Review

How Enzymes, Proteins, and Antibodies Recognize Extended DNAs; General Regularities

Georgy A. Nevinsky

Institute of Chemical Biology and Fundamental Medicine, Siberian Division of Russian Academy of Sciences, 63009 Novosibirsk, Russia; nevinsky@niboch.nsc.ru

Abstract: X-ray analysis cannot provide quantitative estimates of the relative contribution of non-specific, specific, strong, and weak contacts of extended DNA molecules to their total affinity for enzymes and proteins. The interaction of different enzymes and proteins with long DNA and RNA at the quantitative molecular level can be successfully analyzed using the method of the stepwise increase in ligand complexity (SILC). The present review summarizes the data on stepwise increase in ligand complexity (SILC) analysis of nucleic acid recognition by various enzymes—replication, restriction, integration, topoisomerization, six different repair enzymes (uracil DNA glycosylase, Fpg protein from *Escherichia coli*, human 8-oxoguanine-DNA glycosylase, human apurinic/apyrimidinic endonuclease, RecA protein, and DNA-ligase), and five DNA-recognizing proteins (RNA helicase, human lactoferrin, alfa-lactalbumin, human blood albumin, and IgGs against DNA). The relative contributions of structural elements of DNA fragments "covered" by globules of enzymes and proteins to the total affinity of DNA have been evaluated. Thermodynamic and catalytic factors providing discrimination of unspecific and specific DNAs by these enzymes on the stages of primary complex formation following changes in enzymes and DNAs or RNAs conformations and direct processing of the catalysis of the reactions were found. General regularities of recognition of nucleic acid by DNA-dependent enzymes, proteins, and antibodies were established.

Keywords: different enzymes and proteins; anti-DNA antibodies; general regularities of DNA recognition

1. Introduction

DNA- and RNA-dependent enzymes, proteins, and antibodies play a vital role in many key cellular processes—transcription, replication, recombination, repair, integration, chromosome dynamics, and protection from viruses and bacteria. Therefore, understanding the molecular mechanisms of DNA-dependent enzymes, proteins, and antibody action is critical from fundamental and applied points of view.

It is evident that X-ray analysis data on the structures of complexes between enzymes (proteins) and DNA (or RNA), including conformational changes during their interactions, play an important role in the understanding of mechanisms of nucleic acid recognition [1–16]. However, X-ray analysis data cannot provide quantitative estimations of different molecular contacts relative importance, or the relative contributions of weak, moderate, or strong specific and unspecific contacts to the total affinity of proteins for nucleic acids (NAs). X-ray analysis of sequence-specific enzymes with DNAs or RNAs sometimes leads to a misunderstanding of the real role of some specific contacts. There is a point of view that the specific contacts between proteins and DNAs that are revealed by X-ray analysis can provide a high affinity for specific nucleotides or sequences of nucleic acids (reviewed in [11–16]). There are only a few literature data on the quantitative estimation of the relative individual contributions of individual-specific nucleotides and sequences to thermodynamic (formation of complexes) and kinetic (k_{cat}, constant rates of reaction) steps to the affinity of enzymes and proteins for DNAs or to the specificity of

enzymes actions [11–16]. However, only a detailed quantitative estimation of the relative contributions of all nucleotide units of NAs can provide a correct interpretation of data obtained using an X-ray structure analysis. Therefore, it has been have developed a special approach—stepwise increase in ligand complexity (SILC)—to estimate the relative contributions of all individual nucleotides or specific sequences and their various structural elements to an enzyme's affinity for long DNAs [11–16]. The peculiarity of this approach is the gradual complication of the structure of DNA ligands, starting with the possible minimum ligand (structural elements of DNA), which can form any bonds with proteins or enzymes according to the scheme:

*Ortho*phosphate (Pi) or mononuleotide (as possible minimal ligands of enzymes)→ single-stranded (ss) non-specific homo-d(pN)n ⟶ ss non-specific hetero-d(pN)n→ ss specific hetero-d(pN)n ⟶ double-stranded (ds) non-specific homo-d(pN)n ⟶ ds specific hetero-d(pN)n ⟶ extended specific ds DNA

Using the SILC approach, we analyzed replication, restriction, integration, topoisomerization, six different repair enzymes (uracil DNA glycosylase, Fpg protein from *E. coli*, human 8-oxoguanine-DNA glycosylase, human apurinic/apyrimidinic endonuclease, RecA protein, and DNA ligase), and NA-recognizing proteins (RNA helicase, lactoferrin, lactalbumin, human serum albumin (HSA), and IgGs against DNA).

Analysis of various interactions of these enzymes, proteins, and antibodies with long nucleic acids by the SILC approach has shown that the formation of contacts between them and specific nucleotides or sequences of NAs cannot provide their observed high affinity for DNAs and RNAs. Actually, all nucleotide units "covered" by the DNA-binding clefts interact with enzymes and proteins. High affinity is mainly (5–8 orders of magnitude) provided by many weak, unspecific, additive interactions between the enzymes and proteins with different structural elements of many nucleotide units of DNAs and RNAs, mostly with internucleoside phosphate groups and only sometimes with hydrophobic bases [11–16]. The relative contribution of specific contacts to the total affinity of DNAs is rather small, not exceeding one, and rarely reaches two orders of magnitude. The enzyme specificity is provided by the enzyme-dependent stage of adjustment of DNAs or RNAs to their optimal conformations, which provide an increase in k_{cat} values by 4–8 orders of magnitude ongoing from non-specific to specific DNAs.

2. DNA Polymerases

The first enzymes to study general mechanisms of DNA recognition were DNA-dependent DNA polymerases of pro- and eukaryotes [17–21]. These enzymes recognize the template-primer complex.

The Gibbs free energy ($\Delta G°$) for the complexes of ligands with enzymes and proteins are usually equal to the sum of the $\Delta G°$ values for all individual contacts formed

$$\Delta G°_{sum} \text{ (corresponding to all } n \text{ contacts)} = \Delta G°_1 + \Delta G°_2 + \ldots + \Delta G°_n \tag{1}$$

where all $\Delta G°_{1-n}$ corresponds to all individual contacts of the ligand and $\Delta G°$ corresponds to each individual contact = $-RT \times \ln K_d$ [22]. The total K_d value characterizing the formation of any protein–enzyme complex with the ligand is the product of K_d values for all individual contacts

$$K_d \text{ (total)} = K_d(1) \times K_d(2) \times K_d(n) \tag{2}$$

The mechanism of DNA recognition by DNA polymerases (and other enzymes and proteins) was analyzed at the molecular level using the SILC approach in accordance with the following scheme: *ortho*phosphate (Pi) or mononucleotide (as possible minimum ligands of enzymes) → single-stranded (ss) non-specific homo-d(pN)n → ss non-specific hetero-d(pN)$_n$ → ss specific hetero-d(pN)$_n$ → double-stranded (ds) non-specific homo d(pN)$_n$ → ds specific hetero d(pN)$_n$ → extended specific ds DNA.

For estimation of different ligands affinities to the template binding sites of DNA polymerase from *E. coli* (Klenow fragment) and human placenta DNA polymerase alpha

was used Pt^{2+}-containing oligodeoxyribonucleotide (Pt^2 -ON) [17–21]. The K_d values for all ligands were estimated from inhibition of polymerase activities inactivation with Pt^2 –ON affinity reagent interacting with template binding sites with high affinity. All K_d values obtained are provided in Supplementary Table S1. These data were presented in the form of logarithmic (-Log) dependencies of the K_d values on the number of $d(pN)_n$ mononucleotide units (n). As an example, Figure 1A shows the data for the DNA polymerase I from *E. coli*. Such dependences were linear at $n = 1$–20 for both DNA polymerases, which indicated the additivity of free energies characterizing enzyme interactions with individual nucleotide units of $d(pN)_n$ [17–21]. The n values corresponding to the affinity changes and reaching a plateau at $n = 20$ correlated with the relative size of enzyme globules and the length of DNA fragments protected by the enzymes from their hydrolysis by DNases [17–21].
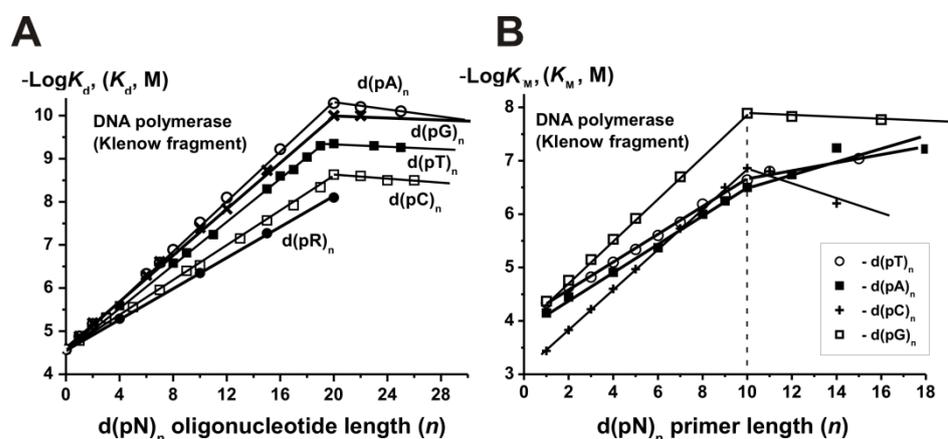
Figure 1. Dependencies of $-\text{Log}K_d$ characterizing the affinity of ss $d(pN)_n$ deoxyribooligonucleotides to template-binding site of DNA polymerase (Klenow fragment, *Escherichia coli*) versus their length (n) (**A**). All $d(pN)_n$ are shown in panel A; $d(pR)_n$ oligomers do not contain any bases [19–21]. Dependencies of $-\text{Log}K_m$ values corresponding to different $d(pN)_n$ primers on their length (n) in the reaction catalyzing by Klenow fragment of *E. coli*. (**B**). The K_m values in each case were determined using the poly(N) templates complementary to the primers used [23]. Figure 1A shows that the minimal ligands of the DNA polymerase template binding site are *ortho*phosphate and various dNMPs. With the lengthening of the ss $d(pN)_n$, the affinity gradually increases up to 20 nucleotide units, which are covered by the enzyme globule. In addition, the growth in the affinity of ss $d(pN)_n$ for the enzyme increases with an increase in the hydrophobicity of their bases: C < T < G < A. From the slopes of the curves for oligonucleotides, one can calculate the coefficient reflecting the increase in affinity with lengthening of each of the $d(pN)_n$ by one nucleotide unit.

It was finally shown that the interaction of single-stranded (ss) DNAs with template binding sites of DNA polymerases leads to an increase in the DNA affinity due to interaction with one nucleotide units with different bases of $d(pN)_n$ by the factor *F*, which include C (1.58), T (1.78), G (1.95), and A (2.0) [17–21]. The K_d value is determined by a superposition of weak additive non-specific electrostatic and hydrogen bonds, and hydrophobic and/or Van der Waals interactions with individual nucleotides is described by a decreasing geometric progression [17–21]

$$K_d[d(pN)_n] = K_d[(P_i)] \times (E)^{1-n} \times (h_C)^{-c} \times (h_T)^{-t} \times (h_G)^{-g} \times (h_A)^{-a} \qquad (3)$$

where K_d [(Pi)] $- K_d$ stands for the minimal *ortho*phosphate ligand. The electrostatic factor $E = 1.52$ reflects the increase in affinity due to an interaction of one internucleoside phosphate group of DNA with DNA polymerase template binding site [17–21]. Factor *F* is the product of factors *E* and h_N ($F = E \times h_N$). The $h_C = 1.04$, $h_T = 1.17$, $h_G = 1.28$, and $h_A = 1.32$ values correspond to hydrophobic factors h_N, reflecting the increase in the efficiency of the enzyme interaction with one of the bases (C, T, G, or A) of $d(pN)_n$, the number of which in

this ligand is *c*, *t*, *g*, and *a*, respectively [18–21]. For both homo- and hetero-oligonucleotides, the contributions of internucleoside phosphates (and different bases) to the affinity of ss DNAs to the template binding sites are the same. The theoretical K_d values calculated using Equation (3) and the K_d values obtained from experiments are nearly completely coincided [18–22]. When passing from one to another DNA polymerase, only very slight changes in the numerical values of K_d [(Pi)], factors *E*, and h_N are observed [18–22].

It was shown that only one DNA internucleoside phosphate group has a high affinity for DNA polymerases (K_d = 26 μM, Klenow fragment) due to forming one Me^{2+}-dependent electrostatic contact ($\Delta G° \approx -1.2$ kcal/mol) and one hydrogen bond ($\Delta G° \approx -4.7$ kcal/mol) with the enzyme. At first glance, all electrostatic contacts 19 of 20 internucleoside phosphate groups and 20 hydrophobic interactions of bases with template binding sites of DNA polymerases are very weak. However, these weak contacts with the enzyme template-binding site, due to their additivity, provide the whole $\Delta G° \approx -4.6$ kcal/mol [18–22].

Complementary to templates, primers play a special role in the primer–template complex interaction with DNA polymerases. The K_m values for $d(pN)_n$ primers were measured as functions of their lengths (Supplementary Table S2) [21,23]. It was found that all mononucleotides (dNMPs) can serve as minimal primers of DNA synthesis. The $-$Log dependencies of K_m values increased linearly until *n* = 10 nucleotide units (Figure 1B) [21,23]. The increase in the length of $d(pT)_n$ and $d(pA)_n$ primers by one nucleotide link led to an increase in their affinity by 1.74-fold, while $d(pC)_n$ and $d(pG)_n$ by a factor of 2.33. Considering two hydrogen bonds between A and T bases of the template and primer and vice versa, or three such bonds between C and G from the 1.74 and 2.33 values square and cube roots respectively were extracted, giving 1.32 value. It was shown that only one 3'-terminal nucleotide of the primer forms strong contacts with DNA polymerases, while 9 of its remaining units only with the complimentary template. The factor 1.32 reflects the increase in the primer's affinity for DNA due to the formation of one Watson–Crick hydrogen bond between the template and the primer. The contribution to the affinity of some internucleoside phosphate groups of the template and primer was estimated using oligonucleotides ethylated at phosphate groups [17–21]. Finally, using all the data obtained, including those presented in Figure 1, the thermodynamic model, i.e., the template–primer binding with DNA polymerases, was proposed and compared with available crystallographic data (Figure 2) [21,23].

Figure 2. A thermodynamic model for the interaction of human DNA polymerase α with $d(pT)_{20} \times d(pA)_{10}$ [20,21]. It can be seen that the globule of DNA polymerase covers 20 nucleotide units of the template and 10 nucleotides of complementary primer. Bases of the template and 19 out of 20 internucleoside groups form weak additive contacts with the template-binding site of the enzyme (total $\Delta G° = -4.6$ kcal/mol), while one internucleoside phosphate group of the template has strong contacts—one hydrogen bond and one electrostatic contact (total $\Delta G° = -4.7$ kcal/mol). Only the 3'- terminal unit of the primer makes contacts with the enzyme, forming ONE hydrogen bond WITH THE 3'-OH primer group and two contacts WITH THE FIRST from the 3'-end internucleoside phosphate group (total $\Delta G° = -4.6$ kcal/mol). The rest of the primer nucleotide units interact only with the complementary template through the formation of Watson–Crick hydrogen bonds.

Later, the mechanisms of DNA recognition by various enzymes and proteins were investigated. It turned out that all enzymes possess one site of an increased affinity for various dNMPs and for only one nucleotide unit of lengthy DNA and they are capable of recognizing both specific and unspecific DNA with high affinity. Equation (3) has been shown to be applicable to describe the interaction of any ss DNAs with any sequence-independent, and non-specific DNAs with any of sequence-dependent proteins and enzymes. When passing from one to another protein or enzyme and from ss to ds DNA, only a strong change in the numerical values of K_d [(Pi)] and a slight change in the factors E and h_N are usually observed [11–16].

3. Uracil DNA Glycosylase

Uracil DNA glycosylases (UDGs) form a ubiquitous, highly conserved specific class of DNA repair enzymes of prokaryotic, eukaryotic organisms, in addition to poxviruses and herpesviruses. They specifically remove uracil base from modified DNA [24,25]. UDGs hydrolyze the glycosidic bond between the deoxyribose sugar and the U base. The resulting abasic site is removed then by apurinic/apyrimidinic endonuclease (APE) and a deoxyribophosphate lyase. The formed gap is filled up by the action of DNA polymerases and DNA ligases. UDGs demonstrate high specificity for efficient removal of uracil from ss and ds DNA, with faster excision from ss DNA [25–27].

The general patterns of non-specific ss DNA recognition by UDG were the same as for DNA polymerases. The $-\text{Log}K_d$ dependencies on the number of nucleotide units (n) in the case of UDG for ss $d(pN)_n$ stop increasing at $n = 9$–10, indicating that UDG globule covers only 9–10 DNA base pairs (Figure 3A) [26,27]. Similar to DNA polymerases, the minimal ligands of UDG were *ortho*phosphate ($K_d = 1.7 \times 10^{-2}$ M) and various dNMPs having different affinity to the enzyme (3.3–15×10^{-3} M) (Supplementary Table S3).

All data on the analysis of UDG and other enzymes using the SILC approach were compared with the published data on the X-ray analysis of these enzymes [28–52]. The main regularities of extended DNA recognition by DNA polymerases and uracil DNA glycosylase were compared with those in the case a large number of other enzymes and proteins [53–69].

UDG formed weak additive contacts with both internucleoside phosphate groups ($E = 1.35$) and bases $h_C = 1.08$, $h_T = 1.22$, $h_G = 1.35$, and $h_A = 1.42$. The affinity of UDG for ss $d(pA)_n$ is ~1.5-fold higher than for ds $d(pA)_n \times d(pT)_n$ (Figure 3A) [25–27]. This is because UDG partially melts base pairing in ds $d(pN)_{10}$, and its contacts with two strands of ds DNAs almost independent [5,28]. According to X-ray data, the structure of the sugar–phosphate backbone of DNA after the binding of UDG with DNA is drastically changed [5,28]. A conservative Leu-272 amino acid of the DNA-binding cleft is positioned in UDG's active center; this Leu intercalating with hydrophobic bases locally melts the DNA helix. A further transformation of DNA structure under the action of the enzyme is related to its melting and destruction of stacking interactions in the double helix covered by the enzyme. Then, the uracil residue "flips outside" [5,28]. As it will be shown below, various enzymes recognize DNA according to similar regularities, but changes in the conformation of DNAs leading to specific catalytically competent states occur in very different ways including deformation of DNA backbone—stretching or compression, partial or complete DNA melting, bending or kinking, eversion of nucleotides from the DNA helix, etc. (see below).

Among all dNMPs, only dUMPs can exist in two almost thermodynamically equivalent conformations, one of which provides the stacking of the bases in ds DNA and the other does not [13,14]. Thus, only UDG can relatively easily melt ds dU-DNAs in comparison with non-specific DNAs promoting effective flipping of the uracil outside the DNA. Only after the uracil residue enters a pocket for its binding, the catalytic process becomes possible.
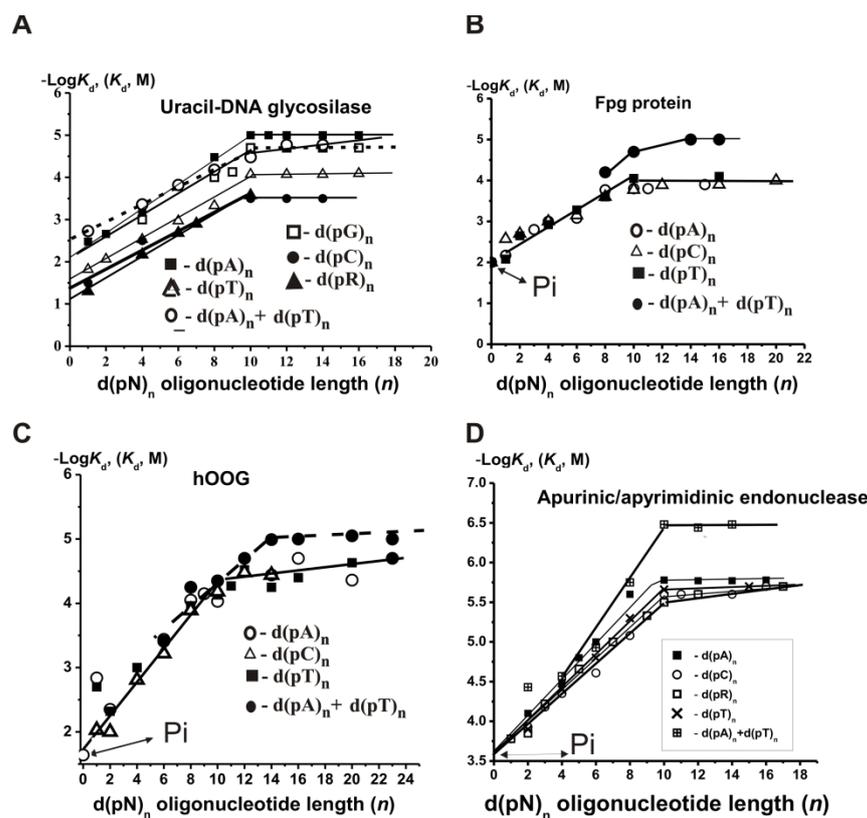
Figure 3. Dependencies of $-\text{Log}K_d$ for ss and ds d(pN)$_n$ deoxyribooligonucleotides versus their length (n) in the case of uracil DNA glycosylase (**A**) [26,27], Fpg protein (**B**) [54–57], hOGG1 (**C**) [59], and apurinic/apyrimidinic endonuclease (**D**) [69]. All oligonucleotides used are shown in panel (**A–D**); d(pR)$_n$ oligomers do not contain any bases.

