ERCC6, a Member of a Subfamily of Putative Helicases, Is Involved in Cockayne's Syndrome and Preferential Repair of Active Genes

Christine Troelstra, Alain van Gool, Jan de Wit,
Wim Vermeulen, Dirk Bootsma,
and Jan H. J. Hoeijmakers
Medical Genetics Center
Department of Cell Biology and Genetics
Erasmus University
3000DR Rotterdam
The Netherlands

Summary

Cells from patients with the UV-sensitive nucleotide excision repair disorder Cockayne's syndrome (CS) have a specific defect in preferential repair of lesions from the transcribed strand of active genes. This system permits quick resumption of transcription after UV exposure. Here we report the characterization of ERCC6, a gene involved in preferential repair in eukaryotes. ERCC6 corrects the repair defect of CS complementation group B (CS-B). It encodes a protein of 1493 amino acids, containing seven consecutive domains conserved between DNA and RNA helicases. The entire helicase region bears striking homology to segments in recently discovered proteins involved in transcription regulation, chromosome stability, and DNA repair. Mutation analysis of a CS-B patient indicates that the gene is not essential for cell viability and is specific for preferential repair of transcribed sequences.

Introduction

The genetic information of an organism is continually subjected to the damaging effects of environmental and internal genotoxic agents, chemical instability, and the inherent imperfection of DNA metabolizing processes. If unrepaired, such DNA injury will lead to mutations, carcinogenesis, general deterioration of cell functioning, and cell death. To counteract these deleterious consequences, an intricate network of DNA repair systems has evolved (for review see Friedberg, 1985). Among the most important and best-studied repair mechanisms, both in prokaryotes and eukaryotes, is the nucleotide excision repair (NER) pathway. This system eliminates a remarkably broad spectrum of structurally unrelated lesions, such as ultraviolet (UV)-induced cyclobutane pyrimidine dimers (CPD) and (6–4) photoproducts, bulky chemical adducts, and DNA cross-links. Removal of these lesions proceeds via a multistep reaction, which is resolved to considerable detail in the bacterium Escherichia coli (van Houten, 1990): a UvrA$_2$B helicase complex scans the DNA for structural distortions by locally unwinding the two strands (Oh and Grossman, 1987) and delivers the UvrB subunit to the lesion (Orren and Sancar, 1989). UvrC interacts with the damage-specific UvrB–DNA complex and catalyzes incision of the damaged strand on both sides of the injury.

Release of the lesion-containing 12–13 nt oligomer, as well as the bound UvrC, is performed by UvrD (helicase II), after which DNA polymerase I turns over UvrB and fills the resulting single-stranded gap (Orren et al., 1992). Finally, the newly synthesized repair patch is ligated to the parental DNA. Knowledge of the molecular mechanism of NER in eukaryotes is rapidly accumulating. Recently, a dual incision has been demonstrated (Huang et al., 1992), and several proteins required for pre- or postincision steps were identified (Shivji et al., 1992).

The prototype repair disorder xeroderma pigmentosum (XP) illustrates the phenotypic consequences of defective NER. Patients suffering from this rare recessive inherited disease characteristically exhibit extreme sun (UV) sensitivity, pigmentation abnormalities, predisposition to skin cancer, and frequently progressive neurological degeneration (Cleaver and Kraemer, 1989). Extensive genetic heterogeneity was disclosed by cell hybridization with the identification of at least seven excision-deficient complementation groups (XP-A to XP-G) (de Weerd-Kastelein et al., 1972; Vermeulen et al., 1991). The biochemical basis for the XP NER defect probably resides in undefined early steps of the pathway preceding incision and repair synthesis (Shivji et al., 1992).

An immediate and urgent effect of the presence of lesions in DNA is interruption of the vital process of transcription (Mayne and Lehmann, 1982). To release this otherwise lethal block, a sophisticated NER subpathway appears to operate that primarily focuses on the rapid and preferential removal of certain lesions (from the transcribed strand of) active genes. This NER process was first discovered in mammalian cells (Bohr et al., 1985) and found to explain the apparent paradox that rodent cells perform only a fraction of CPD removal compared with human cell lines, yet they exhibit the same degree of UV resistance (Zelle et al., 1980; van Zeeland et al., 1981). It appeared that rodent lines mainly rely on the strand-selective repair of active genes and are largely deficient in the NER subpathway that is responsible for the slower and incomplete removal of CPD from the genome overall (including the nontranscribed strand) (Bohr et al., 1985; Mellon et al., 1987). Human cells are proficient for both processes (Mellon et al., 1986, 1987; Venema et al., 1990b). This demonstrates the importance for cell survival of preferential repair of the small but vital active compartment of the genome. Recent analysis of mutations in various genes has provided evidence that strand-selective repair also contributes significantly to prevention of mutagenesis originating from this strand in cultured cells (Vrieling et al., 1991). The universality of preferential repair is demonstrated by the recent discovery that in yeast and E. coli also the transcribed strand is more rapidly cleared from CPD lesions than the nontranscribed strand (Mellon and Hanawalt, 1989; Smerdon and Thoma, 1990). These observations suggest that a direct coupling or a feedback mechanism exists between transcription and NER.

