correct specific interaction of homologous strands of the freshly-formed duplex is enough for homology recognition, which is consistent with our data. In this case, nucleoprotein interactions in the second site should provide for rigid fixation of the leaving strand for prevention of the reverse reaction, which is consistent with data of filament analysis using rapid kinetics technique [98].

The extremely important role in DNA recognition and its specific transformation may belong to the pathways of DNA adaptation to catalytically active state, which are individual for each repair enzyme and are described below.

SEARCH MECHANISMS FOR SPECIFIC SEQUENCES

As shown above, the stage of complex formation is not able to provide for specific DNA highly selective recovery by enzymes. If enzymes had high affinity to specific sequences and low affinity to nonspecific ones, the search for such sequences in extended DNA of the living organism genomes would be carried out via multiple association–dissociation processes by a distributive mechanism and would be slow. However, *in vitro* search for specific sequences even in large DNA proceeds at the rate comparable with that of diffusion [66-68]. As shown above, the nucleic acid-dependent enzymes are characterized by high affinity to both specific and nonspecific DNA. This means that after binding to any site of nonspecific DNA, the enzymes do not dissociate from the

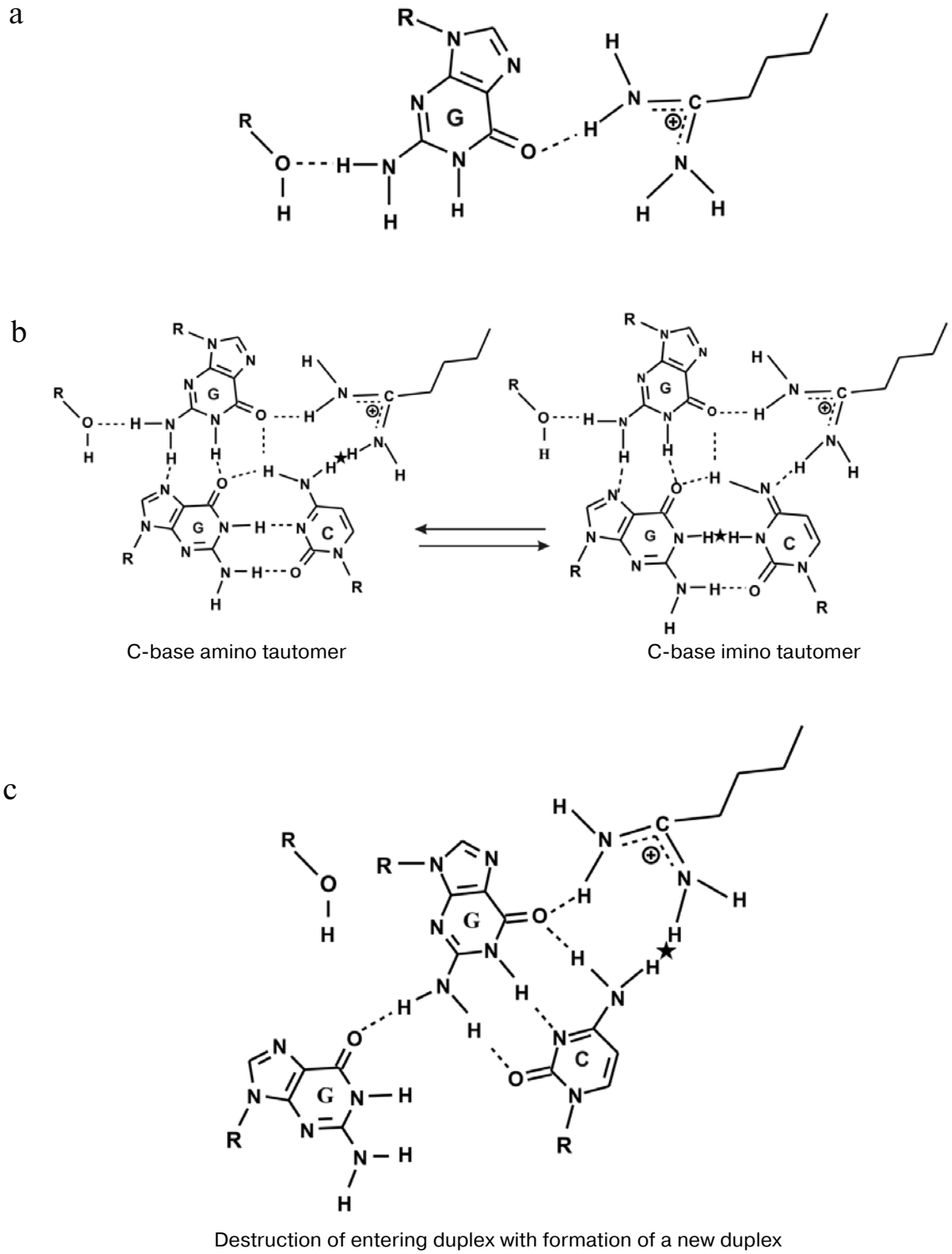

Fig. 11. Hypothetical mechanism of homologous exchange catalyzed by RecA protein considered on the example of G-C pairs. The ssDNA binding and filament formation are followed by formation of contacts of amino acid residues (most probable are Tyr and Arg residues) of RecA protein with G-base of DNA (a). The subsequent binding of DNA duplex by the RecA·ssDNA complex and search of homology are provided by formation of specific hydrogen bonds upon formation of DNA R-triplex (b). In this case RecA protein can catalyze the rupture of Watson–Crick bonds in the duplex due to shifting the amino-imine equilibrium after formation of contact between C-base and an amino acid residue of the DNA-recognizing site of RecA protein. Breaking of hydrogen bonds in the initial duplex results in formation of new dsDNA (c). The impossibility of hydrogen bonding is marked by asterisk. Tyr and Arg are used in the model as amino acid residues most suitable for this role.

