Published online December 1, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 21 ***6251–6259***
doi:10.1093/nar/gkh962

DmGEN, a novel RAD2 family endo-exonuclease from *Drosophila melanogaster*

Gen Ishikawa, Yoshihiro Kanai, Kei-ichi Takata, Ryo Takeuchi, Kaori Shimanouchi, Tatsushi Ruike, Tomoyuki Furukawa, Seisuke Kimura and Kengo Sakaguchi*

Department of Applied Biological Science, Faculty of Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda-Shi, Chiba-ken 278 8510, Japan

Received October 19, 2004; Revised and Accepted November 8, 2004 DDBJ/EMBL/GenBank accession no. AB103508

ABSTRACT

A novel endo-exonuclease, *DmGEN* (*Drosophila Melanogaster* XPG-like endonuclease), was identified in *D.melanogaster*. *DmGEN* is composed of five exons and four introns, and the open reading frame encodes a predicted product of 726 amino acid residues with a molecular weight of 82.5 kDa and a pI of 5.36. The gene locus on *Drosophila* polytene chromosomes was detected at 64C9 on the left arm of chromosome 3 as a single site. The encoded protein showed a relatively high degree of sequence homology with the RAD2 nucleases, especially XPG. Although the XPG-N- and XPG-I-domains are highly conserved in sequence, locations of the domains are similar to those of FEN-1 and EXO-1, and the molecular weight of the protein is close to that of EXO-1. *In vitro*, DmGEN showed endonuclease and 3′–5′ exonuclease activities with both single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA), but the endonuclease action with dsDNA was quite specific: 5′–3′ exonuclease activity was found to occur with nicked DNA, while dsDNA was endonucleolytically cut at 3–4 bp from the 5′ end. Homologs are widely found in mammals and higher plants. The data suggest that DmGEN belongs to a new class of RAD2 nuclease.

INTRODUCTION

DNA repair has evolved to protect cells against the mutagenic and cytotoxic effects of DNA damage. Central to most DNA repair processes is a nucleolytic step that is required for the removal of the damaged section. Nucleases are essential for nucleolytic activities and act in a variety of structural frameworks, ranging from site-specific (e.g. abasic endonuclease) to structure-specific (e.g. members of the RAD2 nuclease family) and non-specific (e.g. DNase I). We have concentrated on investigating enzymes belonging to the RAD2 nuclease family, in the past several years, and elucidating their relationships to DNA repair (1–4).

The RAD2 family includes XPG (Class I), FEN-1 (Class II), EXO-1 (Class III) and SEND-1 (Class IV), and the proteins exhibit a range of substrate-specific exo- and endonuclease activities. Many of these proteins act as nucleases that contribute to DNA repair, replication, and recombination. RAD2 Class I consists of XPG-like proteins that cleave at the 3′ side of the damage-containing bubble structure formed during nucleotide excision repair (5–13). Class II comprises of FEN-1-like proteins, structure-specific nucleases that remove 5′ DNA flaps produced by polymerase strand displacement and are also responsible for the resolution of Okazaki fragments during DNA replication (1,2,10,14–17). Class III is made up of the Exo-1-like proteins found in yeast, Drosophila and mammals. Exo-1-like proteins appear to play roles in DNA recombination, mismatch repair and DNA replication (10,18,19). Much of the amino acid sequence conservation among the RAD2 family is concentrated within two domains, termed the N (N-terminal) and I (internal) nuclease domains. Class IV is a new RAD2 category which was originally found in higher plants (4), and knowledge about its biochemical features is limited.

To effect a deeper comprehension of the nature of RAD2 family nucleases, we have studied their characteristics in *Drosophila melanogaster*, a species for which new mutants in the many genes known to be associated with DNA metabolism and the morphogenesis (20) can be readily prepared. The available mutant collection, coupled with a refined system for genetic analysis, provides a valuable tool for investigation. Since little is known about the proteins, gene expression or *in vivo* functions of RAD2 family nucleases in *Drosophila*, it is necessary to fill this knowledge gap.

We have found a new class of RAD2 nuclease in *D.melanogaster*, and have succeeded in cloning the gene and characterizing the enzyme. In this report, we describe the discovery and biochemical characteristics of this novel enzyme.

MATERIALS AND METHODS

Identification of XPG-like endonuclease (GEN) in *D.melanogaster*

The *Drosophila* expressed sequence tag (EST) database was searched using the BLAST program to identify cDNA clones

*To whom correspondence should be addressed. Tel: +81 4 7124 1501 (Ext. 3409); Fax: +81 4 7123 9767; Email: kengo@rs.noda.tus.ac.jp

with homology to RAD2 family proteins. *Drosophila* EST clone RE33588 was found to be one example. Its insert DNA size was ~2.1 kb and it contained a full-length cDNA. The clone showed significant homology with all of the known RAD2 family proteins, but was not thought to be homologous to any member of the RAD2/FEN-1 family. Therefore, it was designated as *D.melanogaster XPG-like endonuclease* (*DmGEN*) in this report. The nucleotide sequence data reported in this article have been registered under the accession number AB103508.

Overexpression and purification of DmGEN protein

PCR was carried out using the EST clone RE33588, containing the entire *Drosophila GEN* (*DmGEN*) coding sequence. The primers synthesized chemically were 5′-CAT ATG GGC GTC AAG GAA TTA TG-3′ and 5′-GGA TCC CTT AAT CAC TAA TCA CCA CCA-3′. The resultant 2181 bp NdeI/BamHI DNA fragment was cloned into the pET28a expression vector (Novagen). Protein expression was performed by transforming pET28a-*DmGEN* into BL21(DE3) (Novagen) and growing the bacteria in 1000 ml of Luria–Bertani medium. Cells were grown to an OD of 0.7 and isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. After 3 h incubation at 30°C, cells were harvested by centrifugation at 6000 *g* for 10 min. Cell pellets were resuspended in 130 ml of ice-cold binding buffer (20 mM Tris–HCl, pH 7.9, 0.5 M NaCl, 5 mM imidazole and 0.1% NP-40) and sonicated with thirty 10 s bursts. Cell lysates were centrifuged at 15 000 *g* for 30 min and the soluble protein fraction was collected as the crude extract. This was then loaded onto 4 ml of His-Bind Resin (Novagen). The column was washed with 40 ml of binding buffer, and then the bound protein was eluted with 30 ml of elution buffer (20 mM Tris–HCl, pH 7.9, 0.5 M NaCl, 1 M imidazole and 0.1% NP-40) and dialyzed against TEMG (50 mM Tris–HCl, pH 7.5, 1 mM EDTA, 5 mM 2-mercaptoethanol and 15% glycerol) two times. The dialysate was loaded onto a Hitrap-Heparin column (Amersham Pharmacia Biotech) equilibrated with TEMG and washed. Fractions were collected with 30 ml of a linear gradient of 0–1.5 M NaCl in TEMG. Then, the nuclease analyses described below were performed for each fraction. The activated fractions eluted at 0.7–0.8 M NaCl were collected, dialyzed against TEMG and loaded onto a SP Sepharose Fast Flow column (Amersham Pharmacia Biotech) equilibrated with TEMG. After washing, the fraction was collected with 20 ml of a linear gradient of 0–2.0 M NaCl in TEMG. The active fractions were detected by nuclease analysis and then, dialyzed against TEMG. Those fractions were used in the subsequent experiments.