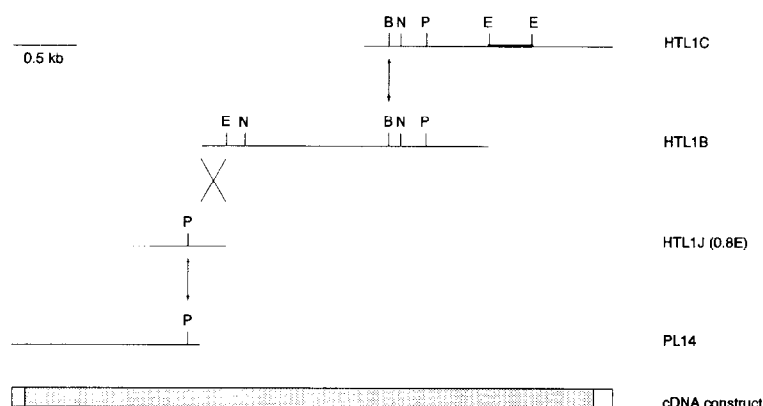The notion that this important process constitutes a dis-

Figure 1. Construction of a Functional *ERCC6* cDNA

The inserts of HTL1B and HTL1C (human testis cDNA library) were isolated from the phage arms via PCR and joined by ligation (vertical arrows) at the unique BglII site (B). The fragment HTL1J(0.8E) is added through homologous recombination (X), and finally PL14 (human placenta cDNA library) is added to the resulting molecule by ligation at the PflMI (P) site. The parts obtained via PCR or homologous recombination were checked by sequence analysis. The thick line in HTL1C corresponds to the probe used for Southern blot analysis in Figure 6B. Other symbols: E, EcoRI; N, NsiI; dotted line, sequence not representing *ERCC6* (artificially ligated to *ERCC6*); stippled region, the ORF. See Experimental Procedures for details.

tinct subpathway of NER came from the finding that patients suffering from Cockayne's syndrome (CS) carry specific defects in this system. CS is a rare genetic disorder that manifests UV hypersensitivity (but no other cutaneous abnormalities), retarded growth, and severe progressive neurological degeneration. Remarkably, in striking contrast with XP, no significant increase in skin cancer is noted (Lehmann, 1987; Nance and Berry, 1992). Cell fusion experiments have revealed the presence of at least two complementation groups, group A (CS-A) and group B (CS-B), within the classical form of the disease (Tanaka et al., 1981; Lehmann, 1982). A characteristic hallmark of the disease is the observation that CS fibroblasts perform rates of repair synthesis in the normal range yet, in contrast with repair-competent cells, are unable to recover their RNA synthesis after UV irradiation (Mayne and Lehmann, 1982). The early suggestion put forward by Mayne and Lehmann (1982) that the CS defect resides in the repair of transcribed genes was only borne out by the recent finding of Venema et al. (1990a) that CS fibroblasts have lost the preferential repair of active genes but are proficient in overall genome repair. For CS-A, it was demonstrated that the absence of preferential repair correlates with a complete deficiency of selective repair of the transcribed strand (Venema, 1991).

Here we report a key step toward defining this system at the molecular level by the analysis of a gene implicated in the preferential repair pathway. This human gene, termed *ERCC6*, was cloned by virtue of its ability to correct the NER defect of a UV-sensitive Chinese hamster ovary (CHO) mutant (Troelstra et al., 1990). The mutant, UV61, is a representative of complementation group 6 of the laboratory-generated, excision-deficient rodent cell lines (Busch et al., 1989). We show that the *ERCC6* gene at the same time corrects the repair defect of CS-B, the most common form of the disease, and that the gene encodes a protein with a presumed DNA unwinding function. The helicase segment bears striking resemblance to similar domains in transcription regulators and proteins involved in various repair processes and chromosome stability recently identified in yeast and Drosophila.

Results

Isolation of the *ERCC6* cDNA

The human excision repair gene *ERCC6* was recently isolated, after large-scale transfection experiments of human (HeLa) genomic DNA to UV-sensitive CHO mutant UV61 (rodent complementation group 6 [CG-6]) and rescue of the correcting human sequences from a secondary transformant (Troelstra et al., 1990). The gene appeared to have a respectable size (for transfection cloning) of close to 100 kb. Unique conserved genomic *ERCC6* sequences were identified and used as probes to hybridize to a human testis cDNA library. This resulted in the isolation of a number of overlapping cDNA clones, together having an insert size of 3 kb with a polyadenylation signal 20 nt from the 3' end. Northern blot analysis with part of this 3 kb cDNA as a probe revealed the presence of two messenger RNAs (mRNAs) of about 5 kb and 7–7.5 kb in length. Sequence analysis and restriction mapping indicated that the difference in size is due to alternative polyadenylation (unpublished data). Both mRNAs are very lowly expressed, an observation upheld by the recovery of very low numbers of *ERCC6* cDNA clones from various cDNA libraries. After screening of several cDNA libraries, an almost full-length molecule of 4.7 kb could be constructed (see Figure 1 for details).

