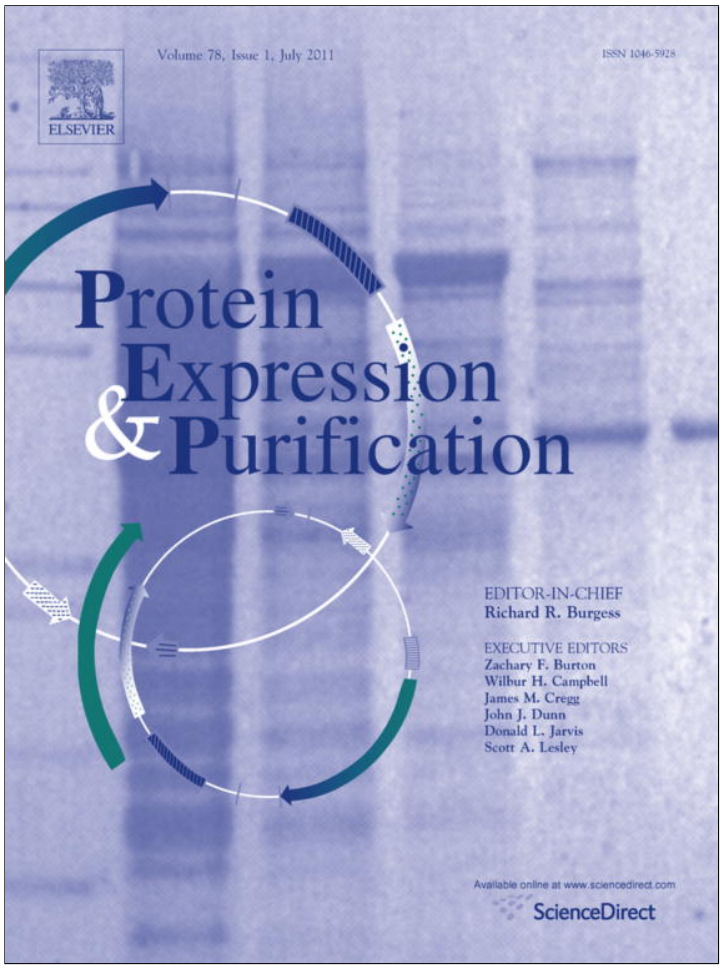

Protein Expression and Purification 78 (2011) 94–101

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Stable isotope-labeling of DNA repair proteins, and their purification and characterization

Prasad T. Reddy [a,*], Pawel Jaruga [a], Bryant C. Nelson [a], Mark Lowenthal [b], Miral Dizdaroglu [a,*]

[a] Biochemical Science Division, National Institute of Standards and Technology, Gaithersburg, MD, USA
[b] Analytical Chemistry Division, National Institute of Standards and Technology, Gaithersburg, MD, USA

ARTICLE INFO

Article history:
Received 7 January 2011
and in revised form 18 February 2011
Available online 26 February 2011

Keywords:
DNA repair
DNA glycosylases
Isotope-dilution
Mass spectrometry
OGG1

ABSTRACT

Reduced DNA repair capacity is associated with increased risk for a variety of disease processes including carcinogenesis. Thus, DNA repair proteins have the potential to be used as important predictive, prognostic and therapeutic biomarkers in cancer and other diseases. The measurement of the expression level of these enzymes may be an excellent tool for this purpose. Mass spectrometry is becoming the technique of choice for the identification and quantification of proteins. However, suitable internal standards must be used to ensure the precision and accuracy of measurements. An ideal internal standard in this case would be a stable isotope-labeled analog of the analyte protein. In the present work, we over-expressed, purified and characterized two stable isotope-labeled DNA glycosylases, i.e., ^{15}N-labeled *Escherichia coli* formamidopyrimidine DNA glycosylase (Fpg) and ^{15}N-labeled human 8-oxoguanine-DNA glycosylase (hOGG1). DNA glycosylases are involved in the first step of the base excision repair of oxidatively induced DNA damage by removing modified DNA bases. The measurement by MALDI-ToF mass spectrometry of the molecular mass and isotopic purity proved the identity of the ^{15}N-labeled proteins and showed that the ^{15}N-labeling of both proteins was more than 99.7%. We also measured the DNA glycosylase activities using gas chromatography/mass spectrometry with isotope-dilution. The enzymic activities of both ^{15}N-labeled Fpg and ^{15}N-labeled hOGG1 were essentially identical to those of their respective unlabeled counterparts, ascertaining that the labeling did not perturb their catalytic sites. The procedures described in this work may be used for obtaining stable isotope-labeled analogs of other DNA repair proteins for mass spectrometric measurements of these proteins as disease biomarkers.

Published by Elsevier Inc.

Exogenous and endogenous sources generate oxidatively induced DNA damage with a plethora of lesions [1]. This type of DNA damage is thought to play an important role in disease processes such as carcinogenesis and aging [2]. Elaborate repair pathways exist in living organisms that repair DNA damage. DNA repair is critical for maintaining the genomic stability and thus for preventing disease development including carcinogenesis [2,3]. Oxidatively induced DNA damage is mainly repaired by base excision repair (BER),[1] and also by nucleotide excision repair (NER), albeit to a lesser extent [2]. In the first step of BER, DNA glycosylases hydrolyze the N-glycosidic bond releasing the damaged base and generating an abasic site, followed by the action of a series of other BER enzymes [4,5]. Formamidopyrimidine DNA glycosylase (Fpg, also called MutM) is one of the main DNA glycosylases in *Escherichia coli*, which specifically excises 4,6-diamino-5-formamidopyrimidine (FapyAde), 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyGua) and 8-hydroxyguanine (8-OH-Gua) from DNA containing multiple lesions [6,7]. In eukaryotes, 8-oxoguanine-DNA glycosylase (OGG1), which is a functional homolog of Fpg, exhibits a strong specificity for excision of FapyGua and 8-OH-Gua, but does not act on FapyAde [8–10].

Reduced DNA repair capacity appears to be associated with increased risk for a variety of cancers [3,11–13]. Thus, DNA repair proteins are emerging as important predictive, prognostic and therapeutic factors in cancer [14]. The measurement of the expression and level of DNA repair proteins may be an important tool for determining the risk for the development of cancer. Furthermore, this may lead to the use of DNA repair proteins as cancer and other disease biomarkers. Expression of DNA repair proteins have been measured mainly by real time quantitative PCR. The use of the reverse-protein microarray assay has also been reported to be a useful tool for such measurements [15]. Liquid chromatography–tandem mass spectrometry (LC–MS/MS) is generally used for separation, positive identification and

* Corresponding authors. Fax: +1 301 975 8505.
E-mail addresses: prasad.reddy@nist.gov (P.T. Reddy), miral@nist.gov (M. Dizdaroglu).