Preparation of substrates for nuclease assays

The nucleotide sequences of the oligonucleotides used in these studies as DNA substrate are shown in Table 1. According to the manufacturer's protocol, T4 polynucleotide kinase was used to 5′ phosphorylate specified oligonucleotides with [γ-^{32}P]ATP (Amersham Biosciences). For 3′ labeling, the oligonucleotides were incubated with terminal deoxynucleotidyl transferase and [α-^{32}P]ddATP (Amersham Biosciences). Excess ATP was removed using a Microspin G-25 column. Annealing reactions were carried out by mixing the oligonucleotides in an annealing buffer (20 mM Tris–HCl, pH 7.4 and 150 mM NaCl) and the labeled strand was annealed to an equal molar concentration of the complementary oligonucleotide by heating to 95°C for 2.5 min, followed by gradual cooling to 4°C. A schematic diagram of the DNA structures generated is shown in Table 2.

Table 1. Nucleotide sequences for substrate oligonucleotides

DNA	Nucleotide sequence
A-1, 30mer	5′-GTC GAC CTG CAG CCC AAG CTT GCG TTG CTG-3′
A-2, 30mer	5′-CAG CAA CGC AAG CTT GGG CTG CAG GTC GAC-3′
B-1, 40mer	5′-ATG CGG ATC CGG AAT CGA AGG TCG TCA TAT GGA ATT CAT A-3′
B-2, 40mer	5′-TAT GAA TTC CAT ATG ACG ACC TTC GAT TCC GGA TCC GCA T-3′
C-1, 60mer	5′-ACG CTG CCG AAT TCT ACC AGT GCC TTG CTA GGA CAT CT T TGC CCA CCT GCA GGT TCA CCC-3′
C-2, 60mer	5′-GGG TGA ACC TGC AGG TGG GCA AAG ATG TCC TAG CAA GGC ACT GGT AGA ATT CGG CAG CGT-3′
D-1, 90mer	5′-CCA GTG ATC ACA TAC GCT TTG CTA GGA CAT CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CAG TGC CAC GTT GTA TGC CCA CGT TGA CCG-3′
D-2, 90mer	5′-CGG TCA ACG TGG GCA TAC AAC GTG GCA CTG GGG GGG GGG GGG GGG GGG GGG GGG GGG GGG ATG TCC TAG CAA AGC GTA TGT GAT CAC TGG-3′
D-3, 90mer	5′-CGG TCA ACG TGG GCA TAC AAC GTG GCA CTG GGG GGG GGG GGG GGG GGG GGG GGG GGG GGG ATG TCC TAG CAA AGC GTA TGT GAT CAC TGG-3′
E-1, 33mer	5′-ATG TGG AAA ATC TCT AGC AGG CTG CAG GTC GAC-3′
E-2, 16mer	5′-CAG CAA CGC AAG CTT G-3′
F-1, 21mer	5′-TAG AGG ATC CCC GCT AGC GGG-3′
F-2, 21mer	5′-TAC CGA GCT CGA ATT CAC TGG-3′
F-3, 42mer	5′-CCA GTG AAT TCG AGC TCG GTA CCC GCT AGC GGG GAT CCT CTA-3′
F-4, 20mer	5′-ACC GAG CTC GAA TTC ACT GG-3′
F-5, 20mer	5′-TAG AGG ATC CCC GCT AGC GG-3′
F-6, 18mer	5′-TAG AGG ATC CCC GCT AGC-3′

Nuclease assay (with circular DNA as the substrate)

Reactions (20 µl) were carried out in buffer (25 mM HEPES–KOH, pH 6.8, 2 mM $MgCl_2$, 1 mM DTT, 50 mg/ml BSA, 20 mM KCl, 10% glycerol and 0.1% NP-40) plus 100 ng of single-stranded M13mp18DNA and DmGEN protein. After incubation at 37°C for 60 min, the reaction was stopped with 1 µl of 0.5 M EDTA. Samples were analyzed by electrophoresis on 0.7% agarose gels containing 0.5 µg/ml ethidium bromide.

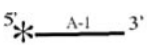
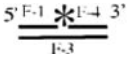
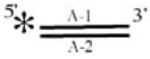
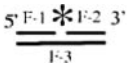
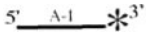
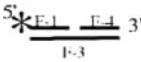
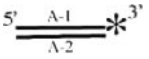
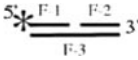
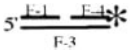
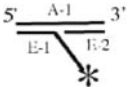
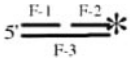
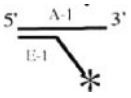
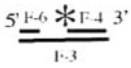
Nuclease assay (with linear DNA as the substrate)

DmGEN protein (16.2 ng) was incubated with 100 fmol of ^{32}P-labeled DNA substrate in a 20 µl reaction mixture containing 50 mM Tris–HCl, pH 7.5, 5 mM $MgCl_2$ and 1 mM DTT at 37°C. The reaction was stopped by adding 10 µl of gel loading buffer (90% deionized formamide, 5 mM EDTA, 0.1% bromophenol blue and 0.1% xylenecyanol), the sample was heated for 5 min at 95°C, and a fraction was loaded onto a 20% polyacrylamide gel containing 7 mM urea in TBE buffer. Electrophoresis was carried out for 2.5 h and reaction products were visualized by autoradiography.

