ISSN 0006-2979, Biochemistry (Moscow), 2012, Vol. 77, No. 3, pp. 270-279. © Pleiades Publishing, Ltd., 2012.
Original Russian Text © A. G. Derevyanko, A. V. Endutkin, A. A. Ishchenko, M. K. Saparbaev, D. O. Zharkov, 2012, published in Biokhimiya, 2012, Vol. 77, No. 3, pp. 342-353.
Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM11-230, January 22, 2012.

Initiation of 8-Oxoguanine Base Excision Repair within Trinucleotide Tandem Repeats

A. G. Derevyanko[1,2], A. V. Endutkin[1,2], A. A. Ishchenko[3], M. K. Saparbaev[3], and D. O. Zharkov[1,2]*

[1]*Institute of Chemical Biology and Fundamental Medicine, Siberian Division of the Russian Academy of Sciences, pr. Lavrentieva 8, 630090 Novosibirsk, Russia; fax: (383) 363-5153; E-mail: dzharkov@niboch.nsc.ru*
[2]*Novosibirsk State University, ul. Pirogova 2, 630090 Novosibirsk, Russia*
[3]*Universite Paris-Sud XI, UMR 8126 C.N.R.S., Institut Gustave Roussy, Villejuif Cedex F-94805, France*

Received August 9, 2011
Revision received September 29, 2011

Abstract—Trinucleotide repeat expansion provides a molecular basis for several devastating neurodegenerative diseases. In particular, expansion of a CAG run in the human *HTT* gene causes Huntington's disease. One of the main reasons for triplet repeat expansion in somatic cells is base excision repair (BER), involving damaged base excision and repair DNA synthesis that may be accompanied by expansion of the repaired strand due to formation of noncanonical DNA structures. We have analyzed the kinetics of excision of a ubiquitously found oxidized purine base, 8-oxoguanine (oxoG), by DNA glycosylase OGG1 from the substrates containing a CAG run flanked by AT-rich sequences. The values of k_2 rate constant for the removal of oxoG from triplets in the middle of the run were higher than for oxoG at the flanks of the run. The value of k_3 rate constant dropped starting from the third CAG-triplet in the run and remained stable until the 3′-terminal triplet, where it decreased even more. In nuclear extracts, the profile of oxoG removal rate along the run resembled the profile of k_2 constant, suggesting that the reaction rate in the extracts is limited by base excision. The fully reconstituted BER was efficient with all substrates unless oxoG was near the 3′-flank of the run, interfering with the initiation of the repair. DNA polymerase β was able to perform a strand-displacement DNA synthesis, which may be important for CAG run expansion initiated by BER.

DOI: 10.1134/S0006297912030054

Key words: base excision repair, trinucleotide repeat expansion, CAG triplets, Huntington's disease, 8-oxoguanine, OGG1

Dynamic mutations are an important type of genome alterations, their inheritance differing from the classical Mendelian inheritance. Dynamic mutations are characterized by genetic anticipation describing the tendency of age of onset to decrease and severity of symptoms to increase through successive generations of an affected family. The major source of dynamic mutations is expansion of trinucleotide tandem repeats, which are found in many genes of many organisms. When the tandem repeat run is short enough, its further elongation is next to impossible, but the elongation rate drastically increases when the run achieves a certain threshold. Expansion of trinucleotide tandem repeats in humans is associated with many severe genetic disorders, especially those affecting the nervous system. In many cases (Huntington's disease, dentatorubropallidoluysian atrophy, spinobulbar muscular atrophy (Kennedy disease), and several types of inherited cerebellar ataxia), expansion of CAG triplets encoding glutamine residues occurs. The protein product of such a gene becomes nonfunctional and tends to form insoluble deposits within the cell [1, 2]. Expansion of other trinucleotides can cause other diseases (CGG – fragile X syndrome; GAA – Friedreich's ataxia; CTG – myotonic dystrophy); these triplets can also be located in a noncoding region of the gene [2].

Huntington's disease (G10 according to ICD-10) is a neurodegenerative genetic disorder caused by an autosomal dominant mutation associated with expansion of CAG trinucleotide tandem repeats in the first exon of the *HTT* gene. This leads to elongation of a polyglutamine (polyQ) peptide in the N-terminal region of huntingtin, the *HTT* gene protein product. The wild type of the gene

Abbreviations: AP, apurine/apyrimidine; BER, base excision repair; DTT, dithiothreitol; NE, nuclear extract; ODN, oligodeoxyribonucleotide; oxoG, 8-oxoguanine.

* To whom correspondence should be addressed.

contains from 9 through 35 CAG repeats, the obligate development of pathology is observed when the number of the repeats is more than 41, and carriers with alleles containing 36-40 repeats comprise the risk group with incomplete (reduced) penetrance [3, 4]. The onset time depends on the run length; the disease manifests in persons younger than 17-18 years when the number of repeats is more than 70.

Until recently, it was assumed that dynamic mutations arise exclusively during replication, and, hence, gradual elongation of repeat run to its critical value occurs when genes are inherited through successive generations. However, when studying heterogeneity of CAG run length in the *HTT* gene within one organism, a drastic expansion of the run in patients with Huntington's disease was found in cells of specifically affected regions: brain cortex and striatum (the basal ganglion of the brain is particularly responsible for motor functions) [5]. This expansion is observed even at very early stages of the disease, with the run length achieving more than 1000 repeats [5]. It is obvious that replicative mechanisms can hardly provide this magnitude of the run expansion in nonproliferating somatic tissue. Two processes associated with DNA repair, namely base excision repair (BER) and mismatch repair, were recently found to be able to provide the source for the CAG run expansion in somatic cells. In particular, CAG trinucleotide expansion can be initiated during BER when 8-oxoguanine (oxoG) is present in DNA [6].

BER is responsible for removal of the majority of damaged bases from DNA [7]. In the case of oxoG, at the initial stage it is recognized by 8-oxoguanine DNA glycosylase (OGG1), which cleaves the N-glycosidic bond between oxoG and the C1′ deoxyribose to form a nick at the 3′-end from the damaged nucleotide by the mechanism of β-elimination (Fig. 1). The formed 3′-terminal unsaturated fragment is eliminated from DNA by the phosphodiesterase activity of apurine/apyrimidine (AP)-endonuclease APEX1. Thus a single-nucleotide gap arises framed by 3′-OH and 5′-phosphate groups. BER can further develop via two general pathways – short- and long-patch ones. In the first case, DNA polymerase β (POLβ) inserts a dGMP residue, and the remaining single-stranded nick is ligated by DNA-ligase IIIα (LIG3α) in a complex with XRCC1 protein (Fig. 1a). In the long-patch BER variant, the repair synthesis continues after insertion of the first dNMP with the displacement or degradation of the forthcoming strand (Fig. 1b). In higher eukaryotes, DNA-polymerases δ, ε, or β with participation of FEN1 endonuclease and auxiliary factors PCNA, RFC, and RPA are responsible for this synthesis. Then the single-stranded nick is ligated by DNA ligase 1 (LIG1) [7].

