7384–7392 Nucleic Acids Research, 2012, Vol. 40, No. 15 Published online 14 May 2012
doi:10.1093/nar/gks423

Thermodynamics of the multi-stage DNA lesion recognition and repair by formamidopyrimidine-DNA glycosylase using pyrrolocytosine fluorescence—stopped-flow pre-steady-state kinetics

Nikita A. Kuznetsov[1], Yuri N. Vorobjev[1], Lev N. Krasnoperov[2] and Olga S. Fedorova[1,*]

[1]Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090, Russia and [2]New Jersey Institute of Technology, Newark NJ 07102, USA

Received February 15, 2012; Revised and Accepted April 24, 2012

ABSTRACT

Formamidopyrimidine-DNA glycosylase, Fpg protein from *Escherichia coli*, initiates base excision repair in DNA by removing a wide variety of oxidized lesions. In this study, we perform thermodynamic analysis of the multi-stage interaction of Fpg with specific DNA-substrates containing 7,8-dihydro-8-oxoguanosine (oxoG), or tetrahydrofuran (THF, an uncleavable abasic site analog) and non-specific (G) DNA-ligand based on stopped-flow kinetic data. Pyrrolocytosine, highly fluorescent analog of the natural nucleobase cytosine, is used to record multi-stage DNA lesion recognition and repair kinetics over a temperature range (10–30°C). The kinetic data were used to obtain the standard Gibbs energy, enthalpy and entropy of the specific stages using van't Hoff approach. The data suggest that not only enthalpy-driven exothermic oxoG recognition, but also the desolvation-accompanied entropy-driven enzyme-substrate complex adjustment into the catalytically active state play equally important roles in the overall process.

INTRODUCTION

Reactive oxygen species damage DNA to produce a variety of genotoxic lesions (1). In particular, 7,8-dihydro-8-oxoguanosine (oxoG) is one of the most common pre-mutagenic products of base oxidation in DNA (2). A variety of pathways for repair of damaged DNA exist (3). Among them, the predominant pathway is the base excision repair (BER) that processes small base lesions derived from the oxidation damage (4,5). DNA glycosylases play a key role in the BER pathway. They recognize a variety of modified or mismatched bases and release inappropriate bases from the deoxyribose phosphate chain (6,7).

OxoG is repaired through excision by formamidopyrimidine-DNA glycosylase (Fpg protein) in bacteria (8). Fpg protein possesses three types of activity: hydrolysis of the *N*-glycosidic bond with transient formation of an abasic (apurinic/apyrimidinic, AP) site (DNA glycosylase activity), elimination of the 3′-phosphate of the nascent AP-site (AP-lyase; β-elimination) and elimination of the 5′-phosphate of this modified AP-site (δ-elimination) (9). Consecutive execution of these three activities by Fpg removes the lesion from duplex DNA, where a single nucleotide gap in the damaged strand is left flanked by the phosphate residues. Catalysis by Fpg proceeds through the formation of a covalent imine (Schiff base) intermediate between Pro-1 residue of the enzyme and C-1′ of the damaged nucleotide (10,11).

It is still unclear how DNA glycosylases efficiently select sparse lesions among the enormous excess of normal DNA (12–14). The three-dimensional structure of *Escherichia coli* Fpg shows that DNA binding is accompanied with drastic conformational changes, including DNA bending, eversion of oxoG from DNA and insertion of Met-73, Arg-108 and Phe-110 residues into DNA (15). Similar structural changes are observed in other DNA glycosylases (12).

One possibility is that the conformational transitions during the substrate binding occur sequentially, with at least some steps structurally optimized for the selection of correct substrates (14). This hypothesis is supported by the stopped-flow studies of *E. coli* Fpg protein that used tryptophan (Trp) and 2-aminopurine (2-aPu) as fluorescence reporters, which reveal several conformational changes preceding the irreversible steps in the reaction pathway (16–18).

The five-step binding kinetic scheme was revealed by the enzyme's internal Trp fluorescence for oxoG-substrate

*To whom correspondence should be addressed. Tel: +7 383 3635174; Fax: +7 383 3635153; Email: fedorova@niboch.nsc.ru

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E{\bullet}S_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E{\bullet}S_2 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} E{\bullet}S_3 \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} E{\bullet}S_4 \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} E{\bullet}S_5 \xrightarrow{k_6} E{\bullet}P \overset{K_P}{\rightleftharpoons} E + P$$

Scheme 1. Binding and cleavage of the oxoG/C-substrate by Fpg (Trp and C^{py} fluorescence detection).

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E{\bullet}S_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E{\bullet}S_2 \xrightarrow{k_3} E{\bullet}P \overset{K_P}{\rightleftharpoons} E + P$$

Scheme 2. Binding and cleavage of the oxoG/C-substrate by Fpg (2-aPu fluorescence detection).

(Scheme 1). The parallel analysis of 2-aPu fluorescence showed that the five reversible steps in Scheme 1 correspond to only two fluorescently discernible conformational changes in DNA (Scheme 2).

Additional information on the dynamics of Fpg protein was provided in the experiments with the F110W and F110A mutants (18) The substitution of Phe with a fluorescent Trp residue allowed identification of the reaction step(s) corresponding to the movement of Phe-110. It was shown that in the case of oxoG/C-substrate, the second step ($\tau_{1/2}$ of ca. 50 ms) and the third step ($\tau_{1/2}$ of ca. 500 ms) were specifically sensitive to this substitution which revealed the participation of Phe-110 in these recognition steps. These steps occur after the first step of non-specific binding but before the eversion of oxoG in the active site of the enzyme. On the fluorescence curves observed for the non-specific G/C-ligand, the second and the third steps are observed indicating participation of Phe-110 residue in the interrogation of lesions in DNA. Most likely, the second step represents insertion of Phe-110 in the DNA chain (18). This interpretation has been corroborated recently in a single-molecule kinetics study (19). It was suggested (18,20) that the third and fourth steps are eversion of oxoG in the Fpg active site, insertion of Met-73 and Arg-108 residues in the DNA void and DNA kink formation. The fifth step is the formation of a catalytically competent enzyme-substrate intermediate (18,20). Previously, the active intrahelical interrogation and extrusion mechanism was proposed by Qi *et al.* (13) and Banerjee *et al.* (21) for a different bacterial 8-oxoguanine-DNA glycosylase, MutM enzyme from *Geobacillus stearothermophilus*. This mechanism differs from the 'passive' mechanism proposed for uracil DNA glycosylase (UNG), in which the protein captures a spontaneously extruded lesion (14) in contrast to the 'active' mechanism discussed above, where Phe-110 'wedge' formation is required in the lesion search.