Effects of biotin adducts and streptavidin on nucleolytic activity

In an attempt to avoid the influence of the exonuclease activity of DmGEN protein, substrates with primers having

Table 2. DNAs employed for DmGEN protein substrate-specificity studies

5′-labeled single-stranded	A-1, A-2, B-1, C-1, D-1		Gapped double-stranded 1	(F-1, F-3, F-4)	
5′-labeled double-stranded	(A-1, A-2), (B-1, B-2), (C-1, C-2), (D-1, D-2)		Nicked double-stranded 1	(F-1, F-2, F-3)	
3′-labeled single-stranded	A-1		Gapped double-stranded 2	(F-1, F-3, F-4)	
3′-labeled double-stranded	(A-1, A-2)		Nicked double-stranded 2	(F-1, F-2, F-3)	
5′ bubble	(D-1, D-3)		Gapped double-stranded 3	(F-1, F-3, F-4)	
5′ flap	(A-1, E-1, E-2)		Nicked double-stranded 3	(F-1, F-2, F-3)	
5′ pseudo flap	(A-1, E-1)		Gapped double-stranded 4	(F-3, F-4, F-5)	
			Gapped double-stranded 5	(F-3, F-4, F-6)	

All substrates were generated as described in Materials and Methods. The asterisk indicates the position of the ^{32}P label. Oligonucleotide that was labeled in the experiment is underlined, and the nucleotide sequences are shown in Table 1.

a biotinylated base at the 3′ end (Sigma) were used. Substrates were incubated with streptavidin (Promega), a 60 kDa protein that recognizes and tightly binds biotin (21), in 25-fold molar excess. Then, nuclease analyses were performed with the above-described procedure. The Klenow flagment used as a positive control was purchased from Toyobo.

RESULTS AND DISCUSSION

Identification of a novel RAD2 family endonuclease from *D.melanogaster*

A search of the *Drosophila* EST database for RAD2 family proteins generated a clone, RE33588, with significant homology to *Drosophila mus201* (XPG homolog). Therefore, we named the new RAD2 family gene *D.melanogaster XPG-like* endonuclease (*DmGEN*).

DmGEN consists of a 2181 bp open reading frame encoding a protein of 726 amino acids with a calculated molecular weight of 82.5 kDa. Its genome sequence is composed of 5 exons in a region of 2503 bp. The predicted length of the DmGEN protein is between those of XPG (Class I) and FEN-1 (Class II), and similar to those of Exonuclease-1 (Class III) and OsSEND-1 (Class IV) proteins (Figure 1).

A phylogenetic tree was drawn based on the neighbor-joining method. The deduced amino acid sequence could be shown to share a high degree of homology with RAD2 nuclease family proteins, with highly conserved N- and I-domains that are characteristic of the RAD2 family (data not shown). The results indicated the DmGEN protein to differ from the proteins in any of the known RAD2 classes. *DmGEN*, registered under accession number AB103508 in the DDBJ/EMBL/GenBank nucleotide sequence data bases, was mapped onto the left arm of chromosome 3, and the location of cDNA of DmGEN on the physical map of *D.melanogaster* was found to be 3L:5,109,154.-5,111,656. The genomic gene contains five exons and four introns.

Figure 1 shows a comparison of the predicted amino acid sequence of DmGEN with the sequences of other RAD2 family proteins. DmGEN is relatively independent from the Class I, II, III or IV proteins, and closer to XPG/RAD2 (Class I), implying that it belongs to a new RAD2 category. In *Drosophila*, mus201 protein (Class I), FEN-1 homolog protein (Class II) and tosca protein (Class III) are the reported RAD2 family proteins, all differing from DmGEN. Although in size DmGEN is closest to the Class III proteins, its N- and I-domains are more homologous with Class I domains. Therefore, the DmGEN protein was concluded to be a novel endonuclease belonging to a new class of RAD2 family proteins present in *D.melanogaster*. This conclusion was also supported by the data on the biochemical properties described below. Actually, DmGEN homologs exist in other species such as *Homo sapiens*, *Mus musculus* and *Arabidopsis*, composing class V. According to the genome database of *D.melanogaster* (FlyBase), the only RAD2 family protein other than the Classes I, II and III

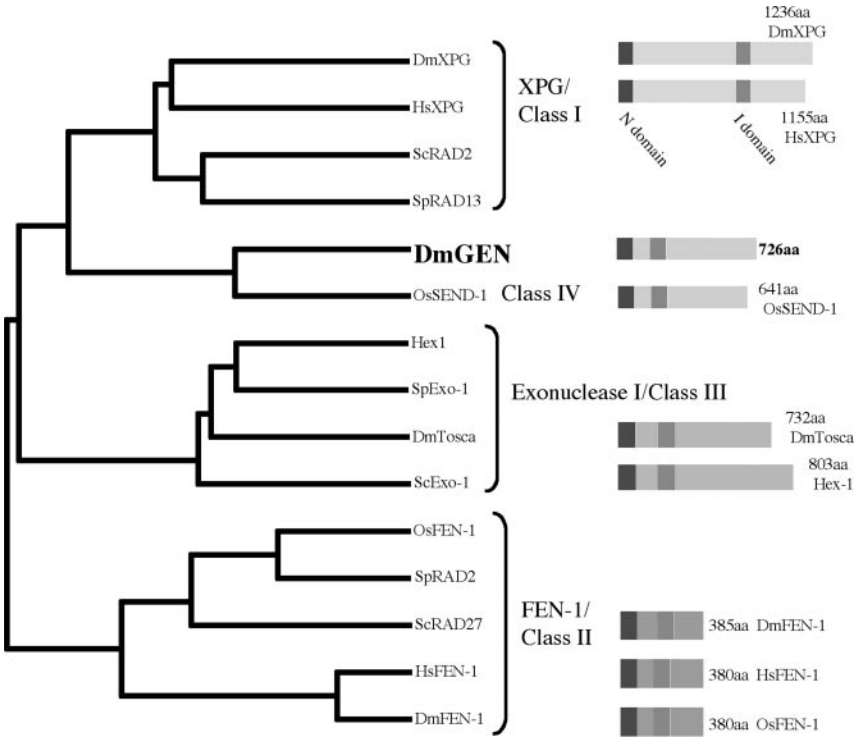


Figure 1. The dendrogram on the left depicts relationships among FEN-1/RAD2 family members. On the right are depicted the protein structures with highly conserved domains (XPG-N domain and XPG-I domain) indicated.