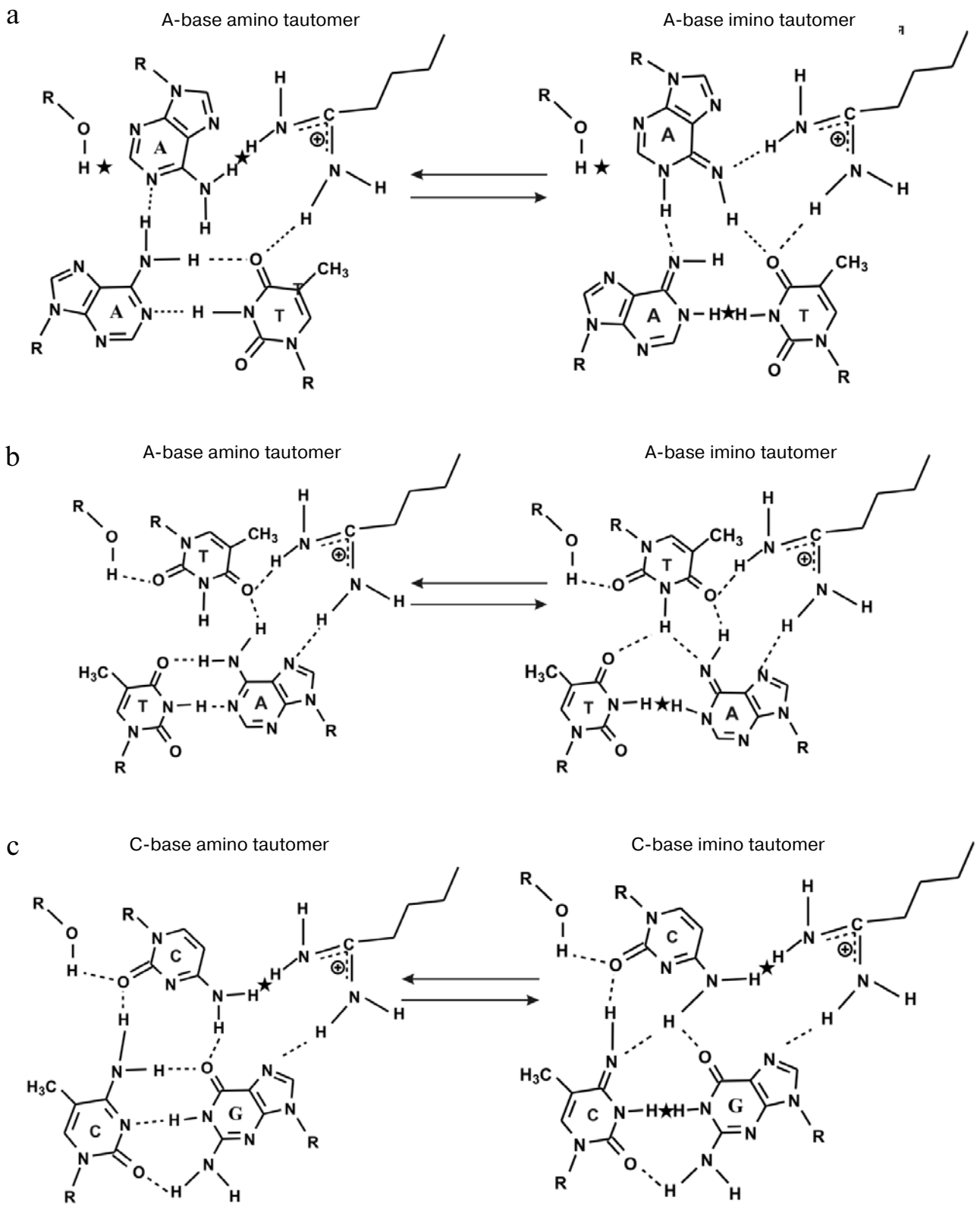

Fig. 12. Hypothetical mechanism of interaction of RecA with homologous DNA sequences within a filament that provides for destruction of initial and formation of new duplex due to formation of additional contacts between bases of the first DNA strand and RecA protein, which finally result in shifting amino-imine equilibrium in the case of A·A·T (a), T·T·A (b), and C·C·G (c) triplexes. The impossibility of hydrogen bonding is marked by asterisks.

complex but somehow slide at high speed along the DNA in the search for DNA fragments containing specific sequences.

The affinity of the enzyme to various nonspecific, and especially to hetero-sequences, is practically the same. Since the energy of enzyme binding to any fragments of extended DNA is comparable, the energy barrier for enzyme "transfer" from one to another neighboring DNA region should be low enough for its being overcome at 25-37°C [66-68]. The isothermodynamic situation of enzyme transfer along DNA should allow enzymes to "slip" along DNA at sufficiently high speed, comparable with that of diffusion of large molecules.

The first data in the literature indicative of DNA-dependent protein slipping along DNA in search for specific sequence were obtained for *lac*-repressor [99, 100]. It was shown that the rate of specific sequence searching by the protein corresponds to high-speed protein movement in one direction along DNA, rather than by a distributive mechanism. Later the processivity of UDG action was shown [101-103]. Starting from the moment of DNA binding to the moment of dissociation, APE can slip along 200 units of DNA and remove 7-8 modified units [104]. Repair enzymes MutY and Fpg are able to remove modified units depending on DNA length and salt concentration both by processive and distributive mechanisms [105]. The processive mechanism of action was also shown for RecA protein. The exchange of DNA strands catalyzed by RecA protein was studied using rapid kinetics. The rate of formation and dissociation of poly(dN) triplexes containing A-T pairs was estimated [98]. The rate of searching homology with formation of A-T pairs and the rate of strand exchange were very high and approximately five times exceeded rates of the newly-formed ssDNA dissociation from the complex. These data suggest the rapid slipping of RecA filament along dsDNA in search for the ss- and dsDNA strand homology. Obviously, the processive mechanism of action is very important and is common for all enzymes interacting with extended DNA.