[1] *Abbreviations used:* hOGG1, human 8-oxoguanine-DNA glycosylase; BER, base excision repair; NER, nucleotide excision repair; FapyAde, 4,6-diamino-5-formamidopyrimidine; FapyGua, 2,6-diamino-4-hydroxy-5-formamidopyrimidine; 8-OH-Gua, 8-hydroxyguanine; LC–MS/MS, liquid chromatography–tandem mass spectrometry; GC/MS, gas chromatography/mass spectrometry.

1046-5928/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.pep.2011.02.011

quantification of peptides released from a protein by a proteolytic enzyme such as trypsin [16]. In order to ensure the high precision and accuracy of quantitative measurements in mass spectrometry, suitable internal standards must be used. An ideal internal standard for accurate quantification of a protein would be its analog fully labeled with stable isotopes such as ^{15}N and/or ^{13}C. The labeled whole protein will have identical chemical and physical properties as the analyte protein, thus compensating for losses that may occur during all stages of the analysis. Trypsin digestion can often be inefficient, leading to incomplete yields of tryptic peptides and consequently to potential measurement bias. Therefore, the fully labeled analog of the analyte protein is essential as an internal standard for accurate measurements of tryptic peptides resulting from trypsin digestion, which is performed prior to analysis by LC–MS/MS. A single labeled tryptic peptide will not meet these important requirements.

In the present study, we aimed to develop methods for the overexpression, purification and characterization of stable isotope-labeled DNA repair proteins. We chose Fpg and hOGG1 as examples. We over-expressed, isolated and purified ^{15}N-labeled Fpg and ^{15}N-labeled hOGG1. In addition, we determined the isotopic purity of these labeled proteins and measured their DNA glycosylase activities to ascertain that the labeling does not perturb the respective catalytic site.

Materials and methods

Materials

Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase, and Calf intestinal alkaline phosphatase were purchased from New England Biolabs (Beverly, MA). Acrylamide, bisacrylamide, protease inhibitor cocktail tablets, and calf thymus DNA were obtained from Sigma–Aldrich (St. Louis, MO). Deoxynucleoside triphosphates were purchased from Pharmacia LKB Biotechnology Inc. (Piscataway, NJ). SeaKem GTG Agarose and NuSieve GTG Agarose were purchased from FMC BioProducts (Rockland, ME). Nickel agarose resin, PCR purification kit, gel extraction kit, and plasmid purification kit were from Qiagen (Valencia, CA). Diethylaminoethyl (DEAE) cellulose (DE52) was from Whatman Inc. (Clifton, NJ). Shodex carboxymethyl cellulose HPLC preparative column (2 cm × 20 cm) was from Phenomenex (Torrance, CA). Dye-dideoxy terminator cycle sequencing kits were purchased from Applied Biosystems Inc., Perkin–Elmer Cetus (Foster City, CA). Modified DNA bases, their stable isotope-labeled analogs and other materials for gas chromatography/mass spectrometry (GC/MS) were obtained as described [17]. Oligodeoxynucleotides were from Operon (Huntsville, AL).

E. coli strains, plasmids and preparation of minimal medium

The pET11a expression vector, which was used for the native protein production of Fpg based on the IPTG induction, was obtained from Novagen (Madison, WI). The relevant strains for cloning and expression were *E. coli* Novablue (K12) and BL21(DE3), respectively. The *hogg1* cDNA sub-cloned into the expression vector pET15b for His-tagged production of the protein was kindly provided by Dr. Dmitry Zharkov (SB RAS Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia). Minimal medium was prepared as described [18]. The composition of the medium was: 6 g NaH_2PO_4, 3 g K_2HPO_4, 0.5 g NaCl, and 1 g ^{15}N-NH_4Cl, 5 g glucose, 246 mg $MgSO_4{\cdot}7H_2O$ per L. Ampicillin was added to a final concentration of 50 μg/mL.

DNA procedures

E. coli Novablue (K12) harboring a recombinant plasmid was grown at 37 °C overnight in 10 mL Luria–Bertani medium containing ampicillin (50 μg/mL) [19]. Minipreparations of plasmid DNA were purified using a Qiagen kit. Digestion of DNA with restriction enzymes was performed according to the manufacturer's recommendation. DNA fragments were separated by electrophoresis on SeaKem GTG agarose or NuSieve GTG agarose. DNA was purified using a Qiagen kit. Ligation of DNA fragments and transformation of the *E. coli* strains were performed as described [20].

Cloning of Fpg into pET11a vector, and production and purification of ^{15}N-labeled Fpg

The coding sequence for *E. coli* Fpg gene from the plasmid pFpg, gift from Dr. Timothy R. O'Connor (Beckman Research Institute of the City of Hope, Duarte, CA), was amplified by PCR with Pfu polymerase using a 5′ end primer (5′-GGAATTC CAT ATG CCT GAA TTA CCC GAA G-3′) and a 3′ end primer (5′-CCG CTC GAG TTA CTT CTG GCA CTG CCG AC-3′). The 5′ end primer sequence contained the NdeI restriction recognition sequence CATATG wherein ATG served as the initiation codon for protein expression. The 3′ end primer contained the TAA translation stop codon. The amplified product was digested with NdeI to produce the gene with 5′ NdeI protruding end and 3′ blunt end. The DNA fragment was purified from a 1% agarose gel, and cloned into the NdeI and BamHI (first digested with BamHI, filled in with DNA polymerase to make this blunt end, then digested with NdeI) sites of the pET11a expression vector. A recombinant plasmid pET11a/Fpg was isolated from *E. coli* Nova Blue cells. *E. coli* BL21(DE3) was transformed with the recombinant plasmid to induce protein expression with IPTG. The Fpg portion of the plasmid was sequenced and found to have no mutations.

E. coli BL21(DE3) harboring pET11a/Fpg recombinant plasmid was grown at 37 °C for 20 h on LB agar plate containing 100 μg ampicillin/mL. A colony was carefully (without touching into the LB medium) transferred into 10 mL minimal medium containing 1 mg ^{15}N-NH_4Cl/mL and 50 μg ampicillin/mL. Cells were grown for 1 h at 37 °C at 250 RPM in a 50 mL tube. This inoculum was transferred to 140 mL minimal medium, in a 500 mL baffled flask, containing ^{15}N-NH_4Cl and ampicillin as above. This culture was grown at 37 °C for 16 h. Next, 3 × 50 mL of this seed culture was transferred to 3 × 450 mL of minimal medium containing ^{15}N-NH_4Cl and ampicillin in 3 × 2 L baffled flasks. IPTG was added to a final concentration of 30 μM to induce Fpg production at 37 °C for 1 h. Next, the culture was shifted to 42 °C for 5 h to continue Fpg production. Cells were harvested at 6000*g* for 20 min and washed with 25 mM Tris buffer (pH 7.5). The wet weight of cells obtained in this procedure was ~3 g.