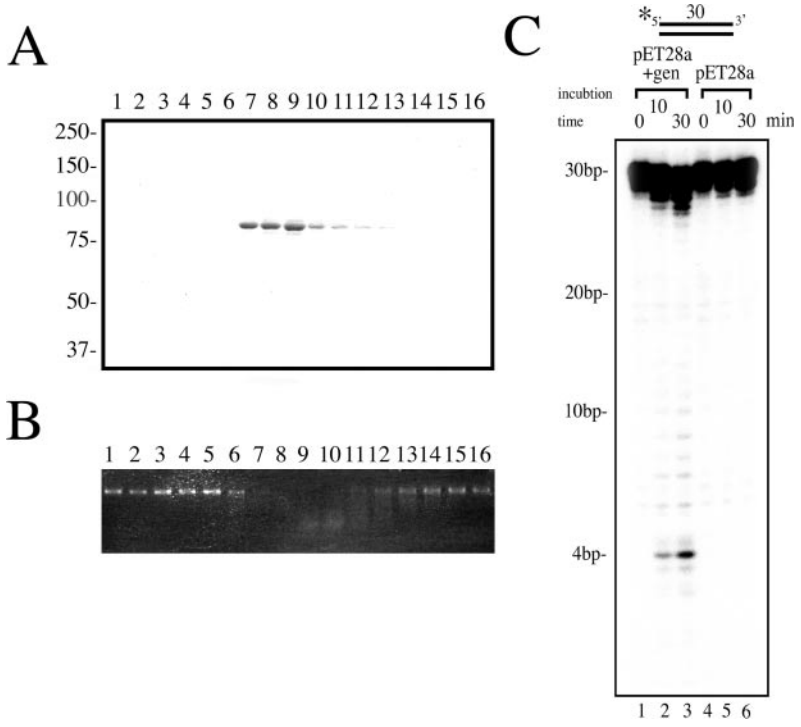


Figure 2. (**A**) Purification of DmGEN protein. DmGEN protein overexpressed in *E.coli* was separated by 7.5% SDS–PAGE and stained with Coomassie brilliant blue. (**B**) Endonuclease activity of each fraction from (A). The substrate was single-strand circular DNA. Incubation was at 37°C for 30 min. (**C**) Nuclease activity on double-stranded DNA (dsDNA) to confirm expression of the DmGEN protein. The total reaction volume was 20 µl. 5′ end-labeled dsDNA (100 fmol) was incubated with the protein (16.2 ng) extracted from *E.coli*. Time course experiments were performed. Schematic representations of the substrates are shown. The asterisk indicates the position of the radiolabel. Substrate and cleavage product sizes were as indicated. Lanes 1–3 included extracts from *E.coli* carrying the pET28a^{+} vector with the DmGEN gene. Lanes 4–6 included extracts from *E.coli* with the pET28a^{+} vector alone. Reaction products were resolved on 20% polyacrylamide/8 M urea gels and visualized using autoradiography.

members is DmGEN. The *Drosophila* genome is relatively simple and almost all genes exist as single copies, and *Drosophila* is an early multicellular organism in evolution. Therefore, Class V may be a major class of the RAD2 family.

Overexpression and purification of the DmGEN protein

The DmGEN coding region was amplified by PCR, cloned into the pET28a expression vector and transformed into *Escherichia coli* for protein induction. Extracts prepared from cells induced for 3 h were shown to contain a 6-histidine C-terminal-tagged DmGEN fusion protein. This was obtained in a soluble state and purified to near homogeneity by His-Bind resin column chromatography, followed by passage through Hitrap-Heparin and SP Sepharose Fast Flow columns. The final product was DmGEN as shown by SDS–PAGE analysis of the SP Sepharose Fast Flow column fractions (Figure 2A). The DmGEN-rich fraction demonstrated endonuclease activity in assays using circular DNA (M13mp18 DNA) as the substrate (Figure 2B), and was used in subsequent experiments.

We next performed nuclease assays with 5′ radiolabeled double-stranded DNA as the substrate. The mixture was run, along with radiolabeled DNA size markers, on a 20% polyacrylamide gel containing 7 M urea to prevent rehybridization of the DNA strandsduring electrophoresis. The presence of DNA size markers in the polyacrylamide gel enabled us to determine the size of the cleavage products precisely. There were products resulting from two different nuclease activities of DmGEN (Figure 2C), endonuclease and exonuclease.

Basic biochemical characteristics of the DmGEN protein

At first, we used circular single-stranded DNA (M13mp18 ssDNA) and circular double-stranded DNA (M13mp18 RF-DNA) as substrates to clarify the basic biochemical features. The fraction of intact substrate decreased with an increase in the amount of DmGEN. The mode of degradation was endonucleolytic rather than exonucleolytic and the endonucleolytic activity toward M13mp18 RF-DNA rose sharply when the concentration of DmGEN protein increased. On the other hand, when the substrate was M13mp18 ssDNA, only a slight increase was evident, indicating some specificity of the DmGEN protein for double-stranded DNA. Monovalent (KCl) ions affected the nuclease activity using M13mp18 ssDNA and M13mp18 RF-DNA as the substrates (Figure 3A and B). DmGEN was optimally active at 75–100 mM KCl for degrading M13mp18 ssDNA, and at 150–300 mM KCl for degrading M13mp18 RF-DNA. The nuclease activity of DmGEN was also greatly influenced by the presence of divalent metal ions (Figure 3C). Mg^{2+}, Mn^{2+}, Zn^{2+} and Ca^{2+} ions all enhanced the activity for cleaving double-stranded DNA. DmGEN prefered Mn^{2+} ions when the substrate was M13mp18 RF-DNA. When the substrate was ssDNA, the enhancement of nuclease activity by the divalent cations was relatively weak. The optimal temperature for DmGEN endonuclease activity was 42°C.