Reconstitution of CAG run expansion *in vitro* has shown a requirement for OGG1, APEX1, POLβ, and either LIG3α/XRCC1 complex or DNA ligase 1 [6, 8]. FEN1 endonuclease and HMGB1 stimulating endonuclease activities of APEX1 and FEN1 facilitate expansion, but they have no influence on the length of the final product [8]. It was shown using a transgenic murine model that the effectiveness of CAG run expansion depends on the ratio of distinct BER enzymes in the cell [9].

Thus, possible implication of OGG1 in CAG tandem repeats in somatic cells has been shown to date. Nevertheless, data are completely absent on the kinetic parameters of oxoG excision from such substrates and on efficiency of the enzyme depending on the damaged base position in the CAG run. Since oxoG repair is limited by the efficiency of its initiation by OGG1 [10], it is relevant to study the OGG1 activity in cleavage of DNA substrates modeling the putative base triplet expansion site, as well as to study efficiency of the full BER cycle for these substrates.

MATERIALS AND METHODS

Oligodeoxyribonucleotides and enzymes. Activity and kinetic parameters of the OGG1 – crude extract and purified enzyme – were determined using the following oligodeoxyribonucleotides (ODNs):

CAG1, 5′-TTTTCA**X**CAGCAGCAGCAGCAGCAGAAAA-3′;
CAG2, 5′-TTTTCAGCA**X**CAGCAGCAGCAGCAGAAAA-3′;
CAG3, 5′-TTTTCAGCAGCA**X**CAGCAGCAGCAGAAAA-3′;
CAG4, 5′-TTTTCAGCAGCAGCA**X**CAGCAGCAGAAAA-3′;
CAG5, 5′-TTTTCAGCAGCAGCAGCA**X**CAGCAGAAAA-3′;
CAG6, 5′-TTTTCAGCAGCAGCAGCAGCA**X**CAGAAAA-3′;
CAG7, 5′-TTTTCAGCAGCAGCAGCAGCAGCA**X**AAAA-3′;
CGG2-1, 5′-TTTTCGGC**X**GCGGCGGCGGCGGCGGAAAA-3′;
CGG2-2, 5′-TTTTCAGCG**X**CGGCGGCGGCGGCGGAAAA-3′;
CAGcomp, 5′-TTTTCTGCTGCTGCTGCTGCTGCTGAAAA-3′;
CGGcomp, 5′-TTTTCCGCCGCCGCCGCCGCCGCCGAAAA-3′,
(where X is oxoG).

The ODNs were synthesized by a solid-phase phosphoramidite method from commercially available reagents on an automated ASM-800 synthesizer (Biosset, Russia) and purified by reverse-phase HPLC on a PRP-1 column (Hamilton, USA). The ODNs CAG1-7 were radioactively labeled at the 5′-end using T4 polynucleotide kinase (Biosan, Russia) and γ-[^{32}P]ATP (Laboratory of Biotechnology, Institute of Chemical Biology and Fundamental Medicine, Siberian Division of the Russian Academy of Sciences) according to the protocol described in [11]. To synthesize dsDNA substrates, the labeled ODNs were annealed with twofold molar excess of the complementary strand, CAGcomp. T4 DNA ligase was purchased from Biosan, and recombinant enzymes Fpg, OGG1, APEX1, and POLβ were isolated from *E. coli* overproducing strains according to earlier described protocols [12-15].

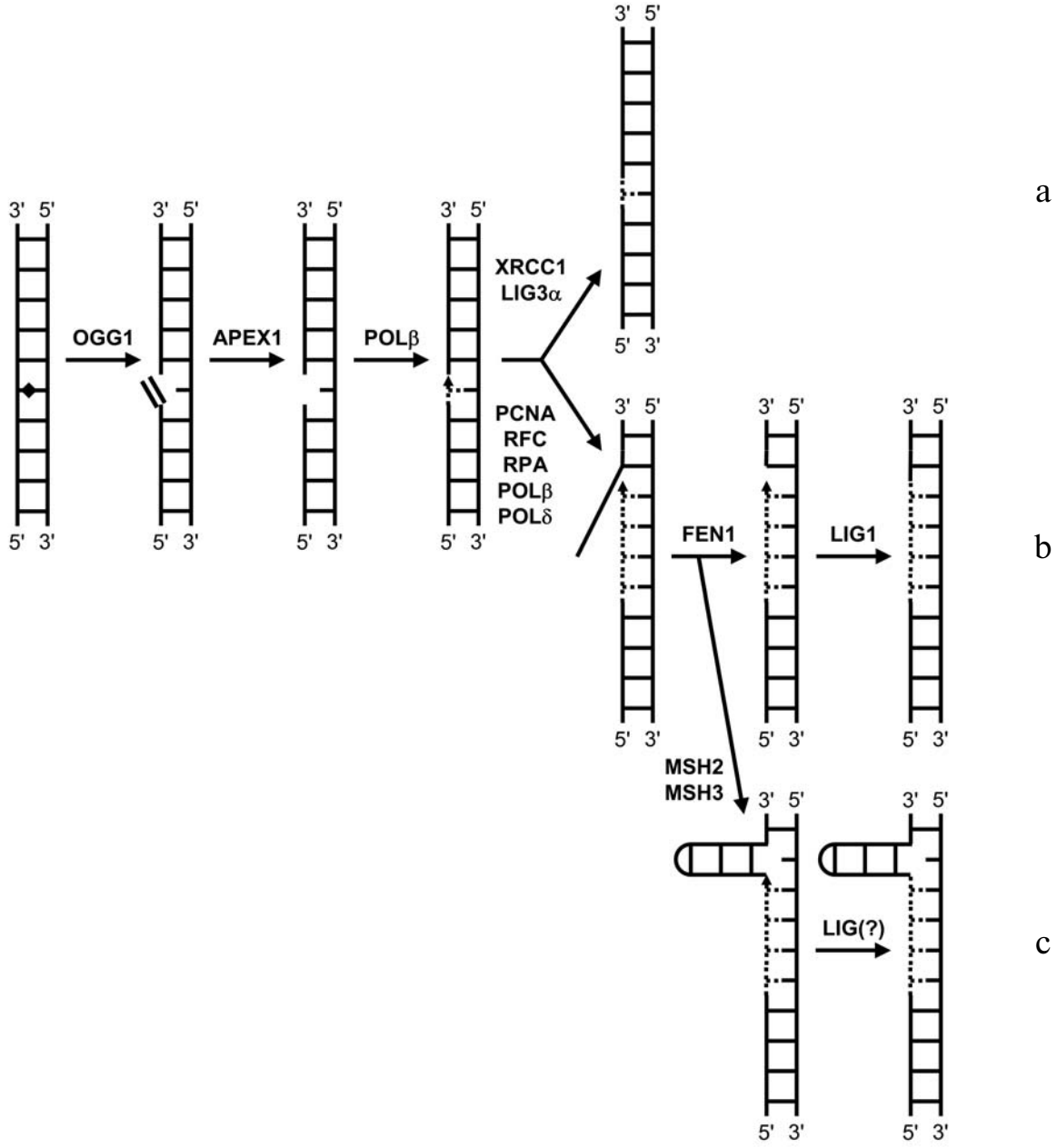

Fig. 1. General scheme of oxoG base excision repair (BER) in human cells. The diamond denotes the damaged DNA base; double line, α,β-unsaturated aldehyde, a product of β-elimination; dashed line, newly synthesized DNA patch; arrow, free 3′-end. The scheme shows the short-patch BER (a), long-patch BER (b), and long-patch BER with stable hairpin formation and trinucleotide repeat expansion (c). In the latter case, the nature of DNA ligase that acts at the final repair stage is unknown.