Correction of UV61 by the *ERCC6* cDNA

The 4.7 kb cDNA molecule was inserted in mammalian expression vector pSLM (see Experimental Procedures) under the control of the SV40 late promoter. Cotransfection of this plasmid with pSV2neo to UV61 cells yielded UV-resistant clones with very high efficiency (transfection frequency of UV and G418 resistance was similar to that of G418 resistance alone). UV survival of these transformants appeared within the wild-type range (Figure 2A).

To verify whether other DNA repair endpoints were corrected by *ERCC6*, we examined the recovery of RNA synthesis after UV exposure. In normal cells, transcription drops rapidly after UV irradiation (owing to the appearance of blocking lesions in the DNA template) but resumes
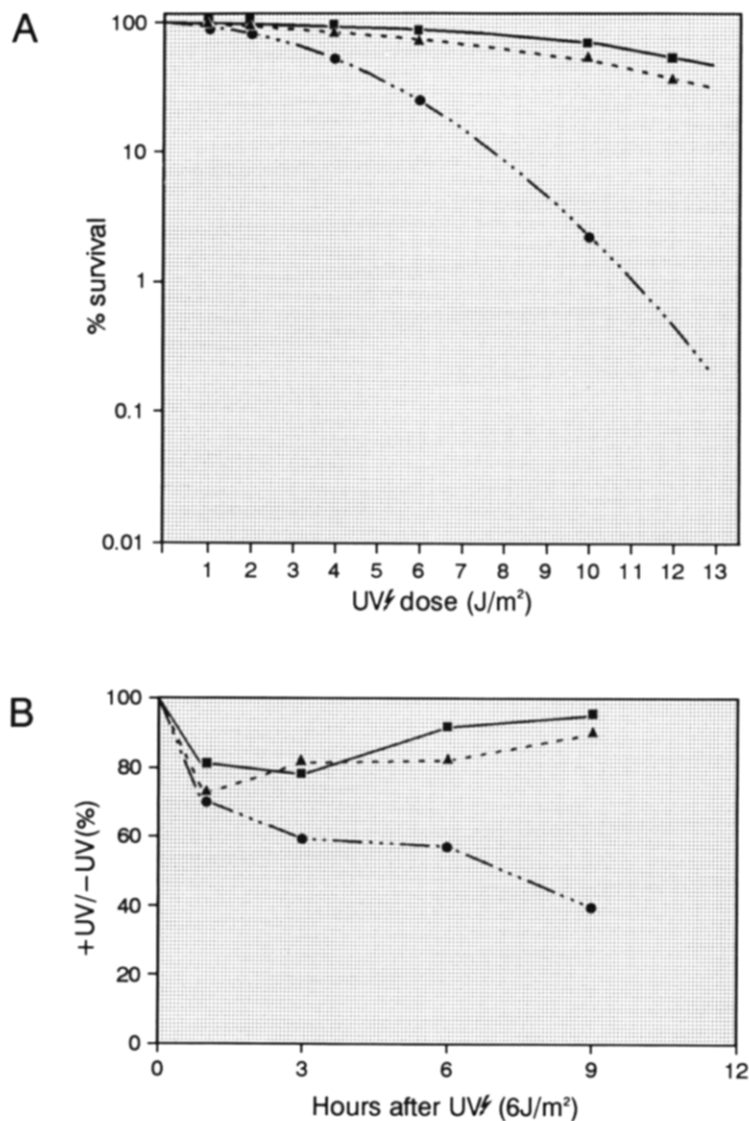
Figure 2. Correction of the UV61 Repair Defect by ERCC6 cDNA

(A) UV survival curves of wild-type CHO cell line AA8 (closed boxes), mutant UV61 (closed circles), and ERCC6 cDNA transformant 61E6a (closed triangles).

(B) Recovery of RNA synthesis in wild-type cell line AA8, mutant UV61, and ERCC6 cDNA transformant 61E6a. Cells unirradiated or UV irradiated with 6 J/m^2 were pulse labeled with [^{3}H]uridine at different times after irradiation; the incorporated radioactivity was determined by liquid scintillation counting (see Experimental Procedures for details on the procedure). For symbols see (A). The standard error of the mean (three duplicate plates) falls within the size of the symbols.

within several hours as the result of preferential (strand-specific) repair. In excision-deficient cells such as XP and CS, recovery of RNA synthesis does not occur (Mayne and Lehmann, 1982). Since this repair parameter has never been investigated for rodent mutants, we decided to assess the behavior of RNA synthesis in CHO wild-type AA8, mutant UV61, and cDNA transformant 61E6a upon UV irradiation. As shown in Figure 2B, RNA synthesis (expressed as the ratio of [^{3}H]uridine incorporation in UV-irradiated cells to that in unirradiated cells) initially decreases within 1 hr to 70%–80% but quickly resumes in AA8, whereas it remains low in UV61. Recovery of RNA synthesis, as in wild type, is observed in ERCC6 transformant 61E6a, indicating that the human gene also corrects this repair endpoint.