Considering all weak and strong unspecific and specific contacts, the interaction of UDG with unspecific DNAs was described using a thermodynamic model (Figure 4A). According to X-ray analysis [5,28], the DNA dU unit forms five hydrogen bonds with UDG (Figure 4B). However, all these hydrogen bonds between UDG and specific U-base of DNA provide only ~1 order of affinity ($\Delta G° \approx -1.4$ to -1.8 kcal/mol, (Figure 4B) [26,27]. According to the thermodynamic analysis of modified uracil's contribution using analogs of an oligonucleotide containing 2′-fluoro-2′-deoxyuridine, its interaction with UDG was characterized by $\Delta G° < -2$ kcal/mol [29].

UDG specifically recognizes the uracil base ($\Delta G° \approx -1.4 \ldots -1.8$ kcal/mol; Figure 4B), but the contribution of its weak additive unspecific interactions with ss DNA links is quite high ($K_d = 1.0 \times 10^{-5}$ M; $\Delta G° \approx -5.9$ kcal/mol; Figure 4A). UDG's affinity for specific ds or ss homo- and hetero-oligonucleotides (ODNs) depending on their sequences, is only ~10–30 times higher than for non-specific ODNs [26,27]. The efficiency of excision of A-, C-, T-, and G-bases by UDG is about five orders of magnitude lower than U-base [26,27]. Thus, the stage of complexation of UDG with specific DNA provides only approximately one order of specificity of its action. The formation of specific contacts between different enzymes and cognate DNAs, including UDG, is not very important on the complexation stage but exclusively important on the stage of DNAs and enzyme conformations adaptations [11–16,26,27]. This adaptation stage provides very precise alignment of the electronic orbitals—"orbital steering" of the reacting atoms and it occurs only in the case of specific DNAs. The specificity of UDG action is provided by the stages of the UDG-dependent dU-DNA adaptation to the optimal conformation, which leads to the increase in catalytic constant (k_{cat}) by ~5 orders of magnitude in comparison with unspecific DNA [26,27]. It is interesting that the stage of complexation of a large number of enzymes

provides only 1–2 orders of magnitude of their specificity, while in the case of specific DNA the k_{cat} increases by 4–7 orders of magnitude (see below) [11–16].
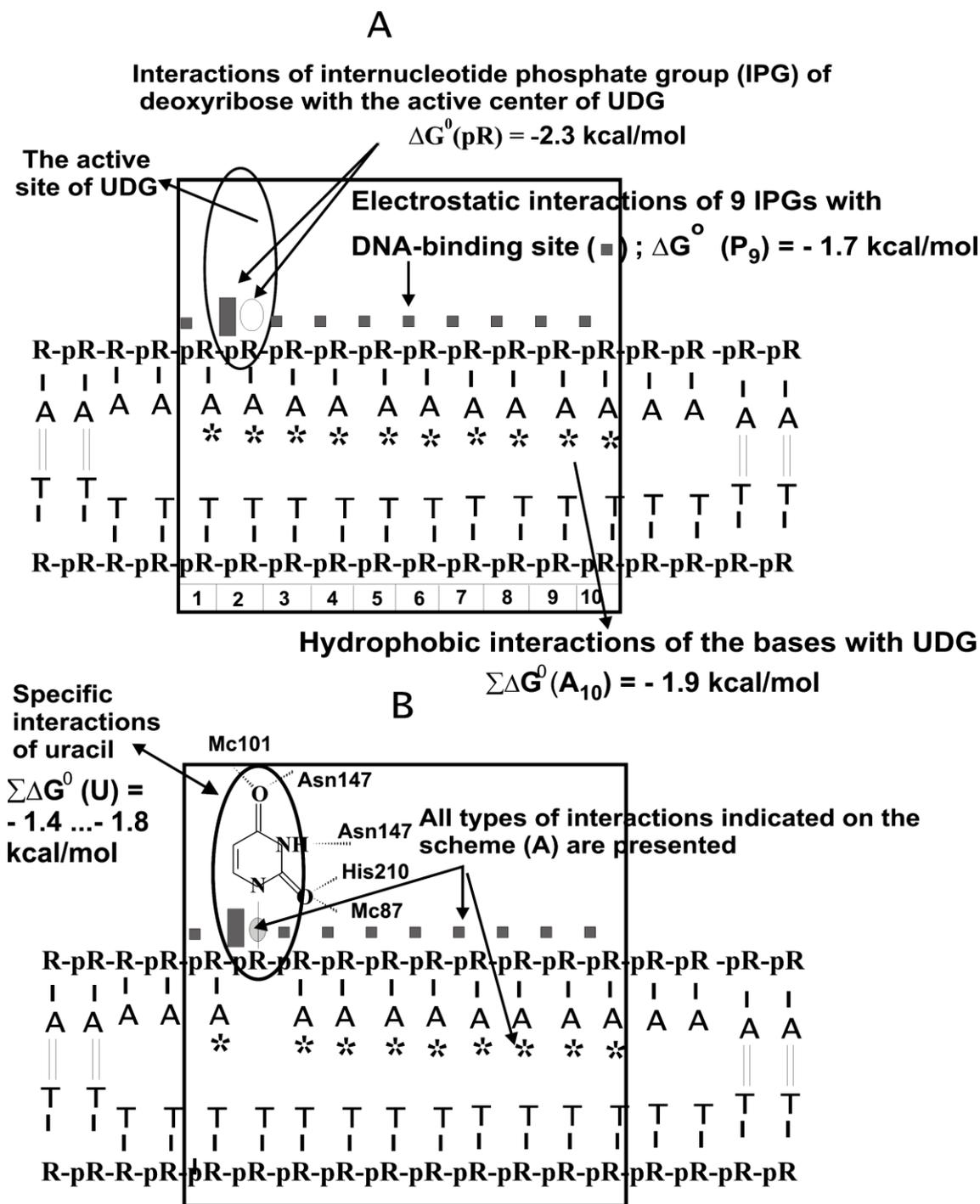
Figure 4. The thermodynamic model of human UDG interaction with unspecific (**A**) and specific DNA containing dU nucleotide unit (**B**). The $\Delta G°$ value characterizing of five specific hydrogen bonds between UDG and U-base was evaluated to be approximately −1.3 kcal/mol (**B**), while unspecific electrostatic (■) and hydrophobic (★) interactions provide $\Delta G° = −5.9$ kcal/mol (**A**) [27].

4. 8-Oxoguanine DNA Glycosylases

Oxidative damage to the cell including different pre-mutagenic modifications of DNA bases is important for both carcinogenesis and aging [30–34]. One of the most common DNA lesions is 7,8-dihydro-8-oxoguanine (8-oxoG). During the synthesis of DNA, the appearance of 8-oxoG leads to a replacement of a G:C pair with a T:A pair [35–38]. 8-oxoG is removed from DNA in *E. coli* by Fpg protein [39–44]. In addition to its *N*-glycosylase activity, Fpg has a nicking activity, cleaving both the 3'- and 5'-phosphodiester bonds at an apurinic/apyrimidinic (AP) site by successive δ- and β-elimination, resulting in both the 3'- and 5'-end of the gap phosphorylated [45,46].

The lesion 8-oxoG is removed from DNA in human cells by oxoguanine-DNA glycosylase (hOGG1). Fpg protein is not homologous to hOGG1 [47,48], and the X-ray structure of hOGG1 and its complexes with DNA also have no similarity with Fpg [50–53]. Human OGG1posses only two catalytic activities—DNA glycosylase and a lyase eliminating the 3'-phosphate at the resulting abasic site [53]. Thus, it was interesting to compare the thermodynamic aspects of 8-oxoG recognition by Fpg and hOGG1.

SILC analysis of oxoguanine-DNA glycosylases was described in [54–62]. The minimal ligands of Fpg [55–59] (Supplementary Table S4) and hOGG1 [60–62] are *ortho*phosphate (1.0×10^{-2} and 6.8×10^{-2} M), respectively) and various dNMPs ((3.3–8.3) $\times 10^{-3}$ and (1.5–9.4) $\times 10^{-3}$ M, respectively), having comparable affinity to these enzymes (Supplementary Table S4). In contrast to DNA polymerases (Figure 1A), the affinity of unspecific ss homo- and hetero-ODNs for Fpg [55–58] (Supplementary Table S5) and hOGG1 [59] did not remarkably depend on the oligonucleotide bases. The $-\text{Log}K_d$ dependencies for $d(pR)_n$ oligomers containing only the sugar–phosphate backbone (R is a tetrahydrofuran analog of abasic deoxyribose) were comparable with those for all $d(pN)_n$ containing different bases (Figure 3B,C). In the case of Fpg, the factor *E* is equal to 1.56 [55–59]. The affinity of specific 8-oxoG units within specific ss and ds DNA for Fpg protein is comparable with that for the enzyme's affinity for free 8-oxo-dGMP [58]. The affinity of specific ss and ds ODNs for Fpg was ~10-fold higher than for unspecific ones (Supplementary Table S6).

The K_d values of hOGG1 for different ligands are presented in Supplementary Table S7. The cognate ss ODN is bound by hOGG1 250-fold stronger than unspecific ss oligonucleotides [59]. In contrast, the affinity of hOGG1 for ds ODN containing 8-oxoG ($K_d = 0.011$ μM) is higher than for different ds ODNs containing G by a factor of 790–890, while the difference in the affinity of free dGMP and oxo-dGMP is only 4.1-fold (Supplementary Table S7) [59]. Thus, the relative contribution of specific interactions of the hOGG1 with 8-oxoG unit within specific DNA is significantly higher than with free 8-oxo-dGMP. However, one cannot exclude that in contract to Fpg, the specific interaction between the 8-oxoG unit of ss and ds DNAs and the active site of hOGG1 can provide a cooperative effect on the specific DNAs binding. For example, some previously formed unspecific weak contacts with nucleotide units of non-cognate ODNs may probably be greatly strengthened when cognate DNA is bound.

Figure 5 demonstrates the thermodynamic model of Fpg interaction with unspecific DNAs [57], while Figure 6 shows the same for hOGG1. Figure 7 shows the thermodynamic model of hOGG1 interaction with specific DNA [59].

The X-ray structure of non-cognate and cognate hOGG1 × DNA complexes and stopped-flow data indicate that this enzyme distorts any specific and unspecific DNAs, creating a sharp kink (bends into an arc; Figures 5–7), nevertheless, it fails to insert the unspecific bases into the active site pocket [49,50,60,61].

All contacts of Fpg with cleaved strand provide ΔG^0 = - 6.0.. -6.8 kcal/mol

G-base, ΔG^0 = - 0.54 kcal/mol

Sugar moiety , ΔG^0 = - 0.12 kcal/mol

Each internucleotide phosphate, group (★), ΔG^0 = - 0.26 kcal/mol

Totally ΔG^0 = - 2.6.. - 3.4 kcal/mol

ΔG^0 = -2.7...2.8 kcal/mol

All type of unspecific contacts of Fpg with backbone and C-base on noncleaved strand together with complementary interactions betwwen the strands can provide ΔG^0 = - 2.7...-2.9 kcal/mol

All type of unspecific interactions of Fpg with nonspecific DNA can be characterized by ΔG^0 = - 9.7...-9.8 kcal/mol

Figure 5. The thermodynamic model of *E. coli* Fpg interaction with unspecific DNA. All types of unspecific contacts are shown on the Panel [57]. The enzyme forms weak additive contacts only with internucleoside phosphate groups of DNA and there are contacts only of the active center with one of the DNA bases.

All contacts of hOGG1 with two strands of nonspecific ds DNA provide overall total ΔG^O of approximately - 6.7 kcal/mol

Deoxyribose and phosphate contribution of approxiomalely ΔG^O = - 2.6 kcal/mol

Each internucleotide phosphate group ΔG^O = - 0.34 kcal/mol, totally ΔG^O of approximately - 3.4 kcal/mol

The sum of all nonspecific contacts of the enzyme with the second strand together with complementary interaction between the strands can provide the increase in ΔG^O of approximately - (0.5 - 0.7) kcal/mol

Figure 6. The thermodynamic model of hOGG1 interaction with unspecific DNA. All types of unspecific contacts are shown on the Panel [59]. The enzyme forms weak additive contacts only with internucleoside phosphate groups of DNAs except for one deoxyribose phosphate.
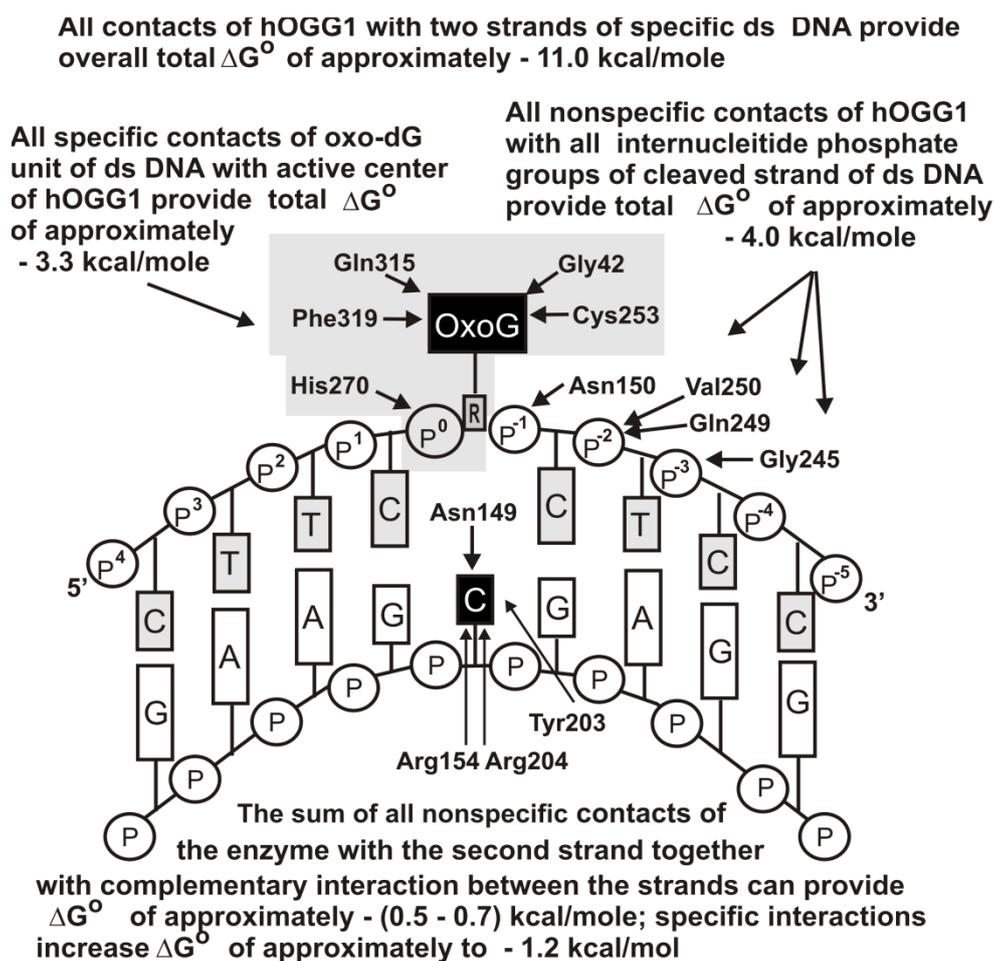
Figure 7. The thermodynamic model of hOGG1interaction of with specific DNA containing 8-oxoG. All types of unspecific contacts are shown on the Panel [59].

As in the case of UDG, the contribution of a cognate 8-oxoG nucleotide to the Fpg affinity for DNA is close to one order of magnitude [54–58]. The affinity of hOGG1 for specific DNA is higher than that for non-specific DNAs by about two orders of magnitude [59]. However, in both cases, the complexation stage cannot provide 8–9 orders of these enzymes' action specificities. At the same time, the transition from unspecific to specific DNAs leads to an increase in the k_{cat} values by 6–8 orders of magnitude. Thus, the specificity of these enzymes' actions (8–9 orders of magnitude) is realized due to two factors, including the stage of specific DNA recognition (1–2 orders of magnitude) and direct catalysis (6–8 orders of magnitude) [54–61].

5. Apurinic/Apyrimidinic Endonuclease

Apurinic/apyrimidinic or abasic (AP) sites can be formed in DNA due to spontaneous base loss, as a result of DNA treatment with different chemicals or physical mutagens (UV or ionizing radiation) and removing damaged bases by several DNA glycosylases [24]. AP-endonucleases recognize AP sites, and cleaving the DNA phosphodiester backbone creates a free 3'-OH terminus suitable for the action of DNA polymerases. The major human apurinic/apyrimidinic endonuclease (APE1) is homologous to the *E. coli* AP endonuclease Xth [62–66]. Crystal structures of wild-type and mutant APE1 bound to DNA containing AP site provide many details concerning possible mechanisms of AP site recognition and its removal due to the backbone significant compression-squeeze for placing the sugar residue of AP–DNA in the active center of the enzyme [62–64,67,68]. This is another way of DNA adaptation to a catalytically active state.

According to SILC analysis data, APE1 interacts with D-ribose ($K_d > 0.17$ M), *ortho*phosphate ($K_d = 360$ μM), different dNMPs ($K_d = 163$ μM) as minimal ligands, and 9–10 nucleotide units of ss $d(pN)_n$ or base pairs of ds ODNs of different lengths and sequences (Supplementary Tables S8 and S9; Figure 3D). The thermodynamic model of DNA recognition by APE1 is presented in Figure 8.

