Online journal ISSN - 1676-5680
GMR Genetics and Molecular Research
EVOLUTION AND TECHNOLOGY

Review

Escherichia coli as a model system to study DNA repair genes of eukaryotic organisms

Luiz Augusto-Pinto, Carlos Gustavo Regis da Silva, Débora de Oliveira Lopes, Alice Machado-Silva and Carlos Renato Machado

Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, 30161-970 Belo Horizonte, MG, Brasil
Corresponding author: C.R. Machado
E-mail: crmachad@mono.icb.ufmg.br

Genet. Mol. Res. 2 (1): 77-91 (2003)
Received November 27, 2002
Published March 31, 2003

ABSTRACT. The bacteria *Escherichia coli* has been widely employed in studies of eukaryotic DNA repair genes. Several eukaryotic genes have been cloned by functional complementation of mutant lineages of *E. coli*. We examined the similarities and differences among bacterial and eukaryotic DNA repair systems. Based on these data, we examined tools used for gene cloning and functional studies of DNA repair in eukaryotes, using this bacterial system as a model.

Key words: DNA repair, Functional complementation, *Escherichia coli*

INTRODUCTION

The bacterial genome encodes approximately 115 proteins involved in at least one aspect of DNA repair (Aravind et al., 1999). This metabolic scenario has been maintained throughout evolution, since all eukaryotic organisms, as well as bacteria, must safeguard against DNA damage, which would be incompatible with life if DNA repair machinery did not make the necessary corrections. When we analyze bacterial and eukaryotic DNA repair systems, we can see that there is a great deal of similarity in the ways these groups of organisms remove damage from DNA. The genome of both kinds of organisms is subjected to the same chemical reactions, which generate the same damaged substrates for the proteins involved with DNA repair. It is therefore expected that the biochemical pathways employed by eukaryotic cells to respond to such damage would be similar to the bacterial ones. For instance, the loss of purines and pyrimidines from DNA produces apurinic or apyrimidinic (AP) sites that are repaired after the action of AP endonucleases, which catalyze the incision of DNA exclusively at AP sites, thereby preparing DNA for subsequent excision, repair synthesis, and strand sealing. This pathway is the same in bacterial and eukaryotic organisms and, furthermore, there is considerable conservation of amino acids among the AP endonucleases of these two types of organisms; this degree of conservation can be as high as 41% (Popoff et al., 1990). However, there are also differences in the DNA repair systems of bacteria and eukaryotes. In the latter the metabolic scenarios are more complex, since some proteins have multiple functions in cell cycle checkpoints, chromatin assembly, DNA repair and DNA replication. In comparison, a pathway that in bacteria involves only five major proteins is constituted of at least 15 other proteins in eukaryotes, as observed in the nucleotide excision repair pathway (Cleaver et al., 2001).

In this review, we describe the use of bacteria as a tool for cloning and for functional studies of eukaryote DNA repair genes.

THE MAJOR PATHWAYS OF DNA REPAIR

All proteins, from bacterial to human, involved in DNA repair can be grouped into the following major functional categories: a) damage reversal, b) base excision repair (BER), c) nucleotide excision repair (NER), and d) mismatch repair (MMR). We describe the fundamental and most recently discovered molecular aspects of each pathway, using the bacterial system as a model for DNA repair in eukaryotic studies, which are considered later in this review.

DAMAGE REVERSAL

The direct reversal of DNA damage is by far the simplest repair mechanism described until now. It consists of a single-polypeptide chain, with enzymatic properties, binding to the damage and restoring the DNA genome to its normal state in a single-reaction step. The major polypeptides involved in this pathway are: a) DNA photolyase, which is the enzyme responsible for removing cyclobutane pyrimidine dimers from DNA in a light-dependent process denominated photoreactivation (Carell et al., 2001), and b) O6-methylguanine-DNA methyltransferase I and II, also called DNA-alkyltransferases, which remove the modified bases O6-alkylguanine and O4-alkylthymine and the backbone-modified phosphate alkylphosphotriesters (methylphosphotriesters) (Pegg, 2000) from DNA. The photolyase protein is not found in all

living cells. Most studies have consistently failed to provide biological or biochemical evidence of photoreactivation in primates. However, the DNA-alkyltransferases are widespread in nature. Activity of this group of enzymes has been identified in extracts of *Aspergillus nidulans*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, fish and in mammalian cells (Friedberg et al., 1995).

EXCISION OF DNA DAMAGE

There are three major pathways of DNA excision repair: i) BER, ii) NER and iii) MMR. In these reactions a nucleotide segment containing base damage, double-helix distortion or mispaired bases is replaced by the normal nucleotide sequence in a new DNA polymerase synthesis process. All of these pathways have been characterized in both bacterial and eukaryotic organisms.

Base excision repair

BER is initiated by DNA glycosylases, which catalyze the hydrolysis of the N-glycosylic bonds, linking particular types of chemically altered bases to the desoxyribose-phosphate backbone. Thus, DNA damage is excised as free bases, generating sites of base loss called apurinic or apyrimidinic (AP) sites. Another means of AP site generation is the depurination or depyrimidation of DNA, due to spontaneous hydrolysis of N-glycosylic bonds. The AP sites are substrates for AP endonucleases. These enzymes produce incisions in duplex DNA as a result of the hydrolysis of a phosphodiester bond immediately 5' or 3' to each AP site. The ribose-phosphate backbone is then removed from the DNA through the action of a specific exonuclease called deoxyribophosphodiesterase or dRpase. Finally, the DNA polymerase and a ligase catalyze the incorporation of a specific deoxyribonucleotide into the repaired site, enabling correct base pairing (Figure 1) (Friedberg et al., 1995).

Base excision repair is a multiple enzymatic DNA repair mechanism in which each enzyme works separately; this scenario has been maintained during evolution. Known from bacteria, several DNA glycosylases and AP endonucleases have also been described in lower and higher eukaryotes, such as *Trypanosoma cruzi* (Perez et al., 1999), *Leishmania major* (Perez et al., 1999), yeast, fish, plants, *C. elegans*, *Drosophila* and mammals (Friedberg et al., 1995). BER is essential to protect DNA from various types of lesions, such as: uracil, hydroxymethyluracil, methylcytosine, hypoxanthine, G-T mispairs, 3-methyladenine, 7-methylguanine, 3-methylguanine, formamidopyrimidine, 8-hydroxyguanine, 5,6-hydrated thymine, and pyrimidine dimers (Friedberg et al., 1995). Each of these is recognized by a specific DNA glycosylase.