The cell pellet was suspended in 40 mL of 50 mM Tris buffer (pH 8.0), 10 mM β-mercaptoethanol, 1 mM EDTA, and 50 mM NaCl (lysis buffer 1) containing one tablet of protease inhibitors. Cell suspension was passed through a French Press at 7×10^4 kPa. The cell-free extract was centrifuged at 10,000*g* for 20 min. The supernatant was mixed with 5 g of DE52 anion exchange resin equilibrated with the lysis buffer 1 in a 500 mL bottle at 80 RPM for 1 h at 4 °C and then poured into a column. The flow through containing nearly all the Fpg and fewer cellular proteins was collected. The resin was washed with 10 mL of the lysis buffer 1 and added to the flow through. The Fpg enriched pool (50 mL) was dialyzed overnight against 1 L of 20 mM K-phosphate buffer (pH 7.4). The dialyzed pool was centrifuged at 100,000*g* for 1 h to remove any particulate material. The supernatant fraction was chromatographed on a HPLC-Shodex carboxymethyl cellulose column (2.0 cm × 20 cm) equilibrated with 20 mM K-phosphate buffer

(pH 7.4). The column was washed with 100 mL of 20 mM K-phosphate buffer, pH 7.4 until the A_{280} stabilized. Then Fpg was eluted with a 0–0.5 M KCl gradient (250 mL each) in 20 mM K-phosphate buffer (pH 7.4). Pure Fpg was eluted as a sharp peak at ~0.3 M KCl. Fractions containing Fpg were pooled, concentrated on YM3 membrane filter, and dialyzed overnight against 500 mL of 50% glycerol, 20 mM Tris, 1 mM EDTA, and 10 mM β-mercaptoethanol (pH 7.5). Protein concentration was determined by the Lowry method using BSA as a standard [21].

Production and purification of ^{15}N-labeled hOGG1

E. coli BL21(DE3) harboring pET15b/hOGG1 recombinant plasmid was grown at 37 °C for 20 h on LB agar plate containing 100 μg ampicillin/mL. A colony was carefully (without touching into the LB medium) transferred into 100 mL minimal medium containing 1 mg ^{15}N-NH_4Cl/mL and 50 μg ampicillin/mL. Cells were grown overnight (16 h) at 37 °C at 250 RPM in a 250 mL baffled flask. This inoculum was transferred to 1000 mL minimal medium, in a 2 L baffled flask, containing ^{15}N-NH_4Cl and ampicillin as above. This culture was grown at 37 °C for 1 h and briefly cooled in ice water to room temperature. Next, IPTG was added to a final concentration of 100 μM to induce hOGG1 production at 24 °C overnight. Cells were harvested at 6000*g* for 20 min and washed with 25 mM Tris–HCl (pH 7.5). The wet weight of cells obtained in this procedure was ~2 g.

The cell pellet was suspended in 20 mL of 50 mM Tris–HCl, pH 8.0, 10 mM β-mercaptoethanol, 300 mM NaCl, and 10 mM imidazole (lysis buffer 2) containing one tablet of protease inhibitors. Cell suspension was passed through a French Press at 7×10^4 kPa. The cell-free extract was centrifuged at 40,000*g* for 1 h. Meanwhile, 2 mL of nickel-agarose slurry (1 mL resin) was washed with the lysis buffer 2 in a 30 mL Bio-Rad polypropylene column. The supernatant was added to the resin and mixed on a rocker for 2 h at 4 °C. The flow through was collected, and the column was washed successively with three 10 mL aliquots of the lysis buffer 2. Next, the resin was washed three times with 3 mL aliquots of the lysis buffer 2 containing an additional 20 mM imidazole (total 30 mM). Next, the hOGG1 was eluted with five 5 mL aliquots of the lysis buffer 2 containing 100 mM imidazole. The first 5 mL contained about 50% of hOGG1 (~95% pure) with minor contaminants as judged by SDS–PAGE. The subsequent four elutions contained the remaining hOGG1 with higher purity (~98%). All the five elutions containing hOGG1 were pooled, concentrated on YM3 membrane filter, and dialyzed overnight against 500 mL of 50% glycerol, 20 mM Tris, 1 mM EDTA, and 10 mM β-mercaptoethanol, pH 7.5. Protein concentration was determined by the Lowry method using BSA as standard [21].

Molecular mass determination of unlabeled and ^{15}N-labeled Fpg and hOGG1 by MALDI-ToF mass spectrometry

MALDI-ToF mass spectra were collected and analyzed using an Applied Biosystems 4700 Proteomics Analyzer (Foster City, CA) with a nitrogen laser (337 nm) operated in linear, high mass mode. Mass spectra were collected in positive ion mode using a source voltage of 20.0 kV. The grid voltage, linear detector voltage and low mass gate were set to 18.4 kV, 2.0 kV and 0.0 *m*/*z*, respectively. Each sample spectrum was obtained by averaging 3000 laser shots from three sequential sample accumulations. Three sample spectrums were acquired for each sample. The default instrument mass calibration was utilized for all data acquisitions.

For the MALDI-ToF MS Analysis of Fpg, the acquisition mass range, focus mass and laser intensity were set to 10,000 *m*/*z*–40,000 *m*/*z*, 30,000 *m*/*z* and 5800 (arbitrary units), respectively. The laser intensity mode, search pattern and search pattern source were set to fixed, uniform and random, respectively. The MALDI matrix solvent was 1:1 (volume fractions) water/acetonitrile in 0.1% trifluoroacetic acid. The MALDI matrix was a saturated solution (10 mg/mL) of sinapinic acid dissolved in matrix solvent. The ^{15}N-labeled (3 μg/μL) and unlabeled (5 μg/μL) Fpg samples were initially prepared in 50% glycerol, 20 mM Tris–HCl, 1 mM EDTA, 10 mM β-mercaptoethanol, pH 7.5 buffer. Each Fpg sample was prepared for MALDI-ToF MS analysis by premixing 10 μL of the sample with 90 μL of MALDI matrix. An aliquot (1 μL) of each premixed sample was deposited onto a stainless steel MALDI target. Each deposit was allowed to air dry at room temperature for 10 min before MALDI-ToF MS analysis.