Deoxyribose
$\Delta G^\circ = -0.83$ kcal/mol
$\Delta\Delta G^\circ = -0.40$ kcal/mol

$\Delta G^\circ = -5.54$ kcal/mol
$\Delta\Delta G^\circ = -0.77$ kcal/mol

All nonspecific interactions of internucleotide phosphate groups (*) and bases provide $\Delta G^\circ = -3.0$ kcal/mol

The sum of all type of nonspecific and specific contacts of APE1 with backbone of noncleaved strand together with complementary interactions between the strands can provide ΔG^0 of approximately **- 0.2 to - 1.0 kcal/mol**
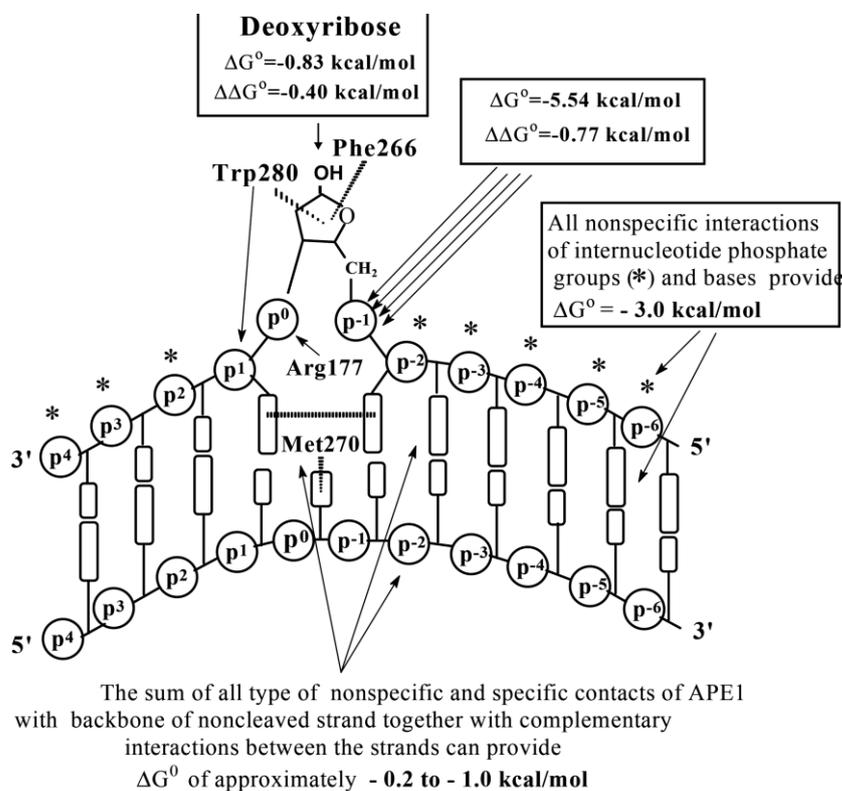
Figure 8. The thermodynamic model of apurinic/apyrimidinic endonuclease (APE1) interaction with specific DNA containing abasic (AP)-site. All types of contacts are shown on the Panel [69].

The affinity of unspecific ss homo- and hetero-$d(pN)_n$ for human APE1 depends very moderate on the sequences of ODNs (Figure 3D) [69]. The interactions occur mainly due to weak additive contacts with internucleoside phosphate groups (factor $E = 1.51$) [69]. These non-specific interactions of APE1 with every nucleotide unit provide a high affinity for ss ($K_d = 1.7$ μM) and increased by a factor of 5.2 for ds DNA ($K_d = 0.33$ μM). The affinity of AP-DNA containing the abasic site is higher compared to unspecific DNA only by one order of magnitude ($\Delta G^\circ \approx -1.1$ kcal/mol to -1.5 kcal/mol). Therefore, the formation of the enzyme-DNA complex cannot alone provide the observed specificity of enzyme catalysis. The specificity of APE1, as in the case of other enzymes, lies in the k_{cat} value, which is elevated by 6–7 orders of magnitude upon transition from unspecific to specific DNAs [69]. The thermodynamic model of DNA recognition by APE1 is presented in Figure 8.

6. DNA Ligase

Mammalian cells contain DNA ligase I and II [2]. We studied using the SILC approach human DNA ligase I [12,13]. Figure 9 demonstrates the K_d values' $-$Log-dependency on the length (n) of ODNs for DNA ligase. Like for other enzymes, *ortho*phosphate ($K_d = 1 \times 10^{-3}$ M) and dNMPs are the minimal ligands of this enzyme. The additive interactions of DNA ligase with internucleoside phosphate groups of ss DNAs remarkably stronger (factor $E = 2.14$). The increase in affinity for ODNs and F factors up to $n = 10$ proceeds in the same order as for DNA-polymerases with the increase in the relative hydrophobicity of the bases C < T < G < A.
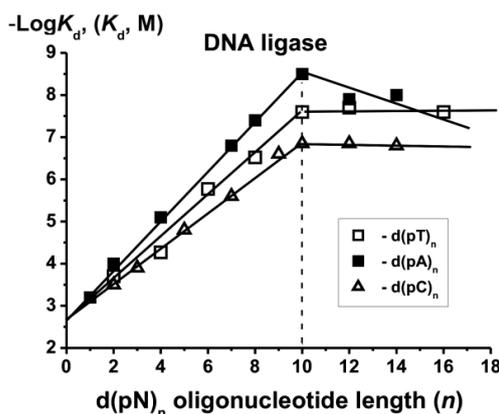
Figure 9. Dependencies of $-\mathrm{Log}K_{\mathrm{d}}$ for ss and ds d(pN)$_n$ deoxyribooligonucleotides versus their length (n) in the case of DNA ligase [12,13].

The change in the K_{d} values for different d(pN)$_n$ is described by a progression

$$K_{\mathrm{d}}[\mathrm{d(pN)}_n] = [K_{\mathrm{d}}(\mathrm{Pi}) = 1 \times 10^{-3}] \times [e = 2.14]^{-n} \times [\mathrm{h_C} = 1.1]^{-c} \times [\mathrm{h_T} = 1.29]^{-t} \times [\mathrm{h_G} = 1.51)]^{-g} \times [\mathrm{h_A} = 1.62]^{-a}.$$

The transition from ss to ds DNA leads to an increase of the ligand affinity approximately by one order of magnitude. DNA ligase I recognizes DNA similar to other analyzed enzymes.

7. RecA Protein

Homologous recombination is a very important process required to maintain genetic diversity and repair damaged DNA in all living organisms. RecA is the protein of SOS-reparation (a form of induced repair, manifested in the ability of a cell to respond to large DNA damage) playing a central role in homologous recombination in *E. coli*. RecA protein provides a search for homology between two DNA molecules and realizes the exchange of homologous strands [70]. RecA is an ATP-dependent, DNA-binding protein (37.8 kDa) [70]. The process of RecA binding to DNAs occurs in three stages, which are (1) *presynapsis*, during which RecA is polymerized on ss DNA, forming a right-helical nucleoprotein filament; (2) *synapsis*, when ds DNA binds to the *presynaptic* complex, and a search for homology with ss DNA occurs and, finally, (3) *strand exchange*, during which a new DNA duplex is formed and ss DNA is displaced. The binding of RecA protein to ss DNA occurs in a non-specific manner; however, there is a preference for protein filamentation on poly(dT) and on GT-rich sequences [71–74]. In the presence of ATP, RecA forms a right-handed 100 Å diameter filament with DNA [75]. This filament assembly occurs cooperatively in the direction of 5′→ 3′ ss DNA [76,77]. In a complex with a protein, DNA is stretched by about 50% (another way of DNA adaptation) the distance between nucleotides increases to 5.1 Å and there are three RecA monomers and nine DNA nucleotide residues per filament helix step [75]. Upon binding to ds DNA, the filament parameters are the same as for ss DNA, and the duplex in the complex turns out to be untwisted compared to the B-form of DNA [78,79]. In the absence of ATP, a more compact inactive RecA filament is formed with a step of 64 Å and a distance between nucleotides of 2.1 Å [80].

After filament formation, RecA can bind ds DNA, which interacts with the protein's second DNA-recognition site. The second RecA center can also bind ss DNA interacting with a filament even more efficiently than ds DNA. The role of the second site of the RecA protein after the homologous strand exchange reaction is to bind the leaving strand after the formation of a new duplex [81].

After binding of ds DNA by the filament, a search for homology between the corresponding strands takes place and the subsequent exchange of strands. Several models have been proposed for the search for homology and chain exchange [82–85], which have not received further experimental confirmation, however (see below).

To understand the subtle mechanisms of the interaction of any enzymes with specific and non-specific DNAs, it was very important to have quantitative data on the effectiveness of the interaction of enzyme globules with each of the nucleotide links of extended DNA covered by them. The mechanism of the search for homology between DNA strands with a subsequent exchange of strands by RecA was of particular interest.

It was shown that the first RecA site, similar to other DNA-recognizing enzymes, interacts with *ortho*phosphate (K_d = 0.5 M) and different dNMPs ((1.2–2.0) $\times 10^{-2}$ M)) as minimal ligands (Supplementary Table S10 [86,87]). The initial filamentation stage proceeds with nearly the same efficiency on ss d(pN)$_n$ of different sequences. The filamentation process on the second site of RecA protein is more specific and leads to the formation of stronger contacts of RecA with d(pT)$_n$ and d(pC)$_n$ as compared to d(pA)$_n$ (Figure 10) [86,87].
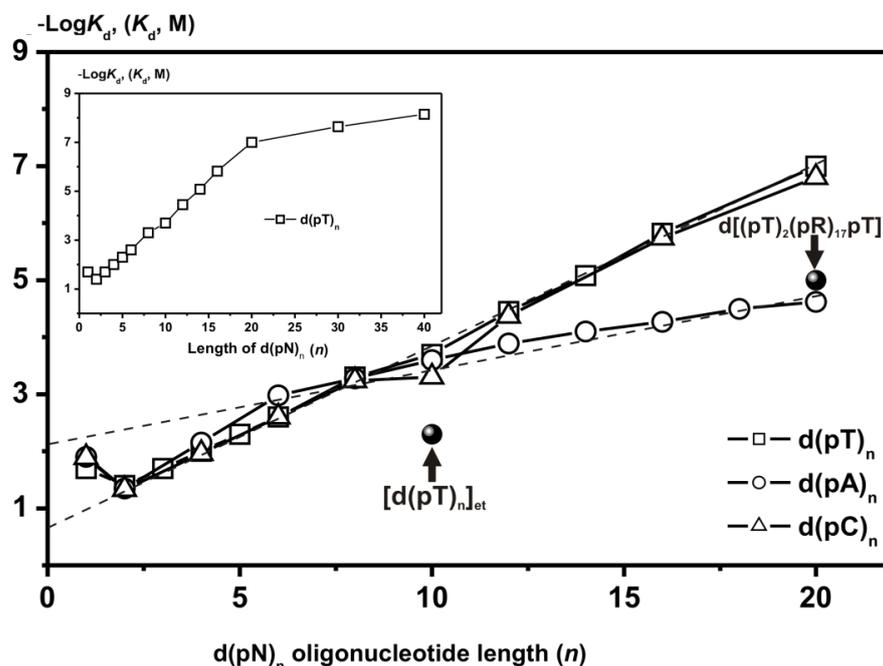
Figure 10. Dependencies of $-$LogK_d for ss and ds d(pN)$_n$ deoxyribooligonucleotides versus their length (*n*) in the case of the second site of RecA protein [86–88]. For comparison, the inset shows the Log-dependences of the K_d values characterizing the affinity of the first site of the RecA filament to d(pT)$_n$. All types of ODNs including those containing not many bases [d(pT)$_2$(pR)$_{17}$pT]) and completely ethylated at internucleoside phosphate groups ([d(pT)$_n$]$_{et}$ are shown on the Panel.

In the case of some other DNA-recognizing enzymes, the interaction efficiency increases with an increase in the bases' relative hydrophobicity [11–16]. RecA demonstrates an inverse relationship—more hydrophobic d(pA)$_n$ ligands interact with the second site of RecA worse than less hydrophobic d(pT)$_n$ and d(pC)$_n$ oligonucleotides [86,87]. Based on a comparison of the affinity of RecA for d(pA)$_n$ with that for ODNs ethylated at internucleoside phosphate groups, and for d(pR)$_n$ oligomers containing no bases, it was concluded that, in an extended filament, RecA does not form effective contacts with A-bases, and the affinity d(pA)$_n$ is provided mainly only due to RecA interaction with the sugar-phosphate backbones [86,87].

Because the degree of RecA-dependent ATP hydrolysis induced by different d(pN)$_{20}$ correlated well with the efficiency of protein filamentation on different ss DNA, the effect of different bases on the efficiency of the interaction of RecA with various ss poly(N) was assessed by the efficiency of ATP hydrolysis (Supplementary Table S11) [86–88]. Despite the fact that guanosine has the same O6 acceptor and HN1 donor groups as inosine, poly(dG), such as poly(dA), weakly interacted with RecA. Deamination of poly(dG) or poly(dA) led to an increase in the efficiency of ATP hydrolysis, and, accordingly, to more efficient

filamentation of the modified DNAs [88]. The same increase occurred when moving from poly(dAG) to poly(dIX). Mixed purine–pyrimidine DNA interacted with RecA in different ways. Thus, poly(dAC) and poly(dTG) bind well to RecA; poly(dAT) showed an intermediate efficiency between poly(dA) and poly(dT) ligands, while poly(dCG) hardly stimulated RecA-dependent ATP hydrolysis (Supplementary Materials, Table S11).

It is known that the A-form of RNA is characterized by the 3'-endo-conformation of ribose [89], which might be the main factor of the absence of effective interactions between RecA and RNA [86–88]. Since DNA significantly unfolds and stretches by 50% during the filamentation with RecA, the 2'-endo-conformation of the sugar-phosphate backbone should be more preferable for this process, because in this case, the helix is less compact, and the distance between nucleotides along the helix axis is 3.03–3.37 Å, in contrast to RNA, where it is 2.56–3.29 ÅB [89]. Thus, the preference for one of the ribose conformations allows RecA to distinguish between DNA and RNA, and thereby hinder the homologous exchange only between DNAs [86–88].

After binding ds DNA with RecA − ss DNA filament, a homology search between the corresponding strands occurs, which is followed by the exchange. The mechanism of this process is still unclear. Several possible homology search mechanisms have been described in the literature. It was assumed that ss DNA can replace one of the strands through the duplex's major or minor groove [70]. It is believed that in the first case, before the exchange reaction, a DNA triplex is formed, which was named the R-form of DNA [82,83]. In the second, alternative case, the homology search may be carried out only due to the formation of standard Watson–Crick hydrogen bonds, which are formed after base eversion [84]. It is assumed that because the bound ss and ds DNA molecules are substantially stretched and untwisted, the bases can easily leave the bodies of this DNA and be checked for homology. However, these models have not received further experimental confirmation [82–88].

Based on the data on the analysis of the processes of the filamentation of the RecA protein on various ss and ds DNA, we proposed a new model describing the formation of specific weak additive hydrogen bonds of the RecA protein with T-, C-, and G-bases, which cannot be in the case of A-bases [86–88]. The assumption of the presence in the vicinity of the acceptor C = O groups of pyrimidines and NH_2- and C = O groups of the G base, in addition to amino acid residues of the RecA filament (OH, COOH, etc.), which are proton donors, made it possible to construct a model for recognizing the nucleotide sequence DNA in the second site, which is linked to the first site of the RecA filament [87,88]. In the presence of appropriate proton-donor filament amino acid residues for any combination of chain homology (A • A • T, or C • C • G; and G • G • C), first the keto-enol and then the amino-imine shifts can occur, which results in reorganization of hydrogen bonds leading to the formation of a new duplex between two DNA strands linked to the first and second sites of the filament (Figure 11) [87,88]. After the strand exchange reaction, the second site binds the DNA strand that is displaced from the homologous duplex, stabilizing the newly formed duplex. It should be assumed that the interaction of ss DNA with the second DNA-recognition center may not have a strictly specific character, since the correct specific interaction between homologous strands of the newly formed duplex is sufficient for homology recognition, which is consistent with our data. In this case, the nucleoprotein interactions at the second site should provide a strong anchorage of the outgoing strand to prevent the reverse reaction, which is consistent with the data of the filament analysis using the rapid kinetic method [90].
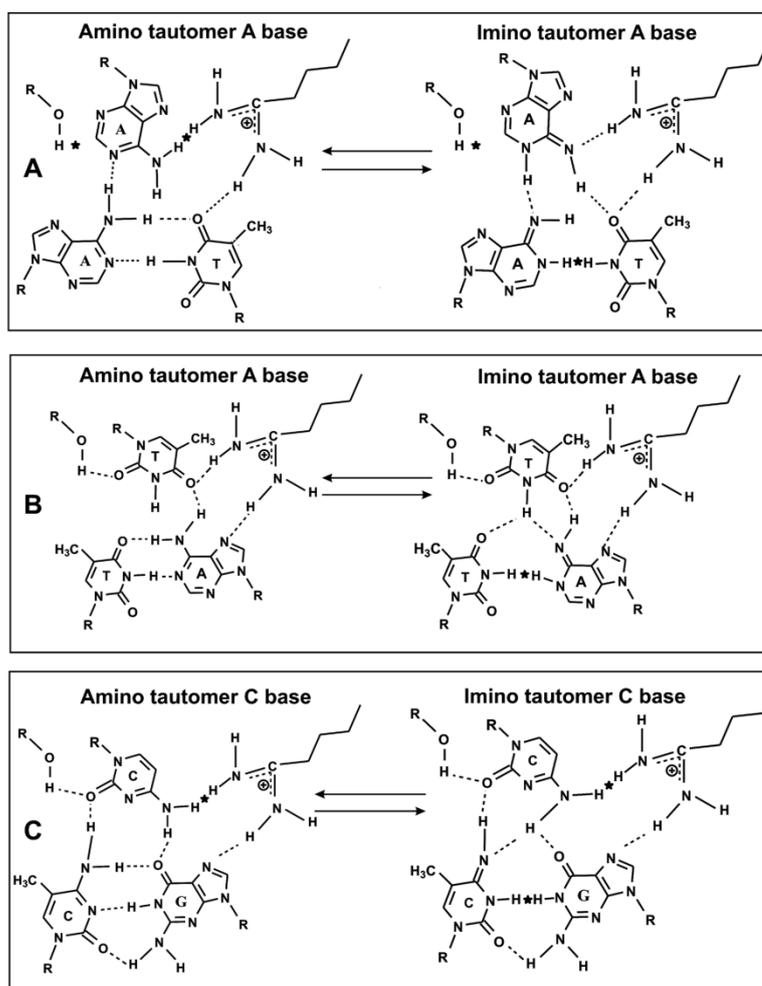
Figure 11. The hypothetical mechanism of homologous exchange catalyzed by RecA, proposed in [87,88]. The binding of ss DNA and the formation of a filament is accompanied by the formation of contacts of the Tyr and Arg RecA residues of the protein with one of the bases of DNA (**A**–**C**). Subsequent binding of the duplex by the RecA-ss DNA complex and the search for homology is provided due to the formation of specific hydrogen bonds during the formation of the R-DNA triplex. In this case, the RecA protein can catalyze the destruction of Watson–Crick bonds in the duplex due to a shift in the amino–imine equilibrium after the formation of contacts from the one of base with the amino acid residue of the DNA recognition center RecA. The destruction of hydrogen bonds in the original duplex leads to the formation of new ds DNA. The impossibility of hydrogen bonding is indicated using (★). Tyr and Arg are used in the model as the most appropriate amino acid residues for this role.

8. RNA–DNA Helicase

Unlike protein recognizing "spatial structure elements" of definite sequences, the enzymes responsible for reproducing genetic information such as RNA–DNA helicases are functioning independently of the NA primary, secondary, and tertiary structure peculiarities. The role of RNA and DNA helicases in replication is the denaturation of ds NAs leading to the appearance of their ss regions [91]. RNA rearrangement proceeds via protein-induced formation and subsequent resolution of RNA branch migration structures, whereby the latter step is dependent on ATP hydrolysis. Despite RNA helicase more specific in the reaction of the unwinding of ds RNA substrates, it can interact with ss RNAs and with ss and ds DNA ligands. We have studied p68 RNA helicase using the SILC approach [12,13].