To gain deeper insight in the mechanism [dubbed in (19) as the 'phonograph needle' mechanism] that permits checking for damages at random locations and to answer the question of how glycosylases find single damages among an overwhelming number of undamaged DNA bases, we perform thermodynamic analysis of the interaction of Fpg with specific (oxoG, THF) and non-specific (G) DNA-substrates based on the stopped-flow kinetic data. The multi-stage kinetic mechanism, the corresponding rate constants of the forward and reverse reactions and the resultant equilibrium constants are derived based on the temporal fluorescence traces of the fluorescent analog of a natural base, pyrrolocytosine, introduced in the DNA-substrates, over a very wide (six orders of magnitude, from milliseconds to thousand seconds) time and an extended temperature ranges. Pyrrolocytosine or 3-[b-D-2-ribofuranosyl]-6-methylpyrrolo[2,3-d]pyrimidin-2(3 H)-one (C^{py}), a fluorescent analog of the nucleoside cytidine retains its Watson–Crick base-pairing capacity with G and is only moderately affected by temperature, making it a versatile probe for fluorescent measurements of the kinetics at different temperatures (22).

The thermodynamic parameters of the specific stages obtained via the temperature-dependent study proved to be very helpful in the reliable identification of specific stages in the overall repair process based on the information that was not available earlier in the single temperature studies.

MATERIALS AND METHODS

Oligodeoxynucleotides and enzymes

Electrophoretically homogeneous *E. coli* Fpg protein was over-expressed, purified, quantified, assayed and stored as described previously (15,17). The fraction of the active enzyme (~90%) was determined by borohydride trapping. The reaction mixture included 2 μM enzyme, 25 mM potassium phosphate (pH 6.8), 100 mM NaCl, 100 mM $NaBH_4$ and varying amounts of oligonucleotide duplex containing an oxoG. The samples were incubated for 1 h at 25°C, mixed with the loading buffer [50 mM Tris–HCl (pH 6.8), 2% SDS, 10% glycerol, 1% 2-mercaptoethanol, 0.1% bromophenol blue] and separated by 12% SDS–PAGE. The gel was stained with coomassie and quantified using Gel-Pro Analyzer 4.0 software (Media Cybernetics, Silver Spring, MD, USA). Oligodeoxynucleotides d(TCTCTCTCXCCTTCCTT)/d(AAGGAAGGC^{py}GAGAGAGA), where X is G, THF and oxoG, C^{py} is fluorescent cytosine analogue pyrrolocytosine (Figure 1) were synthesized by established phosphoramidite methods on an ASM-700 synthesizer (BIOSSET Ltd., Novosibirsk, Russia) from phosphoramidites purchased from Glen Research (Sterling, VA, USA). The oligonucleotide containing oxoG was cleaved and deprotected with ammonium hydroxide containing 0.25 M 2-mercaptoethanol to avoid further oxidation of oxoG. Deprotected oligonucleotides were purified by ion-exchange HPLC on a Nucleosil 100-10 $N(CH_3)_2$ column followed by reverse-phase HPLC on a Nucleosil 100-7 C_{18} column (both 4.6 × 250 mm, purchased from Macherey-Nagel, Düren, Germany) to >98% homogeneity. The purity of oligonucleotides exceeded 98%, as estimated by electrophoresis in 20% denaturing

5'-d(TCTCTCTCXCCTTCCTT)-3' X = G, THF or oxoG
3'-d(AGAGAGAGC^{py}GGAAGGAA)-5'

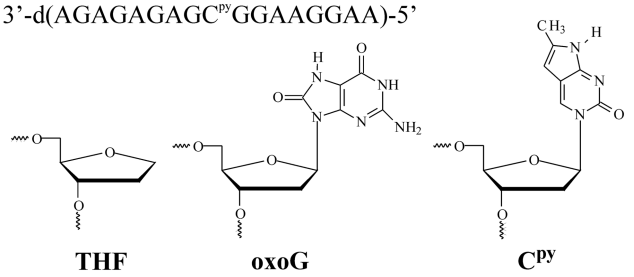

Figure 1. Sequences of oligodeoxynucleotides used in this work and the structures of modified residues.

polyacrylamide gel after staining with the Stains-All dye (Sigma-Aldrich, St. Louis, MO, USA). Concentrations of oligonucleotides were determined from their absorbance at 260 nm. Oligonucleotide duplexes were prepared by annealing modified and complementary strands at a 1:1 molar ratio in the reaction buffer [50 mM Tris–HCl (pH 7.5), 50 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, 9% glycerol].

Stopped-flow measurements

Stopped-flow measurements with fluorescence detection were carried out using a model SX.18MV stopped-flow spectrometer (Applied Photophysics) as described in (16,17). All experiments were carried out in a buffer containing 50 mM Tris–HCl (pH 7.5), 50 mM KCl, 1 mM Na_2EDTA, 1 mM dithiothreitol, 9% glycerol (v/v) at different temperatures (10–30°C). To detect C^{py} fluorescence, $\lambda^{ex} = 344$ nm was used and $\lambda^{em} > 370$ nm was monitored (Corion filter LG-370). The dead time of the instrument was 1.4 ms. Typically, each trace is the average of four or more individual experiments. As an example, one set of the original experimental data are presented in Supplementary Data. The concentration of oligodeoxyribonucleotide duplexes in all experiments was 1.0 μM, and concentrations of Fpg protein were varied in the range 0.25–2.0 μM.