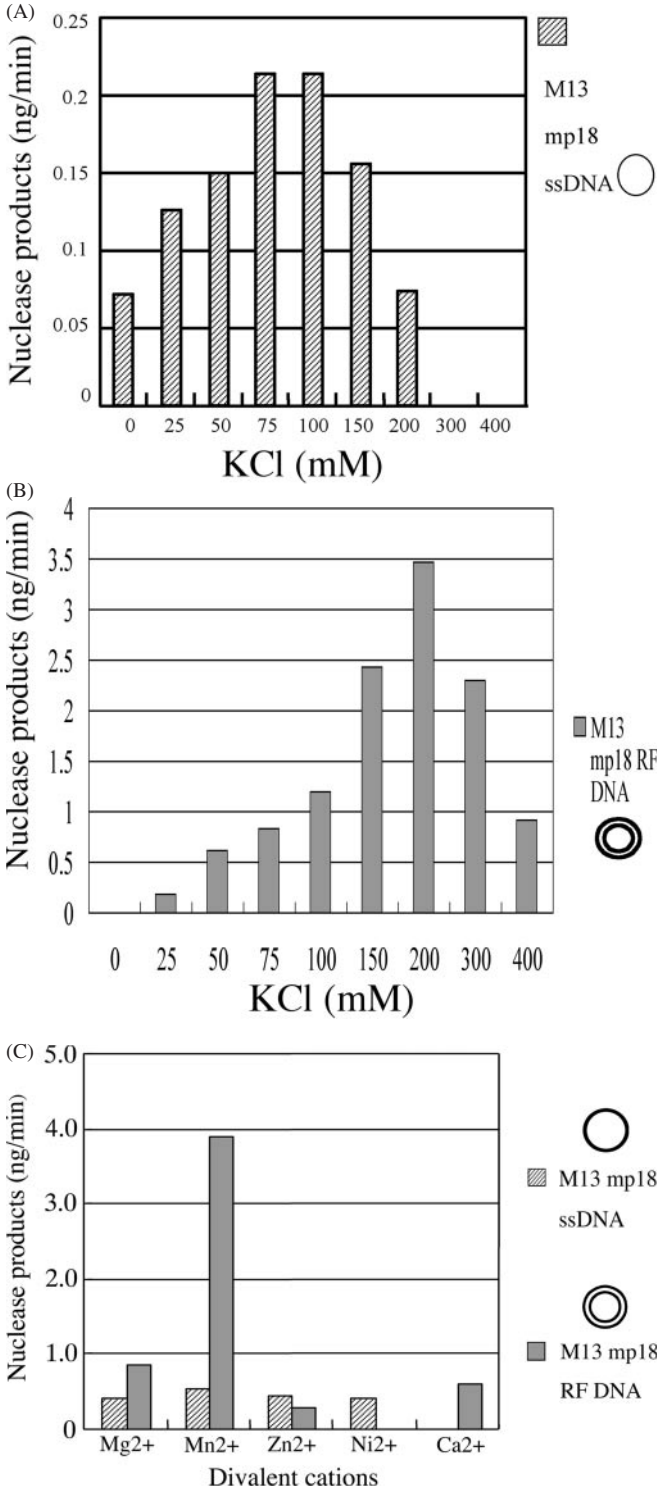


Figure 3. The basic biochemical properties of the DmGEN protein. (**A**) Effects of salt concentration on the endonuclease activity of DmGEN protein (32.4 ng) toward ssDNA substrate. Incubation was conducted at 37°C for 60 min with or without DmGEN proteins. (**B**) Effects of salt concentration on the endonuclease activity of DmGEN protein (32.4 ng) toward dsDNA substrate. Incubation was conducted at 37°C for 10 min with or without DmGEN proteins. (**C**) Effects of divalent cations on the endonuclease activity of the DmGEN protein (32.4 ng) toward ssDNA and dsDNA substrates. Incubation was conducted at 37°C for 30 min with DmGEN proteins. Samples were analyzed by electrophoresis on 0.7% agarose gels. Amounts of nuclease products were calculated with the aid of a luminescent image analyzer (Fujifilm).

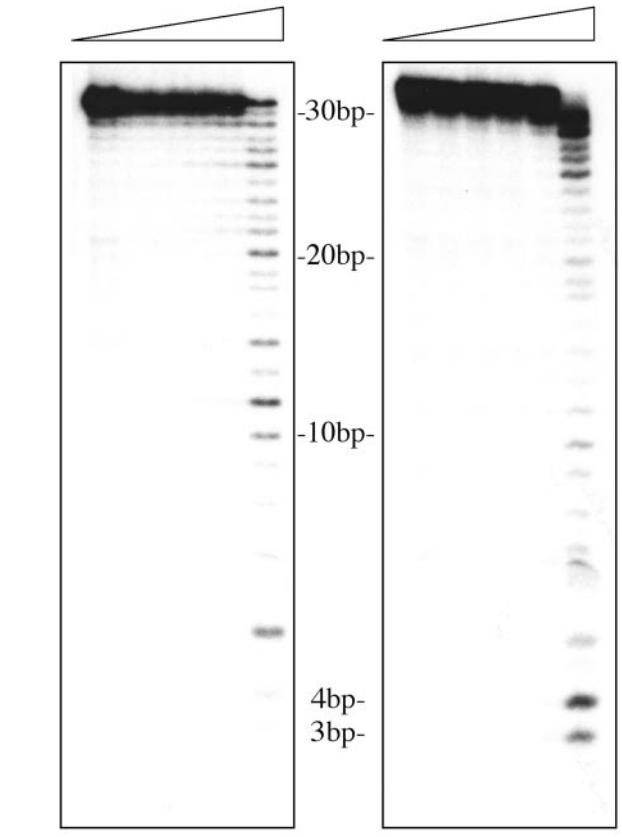


Figure 4. Incubation of different amounts of DmGEN protein in nuclease assays. Ramps indicate increasing amounts of DmGEN in 20 µl of reaction mixture (0, 1.0, 2.0, 4.1, 8.1, 16.2 ng). DmGEN protein was incubated with 100 fmol of ssDNA (**A**) or dsDNA (**B**) for 30 min at 37°C.

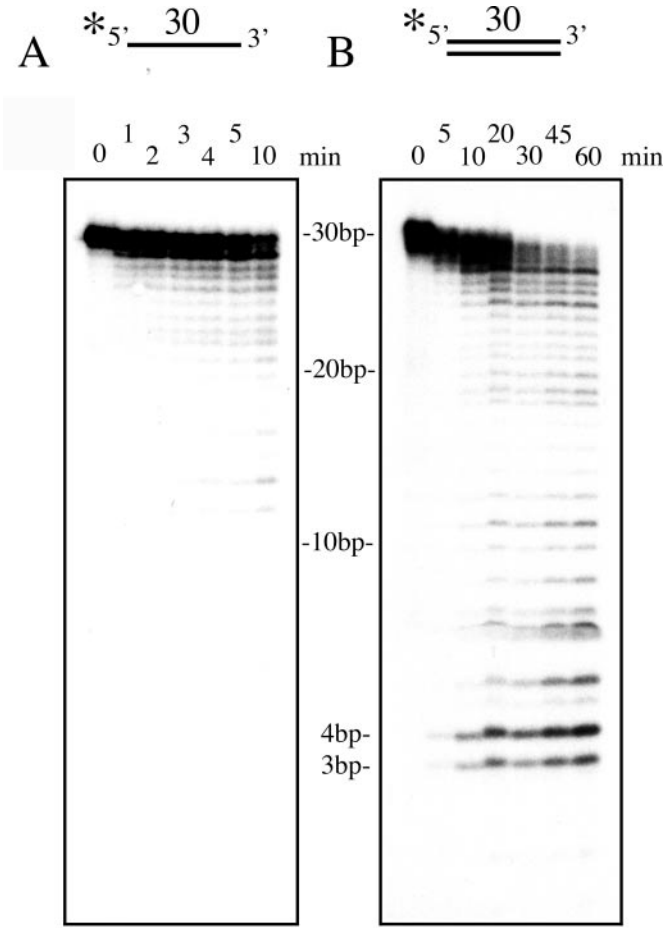


Figure 5. Nuclease activity of DmGEN protein (16.2 ng) on linear DNA substrates. Time course experiments were performed. The substrates are depicted schematically in each panel. The asterisk indicates the position of the radiolabel. Substrate and cleavage product sizes were as indicated. (**A**) The substrate was single-strand labeled at the 5′ end. (**B**) The substrate was double-strand labeled at the 5′ end.