Determination of kinetic parameters of the reaction catalyzed by OGG1. The rate constant of oxoG base excision from DNA (k_2) by OGG1 was determined under conditions of single-turnover kinetics. The reaction was carried out in 220 μl (final volume) of reaction mixture containing 50 mM Tris-HCl buffer, pH 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 10 nM of radioactively labeled dsODN as a substrate, and 200 nM OGG1. Before addition, the enzyme was diluted in 25 mM Tris-HCl buffer, pH 7.5, containing 50 mM NaCl, 0.5 mM EDTA, 0.5 mM DTT, and 0.5% BSA. The reaction was carried out at 13°C for 20 sec-120 min. To determine the product release rate constant (k_3), experiments were carried out under the burst-phase kinetics conditions: the reaction was carried out in the same buffer at 37°C for 1-30 min, concentration of the ODN-substrate was 100 nM, and the concentration of OGG1 was 5 nM. In both cases, an aliquot (20 μl) was taken at the needed

moment, and the reaction was terminated by addition of 2 μl of 1 M NaOH and heating at 95°C for 2 min followed by neutralization by addition of 2 μl of 1 M HCl. Then 12 μl of a solution containing 80% formamide, 20 mM Na-EDTA, 0.1% (w/v) xylene cyanole, and 0.1% (w/v) bromophenol blue were added. The reaction products were analyzed electrophoretically in 20% polyacrylamide gel in presence of 7.2 M urea. Radioactivity of bands was determined by phosphorimaging of an Image Screen K (Kodak, USA) using a Molecular Imager FX system (Bio-Rad Laboratories, USA). The rate constants k_2 and k_3 were determined by nonlinear or linear regression of parameters of the equations $[P] = [S]_0(1 - e^{-k_2 t})$ and $v = k_3[E]_0$, respectively, using the SigmaPlot 8.0 program package (SPSS, USA) from the data of 3-5 independent experiments.

Preparation of cell extracts. Fractionated cell extracts containing DNA glycosylase activities [16-20] were prepared from cultured MRC-5 and LICH cells. The cell layers were grown on DMEM medium containing 10% fetal calf serum 1% sodium pyruvate, 1% penicillin, and 1% streptomycin until the confluent state. The dishes with the cells were washed twice with 20 ml of cold PBS, and the cells were rubbed off the dishes, resuspended in 1.5 ml of PBS, and centrifuged at 1000*g* and 4°C for 5 min. The pellet was resuspended in an equal volume of 20 mM HEPES-KOH buffer, pH 7.6, containing 0.15 mM EDTA, 10 mM KCl, 0.15 mM spermine, 0.75 mM spermidine, and protease inhibitor mix (Roche, Switzerland). Then Nonidet P40 was added to the final concentration of 0.1%, and the cells were incubated on ice for 15 min and centrifuged at 300*g* and 4°C for 5 min. The pellet was stored, and the supernatant was centrifuged at 10,000*g* and 4°C for 10 min. The resulting supernatant is a cytoplasmic cell extract. Then, both the pellets from the previous stages were resuspended in equal volumes of 20 mM HEPES-KOH buffer, pH 7.6, containing 5 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, and the protease inhibitor mix and homogenized with a Teflon pestle. The lysates were centrifuged at 16,000*g* and 4°C for 15 min to prepare nuclear (NE-1) and mitochondrial extracts, respectively, used in further experiments. To prepare the nuclear extract NE-2, the pellet of the NE-1 extract was resuspended in an equal volume of 20 mM HEPES-KOH buffer, pH 7.6, containing 400 mM NaCl, 1 mM DTT, and protease inhibitor mix, vigorously shaken for 30 sec, and centrifuged as described above. All four supernatants were conserved by addition of equal volumes of 10% glycerol as a cryoprotectant and stored at −70°C before use. Protein was determined by the Bradford method using a commercial reagent (Bio-Rad Laboratories).

Determination of repair initiation rate in nuclear extracts. The reaction was carried out in 20 μl (total volume) of 20 mM HEPES-KOH buffer, pH 7.5, containing 0.1 mM EDTA, 50 mM KCl, 0.01% (w/v) BSA, 1 mM DTT, 20 nM dsODN substrate, and 0.025-0.5 μg/μl total protein at 37°C for 10-240 min. Then the reaction was terminated as described above and desalted on a Sephadex G-25 microcolumn equilibrated with 5 M urea containing 0.1% bromophenol blue. Reaction products were analyzed by electrophoresis followed by phosphorimaging. The reaction rate was determined from the slope of the initial linear segment on the plot of product concentration vs. time.

Reconstitution of base excision repair *in vitro*. To reconstitute oxoG removal from DNA by the complete BER system, the reaction was carried out in 20 μl (total volume) of 50 mM Tris-HCl buffer, pH 7.8, containing 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, either 0.25 mM dNTP mix or 1 mM dGTP, 100 nM OGG1, 100 nM APEX1, 100 nM POLβ, and 5 U of T4 DNA ligase. If necessary, enzymes were preliminary diluted in 25 mM Tris-HCl buffer, pH 7.8, containing 5 mM $MgCl_2$, 5 mM DTT, 0.5 mM ATP, 0.05% BSA, and 5% (v/v) glycerol. The reaction was carried out at 37°C for 30 min and terminated by addition of 20 μl of solution containing 80% formamide, 20 mM Na-EDTA, 0.1% (w/v) xylene cyanole, and 0.1% (w/v) bromophenol blue followed by heating at 95°C for 5 min. Reaction products were analyzed by electrophoresis followed by phosphorimaging.

RESULTS

Kinetic parameters of oxoG excision from model CAG-substrates by DNA glycosylase OGG1. Initially, we determined kinetic parameters describing the interaction between OGG1 and substrates modeling CAG runs. These substrates were 29-bp dsODNs containing seven CAG repeats and flanking 4-bp sequences. One of the G-bases was replaced for oxoG in each substrate (see "Materials and Methods"). OGG1 binds to DNA asymmetrically by forming bonds with three phosphate groups on the 3′-side of the damage and with one phosphate group on the 5′-side and partially shielding several other base pairs from the solvent; totally the enzyme binds ~9-10 bp [21, 22]. OGG1 cleaves substrates with oxoG at the fifth position from any of the ends and those with central location of oxoG with similar efficiency [14]. So, one can surmise that substrates with oxoG in repeats 1 and 2 (CAG1//CAGcomp and CAG2//CAGcomp) are the model of oxidative damage on the 5′-border of the CAG run, substrates with oxoG in repeats 6 and 7 (CAG6//CAGcomp and CAG7//CAGcomp) are the model of the damage on the 3′-border, and the other substrates are the model of the damage in the middle of the run. The difference is that in the first case the enzyme interacts with non-CAG sequences with the part located on the 5′-side of the damage, in the second one on the 3′-side, while in the third case the interaction area only encompasses CAG trinucleotides.