Data processing

The approach is based on the fluorescence intensity variation in the course of the reaction due to the sequential formation and further transformation of DNA–enzyme complex and conformers. The kinetic parameters (both the rate constants and the response factors) were obtained by global non-linear fits using the DynaFit software (BioKin, Pullman, WA, USA) (23). The software performs numerical integration of a system of ODE with subsequent non-linear least-squares regression analysis. The response factors of the intermediates (that are essentially the products of the extinction coefficients and the fluorescence quantum yields) are treated as fitting parameters in the data processing.

In the evaluated mechanisms, except for the first bimolecular step, all other reactions are first order reactions. In the data processing, the kinetic information is obtained from the 'temporal behavior' of the fluorescence intensity, not from the 'amplitudes' of specific signal contributions. Only the rate constants obtained in non-linear fits were used to derive the thermochemical parameters. The 'response factors' for different conformers resulting from the fits were not used in the determination of the equilibrium constants, but rather provided additional information on the fluorescence intensity variations in different complexes and conformers.

Processing of individual kinetic curves does not unambiguously provide the kinetic parameters; therefore, global fits of sets of kinetic curves obtained at different concentrations of the reactants at each temperature was used. In the fits all relevant rate constants for the forward and reverse reactions, as well as the specific molar responses for all intermediate complexes were optimized.

Several mechanisms containing N binding reversible steps were evaluated. It was found that a binding mechanism with at minimum five reversible steps is required to provide fits with acceptable residuals comparable with the signal noise. The procedure of the mechanism determination was described in detail in previous publications (16,17).

Molecular modeling

The X-ray data obtained for the complexes of Fpg protein with identical 13-mer oxoG-DNA from *Geobacillus stearothermophilus*, pdb code 3GPP (13) and 3JR5 (24), respectively, are taken as models of atomic structures of protein–DNA complexes $(E•OG)_1$ and $(E•OG)_5$ in Scheme 1. Inspection of these structures revealed that the structure 3GPP represents an example of a non-specific protein–DNA complex $(E•OG)_1$, with unbend DNA interacting with the protein over the minor groove. The 3JR5 represents specific protein–DNA complex $(E•OG)_5$ with oxoG base everted from the double helix into the enzyme's active site, with filled void in the double helix by Arg-112 and enzyme forming a tight complex with the DNA-substrate in the catalytically active conformation. The contact areas ΔMS (molecular surface) of the protein–DNA interface were calculated as the difference of the MS confining the solvent-excluded volume (25) of the isolated protein and the DNA and protein–DNA complex, ΔMS = MS (protein–DNA complex) – MS (protein) – MS (DNA), where the buried MS has negative sign. The MS surface was calculated by the SIMS method (26). The free energy of the cavity formation in water solvent in the process of dissolution has entropic nature and is approximated by the linear equation $\Delta G_{cav} = \gamma \Delta MS$, where the parameter γ is in the range of 70–117 cal/mol/Å^2 (25). For estimation of the entropy loss we took the value of the parameter $\gamma = 80$ cal/mol/ Å^2 that approximates the entropy of dissolution of non-polar species in water (27).

RESULTS AND DISCUSSION

The thermodynamic analysis of the interaction of Fpg with DNA-substrates based on the stopped-flow kinetic data was performed. Using the 17-nt duplexes (Figure 1) multiple transient changes in the fluorescence were

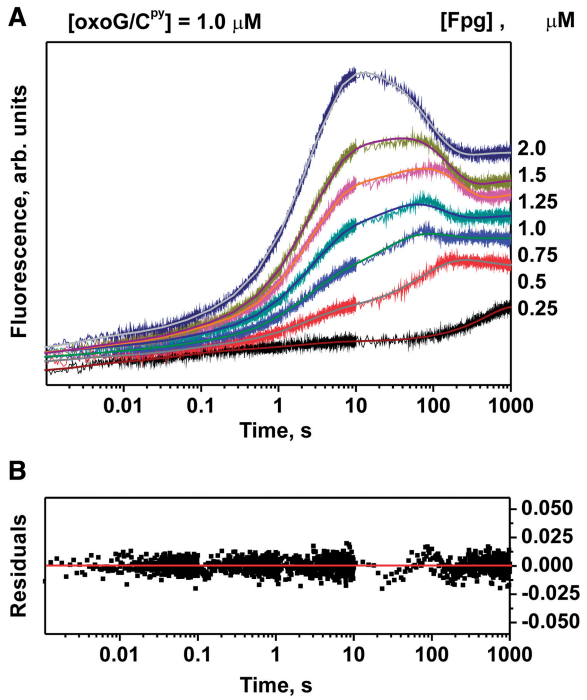

Figure 2. Changes in C^{py} fluorescence intensity during interaction of Fpg with (**A**) oxoG-substrate (1 μM) at varied concentrations of enzyme at 25°C (solid lines represent the fitted curves) and (**B**) residuals for fit by Scheme 1. The concentrations of the enzyme and the oxoG/C^{py}-substrate are 1 μM.

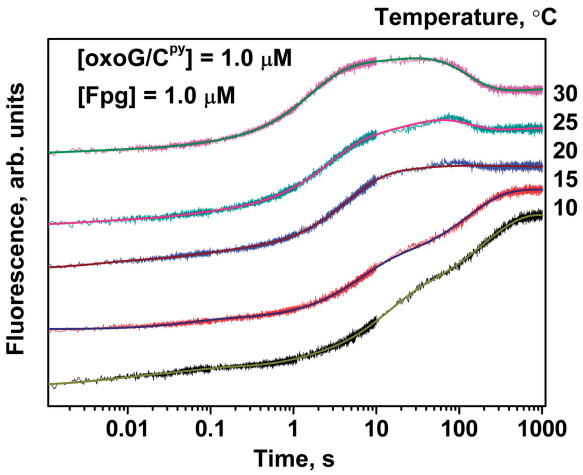

Figure 3. Changes in the C^{py} fluorescence intensity during the interaction of Fpg with oxoG-substrate at different temperatures. The concentrations of the enzyme and the substrate were 1 μM.

observed at single-turnover conditions, indicating conformational transitions in the DNA molecule. In contrast to our previous studies (16–18), where 12-nt duplexes as DNA-substrates were studied, the longer DNA permitted investigation of the process over a wide temperature range (10–30°C).