To investigate the specificity of correction by ERCC6, the gene was transfected to mutants from rodent CG-4, CG-5, and CG-7 and to two other members of group 6 (CG-1, CG-2, and CG-3 were not tested since the correct-

ing genes for these mutants have been cloned [van Duin et al., 1986; Weber et al., 1990; Weeda et al., 1990a, 1990b]). No UV-resistant transformants were obtained in the heterologous CG, whereas numerous UV-surviving colonies were observed in the other representatives of CHO group 6, demonstrating the specificity of ERCC6 correction and validating the complementation studies.

Nucleotide and Amino Acid Sequence of ERCC6 cDNA

The nucleotide sequence of the ERCC6 cDNA is shown in Figure 3A. The first ATG codon (nucleotide 80), preceded by two in-frame stop codons, starts an open reading frame (ORF) of 4479 bp. The purine often found at the important −3 position before a start codon (Kozak, 1991) is present, as is the G at +4. The other nucleotides within the start site, however, deviate from the optimal initiation start site GCC(A/G)CCATGG (Kozak, 1991). All other ORFs present encode much smaller polypeptides. Therefore, we con-
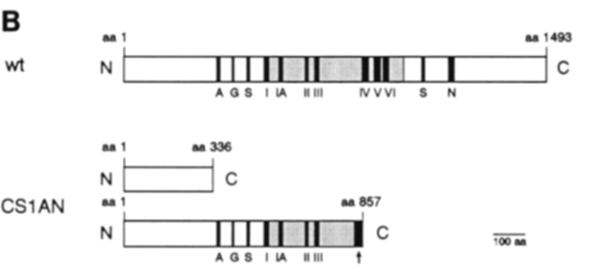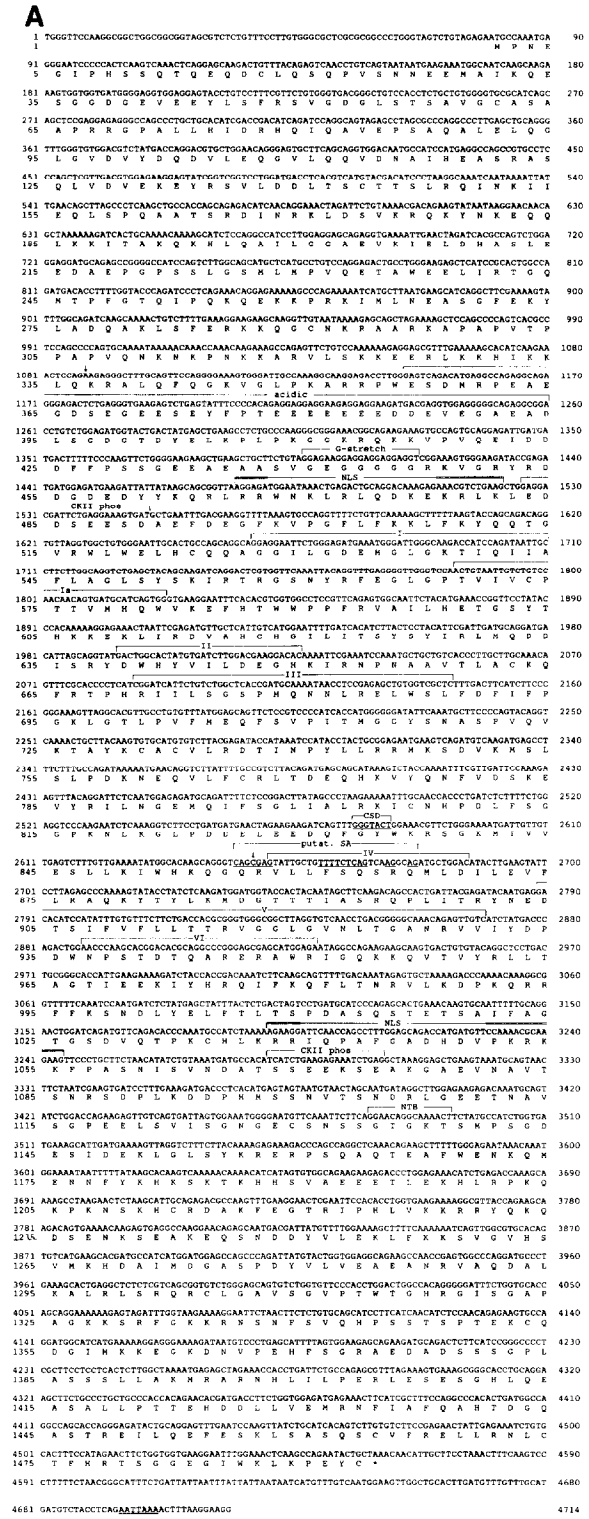
A

B

Figure 3. Nucleotide and Predicted Amino Acid Sequence of the *ERCC6* cDNA

(A) *ERCC6* nucleotide and predicted amino acid sequence. The sequence of 4714 nt has been obtained from the cDNA fragments, as depicted in Figure 1, together with a cDNA clone from an adenovirus-transformed retina cell library (the first 16 nt). Putative functional units are overlined: CKII phos, casein kinase II phosphorylation site; NLS, nuclear location signal (basic regions are double overlined); NTB, nucleotide binding fold; numbers I, IA, and II–VI refer to the corresponding helicase motifs. The AATTAAA polyadenylation signal is underlined. Arrows indicate the mutations in patient CS1AN; CSD means the cryptic splice donor, while putat. SA indicates the putative splice acceptor; the putative branch point, the pyrimidine stretch, and the AG dinucleotides are underlined; the third AG represents the normal exon 13 border.