In the absence of other protein factors, BER catalyzed by OGG1 occurs via formation of a kinetically stable complex of the enzyme with AP-product (EP) following excision of oxoG (scheme (1), $k_2 >> k_3$):

$$E \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} ES \xrightarrow{k_3} E + P. \quad (1)$$

For this kinetic scheme the Michaelis constant can be determined formally, but it represents a combination of rate constants that gives little information on reaction mechanism. Nevertheless, for this mechanism, the rate constants k_2 and k_3 can be determined separately. The first describes the reaction stage preceding the rate-limiting stage (in the case of OGG1 – hydrolysis of *N*-glycosidic bond), while the second one gives the rate-limiting stage (β-elimination and following release of product).

We examined the dynamics of the product accumulation under single-turnover conditions (concentration of substrate in the reaction mixture is far less than that of the enzyme) and determined k_2 values (rate constants of oxoG excision from ODN substrates by OGG1) (table). Maximum efficiency of repair initiation was observed when the damaged base was located in the middle of CAG run; excision of oxoG from the left- and rightmost triplets was less effective, especially when the lesion was at the 3′-end of the run.

Then, k_3 (the rate constant of DNA product release from the complex with OGG1 – see table) values were calculated for each substrate. For the reaction described by scheme (1), this constant is determined by the slope of the plot of product accumulation vs. time at substrate concentrations far exceeding concentration of the enzyme. A certain decrease in k_3 was observed upon the displacement of oxoG from 5′-end of the run to its 3′-end. Since drastic decrease in k_3 begins from CAG3 substrate and remains virtually unchanged up to CAG7, this suggests that tandem repeats on the 5′-side of the damaged base hinder either β-elimination or enzyme release from its complex with the product.

Kinetic parameters of oxoG excision from trinucleotide tandem repeats CAG and CGG

Substrate*	k_2, min^{-1}	k_3, min^{-1}
CAG1	2.3 ± 0.3	0.21 ± 0.03
CAG2	3.4 ± 1.1	0.18 ± 0.02
CAG3	3.3 ± 1.0	0.074 ± 0.012
CAG4	4.4 ± 1.4	0.050 ± 0.004
CAG5	8.2 ± 2.3	0.047 ± 0.007
CAG6	5.6 ± 0.9	0.057 ± 0.012
CAG7	0.06 ± 0.01	0.010 ± 0.004
CGG2-1	2.1 ± 0.2	0.11 ± 0.01
CGG2-2	3.0 ± 0.3	0.15 ± 0.02

* dsODN with denoted damaged strand and corresponding complementary strand.

To compare repair initiation efficacy in runs of CAG and other repeats, we determined kinetic parameters k_2 and k_3 for two substrates, CGG2-1 and CGG2-2, containing seven CGG triplets; in the latter either the first or the second G base was replaced by oxoG. CGG repeats are characteristic of fragile X syndrome (Martin–Bell syndrome; Q99.2 according to ICD-10) [2, 23]. The data show that parameters k_2 and k_3 of the reaction catalyzed by OGG1 are close for the substrates CAG2, CGG2-1, and CGG2-2 (table).

Repair initiation rate in cell extracts. Efficacy of BER initiation by DNA glycosylases can be modulated by other repair proteins, such as AP-endonucleases and many accessory enzymes and noncatalytic factors (XRCC1, PARP1, etc.) [7, 24]. In connection with this, we felt it necessary to examine 8-oxoguanine-DNA glycosylase activity towards DNA containing the damaged CAG runs in human cell extracts. The protocols we applied in the work for fractionating eukaryotic cells are widely used for preparation of extracts that are competent at distinct stages of replication, transcription, translation, and repair [16-20]. We prepared nuclear, mitochondrial, and cytoplasmic extracts from MRC-5 (human lung fibroblast) and LICH (human hepatoma) cells and tested them for ability to cleave the substrates CAG1-CAG7. Activity was found in all specimens with the greatest specific activity in nuclear extracts (Fig. 2, a-d). However, fractions prepared by extraction of nuclear pellet with low-ionic buffer (NE-1; 20 mM HEPES-KOH, 5 mM KCl, 1.5 mM $MgCl_2$) demonstrated a significant unspecific degradation of ODNs that was absent in fractions extracted with medium-ionic buffer (NE-2; 20 mM HEPES-KOH, 400 mM NaCl, 1 mM DTT) (Fig. 2, a-d).

The measured repair initiation rate of substrates CAG1-CAG7 by nuclear extract NE-2 from LICH cells is shown on the Fig. 2e. The initiation efficacy was low when the damaged base was located near the 3′- or 5′-end of the run and increased when oxoG was nearer the middle (triplets 4 and 5). This activity profile resembles the dependence of the rate constant of damaged base excision (k_2) by the purified OGG1 on damage position in the run (table and Fig. 2e), but it differs significantly from the dependence of rate constant k_3, which decreases from the 5′- to the 3′-end of the run. Thus, it is likely that the cleavage rate of oxoG-containing DNA in cell extract is limited by the catalytic step of the reaction (as shown earlier for OGG1 activity in the presence of AP-endonuclease APEX1 [14]) rather than by OGG1 release from the complex with the product, as in the case of pure protein. This is also evident from the character of products: the