Nucleotide excision repair

Several types of agents generate bulky base adducts in DNA, leading to a significant distortion of the DNA helix. The most widely studied of these DNA damaging agents is UV radiation, responsible for thymine dimers, which produce a bend of ~30^{o} in the DNA (Husain et al., 1988). Some chemical agents form DNA cross-links, which are particularly hazardous. These cross-links produce conformational distortions in DNA; they are substrates for DNA endonucleases that make an incision in DNA, several nucleotides to each side of the damage,

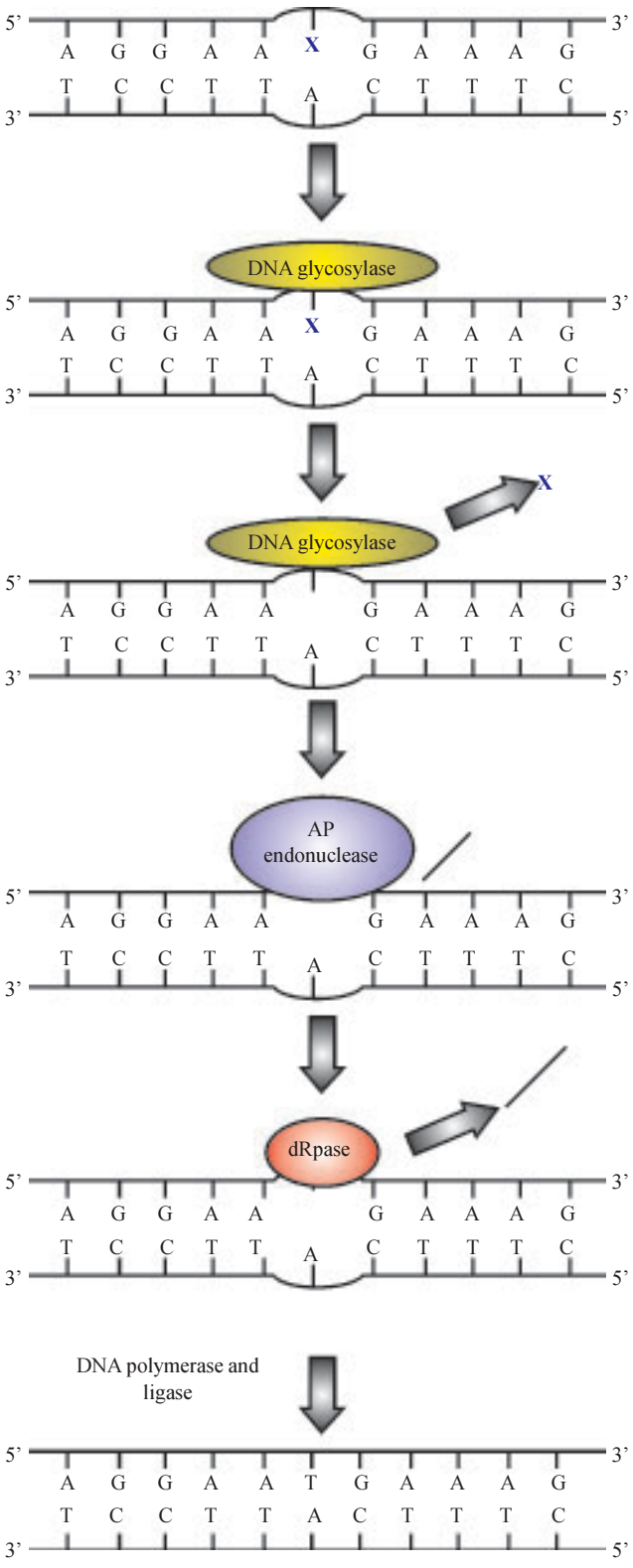

Figure 1. Schematic representation of base excision repair. Base damage (represented as X) is recognized and removed by specific DNA glycosylases (in yellow), generating an AP site. AP endonuclease (shown in blue) then hydrolyses the phosphodiester bond immediately 5' or 3' to each AP site and the ose-phosphate backbone is removed from DNA through the action of dRpase (in red). Finally, the resulting single nucleotide gap is filled by the action of DNA polymerase, and a ligase seals the repaired strand. dRpase = deoxyribophosphodiesterase.

generating a potential oligonucleotide fragment. Subsequent helicase reactions promote the excision of this fragment. The resulting gap is filled by DNA polymerase synthesis and covalently sealed by DNA ligase. These sequential enzymatic reactions, initiated by a specific endonuclease that recognizes the DNA distortion, are called NER. This pathway was originally described in the repair of DNA in cells exposed to UV radiation, which produces lesions such as pyrimidine dimers and 6-4 photoproducts that are enzymatically excised from DNA as intact nucleotides rather than as free bases. NER is a much more complex biochemical process than BER, especially in eukaryotic cells. Several gene products are required in a multiple step process, during which the ordered assembly of DNA proteins provides an enzymatic complex that discriminates damaged from undamaged DNA.

In *Escherichia coli* there are three specific proteins, called UvrA, B and C, involved in lesion recognition and endonuclease incision. An assembly of two UvrA and one UvrB protein binds to DNA nonspecifically; this assembly translocates unidirectionally in the genome, driven by the energy of ATP hydrolysis (Grossman and Yeung, 1990). This mechanism allows constant monitoring for base damage in living cells. It has been suggested that when a damaged site is found, UvrA proteins dissociate and a stable UvrB-DNA complex is formed (Grossman and Yeung, 1990). Then the UvrC protein associates with high affinity to the UvrB-DNA site and induces a conformational change that enables bound UvrB protein to nick the DNA at the fourth nucleotide, 3' to the site of damage (Grossman and Yeung, 1990). This reaction requires the binding of ATP by UvrB protein, however, no ATP hydrolysis occurs at this time (Grossman and Yeung, 1990). Following the 3' incision, UvrC protein catalyzes nicking of the DNA at the seventh nucleotide, 5' to the damage (Grossman and Yeung, 1990). Thus, a stretch of 12 nucleotides is held in the DNA by only hydrogen bonds. This fragment is released by UvrD helicase action, generating a gap that is finally submitted to repair synthesis (Figure 2) (Grossman and Yeung, 1990).

Mismatch repair

The long patch MMR machinery is present in several organisms. Its function is to remove base substitution and frameshift mismatches that escape from DNA polymerase proof-reading activity after DNA replication, increasing DNA replication fidelity 100- to 1000-fold (Modrich and Lahue, 1986). In *E. coli*, the factors that are exclusively involved in MMR are encoded by *mutS*, *mutL* and *mutH* genes (Lahue et al., 1989).

The MutS protein homodimers recognize and bind specifically to base-base mispairing and insertion/deletion loop-outs (IDL). Then, MutS, in association with MutL protein homodimers, activates the MutH protein to make an excision-initiating nick in the unmethylated, newly synthesized strand. The nicked strand containing the mismatch or IDL is excised by exonucleases and resynthesized by DNA polymerase and DNA ligase (Figure 3) (Marra and Schär, 1999).