Similar to the case of other enzymes, the minimal ligands of helicase were *ortho*phosphate (K_d = 0.1 M) and different dNMPs ($\sim 6.3 \times 10^{-2}$ M). The dependency of $-\text{Log}K_d$ upon the number of nucleotide units (n) is given in Figure 12A. The K_d values change for different d(pN)$_{1-10}$ can be described by the following progression

$$K_d[\text{d(pN)}_n] = [K_d(Pi) = 0.1\text{M}] \times [E = 1.61]^{-n} \times [h_C = 1.0.5]^{-c} \times [h_T = 1.26]^{-t} \times [h_G = 1.48)]^{-g} \times [h_A = 1.59]^{-a}.$$
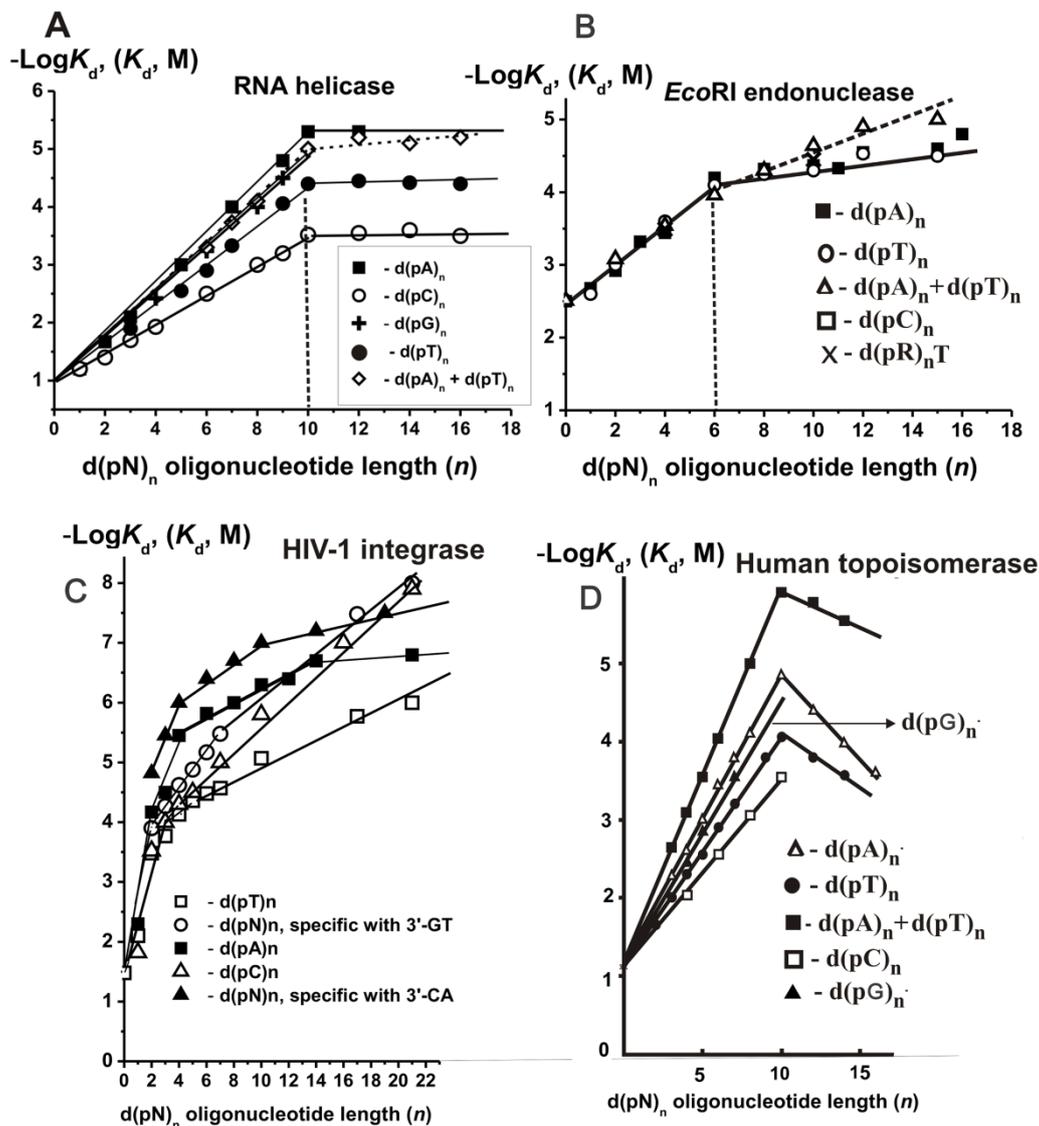
Figure 12. Dependencies of $-\text{Log}K_d$ for ss and ds d(pN)$_n$ deoxyribooligonucleotides versus their length (n) in the case of RNA–DNA helicase (**A**) [12,13], *Eco*R1 endonuclease (**B**) [92,93], HIV-1 integrase (**C**) [94,95], and human topoisomerase I (**D**) [96–98]. All types of enzymes and ODNs are shown on the panels; d(pR)$_n$pT oligomers contain only one T base.

The transition from ss to ds ODNs leads to an increase in the affinity $\sim$10-fold (Figure 12A).

9. *Eco*RI Endonuclease

Among the type II restriction endonucleases, the dimeric form of *Eco*RI (31 kDa) is the best-studied enzyme. Endonuclease *Eco*RI cleaves a bond between G and A bases of the hexanucleotide sequence in a duplex DNA shown below.

Points of cleavage:

5′-G ↓ AATT C-3′
3′-C TTAA ↑ G-5′

The enzyme crystal structure was determined [1,99,100]. It was shown that the interaction of *Eco*RI with ODNs containing a specific recognition sequence leads to conformational changes in both components. The structural features of the B-DNA double helix of DNA regions outside the enzyme are retained, but there are changes at a specific region associated with the enzyme. The enzyme also undergoes conformational changes of its quaternary structure. Certain amino acid residues of the protein form a so-called arm, which "twists" around the DNA, fixing it on the enzyme. Two symmetric "channels" of two subunits are located at a distance of 3.5 Å [1,99,100]. This seems to be the main factor responsible for specific DNA melting and bending in the central part of the DNA recognition site. In addition to electrostatic contacts with internucleoside phosphate groups of DNA, each of the *Eco*RI subunits forms hydrogen bonds with the bases of specific DNA. Thus, the guanidinium group Arg-200 forms two bonds with N-7 and O-6 atoms of G nucleotide. Glu-144 and Arg-145 interact, respectively, with N-6 and N-7 atoms of neighboring A nucleotide, forming also two hydrogen bonds. A total of 12 hydrogen bonds are formed between the specific ds DNA and the two subunits of protein, which are believed to provide specific recognition of the substrate [1,99,100]. These data are of a qualitative nature.

The quantitative data were found using the SILC approach (Supplementary Table S12), and the dependencies of $-\text{Log}K_d$ values on the length (n) of ss d(pN)$_n$ and their duplexes are presented in Figure 12B [92,93].

One of 6–7 subsites of *Eco*RI subunits has a high affinity to P$_i$ (K_d = 31 mM) and deoxyribosophosphate d(pR) (K_d = 4.6 mM) [92,93,101]. This suggests the formation of additional contacts of one *Eco*RI subsite with the sugar residue of the d(pR). The enzyme has an affinity for thymine (K_d = 450 mM) and thymidine (K_d = 65 mM). Additivity of $\Delta G°$ values is retained: K_d value for dTMP can be calculated as the product of K_d values for *ortho*phosphate and thymidine as well as for d(pR) and thymine (Supplementary Table S12). Consequently, one specific subsite of *Eco*RI recognizes dNMP (or one nucleotide unit of DNA) by additive interactions with all of its structural elements—sugar residue, phosphate group, and a base. The affinity of *Eco*RI to *ortho*phosphate is about 24-fold higher than that of non-charged triethylphosphate (K_d = 740 mM), which would efficiently interact with the enzyme only through the oxygen atom of the P = O group. Thus, the enzyme most probably interacts with both oxygen atoms (O = P-O$^-$) of the phosphate group of dNMPs as well as with one internucleoside phosphate group of specific DNA [93]. The slopes of the $-\text{Log}K_d$ dependencies on the number of nucleotide units (n) for *Eco*RI are nearly coincide for all homo- and hetero-d(pN)$_n$ (Figure 12B). The K_d value for oligomers without bases d[(pR)$_n$], where R is a chemically stable analog of deoxyribose, are comparable with K_d values for d(pT)$_n$ and d(pA)$_n$ of the same length (Supplementary Table S12) [93,101]. Thus, *Eco*RI interacts only with the sugar-phosphate backbone of unspecific DNAs.

Taking into account X-ray data [1,99,100], one can conclude that about 5–6 subsites of each subunit of *Eco*RI can form weak electrostatic contacts with internucleoside phosphate groups of d(pN)$_n$. *Eco*RI use as substrates specific ds DNAs. The second chain slightly decreases the *Eco*RI affinity to ds ODNs in comparison with ss d(pN)$_n$. Therefore, it seems that the interaction of two *Eco*RI subunits with DNA chains may significantly weaken the complementary interactions between two chains. The enzyme affinity for different hetero-d(pN)$_n$ completely or partially corresponding to the specific DNA sequence was determined (Supplementary Table S12) [92,93]. The K_d values for various short homo- and hetero-ODNs of the same length (n = 3–5) were comparable. The increase in the affinity ~10-fold was observed only for ODNs containing all six bases of the specific sequence. These results well correlate with X-ray data [1], according to which *Eco*RI should form 12 hydrogen bonds with the bases of a specific sequence of ds DNA. *Eco*RI is a very sequence-specific enzyme. As mentioned above, most repair enzymes' affinity for specific DNAs is usually higher than for unspecific ones only by one order of magnitude.

Interestingly, *Eco*RI affinity for specific 16-mer ds ODN is only 50–100-fold higher than for unspecific ds ODNs of the same length (Supplementary Table S12). Weak additive non-specific electrostatic contacts of *Eco*RI with phosphate groups of ODNs provide approximately five orders of magnitude in DNA affinity, whereas the contribution of specific interactions (12 hydrogen bonds) is no more than two orders of magnitude of a total ds ODN's affinity. Thus, as for other enzymes, the complex formation between *Eco*RI and specific ds DNA cannot provide specificity of its action. At the same time, *Eco*RI-dependent cleavage of $d(pT)_{15} \times d(pA)_{15}$ is 6–7 orders of magnitude slower than that for a 16-mer specific substrate [92,93,101]. Consequently, the *Eco*RI-dependent catalysis step is more sensitive to DNA structure than the step of the complex formation.

10. HIV-1 Integrase

Replication of HIV-1 DNA depends on integrating ds DNA copy of the virus genome into the human cell nuclear genome (reviewed in [102]). The retroviral integrase (IN) catalyzes the DNA integration; the recognition sequence is located at the ends of DNA long terminal repeats (LTRs). This LTR-specific sequence is critical for site-specific DNA cleavage and integration [103,104]. To insert both ends of the viral DNA into the host cell genome, HIV-1 integrase catalyzes two reactions—(1) 3'-processing reaction, when the two GT nucleotides from the 3'-ends of each linear viral DNA strands are removed, leaving the CA dinucleotide at 5-ends and (2) the strand transfer of DNA joining reaction in which the modified viral DNA ends are inserted into the host DNA [105,106].

The IN × DNA complex formation was analyzed by the SILC approach [94,95]. First, it was shown that IN recognizes free dNMPs (K_d = 5 mM) through interactions with all its structural elements (base (>50 mM), sugar (>50 mM), and phosphate (16 mM)), the phosphate makes the major contribution (Supplementary Table S13).

All −Log-dependencies for unspecific ss $d(pN)_n$ were biphasic (Figure 12C). The transition from dTMP, dCMP, or dAMP to the corresponding dinucleotides leads to the affinity increase by a factor of 2.5, 50, and 62, respectively. IN demonstrated two parts of the dependencies close to linear at n = 0–2 and n = 4–21, respectively (Figure 12C). The region between n = 2–4 has a transitional character. Monotonic decrease in K_d (n > 4; K_d = 1/F = 0.79–0.33 M)), values characterized the interaction of IN with each $d(pN)_n$ nucleotide unit. The affinity of IN for one of the free dNMPs was ~1000-fold higher than that of other nucleotide units at n > 4. The linear −LogK_d dependencies for ss ODNs provide evidence of $\Delta G°$ values close to an additive for the interaction of some of the 21 individual units of $d(pN)_{21}$ with the integrase. For HIV-1 integrase, the affinity for unspecific $d(pC)_n$, $d(pT)_n$, and $d(pA)_n$ did not follow the relative hydrophobicity of the bases. The affinities of $d(pA)_{21}$ and $d(pC)_{21}$ with minimal and maximal relative hydrophobicities are similar, while the affinity of $d(pT)_{21}$ is significantly lower (Figure 12C; Supplementary Tables S13 and S14). Interestingly, a faster increase in the affinity for $d(pC)_n$ as compared with $d(pT)_n$ resulted in a difference of ~2 orders of magnitude for their 21-mer ODNs (Figure 12C). The higher affinity of $d(pC)_n$ can be explained by a better adaptation of $d(pC)_n$ to a specific conformation in the viral DNA × integrase complex since of a very flexible structure of $d(pC)_n$, while $d(pT)_n$ usually adopts a rigid non-flexible structure [107]. Therefore, we have compared the affinity of IN for $d(pT)_n$ and oligothymidylates containing several $d(pC)$ links in different positions (Supplementary Table S14). The affinities such ODNs were comparable with that of $d(pC)_n$ but not for $d(pT)_n$ of the same length. Thus, the flexible structure of $d(pC)_n$ may be important for its productive changes during complex formation [94,95].

The 3'-end processing catalyzed by HIV-1 integrase was assayed with the specific substrate of IN ds GT-ODN$_{21}$ shown below.

Point of cleavage:

5'- G-T-G-T-G-G-A-A-A-A-T-C-T-C-T-<u>A-G-C-A</u>↓**G-T** -3'
3'- C-A-C-A-C-C-T-T-T-T-A-G-A-G-A-T-C-G-T C-A -5'

An analysis of the relative affinity of individual strands and of the specific duplex itself with and without 3'-terminal CT dinucleotide and different length sequences corresponding to both chains of the duplex was carried out (Figure 12C; Supplementary Table S15). It turned out that the contribution of GT terminal nucleotides to the total affinity of IN for ds GT-ODN$_{21}$ before the processing reaction is significantly higher than that of any other nucleotide of ds GT-ODN$_{21}$. However, after removing the 3'-GT-terminal dinucleotide, IN forms new strong contacts with four 3'-AGCA-terminal nucleotides of the first CA-ODN$_{19}$ chain of the shortened duplex. With this in mind, the recognition of a specific ds DNA by integrase was described using two thermodynamic models corresponding to the formation of complexes with ds GT-ODN$_{21}$ before (Figure 13) and after removing the terminal GT nucleotide (Figure 14). The affinities of IN for ss GT-ODN$_{21}$ (10 nM) and ss CA-ODN$_{19}$ (30 nM) were nearly the same and only respectively 7.7- and 2.5-fold lower than those of their duplexes (Supplementary Table S15). The complex formation provides only for a ~30-fold difference in the affinity between specific to non-specific DNAs. At the transition from unspecific to specific ODNs, the IN-dependent processing reaction's reaction rate is increased by more than 5–6 orders of magnitude. Thus, the catalytic step is significantly more sensitive to the specific viral DNA sequence than the IN × DNA complexation step.
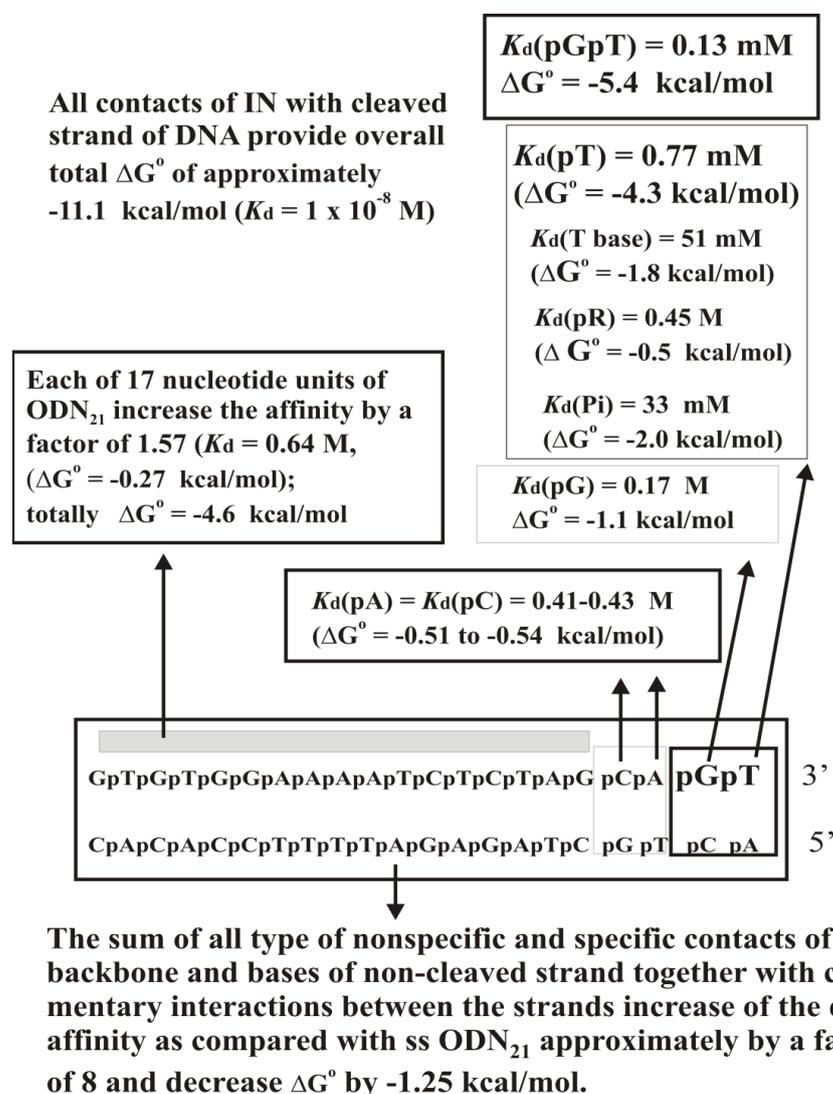
Figure 13. Thermodynamic model of HIV-1 IN interaction of with cognate ds GT-ODN$_{21}$ containing 3'-terminal-specific GT dinucleotide. $\Delta G°$ values characterizing different contacts between the enzyme and DNA strands are shown [95].

All contacts of IN with cleaved strand of DNA after removal of GT dinucleotide provide overall total ΔG^o of approximately -10.4 kcal/mol (Kd = 3 x 10^{-8} M).

K_d**(pCpA) = 15 μ M**
ΔG^o **= -6.7 kcal/mol**

Each of nine 5'-terminal nucleotides ODN19 provides ΔG^o = -0.08 kcal/mol; totally ΔG^o = -0.72 kcal/mol

K_d**(pA) = 1.5-2.0 mM**
(ΔG^o = -3.2 to -3. 9 kcal/mol)

Each of six nucleotides provides ΔG^o = -0.23 kcal/mol; totally ΔG^o = -0.88 kcal/mol

K_d**(pC) = 1.5-2.0 mM**
(ΔG^o = -2.7 to -3. 5 kcal/mol)

ΔG^o **= -0.88 kcal/mol)**

ΔG^o **= -0.75 kcal/mol**

GpTpGpTpGpGpApApApApTpCpTpCpTpA pG | **pCpA** | **3'**

CpApCpApCpCpTpTpTpTpApGpApGpApT pC | **pG** | **pT** | **pC** | **pA** | **5'**
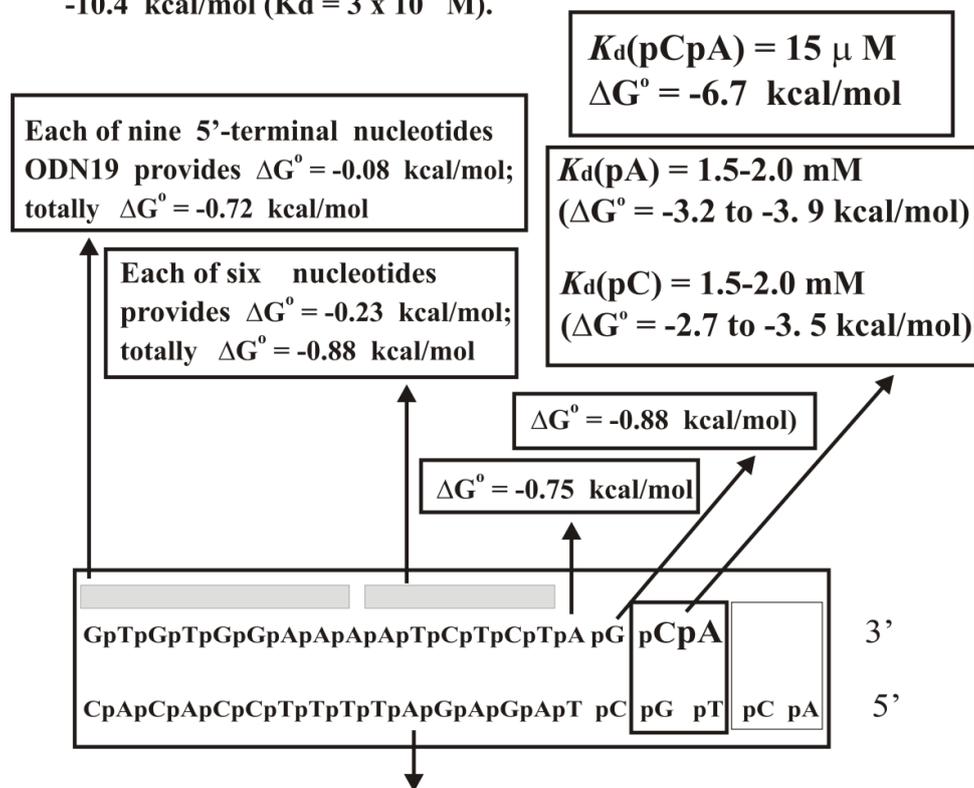
The sum of all type of nonspecific and specific contacts of IN with backbone and bases of non-cleaved strand together with complementary interactions between the strands increase of the duplex affinity as compared with ss ODN21 approximately by a factor of 7.5 and decrease ΔG^o by -1.20 kcal/mol.