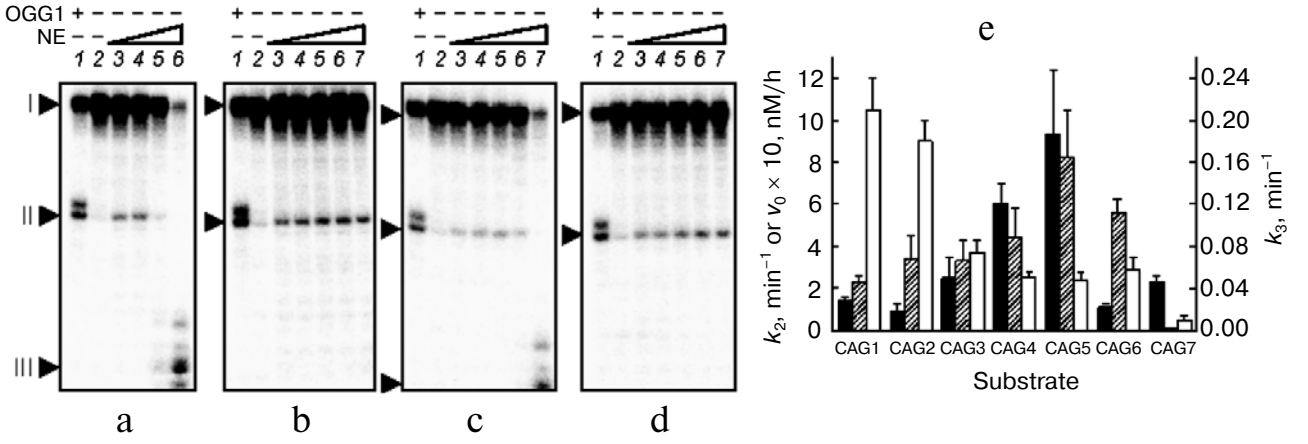

Fig. 2. Radioautography of electrophoretically separated products of CAG4 cleavage by nuclear extracts of MRC-5 (a, b) and LICH (c, d) cells. Data are shown for low-ionic-strength (NE-1) (a, c) and medium-ionic-strength (NE-2) (b, d) extracts. Arrows indicate mobility of ODN substrate (I), product of substrate cleavage at oxoG (II), and products of unspecific degradation by nucleases (III). In all cases the reaction mixture contained either 100 nM (lane *1*) or 0 nM (lanes *2-7*) recombinant OGG1 or 0.5, 1, 2, 5, and 10 μg nuclear extract (lanes *3-7*, respectively). All other reaction conditions are described in "Materials and Methods". The diagram (e) represents the rates of CAG1-CAG7 cleavage by the NE-2 extract of LICH cells (black columns) in comparison with the rate constants k_2 (shaded columns) and k_3 (white columns) of oxoG excision from DNA by OGG1.

recombinant OGG1 cleaves DNA with formation of β-elimination and β,δ-elimination product mixture (double band after electrophoresis, Fig. 2, a-d, lanes *1*), whereas the product of β-elimination does not accumulate in the extract because of the 3′-phosphodiesterase activity of APEX1 (Fig. 2, a-d, lanes *3-7*).

Reconstitution of base excision repair *in vitro*. To determine the ability of the minimal BER enzyme set to repair oxoG in CAG runs, we reconstituted the BER process *in vitro* with human DNA-glycosylase OGG1, AP-endonuclease APEX1, and DNA-polymerase β. The ligation ability of DNA-polymerase reaction products was tested using T4 DNA-ligase. When dGTP alone was used as a substrate for polymerase stage, BER exclusively occurred via the short-patch pathway (Fig. 3; the substrates CAG4-CAG7 are shown). Depending on the enzymes added into the reaction mixture, one can observe: the cleavage of ODN substrate at oxoG link by OGG1 (Fig. 3, lanes *3*; the product of β-elimination compared with that of β,δ-elimination catalyzed by Fpg enzyme is less mobile), further shortening of the product due to removal of 3′-terminal unsaturated pentose fragment by APEX1 (Fig. 3, lanes *4*), and elongation of repair product by one dGMP link by DNA-polymerase β (Fig. 3, lanes *5*). In the presence of T4 DNA-ligase the polymerase reaction product was able to restore the full-length ODN (Fig. 3, lanes *6*). Products with low electrophoretic mobility were also formed. It is likely that they are the ligation products of labeled blunt-ended dsODNs that are not associated with the repair process. The reconstitution efficiency determined by the amount of reaction product after the action of APEX1 or POLβ achieved a maximum when the damaged base was located in the middle of the run, and it decreased at its ends, which correlates with both the OGG1 activity profile in cell extracts and k_2 values of purified OGG1.

When the mixture of all dNTPs was taken instead of dGTP alone, POLβ incorporated not a single nucleotide, but fulfilled synthesis with strand displacement during the repair, because the length of polymerase reaction products ranged from the length of AP-endonuclease product through the length of initial ODN substrate (Fig. 4). The synthesis of full-length product effectively occurred for all substrates except for CAG1. In the latter case, it is likely that the POLβ–ODN complex dissociates after incorporation of the first dNMP, and the insufficiently long primer part of the ODN is unstable at the reaction temperature (37°C).

DISCUSSION

Microsatellites, also known as simple sequence repeats (SSRs) or short tandem repeats (STRs), are repeating sequences of 1-6 bp of DNA that comprise ~3% of the human genome with mean density of ~14,000 bp/Mb across all chromosomes [25]. In the human genome, the density of trinucleotide repeats is about threefold lower than that of di- and tetranucleotide motifs [25]. Distribution of microsatellites along the genome is not random: their localization and density are associated with their length and ability of forming stable secondary structures. Microsatellites can fulfill important biological functions. In particular, microsatellites located in promoters, non-translated regions, and introns are implicated in regulation of gene expression on the transcriptional

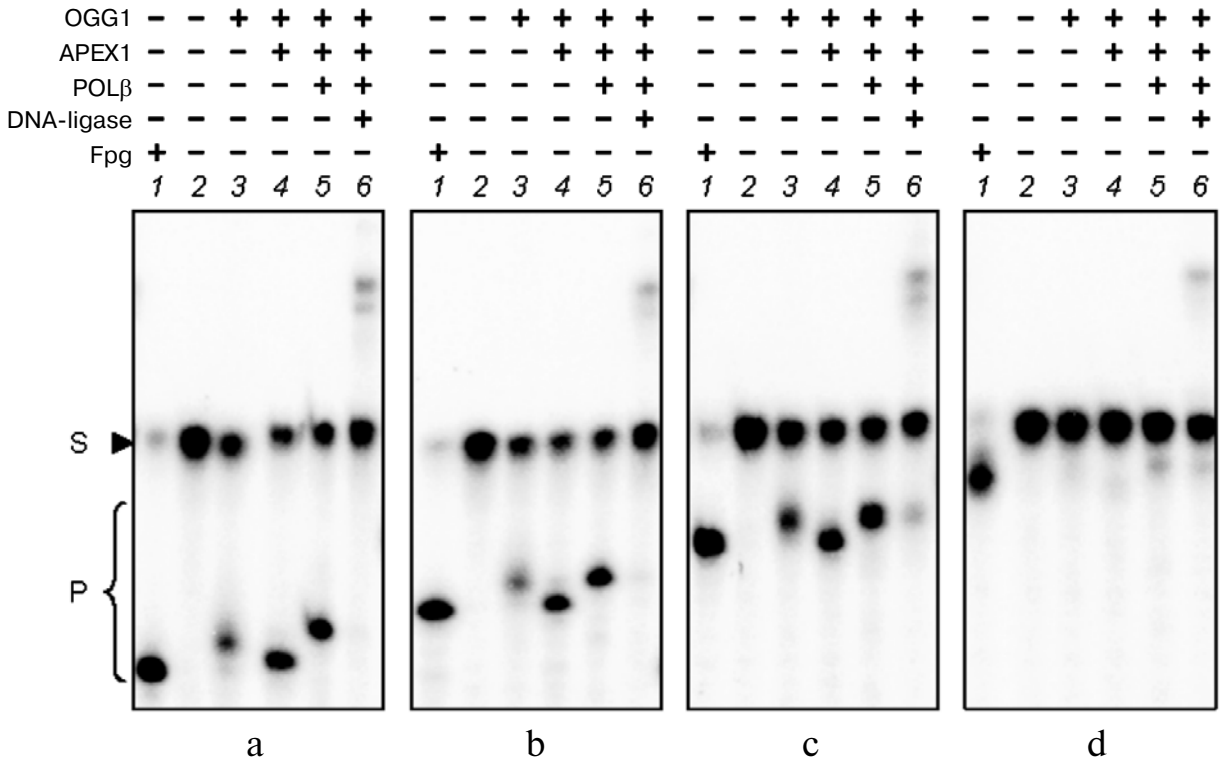

Fig. 3. Radioautography of electrophoretically separated products of reconstituted short-patch BER with enzymes OGG1, APEX1, POLβ, and T4 DNA-ligase. S – dsODNs: a) CAG4; b) CAG5; c) CAG6; d) CAG7; P – BER intermediate products of different length (see Fig. 1 and explanations in the text).

level, RNA stability, splicing, and RNA–protein interactions [26, 27]. Microsatellites located in coding regions of genes lead to appearance of amino acid or short peptide repeats; these repeats are important for the function of proteins implicated in development and nervous system function, structural proteins of extracellular matrix, etc. [28].