The current picture of MMR in eukaryotic cells resembles that of *E. coli* to a great extent, but with two important differences. The first is related to strand discrimination. In eukaryotes the hemimethylation status of newly replicated DNA does not play a role in directing the MMR. No MutH homologue has been identified and it has been proposed that strand discrimination is mediated by strand discontinuities in the newly synthesized DNA. The second fundamental difference is that the MutS and MutL functional homologues are heterodimeric

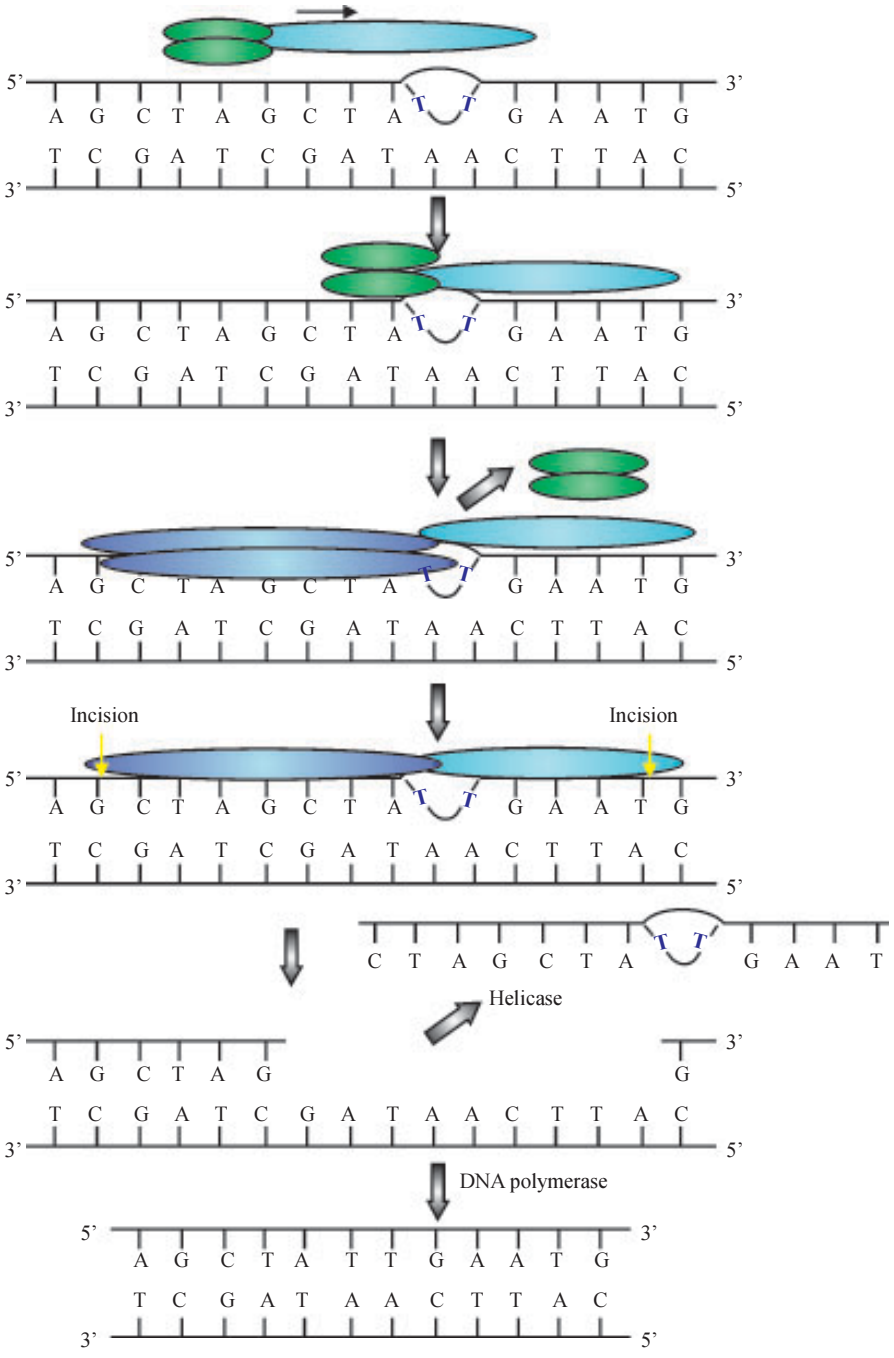

Figure 2. Schematic representation of nucleotide excision repair. The UvrAB heterodimer scans the DNA searching for large distortions in the helix such as the ones caused by pyrimidine dimers. Once a damaged site is found, UvrA proteins (dark green) dissociate, and a stable UvrB-DNA (light green) complex is formed. UvrC (blue) associates to bound UvrB and enables UvrB protein to nick the DNA at the fourth nucleotide 3' to the site of damage. Following the 3' incision, UvrC protein catalyzes nicking of the DNA at the seventh nucleotide, 5' to the damage. The potential oligonucleotide fragment that is generated is removed by a helicase. The remaining gap is filled up by polymerase synthesis and repair is completed by ligase.