Endonuclease and exonuclease activity of the DmGEN protein

First, we determined the optimal concentration of DmGEN for the nuclease assay to be 16.2 mg/ml (Figure 4). We then measured the nuclease activity of DmGEN against various radiolabeled DNA substrates in more detail. In this assay, terminally labeled substrate was incubated with the enzyme and activity was monitored by the appearance of bands. DmGEN cleaved single- and double-stranded linear DNA substrates to generate a ladder of labeled products resulting from 3′–5′ exonuclease digestion (Figures 5 and 6). One ^{32}P-labeled nucleotide was added to the 3′ end using TdT and [α-^{32}P]ddATP (Table 1). Based on the release of mononucleotides, these products were generated by the 3′–5′ exonuclease activity of DmGEN (Figure 6).

There were also some bands of 3 and 4 bp considered to be generated by endonuclease activity when the substrates were double-stranded DNA (Figures 4B and 5B). To determine their nature more precisely, we attempted to inhibit the exonuclease activity by using a substrate with a primer having a biotinylated base at the 3′ end. Quantitative biotinylation of the substrate is evident because the biotinylation changes the gel mobility of the substrate. Streptavidin is a 60 kDa protein that recognizes and tightly binds biotin (21). Substrates were incubated with streptavidin in 25-fold molar excess, resulting in quantitative gel shifts of the biotinylated substrates. This demonstrated that the protein was bound to all of the biotinylated oligonucleotides. Then, we conducted this experiment for DmGEN on single- and double-stranded DNA substrates with biotin and streptavidin (Figure 7). As a result, the 3′–5′ exonuclease activity of DmGEN was completely inhibited and there were only endonucleolytic products. DmGEN was confirmed to cleave double-stranded DNA 3 or 4 bases from the 5′ end by endonuclease activity (Figure 7B). The cleavage was not sequence-specific, because the results were relatively similar among nuclease assays using substrates with the same length but different sequences (Tables 1 and 2).

Based on previous studies, RAD2 family proteins have structure-specific endonuclease activity. For example, XPG (Class I) incises the target strand 3′ to the bubble-like, damage-containing structure. FEN-1 (Class II) exhibits a flap endonuclease activity for bifurcated DNA structures. Exo-1 (Class III) operates as a 5′-nuclease (i.e. either a 5′-flap endonuclease or a 5′–3′ exonuclease). However,

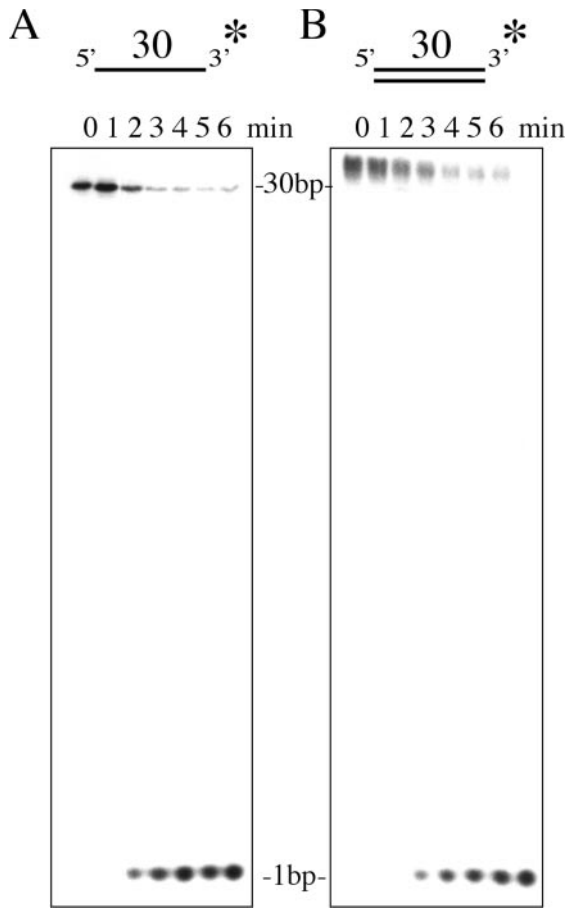


Figure 6. Nuclease activity of DmGEN protein (16.2 ng) on linear DNA substrates. Time course experiments were performed. The substrates are depicted schematically in each panel. The asterisk indicates the position of the radiolabel. Substrate and cleavage product sizes were as indicated. (**A**) The substrate was single-strand labeled at the 3′ end. (**B**) The substrate was double-strand labeled at the 3′ end.

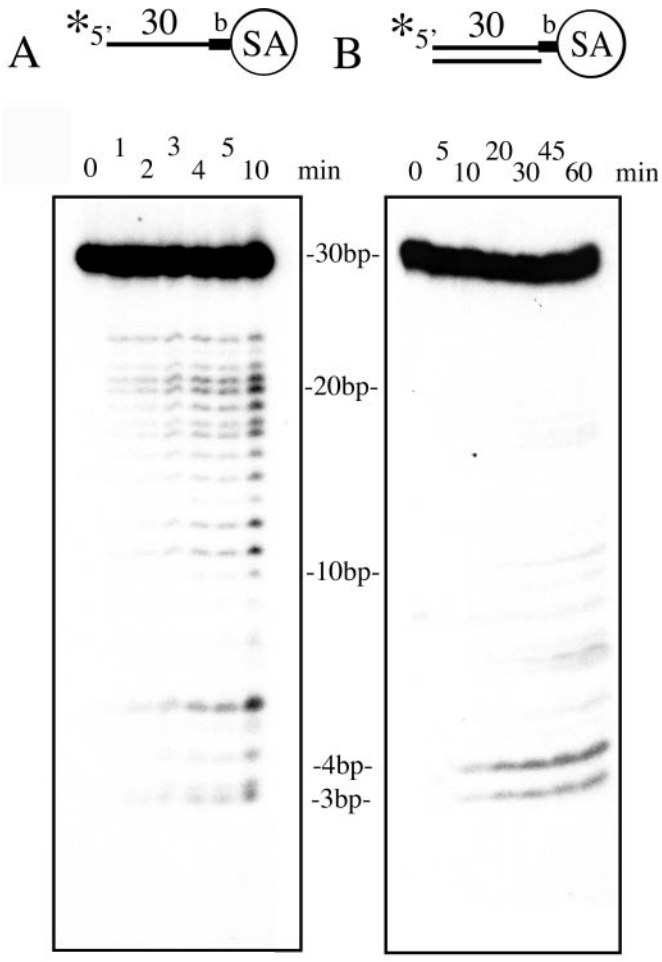


Figure 7. Effects of biotin adducts and streptavidin binding to single- and double-stranded substrates on the exonuclease activity of DmGEN protein (16.2 ng). Time course experiments were performed. The substrates are depicted schematically in each panel. The asterisk indicates the position of the radiolabel. Substrate and cleavage product sizes were as indicated. (**A**) The substrate was single-strand labeled at the 5′ end. (**B**) The substrate was double-strand labeled at the 5′ end.