The C^{py} fluorescence traces were recorded for a set of Fpg concentrations at each temperature (10, 15, 20, 25, 30°C). The individual rate constants for all recognition steps under interaction with non-damaged DNA (G), with analog of AP-containing site (THF), and with oxoG-containing DNA (oxoG) were obtained from the fits.

The fluorescence of pyrrolocytosine is quenched in a double-stranded DNA compared with a single-stranded DNA (28,29). Figure 2A demonstrates, that the C^{py} fluorescence intensity is larger for the enzyme–DNA complex (the middle part of the traces, 10–100 s) compared with the free double-stranded DNA (the initial parts of the traces, short times) and to the final non-specific DNA–product–enzyme complex formed after the removal of oxoG (the final part of the traces, long times). The increase of the plateau level with the increase of Fpg concentration (Figure 2A, long times) indicates that C^{py} fluorescence intensity is higher when DNA is in the complex with the enzyme, compared with free dsDNA.

The measured fluorescence traces (Figure 2A) were fitted, as in the previous studies (16,17), by a kinetic mechanism (Scheme 1) containing five reversible steps, representing the sequential recognition of the damaged site and formation of the enzyme catalytically competent conformation. The sixth irreversible step corresponding to the catalytic reactions and the decomposition of the enzyme/product complex is clearly recognizable at high temperatures (Figure 3). The equilibrium constants were determined for each step in the reaction mechanism (Supplementary Table S1). The residuals of the fits by Scheme 1 are shown in Figure 2B. Figure 3 demonstrates the fluorescence traces obtained for 1 μM Fpg and 1 μM oxoG/C^{py}-containing DNA-substrate at different temperatures.

To detect the fluorescence changes C^{py} responsible to non-specific dsDNA binding, we used an undamaged duplex (G-ligand) (Figure 4A). The process of binding the non-specific G-ligand was completed by 0.1 s. An increase in the fluorescence intensity of C^{py} was observed indicating destabilization of Watson–Crick or stacking interactions in the primary non-specific enzyme–DNA complex. The 'melting' of DNA structure was shown in our previous work using undamaged duplex with 2-aPu residue (17). According to Qi *et al.* studies (24) three amino acid residues [Arg-77, Met-112 and Phe-114 for Fpg from *G. stearothermophilus*, which are analogs of Met-73, Arg-108 and Phe-110 in Fpg from *E. coli* (30)] penetrate DNA duplex and stabilize the bended DNA in non-specific complex. In this complex normal G base is partially extruded from the DNA, but still does not attain the catalytic site of the enzyme. In the recent work of Dunn *et al.* (19) it was shown that Phe-111 residue of *E. coli* Fpg (corresponds to Phe-110 in our work) might play a crucial role in the base interrogation. Measuring the rate of the enzyme scanning along DNA, it was found that when this residue was mutated in the *E. coli* Fpg gene to an alanine, there was a dramatic increase in the overall diffusion rate of the variant compared with the wild-type protein.

Fitting the experimental data to the one-step binding model (Scheme 3) gave the values for the forward and reverse rate constants and dissociation constant at different temperatures (Supplementary Table S2).

To gain a better understanding of the recognition of specific substrate by *E. coli* Fpg protein, we used the

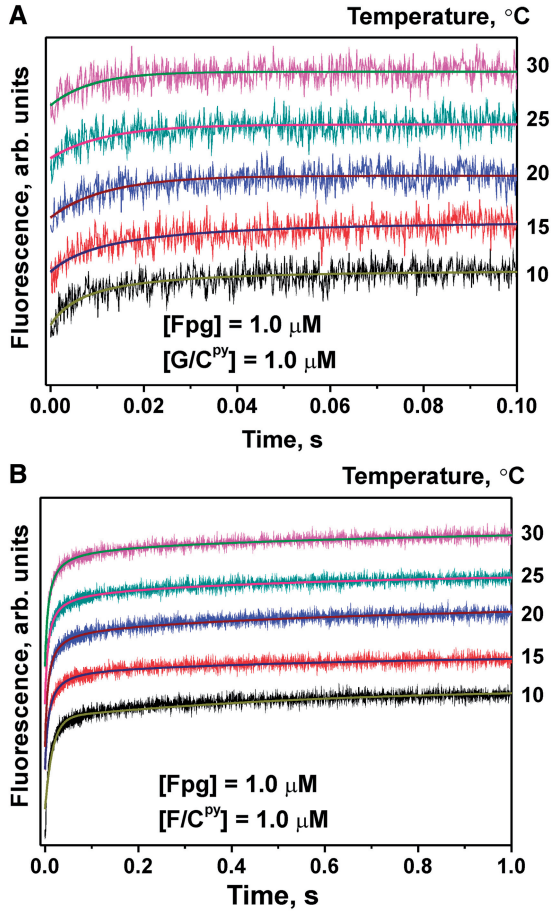

Figure 4. Changes in the C^{py} fluorescence intensity during the interaction of Fpg with (**A**) G- and (**B**) THF-ligand at different temperatures. The concentrations of the enzyme and the DNA were 1 μM.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\bullet S_1$$

Scheme 3. Binding of the G-ligand by Fpg (C^{py} fluorescence detection).

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\bullet S_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E\bullet S_2$$

Scheme 4. Binding of the THF-ligand by Fpg (C^{py} fluorescence detection).