Figure 14. Thermodynamic model of HIV-1 IN interaction of with cognate ds CA-specific ODN$_{19}$ after removal of 3'-terminal-specific GT dinucleotide from the cleavable chain of ds ODN$_{21}$ [95]. $\Delta G°$ values characterizing various contacts between the enzyme and DNA strands are shown.

11. Human DNA Topoisomerase I

Eukaryotic DNA topoisomerase I (Topo) changes the topological state both positively and negatively supercoiled (sc) DNA and plays an important role in key biological processes of the cell: replication, repair, transcription, integration, and recombination [108]. Topo consists of the N-, C-terminal, core domains, and a linker region [109]. The N-terminal domain is highly charged [109]; it is responsible for the localization of the enzyme in the nucleus [110] but is not important for the enzyme activity [109]. Topo is not absolutely sequence-specific for DNA sequences, but the below sequence is most preferred; the reaction rate in the case of this sequence is 3–4 orders of magnitude higher than for several other DNA sequences cleaved by the enzyme [111,112]. According to X-ray analysis, human Topo forms a large number of contacts with DNA internucleoside phosphate groups and specifically interacts with the T-base of the cleaved DNA strand:

Site of cleavage

```
-10- 9 -8 -7 -6 - 5 - 4 - 3 - 2 -1 │ 1  2  3  4  5 6  7 8 9
-T-A-A- A-A-G-A-C- T   T▼A-G-A-A-A-A-A-T-T
-A-T-T-T- T-C-T- G- A-A   T -C- T-T-T- T-T-A-A
          A                    B
```
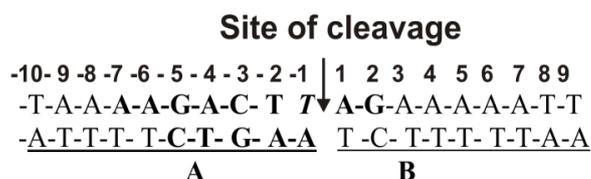
Recognition and transformation of the Topo substrate include the following stages: (a) binding of supercoiled DNA, (b) cleavage of one DNA strand, (c) shedding of supercoils, (d) ligation of the cleaved DNA strand, (e) changing the protein conformation for the following relaxation stages, and (f) dissociation of topo from relaxed DNA [113]. However, the X-ray data qualitative nature does not allow understanding which of these contacts make the main contribution to the total affinity of the enzyme for a specific DNA sequence.

First, we have analyzed the interaction of Topo with unspecific ODNs. *Ortho*phosphate (K_d = 0.38 M) was a minimal ligand of Topo (Supplementary Table S16) [96–98]. Figure 12D shows the $-$Log-dependencies K_d values on the length (n) of the used $d(pN)_n$. It was shown that there are 20 DNA base pairs within the protein globule of the enzyme [114]. However, the affinity ss and ds $d(pN)_n$ increases linearly up to n = 10 (Figure 12D). Data on the Topo's affinity for different short ONs corresponding to different parts of the specific DNA sequence are summarized in Supplementary Materials, Table S17. The factor F corresponding the increase in the affinity upon $d(pN)_n$ lengthening by one nucleotide unit were calculated as 1.71—$d(pC)_n$, 1.93—$d(pT)_n$, 2.12—$d(pG)_n$, 2.33—$d(pA)_n$, and 2.89—for duplex $d(pT)_n \times d(pA)_n$, and electrostatic factor E is equal to 1.67 (K_d = 0.6 M, $\Delta G°$ = -0.3 kcal/mol) [96,97]. Thus, the h_N coefficients of increase in Topo's affinity due to hydrophobic interactions with different bases were calculated (C—1.02, N—1.16, G—1.27, and A—1.40). It is obvious that the relative contribution of additive "electrostatic" interactions exceeds the contribution of "hydrophobic" contacts, and only in the case of $d(pA)_n$, they become comparable. Overall, the interaction of Topo with 10 units of ss DNAs is described by the geomantic progression common for all studied enzymes given above.

From the X-ray data [115], the internucleoside phosphate groups of specific ODNs form with Topo from 1 to 4-hydrogen bonds. The value $\Delta G°$ = -0.3 kcal/mol is characteristic of weak ion–dipole and dipole–dipole interactions [22]. Therefore, it is reasonable to consider that internucleoside phosphate groups of unspecific ODNs do not form hydrogen bonds with the enzyme and their weak additive interactions with Topo refer to ion-dipole and dipole–dipole interactions. The interactions of negatively charged phosphate groups of unspecific DNAs with a positively charged surface of Topo DNA-recognition site can also be attributed to interactions not of directly contacting groups, but of oppositely charged surfaces of these biopolymers.

Then, an analysis of the complexation of Topo with other different hetero-oligonucleotides of various structures and lengths corresponding to different parts of the specific DNA sequence was carried out (Supplementary Table S18). The transition from unspecific to various specific ODNs, the effectiveness of some contacts increases ~6–30 times, and the K_d values for several specific nucleotide units decreased to 0.1 to 0.02 M. Figure 15 demonstrates the efficiencies of contacts of several structural elements corresponding to the central part of specific DNA with Topo estimated by us using different ss and ds hetero-ODNs. We came to the conclusion that hydrogen bonds can be formed only after the first step of a specific change in the structure of DNA and Topo. In addition, the efficiency of most hydrogen bonds detected by X-ray analysis [115] may be relatively low. For example, if to assume that all 4 hydrogen bonds between the (-1-2) -phosphate group of AA residues of specific DNA and Topo (Figure 15) are comparable in their efficiency, then the formation of each of them will lead to an increase in the enzyme's affinity for the ligand only by ~2.1–2.7 times ($\Delta G°$ = $-0.4 \ldots -0.6$ kcal/mol). These values are quite comparable with those for weak additive contacts of Topo with internucleoside phosphate groups of ss ODNs. It should be noted that Topo stimulates several successive stages of changes in the conformation of specific DNA to its catalytically active state. For example, the affinity

of specific ODNs for Topo increases after their mixtures preincubation for 10–15 min (Supplementary Table S18) [116,117].
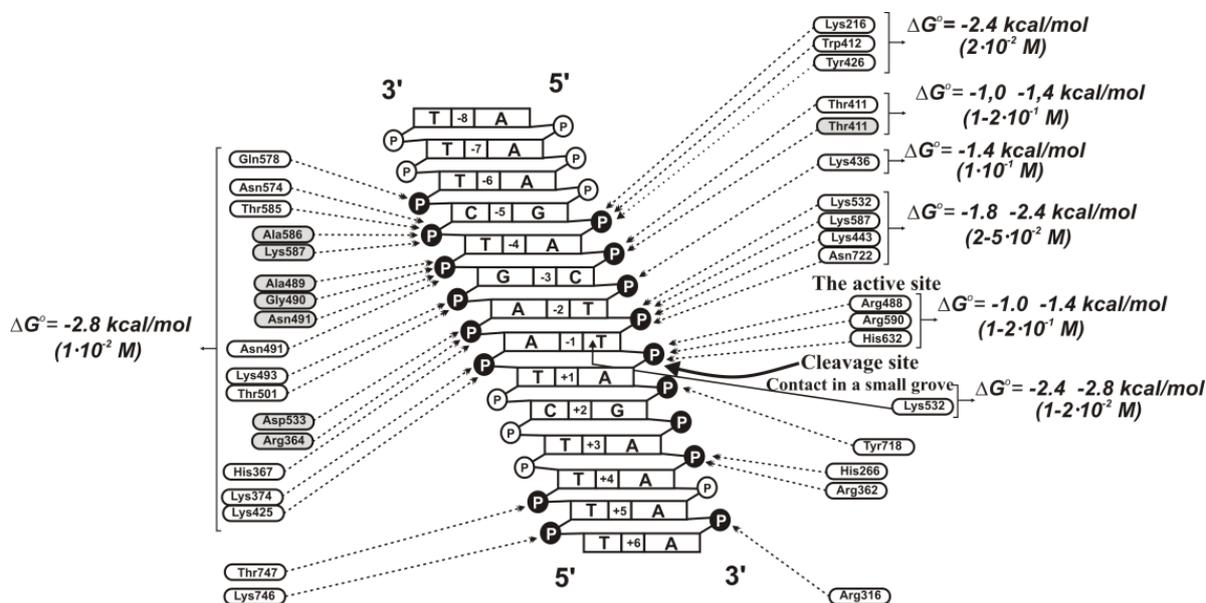
Figure 15. Schematic representation of the relative efficiency of different contacts formed by amino acid residues of human Topo I with internucleoside phosphate groups and the T-base of a specific sequence, revealed by X-ray analysis. The data on the right and left show approximate K_d values and $\Delta G°$ estimated using the SILC method [11–13].

Based on the analysis of the data obtained for a large number of specific and unspecific, short and long ss and ds ODNs, and the interaction of Topo with supercoiled DNA, the recognition of specific DNA by Topo can be described as follows.

At the first stage of the formation of the primary complex with unspecific ss and ds DNA, Topo binds these ligands mainly due to weak additive contacts, the affinity of enzymes for these DNAs is quite high; the K_d values, for ss d(pA)$_{10}$ and ds d(pA)$_{10}$ × d(pT)$_{10}$ are $8.5 × 10^{-5}$ and $1.0 × 10^{-5}$ M, respectively [96–98].

Thus, the contribution to the affinity of non-specific additive contacts of topo for 10 base pairs of the central part of the duplex can be estimated to be close to five orders of magnitude ($K_d = 10^{-5}$ M). Strengthening these contacts and/or the formation of additional contacts of the enzyme with 10 base pairs of specific d(pN)$_9$ leads to the increase in the affinity by about one order of magnitude ($K_d = 10^{-6}$ M). Additional interactions of Topo with AT-flanking sequences of extended specific 27-mer ds ODNs increase the affinity by approximately one order of magnitude ($K_d = 10^{-7}$ M). The increase in the affinity of various unspecific and specific ODNs after their preincubation with Topo provides a change of Topo-DNA complex, resulting in the affinity increase by an additional one order of magnitude ($K_d = 10^{-8}$ M). Interestingly, the affinity for 27-mer specific ds ODN ($K_d = 10^{-8}$ M) is comparable to that for relaxed plasmid DNA, while the difference in the affinity between supercoiled and relaxed DNA is approximately two orders of magnitude [116]. Thus, the contribution of specific interactions of Topo with the structural elements of supercoiled DNA ($K_d = 10^{-10}$ M) does not exceed two orders of magnitude. These interactions with supercoiled DNA may depend on both additional factors of the curvature of the Topo-binding site and on the structural features of supercoiled DNA. In addition, it cannot be ruled out that only supercoiled DNA is capable of forming any additional contacts with nucleotide units distant from the central part of the DNA specific duplex.

In general, the high affinity of Topo for supercoiled DNA is provided by the sum of all the factors listed above, which is 10^{-5} M × 10^{-1} × 10^{-1} × 10^{-1} × 10^{-2} = 10^{-10} M (Supplementary Table S19) [97,98]. In addition, as in the case of other enzymes described above, the high specificity of the action of eukaryotic topoisomerase I is provided mainly

by the catalysis stage. When passing from absolutely non-specific to specific DNA, the reaction rate increases by 5–6 orders of magnitude [96–98].

12. Human Lactoferrin

Lactoferrin (LF) is a protein with a single polypeptide chain (76–80 kDa) containing two lobes, each of which binds one Fe^{3+} ion and contains one glycan chain [118,119]. Many different functions were attributed to LF—immunomodulation and cell growth regulation, protection from iron-induced lipid peroxidation, transcriptional activation of specific DNA sequences, etc. (for review see [119]). For a long time, LF was considered as a protein that does not possess any enzymatic functions. First, it was shown that LF efficiently splits RNAs [120]. LF is protein interacting with DNA specific sequence and activating transcription [121]. In addition, LF is an enzyme possessing five different catalytic activities— DNase, RNase, ATPase, phosphatase, and amylase [122]. LF possesses two DNA-binding sites that interact with specific and non-specific ODNs [123]. Thus, LF can be considered as a very interesting polyfunctional protein–enzyme possessing several enzymatic functions.

Supplementary Table S20 demonstrates that the minimal ligands of the first LF DNA-recognizing site are *ortho*phosphate (K_d = 5 mM), deoxyribosephosphate (K_d = 3 mM), and different dNMPs (K_d = 0.56–1.6 mM). The affinity of LF for dCMP is 1.8–2.9-fold higher than for other dNMPs (Supplementary Table S20). The affinity for deoxyribose is very low ($K_d \approx 0.6$ M). The K_d characterizing of the protein affinity for the C base of dCMP is 0.19 M. Thus, one site of LF recognizes free dNMPs through interactions with all their structural elements (phosphate, sugar, and base); the phosphate making the major contribution to the affinity.

The linear $-$Log dependencies of K_d (Supplementary Table S21) for ss $d(pN)_n$ ($0 \leq n \leq 10$–11, $n = 0$ corresponds to *ortho*phosphate), demonstrate the additivity of $\Delta G°$ values for the interaction of LF with 10–11 individual nucleotide units of $d(pN)_n$ (Figure 16A).

Values of factor *F* and K_d (2.28 ± 0.02 M; K_d = 0.44 ± 0.004 M) for $d(pT)_n$ and $d(pA)_n$ are lower than that for $d(pC)_n$ (2.36 ± 0.03 M; K_d = 0.42 ± 0.005 M). Thus, the K_d values reflecting the LF affinity for different dNMPs (0.56–1.6 mM) are approximately 81–770-fold lower than the K_d (average ~0.43 M) characterizing the enzyme interaction with other from 9 to 10 nucleotide units of an extended ODNs.

In contrast to DNA polymerases, human UDG, APN1, and DNA topoisomerases but similarly to *E. coli* Fpg, human hOGG1, and *Eco*RI endonuclease, the affinity of LF for different $d(pN)_n$, does not remarkably depend on the relative hydrophobicity of ODNs bases (Figure 16A). The *E* factors (2.28–2.36) and K_d values (~0.42–0.44 M) characterizing the weak interactions of LF DNA-binding site with each nucleotide unit of different ss $d(pN)_n$ were nearly the same. Oligomers with and without bases have nearly the same affinity for LF (Supplementary Table S21). The affinity of $d(pT)_{10}$ (0.83 μM) was ~12-fold higher than that for ethylated $d[p(Et)T]_{10}$ (K_i = 10.0 μM) (Figure 16A; Supplementary Table S21) that points to an important role of internucleoside phosphates negative charges for LF complexation with DNA. Negatively charged internucleoside phosphate groups of unspecific ODNs could interact with LF as in the case of other various enzymes through dipolar electrostatic forces rather than electrostatic interactions between point charges.

Figure 16A shows that the minimal ODN exhibiting duplex properties toward LF is $d(pT)_{12} \times d(pA)_{12}$, for which the melting point is higher than the reaction temperature (25 °C). Thus, there is no effective stabilization of short ODN duplexes at their interaction with LF. The relative affinity of the first DNA chain is ~6–6.5 orders of magnitude, while the addition of the second chain results in an affinity increase of only 16–33-fold. Thus, similar to other enzymes, the contribution of the second DNA strand for LF is significantly less than that of the first one.
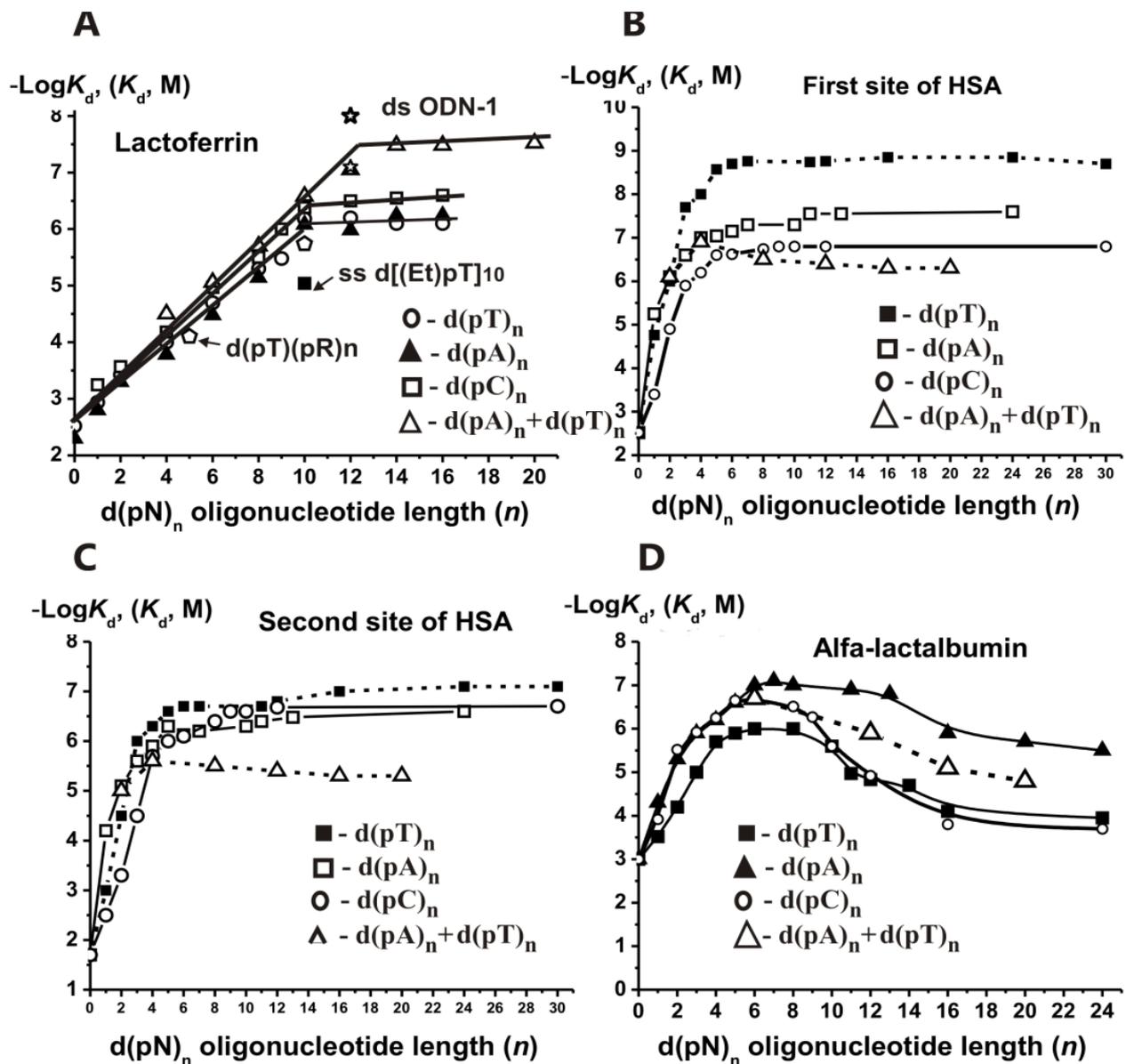
Figure 16. Dependencies of $-\text{Log}K_d$ for ss and ds $d(pN)_n$ deoxyribooligonucleotides versus their length (n) in the case of human lactoferrin (**A**) [124], first (**B**), and second (**C**) DNA-binding sites of human albumin [125], and human α-lactalbumin (**D**) [126]. All types of enzymes and ODNs are shown on the Panels; $d(pT)(pR)_n$ oligomers contain only one T base, $d[(Et)pT]_{10}$ is oligonucleotide with ethylated internucleoside phosphate groups.