DmGEN did not have such structure-specific endonuclease activity as a flap endonuclease (data not shown). In addition, DmGEN could not cleave a 'bubble' structure consisting of a 30-nucleotide unpaired region flanked by duplex regions of 30 bp (data not shown). These results provide further evidence that DmGEN belongs to a new class of RAD2 family nucleases.

DmGEN exhibits 5′–3′ exonuclease activity on gapped or nicked double-stranded substrates

To determine the substrate specificity of DmGEN, we produced gapped and nicked double-stranded DNA substrates. DmGEN was found to exhibit 5′–3′ exonuclease activity on both, as determined by the release of mononucleotides (Figure 8). Moreover, the 5′–3′ exonuclease activity gradually rose as the gap narrowed. This is a novel characteristic, because no other RAD2 family proteins have such a nuclease function except FEN-1, which has been demonstrated to act (i) as a flap endonuclease and (ii) as a nick-specific 5′–3′ exonuclease.

To determine the activity of DmGEN for gapped and nicked substrates in more detail, we produced substrates radiolabeled at the other side of the gap and nick. It is shown that both 3′–5′ exonuclease activity and endonuclease cleavage 3 or 4 bp from the 5′ end were still exhibited (data not shown). This suggests that DmGEN cleaves in both directions, 5′–3′ and 3′–5′, from gapped and nicked positions.

Function of the DmGEN protein

Based on the data from this study, the cleavage sites on various substrates of DmGEN are summarized in Figure 9. DmGEN has several nuclease functions, acting as an endonuclease and a 5′–3′ or 3′–5′ exonuclease, and these activities vary depending on the substrate, contrasting with the findings from previous nuclease studies. For example, with single-stranded breaks as occur *in vivo*, DmGEN exhibits nick- or gap-dependent 5′–3′ and 3′–5′ endonuclease activity for excising damaged or mismatch regions. With double-stranded breaks, DmGEN exhibits specfic 5′–3′ endonuclease activity 3 or 4 bases from the 5′ end, and modifies the break point. This modification may give signals to the monitors of repair pathways, or for other biological functions such as recombination and

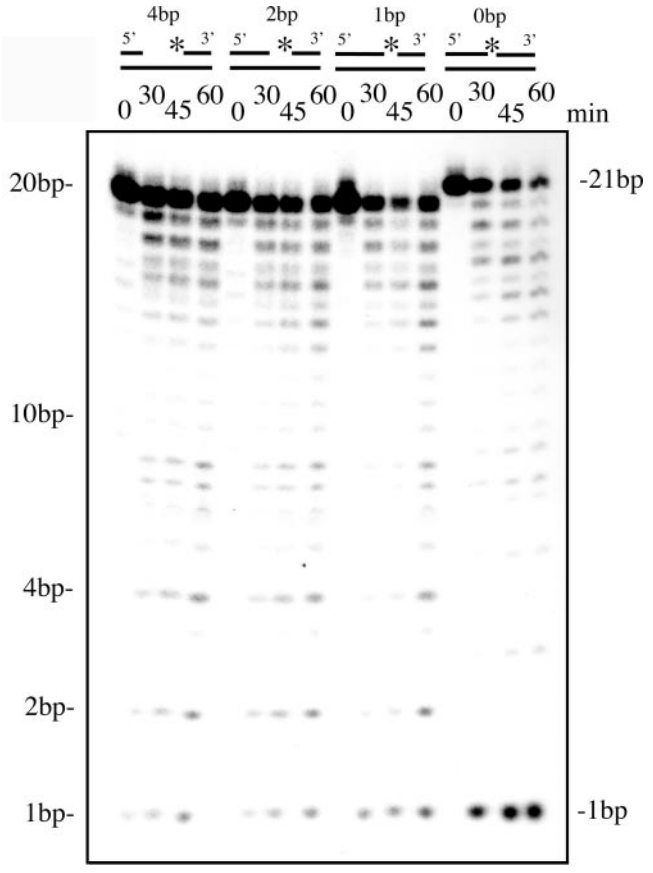


Figure 8. Nuclease activity of DmGEN protein (16.2 ng) on different substrates. Time course experiments were performed. Ramps indicate an increase in the reaction time (0, 30, 45 and 60 min). The substrates are depicted schematically in each panel. The asterisk indicates the position of the radiolabel. The sizes of gaps and cleavage products were as indicated.

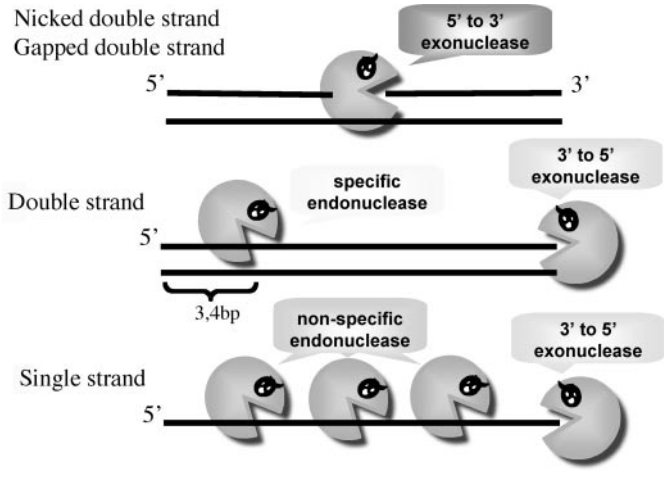


Figure 9. Summary of cleavage sites.

apoptosis. Double-stranded breaks in recombination are well known to occur (22–24), and apoptosis and nucleases are also closely related (25,26). It is seen from the present findings on the biochemical properties of this novel RAD2 family endonuclease that DmGEN may play specific roles in repair processes.

ACKNOWLEDGEMENTS

We thank M. Kubota for his expert advice. We are also grateful to M. Oshige, I. Sakimoto, H. Tiku, Y. Uchiyama and N. Kasai of Tokyo University of Science for technical advice and helpful discussions.