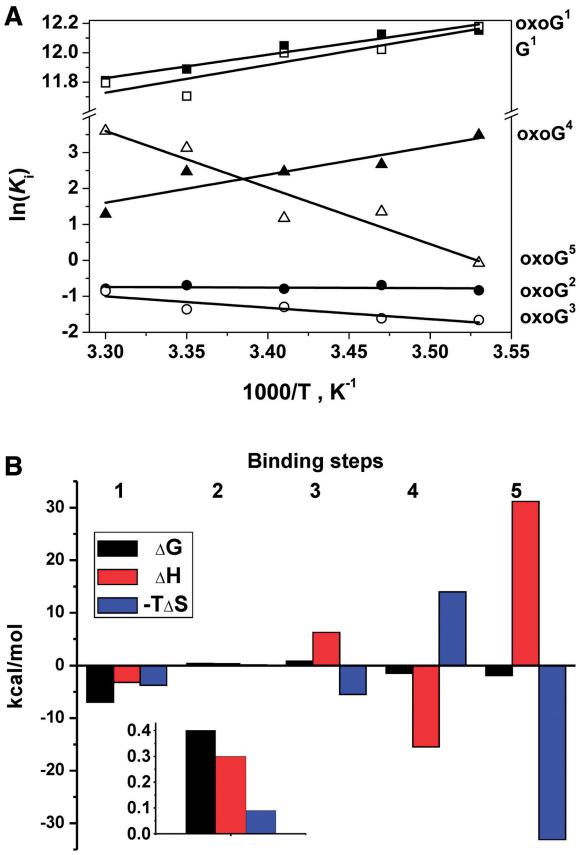

Figure 5. (**A**) Van't Hoff analysis of the temperature dependence of *Ki* for G- and oxoG-substrates. Upper indexes indicate the binding step number. (**B**) Thermodynamics signatures of oxoG-substrate binding to Fpg. The inset shows the thermodynamic effects for the second binding step.

model duplex containing the non-cleavable abasic site analog (THF-ligand). THF moiety (Figure 1) lacks the C-1′ hydroxyl group of the 2′-deoxyribose. This feature allows elimination of the catalytic event that occurs subsequent to abasic site recognition and to use THF-ligand for specific recognition studies. According to the X-ray data, the DNA in this complex is kinked at the point of the lesion, and Met-73, Arg-108 and Phe-110 are inserted in the DNA to fill the abasic void (30).

The changes the C^{py} fuorescence under interaction of Fpg with THF-ligand were observed (Figure 4B). The kinetic curves for the THF-ligand were characterized by an increase in C^{py} fluorescence intensity during 1 s. Therefore, filling of the abasic void in the DNA duplex by amino acids of Fpg leads to the increase in the C^{py} fluorescence. The experimental curves were fitted by a two-step binding model (Scheme 4). The rate constants of the elementary steps and the total binding constant of Fpg with the THF-ligand determined using this kinetic scheme are listed in Supplementary Table S2.

Using measured rate constants the equilibrium constants K_i (k_i/k_{-i}, i is the step number) were determined for G- and THF-ligands and oxoG-substrate. The ΔH^o_i and ΔS^o_i were calculated using the relationship: $ln(K_i) = \Delta S^o_i/R - \Delta H^o_i/RT$. As shown in Figure 5A, the dependence $ln(K_i)$ versus 1/T is linear, as expected for the relatively narrow temperature range of the study. The Gibbs free energies ΔG^o_i at 25°C were calculated from $\ln(K_i) = -\Delta G^o_i/RT$ (Table 1). Visualization of the inter-relationship of the thermodynamic parameters of the binding steps of oxoG-substrate is presented in Figure 5B.

Inspection of the thermodynamic data summarized in Table 1 reveals a clear qualitative difference in the

Table 1. Thermodynamics parameters of Fpg interactions with DNA[a]

DNA	Step number	Parameter				
		$\Delta G°_{i298}$, kcal/mol	$\Delta H°_i$, kcal/mol	$\Delta S°_i$, cal/K*mol	Equilibrium constants K_i (298K)	Process
G	1	−7.0	−3.8 ± 0.9	10.9 ± 3.2	$1.2 \times 10^5\,M^{-1}$	Non-specific binding, DNA distortion
THF	1	−7.2	−4.0 ± 0.3	10.8 ± 1.0	$1.9 \times 10^5\,M^{-1}$	Non-specific binding, DNA distortion
	2	0.7	6.7 ± 0.3	20.3 ± 0.9	0.5	DNA kinking together with amino acids insertion
	$\sum_{i=1}^{i=2}$ or $\prod_{i=1}^{n} K_i$	−6.5	2.7 ± 0.4	31.1 ± 1.3	$0.95 \times 10^5\,M^{-1}$	
oxoG	1	−7.0	−3.2 ± 0.4	12.7 ± 1.5	$1.4 \times 10^5\,M^{-1}$	Non-specific binding, DNA distortion
	2	0.4	0.3 ± 0.8	−0.3 ± 2.7	0.5	Enzyme conformational change into interrogation capable Phe-110 'wedge' state
	3	0.8	6.3 ± 1.7	18.4 ± 5.8	0.25	DNA kinking
	4	−1.5	−15.5 ± 3.9	−46.9 ± 13.5	12.7	Eversion of oxoG into enzyme active center together with Arg-108 and Met-73 insertion
	5	−1.9	31.2 ± 5.5	111.1 ± 18.6	25.0	Final adjustment of enzyme active center, DNA desolvation
	$\sum_{i=1}^{i=5}$ or $\prod_{i=1}^{n} K_i$	−9.2	19.1 ± 12.3	95.0 ± 42.1	$5.6 \times 10^6\,M^{-1}$	
	Transition state of the catalytic step 6[b]	19.6	6.0 ± 1.4	−45.5 ± 4.7		Irreversible catalytic step
	Product dissociation	−7.0	−4.1 ± 0.3	9.6 ± 0.9	$1.3 \times 10^5\,M^{-1}$	

[a]The errors indicated are ±1 SD $\Delta\Delta G°_{i298} = RT(\Delta K_i/K_i) \leq 0.1$ kcal/mol.
[b]Apparent thermodynamic parameters of the 'transition state' calculated as $\Delta G_6^{\ddagger} = RT\ln(k_B T/hk_6)$, where k_B and h are Boltzmann and Planck's constants, respectively, R is the gas constant, T is absolute temperature in Kelvins.

thermodynamics of Fpg binding to non-specific and specific DNA substrates that is important for the understanding of the mechanisms of specific sites recognition by the DNA glycosylases.