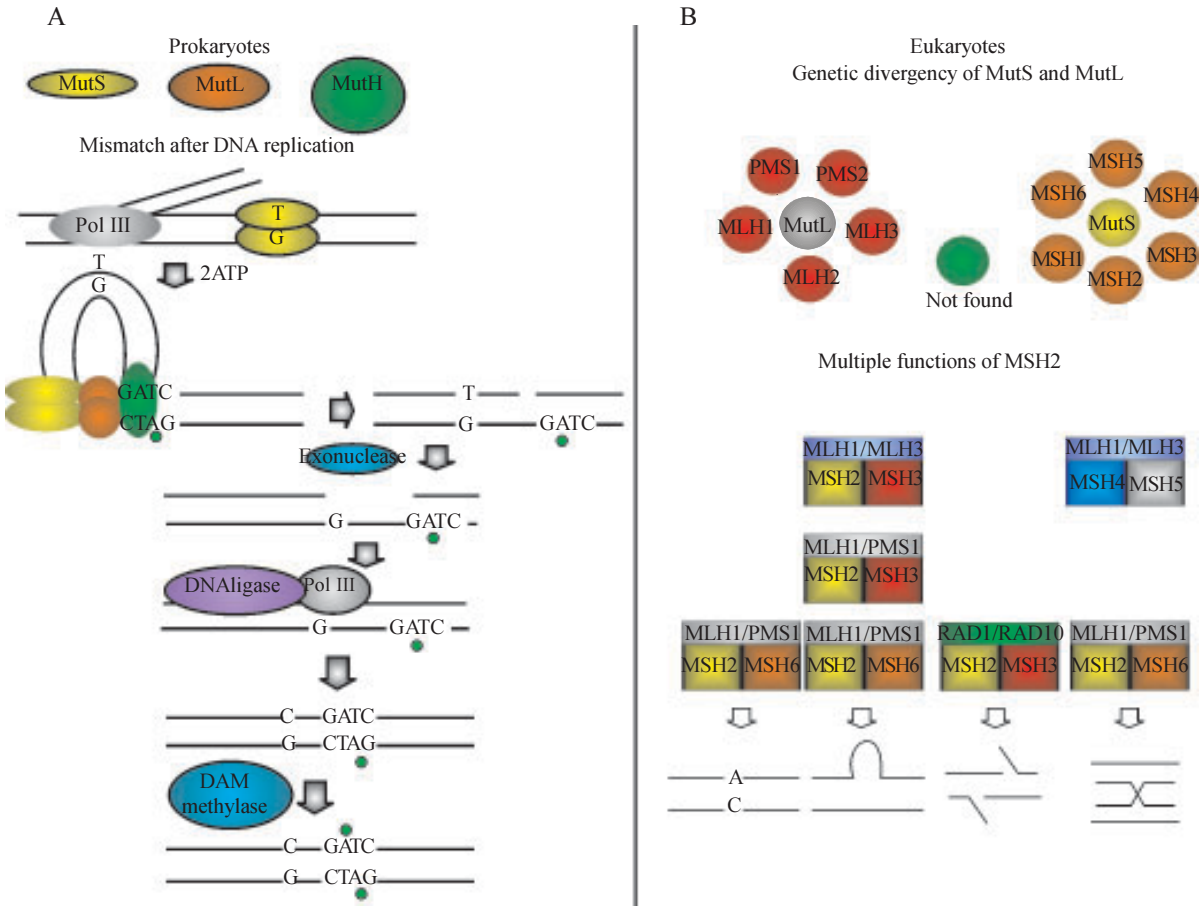

Figure 3. A, Schematic representation of mismatch repair in prokaryotes. The MutS homodimer protein binds to the DNA mismatch and makes a loop in DNA using the energy of hydrolysis of two ATP molecules. The MutL homodimer protein then associates with the bottom of this loop and activates the endonuclease MutH. The MutH protein only nicks the unmethylated strand, which contains the incorrect base. Afterwards, the cleaved strand is submitted to exonuclease activity, DNA resynthesis and ligation. **B,** Schematic representation of the eukaryotic mismatch repair system. The MutS proteins diverged into six orthologous genes, while MutL diverged into five other genes; these are denominated MSH and MLH, respectively. The MutH protein is not found in eukaryotes. The MSH and MLH proteins interact as a functional heterocomplex and repair several types of substrates, such as mismatches, single-strand loops generated during microsatellite replication, DNA double-strand breaks, and holiday junctions from meiotic crossing-over.

rather than homodimeric. At least six MutS homologues and five MutL homologues, referred to as MSH and MLH, respectively, have been identified in eukaryotes. The best characterized of these factors are MSH2, MSH3 and MSH6, which are involved in MMR in the nucleus. MSH2-MSH6 heterodimers recognize and repair base mismatches and loops of up to two bases, whereas MSH2-MSH3 heterodimers recognize loop-outs of different sizes (Drumond et al., 1995; Marsischky et al., 1996). MSH4 and MSH5 proteins constitute another MSH heterodimer, which, however, does not participate in MMR, but instead is involved in meiotic crossing-over and chromosome segregation (Figure 3) (Nakagawa et al., 1999). MSH1 is targeted to the mitochondria and is necessary for mitochondrial stability in yeast (Reenan and Kolodner, 1992).

CLONING AND CHARACTERIZATION OF EUKARYOTIC DNA REPAIR GENES USING *ESCHERICHIA COLI*

Complementation with homologous genes

In 1985, Sancar described the use of *E. coli* as an instrument to study eukaryotic DNA repair genes. A fragment carrying the photolyase gene (PHR1) of *Saccharomyces cerevisiae* was cloned into an *E. coli* expression vector and introduced into *E. coli* strains deficient in DNA photolyase. Complementation of the *E. coli* phr-1 mutation was observed. This was the first time that a complementation approach was used to study eukaryotic DNA repair genes in *E. coli*.

After this, several other eukaryotic genes were identified and characterized by complementation assays in bacteria deficient in DNA repair. They were isolated from *E. coli* mutant cells transformed with a cDNA expression library, due to their ability to restore, at least partially, the resistance of the mutants to treatment with DNA damaging agents (Figure 4).

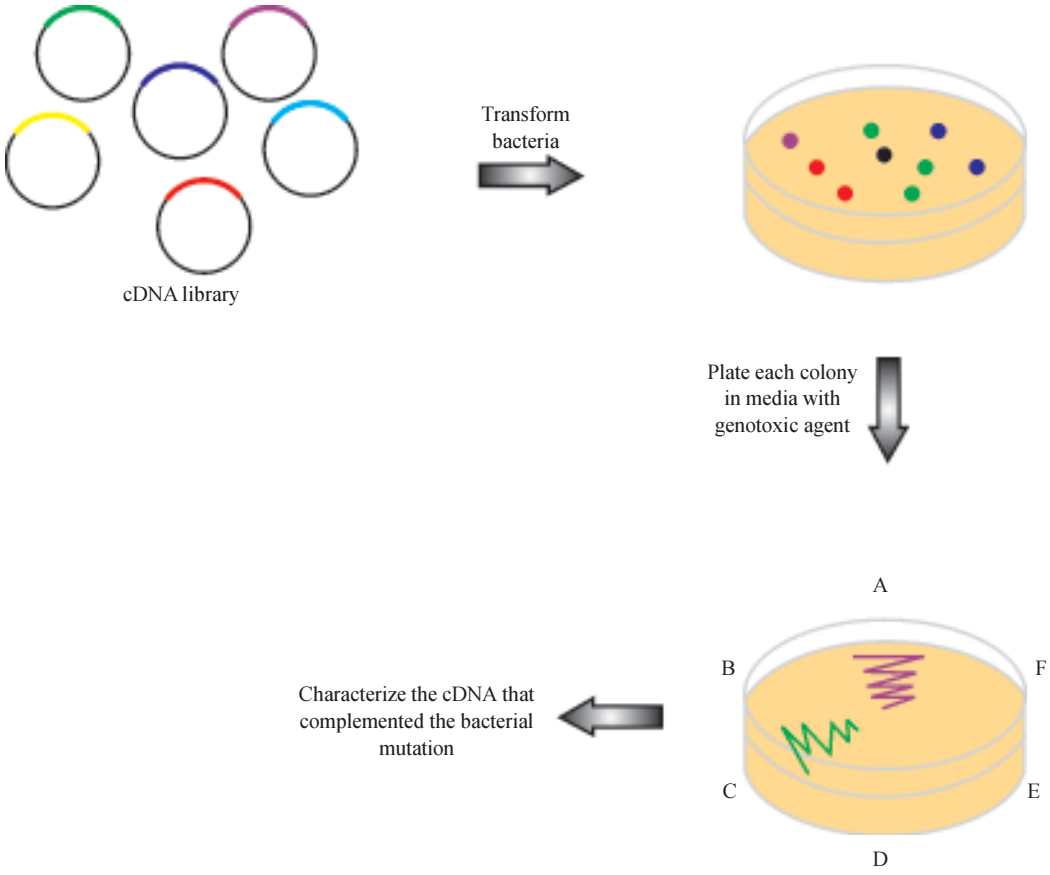

Figure 4. Schematic representation of a functional complementation assay. A cDNA library is used to transform bacteria deficient in DNA repair. The colonies are plated in a selective media containing a genotoxic agent. Clones are thus selected by their ability to restore, at least partially, the resistance of the mutants to treatment with DNA damaging agents. The cDNAs that complemented the bacterial mutation are then characterized.