The acquisition mass range, focus mass and laser intensity at 20,000 *m*/*z*–80,000 *m*/*z*, 40,000 *m*/*z* and 5800 (arbitrary units), respectively, were used for the MALDI-ToF MS Analysis of hOGG1. The laser intensity mode, search pattern and search pattern source were set to fixed, center bias and random, respectively. The MALDI matrix solvent was 1:1 (volume fractions) water/acetonitrile in 0.1% trifluoroacetic acid. The MALDI matrix was a saturated solution (20 mg/mL) of sinapinic acid dissolved in matrix solvent. The ^{15}N-labeled (1.5 μg/μL) and unlabeled (1.5 μg/μL) hOGG1 samples were prepared in a buffer (pH 7.5) containing 50 mM Tris–HCl, 100 mM NaCl, 100 mM imidazole, 100 mM β-mercaptoethanol. Each hOGG1 sample was initially desalted using Zeba™ Spin (7 kDa molecular mass cutoff) desalting columns (Thermo Scientific) following the column manufacturer's suggested protocols. Each sample was subsequently prepared for MALDI-ToF MS analysis by premixing 3 μL of sample with 3 μL of MALDI matrix. An aliquot (0.5 μL) of each premixed sample was deposited onto a stainless steel MALDI target. Each deposit was allowed to air dry at room temperature for 10 min before MALDI-ToF MS analysis.

Preparation of DNA samples, enzymic assays and gas chromatography/mass spectrometry

Calf thymus DNA was dissolved in 10 mM phosphate buffer (pH 7.4, 0.3 mg/mL) at 4 °C, saturated with N_2O and irradiated in a ^{60}Co γ-source at a dose of 5 Gy (dose rate 10 Gy/min). After dialysis against water for 18 h at 4 °C, aliquots of 50 μg of DNA samples were dried in a SpeedVac under vacuum. DNA samples (50 μg) were dissolved in 50 μL of an incubation buffer consisting of 50 mM phosphate buffer (pH 7.4), 100 mM KCl, 1 mM EDTA, and 0.1 mM dithiothreitol. Aliquots of FapyAde-^{13}C,$^{15}N_2$, FapyGua-^{13}C,$^{15}N_2$ and 8-OH-Gua-$^{15}N_5$ were added as internal standards. Samples were incubated with 2 μg of an unlabeled or labeled protein for 1 h at 37 °C in a water bath and then processed and analyzed by GC/MS for the identification and quantification of FapyAde, FapyGua and 8-OH-Gua released by Fpg and hOGG1 [17,22].

Results and discussion

Production and purification of ^{15}N-labeled E. coli Fpg

We have previously described an efficient production of Fpg and a method for its rapid purification [17]. The production protocol employed a heat inducible lambda P_L promoter based expression vector pRE [23], and an *E. coli* strain MZ1 that carries a temperature sensitive repressor [24]. This expression system was suitable for production of unlabeled Fpg by growing cells in LB medium. However, the pRE expression system is unsuitable for production of ^{15}N-labeled Fpg because the MZ1 strain has a deletion for *his ilv* biosynthetic pathway and growth in minimal medium containing ammonium chloride is not achievable. Hence, we sub-cloned *E. coli fpg* gene into an IPTG inducible native expression vector

pET11a that utilizes *E. coli* BL21(DE3) cells for protein production. Just a brief reminder that Fpg production in expression vectors such as pET15b that would result in addition of a His tag is detrimental to the activity because any modification at the amino terminus of Fpg will interfere with the catalytic function at the residue Pro_2 [17]. Hence, Fpg production in native expression vector was chosen as the correct path.

We carried out optimization of the production of Fpg in minimal medium with respect to IPTG concentration, time and temperature of induction. The following conditions were tested: induction at 20 °C, 24 °C, 37 °C, and 42 °C for various times (16 h, 32 h, 42 h) after addition of IPTG to 10 μM, 30 μM, 100 μM and 300 μM concentration. Although induction with 100 μM IPTG and growth at any of the temperatures and times produced ~20% of the total cellular protein as Fpg, only 2–3% represented the soluble Fpg and the remaining was in "inclusion bodies". After reviewing our expression conditions in the heat inducible pRE/MZ1 expression system, it was determined that the Fpg production was at 42 °C and ~15% of the total cellular protein was soluble Fpg [17]. Thus, we tested Fpg production in minimal medium at 42 °C with 30 μM IPTG and found that this growth condition yielded ~7% of the total cellular protein as soluble Fpg. A probable reason for the production of more soluble Fpg at 42 °C may be that heat inducible chaperones produced at 42 °C help fold Fpg into its native conformation. For scale up production and purification of ^{15}N-labeled Fpg, we selected the induction with 30 μM IPTG and growth at 42 °C for 5 h.

^{15}N-labeled Fpg was purified by a rapid procedure (Fig. 1). Briefly, Fpg enriched soluble extract was mixed with DEAE cellulose resin. Since Fpg has a high isoelectric point (Pi = 8.7), Fpg would have a polycationic character and would not bind to the DEAE resin; however, numerous other *E. coli* proteins in the preparation would bind strongly to the resin and thus eliminated. The second step employed strong Fpg binding to carboxylmethyl cellulose and elution with salt gradient. This procedure yielded 1 mg of nearly homogenous Fpg/L culture. Purified ^{15}N-labeled Fpg was more than 95% pure and contained some low molecular mass proteins as contaminants.

Production and purification of ^{15}N-labeled hOGG1

There are several isoforms of hOGG1 (http://www.uniprot.org/uniprot/O15527). Two main isoforms, the isoform 1A, also called α-hOGG1 with 345 amino acids and the isoform 2A, also called β-hOGG1, with 424 amino acids are located to the nucleus and the mitochondrion, respectively [8,25–29]. These two isoforms result from an alternative splicing after the transcription of the *hogg1* gene, which is localized on chromosome 3p25 [27]. Starting from the N-terminus, these proteins have an identical sequence with 316 amino acids, with the rest of the molecules exhibiting a different sequence [27]. As a result, the 30 theoretical tryptic peptides of both α-hOGG1 and β-hOGG1 that would result from trypsin digestion are identical. α-hOGG1 is the most abundant among the isoforms [30]. In the present work, a recombinant clone of α-hOGG1 in pET15b expression vector, which would result in an N-terminal His-tag, was used for production and purification of the protein. Unlike Fpg, the addition of the His-tag at the amino terminus of hOGG1 does not affect the activity of hOGG1, since this enzyme uses Lys249 as the catalytic residue [31,32]. The theoretical molecular mass of α-hOGG1 (called hOGG1 from here on) is 38782.2 Da. The His-tag MGSSHHHHHHSSGLVPRGSHMEL replaces Met at the N-terminal of the protein, giving rise to a molecular mass of 41187.8 Da.