(B) Schematical representation of predicted functional domains in the ERCC6 protein. The truncated ERCC6 proteins encoded by the CS1AN alleles are indicated. A, acidic amino acid stretch; G, glycine-rich region; S, serine phosphorylation site preceded by a nuclear location signal; N, nucleotide binding fold; arrow, nonsense sequence due to a frameshift in allele b of patient CS1AN.

computer search did not identify any protein with overall homology to ERCC6, a close inspection of the sequence did reveal several regions of interest and similarity to other polypeptides.

First, there was an acidic part from residue 356–394 harboring close to 60% glutamic or aspartic acids including a stretch of 10 acidic residues. Acidic regions have been found in a number of nuclear proteins that associate with chromatin or histones (Lapeyre et al., 1987; Ptashne, 1988; Sung et al., 1988; Wen et al., 1989). This part is followed by a stretch of glycin residues (amino acids 442–446) and a very hydrophilic segment (from residue 447–495) containing >60% charged amino acids.

Second, two regions with the hallmarks of a bipartite nuclear location signal (amino acids 466–481 and 1038–1055; Robbins et al., 1991) were found in close vicinity to potential casein kinase II phosphorylation sites (Kemp and Pearson, 1990, amino acids 483–490 and 1068–1074). Recently, it was noticed that putative casein kinase II phosphorylation sites are located nearby many nuclear location signal sequences (Rihs et al., 1991).

Third, the region from amino acid 527 to amino acid 950 contains seven sequences (starting with a nucleotide binding fold) matching seven consecutive motifs recognized by several investigators to be conserved between two superfamilies of RNA and DNA helicases (Gorbalenya et al., 1989). This region (amino acids 527–1007) exhibits, in addition, a very high overall homology with parts of eight other proteins containing the same seven motifs: RAD16 (Bang et al., 1992; Schild et al., 1992), RAD54, SNF2

clude that the ORF of 4479 bp, depicted in Figure 3A, defines the ERCC6 protein.

The deduced amino acid sequence specifies a protein of 1493 amino acids (see Figures 3A and 3B), with a predicted molecular size of 168 kd. Although a systematic

(SWI2), STH1 (Laurent et al., 1992), MOT1 (Davis et al., 1992), FUN30 (Clark et al., 1992), brm (Tamkun et al., 1992), and lodestar (Girdham and Glover, 1991) (Figure 4B). The similarity is confined to this particular segment (Figure 4A) and is far higher than the motif-restricted homology normally detected between different helicases (see Figure 4B for the consensus sequence of the seven helicase motifs).

Fourth, a second putative nucleotide binding domain (Saraste et al., 1990) is found in the C terminus (amino acids 1134–1138).

Involvement of ERCC6 in CS

The ERCC6 gene, correcting a CHO mutant, could possibly also be implicated in one of the forms of XP or CS. The first clue concerning this possibility was presented by a patient with symptoms of late onset CS, being heterozygous for an interstitial 10q211 deletion in all cells (Fryns et al., 1991), i.e., a chromosomal region to which we have assigned ERCC6 (Troelstra et al., 1992). Southern blot analysis of DNA from the patient suggested that ERCC6 resides within the deleted segment (data not shown). Since no cells of this patient were available and the complementation group to which he belonged was unknown, we decided to transfect the ERCC6 cDNA expression plasmid together with the dominant selectable marker pSV2neo into CS3BE.S3.G1 and CS1AN.S3.G2, immortalized cell lines representing CS CG-A and CG-B, respectively (Mayne et al., 1986). G418-resistant transformants of each cell line were grown in mass culture. In view of the poor transfectability of many SV40-immortalized human fibroblasts, Southern blot analysis was performed to verify whether a significant fraction of the G418-resistant transformants had indeed incorporated the cotransfected ERCC6 cDNA. For CS3BE.S3.G1, this fraction was very low, but detectable, on Southern blot (average incorporation of <0.1 intact ERCC6 cDNA copy per cell). CS1AN.S3.G2 cells had an average integration of one to two copies per transformant (data not shown). To test whether ERCC6 restores the DNA repair capacity of one of the CS complementation groups, the G418-resistant mass populations (consisting of ±100 and 800–1000 independent clones for CG-A and CG-B, respectively) were subjected to a UV dose lethal to the parental CS cells (8 or 6 J/m², which gives a 0.25% survival for CS-A and CS-B, respectively, and 15% or 25% survival for wild-type cells; the treatment was repeated after 48 hr). For CS3BE.S3.G1, none of the G418-resistant cells survived the UV selection, and neither did the transfected mass population show enhanced survival (at different doses). This suggests that ERCC6 is unable to correct the repair defect of CS-A (although it should be kept in mind that less than 10% of the transformants had incorporated an intact copy of the ERCC6 cDNA). ERCC6 did confer UV resistance to CS1AN.S3.G2 cells; the transfected mass population efficiently survived the UV selection. The UV survival of two independent transformants (E61ANa and E61ANd) is depicted in Figure 5A and shows that both have regained a wild-type UV resistance. To establish the specificity of correction further, ERCC6 cDNA was introduced (either by DNA transfection