Expansion of trinucleotide repeat runs in some genes is the major cause of severe genetic disorders [2, 27]. These disorders are characterized by genetic anticipation, that is, the age of onset tends to decrease and severity to increase is each following generation in a family with the mutant gene. This is because replication of the run in germ line cells leads to its elongation, resulting in either alteration in function of the encoded protein (when the run is localized in an exon) or misregulation of gene expression (when the run is localized in a noncoding region). Recently, it was found that a dramatic expansion of trinucleotide repeat runs can also occur in somatic cells, particularly in the cells affected in the disease. For instance, expansion of CAG runs in Huntington's disease occurs in striatum before the onset of the disease [5]. In murine line R6/1 containing >100 CAG-repeats in the *HTT* gene, the length of the run varies to a small extent during the life of an individual in most tissues except for striatum and cortex, and in the latter quick and irreversible expansion is observed [29, 30]. In model mouse line Hdh^{Q111}, the CAG run in the *HTT* gene expands in striatum and liver [31]. Instability of the run length in non-dividing somatic cells was also reported for CGG runs in fragile X syndrome [32, 33] and GAA runs in Friedreich's ataxia [34, 35].

Since expansion of trinucleotide repeats cannot be associated with replication in terminally differentiated neurons, it is supposed that it is initiated by repair-associated DNA synthesis. In fact, expansion of CAG run in R6/1 mice requires activity of DNA glycosylase OGG1 [6]. A model was proposed in which BER initiated during repair of oxoG with implication of OGG1 occurs by the long-patch pathway with displacement of the forthcoming DNA strand, its folding into a hairpin, and ligating of this intermediate to elongate the strand containing oxoG [6, 9] (Fig. 1c). The appearing hairpin can be stabilized by proteins MSH2 and MSH3, which are participants of the mismatch repair system [6, 36, 37]. On the other hand, some reports do not agree with this model or, at least, do not allow its acceptance as a universal one. For example, Huntington's disease patients carrying the $OGG1^{326C}$ allele encoding the less active enzyme have mean length of CAG run in *HTT* gene not less (as one might expect), but more than patients homozygous at the wild-type allele $OGG1^{326S}$ [38]. This correlates with earli-

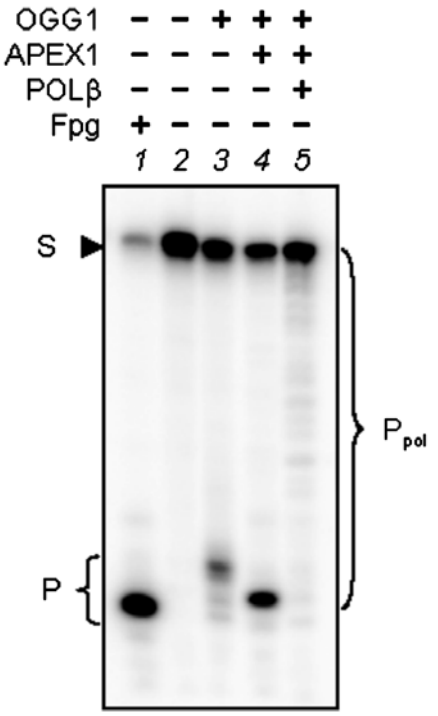

Fig. 4. Radioautography of electrophoretically separated repair products under reconstitution of long-patch BER with OGG1, APEX1, and POLβ. S – dsODN substrate CAG2; P – products of DNA-glycosylase and AP-endonuclease reactions catalyzed by OGG1, Fpg, and APEX1 enzymes; P_{pol} – partially or completely elongated products of POLβ-catalyzed reaction.

er onset of the disease in *OGG1*^{326C} carriers [38]. In human cell culture, a decrease in *OGG1* or *APEX1* expression by RNA interference does not cause a tendency to run shortening [39]. In a murine model of fragile X syndrome, the oxidant $KBrO_3$, which causes appearance of oxoG base in DNA under chronic alimentary introduction, induced hereditary expansion of CGG runs, but not additional somatic expansion [33]. Thus, it is likely that the expansion initiation efficacy upon oxoG repair involving OGG1 can vary depending on the composition of the trinucleotide repeat, position of damaged base in the run, and the presence or absence of other enzymes of the BER system.

The change in values of k_2 and k_3 constants for the reaction catalyzed by OGG1 was the most prominent consequence of altered oxoG position in the run (scheme (1), Fig. 2e, and table). These values usually differed for substrates with damaged CAG-triplets located on the run borders and within the run; k_2 achieved the maximum within the run, and k_3 was maximal at the 5′-end and minimal at the 3′-end of the run. Kinetic parameters of OGG1 are known to depend on local DNA conformational and thermodynamic parameters; for instance, the angle of DNA double helix twist (conformational parameter *twist* [40]) and ΔG° of duplex melting [41]. Such influence is explained by substantial alterations that DNA structure undergoes upon OGG1 binding – a kink of double helix axis and local DNA melting occur [21]. As evident from X-ray structural data, the area of the enzyme–DNA contact is asymmetric [21]: different regions of the OGG1 molecule should contact trinucleotides of the run when the damaged base is localized near 5′-end, 3′-end, or in the middle of the run. Hence, one would anticipate that the difference, if one exists, between local conformational and thermodynamic parameters of DNA in CAG-tracts and those in oligo(T) and oligo(A) flanking sequences in the used substrates, can explain, in part, the variable OGG1 activity in relation to them.