The first step of oxoG-recognition has similar thermodynamic parameters as the binding of non-specific DNA (G-ligand) and the first step of THF-ligand. The Gibbs free energy of binding (ca. −8 kcal/mol) is typical for protein–DNA complexes (31). Formation of the protein–DNA complex restraints translational, rotational and conformational motions of the protein and DNA leading to the entropy loss. The entropy gain is provided by water molecules release from the protein–DNA interface (31), which has been shown to accompany molecular complexation in many systems, including protein–DNA complexes.

To analyze the relationship between species in Schemes 1 and 2 registered by the tryptophan (16), 2-aminopurine (17) and pyrrolocytosine fluorescence, we performed a kinetic simulation of the reaction pathway using the rate constants determined by detection of all fluorescence types for the oxoG-substrate, and compared the time courses of the formation and disappearance of the intermediate species (Figure 6).

The comparison of the characteristic times of the appearance and disappearance of various enzyme–DNA complexes (Figure 6) shows that the first step of binding leads to simultaneous formation of ES_1^{Cpy} and ES_1^{Trp}. However, conversion of ES_1^{Trp} to ES_2^{Trp} occurs faster than the conversion of ES_1^{Cpy} to ES_2^{Cpy} indicating that the conformational changes of the enzyme occur before the binding with the DNA-specific site. This observation is in good agreement with the data (18,19) suggesting an

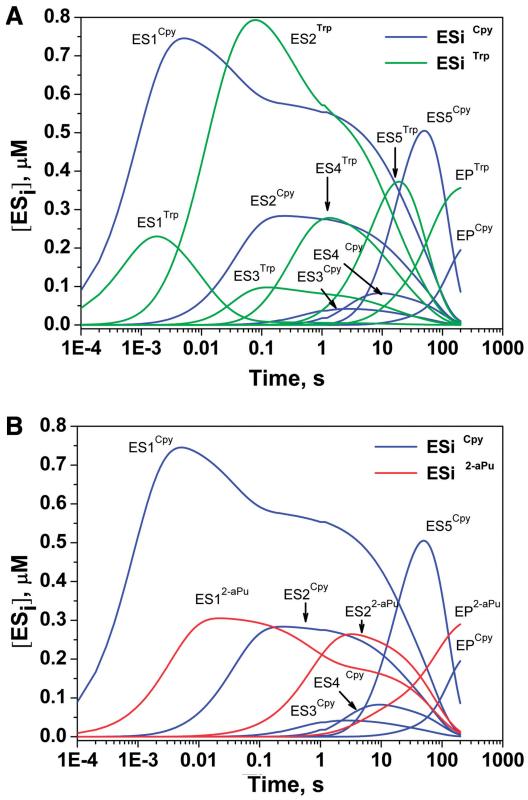

Figure 6. Kinetic simulation of the time-courses of the appearance and disappearance of different intermediates evident from the C^{py} fluorescence and the intrinsic Trp fluorescence (**A**) or 2-aPu fluorescence (**B**): blue line, C^{py} fluorescence; green line, intrinsic Trp fluorescence; red line, 2-aPu fluorescence. The intermediates are labeled according to Schemes 1 and 2, respectively. The simulations were run at $[E]_0 = [S]_0 = 1\,\mu M$.

active mechanism of the lesion search by Phe-110 at the first step of recognition. It should be noted that the formation of the ESi complexes detected by C^{py} fluorescence occurs after the formation of corresponding ESi complexes detected by Trp fluorescence, whereas only two complexes were detected by measurements of 2-aPu fluorescence. The $ES_1^{2\text{-aPu}}$ likely corresponds to ES_2 and ES_3, whereas $ES_2^{2\text{-aPu}}$ corresponds to ES_4 and ES_5, both detected by the Trp and C^{py} fluorescence.

We suggested that formation of ES_2^{Trp} complex leads to destabilization of the local DNA structure, this process also detected by C^{py} fluorescence with a delay of a few millisecond (formation of ES_2^{Cpy}). At this stage, 2-aPu fluorescence indicates a decrease in the level of stacking of the 2-aPu reporter base adjacent to the damaged base. The next step of the recognition accompanied with the formation of ES_3 complex, was attributed to the bending of the DNA chain. The ES_4 complex was ascribed to the oxoG eversion and the amino acids plugging into the DNA chain because of the decrease in the 2-aPu and C^{py} fluorescence intensity in this step. In the last recognition step detected by Trp and C^{py} fluorescence all specific contacts are formed in the complex ES_5 and the next catalytic reaction step can proceed.

The thermodynamic parameters of the steps of oxoG recognition, together with our previous data (16–18) on the Trp and 2-aminopurine fluorescence changes upon Fpg complexation with DNA, suggest that the sequential reaction steps in Scheme 1 correspond to: (i) non-specific primary encounter; (ii) initial recognition with the destabilization of the DNA around the lesion with the insertion of Phe-110; (iii) formation of a kink in the DNA chain; (iv) eversion of oxoG base from the double helix into the enzyme's active site; filling the resulting void in the double helix by Arg-108 and Met-73 and (v) isomerization of the enzyme to form a tight complex with the DNA-substrate and to produce catalytically active conformation.

The first step has similar thermodynamic parameters for all DNA ligands containing G, THF and oxoG. In this step the moderate enthalpy gain is accompanied by an increase in entropy most probably due to the DNA 'melting' in the point of contact and the conformational changes of the DNA binding site in the Fpg molecule with the emphasis on the contribution of the Phe-110 wedge movement (13,16–19).

The second step is energetically neutral that is very important for the 'active' mechanism of the lesion chasing using the Phe-110 'wedge'.

The third step is endothermic but is accompanied by the increase in entropy. This is consistent with the energy required to kink the DNA and the water molecules release in the kink position.

The fourth step has favorable enthalpy but unfavorable entropy, which makes us to modify our previous interpretation (18,20) and to assign this step to extrusion of oxoG base from the DNA chain and the insertion in the enzyme active site, as well as the insertion of Arg-108 and Met-73 residue in the DNA chain (Figure 7). In this step, new contacts between aminoacids of the active center and oxoG residue, as well as between the aminoacids of Fpg with the DNA in the void are formed that leads to the energy gain. However, the resulting structure is more rigid, which explains the entropy loss.