It appears that many enzymes contain "specialized" amino acid residues used for detection of specific DNA units or sequences. After searching for a specific site while slipping along DNA, the enzymes stop due to a minor increase in affinity to this site, and then there is efficient change in DNA and enzyme structures [68]. Conservative Leu272, whose hydrophobic moiety intercalates between bases during the search for a dU unit, is localized in the active site of UDG. In this case only specific DNA undergoes local (but practically complete) melting of the DNA helix [7-9]. Fpg, hOOG1 [29, 87, 88, 93], and APE [90] use different hydrophobic or aromatic amino acids for the search for specific units with subsequent partial melting and destruction of complementary interactions between strands of specific DNA and its bending. Formation of catalytically competent complex due to interdependent change in DNA and protein conformation for each DNA-dependent enzyme is exclusively individual and very strongly depends on its plasticity. Thus, the protein globule of *lac*-repressor is extremely plastic, and it binds specific DNA in a conformation absolutely different from nonspecific DNA [106]. In the case of DNA-dependent enzymes, hydrophobic, hydrophilic, and charged zones in the DNA binding site are so distributed that their interaction with DNA results in formation of absolutely its structure that is not characteristic of that in solution. Change in DNA structure due to enzyme interactions can be expressed in partial or almost complete melting, change in sugar-phosphate skeleton structure, extension or compression of DNA as a whole, dilation or narrowing of the small or large groove, formation of fractures and bendings, base flipping out the DNA body, etc. Individual combination of these elements and the extent of their realization are the basis of peculiarities of assurance of specificity of action of each particular enzyme [66-68]. Selection of specific DNA by enzymes is based on the ability of the initial DNA structure characteristic of solution to change efficiently in the enzyme-assigned direction towards conformation optimal for the enzyme [66-68].

According to data of thermodynamic analysis, fixation of DNA strands in DNA-binding regions of repair enzymes results in lowering the complementary interactions between them, and the DNA appears melted to different extent [12, 36, 45, 86]. According to X-ray data on UDG, APE, Fpg, and hOGG1 [7-9, 29, 89-91], partial or complete enzyme-dependent DNA melting is an inherent part of the structural adjustment of DNA substrates. Features of adjustment processes are strongly dependent on the structure of the modified nucleotide units. Thus, among all natural dNMP only dUMP can exist in two thermodynamically almost equivalent conformations, one of which stimulates stacking of its base within DNA, while the other does not [107]. A relatively easy dU unit transition from C1′-exo-conformation to C2′-endo-conformation with preservation of the base *anti*-conformation necessary for efficient UDG catalysis stimulates practically complete melting of dU-DNA caused by the enzyme. This results in efficient U-base flipping out of dsDNA without substrate bending compared to other nucleotide units: the *trans-Gauche*-conformation for the C4′–C5′ bond of 5′-dUMP unit is more optimal for catalysis than its *Gauche-Gauche*-conformation characteristic of dsDNA [12]. In the course of "adjustment" of DNA and UDG conformations, significant structural changes of the substrate sugar-phosphate skeleton occurs, after which it cannot be assigned to either A and B form [7-9, 108]. After almost complete melting, one DNA strand firmly binds one and the other strand binds the other wall of the DNA-binding channel, and an uracil residue is turned inside-out into the "pocket" designated for its binding, which results in glycoside

bond hydrolysis. A similar situation is observed for *Hha1* (cytosine-5′)-methyl transferase [7].

Some repair enzymes carry out only partial and local DNA "melting" just to achieve its efficient bending to the different extent revealing "horseshoe" structure [26-29, 83]. Such bending of a DNA molecule makes possible destruction of stacking bases and removal of a modified unit from the DNA. Such structural alterations in DNA can be used by several repair enzymes: APE [40, 87], Fpg [29], and hOGG1 [26-28]. It should be noted that in each case specific DNA bending is an exclusively individual process, and both DNA and enzyme change during adjustment.

As shown by X-ray analysis [40, 90], APE first electrostatically orients DNA in the DNA-binding channel and then forms several additional contacts with internucleoside phosphate groups of cleaved and uncleaved strands and incorporates several amino acid residues into the large and small DNA grooves (Fig. 7b). This allows APE to bend DNA several times more efficiently and to cause its contraction near the AP site followed by deoxyribose residue flipping out of the DNA body into a specific "pocket" of the enzyme. The APE-bound DNA is bent, forming an angle of ~35°, the helix axis is displaced approximately for 5 Å, while the AP site is fixed in the active site by hydrophobic and aromatic amino acid residues of the enzyme (Fig. 7b).

According to X-ray data, a very similar mechanism of deoxyribose residue flipping out is observed in the case of another AP endonuclease, Endo IV, which has no structural similarity with APE as well as with pro- and eukaryotic glycosylases eliminating 8-oxoguanine from DNA [40]. However, processes of DNA adaptation by the latter enzymes have certain similarities with those for APE and with each other.