Chen et al. (1989) used this system to clone an *S. cerevisiae* gene. They used *E. coli* strains MV1932 (*alk*A1 and *tag*) and MV1902 (*alk*A150::λpSG1 and *tag*), which have mutant *alk*A and *tag* genes. These bacterial cells are sensitive to DNA damage by methylating agents, such as methyl methanesulfonate (MMS), due to the absence of either the *AlkA* or the *Tag* gene product, respectively. The 3-methyladenine DNA glycosylase (*MAG*) from *S. cerevisiae* (*ScMAG*) was cloned from a genomic library, as it complemented the bacterial gene function, restoring the ability of the bacteria to survive in the presence of MMS. An additional enzymatic assay showed that extracts from deficient bacteria harboring the cloned yeast gene were able to excise modified bases from [H^3]-DMS-treated calf thymus DNA. Chromatographic procedures demonstrated that the excised base was 3-methyladenine. Sequence analyses established the identity of this gene as *ScMAG* and its expression in yeast-protected cells against the lethal effects of alkylating agents (Chen et al., 1990). This bacterial complementation strategy has also allowed the cloning of other *MAG* genes. Human *MAG* (*hsMAG*) was isolated from a human liver cDNA library (Samson et al., 1991). Its cDNA sequence has an open-reading frame of 894 base pairs and a sequence of 176 amino acids, sharing 85% identity with a rat glycosylase polypeptide. Enzymatic analyses showed the ability of three variants of the HsMAG protein to release methylated bases from DNA (O'Connor, 1993). HPLC chromatography identified the modified bases as 3-methyladenine, 7-methylguanine and 3-methylguanine. *Arabidopsis thaliana*, *Schizosaccharomyces pombe* and rat *MAGs* have also been cloned and characterized using this bacterial system (O'Connor and Laval, 1990; Santerre and Britt, 1994; Memisoglu and Samson, 1996). In addition to *MAG* genes, an *S. cerevisiae* akyl DNA glycosylase has also been cloned from a genomic DNA library, using an *E. coli* strain with a phenotype similar to the *alk*A *tag* double mutant (Berdal et al., 1990).

Genetic manipulation of *E. coli* to create mutator phenotype strains could also be used to clone DNA repair genes. Hydroxyl radical ($OH^{\bullet}$) and singlet oxygen ($^1O_2$), produced endogenously by oxidative stress can generate complex modifications in DNA bases (Boiteux et al., 1992). One of these is the oxidized guanine, 7,8-dihydro-8-oxoguanine (8-OxoG), which can pair up with adenine instead of cytosine, causing a GC → TA transition. The 8-OxoG residues can be removed from DNA by the action of FPG protein and the adenine residue opposite 8-OxoG can be excised by MutY protein. Hence, disruption of *fpg* and *mutY* genes could be used to create a mutator phenotype strain in order to select for DNA repair genes involved in 8-OxoG excision. The *S. cerevisiae OGG1* gene (*ScOGG1*) was cloned this way (van der Kemp et al., 1996). A *fpg mutY*-double mutant *E. coli* strain was transformed with a yeast genomic DNA library and clones that partially suppressed the bacterial mutator phenotype were selected. The loss of the ability to generate mutants was visualized by the reduction in the number of colonies resistant to the antibiotic rifampicin. ScOGG1 protein purified from over expression in *E. coli* was used to confirm the enzymatic nature of this gene product in experiments with 8-OxoG excision from calf thymus DNA. Mouse and human homologues involved in 8-OxoG excision have also proven their ability to promote the loss of a mutator phenotype in a rifampicin-resistance assay, providing additional information for the characterization of OGG1 (Rosenquist et al., 1997; Aburatani et al., 1997; Arai et al., 1997; Sugimura et al., 1999). Kohno et al. (1998) have also used this bacterial complementation assay to study differences in 8-OxoG repair caused by human polymorphic *OGG1* genes. The gene product with a Ser at codon 326 had an increased efficiency of complementation compared to its Cys counterpart.

An interesting approach to study eukaryotic DNA repair genes is the employment of *E. coli* mutator or sensitive phenotype reversion in order to select and characterize mutant genes. Christians and Loeb (1996) have used a Δ*ada*-25 Cm^r *ogt*-1::Kan^r *E. coli* strain to search for mutations in human O^6-methylguanine DNA methyltransferase (*hsMGMT*) that could maintain the complementation phenotype in the deficient bacterial strain. This bacterial strain, which lacks the *ada* and *ogt* gene products, is sensitive to alkylating agents such as *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) and to the chemotherapeutic nitrosoureas. A plasmid library harboring the human MGMT gene, randomly mutated in a region surrounding the Cys active residue, was used to transform these deficient bacteria. Clones were selected by their ability to grow in the presence of MNN and also to reverse the mutator phenotype. As the chosen region is evolutionary conserved, few amino acid changes were found. Using this strategy of mutant identification in an *ada ogt* deficient bacterial strain, Xu-Welliver et al. (1999) tested for *hsMGMT* mutation at codon 160, which rendered the bacteria resistant to O^6-benzylguanine (BG). A cDNA from this gene was randomly mutated at this position and used to transform bacterial strains lacking alkyltransferase activity. As it is an inhibitor of human MGMT protein (Dolan et al., 1990; Dolan and Pegg, 1997), this drug has been used in clinical trials to improve the efficiency of akylating drugs such as BCNU and temozolomide (Friedman et al., 1998). Polymorphic G160R *hsMGMT* has been found in some individuals (Imai et al., 1995), and this polymorphism has been associated with BG resistance (Edara et al., 1996). Fourteen amino acid substitutions at codon 160 have been related to BG resistance (Xu-Welliver et al., 1999), all of them protecting the *E. coli* mutant strain against MNNG damage. Enzymatic analysis of bacterial extracts or purified protein have proven the effectiveness of these mutant proteins in repairing O^6-[^{3}H]-methylguanine, although with a lower activity than wild type extracts.