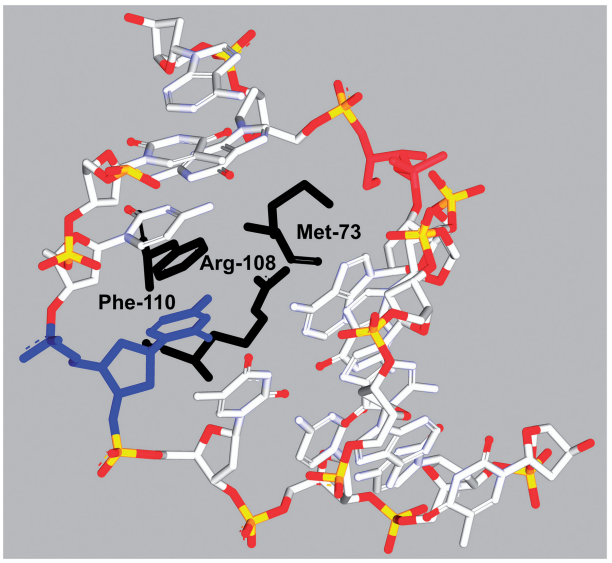

Figure 7. Representation of the aminoacids that surround the cytosine base (colored blue) in the Fpg enzyme complex with DNA substrate (PDB ID 1K82). The triad Phe-110, Arg-108 and Met-73 are inserted into the duplex.

The fifth reversible recognition step is characterized by an unfavorable enthalpy that is compensated by favorable entropic contribution. The large positive entropy of protein binding originates from the dehydration of the DNA grooves. Changes in the solvation are crucial in assisting the site discrimination during the catalysis, as well as in the directing sequence-specific DNA binding by numerous enzymes [see, e.g. (31)].

Using X-ray data obtained for the complexes of Fpg protein from *G. stearothermophilus* with oxoG-DNA, [pdb code 3GPP (13) and 3JR5 (24)] as atomic models for the Fpg–DNA complexes of the Stage 1 and Stage 5, respectively, change of the area of the protein–DNA interface during the formation of the catalytically competent complex ES_5 from the initial non-specific complex ES_1 was estimated. The DNA bending and the insertion of the oxoG into protein cavity in the complex ES_5 are accompanied by the increase of the contact surface interface by 680 $Å^2$ (from 894 to 1583 $Å^2$ in complexes ES_1 and ES_5, respectively). Up to five additional base pairs of DNA are involved in the contact with the protein over the major groove of DNA. Burring of 680 $Å^2$ of the molecular surface of the tight protein–DNA complex gives entropy gain of about 180 cal/K*mol. This number is in a qualitative agreement with the data in Table 1, taking into account that formation of the tight specific protein–DNA complex leads to a decrease of the conformational entropy of the complex.

Total change of Gibbs energy, $\Delta G_{298} = \Sigma\ \Delta G_i = -9.2$ kcal/mol within the experimental error agrees with our previous value of -10.9 kcal/mol obtained using the 'stepwise increased ligand complexity' (SILC) method (32). The energies of the first steps of oxoG-substrate

and G- and THF-ligands binding, as well as the product dissociation step are in good agreement. It means that this stage of binding is characterized by the same processes that are independent of the different structure of DNA duplexes. Our data obtained for the binding of the THF-ligand are in agreement with $\Delta H = 0.8$ kcal/mol, $\Delta S = 35.2$ cal/K*mol (33), and show that the complexation of Fpg with the THF-containing duplex is exclusively an entropy-driven process. The sum of the ΔH_i and ΔS_i values (Table 1) characterizing the overall process of the Fpg binding with oxoG-containing DNA substrate differs from the data obtained for the interaction of Fpg protein with the THF-ligand, which model the apurinic/apyrimidinic site, the intermediate of oxoG conversion. Comparison of these values leads to the conclusion that recognition of oxoG residue versus AP-site has unfavorable enthalpy ($\Delta\Delta H = 16.4$ kcal/mol) but favorable entropy ($\Delta\Delta S = 63.9$ cal/K*mol). This difference can be explained by the very unfavorable enthalpy of adjustment of enzyme–DNA complex to catalytically competent state. On the other hand, the binding of oxoG-containing duplex leads to the formation of additional contacts in the protein–DNA complex. This induces additional desolvation of the protein and DNA surfaces that increases entropy. Therefore, essentially the desolvation-related entropic forces drive protein into the specific binding with DNA and to adjustment to the catalytically competent conformation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2 and Supplementary Experimental Data.

FUNDING

Russian Foundation for Basic Research [RFBR 10-04-00070, 12-03-93180 and 12-04-00135]; Russian Ministry of Education and Science [SS-64.2012.4]; Program of Russian Government to support leading scientists [11.G34.31.0045]; Siberian Branch of Russian Academy of Sciences. Funding for open access charge: Russian Foundation for Basic Research [RFBR 10-04-00070].

Conflict of interest statement. None declared.