Specific interactions of DNA with Fpg and hOGG1 (X-ray data) [26-29] are shown in Figs. 8 and 9. Proteins Fpg and hOGG1, like APE, first orient any DNA in DNA-binding channels due to the specific electrostatic profiles of these channels. After formation of additional contacts of these enzymes mainly with internucleoside units of the cleavable strand and relatively low number of contacts with the uncleavable strand, partial DNA melting and bending occurs. Such DNA transformations are more efficient in the case of specific DNA. Since unlike canonical nucleotides existing in DNA in *anti*-conformation, the 8-oxoGMP unit is in *syn*-conformation, this causing local lowering of DNA melting temperature and easy flipped out from DNA body upon its bending. Stabilization in the enzyme active site of the flipped out base is achieved due to formation of a number of direct contacts of the enzyme with the base, as well as due to stacking of the base localized opposite 8-oxoGMP, with aromatic residues of the enzyme (Fig. 9). These data show unambiguously that upon adaptation each repair enzyme changes DNA structure in an individual way.

EFFICIENCY OF SPECIFIC CONTACTS OF REPAIR ENZYME WITH DNA

Recognition of small substrate molecules by enzymes is usually due to formation of several strong contacts like hydrogen bonds, electrostatic contacts, or stacking interactions [65-68]. The contribution of a large number (from 4 to 12) of specific contacts of the enzyme with specific units of extended DNA to its total affinity to the enzymes is usually one order of magnitude and rarely increases to two orders (see above). However, if specific enzyme contacts with long DNA were as firm as those for low molecular weight ligands, they could not discriminate extended DNA and their free low molecular weight structural components. This is especially important for repair enzymes for which similar efficiency of specific interactions with one unit at the level of free dNMP and polymeric DNA is observed [12, 36, 45, 86].

In this connection, it is interesting to estimate the efficiency of specific contacts between repair enzymes and modified units of extended DNA. In the case of formation of the specific complex of dsDNA with UDG, five hydrogen bonds with U-base are formed [7-9], which provide less than two orders of affinity (Fig. 6; ΔG^o = −1.4 ... −1.8 kcal/mol) [12]. If it is supposed that all five hydrogen bonds are approximately comparable in energy, then the energy of each such bond is rather low ($\Delta G^o \approx -0.38$ kcal/mol) and comparable with energy of weak additive nonspecific interactions ($\Delta G^o \approx -0.2$ ... −0.5 kcal/mol). Even lower ΔG^o values characterize separate specific bonds of Fpg with 8-oxoG nucleotide in DNA [36]. It should be assumed that weak contacts of sequence-dependent enzymes with specific structural elements of DNA are not very important at stage of complex formation but very important at stages of adaptation of DNA and enzyme structures and "orbital steering", which provide for increase in reaction rate upon transition from nonspecific to specific DNA by 5-8 orders of magnitude.

As mentioned above, unlike the various considered repair enzymes, upon transition from pairs (ns)-dNMP–(s)-dNMP to pairs (ns)-ssDNA–s-DNA and then to (ns)-dsDNA–ssDNA the difference in affinity of these ligand pairs to enzyme in the case of hOGG1 increases by 17- and 110-fold, respectively (Table 4) [86]. This means that among all studied repair enzymes there is no additivity, and cooperative interactions are observed only in the case of hOGG1. However, it should be assumed that such increase in efficiency of interaction between specific ss- and dsDNA is due not only to formation of specific contacts of the enzyme active site directly with oxoG unit, but also to strengthening the enzyme contacts with nonspecific DNA fragments.

According to X-ray data, APE, Fpg, and hOGG1 form in the case of specific DNA mainly hydrogen bonds with internucleoside phosphate groups (Figs. 7 and 9), while no such bonds were detected for UDG, the latter

being characterized by electrostatic contacts. Nevertheless, *e* values for all studied enzymes are low and comparable (1.35-1.78; Table 4). Thus, these enzymes can interact with internucleoside phosphate groups of nonspecific and specific DNA due to hydrogen bonds and electrostatic, ion–dipole, and dipole–dipole interactions, but they are very weak and of comparable efficiency. Upon transition to specific DNA these weak nonspecific contacts are evidently strengthened and possibly may acquire the character of specific interactions due to interdependent change in DNA and enzyme conformations causing approach of interacting groups [66-68].

FACTORS PROVIDING FOR HIGH REACTION RATE

As mentioned above, each enzyme uses its own mechanisms for changing the DNA conformation, including cooperative interactions to transform DNA structure to catalytically competent form. It is known that, although at a very low efficiency, one per 10^5-10^8 catalysis acts, repair enzymes may err and remove unmodified canonical bases [1, 12, 36, 38, 45, 86]. According to data of rapid kinetics, nonspecific DNA, unlike specific DNA, does not undergo further conformational changes after primary complex formation with Fpg [108, 109]. Nevertheless, for DNA containing 8-oxoG-unit several stages of structural changes in DNA and enzyme are observed after primary complex formation, which include incorporation of amino acid residues into the DNA body with following base flipping out, formation of the transitory state of the complex, and base removal. Similar data concerning several stages of structural transformation only for specific (unlike nonspecific) DNA were obtained by rapid kinetic technique for hOGG1 [110] and APE [111].

Since efficient transformation of DNA and enzyme is possible only for specific DNA, then upon transition from nonspecific to specific DNA the rate of repair enzyme reactions increases by 4-8 orders of magnitude (Table 4). One of the especially important factors causing the increase in enzyme reaction rates is "orbital steering" or adjustment of the enzyme and substrate reactive groups with accuracy up to 10-15° [65]. Stress caused by bond deformation by 10° constitutes 2.7 kcal/mol, and such an energy barrier is easily overcome already at 25-30°. Just at the stage of adaptation the enzyme discriminates between specific and nonspecific DNA sequences, which is defined by the possibility of "adjustment" of the DNA structure to the optimal enzyme conformation (with accuracy up to 10-15°). As a result, the reaction rate in the case of specific DNA increases by 4-8 orders of magnitude [12, 36, 45, 66-68, 86].