LF interacts with the specific sequence of ds DNAs and activates transcription TAGAA-GATCAAA (ODN1; Figure 16A) [121]. The affinity of ss 12-mer ODN1 ($K_d = 8.0 \times 10^{-8}$ M) was 4.0–7.9-fold higher than different unspecific ss 12-mer ODNs ($K_d = (3.2–6.3) \times 10^{-7}$ M), while ds ODN1 ($K_d = 1.0 \times 10^{-8}$ M) demonstrated nine-fold higher affinity than ds homo-$d(pN)_{12}$ of the same length ($K_d = 9.0 \times 10^{-8}$ M). Overall, the types of LF interaction with unspecific and specific DNAs can be summarized using the thermodynamic model (Figure 17).
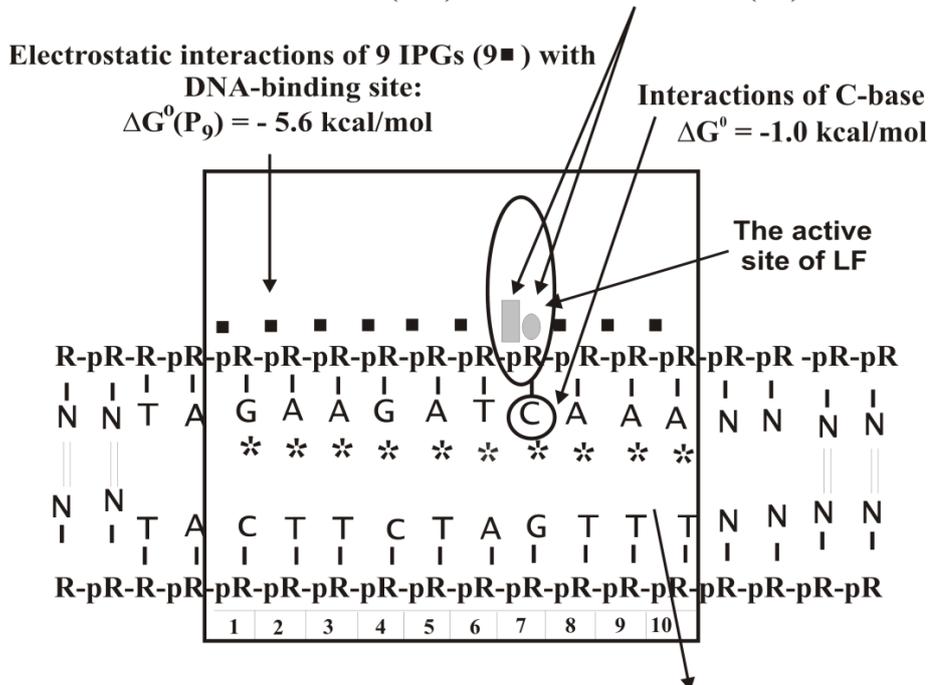
Figure 17. Thermodynamic model of human lactoferrin interaction with specific oligonucleotide. $\Delta G°$ values characterizing various contacts between the protein and DNA are shown [124].

13. Human Serum Albumin

Human serum albumin (HSA) is a major protein of human blood possessing very important biochemical and physiological functions [127]. HSA is a very universal protein interacting with many endogenous and exogenous compounds—fatty acids, metabolites, metal ions, pharmaceuticals, and many other blood components. Because of essential diversity in the ligands binding, HSA is characterized by significant structural diversity [127]. HSA has also long been regarded as a protein lacking any enzymatic activities. Later, it was shown that HSA has several catalytic activities (the hydrolysis of amides, esters, and organophosphorus compounds) [127], in addition to RNase and DNase ones [128,129]. However, there were no data on the mechanisms of DNA recognition by HSA.

It was shown that HSA possesses two DNA-binding sites. The relative contributions of $d(pN)_n$ to the total affinity for these two binding sites with lower and higher affinity for DNA were estimated (Supplementary Table S22). The minimal ligands of these two binding sites are *ortho*phosphate (K_d = 3.0 Mm and 20.0 mM in the case of two sites, respectively) [125]. The first binding site discriminates different dNMP, meaning the affinity decrease is in the order (Mole/Liter): dAMP (5.6×10^{-6}) > dTMP (1.7×10^{-5}) > dCMP (4.0×10^{-4}). The second site also discriminates various dNMPs with the order dAMP (6.3×10^{-5}) > dTMP (1.0×10^{-3} M) > dCMP (1.8×10^{-2}), but it possesses for

these mononucleotides significantly lower affinity than the first site, i.e., d(pC) (222-fold) > d(pT) (170-fold) > d(pA) (8.9-fold) (Supplementary Table S22). Maximal contribution to the total affinity of all $d(pN)_n$ in the case of both sites was observed for their one nucleotide and remarkably lower for three additional nucleotide units of $d(pN)_n$ ($n = 1–4$) with a significant decrease in the contribution at $n = 5–6$, and at $n \geq 7–8$ all dependencies reached plateaus (Figure 16B,C). All $d(pN)_{n>10}$ finally demonstrated high affinity for HSA first (1.4–150 nM) and second (80–2400 nM) sites. The thermodynamic parameters characterizing the contribution of different nucleotide units of all ss $d(pN)_{1–9}$ ($\Delta G°$) to their total affinity for HSA were estimated. Overall, two HSA DNA-binding sites recognize ss DNA by forming additive contacts with 6–8 nucleotide units of ss $d(pN)_n$, but in contrast to other enzymes, the relative contributions of $d(pN)_n$ nucleotide units ($\Delta G°$) at $n = 2–8$ are not the same (Figure 18) [125].
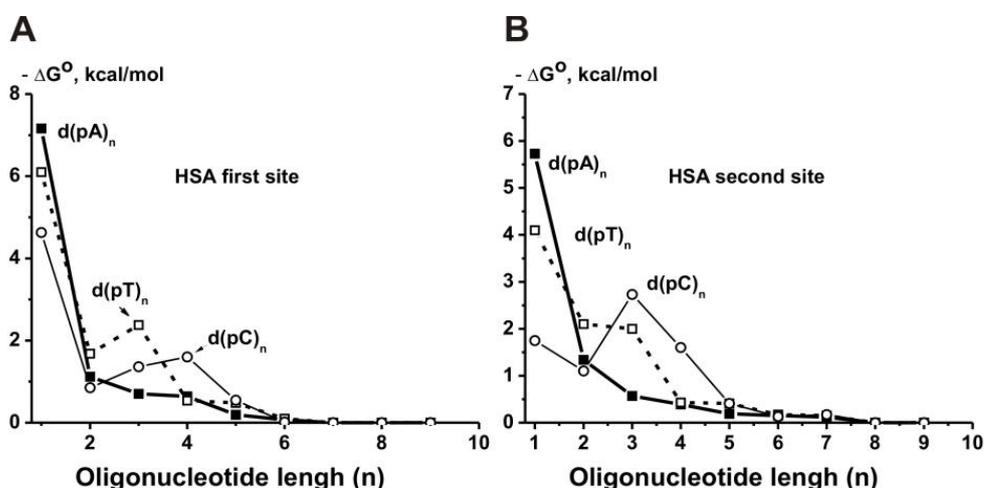
Figure 18. Thermodynamic values ($-\Delta G°$) characterizing a relative contribution of various mononucleotides of ss ODNs to the total affinity of these $d(pN)_8$ for the first (**A**) and the second (**B**) sites of human serum albumin (has). Different ss $d(pN)_n$ are shown on panels (**A,B**) [125].

The second strand's contribution to the affinity of ds DNA for all analyzed enzymes is usually much smaller than that for the first one. HSA demonstrated a significantly lower affinity for complementary ds $d(pN)_{n>10}$ with T_m melting points higher than the temperature used (Figure 16B,C; Supplementary Tables S22 and S23). However, in general, DNA recognition by albumin occurs in accordance with the same regularities as for other enzymes and proteins [125].

14. Human α-lactalbumin

Human milk α-lactalbumin (LA, 14.1 kDa) is the major protein of milk [130]. LA does not possess nuclease activities. It consists of a small β-fold and a large α-helical domain. It was showed that α-lactalbumin is a protein that can acquire various functions depending on its folding state [130]. LA interacts with DNA bound with histones in tumor cells and impaired the chromatin structure [131]. In the literature, there was no information on how LA recognizes DNA.

The SILC approach was used for the analysis of human LA (Supplementary Table S24) [126]. The LA DNA-binding site recognizes *ortho*phosphate and all dNMPs ($K_d = (5.0–43) \times 10^{-5}$ M) as minimal ligands [126]. For LA, an increase in the affinity ($-\text{Log}K_d$) was observed only for 5–6 nucleotides of ODNs (Figure 16D). Maximal contribution to LA's total affinity was observed for three-nucleotide units of all $d(pN)_n$ with a considerable decrease in the order 1 > 2 > 3 nucleotides. At $n = 4–6$, the affinity was remarkably lower, while at $n \geq 6–7$ all $-\log K_d$ dependencies upon n reached plateaus (Figure 16D). Double-stranded $d(pN)_n$ showed significantly lower affinity in comparison with ss $d(pN)_n$ for both sites

of LA (Figure 16D). The ΔG° values characterizing the specific contribution of $d(pN)_{1-6}$ every nucleotide unit to the total affinity for the DNA-binding site of LA were estimated (Figure 19). The spatial model of the LA–DNA complex was calculated. The protein sequence of LA has homology with five histones (H1, H2A, H2B, H3, and H4) involved in the chromatin in the interactions between themselves and their complex with DNA. It is proposed that such homology may be the main reason for the interaction of LA with DNA of chromatin, leading to a violation in its structure, as well as the proper binding of histones between themselves and with DNA [126].

The relative contribution of six oligonucleotide units to their common affinity for LA

			$-\Delta G^\circ$, kcal/mol					
$d(pA)_n$:	5.8	1.32	0.38	0.85	0.59	0.59	0.0	sum = 9.53
$d(pT)_n$	4.76	0.89	1.14	0.97	0.15	0.25	0.0	sum = 8.16
$d(pC)_n$	4.55	0.72	0.70	0.20	0.60	0.44	0.0	sum = 7.21

	1	2	3	4	5	6	7	
$r(pA)_n$	3.57	2.37	0.81	0.79	1.21	0.47	0.0	sum = 9.22
$r(pU)_n$	4.55	0.72	0.70	0.20	0.60	0.44	0.0	sum = 7.21
$r(pC)_n$	4.76	1.24	0.35	0.65	0.60	0.67	0.0	sum = 7.21

Figure 19. Thermodynamic model ($-\Delta G^\circ$ values) characterizing a relative contribution of different mononucleotides of ss $d(pN)_7$ and ribo$(pN)_7$ to the total affinity of these $(pN)_7$ for human milk α-lactalbumin (LA) [126]. Different ss $(pN)_7$ are shown on the Panel.

15. Human Antibodies against DNA

The generation of autoantibodies (auto-Abs) to self-antigens such as DNA and RNA usually proceeds not only in patients with autoimmune diseases but also in healthy donors [132–136]. During the last three decades, it has become clear that autoantibodies from the sera of patients with various autoimmune diseases can possess DNase and RNase activities [132]. Similar to artificial abzymes produced against chemically stable analogs of transition states of catalytic reactions, natural abzymes of human blood may be auto-Abs raised directly against enzyme substrates, which act similar to haptens and mimic transient states of catalytic reactions [132–136]. Moreover, the second anti-idiotypic auto-Abs against catalytic centers of enzymes can also possess catalytic activities [132–136]. There were no data in the literature on how anti-DNA autoantibodies can recognize DNA.

To analyze the molecular mechanism of DNA recognition, we used anti-DNA IgGs from sera of patients with multiple sclerosis [137]. The data for all used ss and ds $d(pN)_n$ are summarized in Supplementary Table S25. It shows that IgGs have two DNA-binding sites; the first site corresponds to IgG heavy, while the second corresponds to the light chain of auto-Abs.

The minimal ligand of the first (1.0×10^{-3} M) and the second (3.3×10^{-3} M) IgG sites is *ortho*phosphate (Figure 20) [137]. The affinity of the IgG first binding site for deoxyribose ($K_d \approx 0.14$ M) and deoxyribosephosphate (d(pR)), $K_d \approx 1.4 \times 10^{-4}$ M), and the second site for deoxyribose ($K_d \approx 0.13$ M) and deoxyribosephosphate (4.3×10^{-4} M) were estimated. The transition from deoxyribosephosphate to dCMP, dTMP, and dAMP in the case of the first binding site results in the increase in the affinity respectively by a factor of ~45.2-, 73.7-, and 222.2-fold; the K_d values 2.21×10^{-2} M, 1.36×10^{-2} M, and 4.5×10^{-3} M characterize the affinity of this site for C, T, and A bases of dNMPs. The affinity of the second site corresponding to the IgGs light chains for dCMP (2.8×10^{-4} M), dTMP (1.2×10^{-5} M), and dAMP (1.7×10^{-4} M) are respectively, 73.7-, 15.8-, and 269.8-fold lower than that in the case of the first DNA binding site corresponding to the heavy chains of IgGs (Supplementary Table S25). The affinity A, C, and T bases of dNMPs to the second binding site of IgGs was calculated as 0.4 M, 0.65 M, and 0.028 M. The affinity of A and C bases to the second site of IgGs is respectively 88.9- and 29.4-fold lower than that to the first site, while for T base only 2.0-fold. The affinity of the mononucleotides to the first IgG site is increased

with an enhancement in the relative C < T < A bases hydrophobicity: their phosphate groups make the major contributions (1.0×10^{-3} M) in comparison with deoxyribose (0.14 M), but contributions of the bases to some extent comparable (from 1.36×10^{-2} to 4.5×10^{-3} M). However, the efficiencies of the bases (C < A < T) interactions with the second IgG site do not correlate with their relative hydrophobicity. Overall, both DNA binding sites of antibodies can recognize free dNMPs through the interactions with all structural elements (phosphate, sugar, and base) with the maximal contributions to the total affinity of their phosphate groups. The affinity of different homo-$d(pN)_n$ to the first and the second IgG sites significantly depends on the length only for ODNs containing from one to four nucleotides (Figure 20A,B).
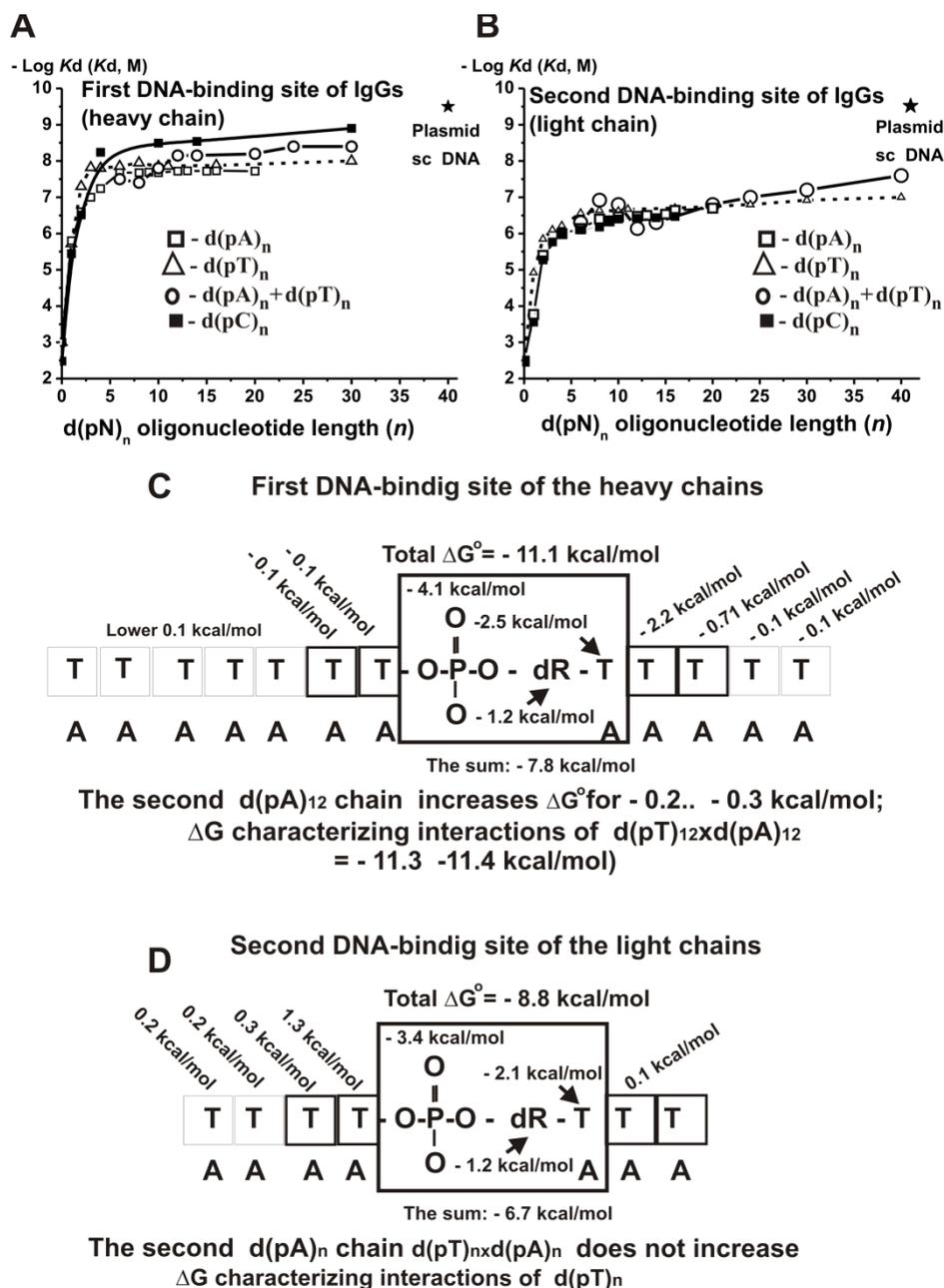
Figure 20. The $-$Log dependences of K_d values for $d(pN)_n$ corresponding to the first (**A**) and second (**B**) sites of anti-DNA IgGs upon ODNs length [137]. Different ss and ds ODNs are shown on the Panels (**A,B**). The average error in the complex formation from two independent experiments did not exceed 7–10%. Thermodynamic models of $d(pT)_n$ interaction with the heavy (**C**) and light (**D**) chains of IgGs. Values $\Delta G°$ are given on the Panels (**C,D**).

In addition, there was no observed correlation of the affinity with the relative hydrophobicity of ODNs bases at $n \geq 6–8$; all $d(pN)_n$, demonstrating comparable affinity (Supplementary Table S25).

The affinity of two IgG binding sites to several hetero-ODNs was also analyzed (Supplementary Table S24). The interaction of different hetero- $d(pN)_n$ with the first site of IgGs is, to some extent, similar to that for homo- ODNs (Supplementary Table S25), and it is to some extent comparable with the interaction of this site with $d(pA)_n$. However, several hetero-$d(pN)_n$ of different sequences but of the same length show a remarkable difference in the affinity to the second DNA-binding site (for example, two 4-mer and two 9-mer ODNs; Supplementary Table S26). On average, the affinity of hetero-ODNs of different lengths to the second DNA-binding site (Supplementary Table S26) is significantly lower; for example, 13-mer hetero-ODN demonstrates 116-fold lover affinity to the second in comparison with the first site.