Fifty calculated conformational and thermodynamic parameters of dsODN whose sequence was identical to those of CAG1-CAG7 substrates except for oxoG being substituted for G were analyzed using the DNAtool software [42]. Drastic change in three parameters was observed on the border between the run and flanking sequences: 1) the angle of DNA axis kink [43] (Fig. 5a); 2) DNA rigidity (Young's modulus *E* calculated from the anisotropic elastic rod bending model [44], Fig. 5b); 3) dihedral angle between the planes of two adjacent bases (conformational parameter *tilt* [40] calculated by several different methods, Fig. 5c). For comparison, the same parameters were calculated for random sequences, either of the same nucleotide composition as CAG-substrates or with equal contents of canonical DNA bases. The values of the angle of DNA axis kink and DNA rigidity inside the CAG-tract were far lower than their average values for random sequence, and on the borders between the run and flanking sites they were much higher than their average (Fig. 5, a and b). The calculated tilt values in the run were within the range of average values for random sequence, and in flanking sites they were substantially higher (in the 5′-region) or lower than the average (in the 3′-region; Fig. 5c). The binding of OGG1 with DNA is known to result in substantial kink of the latter, achieving ~70° [21], and this kink changes during the process of recognition and excision of the damaged base [21, 45-50]. This suggests that the difference in k_2 values within the CAG-tract and on its borders is due to the fact that these DNA regions have different rigidity and different initial kink angles. The constant k_3 characterizes the process of β-elimination with subsequent release of DNA product from the complex with OGG1. Changes in the enzyme–DNA complex structure, if present during these events, are still not understood, but the influence of local DNA conformation, including *tilt* parameter, cannot be excluded.

Thus, we have shown that the repair initiation efficacy of oxoG base located in trinucleotide repeat runs depends on position of the damaged base in the run, namely, lowered on its 5′- and 3′-borders. This dependence concerns both the rate constant of oxoG excision

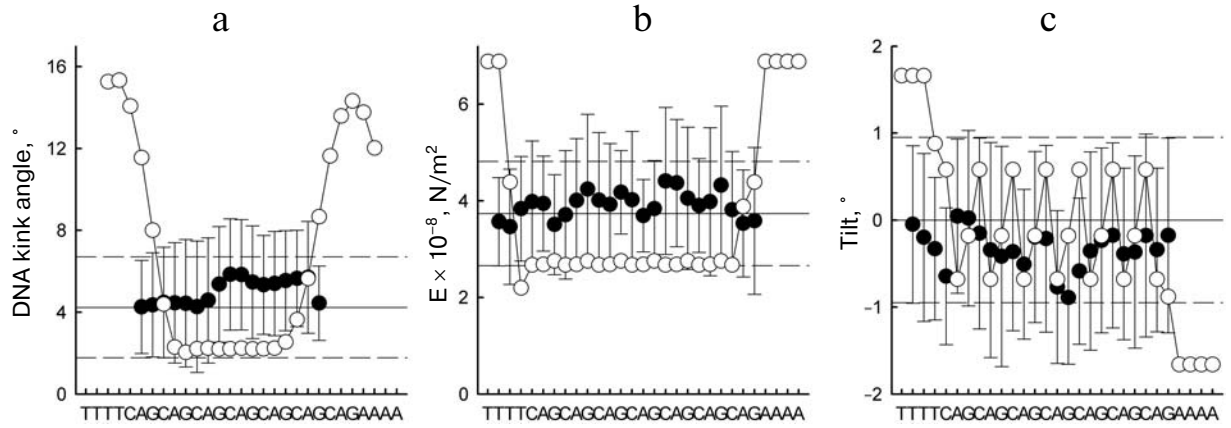

Fig. 5. Calculated values of the angle of DNA axis kink (a), Young's modulus for DNA rigidity (b), and conformational parameter *tilt* calculated from all known spatial structures of DNA B-conformation [51] (c). Data are presented for non-damaged DNA with sequence identical to that of substrates CAG1-CAG7 (open circles), the mean and standard deviation for 20 random sequences with the same nucleotide composition (closed circles), and the mean (solid line) and standard deviation (dashed line) for 20 random sequences with equal nucleotide contents.

by pure DNA-glycosylase OGG1 and the rate of cleavage of damaged DNA in cell extracts and efficacy of initial repair stages in the reconstituted BER system. So, long runs can expand easier than the short ones, not only because of more probable appearance of the damage, but also due to more effective repair within the run. As we have shown in this work, DNA polymerase β can effectively fulfill synthesis with strand displacement in the context of CAG runs. Because of lowered level of FEN1 endonuclease in striatum cells [9], this can provide synthesis of a long hanging single-stranded fragment capable of stable hairpin formation and, as a result, lead to elongation of the DNA strand under repair. The system of substrates and enzymes we have developed here for BER reconstitution can be used for estimating effectiveness of substances inhibiting distinct BER stages as promising therapeutic drugs, which decrease the risk of trinucleotide repeat expansion in somatic cells and thus prevent progression of Huntington's disease.

The authors are grateful to G. A. Maksakova (Institute of Chemical Biology and Fundamental Medicine, Siberian Division of the Russian Academy of Sciences) for synthesis of oligonucleotides used in the study.

This study was supported by grants of the Presidium of the Russian Academy of Sciences (program "Molecular and Cell Biology", No. 6.14), the Russian Federation Ministry of Education and Science (No. 14.740.11.1195), the Russian Foundation for Basic Research (10-04-91058-PICS_a), the National Research Center of France (PICS No. 5479 – Russie), and the Fondation de France (No. 12091).

REFERENCES

1. Perutz, M. F., and Windle, A. H. (2001) *Nature*, **412**, 143-144.
2. Orr, H. T., and Zoghbi, H. Y. (2007) *Annu. Rev. Neurosci.*, **30**, 575-621.
3. Duyao, M., Ambrose, C., Myers, R., Novelletto, A., Persichetti, F., Frontali, M., Folstein, S., Ross, C., Franz, M., Abbott, M., Gray, J., Conneally, P., Young, A., Penney, J., Hollingsworth, Z., Shoulson, I., Lazzarini, A., Falek, A., Koroshetz, W., Sax, D., Bird, E., Vonsattel, J., Bonilla, E., Alvir, J., Bickham Conde, J., Cha, J.-H., Dure, L., Gomez, F., Ramos, M., Sanchez-Ramos, J., Snodgrass, S., de Young, M., Wexler, N., Moscowitz, C., Penchaszadeh, G., MacFarlane, H., Anderson, M., Jenkins, B., Srinidhi, J., Gusella, G. B. J., and MacDonald, M. (1993) *Nat. Genet.*, **4**, 387-392.
4. Walker, F. O. (2007) *Lancet*, **369**, 218-228.
5. Kennedy, L., Evans, E., Chen, C.-M., Craven, L., Detloff, P. J., Ennis, M., and Shelbourne, P. F. (2003) *Hum. Mol. Genet.*, **12**, 3359-3367.
6. Kovtun, I. V., Liu, Y., Bjoras, M., Klungland, A., Wilson, S. H., and McMurray, C. T. (2007) *Nature*, **447**, 447-452.
7. Zharkov, D. O. (2008) *Cell. Mol. Life Sci.*, **65**, 1544-1565.
8. Liu, Y., Prasad, R., Beard, W. A., Hou, E. W., Horton, J. K., McMurray, C. T., and Wilson, S. H. (2009) *J. Biol. Chem.*, **284**, 28352-28366.
9. Goula, A.-V., Berquist, B. R., Wilson, D. M., III, Wheeler, V. C., Trottier, Y., and Merienne, K. (2009) *PLoS Genet.*, **5**, e1000749.
10. Sokhansanj, B. A., Rodrigue, G. R., Fitch, J. P., and Wilson, D. M., III (2002) *Nucleic Acids Res.*, **30**, 1817-1825.
11. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
12. Gilboa, R., Zharkov, D. O., Golan, G., Fernandes, A. S., Gerchman, S. E., Matz, E., Kycia, J. H., Grollman, A. P., and Shoham, G. (2002) *J. Biol. Chem.*, **277**, 19811-19816.