The optimization of the production of hOGG1 in minimal medium containing ^{15}N-NH_4Cl was carried out as described for Fpg. Of all the conditions tested: such as induction at 24 °C overnight, 37 °C for 5 h, and 42 °C for 5 h after addition of IPTG to 10 μM, 30 μM, and 100 μM concentration, induction with 100 μM IPTG and growth at any of the temperatures and times produced 20–30% of the total cellular protein as hOGG1. However, only induction at 24 °C produced hOGG1 as mostly soluble protein amounting to about 20% of the total cellular protein. Induction at 37 °C or 42 °C produced hOGG1 as inclusion bodies. The optimal temperature of induction for the production of soluble hOGG1, 24 °C, is quite opposite to the optimal temperature (42 °C) of induction for the production of soluble Fpg. Hence, the conditions for the optimization of overproduction of recombinant proteins in soluble form must be tested for every protein of interest. ^{15}N-labeled hOGG1 was purified to near homogeneity (~99%) from

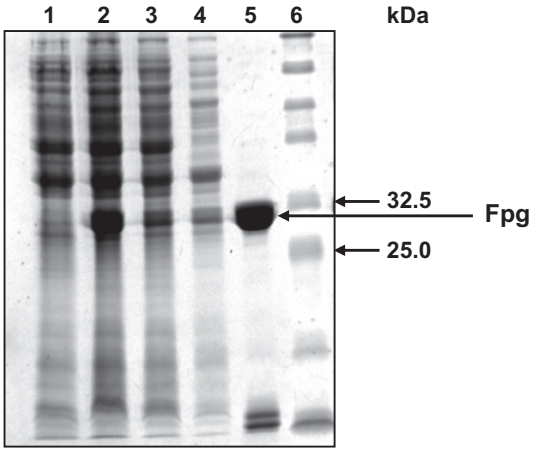

Fig. 1. Purification of ^{15}N-labeled *E.coli* Fpg. Sodium dodecylsulfate–13% polyacrylamide gel electrophoresis analysis. Lane 1: 21 μg protein of the extract of BL21(DE3) cells harboring pET11a plasmid (control); Lane 2: 31 μg protein of the extract of BL21(DE3) cells harboring Fpg/pET11a plasmid induced with IPTG; Lane 3: 26 μg protein of 10,000*g* supernatant; Lane 4: 8.6 μg flow through from DEAE cellulose column; Lane 5: 8.5 μg Fpg eluted from CM cellulose column; Lane 6: Molecular mass markers from top to bottom in kDa – 175, 83, 62, 47.5, 32.5, 25, 16.5 and 6.5.

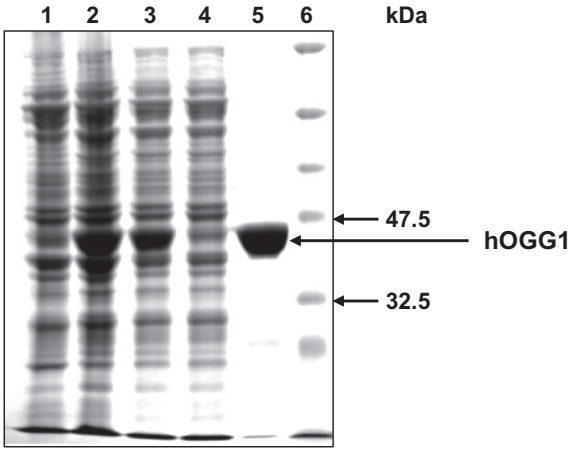

Fig. 2. Purification of ^{15}N-labeled hOGG1. Sodium dodecylsulfate–10% polyacrylamide gel electrophoresis analysis: Lane 1: 20 μg protein of the extract of BL21(DE3) cells harboring pET15b plasmid (control); Lane 2: 21 μg protein of the extract of BL21(DE3) cells harboring hOGG1/pET15b plasmid induced with IPTG; Lane 3: 17 μg protein of 48,000*g* supernatant; Lane 4: 15.5 μg protein from Ni-agarose flow through; Lane 5: 8.3 μg hOGG1 eluted from Ni-agarose column; Lane 7: Molecular mass markers from top to bottom in kDa – 175, 83, 62, 47.5, 32.5, 25, and 16.5.

the soluble extract in a single step on nickel agarose resin (Fig. 2). The yield of hOGG1 was 5 mg protein/L of the culture.

Mass spectrometric analysis of labeled proteins

MALDI-ToF mass spectrometry was used to check the isotopic purity of ^{15}N-labeled Fpg and hOGG1. We performed three measurements of the molecular mass of unlabeled and ^{15}N-labeled Fpg and hOGG1 proteins. As examples, Fig. 3A and B illustrates the ion-current profiles the protonated molecular ion (MH^+) of unlabeled and ^{15}N-labeled Fpg proteins, respectively. The theoretical average molecular mass of Fpg is 30,290 Da. For unlabeled Fpg, a mean value of 30,086 ± 6 Da (the uncertainty is the standard deviation) was found. This is 99.3% of the theoretical value. ^{15}N-labeled Fpg contains 383 nitrogen atoms and thus its theoretical average molecular mass is 30,673 Da. The measurement yielded a value of 30,442 ± 2 Da, which is 99.2% of the theoretical value. Thus, the ^{15}N-incorporation into Fpg was approximately 99.9%. The measured molecular mass of hOGG1 was 41,440 ± 12 Da. The His-tagged hOGG1 has a theoretical average molecular mass of 41,187.8 Da, which amounts to 99.4% of the measured value. This protein contains 548 nitrogen atoms and the average molecular mass of ^{15}N-labeled hOGG1 would be 41,735.8 Da. We found a value of 41,847 ± 14 (the uncertainty is the standard deviation). The theoretical average molecular mass of ^{15}N-labeled His-tagged-hOGG1 amounts to 99.7% of the measured value. The calculation using these values showed that the ^{15}N-incorporation into hOGG1 amounted to approximately 99.7%. As examples, the ion-current profiles of the MH^+ of unlabeled and ^{15}N-labeled hOGG1 proteins are shown in Fig. 4A and B, respectively. These results suggest that