We have estimated the K_d values for ds ODNs (Figure 20A,B); the minimal ligands containing two complementary chains and demonstrating remarkably higher affinity (duplex properties) toward the first site of IgGs is $d(pT)_{12} \times d(pA)_{12}$, for which the value of the melting point is higher than the used temperature (25 °C). Thus, no remarkable stabilization of short duplexes occurs by their interaction with the first site of IgGs. Ds and ss ODNs show the same affinity for the second site of IgGs. Taken together, the relative contribution of the $d(pN)_n$ due to interaction with the first DNA chain is high (10^{-8}–10^{-9} M). The addition of the second complementary chain leads to an average increase in the affinity only by factors of 2.4–2.6 in the case of the first and 1.0–1.3 for the second site. Thus, the contributions of the second strands are significantly less than that of the first one. This situation is very similar for all other analyzed enzymes, for which the difference between ss and ds ODNs varies two to five times (see above). Long supercoiled DNA interacts with both light and heavy chains with an affinity ~10-fold higher than that for short ODNs. The thermodynamic models were used to describe the interactions of IgGs light and heavy chains with DNA (Figure 20C,D).

16. Conclusions

It is known that structural elements of some low-molecular-weight ligands and substrates form very strong bonds with enzymes, 10^{-5}–10^{-6} M [11–21]. However, a very high affinity of enzymes and proteins for specific sequences or nucleotide units of extended DNAs can be dangerous for living organisms. For example, repair enzymes show a high affinity for extended DNA, 10^{-7}–10^{-9} M [11–21,25–27,54–61,69,86–88]. If such a high affinity is provided by the interaction of enzymes with only specific modified nucleotide units, after their removal from DNA, enzymes could be significantly inhibited by free mononucleotides. It is known that the rate of action of many enzymes is very high. Many studies of enzymes have shown that to ensure a high reaction rate, enzymes slide along DNA when searching for specific sequences and/or structural elements (specific sequences, single-stranded DNA fragments, modified nucleotides, breaks, etc.) [11–21,25–27,54–61,69,86–88]. Such sliding can only be achieved if the enzymes can bind to the DNA of any sequence with sufficiently high but with a close affinity for adjacent DNA fragments (isothermodynamic situation). Therefore, significant differences in the enzymes' affinity for unspecific and specific DNA can lead to a strong suppression in the sliding speed and, as a consequence, to the decrease in reaction rate.

To prove that the most important factors are providing specificity, we first analyzed many enzymes—DNA repair, DNA replication, topoisomerization, integration, recombination, and other ones using different physiochemical approaches, including the SILC method [11–21,25–27,54–61,69,86–88]. It was demonstrated that a high affinity of long DNA to all enzymes and proteins is provided by forming many weak additive hydrophobic and/or van der Waals interactions with all nucleotide links covered by enzyme and protein globules. Depending on the enzyme or protein, the sum of weak unspecific additive contacts provides 5–8 orders of the enzyme affinity for specific and unspecific

DNAs [11–21,25–27,54–61,69,86–88]. In contrast to enzymes interacting with small ligands, all interactions of enzymes with specific structural elements of long DNAs are usually weak. Their efficiencies are comparable with weak additive non-specific contacts. All specific interactions between DNA and enzymes usually provide approximately only one [11–16] and rarely about two orders of the affinity [59,92,93]. In addition, all the above-considered enzymes and proteins form complexes with DNAs and RNAs with an affinity of only about 1–1.5 orders of magnitude lower [25–27,53–59,69,86–88,92–98,101,124,125,137]. However, these DNA-dependent enzymes do not catalyze the reactions in the case of sc and ds RNA, corresponding to sequences of DNA substrates.

According to X-ray data, after DNA binding with different enzymes, there is a stage of very specific DNA and enzyme conformation adjustment (for review see [11–16]). Depending on the enzyme, there may be deformation of the DNA backbone, stretching or compression, partial or complete DNA melting, bending or kinking, eversion of nucleotides from the DNA helix, etc. [11–16]. These changes in DNA are very specific for each individual enzyme and protein. Enzyme-dependent specific changes in DNAs' conformation are required for effective adjustment of DNA-dependent enzymes and DNAs' reacting orbitals with an accuracy of about 10–15 degrees [22], which is possible only for specific DNAs but not RNAs [11–16]. The transition from unspecific to specific DNAs usually leads to the rise in the reaction rate (k_{cat}) by 5–8 orders of magnitude. Taken together, the stages of enzyme-dependent adjustment of DNA conformation and direct processing of the catalysis provide the high specificity of enzymes' actions [11–16]. Moreover, we have shown that DNA recognition by canonic enzymes and proteins with low activity and without enzymatic activities occur under the same regularities.

Supplementary Materials: They are available online at https://www.mdpi.com/1422-0067/22/3/1369/s1. 26 Supplementary Tables: Tables S1–S26 demonstrating affinity (K_d values) of different single- and double-stranded homo- and hetero-oligonucleotides to various analyzed enzymes and proteins.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

AP, apurinic/apyrimidinic; APE1, human apurinic/apyrimidinic endonuclease; Fpg, formamidopyrimidine or 8-oxoguanine DNA glycosylase; hOGG, human oxoguanine-DNA glycosylase; UDG, uracil DNA glycosylase; NA, nucleic acid; ODN, oligodeoxyribonucleotide; oxoG, 8-oxoguanine; SILC, stepwise increase in ligand complexity; Pi, *ortho*phosphate; ss and ds, single- and double-stranded, respectively; Topo, DNA topoisomerase I.

References

1. McClarin, J.A.; Frederick, C.A.; Wang, B.-C.; Greene, P.; Boyer, H.W.; Grable, J.; Rosenberg, J.M. Structure of the DNA-*Eco*RI endonuclease recognition complex at 3 A resolution. *Science* **1986**, *234*, 44–59. [CrossRef] [PubMed]
2. Freemont, P.S.; Lane, A.N.; Sanderson, M.R. Structural aspects of protein-DNA recognition. *Biochem. J.* **1991**, *278*, 1–23. [CrossRef] [PubMed]
3. Steitz, T.A. Structural studies of protein-nucleic acid interaction: The sources of sequence-specific binding. *Q. Rev. Biophys.* **1990**, *23*, 205–280. [CrossRef] [PubMed]
4. Jacobo-Molina, A.; Arnold, E. HIV reverse transcriptase structure-function relationships. *Biochemistry* **1991**, *30*, 6351–6356. [CrossRef] [PubMed]
5. Savva, R.; McAuley-Hecht, K.; Brown, T.; Pearl, L. The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature* **1995**, *373*, 487–489. [CrossRef] [PubMed]
6. Luscombe, N.M.; Austin, S.E.; Berman, H.M.; Thornton, J.M. An overview of the structures of protein-DNA complexes. *Genome Biol.* **2000**, *1*, 1–37. [CrossRef]
7. Schlundt, A.; Niessing, D.; Heissmeyer, V.; Sattler, M. RNA recognition by Roquin in posttranscriptional gene regulation. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 455–469. [CrossRef]

8. Schlundt, A.; Tants, J.N.; Sattler, M. Integrated structural biology to unravel molecular mechanisms of protein-RNA recognition. *Methods* **2017**, *118–119*, 119–136. [CrossRef]

9. Schlundt, A.; Buchner, S.; Janowski, R.; Heydenreich, T.; Heermann, R.; Lassak, J.; Geerlof, A.; Stehle, R.; Niessing, D.; Jung, K.; et al. Structure-Function analysis of the DNA-binding domain of a transmembrane transcriptional activator. *Sci. Rep.* **2017**, *7*, 1051. [CrossRef]

10. Schlundt, A.; Heinz, G.A.; Janowski, R.; Geerlof, A.; Stehle, R.; Heissmeyer, V.; Niessing, D.; Sattler, M. Structural basis for RNA recognition in roquin-mediated post-transcriptional gene regulation. *Nat. Struct. Mol. Biol.* **2014**, *21*, 671–678. [CrossRef]

11. Nevinsky, G.A. The essential role of weak interactions in enzyme recognition of long DNA and RNA molecules. *Mol. Biol.* **1995**, *29*, 16–37.

12. Bugreev, D.V.; Nevinsky, G.A. Potential of the method of stepwise increase in ligand complexity in studies of protein–nucleic acid interactions: Mechanisms of the functioning of some replication, repair, topoisomerization, and restriction enzymes. *Biokhimiya* **1999**, *64*, 291–305.

13. Nevinskii, G.A. The role of weak specific and nonspecific interactions in recognition and conversation by enzymes of long DNA. *Mol. Biol.* **2004**, *38*, 756–785.

14. Nevinsky, G.A. Structural, thermodynamic, and kinetic basis of DNA- and RNA-dependent enzymes functioning: Important role of weak nonspecific additive interactions between enzymes and long nucleic acids for their recognition and transformation. In *Protein Structures: Kaleidoscope of Structural Properties and Functions*; Uversky, V.N., Ed.; Research Signpost: Kerala, India, 2003; pp. 133–222.

15. Zharkov, D.O.; Mechetin, G.V.; Nevinsky, G.A. Uracil-DNA glycosylase: Structural, thermodynamic and kinetic aspects of lesion search and recognition. *Mutat. Res.* **2010**, *685*, 11–20. [CrossRef] [PubMed]

16. Nevinsky, G.A. Structural, thermodynamic, and kinetic basis for the activities of some nucleic acid repair enzymes. *J. Mol. Recognit.* **2011**, *24*, 656–677. [CrossRef]

17. Nevinskii, G.A.; Levina, A.S.; Podust, V.N.; Lavrik, O.I. Eukaryotic and prokaryotic DNA-polymerase. II. The role of internucleotide phosphate groups of a template in its binding with the enzyme. *Bioorg. Khimiia* **1987**, *13*, 58–68.

18. Volchkova, V.A.; Gorn, V.V.; Kolocheva, T.I.; Lavrik, O.I.; Levina, A.S. Klenow fragment of DNA-polymerase I from E. coli. III. The role of internucleotide phosphate groups of the matrix in its binding with the enzyme. *Bioorg. Khimiia* **1989**, *15*, 78–89.

19. Kolocheva, T.I.; Nevinsky, G.A.; Volchkova, V.A.; Levina, A.S.; Khomov, V.V.; Lavrik, O.I. DNA polymerase I (Klenow fragment): Role of the structure and length of a template in enzyme recognition. *FEBS Lett.* **1989**, *248*, 97–100. [CrossRef]

20. Lavrik, O.I.; Nevinskii, G.A. Protein-Nucleic acid interactions in reactions catalyzed by eukaryotic and prokaryotic DNA-polymerases. *Biokhimiia* **1989**, *54*, 757–764.

21. Kolocheva, T.I.; Nevinsky, G.A.; Levina, A.S.; Khomov, V.V.; Lavrik, O.I. The mechanism of recognition of templates by DNA polymerases from pro- and eukaryotes as revealed by affinity modification data. *J. Biomol. Struct. Dyn.* **1991**, *9*, 169–186. [CrossRef]

22. Fersht, A. *Enzyme Structure and Mechanism*, 2nd ed.; W.H. Freeman Co.: New York, NY, USA, 1985.

23. Nevinsky, G.A.; Veniaminova, A.G.; Levina, A.S.; Podust, V.N.; Lavrik, O.I.; Holler, E. Structure-function analysis of mononucleotides and short oligonucleotides in the priming of enzymatic DNA synthesis. *Biochemistry* **1990**, *29*, 1200–1207. [CrossRef] [PubMed]

24. Friedberg, E.C. *DNA Repair*; W.Y. Freeman and Co.: New York, NY, USA, 1985.

25. Vinogradova, N.L.; Nevinsky, G.A. Enzymes of direct, excision, and corrected systems of repair of higher and lower organisms and their biological roles. *Mol. Biol.* **2003**, *37*, 944–960.

26. Vinogradova, N.L.; Iamkovoi, V.I.; Tsvetkov, I.V.; Nevinskii, G.A. Interaction of uracil-DNA-glycosylase from human placenta with single-stranded deoxy- and oligoribonucleotides and their complexes. *Mol. Biol.* **1996**, *30*, 209–219.

27. Vinogradova, N.L.; Bulychev, N.V.; Maksakova, G.A.; Johnson, F.; Nevinsky, G.A. Uracil-DNA glycosylase: Interpretation of X-ray data in the light of kinetic and thermodynamic studies. *Mol. Biol.* **1998**, *32*, 489–499.

28. Mol, C.D.; Arvai, A.S.; Sanderson, R.J.; Sluppaug, G.; Kavli, B.; Krokan, H.; Mosbaugh, D.W.; Tainer, J.A. Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: Protein mimicry of DNA. *Cell* **1995**, *82*, 701–708. [CrossRef]

29. Stivers, J.T.; Pankiewicz, K.W.; Watanabe, K.A. Kinetic mechanism of damage site recognition and uracil flipping by *Escherichia coli* uracil DNA glycosylase. *Biochemistry* **1999**, *38*, 952–963. [CrossRef]

30. Ames, B.N. Dietary carcinogens and anticarcinogens. Oxygen radicals and degenerative diseases. *Science* **1983**, *221*, 1256–1264. [CrossRef]

31. Cathcart, R.; Schwiers, R.L.; Saul, E.; Ames, B.N. Thymine glycol and thymidine glycol in human and rat urine: A possible assay for oxidative DNA damage. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 5633–5637. [CrossRef]

32. Richter, C.; Park, J.W.; Ames, B.N. Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6465–6467. [CrossRef]

33. Ames, B.N. Endogenous DNA damage as related to cancer and aging. *Mutat. Res.* **1989**, *214*, 41–46. [CrossRef]

34. Cutler, R.G. Antioxidants and aging. *Am. J. Clin. Nutr.* **1991**, *53*, 373S–379S. [CrossRef] [PubMed]

35. Wood, M.L.; Dizdaroglu, M.; Gajewski, E.; Essigmann, J.M. Mechanistic studies of ionizing radiation and oxidative mutagenesis: Genetic effects of a single 8-hydroxyguanine (7-hydro-8-oxoguanine) residue inserted at a unique site in a viral genome. *Biochemistry* **1990**, *29*, 7024–7032. [CrossRef] [PubMed]

36. Shibutani, S.; Takeshita, M.; Grollman, A.P. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* **1991**, *349*, 431–434. [CrossRef] [PubMed]
37. Moriya, M.; Ou, C.; Bodepudi, V.; Johnson, F.; Takeshita, M.; Grollman, A.P. Site-Specific mutagenesis using a gapped duplex vector: A study of translesion synthesis past 8-oxodeoxyguanosine in *E. coli. Mutat. Res.* **1991**, *254*, 281–288. [CrossRef]
38. Cheng, K.C.; Cahill, D.S.; Kasai, H.; Nishimura, S.; Loeb, L.A. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G-T and A-C substitutions. *J. Biol. Chem.* **1992**, *267*, 166–172. [CrossRef]
39. Floyd, R.A.; Watson, J.J.; Harris, J.; West, M.; Wong, P.K. Formation of 8-hydroxydeoxyguanosine, hydroxyl free radical adduct of DNA in granulocytes exposed to the tumor promoter, tetradecanoylphorbolacetate. *Biochem. Biophys. Res. Commun.* **1986**, *137*, 841–846. [CrossRef]
40. Pryor, W.A. Why is the hydroxyl radical the only radical that commonly adds to DNA? Hypothesis: It has a rare combination of high electrophilicity, high thermochemical reactivity, and a mode of production that can occur near DNA. *Free Radic. Res. Commun.* **1988**, *4*, 219–223. [CrossRef]
41. Kasai, H.; Okada, Y.; Nishimura, S.; Rao, M.S.; Reddy, J.K. Formation of 8-hydroxydeoxyguanosine in liver DNA of rats following long-term exposure to a peroxisome proliferator. *Cancer Res.* **1989**, *49*, 2603–2605.
42. Fraga, C.G.; Shigenaga, M.K.; Park, J.W.; Degan, P.; Ames, B.N. Oxidative damage to DNA during aging: 8-hydroxy-2′-deoxyguanosine in rat organ DNA and urine. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 4533–4537. [CrossRef]
43. Roy, D.; Floyd, R.A.; Liehr, J.G. Elevated 8-hydroxydeoxyguanosine levels in DNA of diethylstilbestrol-treated Syrian hamsters: Covalent DNA damage by free radicals generated by redox cycling of diethylstilbestrol. *Cancer Res.* **1991**, *51*, 3882–3885.
44. Weitzman, S.A.; Turk, P.W.; Milkowski, D.H.; Kozlowski, K. Free radical adducts induce alterations in DNA cytosine methylation. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 1261–1264. [CrossRef] [PubMed]
45. Bailly, V.; Verly, W.G.; O'Connor, T.R.; Laval, J. Mechanism of DNA strand nicking at apurinic/apyrimidinic sites by *Escherichia coli* [formamidopyrimidine]DNA glycosylase. *Biochem. J.* **1989**, *262*, 581–589. [CrossRef] [PubMed]
46. Graves, R.J.; Felzenszwalb, I.; Laval, J.; O'Connor, T.R. Excision of 5′-terminal deoxyribose phosphate from damaged DNA is catalyzed by the Fpg protein of *Escherichia coli. J. Biol. Chem.* **1992**, *267*, 14429–14435. [CrossRef]
47. Nash, H.M.; Bruner, S.D.; Schärer, O.D.; Kawate, T.; Addona, T.A.; Spooner, E.; Lane, W.S.; Verdine, G.L. Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr. Biol.* **1996**, *6*, 968–980. [CrossRef]
48. Rosenquist, T.A.; Zharkov, D.O.; Grollman, A.P. Cloning and characterization of a mammalian 8-oxoguanine DNA glycosylase. *Proc. Natl Acad. Sci. USA* **1997**, *94*, 7429–7434. [CrossRef] [PubMed]
49. Bruner, S.D.; Norman, P.G.; Verdine, G.L. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature* **2000**, *403*, 859–866. [CrossRef]
50. Bjørås, M.; Seeberg, E.; Luna, L.; Pearl, L.H.; Barrett, T.E. Reciprocal 'flipping' underlies substrate recognition and catalytic activation by the human 8-oxo-guanine DNA glycosylase. *J. Mol. Biol.* **2002**, *317*, 171–177. [CrossRef]
51. Yang, W.; Karplus, M.; Verdine, G.L. Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* **2005**, *434*, 612–618.
52. Gilboa, R.; Zharkov, D.O.; Golan, G.; Fernandes, A.S.; Gerchman, S.E.; Matz, E.; Kycia, J.H.; Grollman, A.P.; Shoham, G. Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *J. Biol. Chem.* **2002**, *277*, 19811–19816. [CrossRef]
53. Zharkov, D.O.; Rosenquist, T.A.; Gerchman, S.E.; Grollman, A.P. Substrate specificity and reaction mechanism of murine 8-oxoguanine-DNA glycosylase. *J. Biol. Chem.* **2000**, *275*, 28607–28617. [CrossRef]
54. Ishchenko, A.A.; Bulychev, N.V.; Maksakova, G.A.; Johnson, F.; Nevinsky, G.A. Recognition and conversion of single- and double-stranded oligonucleotide substrates by 8-oxoguanine-DNA glycosylase from Escherichia coli. *Biochemistry* **1997**, *62*, 204–211. [PubMed]
55. Ishchenko, A.A.; Bulychev, N.V.; Maksakova, G.A.; Johnson, F.; Nevinsky, G.A. Interaction of *Escherichia coli* 8-oxoguanine-DNA glycosylase with single-stranded oligodeoxyribonucleotides and their complexes. *Mol. Biol.* **1998**, *32*, 454–462.
56. Ishchenko, A.A.; Koval, V.V.; Fedorova, O.S.; Douglas, K.T.; Nevinsky, G.A. Structural requirements of double and single stranded DNA substrates and inhibitors, including a photoaffinity label, of Fpg protein from *E. coli. J. Biomol. Struct. Dyn.* **1999**, *17*, 301–310. [CrossRef] [PubMed]
57. Ishchenko, A.A.; Vasilenko, N.L.; Sinitsina, O.I.; Yamkovoy, V.I.; Fedorova, O.S.; Douglas, K.T.; Nevinsky, G.A. Thermodinamic, kinetic, and structural basis for recognition and repair of 8-oxoguanine in DNA by Fpg protein from *E. coli. Biochemistry* **2002**, *41*, 7540–7548. [CrossRef] [PubMed]
58. Zharkov, D.O.; Ishchenko, A.A.; Douglas, K.T.; Nevinsky, G.A. Recognition of damaged DNA by Escherichia coli Fpg protein: Insights from structural and kinetic data. *Mutat. Res.* **2003**, *531*, 141–156. [CrossRef]
59. Kirpota, O.O.; Endutkin, A.V.; Ponomarenko, M.P.; Ponomarenko, P.M.; Zharkov, D.O.; Nevinsky, G.A. Thermodynamic and kinetic basis for recognition and repair of 8-oxoguanine in DNA by human 8-oxoguanine-DNA glycosylase. *Nucleic Acids Res.* **2011**, *39*, 4836–4850. [CrossRef]
60. Kuznetsov, N.A.; Koval, V.V.; Zharkov, D.O.; Nevinsky, G.A.; Douglas, K.T.; Fedorova, O.S. Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase. *Nucleic Acids Res.* **2005**, *33*, 3919–3931. [CrossRef]
61. Kuznetsov, N.A.; Koval, V.V.; Nevinsky, G.A.; Douglas, K.T.; Zharkov, D.O.; Fedorova, O.S. Kinetic conformational analysis of human 8-oxoguanine-DNA glycosylase. *J. Biol. Chem.* **2007**, *282*, 1029–1038. [CrossRef]