REFERENCES

1. Kimura,S., Kai,M., Kobayashi,H., Suzuki,A., Morioka,H., Otsuka,E. and Sakaguchi,K. (1997) A structure-specific endonuclease from cauliflower (*Brassica oleracea* var. *botrytis*) inflorescence. *Nucleic Acids Res*., **25**, 4970–4976.
2. Kimura,S., Suzuki,T., Yanagawa,Y., Yamamoto,T., Nakagawa,H., Tanaka,I., Hashimoto,J. and Sakaguchi,K. (2001) Characterization of plant proliferating cell nuclear antigen (PCNA) and flap endonuclease-1 (FEN-1), and their distribution in mitotic and meiotic cell cycles. *Plant J*., **28**, 643–653.
3. Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and Sakaguchi,K. (2000) Plant homologue of flap endonuclease-1: molecular cloning, characterization, and evidence of expression in meristematic tissues. *Plant Mol. Biol*., **42**, 415–427.
4. Furukawa,T., Kimura,S., Ishibashi,T., Mori,Y., Hashimoto,J. and Sakaguchi,K. (2003) OsSEND-1: a new RAD2 nuclease family member in higher plants. *Plant Mol. Biol*., **51**, 59–70.
5. Calleja,F.M., Nivard,M.J. and Eeken,J.C. (2001) Induced mutagenic effects in the nucleotide excision repair deficient Drosophila mutant mus201(D1), expressing a truncated XPG protein. *Mutat. Res*., **461**, 279–288.
6. Constantinou,A., Gunz,D., Evans,E., Lalle,P., Bates,P.A., Wood,R.D. and Clarkson,S.G. (1999) Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair. *J. Biol. Chem*., **274**, 5637–5648.
7. Davies,A.A., Friedberg,E.C., Tomkinson,A.E., Wood,R.D. and West,S.C. (1995) Role of the Rad1 and Rad10 proteins in nucleotide excision repair and recombination. *J. Biol. Chem*., **270**, 24638–24641.
8. Evans,E., Fellows,J., Coffer,A. and Wood,R.D. (1997) Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein *Embo. J*., **16**, 625–638.
9. Habraken,Y., Sung,P., Prakash,L. and Prakash,S. (1993) Yeast excision repair gene RAD2 encodes a single-stranded DNA endonuclease. *Nature*, **366**, 365–368.
10. Lieber,M.R. (1997) The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays*, **19**, 233–240.
11. O'Donovan,A., Davies,A.A., Moggs,J.G., West,S.C. and Wood,R.D. (1994) XPG endonuclease makes the 3′ incision in human DNA nucleotide excision repair. *Nature*, **371**, 432–435.
12. Sekelsky,J.J., Hollis,K.J., Eimerl,A.I., Burtis,K.C. and Hawley,R.S. (2000) Nucleotide excision repair endonuclease genes in *Drosophila melanogaster*. *Mutat. Res*., **459**, 219–228.
13. Vogel,E.W. and Nivard,M.J. (2001) Phenotypes of Drosophila homologs of human XPF and XPG to chemically-induced DNA modifications. *Mutat. Res*., **476**, 149–165.
14. Alleva,J.L. and Doetsch,P.W. (1998) Characterization of *Schizosaccharomyces pombe* RAD2 protein, a FEN-1 homolog. *Nucleic Acids Res*., **26**, 3645–3650.
15. Hiraoka,L.R., Harrington,J.J., Gerhard,D.S., Lieber,M.R. and Hsieh,C.L. (1995) Sequence of human FEN-1, a structure-specific endonuclease, and chromosomal localization of the gene (*FEN1*) in mouse and human. *Genomics*, **25**, 220–225.
16. Murray,J.M., Tavassoli,M., al-Harithy,R., Sheldrick,K.S., Lehmann,A.R., Carr,A.M. and Watts,F.Z. (1994) Structural and functional conservation of the human homolog of the *Schizosaccharomyces pombe RAD2* gene, which is required for chromosome segregation and recovery from DNA damage. *Mol. Cell. Biol*., **14**, 4878–4888.
17. Reagan,M.S., Pittenger,C., Siede,W. and Friedberg,E.C. (1995) Characterization of a mutant strain of *Saccharomyces cerevisiae*

with a deletion of the *RAD27* gene, a structural homolog of the RAD2 nucleotide excision repair gene. *J. Bacteriol.*, **177**, 364–371.
18. Lee,B.I. and Wilson,D.M.,III (1999) The RAD2 domain of human exonuclease 1 exhibits 5′ to 3′ exonuclease and flap structure-specific endonuclease activities. *J. Biol. Chem.*, **274**, 37763–37769.
19. Wilson,D.M.,III, Carney,J.P., Coleman,M.A., Adamson,A.W., Christensen,M. and Lamerdin,J.E. (1998) Hex1: a new human RAD2 nuclease family member with homology to yeast exonuclease 1. *Nucleic Acids Res.*, **26**, 3762–3768.
20. Spradling,A.C. and Rubin,G.M. (1981) *Drosophila* genome organization: conserved and dynamic aspects. *Annu. Rev. Genet.*, **15**, 219–264.
21. Green,N.M. and Joynson,M.A. (1970) A preliminary crystallographic investigation of avidin. *Biochem. J.*, **118**, 71–72.
22. Gonzalez-Barrera,S., Cortes-Ledesma,F., Wellinger,R.E. and Aguilera,A. (2003) Equal sister chromatid exchange is a major mechanism of double-strand break repair in yeast. *Mol. Cell*, **11**, 1661–1671.
23. Storici,F., Durham,C.L., Gordenin,D.A. and Resnick,M.A. (2003) Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. *Proc. Natl Acad. Sci. USA*, **100**, 14994–14999.
24. Lundin,C., Erixon,K., Arnaudeau,C., Schultz,N., Jenssen,D., Meuth,M. and Helleday,T. (2002) Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells. *Mol. Cell. Biol.*, **22**, 5869–5878.
25. Schafer,P., Scholz,S.R., Gimadutdinow,O., Cymerman,I.A., Bujnicki,J.M., Ruiz-Carrillo,A., Pingoud,A. and Meiss,G. (2004) Structural and functional characterization of mitochondrial EndoG, a sugar non-specific nuclease which plays an important role during apoptosis. *J. Mol. Biol.*, **338**, 217–228.
26. Peitsch,M.C., Polzar,B., Stephan,H., Crompton,T., MacDonald,H.R., Mannherz,H.G. and Tschopp,J. (1993) Characterization of the endogenous deoxyribonuclease involved in nuclear DNA degradation during apoptosis (programmed cell death). *EMBO J.*, **12**, 371–377.