or by microinjection) into representative cell lines of XP-C and XP-E to XP-G; the correcting genes for XP-A (Tanaka et al., 1990), XP-B (Weeda et al., 1990b), and XP-D (Flejter et al., 1992) have already been isolated and are different from ERCC6. Because XP-E cells are only slightly UV sensitive, it was not possible to derive definite conclusions from the transfection experiments to this complementation group. The ERCC6 gene was unable to correct the repair defect of the other XP groups tested. These results confirm the complementation data and demonstrate that no overlap exists between CS-B and any of the known XP groups (with the possible exception of XP-E), in contrast with CS-C, which is identical to XP-B.

RNA Synthesis Recovery in ERCC6 cDNA Transformants of CS1AN.S3.G2

As mentioned above, CS cells are, in contrast with repair-proficient human cells, unable to recover their RNA synthesis after UV exposure (Mayne and Lehmann, 1982). To see whether ERCC6 also complements this CS-typical repair parameter, RNA synthesis was measured after UV irradiation in SV40-transformed normal fibroblasts, CS1AN.S3.G2, and cDNA transformants E61ANa and E61ANd. As shown in Figure 5B, transcription initially drops to ≈60% within 1 hr after irradiation. The RNA synthesis quickly recovers in normal cells, whereas it remains repressed in CS1AN.S3.G2 cells. The cDNA transformant E61ANd resumes transcription with almost wild-type kinetics; E61ANa, on the other hand, recovers more slowly, being retarded at early times after UV but reaching wild-type levels at 16 hr. In Figure 5C the level of transcription 16 hr after UV irradiation is tested at different UV fluences. In both cDNA transformants, RNA synthesis is similar to that in repair-proficient cells for all UV doses, whereas transcription in CS1AN.S3.G2 is consistently lower. The slow (but after 16 hr complete) recovery of E61ANa, in contrast with the fast recovery of E61ANd (Figure 5B), might be attributed to the slow growth rate of E61ANa. Alternatively, expression levels of the cDNA, influenced by, e.g., the integration site, could have an effect on the kinetics.

To demonstrate the universality of CS-B correction, primary fibroblasts of another CS-B patient, CS1BE, were microinjected with ERCC6. Recovery of UV-induced transcription inhibition was, as in the CS1AN cDNA transformants, restored to levels within wild-type range (data not shown). Preliminary data suggesting that microinjection of ERCC6 into primary fibroblasts of CS-A did not restore RNA synthesis recovery after UV irradiation confirm the complementation data.

The ERCC6 Mutation in CS1AN Cells

The previous experiments showed that ERCC6 specifically corrects the UV survival and RNA synthesis recovery after UV exposure of CS-B. To obtain direct proof that ERCC6 is responsible for the CS-B repair defect, we searched for mutations in the gene. The polymerase chain reaction (PCR) was used to amplify ERCC6 cDNA selectively from RNA of the SV40-transformed CS-B cell line CS1AN.-S3.G2. For both cDNA synthesis and PCR, the ERCC6

A

'helicase' domains

SNF2

Lodestar

RAD16

RAD54

ERCC6

Figure 4. Homology of the ERCC6 Helicase Region to the Similar Region of RAD54, RAD16, lodestar, SNF2, STH1, brm, MOT1, and FUN30

(A) Schematical representation of ERCC6, RAD54, RAD16, lodestar, and SNF2 to indicate the location of the homologous segment in the proteins. Stippled boxes represent homology; helicase motifs are indicated by closed boxes.