13. Kuznetsov, N. A., Koval, V. V., Zharkov, D. O., Nevinsky, G. A., Douglas, K. T., and Fedorova, O. S. (2005) *Nucleic Acids Res.*, **33**, 3919-3931.
14. Sidorenko, V. S., Nevinsky, G. A., and Zharkov, D. O. (2007) *DNA Repair*, **6**, 317-328.
15. Beard, W. A., and Wilson, S. H. (1995) *Meth. Enzymol.*, **262**, 98-107.
16. Dignam, J. D. (1990) *Meth. Enzymol.*, **182**, 194-203.
17. Stierum, R. H., Dianov, G. L., and Bohr, V. A. (1999) *Nucleic Acids Res.*, **27**, 3712-3719.
18. Sinitsyna, O., Krysanova, Z., Ishchenko, A., Dikalova, A. E., Stolyarov, S., Kolosova, N., Vasunina, E., and Nevinsky, G. (2006) *J. Cell. Mol. Med.*, **10**, 206-215.
19. Visnes, T., Akbari, M., Hagen, L., Slupphaug, G., and Krokan, H. E. (2008) *DNA Repair*, **7**, 1869-1881.
20. Akbari, M., and Krokan, H. E. (2012) *Mutat. Res.*, in press.
21. Bruner, S. D., Norman, D. P. G., and Verdine, G. L. (2000) *Nature*, **403**, 859-866.
22. Kirpota, O. O., Zharkov, D. O., Buneva, V. N., and Nevinsky, G. A. (2006) *Mol. Biol.*, **40**, 952-960.
23. Garber, K. B., Visootsak, J., and Warren, S. T. (2008) *Eur. J. Hum. Genet.*, **16**, 666-672.
24. Hang, B., and Singer, B. (2003) *Chem. Res. Toxicol.*, **16**, 1181-1195.
25. Subramanian, S., Mishra, R. K., and Singh, L. (2003) *Genome Biol.*, **4**:R13.
26. Li, Y.-C., Korol, A. B., Fahima, T., Beiles, A., and Nevo, E. (2002) *Mol. Ecol.*, **11**, 2453-2465.
27. Li, Y.-C., Korol, A. B., Fahima, T., and Nevo, E. (2004) *Mol. Biol. Evol.*, **21**, 991-1007.
28. Jorda, J., and Kajava, A. V. (2010) *Adv. Protein Chem. Struct. Biol.*, **79**, 59-88.
29. Gonitel, R., Moffitt, H., Sathasivam, K., Woodman, B., Detloff, P. J., Faull, R. L. M., and Bates, G. P. (2008) *Proc. Natl. Acad. Sci. USA*, **105**, 3467-3472.
30. Mollersen, L., Rowe, A. D., Larsen, E., Rognes, T., and Klungland, A. (2010) *PLoS Genet.*, **6**, e1001242.
31. Lee, J.-M., Zhang, J., Su, A. I., Walker, J. R., Wiltshire, T., Kang, K., Dragileva, E., Gillis, T., Lopez, E. T., Boily, M.-J., Cyr, M., Kohane, I., Gusella, J. F., MacDonald, M. E., and Wheeler, V. C. (2010) *BMC Syst. Biol.*, **4**, 29.
32. Entezam, A., and Usdin, K. (2008) *Nucleic Acids Res.*, **36**, 1050-1056.
33. Entezam, A., Lokanga, A. R., Le, W., Hoffman, G., and Usdin, K. (2010) *Hum. Mutat.*, **31**, 611-616.
34. Sharma, R., Bhatti, S., Gomez, M., Clark, R. M., Murray, C., Ashizawa, T., and Bidichandani, S. I. (2002) *Hum. Mol. Genet.*, **11**, 2175-2187.
35. Clark, R. M., De Biase, I., Malykhina, A. P., Al-Mahdawi, S., Pook, M., and Bidichandani, S. I. (2007) *Hum. Genet.*, **120**, 633-640.
36. Manley, K., Shirley, T. L., Flaherty, L., and Messer, A. (1999) *Nat. Genet.*, **23**, 471-473.
37. Owen, B. A. L., Yang, Z., Lai, M., Gajek, M., Badger, J. D., III, Hayes, J. J., Edelmann, W., Kucherlapati, R., Wilson, T. M., and McMurray, C. T. (2005) *Nat. Struct. Mol. Biol.*, **12**, 663-670.
38. Coppede, F., Migheli, F., Ceravolo, R., Bregant, E., Rocchi, A., Petrozzi, L., Unti, E., Lonigro, R., Siciliano, G., and Migliore, L. (2010) *Toxicology*, **278**, 199-203.
39. Lin, Y., and Wilson, J. H. (2007) *Mol. Cell. Biol.*, **27**, 6209-6217.
40. Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H. C. M., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C.-S., Wang, A. H.-J., and Zhurkin, V. B. (1989) *Nucleic Acids Res.*, **17**, 1797-1803.
41. Kirpota, O. O., Endutkin, A. V., Ponomarenko, M. P., Ponomarenko, P. M., Zharkov, D. O., and Nevinsky, G. A. (2011) *Nucleic Acids Res.*, **39**, 4836-4850.
42. Vlahovicek, K., Kajan, L., and Pongor, S. (2003) *Nucleic Acids Res.*, **31**, 3686-3687.
43. Gabrielian, A., Vlahovicek, K., and Pongor, S. (1997) *FEBS Lett.*, **406**, 69-74.
44. Gromiha, M. M., Munteanu, M. G., Gabrielian, A., and Pongor, S. (1996) *J. Biol. Phys.*, **22**, 227-243.
45. Banerjee, A., Yang, W., Karplus, M., and Verdine, G. L. (2005) *Nature*, **434**, 612-618.
46. Banerjee, A., and Verdine, G. L. (2006) *Proc. Natl. Acad. Sci. USA*, **103**, 15020-15025.
47. Lee, S., Radom, C. T., and Verdine, G. L. (2008) *J. Am. Chem. Soc.*, **130**, 7784-7785.
48. Fromme, J. C., Bruner, S. D., Yang, W., Karplus, M., and Verdine, G. L. (2003) *Nat. Struct. Biol.*, **10**, 204-211.
49. Norman, D. P. G., Bruner, S. D., and Verdine, G. L. (2001) *J. Am. Chem. Soc.*, **123**, 359-360.
50. Chung, S. J., and Verdine, G. L. (2004) *Chem. Biol.*, **11**, 1643-1649.
51. Bansal, M., Bhattacharyya, D., and Ravi, B. (1995) *Comput. Appl. Biosci.*, **11**, 281-287.