Bacterial complementation by eukaryotic genes could be used to study genetic polymorphisms or mutants, as has been shown for the *hsMGMT* gene. Another example of successful exploration of this approach involves DNA polymerase β (*pol β*) gene studies. Sweasy and Loeb (1992) showed the ability of rat *pol β* to complement the activity of DNA pol I from an *E. coli* strain, which has a mutation in the DNA pol I gene (*polA12)* that makes bacterial cells unable to grow at high temperatures. This complementation system has opened new perspectives to identify mutations in the rat *pol β* sequence that impair polymerase function. Rat *pol β* cDNA has been used to construct a library of *pol β* cDNA mutants by treatment with nitrous acid, PCR amplification and bacterial cloning (Sweasy and Loeb, 1993). Sensitive bacteria were transformed with the mutant cDNA library and clones with partial or total loss of polymerase activity were selected. A total of 1186 bacterial transformants were tested for their ability to grow in the presence of MMS at 42°C. Polymerase function was altered in 263 clones. Sequence analysis of 10 clones confirmed single-base substitutions. Fifty-six mutants were unable to grow in both conditions and the others only lost the ability to grow at 42°C. This work has demonstrated the possibility of employment of this *polA12* strain in mutant selection. Further mutant characterization was carried out to identify the low fidelity polymerase mutants that conferred the enzymatic mutator phenotype (Washington et al., 1997). Two methods were used to create the mutant library: the first one was nitrous acid treatment, followed by PCR amplification and bacterial cloning, and the other one was PCR under mutagenic conditions. After selection for ability to complement the *polA12* strain, positive clones were tested for promotion of *trp* mutation reversion, which could be associated with the mutator phenotype. Three rat *pol β* mutants were identified and characterized. One of them had a Y265C substitu-

tion and a mutation rate 30-fold higher than the wild type. This mutator enzymatic activity was detected by an *in vitro* activity assay. The other two had a P312S and a Y265H substitution, respectively, and the mutation rates were increased 10-fold in comparison to wild type. The mutation in the three clones that were analyzed is located in the protein's carboxyl terminus at an external position on the DNA-binding cleft. Similar to hMGMT, it was also possible to identify drug-resistant mutants of the rat *pol* β gene (Kosa and Sweasy, 1999). A cDNA library was generated in a mutagenic PCR condition that guaranteed mostly single-base changes and was used to transform a *polA12 E. coli* strain. The transformed bacteria were then challenged for growth in the presence of 3'-Azido-3'-deoxythymidine (AZT) at 42°C. The rat *pol* β is able to incorporate AZT into DNA during synthesis, causing cell death. The clones able to grow in the presence of AZT at 42°C kept their enzymatic activity, but they lost the ability to incorporate AZT. Two mutants were selected, D246V and R253M; they gave a more effective discrimination against AZT than the wild type protein. The D246V mutant had a 10-fold decrease in the catalytic efficiency of AZT incorporation but dTTP incorporation was unaffected. On the other hand, the R253M mutants showed a moderate efficiency in AZT incorporation, combined with a perturbation in dTTP incorporation. These amino acid substitutions are close to the nucleotide-binding pocket and could affect nucleotide discrimination and selection.

All the examples above are from damage reversal or base excision repair. This is due to the fact that the proteins involved in these pathways can work separately and do not need to interact with other proteins to accomplish their function. The study of genes involved in NER and MMR is more difficult. A good example is the study of the MSH2 gene, the eukaryotic homologue of the *E. coli* MMR gene MutS. It has been shown that the expression in *E. coli* of a eukaryotic protein related to the MutS family results in an increased mutation rate in the bacteria, due to lack of interaction of the heterologous protein with the normal bacterial repair machinery (Augusto-Pinto et al., 2001).

Complementation with non-homologous genes

Besides phenotype complementation with homologous genes, the use of DNA-repair-deficient *E. coli* strains to select for genes involved in this DNA metabolism pathway have provided a means for the selection and characterization of new genes that are involved in DNA repair but are not homologous with known DNA repair proteins.

Pang et al. (1992, 1993) employed this technique to obtain clones that complement *E. coli* strain *recA*$^-$, *uvr*$^-$, *phr*$^-$ (genes related to DNA repair). One of the isolated cDNAs partially complements recombination related (RecA$^-$) phenotypes. Another three cDNAs also partially corrected the *E. coli* mutant phenotypes, but their molecular roles were not clearly identified. Most of these cDNAs are apparently targeted to chloroplasts.

Among 840 survivors of heavily UV-irradiated mutants harboring plasmids derived from an Arabidopsis cDNA library, four unique plant cDNAs were identified. Two of them were specific for UV-light damage, and complemented both UvrB- and UvrC- phenotypes in the dark. These cDNAs showed no extensive amino acid homology with known DNA repair proteins. Although the light dependence of one cDNA activity was consistent with its identification as a photoreactivating enzyme, its predicted amino acid sequence did not resemble known photolyase sequences (Pang et al., 1992). The latter cDNA increased the resistance of *RecA-UvrB-Phr-* bacteria to mitomycin C and MMS as well as to UV light. This lack of specificity,

and its ability to increase resistance in both *UvrB*- and *UvrC*- mutants, suggested that cDNA activity might be complementing *RecA*- phenotypes. The partial complementation of three *RecA*- phenotypes could be due the activities of this gene in homologous recombination. The 395-amino acid open-reading frame encodes an apparent N-terminal chloroplast transit peptide and a putative 322-residue mature protein, with a conserved nucleotide-binding motif but otherwise little global homology with bacterial RecA proteins (Pang et al., 1993).

Using the same strategy, Machado et al. (1996) isolated a DNA repair-related gene from *Arabidopsis thaliana* that was able to complement bacteria defective in enzymes that play a part in the initial steps of base excision repair. The *E. coli* strain employed was the triple mutant BW535 (*xth,nfo,nth*), defective in the DNA repair endonucleases: exonuclease III, endonuclease IV and endonuclease III. These enzymes are involved in the first steps of base excision repair of DNA lesions normally mediated by oxidative and alkylation products (Doetsch and Cunningham, 1990). The mutant strain presents a hypersensitive phenotype to agents that produce oxygen radicals, such as hydrogen peroxide, and to alkylating agents, such as MMS (Cunningham et al., 1986). The screening was performed by transfecting a cDNA library from *Arabidopsis thaliana* into BW535 and selecting clones that had increased resistance to MMS. One clone was isolated and characterized. This *Arabidopsis* cDNA also confers MMS resistance to *E. coli* strain BW9109, which is defective only in the product of the *xth* gene (exonuclease III) and it is also able to enhance survival after UV irradiation of the UV-sensitive *uvrA E.coli* strain AB1886, defective in nucleotide excision repair. This lack of specific activities suggests a general role in DNA damage tolerance rather than damage removal. This cDNA is targeted to chloroplasts and mitochondria. Surprisingly, this cDNA was found to be a homologue of yeast thiamine biosynthesis genes, *THI4* from *S. cerevisiae* and *nmt2* from *S. pombe* that are involved in the formation of the thiazole precursor of thiamine. The ability of the *Arabidopsis* cDNA to functionally complement the yeast *thi4* disruption strain suggests that it has a similar role in plant thiamine biosynthesis, so it was named Thi1. In addition, the Thi4-deficient yeast strain also showed a phenotype of mitochondrial DNA instability after treatment with genotoxic drugs, such as MMS and after UV irradiation (Machado et al., 1997). This mitochondrial DNA instability phenotype was complemented by the *Arabidopsis* Thi1 gene (Machado et al., 1997). Thus, it appears that *Arabidopsis* Thi1 and *S. cerevisiae* Thi4 proteins have a dual role in DNA repair and thiamine biosynthesis.