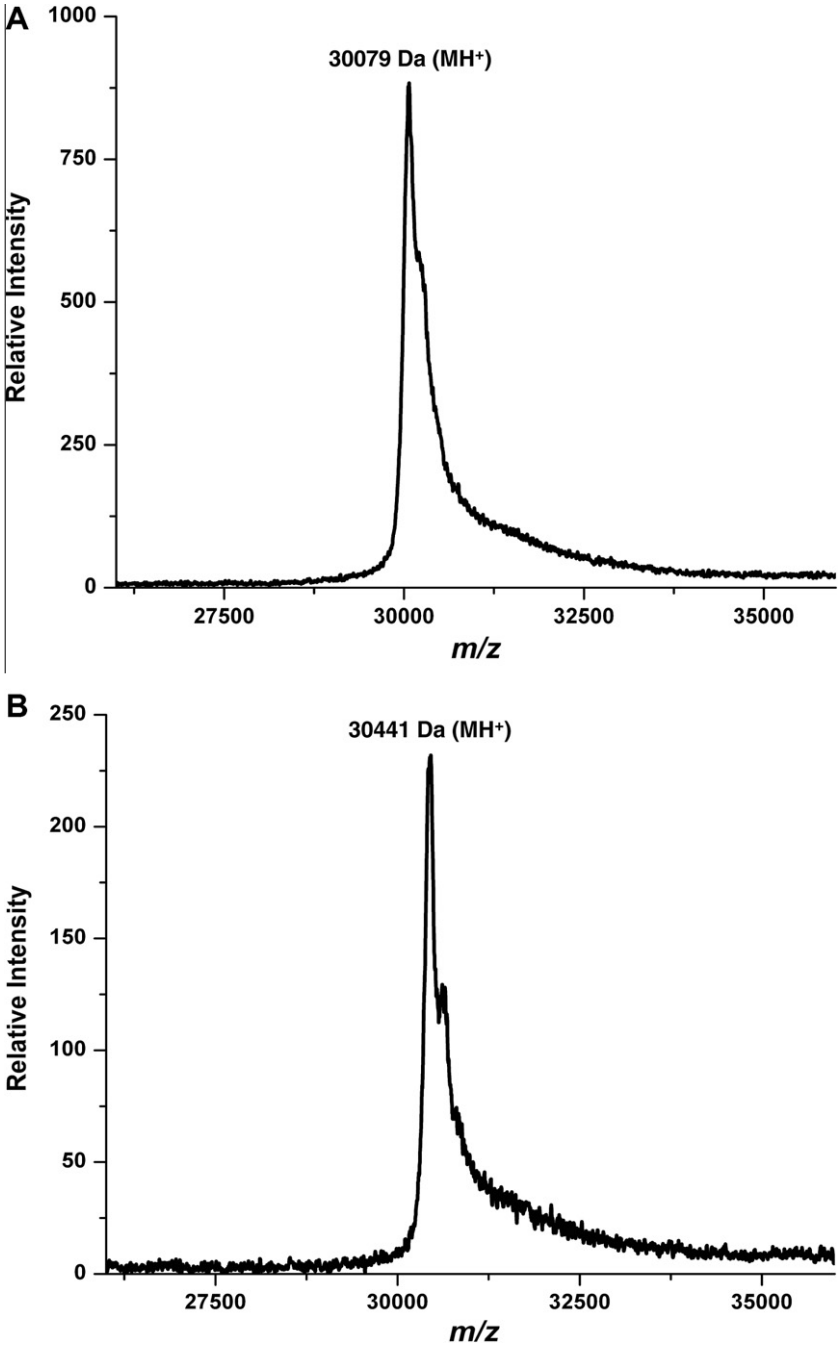

Fig. 3. (A) Ion-current profile of the protonated molecular mass ion of Fpg. (B) Ion-current profile of the protonated molecular mass ion of ^{15}N-labeled Fpg.

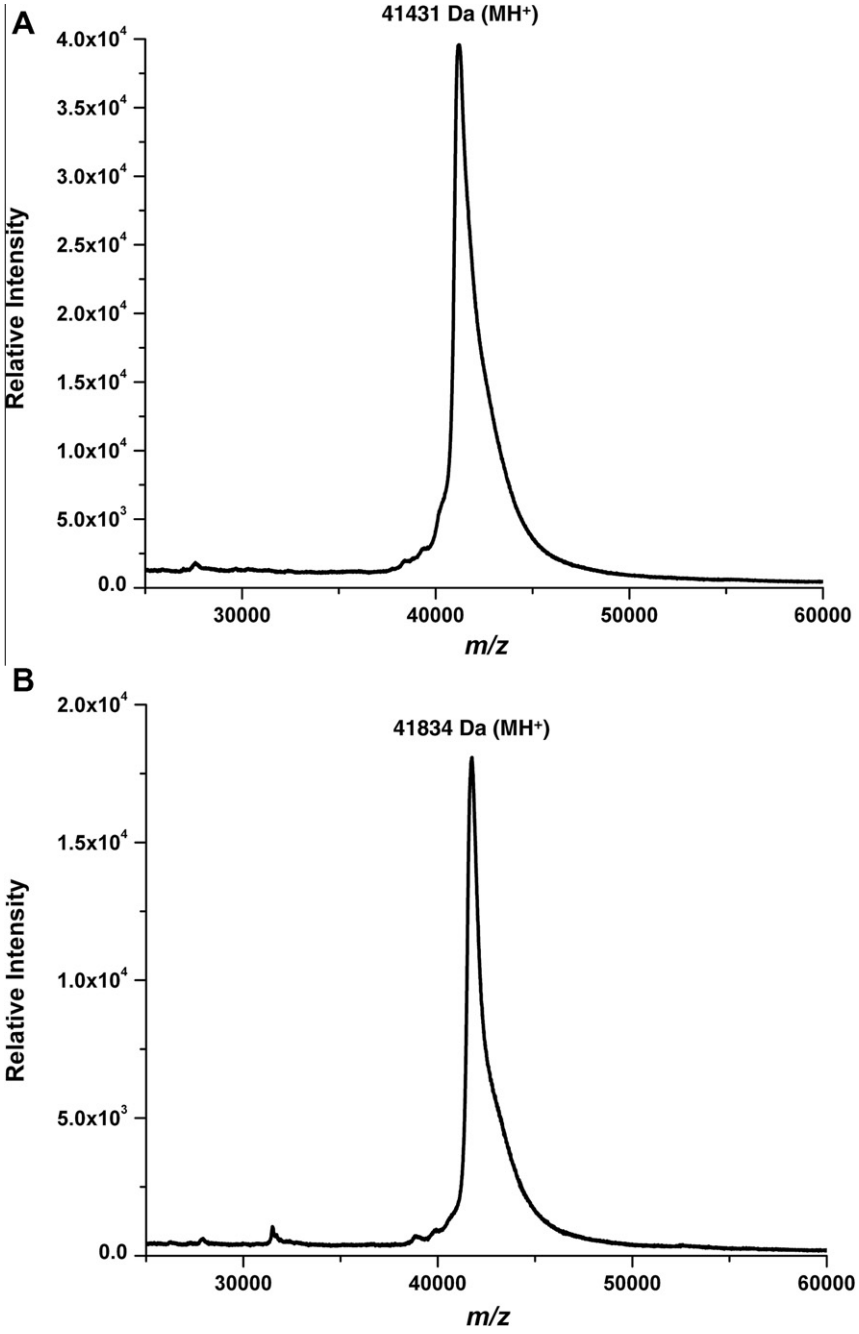

Fig. 4. (A) Ion-current profile of the protonated molecular mass ion of hOGG1. (B) Ion-current profile of the protonated molecular mass ion of ^{15}N-labeled hOGG1.