REFERENCES

1. von Sonntag,C. (2006) *Free-Radical-Induced DNA Damage and Its Repair: A Chemical Perspective*. Springer, Berlin - Heidelberg.
2. Wallace,S.S. (2002) Biological consequences of free radical-damaged DNA bases. *Free Radic. Biol. Med.*, **33**, 1–14.
3. Friedberg,E.C., Walker,G.C. and Siede,W. (1995) *DNA Repair and Mutagenesis*. ASM Press, Washington.
4. Gros,L., Saparbaev,M.K. and Laval,J. (2002) Enzymology of the repair of free radicals-induced DNA damage. *Oncogene*, **21**, 8905–8925.
5. Memisoglu,A. and Samson,L. (2000) Base excision repair in yeast and mammals. *Mutat. Res.*, **451**, 39–51.
6. David,S.S. and Williams,S.D. (1998) Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chem. Rev.*, **98**, 1221–1261.
7. Krokan,H.E., Standal,R. and Slupphaug,G. (1997) DNA glycosylases in the base excision repair of DNA. *Biochem. J.*, **325**, 1–16.
8. Zharkov,D.O., Shoham,G. and Grollman,A.P. (2003) Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair*, **2**, 839–862.
9. Bhagwat,M. and Gerlt,J.A. (1996) 3′- and 5′-strand cleavage reactions catalyzed by the Fpg protein from *Escherichia coli* occur via successive β- and δ-elimination mechanisms, respectively. *Biochemistry*, **35**, 659–665.
10. Tchou,J. and Grollman,A.P. (1995) The catalytic mechanism of Fpg protein. Evidence for a Schiff base intermediate and amino terminus localization of the catalytic site. *J. Biol. Chem.*, **270**, 11671–11677.
11. Zharkov,D.O., Rieger,R.A., Iden,C.R. and Grollman,A.P. (1997) NH2-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein. *J. Biol. Chem.*, **272**, 5335–5341.
12. Hitomi,K., Iwai,S. and Tainer,J.A. (2007) The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. *DNA Repair*, **6**, 410–428.
13. Qi,Y., Spong,M.C., Nam,K., Banerjee,A., Jiralerspong,S., Karplus,M. and Verdine,G.L. (2009) Encounter and extrusion of an intrahelical lesion by a DNA repair enzyme. *Nature*, **462**, 762–766.
14. Friedman,J.I. and Stivers,J.T. (2010) Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry*, **49**, 4957–4967.
15. Gilboa,R., Zharkov,D.O., Golan,G., Fernandes,A.S., Gerchman,S.E., Matz,E., Kycia,J.H., Grollman,A.P. and Shoham,G. (2002) Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *J. Biol. Chem.*, **277**, 19811–19816.
16. Koval,V.V., Kuznetsov,N.A., Zharkov,D.O., Ishchenko,A.A., Douglas,K.T., Nevinsky,G.A. and Fedorova,O.S. (2004) Pre-steady-state kinetics shows differences in processing of various DNA lesions by *Escherichia coli* formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res.*, **32**, 926–935.
17. Kuznetsov,N.A., Koval,V.V., Zharkov,D.O., Vorobiev,Y.N., Nevinsky,G.A., Douglas,K.T. and Fedorova,O.S. (2007) Kinetic basis of lesion specificity and opposite-base specificity of *Escherichia coli* formamidopyrimidine-DNA glycosylase. *Biochemistry*, **46**, 424–435.
18. Koval,V.V., Kuznetsov,N.A., Ishchenko,A.A., Saparbaev,M.K. and Fedorova,O.S. (2010) Real-time studies of conformational dynamics of the repair enzyme *E. coli* formamidopyrimidine-DNA glycosylase and its DNA complexes during catalytic cycle. *Mutat. Res.*, **685**, 3–10.
19. Dunn,A.R., Kad,N.M., Nelson,S.R., Warshaw,D.M. and Wallace,S.S. (2011) Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. *Nucleic Acids Res.*, **39**, 7487–7498.
20. Fedorova,O.S., Kuznetsov,N.A., Koval,V.V. and Knorre,D.G. (2010) Conformational dynamics and pre-steady-state kinetics of DNA glycosylases. *Biochemistry (Moscow)*, **75**, 1225–1239.
21. Banerjee,A., Santos,W.L. and Verdine,G.L. (2006) Structure of a DNA glycosylase searching for lesions. *Science*, **311**, 1153–1157.
22. Tinsley,R.A. and Walter,N.G. (2006) Pyrrolo-C as a fluorescent probe for monitoring RNA secondary structure formation. *RNA*, **12**, 522–529.
23. Kuzmic,P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.*, **237**, 260–273.
24. Qi,Y., Spong,M.C., Nam,K., Karplus,M. and Verdine,G.L. (2010) Entrapment and structure of an extrahelical guanine attempting to enter the active site of a bacterial DNA glycosylase, MutM. *J. Biol. Chem.*, **285**, 1468–1478.
25. Vorobjev,Y.N. (2011) Advances in implicit models of water solvent to compute conformational free energy and molecular dynamics of proteins at constant pH. *Adv. Protein Chem. Struct. Biol.*, **85**, 281–322.

26. Vorobjev,Y.N. and Hermans,J. (1997) SIMS: computation of a smooth invariant molecular surface. *Biophys. J.*, **73**, 722–732.
27. Rose,G.D. and Wolfenden,R. (1993) Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu. Rev. Biophys. Biomol. Struct.*, **22**, 381–415.
28. Rist,M.J. and Marino,J.P. (2002) Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions. *Curr. Org. Chem.*, **6**, 775–793.
29. Berry,D.A., Jung,K.Y., Wise,D.S., Sercel,A.D., Pearson,W.H., Mackie,H., Randolph,J.B. and Somers,R.L. (2004) Pyrrolo-dC and pyrrolo-C: fluorescent analogs of cytidine and 2′-deoxycytidine for the study of oligonucleotides. *Tetrahedron Lett.*, **45**, 2457–2461.
30. Pereira de Jesus,K., Serre,L., Zelwer,C. and Castaing,B. (2005) Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA. *Nucleic Acids Res.*, **33**, 5936–5944.
31. Privalov,P.L., Dragan,A.I., Crane-Robinson,C., Breslauer,K.J., Remeta,D.P. and Minetti,C.A. (2007) What drives proteins into the major or minor grooves of DNA? *J. Mol. Biol.*, **365**, 1–9.
32. Ishchenko,A.A., Vasilenko,N.L., Sinitsina,O.I., Yamkovoy,V.I., Fedorova,O.S., Douglas,K.T. and Nevinsky,G.A. (2002) Thermodynamic, kinetic, and structural basis for recognition and repair of 8-oxoguanine in DNA by Fpg protein from *Escherichia coli*. *Biochemistry*, **41**, 7540–7548.
33. Minetti,C.A., Remeta,D.P., Zharkov,D.O., Plum,G.E., Johnson,F., Grollman,A.P. and Breslauer,K.J. (2003) Energetics of lesion recognition by a DNA repair protein: thermodynamic characterization of formamidopyrimidine-glycosylase (Fpg) interactions with damaged DNA duplexes. *J. Mol. Biol.*, **328**, 1047–1060.