(B) Alignment of conserved regions of the ERCC6, RAD54, RAD16, lodestar, SNF2, STH1, brm, MOT1, and FUN30 amino acid sequences. Sequences that correspond to the seven consensus motifs identified among two super-families of DNA and RNA helicases are overlined. The consensus (consens) sequences given below the corresponding domains are from Gorbalenya et al. (1989); hydrophobic (I, L, V, M, F, Y, W) and charged or polar residues (S, T, D, E, N, Q, K, R) are indicated by plus and 0, respectively. Shaded amino acids are similar in at least five of the proteins compared; residues belonging to one of the following groups are considered to be similar: L, V, I, M; F, Y, W; S, T, A, P, G; K, R, H; and E, D, Q, N. The RAD54 sequence is taken from Emery et al. (1991) and the SNF2 sequence from Laurent et al. (1991); other references are given in the text. A small change has been introduced in the published sequence of the *lodestar* gene (Girdham and Glover, 1991), since we noticed that this results in the presence of all seven helicase domains, instead of the reported three. By introducing a loop of 61 amino acids between domains III and IV in the FUN30 amino acid sequence, we were able to perfectly align domains I–III, in addition to domains IV–VI that were identified by Clark et al. (1992).
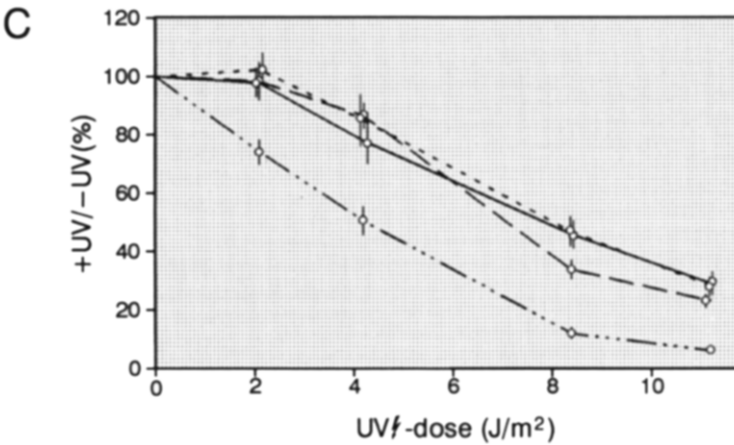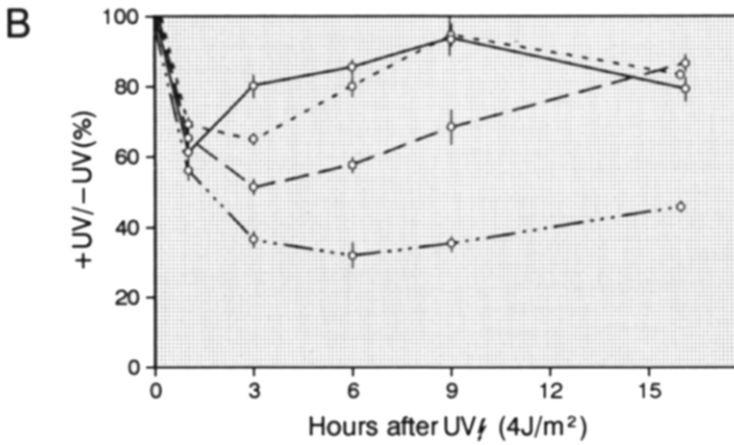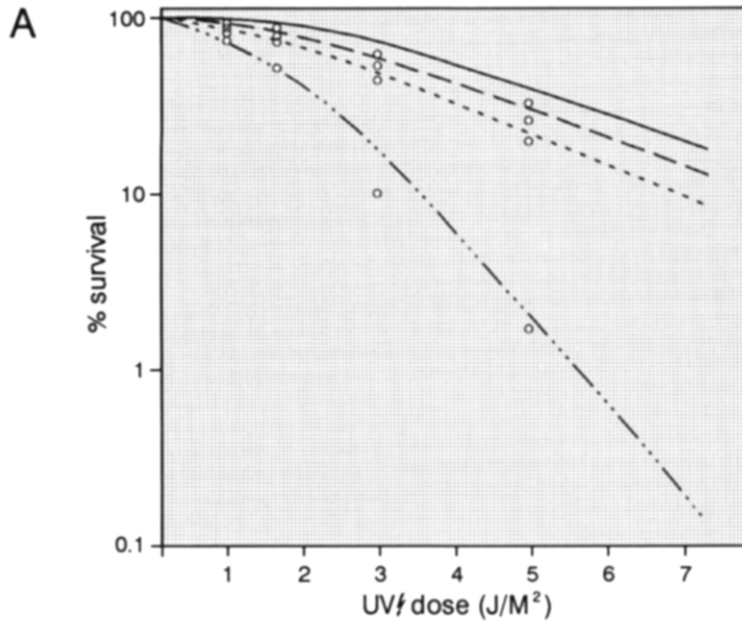
Figure 5. Correction of the CS1AN.S3.G2 Repair Defect by *ERCC6* cDNA

(A) UV survival curves of an SV40-transformed wild-type cell line (solid line), mutant CS1AN.-S3.G2 (dashed and dotted line), and *ERCC6* cDNA transformants E61ANa (dashed line) and E61ANd (dotted line).

(B) Recovery of RNA synthesis at different times after UV irradiation. Cells unirradiated or UV irradiated with 4 J/m² were pulse labeled with [³H]uridine at different times after irradiation; the incorporated radioactivity was determined by liquid scintillation counting. For symbols see (A); the average (open circle) and the standard error of the mean (vertical line) are given (within one experiment).