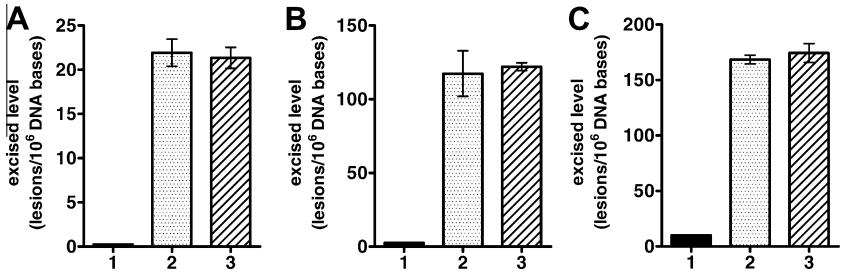

Fig. 5. Measurement of the glycosylase activities of Fpg and ^{15}N-labeled Fpg. Excision of FapyAde (A), FapyGua (B) and 8-OH-Gua (C). 1: no enzyme; 2: Fpg; 3: ^{15}N-labeled Fpg. Uncertainties are standard deviations.

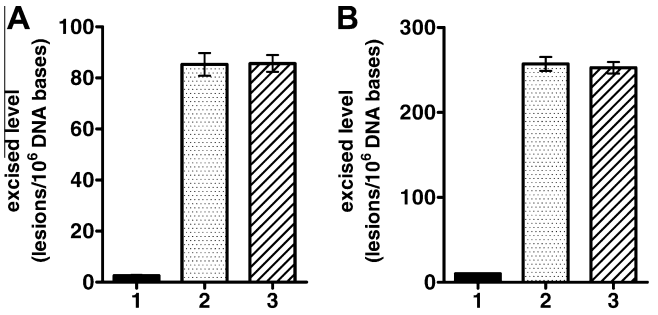

Fig. 6. Measurement of the glycosylase activities of hOGG1 and ^{15}N-labeled hOGG1. Excision of FapyGua (A) and 8-OH-Gua (B). 1: no enzyme; 2: hOGG1; 3: ^{15}N-labeled hOGG1. Uncertainties are standard deviations.

an almost complete ^{15}N-labeling of both Fpg and hOGG1 was achieved, which makes the labeled proteins highly suitable internal standards for the mass spectrometric measurements of Fpg and hOGG1.

DNA glycosylase activities of ^{15}N-labeled Fpg and hOGG1

A stable isotope-labeled protein to be used as an internal standard must have the identical properties as its unlabeled counterpart such as the possession of the enzymic activity. For this reason, we determined the DNA glycosylase activities of ^{15}N-labeled Fpg and ^{15}N-labeled hOGG1, and their unlabeled analogs by using GC/MS with isotope-dilution as described previously [17]. Fpg removes FapyAde, FapyGua and 8-OH-Gua from DNA containing multiple lesions [6,7], whereas hOGG1 acts on FapyGua and 8-OH-Gua, but not on FapyAde [8–10]. The data showed that ^{15}N-labeled Fpg was as active as unlabeled Fpg and efficiently removed FapyAde, FapyGua and 8-OH-Gua from DNA as illustrated in Fig. 5A–C, respectively. Similarly, the glycosylase activity of ^{15}N-labeled hOGG1 was preserved as evidenced by the efficient excision of FapyGua and 8-OH-Gua from DNA (Fig. 6A and B). These results unequivocally show that the ^{15}N-labeling does not affect the enzymic activities of Fpg and hOGG1.

Conclusions

We described a procedure for the efficient production, isolation and purification of large quantities of two ^{15}N-labeled DNA repair proteins. Mass spectrometric measurements showed that these proteins were labeled to a degree of more than 99.7%. We also demonstrated that the labeling did not affect the glycosylase activities of these enzymes. A fully labeled analog of the analyte protein possesses identical properties and can be added to a sample in the first step of the experimental procedure. This will compensate material losses that may occur during all stages of the analysis, enabling accurate quantification. The use of a labeled tryptic peptide as an internal standard would not meet these requirements. This is especially true in the case of the use of two-dimensional gel electrophoresis to isolate a protein prior to analysis by LC–MS/MS. Thus ^{15}N-labeled Fpg and ^{15}N-labeled hOGG1 can be used as ideal internal standards for the measurement of *E. coli* Fpg and hOGG1 by mass spectrometric techniques such as LC–MS/MS with isotope-dilution. Moreover, the tryptic peptides expected from trypsin digestion of the major isoforms α-hOGG1 and β-hOGG1 are identical up to the first 316 amino acids of the N-terminal. This means that ^{15}N-labeled α-hOGG1 would serve as an ideal internal standard for identification and quantification of both isoforms by LC–MS/MS following trypsin digestion. This is also valid for other minor isoforms, whose N-terminal sequences to a great extent are identical to that of α-hOGG1. The procedures described in this work are likely to be applicable to the over-expression, isolation, purification and characterization of stable isotope-labeled analogs of other DNA repair proteins for use as suitable internal standards. However, the protein expression must be optimized for each protein. The availability of such labeled protein standards would facilitate the *in vivo* measurement by mass spectrometric techniques of DNA repair proteins as possible predictive, prognostic and therapeutic biomarkers for cancer and other diseases.

Acknowledgments

We are grateful to Dr. Dmitry Zharkov (SB RAS Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia) for providing the *hogg1* gene in the expression vector pET15b for His-tagged production of the protein, and to Dr. Timothy R. O'Connor (Beckman Research Institute of the City of Hope, Duarte, CA) for the gift of the coding sequence for *E. coli* Fpg gene from the plasmid pFpg.