CONCLUDING REMARKS

These examples demonstrate only part of the potential utilization of bacterial systems for the study of DNA repair genes. These techniques are facilitated by the ease and rapidity of bacterial manipulations and by our knowledge of *E. coli* DNA repair genes. These features have allowed bacteria to be used in cloning and complementation studies of both native and mutated gene products.

ACKNOWLEDGMENTS

Research supported by Conselho Nacional de Pesquisas (Brazil), Comissão de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/COFECUB), the World Health Organization, Programa de Núcleos de Excelência and Pró-Reitoria de Pesquisa, Universidade Federal de Minas Gerais.

REFERENCES

Aburatani, H., Hippo, Y., Ishida, T., Takashima, R., Matsuba, C., Kodama, T., Takao, M., Yasui, A., Yamamoto, K. and **Asano, M.** (1997). Cloning and characterization of mammalian 8-hydroxyguanine-specific DNA glycosylase/apurinic, apyrimidinic lyase, a functional mutM homologue. *Cancer Res. 57*: 2151-2156.

Arai, K., Morishita, K., Shinmura, K., Kohno, T., Kim, S.R., Nohmi, T., Taniwaki, M., Ohwada, S. and **Yokota, J.** (1997). Cloning of a human homolog of the yeast OGG1 gene that is involved in the repair of oxidative DNA damage. *Oncogene 14*: 2857-2861.

Aravind, L., Walker, D.R. and **Koonin, E.V.** (1999). Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res. 27*: 1223-1242.

Augusto-Pinto, L., Bartholomeu, D.C., Teixeira, S.M., Pena, S.D. and **Machado, C.R.** (2001). Molecular cloning and characterization of the DNA mismatch repair gene class 2 from the *Trypanosoma cruzi*. *Gene 272*: 323-333.

Berdal, K.G., Bjoras, M., Bjelland, S. and **Seeberg, E.** (1990). Cloning and expression in *Escherichia coli* of a gene for an alkylbase DNA glycosylase from *Saccharomyces cerevisiae*; a homologue to the bacterial alkA gene. *EMBO J. 9*: 4563-4568.

Boiteux, S., Gajewski, E., Laval, J. and **Dizdaroglu, M.** (1992). 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*: 106-110.

Carell, T., Burgdorf, L.T., Kundu, L.M. and **Cichon, M.** (2001). The mechanism of action of DNA photolyases. *Curr. Opin. Chem. Biol. 5*: 491-498.

Chen, J., Derfler, B., Maskati, A. and **Samson, L.** (1989). Cloning a eukaryotic DNA glycosylase repair gene by the suppression of a DNA repair defect in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA 86*: 7961-7965.

Chen, J., Derfler, B. and **Samson, L.** (1990). *Saccharomyces cerevisiae* 3-methyladenine DNA glycosylase has homology to the AlkA glycosylase of *E. coli* and is induced in response to DNA alkylation damage. *EMBO J. 9*: 4569-4575.

Christians, F.C. and **Loeb, L.A.** (1996). Novel human DNA alkyltransferases obtained by random substitution and genetic selection in bacteria. *Proc. Natl. Acad. Sci. USA 93*: 6124-6128.

Cleaver, J.E., Karplus, K., Kashani-Sabet, M. and **Limoli, C.L.** (2001). Nucleotide excision repair "a legacy of creativity". *Mutat. Res. 485*: 23-36.

Cunningham, R.P., Saporito, S.M., Spitzer, S.G. and **Weiss, B.** (1986). Endonuclease IV (nfo) mutant of *Escherichia coli*. *J. Bacteriol. 168*: 1120-1127.

Doetsch, P.W. and **Cunningham, R.P.** (1990). The enzymology of apurinic/apyrimidinic endonucleases. *Mut. Res. 336*: 173-201.

Dolan, M.E. and **Pegg, A.E.** (1997). O6-benzylguanine and its role in chemotherapy. *Clin. Cancer Res. 3*: 837-847.

Dolan, M.E., Moschel, R.C. and **Pegg, A.E.** (1990). Depletion of mammalian O6-alkylguanine-DNA alkyltransferase activity by O6-benzylguanine provides a means to evaluate the role of this protein in protection against carcinogenic and therapeutic alkylating agents. *Proc. Natl. Acad. Sci. USA 87*: 5368-5372.

Drummond, J.T., Li, G.M., Longley, M.J. and **Modrich, P.** (1995). Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science 268*: 1909-1912.

Edara, S., Kanugula, S., Goodtzova, K. and **Pegg, A.E.** (1996). Resistance of the human O6-alkylguanine-DNA alkyltransferase containing arginine at codon 160 to inactivation by O6-benzylguanine. *Cancer Res. 56*: 5571-5575.

Friedberg, E.C., Walker, G.C. and **Siede, W.** (1995). *DNA Repair and Mutagenesis*. American Society for Microbiology, Washington DC, USA, pp. 91-225.

Friedman, H.S., Kokkinakis, D.M., Pluda, J., Friedman, A.H., Cokgor, I., Haglund, M.M., Ashley, D.M., Rich, J., Dolan, M.E., Pegg, A.E., Moschel, R.C., McLendon, R.E., Kerby, T., Herndon, J.E., Bigner, D.D. and **Schold Jr., S.C.** (1998). Phase I trial of O6-benzylguanine for patients undergoing surgery for malignant glioma. *J. Clin. Oncol. 16*: 3570-3575.

Grossman, L. and **Yeung, A.T.** (1990). The UvrABC endonuclease of *Escherichia coli*. *Photochem. Photobiol. 51*: 749-755.

Husain, I., Griffith, J. and **Sancar, A.** (1988). Thymine dimers bend DNA. *Proc. Natl. Acad. Sci. USA 85*: 2558-2562.

Imai, Y., Oda, H., Nakatsuru, Y. and **Ishikawa, T.** (1995). A polymorphism at codon 160 of human O6-methylguanine-DNA methyltransferase gene in young patients with adult type cancers and functional assay. *Carcinogenesis 16*: 2441-2445.