The stage of adaptation of DNA structure (unlike the stage of complex formation) is extremely sensitive to DNA sequence and initial structure in solution [66-68]. Even slight modifications, like incorporation of a fluorine atom into the base as well as substitution of ribose or different sugars for deoxyribose in the dU unit of DNA resulted in significant decrease in the rate of uracil residue "cleaving out" from DNA ligand, or even in the absence of reaction when affinity is the same or even increased compared to canonical substrate [13, 14]. A similar situation was observed for APE [43]. Upon transition from specific to specific modified DNA, the affinity was retained while reaction rates decreased by 4-7 orders of magnitude. Thus, like for the several different enzymes [65], not complex formation but rather the second stage of reaction, namely DNA adaptation to optimal conformation, and the third stage defining the rate of catalytic process (k_{cat}) are most important for assurance of repair enzyme specificity.

Introduction of several lesions into DNA is of special interest because in this case distortions of canonical spatial structure in DNA can be stronger than for DNA with a single lesion. Authors of works [112, 113] analyzed the effect of two AP sites localized in two complementary strands of dsDNA on efficiency of catalysis by human APE and obtained comparable results. Depending on localization of the second AP site in the complementary strand (directly opposite the first AP site or at a distance from it by several units), the affinity of dsON decreased by 40-125-fold and the reaction rate by 5-25-fold [113].

Interaction of Fpg and OGG1 as well as APE of yeast (Apn1) and *E. coli* (Nfo) with ds-d(pN)$_n$ substrates containing in same-strand oxoG and tetrahydrofuran residues at different distance from each other (R is a chemically stable analog of an apurine site) was studied [114]. The effect of introduction of the second damaged base strongly depended both on mutual arrangement of damaged bases and on the analyzed enzyme. Maximal decrease in affinity (1.6-148-fold) and transformation rate (4.8-58-fold) was observed for Fpg and OGG1 in the case of substrates containing R from the 5′-side or from the 3′-side of oxoG. Introduction of the second modified unit had a small effect on K_m values for substrates in the case of Apn1 and Nfo. In this case transformation rates decreased by 5-10-fold in the case of substrates containing two modified units in neighboring positions of the same strand. For all studied enzymes, elongation of the distance between two modified units within dsDNA resulted in decrease of their effect on K_m and k_{cat}. Thus, two modified bases cause more pronounced structural changes in DNA when they are adjacent to each other in the same or two complementary strands. This may result in pronounced decrease in both affinity and rate of reaction catalyzed by the enzyme. Due to sharp decrease in affinity to such DNA regions, repair enzymes can dissociate into solution or still repair such distortions but at a low rate; in the case of their dissociation from complexes with DNA, such complexes can be later corrected by dif-

ferent repair enzymes specifically recognizing "voluminous" distortions in DNA [1, 38].

It should be stressed that not only repair enzymes but also topoisomerization, restriction, and integration enzymes approximately equally interact with nonspecific DNA of any context due to a very large number of weak and often additive nonspecific contacts. Just these contacts provide for both high enzyme affinity to any DNA (10^{-5}-10^{-8}; Tables 1-4) and their ability to slip along DNA at a high speed in search for specific sequence. Specific interactions between enzyme and DNA play an important role mainly at the final stage of the substrate structure "adjustment" to optimal conformation and directly to the catalysis stage, but not in the primary complex formation. The efficiency of correct substrate selection by enzymes at the expense of complex formation (1-2 orders of magnitude) and following stages of change in DNA conformation and directly of catalysis (4-8 orders) finally cause 5-9 orders of specificity, which corresponds to the enzyme specificity levels observed *in vivo*.

Thus, the basis of specificity of repair enzymes including RecA protein and other enzymes [65] are not the complex formation stage, but rather the following stages that finally define relative constant of reaction rate (k_{cat}) depending on substrate structure.

This work was supported by the Program of Fundamental Investigations of Presidium of the Russian Academy of Sciences "Molecular and Cell Biology" (grant 22.7), the Russian Foundation for Basic Research (grants 07-04-00395 and 08-04-00596), and by the Siberian Branch of the Russian Academy of Sciences.