Certain commercial equipment or materials are identified in this paper in order to specify adequately the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

References

[1] M.D. Evans, M. Dizdaroglu, M.S. Cooke, Oxidative DNA damage and disease: induction, repair and significance, Mutat. Res. 567 (2004) 1–61.
[2] E.C. Friedberg, G.C. Walker, W. Siede, R.D. Wood, R.A. Schultz, T. Ellenberger, DNA Repair and Mutagenesis, ASM Press, Washington, DC, 2005.
[3] J.H. Hoeijmakers, Genome maintenance mechanisms for preventing cancer, Nature 411 (2001) 366–374.
[4] L. Gros, M.K. Saparbaev, J. Laval, Enzymology of the repair of free radicals-induced DNA damage, Oncogene 21 (2002) 8905–8925.
[5] M.L. Hegde, T.K. Hazra, S. Mitra, Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells, Cell Res. 18 (2008) 27–47.
[6] S. Boiteux, E. Gajewski, J. Laval, M. Dizdaroglu, Substrate specificity of the *Escherichia coli* Fpg protein (formamidopyrimidine–DNA glycosylase): excision of purine lesions in DNA produced by ionizing radiation or photosensitization, Biochemistry 31 (1992) 106–110.
[7] M. Dizdaroglu, Base-excision repair of oxidative DNA damage by DNA glycosylases, Mutat. Res. 591 (2005) 45–59.
[8] C. Dherin, J.P. Radicella, M. Dizdaroglu, S. Boiteux, Excision of oxidatively damaged DNA bases by the human alpha-hOgg1 protein and the polymorphic alpha-hOgg1(Ser326Cys) protein which is frequently found in human populations, Nucleic Acids Res. 27 (1999) 4001–4007.
[9] M. Audebert, J.P. Radicella, M. Dizdaroglu, Effect of single mutations in the OGG1 gene found in human tumors on the substrate specificity of the ogg1 protein, Nucleic Acids Res. 28 (2000) 2672–2678.
[10] V.S. Sidorenko, A.P. Grollman, P. Jaruga, M. Dizdaroglu, D.O. Zharkov, Substrate specificity and excision kinetics of natural polymorphic variants and phosphomimetic mutants of human 8-oxoguanine–DNA glycosylase, FEBS J. 276 (2009) 5149–5162.
[11] M. Berwick, P. Vineis, Markers of DNA repair, susceptibility to cancer in humans. An epidemiologic review, J. Natl. Cancer Inst. 92 (2000) 874–897.
[12] S.M. Lippman, M.R. Spitz, Lung cancer chemoprevention. An integrated approach, J. Clin. Oncol. 19 (2001) 74S–82S.
[13] E.L. Goode, C.M. Ulrich, J.D. Potter, Polymorphisms in DNA repair genes and associations with cancer risk, Cancer Epidemiol. Biomarkers Prev. 11 (2002) 1513–1530.
[14] S. Madhusudan, M.R. Middleton, The emerging role of DNA repair proteins as predictive prognostic and therapeutic targets in cancer, Cancer Treat. Rev. 31 (2005) 603–617.
[15] Q. Wei, L.E. Wang, E.M. Sturgis, L. Mao, Expression of nucleotide excision repair proteins in lymphocytes as a marker of susceptibility to squamous cell carcinomas of the head and neck, Cancer Epidemiol. Biomarkers Prev. 14 (2005) 1961–1966.
[16] I. Eidhammer, K. Flikka, L. Martens, S.-O. Mikalsen, Computational Methods for Mass Spectrometric Proteomics, John Wiley & Sons Ltd., West Sussex, England, 2007.
[17] P. Reddy, P. Jaruga, T. O'Connor, H. Rodriguez, M. Dizdaroglu, Overexpression and rapid purification of Escherichia coli formamidopyrimidine–DNA glycosylase, Protein Expr. Purif. 34 (2004) 126–133.
[18] R.W. Davis, D. Botstein, J.R. Roth, A Manual for Genetic Engineering: Advanced Bacterial Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1980.

[19] J.H. Miller, Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972.
[20] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.
[21] O.H. Lowry, N.J. Rosebrough, A.L. Farr, R.J. Randall, Protein measurement with the Folin phenol reagent, J. Biol. Chem. 193 (1951) 265–275.
[22] P. Jaruga, G. Kirkali, M. Dizdaroglu, Measurement of formamidopyrimidines in DNA, Free Radic. Biol. Med. 45 (2008) 1601–1609.
[23] P. Reddy, A. Peterkofsky, K. McKenney, Hyperexpression and purification of *Escherichia coli* adenylate cyclase using a vector designed for expression of lethal gene products, Nucleic Acids Res. 17 (1989) 10473–10488.
[24] M. Zuber, T.A. Patterson, Court DL, Analysis of nutR, a site required for transcription antitermination in phage lambda, Proc. Natl. Acad. Sci. USA 84 (1987) 4514–4518.
[25] J.P. Radicella, C. Dherin, C. Desmaze, M.S. Fox, S. Boiteux, Cloning and characterization of hOGG1, a human homolog of the OGG1 gene of *Saccharomyces cerevisiae*, Proc. Natl. Acad. Sci. USA 94 (1997) 8010–8015.
[26] K. Arai, K. Morishita, K. Shinmura, T. Kohno, S.R. Kim, T. Nohmi, M. Taniwaki, S. Ohwada, J. Yokota, Cloning of a human homolog of the yeast *OGG1* gene that is involved in the repair of oxidative DNA damage, Oncogene 14 (1997) 2857–2861.
[27] S. Boiteux, J.P. Radicella, Base excision repair of 8-hydroxyguanine protects DNA from endogenous oxidative stress, Biochimie 81 (1999) 59–67.
[28] M. Takao, H. Aburatani, K. Kobayashi, A. Yasui, Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage, Nucleic Acids Res. 26 (1998) 2917–2922.
[29] K. Nishioka, T. Ohtsubo, H. Oda, T. Fujiwara, D. Kang, K. Sugimachi, Y. Nakabeppu, Expression and differential intracellular localization of two major forms of human 8-oxoguanine DNA glycosylase encoded by alternatively spliced OGG1 mRNAs, Mol. Biol. Cell 10 (1999) 1637–1652.
[30] Y. Monden, T. Arai, M. Asano, E. Ohtsuka, H. Aburatani, S. Nishimura, Human MMH (OGG1) type 1a protein is a major enzyme for repair of 8-hydroxyguanine lesions in human cells, Biochem. Biophys. Res. Commun. 258 (1999) 605–610.
[31] H.M. Nash, R. Lu, W.S. Lane, G.L. Verdine, The critical active-site amine of the human 8-oxoguanine DNA glycosylase, hOgg1: direct identification, ablation and chemical reconstitution, Chem. Biol. 4 (1997) 693–702.
[32] S. Boiteux, J.P. Radicella, The human OGG1 gene: structure, functions, and its implication in the process of carcinogenesis, Arch. Biochem. Biophys. 377 (2000) 1–8.