62. Mol, C.D.; Hosfield, D.J.; Tainer, J.A. Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: The 3' ends justify the means. *Mutat. Res.* **2000**, *460*, 211–229. [CrossRef]

63. Mol, C.D.; Izumi, T.; Mitra, S.; Tainer, J.A. DNA-Bound structures and mutants reveal abasic DNA binding by APE1 DNA repair and coordination. *Nature* **2000**, *403*, 451–456. [CrossRef]

64. Wilson, D.M., 3rd; Barsky, D. The major human abasic endonuclease: Formation, consequences and repair of abasic lesions in DNA. *Mutat. Res. DNA Repair* **2001**, *485*, 283–307. [CrossRef]

65. Eisen, J.A.; Hanawalt, P.C. A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat. Res.* **1999**, *435*, 171–213. [CrossRef]

66. Aravind, L.; Walker, D.R.; Koonin, E.V. Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res.* **1999**, *27*, 1223–1242. [CrossRef]

67. Gorman, M.A.; Morera, S.; Rothwell, D.G.; de La Fortelle, E.; Mol, C.D.; Tainer, J.A.; Hickson, I.D.; Freemont, P.S. The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. *EMBO J.* **1997**, *16*, 6548–6558. [CrossRef] [PubMed]

68. Beernink, P.T.; Segelke, B.W.; Hadi, M.Z.; Erzberger, J.P.; Wilson, D.M., III; Rupp, B. Two divalent metal ions in the active site of a new crystal form of human apurinic/apyrimidinic endonuclease, Ape1: Implications for the catalytic mechanism. *J. Mol. Biol.* **2001**, *307*, 1023–1034. [CrossRef] [PubMed]

69. Beloglazova, N.G.; Kirpota, O.O.; Starostin, K.V.; Ishchenko, A.A.; Yamkovoy, V.I.; Zharkov, D.O.; Douglas, K.T.; Nevinsky, G.A. Thermodynamic, kinetic and structural basis for recognition and repair of abasic sites in DNA by apurinic/apyrimidinic endonuclease from human placenta. *Nucleic Acids Res.* **2004**, *32*, 5134–5146. [CrossRef]

70. Lusetti, S.L.; Cox, M.M. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu. Rev. Biochem.* **2002**, *71*, 71–100. [CrossRef]

71. Amaratunga, M.; Benight, A.S. DNA sequence dependence of ATP hydrolysis by RecA protein. *Biochem. Biophys. Res. Commun.* **1988**, *157*, 127–133. [CrossRef]

72. Cazenave, C.; Chabbert, M.; Toulme, J.J.; Helene, C. Absorption and fluorescence studies of the binding of the recA gene product from *E. coli* to single-stranded and double-stranded DNA. Ionic strength dependence. *Biochim. Biophys. Acta* **1984**, *781*, 7–13. [CrossRef]

73. McEntee, K.; Weinstock, G.M.; Lehman, I.R. Binding of the recA protein of *Escherichia coli* to single- and double-stranded DNA. *J. Biol. Chem.* **1981**, *256*, 8835–8844. [CrossRef]

74. Tracy, R.B.; Kowalczykowski, S.C. In Vitro selection of preferred DNA pairing sequences by the *Escherichia coli* RecA protein. *Genes Dev.* **1996**, *10*, 1890–1903. [CrossRef] [PubMed]

75. Yu, X.; Egelman, E.H. Structural data suggest that the active and inactive forms of the RecA filament are not simply interconvertible. *J. Mol. Biol.* **1992**, *227*, 334–346. [CrossRef]

76. Register, J.C., 3rd; Griffith, J. The direction of RecA protein assembly onto single strand DNA is the same as the direction of strand assimilation during strand exchange. *J. Biol. Chem.* **1985**, *260*, 12308–12312. [CrossRef]

77. Bork, J.M.; Cox, M.M.; Inman, R.B. RecA protein filaments disassemble in the 5' to 3' direction on single-stranded DNA. *J. Biol. Chem.* **2001**, *276*, 45740–45743. [CrossRef] [PubMed]

78. Stasiak, A.; Di Capua, E. The helicity of DNA in complexes with recA protein. *Nature* **1982**, *299*, 185–186. [CrossRef]

79. Pugh, B.F.; Schutte, B.C.; Cox, M.M. Extent of duplex DNA underwinding induced by RecA protein binding in the presence of ATP. *J. Mol. Biol.* **1989**, *205*, 487–492. [CrossRef]

80. Heuser, J.; Griffith, J. Visualization of RecA protein and its complexes with DNA by quick-freeze/deep-etch electron microscopy. *J. Mol. Biol.* **1989**, *210*, 473–484. [CrossRef]

81. Mazin, A.V.; Kowalczykowski, S.C. The function of the secondary DNA-binding site of RecA protein during DNA strand exchange. *EMBO J.* **1998**, *17*, 1161–1168. [CrossRef]

82. Zhurkin, V.B.; Raghunathan, G.; Ulyanov, N.B.; Camerini-Otero, R.D.; Jernigan, R.L. A parallel DNA triplex as a model for the intermediate in homologous recombination. *J. Mol. Biol.* **1994**, *239*, 181–200. [CrossRef]

83. Kim, M.G.; Zhurkin, V.B.; Jernigan, R.L.; Camerini-Otero, R.D. Probing the structure of a putative intermediate in homologous recombination: The third strand in the parallel DNA triplex is in contact with the major groove of the duplex. *J. Mol. Biol.* **1995**, *247*, 874–889. [CrossRef]

84. Zhou, X.; Adzuma, K. DNA strand exchange mediated by the *Escherichia coli* RecA protein initiates in the minor groove of double-stranded DNA. *Biochemistry* **1997**, *36*, 4650–4661. [CrossRef] [PubMed]

85. Howard-Flanders, P.; West, S.C.; Stasiak, A. Role of RecA protein spiral filaments in genetic recombination. *Nature* **1984**, *309*, 215–219. [CrossRef] [PubMed]

86. Bugreeva, I.P.; Bugreev, D.V.; Nevinsky, G.A. Formation of nucleoprotein RecA filament on single-stranded DNA: Analysis by stepwise increase in ligand complexity. *FEBS J.* **2005**, *272*, 2734–2745. [CrossRef] [PubMed]

87. Bugreeva, I.P.; Bugreev, D.V.; Nevinskii, G.A. Physico-Chemical basis of RecA nucleo-protein filament formation on single-stranded DNA. *Mol. Biol.* **2005**, *39*, 984–998. [CrossRef]

88. Bugreeva, I.P.; Bugreev, D.V.; Nevinskii, G.A. Interaction of single-stranded DNA with the second DNA-binding site of RecA nucleoprotein filament. *Mol. Biol.* **2007**, *41*, 524–534. [CrossRef]

89. Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, NY, USA, 1984.

90. Folta-Stogniew, E.; O'Malley, S.; Gupta, R.; Anderson, K.S.; Radding, C.M. Exchange of DNA base pairs that coincides with recognition of homology promoted by *E. coli* RecA protein. *Mol. Cell* **2004**, *15*, 965–975. [CrossRef]

91. Wessel, F.; Schweizer, J.; Stahl, H. Simian virus 40 T-antigen DNA helicase is a hexamer which forms a binary complex during bidirectional unwinding from the viral origin of DNA replication. *Virology* **1992**, *66*, 804. [CrossRef]

92. Kolocheva, T.I.; Demidov, S.A.; Maksakova, G.A.; Nevinskii, G.A. Interaction of endonuclease *EcoRI* with short specific and nonspecific oligonucleotides. *Mol. Biol.* **1998**, *32*, 1025–1033.

93. Kolocheva, T.I.; Maksakova, G.A.; Bugreev, D.; Nevinsky, G.A. Interaction of endonuclease *Eco*RI with short specific and nonspecific oligonucleotides. *IUBMB Life* **2001**, *51*, 189–195. [CrossRef]

94. Caumont, A.; Jamieson, G.; Richard de Soultrait, V.; Parissi, V.; Fournier, M.; Zakharova, O.D.; Bayandin, R.; Litvak, S.; Tarrago-Litvak, L.; Nevinsky, G.A. High affinity interaction of HIV-1 integrase with specific and non-specific single-stranded short oligonucleotides. *FEBS Lett.* **1999**, *455*, 154–158. [CrossRef]

95. Bugreev, D.V.; Baranova, S.; Zakharova, O.D.; Parissi, V.; Desjobert, C.; Sottofattori, E.; Balbi, A.; Litvak, S.; Tarrago-Litvak, L.; Nevinsky, G.A. Dynamic, thermodynamic, and kinetic basis for recognition and transformation of DNA by human immunodeficiency virus type 1 integrase. *Biochemistry* **2003**, *42*, 9235–9247. [CrossRef] [PubMed]

96. Bugreev, D.V.; Buneva, V.N.; Sinitsina, O.I.; Nevinsky, G.A. Mechanism of recognition of supercoiled DNA by eukaryotic DNA topoisomerases I: Interactions of enzymes with nonspecific oligonucleotides. *Bioorg. Khimiia* **2003**, *29*, 163–174.

97. Bugreev, D.V.; Buneva, V.N.; Sinitsina, O.I.; Nevinskii, G.A. The mechanism of supercoiled DNA recognition by eukaryotic type I topoisomerases. II. A comparison of the enzyme interaction with specific and nonspecific oligonucleotides. *Bioorg. Khimiia* **2003**, *29*, 277–289.

98. Bugreev, D.V.; Buneva, V.N.; Nevinsky, G.A. Mechanism of supercoiled DNA cleavage by human DNA topoisomerase I: Effect of ligand structure on the catalytic stage of the reaction. *Mol. Biol.* **2003**, *37*, 325–339. [CrossRef]

99. Frederick, C.A.; Grable, J.; Melia, M.; Samudzi, C.; Jen-Jacobson, L.; Wang, B.-C.; Greene, P.; Boyer, H.W.; Rosenberg, J.M. Kinked DNA in crystalline complex with *Eco*RI endonuclease. *Nature* **1984**, *309*, 327–331. [CrossRef] [PubMed]

100. Kim, Y.; Grable, J.C.; Love, R. Refinement of *Eco RI* endonuclease crystal structure: A revised protein chain tracing. *Science* **1990**, *249*, 1307–1309. [CrossRef]

101. 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.; Shabarova, Z.A. Modified substrates as probes for studing uracil-DNA glycosylase. *Gene* **1995**, *157*, 167–171. [CrossRef]

102. Asante-Appiah, E.; Skalka, A.M. HIV-1 integrase: Structural organization, conformational changes, and catalysis. *Adv. Virus Res.* **1999**, *52*, 351–369.

103. Katzman, M.; Katz, R.A.; Skalka, A.M.; Leis, J. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. *J. Virol.* **1989**, *63*, 5319–5327. [CrossRef]

104. Roth, M.J.; Schwartzberg, P.L.; Goff, S.P. Structure of the termini of DNA intermediates in the integration of retroviral DNA: Dependence on IN function and terminal DNA sequence. *Cell* **1989**, *58*, 47–54. [CrossRef]

105. Craigie, R.; Fujiwara, T.; Bushman, F. The integrase protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. *Cell* **1990**, *62*, 829–837. [CrossRef]

106. Katz, R.A.; Merkel, G.; Kulkosky, J.; Leis, J.; Skalka, A.M. The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell* **1990**, *63*, 87–95. [CrossRef]

107. Neidle, S. *Principles of Nucleic Acid Structure*, 1st ed.; Academic Press: Cambridge, MA, USA, 2007.

108. Wang, J.C. DNA topoisomerases. *Annu. Rev. Biochem.* **1996**, *65*, 635–692. [CrossRef] [PubMed]

109. Stewart, L.; Ireton, G.C.; Champoux, J.J. The domain organization of human topoisomerase. *J. Biol. Chem.* **1996**, *271*, 7602–7608. [CrossRef] [PubMed]

110. Bharti, A.K.; Olson, M.O.; Kufe, D.W.; Rubin, E.H. Identification of a nucleolin binding site in human topoisomerase I. *J. Biol. Chem.* **1996**, *271*, 1993–1997. [CrossRef]

111. Thomsen, B.; Mollerup, S.; Bonven, B.J.; Frank, R.; Blocker, H.; Nielsen, O.F.; Westergaard, O. Sequence specificity of DNA topoisomerase I in the presence and absence of camptothecin. *EMBO J.* **1987**, *6*, 1817–1823. [CrossRef] [PubMed]

112. Busk, H.; Thomsen, B.; Bonven, B.J.; Kjeldsen, E.; Nielsen, O.F.; Westergaard, O. Preferential relaxation of supercoiled DNA containing a hexadecameric recognition sequence for topoisomerase I. *Nature* **1987**, *327*, 638–640. [CrossRef]

113. Christiansen, K.; Westergaard, O. Characterization of intra- and intermolecular DNA ligation mediated by eukaryotic topoisomerase I. Role of bipartite DNA interaction in the ligation process. *J. Biol. Chem.* **1994**, *269*, 721–729. [CrossRef]

114. Stevsner, T.; Mortensen, U.H.; Westergaard, O.; Bonven, B.J. Interactions between eukaryotic DNA topoisomerase I and a specific binding sequence. *J. Biol. Chem.* **1989**, *264*, 10110–10113. [CrossRef]

115. Stewart, L.; Redinbo, M.R.; Qiu, X.; Hol, W.G.; Champoux, J.J. A model for the mechanism of human topoisomerase I. *Science* **1998**, *279*, 1534–1541. [CrossRef]

116. Nevinsky, G.A.; Bugreev, D.V.; Buneva, V.N.; Yasui, Y.; Nishizava, M.; Andoh, T. High affinity interaction of mammalian DNA topoisomerase I with short single- and double-stranded oligonucleotides. *FEBS Lett.* **1995**, *368*, 97–100. [CrossRef]

117. Bugreev, D.V.; Vasyutina, E.L.; Buneva, V.N.; Yasui, Y.; Nishizava, M.; Andoh, T.; Nevinsky, G.A. High affinity interaction of mouse DNA topoisomerase I with di- and trinucleotides corresponding to specific sequences of supercoiled DNA cleaved chain. *FEBS Lett.* **1997**, *407*, 18–20. [CrossRef]

118. Anderson, B.F.; Baker, H.M.; Norris, G.E.; Rice, D.W.; Baker, E.N. Structure of human lactoferrin: Crystallographic structure analysis and refinement at 2.8 A resolution. *J. Mol. Biol.* **1989**, *209*, 711–734. [CrossRef]

119. Kanyshkova, T.G.; Buneva, V.N.; Nevinsky, G.A. Lactoferrin and its biological functions. *Biochemistry* **2001**, *66*, 1–7. [PubMed]

120. Ramaswamy, H.; Swamy, C.V.; Das, M.R. Purification and characterization of a high molecular weight ribonuclease from human milk. *J. Biol. Chem.* **1993**, *268*, 4181–4187. [CrossRef]

121. He, J.; Furmanski, P. Sequence specificity and transcriptional activation in the binding of lactoferrin to DNA. *Nature* **1995**, *373*, 721–724. [CrossRef]

122. Kanyshkova, T.G.; Babina, S.E.; Semenov, D.V.; Isaeva, N.; Vlassov, A.V.; Neustroev, K.N.; Kul'minskaya, A.A.; Buneva, V.N.; Nevinsky, G.A. Multiple enzymic activities of human milk lactoferrin. *Eur. J. Biochem.* **2003**, *270*, 3353–3361. [CrossRef]

123. Kanyshkova, T.G.; Semenov, D.V.; Buneva, V.N.; Nevinsky, G.A. Human milk lactoferrin binds two DNA molecules with different affinities. *FEBS Lett.* **1999**, *451*, 235–237. [CrossRef]

124. Guschina, T.A.; Soboleva, S.E.; Nevinsky, G.A. Recognition of specific and nonspecific DNA by human lactoferrin. *J. Mol. Recognit.* **2013**, *26*, 136–148. [CrossRef]

125. Alinovskaya, L.I.; Sedykh, S.E.; Ivanisenko, N.V.; Soboleva, S.E.; Nevinsky, G.A. How human serum albumin recognizes DNA and RNA. *Biol. Chem.* **2018**, *399*, 347–360. [CrossRef]

126. Nevinsky, G.A.; Alinovskaya, L.I.; Sedykh, S.E.; Soboleva, S.E. How human alpha-lactalbumin recognize DNA and RNA. *Biochem. Anal. Biochem.* **2018**, *7*, 4.

127. Quinlan, G.J.; Martin, G.S.; Evans, T.W. Albumin: Biochemical properties and therapeutic potential. *Hematology* **2005**, *41*, 1211–1219. [CrossRef]

128. Gerasimova, Y.V.; Knorre, D.D.; Shakirov, M.M.; Godovikova, T.S. Human serum albumin as a catalyst of RNA cleavage: N-homocysteinylation and N-phosphorylation by oligonucleotide affinity reagent alter the reactivity of the protein. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5396–5398. [CrossRef]

129. Gerasimova, Y.V.; Erchenko, I.A.; Shakirov, M.M.; Godovikova, T.S. Interaction of human serum albumin and its clinically relevant modification with oligoribonucleotides. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4511–4514. [CrossRef]

130. Phillips, N.I.; Jenness, R. Isolation and properties of human α-lactalbumin. *Biochim. Biophys. Acta* **1971**, *229*, 407–410. [CrossRef]

131. Köhler, C.; Håkansson, A.; Svanborg, C.; Orrenius, S.; Zhivotovsky, B. Protease activation in apoptosis induced by MAL. *Exp. Cell Res.* **1999**, *249*, 260–268. [CrossRef]

132. Nevinsky, G.A. Autoimmune processes in multiple sclerosis: Production of harmful catalytic antibodies associated with significant changes in the hematopoietic stem cell differentiation and proliferation. In *Multiple Sclerosis*; Conzalez-Quevedo, A., Ed.; InTech: Rijeka, Croatia, 2016; pp. 100–147.

133. Nevinsky, G.A.; Buneva, V.N. Natural catalytic antibodies–abzymes. In *Catalytic Antibodies*; Keinan, E., Ed.; VCH-Wiley Press: Weinheim, Germany, 2005; pp. 505–569.

134. Nevinsky, G.A. Natural catalytic antibodies in norm and in autoimmune diseases. In *Autoimmune Diseases: Symptoms, Diagnosis and Treatment*; Brenner, K.J., Ed.; Nova Science Publishers Inc.: New York, NY, USA, 2010; pp. 1–107.

135. Nevinsky, G.A. Natural catalytic antibodies in norm and in HIV-infected patients. In *Understanding HIV/AIDS Management and Care—Pandemic Approaches the 21st Century*; Kasenga, F.H., Ed.; InTech: Rijeka, Croatia, 2011; pp. 151–192.

136. Nevinsky, G.A. Catalytic antibodies in norm and systemic lupus erythematosus. In *Lupus*; Khan, W.A., Ed.; InTech: Rijeka, Croatia, 2017; pp. 41–101.

137. Andreev, S.K.; Buneva, V.N.; Nevinsky, G.A. How human IgGs against DNA recognize oligonucleotides and DNA. *J. Mol. Recognit.* **2016**, *29*, 596–610. [CrossRef]