REFERENCES

1. Friedberg, E. C., Walker, G. C., Siede, W., and Wood, R. D. (2006) *DNA Repair and Mutagenesis*, ASM Press, Washington.
2. Beckman, K. B., and Ames, B. N. (1998) *Physiol. Rev.*, **78**, 547-581.
3. Krokan, H. E., Drablos, F., and Slupphaug, G. (2002) *Oncogene*, **21**, 8935-8948.
4. Warner, H. R., Duncan, B. K., Garrett, C., and Neuhard, J. (1981) *J. Bacteriol.*, **145**, 687-695.
5. Drake, J., and Baltz, R. M. (1976) *Annu. Rev. Biochem.*, **45**, 11-37.
6. Mosbaugh, D. W. (1988) *Rev. Biochem. Toxicol.*, **9**, 69-130.
7. Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995) *Nature*, **373**, 487-493.
8. Mol, C. D., Arvai, A. S., Sluppaug, G., Kavli, B., Alseth, J., Krokan, H., and Tainer, J. A. (1995) *Cell*, **80**, 869-878.
9. Mol, C. D., Arvai, A. S., Sanderson, R. J., Sluppaug, G., Kavli, B., Krokan, H. E., Mosbaugh, D. W., and Tainer, J. A. (1995) *Cell*, **82**, 701-708.
10. Vasilenko, N. L., Bulychev, N. V., Gorn, V. V., Levina, A. S., and Nevinsky, G. A. (1994) *Mol. Biol. (Moscow)*, **28**, 679-690.
11. Vinogradova, N. L., Yamkovoy, V. I., Tsvetkov, I. V., and Nevinsky, G. A. (1996) *Mol. Biol. (Moscow)*, **30**, 209-219.
12. Vinogradova, N. L., Bulychev, N. V., Maksakova, G. A., Johnson, F., and Nevinsky, G. A. (1998) *Mol. Biol. (Moscow)*, **32**, 498-508.
13. Kubareva, E. A., Volkov, E. M., Vinogradova, N. L., Kanevsky, I. A., Oretskaya, T. S., Kuznetsova, S. A., Brevnov, M. G., Gromova, E. S., Nevinsky, G. A., and Shabarova, Z. A. (1995) *Gene*, **157**, 167-171.
14. Kubareva, E. A., Vasilenko, N. L., Vorobjeva, O. V., Volkov, E. M., Oretskaya, T. S., Korshunova, G. A., and Nevinsky, G. A. (1998) *Biochem. Mol. Biol. Int.*, **46**, 597-606.
15. Krokan, H. E., Otterlei, M., Nilsen, H., Kavli, B., Skorpen, F., Andersen, S., Skjelbred, C., Akbari, M., Aas, P. A., and Slupphaug, G. (2001) *Prog. Nucl. Acid Res. Mol. Biol.*, **68**, 365-386.
16. Krokan, H. E., Drablos, F., and Slupphaug, G. (2002) *Oncogene*, **21**, 8935-8948.
17. Vasilenko, N. L., and Nevinsky, G. A. (2003) *Biochemistry (Moscow)*, **68**, 135-151.
18. Pytel, D., Slupianek, A., Ksiazek, D., Skorski, T., and Blasiak, J. (2008) *Postepy Biochem.*, **54**, 362-370.
19. Zharkov, D. O., Mechetin, G. V., and Nevinsky, G. A. (2009) *Mutat. Res.*, **685**, 11-20.
20. Vasilenko, N. L., and Nevinsky, G. A. (2003) *Mol. Biol. (Moscow)*, **37**, 944-960.
21. Demple, B., and Harrison, L. (1994) *Annu. Rev. Biochem.*, **63**, 915-948.
22. Dherin, C., Radicella, J. P., Dizdaroglu, M., and Boiteux, S. (1999) *Nucleic Acids Res.*, **27**, 4001-4007.
23. Zharkov, D. O., Rosenquist, T. A., Gerchman, S. E., and Grollman, A. P. (2000) *J. Biol. Chem.*, **275**, 28607-28617.
24. Nash, H. M., Bruner, S. D., Scharer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S., and Verdine, G. L. (1996) *Curr. Biol.*, **6**, 968-980.
25. Rosenquist, T. A., Zharkov, D. O., and Grollman, A. P. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 7429-7434.
26. Bruner, S. D., Norman, D. P. G., and Verdine, G. L. (2000) *Nature*, **403**, 859-866.
27. Bjoras, M., Seeberg, E., Luna, L., Pearl, L. H., and Barrett, T. E. (2002) *J. Mol. Biol.*, **317**, 171-177.
28. Banerjee, A., Yang, W., Karplus, M., and Verdine, G. L. (2005) *Nature*, **434**, 612-618.
29. Gilboa, R., Zharkov, D. O., Golan, G., Fernandes, A. S., Gerchman, S. E., Matz, E., Kycia, J. H., Grollman, A. P., and Shoham, G. (2002) *J. Biol. Chem.*, **277**, 19811-19816.
30. Bailly, V., Verly, W. G., O'Connor, T. R., and Laval, J. (1989) *Biochem. J.*, **262**, 581-589.
31. Zharkov, D. O., Ishchenko, A. A., Douglas, K. T., and Nevinsky, G. A. (2003) *Mutat. Res.*, **531**, 141-156.
32. Ishchenko, A. A., Bulychev, N. V., Zharkov, D. O., Maksakova, G. A., Johnson, F., and Nevinsky, G. A. (1997) *Mol. Biol. (Moscow)*, **31**, 332-337.
33. Ishchenko, A. A., Bulychev, N. V., Maksakova, G. A., Johnson, F., and Nevinsky, G. A. (1997) *Biochemistry (Moscow)*, **62**, 204-211.
34. Ishchenko, A. A., Bulychev, N. V., Maksakova, G. A., Johnson, F., and Nevinsky, G. A. (1999) *IUBMB Life*, **48**, 613-618.
35. Ishchenko, A. A., Koval, V. V., Fedorova, O. S., Douglas, K. T., and Nevinsky, G. A. (1999) *J. Biomol. Struct. Dyn.*, **17**, 301-310.