Kohno, T., Shinmura, K., Tosaka, M., Tani, M., Kim, S.R., Sugimura, H., Nohmi, T., Kasai, H. and

Yokota, J. (1998). Genetic polymorphisms and alternative splicing of the hOGG1 gene, that is involved in the repair of 8-hydroxyguanine in damaged DNA. *Oncogene 16*: 3219-3225.

Kosa, J.L. and **Sweasy, J.B.** (1999). 3'-Azido-3'-deoxythymidine-resistant mutants of DNA polymerase beta identified by *in vivo* selection. *J. Biol. Chem. 274*: 3851-3858.

Lahue, R.S., Au, K.G. and **Modrich, P.** (1989). DNA mismatch correction in a defined system. *Science 245*: 160-164.

Machado, C.R., de Oliveira, R.L., Boiteux, S., Praekelt, U.M., Meacock, P.A. and **Menck, C.F.** (1996). Thi1, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements bacterial defects in DNA repair. *Plant Mol. Biol. 31*: 585-593.

Machado, C.R., Praekelt, U.M., de Oliveira, R.C., Barbosa, A.C., Byrne, K.L., Meacock, P.A. and **Menck, C.F.** (1997). Dual role for the yeast THI4 gene in thiamine biosynthesis and DNA damage tolerance. *J. Mol. Biol. 273*: 114-121.

Marra, G. and **Schär, P.** (1999). Recognition of DNA alterations by the mismatch repair system. *Biochem. J. 338*: 1-13.

Marsischky, G.T., Filosi, N., Kane, M.F. and **Kolodner, R.** (1996). Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes Dev. 10*: 407-420.

Memisoglu, A. and **Samson, L.** (1996). Cloning and characterization of a cDNA encoding a 3-methyladenine DNA glycosylase from the fission yeast *Schizosaccharomyces pombe*. *Gene 177*: 229-235.

Modrich, P. and **Lahue, R.** (1986). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem. 65*: 101-133.

Nakagawa, T., Abhijit, D. and **Kolodner, R.D.** (1999). Multiple functions of MutS and MutL related heterocomplexes. *Proc. Natl. Acad. Sci. USA 96*: 14186-14188.

O'Connor, T.R. (1993). Purification and characterization of human 3-methyladenine-DNA glycosylase. *Nucleic Acids Res. 21*: 5561-5569.

O'Connor, T.R. and **Laval, F.** (1990). Isolation and structure of a cDNA expressing a mammalian 3-methyladenine-DNA glycosylase. *EMBO J. 9*: 3337-3342.

Pang, Q., Hays, J.B. and **Rajagopal, I.** (1992). A plant cDNA that partially complements *Escherichia coli recA* mutations predicts a polypeptide not strongly homologous to RecA proteins. *Proc. Natl. Acad. Sci. USA 89*: 8073-8077.

Pang, Q., Hays, J.B., Rajagopal, I. and **Schaefer, T.S.** (1993). Selection of Arabidopsis cDNAs that partially correct phenotypes of *Escherichia coli* DNA-damage-sensitive mutants and analysis of two plant cDNAs that appear to express UV-specific dark repair activities. *Plant Mol. Biol. 22*: 411-426.

Pegg, A.E. (2000). Repair of O(6)-alkylguanine by alkyltransferases. *Mutat. Res. 462*: 83-100.

Perez, J., Gallego, C., Bernier-Villamor, V., Camacho, A., Gonzalez-Pacanowska, D. and **Ruiz-Perez, L.M.** (1999). Apurinic/apyrimidinic endonuclease genes from the Trypanosomatidae *Leishmania major* and *Trypanosoma cruzi* confer resistance to oxidizing agents in DNA repair-deficient *Escherichia coli*. *Nucleic Acids Res. 27*: 771-777.

Popoff, S.C., Spira, A.I., Johnson, A.W. and **Demple, B.** (1990). Yeast structural gene (APN1) for the major apurinic endonuclease: homology to *Escherichia coli* endonuclease IV. *Proc. Natl. Acad. Sci. USA 87*: 4193-4197.

Reenan, R.A. and **Kolodner, R.D.** (1992). Characterization of insertion mutations in the *Saccharomyces cerevisae* MSH1 and MSH2 genes: evidence for separate mitochondrial nuclear functions. *Genetics 132*: 975-985.

Rosenquist, T.A., Zharkov, D.O. and **Grollman, A.P.** (1997). Cloning and characterization of a mammalian 8-oxoguanine DNA glycosylase. *Proc. Natl. Acad. Sci. USA 94*: 7429-7434.

Samson, L., Derfler, B., Boosalis, M. and **Call, K.** (1991). Cloning and characterization of a 3-methyladenine DNA glycosylase cDNA from human cells whose gene maps to chromosome 16. *Proc. Natl. Acad. Sci. USA 88*: 9127-9131.

Sancar, G.B. (1985). Expression of a *Saccharomyces cerevisiae* photolyase gene in *Escherichia coli*. *J. Bacteriol. 161*: 769-771.

Santerre, A. and **Britt, A.B.** (1994). Cloning of a 3-methyladenine-DNA glycosylase from *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA 91*: 2240-2244.

Sugimura, H., Kohno, T., Wakai, K., Nagura, K., Genka, K., Igarashi, H., Morris, B.J., Baba, S., Ohno, Y., Gao, C., Li, Z., Wang, J., Takezaki, T., Tajima, K., Varga, T., Sawaguchi, T., Lum, J.K., Martinson, J.J., Tsugane, S., Iwamasa, T., Shinmura, K. and **Yokota, J.** (1999). hOGG1 Ser326Cys polymorphism and lung cancer susceptibility. *Cancer Epidemiol. Biomarkers Prev. 8*: 669-674.

Sweasy, J.B. and **Loeb, L.A.** (1992). Mammalian DNA polymerase beta can substitute for DNA polymerase I during DNA replication in *Escherichia coli*. *J. Biol. Chem. 267*: 1407-1410.

Sweasy, J.B. and **Loeb, L.A.** (1993). Detection and characterization of mammalian DNA polymerase beta

mutants by functional complementation in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA 90*: 4626-4630.

Xu-Welliver, M., Leitao, J., Kanugula, S., Meehan, W.J. and **Pegg, A.E.** (1999). Role of codon 160 in the sensitivity of human O6-alkylguanine-DNA alkyltransferase to O6-benzylguanine. *Biochem. Pharmacol. 58*: 1279-1285.

van der Kemp, P.A., Thomas, D., Barbey, R., de Oliveira, R. and **Boiteux, S.** (1996). Cloning and expression in *Escherichia coli* of the OGG1 gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. *Proc. Natl. Acad. Sci. USA 93*: 5197-5202.

Washington, S.L., Yoon, M.S., Chagovetz, A.M., Li, S.X., Clairmont, C.A., Preston, B.D., Eckert, K.A. and **Sweasy, J.B.** (1997). A genetic system to identify DNA polymerase beta mutator mutants. *Proc. Natl. Acad. Sci. USA 94*: 1321-1326.