36. Ishchenko, A. A., Vasilenko, N. L., Sinitsina, O. I., Yamkovoy, V. I., Fedorova, O. S., Douglas, K. T., and Nevinsky, G. A. (2002) *Biochemistry*, **41**, 7540-7548.
37. Atamna, H., Cheung, I., and Ames, B. N. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 686-691.
38. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, ASM Press, Washington, D.C.
39. Luscombe, N. M., Austin, S. E., Berman, H. M., and Thornton, J. M. (2000) *Genome Biol.*, **1**, 1-37.
40. Mol, C. D., Hosfield, D. J., and Tainer, J. A. (2000) *Mutat. Res.*, **460**, 211-229.
41. Hadi, M. Z., Ginalski, K., Nguyen, L. H., and Wilson, D. M. (2002) *J. Mol. Biol.*, **316**, 853-866.
42. Burkovics, P., Szukacsov, V., Unk, I., and Haracska, L. (2006) *Nucleic Acids Res.*, **34**, 2508-2515.
43. Beloglazova, N. G., Lokhova, I. A., Maksakova, G. A., Tsvetkov, I. A., and Nevinsky, G. A. (1996) *Mol. Biol. (Moscow)*, **30**, 220-230.
44. Beloglazova, N. G., Petruseva, I. O., Bulychev, N. V., Maksakova, G. A., Johnson, F., and Nevinsky, G. A. (1997) *Mol. Biol. (Moscow)*, **31**, 1104-1111.
45. Beloglazova, N. G., Kirpota, O. O., Starostin, K. V., Ishchenko, A. A., Yamkovoy, V. I., Zharkov, D. O., Douglas, K. T., and Nevinsky, G. A. (2004) *Nucleic Acids Res.*, **32**, 5134-5146.
46. Ishchenko, A. A., Ide, H., Ramotar, D., Nevinsky, G., and Saparbaev, M. (2004) *Biochemistry*, **43**, 15210-15216.
47. Ishchenko, A. A., Sanz, G., Privezentzev, C. V., Maksimenko, A. V., and Saparbaev, M. (2003) *Nucleic Acids Res.*, **31**, 6344-6353.
48. Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P., and Tainer, J. A. (1999) *Cell*, **98**, 397-408.
49. Lusetti, S. L., and Cox, M. M. (2002) *Annu. Rev. Biochem.*, **71**, 71-100.
50. Amaratunga, M., and Benight, A. S. (1988) *Biochem. Biophys. Res. Commun.*, **157**, 127-133.
51. Cazenave, C., Chabbert, M., Toulme, J. J., and Helene, C. (1984) *Biochim. Biophys. Acta*, **781**, 7-13.
52. McEntee, K., Weinstock, G. M., and Lehman, I. R. (1981) *J. Biol. Chem.*, **256**, 8835-8844.
53. Tracy, R. B., and Kowalczykowski, S. C. (1996) *Genes Dev.*, **10**, 1890-1903.
54. Yu, X., and Egelman, E. H. (1992) *J. Mol. Biol.*, **227**, 334-346.
55. Register, J. C., III, and Griffith, J. (1985) *J. Biol. Chem.*, **260**, 12308-12312.
56. Bork, J. M., Cox, M. M., and Inman, R. B. (2001) *J. Biol. Chem.*, **276**, 45740-45743.
57. Stasiak, A., and Di Capua, E. (1982) *Nature*, **299**, 185-186.
58. Pugh, B. F., Schutte, B. C., and Cox, M. M. (1989) *J. Mol. Biol.*, **205**, 487-492.
59. Heuser, J., and Griffith, J. (1989) *J. Mol. Biol.*, **210**, 473-484.
60. Mazin, A. V., and Kowalczykowski, S. C. (1998) *EMBO J.*, **17**, 1161-1168.
61. Zhurkin, V. B., Raghunathan, G., Ulyanov, N. B., Camerini-Otero, R. D., and Jernigan, R. L. (1994) *J. Mol. Biol.*, **239**, 181-200.
62. Kim, M. G., Zhurkin, V. B., Jernigan, R. L., and Camerini-Otero, R. D. (1995) *J. Mol. Biol.*, **247**, 874-889.
63. Zhou, X., and Adzuma, K. (1997) *Biochemistry*, **36**, 4650-4661.
64. Howard-Flanders, P., West, S. C., and Stasiak, A. (1984) *Nature*, **309**, 215-219.
65. Fersht, A. (1980) *Enzyme Structure and Mechanism*, Freeman and Co., San Francisco.
66. Nevinsky, G. A. (1995) *Mol. Biol. (Moscow)*, **29**, 16-37.
67. Bugreev, D. V., and Nevinsky, G. A. (1999) *Biochemistry (Moscow)*, **64**, 237-249.
68. Nevinsky, G. A. (2003) in *Protein Structures: Kaleidoscope of Structural Properties and Functions* (Uversky, V. N., ed.) Research Signpost, Kerala, pp. 133-222.
69. Nevinsky, G. A. (2004) *Mol. Biol. (Moscow)*, **38**, 756-785.
70. Knorre, D. G., Lavrik, O. I., and Nevinsky, G. A. (1988) *Biochimie*, **70**, 655-661.
71. Nevinsky, G. A., Veniaminova, A. G., Levina, A. S., Podust, V. N., Lavrik, O. I., and Holler, E. (1990) *Biochemistry*, **29**, 1200-1207.
72. Kolocheva, T. I., Nevinsky, G. A., Levina, A. S., Khomov, V. V., and Lavrik, O. I. (1991) *J. Biomol. Struct. Dyn.*, **9**, 169-186.
73. Kolocheva, T. I., Maksakova, G. A., Bugreev, D. V., and Nevinsky, G. A. (2001) *IUBMB Life*, **51**, 189-195.
74. Bugreev, D. V., Vasyutina, E. L., Buneva, V. N., Yasui, Y., Nishizava, M., Andoh, T., and Nevinsky, G. A. (1997) *FEBS Lett.*, **407**, 18-20.
75. Bugreev, D. V., Vasutina, E. L., Kolocheva, T. I., Buneva, V. N., Andoh, T., and Nevinsky, G. A. (1998) *Biochimie*, **80**, 303-308.
76. Bugreev, D. V., Buneva, V. N., Sinitsina, O. I., and Nevinsky, G. A. (2003) *Bioorg. Chem. (Moscow)*, **29**, 163-174.
77. Bugreev, D. V., Sinitsina, O. I., Buneva, V. N., and Nevinsky, G. A. (2003) *Bioorg. Chem. (Moscow)*, **29**, 275-288.
78. Bugreev, D. V., Buneva, V. N., and Nevinsky, G. A. (2003) *Mol. Biol. (Moscow)*, **37**, 325-339.
79. Bugreev, D. V., Baranova, S., Zakharova, O. D., Parissi, V., Desjobert, C., Sottofattori, E., Balbi, A., Litvak, S., Tarrago-Litvak, L., and Nevinsky, G. A. (2003) *Biochemistry*, **42**, 9235-9247.
80. Bugreev, D. V., Vasyutina, E. L., Maksakova, G. A., Buneva, V. N., Ando, T., and Nevinsky, G. A. (1997) *Mol. Biol. (Moscow)*, **31**, 418-430.
81. Kolocheva, T. I., Nevinsky, G. A., Volchkova, A. S., Levina, A. S., Khomov, V. V., and Lavrik, O. I. (1989) *FEBS Lett.*, **248**, 97-100.
82. Kornberg, A. (1977) *DNA Synthesis* [Russian translation], Mir, Moscow, p. 71.
83. Kolocheva, T. I., Demidov, S. A., Maksakova, G. A., and Nevinsky, G. A. (1998) *Mol. Biol. (Moscow)*, **32**, 1025-1033.
84. Ishchenko, A. A., Bulychev, N. V., Maksakova, G. A., Johnson, F., and Nevinsky, G. A. (1998) *Mol. Biol. (Moscow)*, **31**, 549-558.
85. Doronin, S. V., Lavrik, O. I., Nevinsky, G. A., and Podust, V. N. (1987) *FEBS Lett.*, **216**, 221-224.
86. Kirpota, O. O., Zharkov, D. O., Buneva, V. N., and Nevinsky, G. A. (2006) *Mol. Biol. (Moscow)*, **40**, 1055-1043.
87. Serre, L., Pereira de Jesus, K., Boiteux, S., Zelwer, C., and Castaing, B. (2002) *EMBO J.*, **21**, 2854-2865.
88. Fromme, J. C., and Verdine, G. L. (2002) *Nat. Struct. Biol.*, **9**, 544-552.

89. Gorman, M. A., Morera, S., Rothwell, D. G., de La Fortelle, E., Mol, C. D., Tainer, J. A., Hickson, I. D., and Freemont, P. S. (1997) *EMBO J.*, **16**, 6548-6558.
90. Mol, C. D., Izumi, T., Mitra, S., and Tainer, J. A. (2000) *Nature*, **403**, 451-456.
91. Beernink, P. T., Segelke, B. W., Hadi, M. Z., Erzberger, J. P., Wilson, D. M., III, and Rupp, B. (2001) *J. Mol. Biol.*, **307**, 1023-1034.
92. Bjoras, M., Luna, L., Johnsen, B., Hoff, E., Haug, T., Rognes, T., and Seeberg, E. (1997) *EMBO J.*, **16**, 6314-6322.
93. Sugahara, M., Mikava, T., Kumasaka, T., Yamomoto, M., Kato, R., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000) *EMBO J.*, **19**, 3857-3869.
94. Bugreeva, I. P., Bugreev, D. V., and Nevinsky, G. A. (2005) *FEBS J.*, **272**, 2734-2745.
95. Bugreeva, I. P., Bugreev, D. V., and Nevinsky, G. A. (2005) *Mol. Biol. (Moscow)*, **39**, 984-998.
96. Bugreeva, I. P., Bugreev, D. V., and Nevinsky, G. A. (2007*) Mol. Biol. (Moscow)*, **41**, 524-534.
97. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, N. Y.
98. Folta-Stogniew, E., O'Malley, S., Gupta, R., Anderson, K. S., and Radding, C. M. (2004) *Mol. Cell*, **15**, 965-975.
99. Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981) *Biochemistry*, **20**, 6928-6948.
100. Winter, R. B., and von Hippel, P. H. (1981) *Biochemistry*, **20**, 6948-6960.
101. Higley, M., and Lloyd, R. S. (1993) *Mutat. Res.*, **294**, 109-116.
102. Bennett, S. E., Sanderson, R. J., and Mosbaugh, D. W. (1995) *Biochemistry*, **34**, 6109-6119.
103. Sidorenko, V. S., Mechetin, G. V., Nevinsky, G. A., and Zharkov, D. O. (2008) *FEBS Lett.*, **582**, 410-414.
104. Carey, D. C., and Strauss, P. R. (1999) *Biochemistry*, **38**, 16553-16560.
105. Francis, A. W., and David, S. S. (2003) *Biochemistry*, **42**, 801-810.
106. Kalodimos, C. G., Bonvin, A. M., Salinas, R. K., Wechselberger, R., Boelens, R., and Kaptein, R. (2002) *EMBO J.*, **21**, 2866-2876.
107. Viswamitra, M. A., and Seshadri, T. P. (1975) *Nature*, **258**, 542-544.
108. Kuznetsov, N. A., Koval, V. V., Zharkov, D. O., Vorobjev, Y. N., Nevinsky, G. A., Douglas, K. T., and Fedorova, O. S. (2007) *Biochemistry*, **46**, 424-435.
109. Kuznetsov, N. A., Koval, V. V., Nevinsky, G. A., Douglas, K. T., Zharkov, D. O., and Fedorova, O. S. (2007) *J. Biol. Chem.*, **282**, 1029-1038.
110. Kuznetsov, N. A., Koval, V. V., Zharkov, D. O., Nevinsky, G. A., Douglas, K. T., and Fedorova, O. S. (2005) *Nucleic Acids Res.*, **33**, 3919-3931.
111. Timofeyeva, N. A., Koval, V. V., Knorre, D. G., Zharkov, D. O., Saparbaev, M. K., Ishchenko, A. A., and Fedorova, O. S. (2009) *J. Biomol. Struct. Dyn.*, **26**, 637-652.
112. Chaudhry, M. A., and Weinfeld, M. (1997) *J. Biol. Chem.*, **272**, 15650-15655.
113. McKenzir, J. A., and Strauss, P. R. (2001) *Biochemistry*, **40**, 13254-13261.
114. Starostin, K. V., Ishchenko, A. A., Zharkov, D. O., Buneva, V. N., and Nevinsky, G. A. (2007) *Mol. Biol. (Moscow)*, **41**, 112-120.