(C) RNA synthesis 16 hr after UV irradiation with increasing dose. Cells unirradiated or UV irradiated with different dose were pulse labeled with [³H]uridine 16 hr after irradiation; the incorporated radioactivity was determined by liquid scintillation counting. For symbols see (A) and (B).
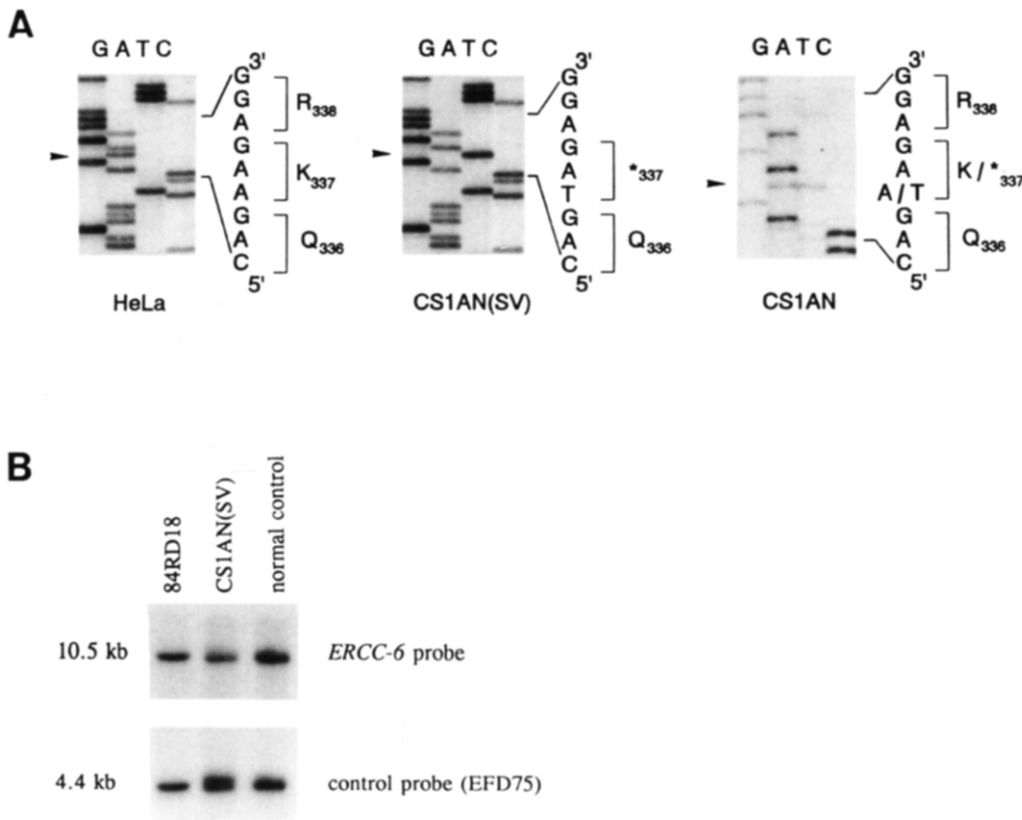
Figure 6. *ERCC6* Mutation in CS1AN.S3.G2 and Allele a of Patient CS1AN

(A) Part of the *ERCC6* nucleotide sequence of total PCR product (from cDNA) from HeLa, CS1AN.S3.G2 (CS1AN(SV)), and the primary fibroblasts of patient CS1AN. The A to T transversion, resulting in a premature stop codon (the mutation in allele a), is indicated by an arrowhead.

(B) Southern blot analysis of CS1AN.S3.G2 and two control cell lines. The same Southern blot was hybridized with two probes. The *ERCC6* probe is a 350 bp EcoRI cDNA fragment (see Figure 1), and EFD75 is a control probe derived from the telomeric end of chromosome 10q (10q26) (Simpson and Cann, 1990). The control probe recognizes two polymorphic bands in DNA of CS1AN.S3.G2, each having an intensity half of that from the single band of both controls. This indicates that at least this telomeric part of chromosome 10 is diploid in CS1AN.S3.G2. For the two controls, the *ERCC6* signal is comparable with that of EFD75. In CS1AN.S3.G2, however, the intensity of the *ERCC6* signal is similar to that of only one of the two polymorphic EFD75 bands, suggesting that only one copy of *ERCC6* is present.

mRNA was divided into six overlapping segments of approximately 1 kb each. Single-stranded DNA for direct sequence analysis of the PCR product was prepared by asymmetric PCR, using amplified cDNA as a template. Sequencing revealed an A to T transversion at nucleotide 1088, resulting in a premature stop at amino acid position 337 (Figure 6A). Only one type of allele was detected, both at the level of RNA and in genomic DNA, implying that CS1AN.S3.G2 is either homozygous or hemizygous for the A^{1088} to T transversion. In the family history of patient CS1AN, consanguinity has not been reported. Southern blot analysis of DNA from CS1AN.S3.G2 confirmed that at least part of one *ERCC6* allele was deleted (Figure 6B).

Since SV40-transformed fibroblasts are chromosomally unstable, sequence analysis of the *ERCC6* gene and cDNA was verified in the primary fibroblasts of patient CS1AN, GM00739. The A to T transversion at nucleotide 1088 was confirmed, but unexpectedly the primary line appeared to possess a second allele with the wild-type A residue at nucleotide position 1088 (Figure 6A). Further

sequence analysis of the total PCR product derived from RNA of GM00739 revealed the appearance of a double sequence ladder when entering from exon 14 backward into exon 13 (Figure 7A). When the normal nucleotide sequence of exon 13 is subtracted, a perfect match is found with a sequence 100 bp upstream in exon 13. The resulting 100 bp deletion is not present in RNA from the SV40-transformed CS1AN.S3.G2 cell line (Figure 7A), thus proving that it is derived from the second allele (allele b) of GM00739. The 3' end of the deletion coincides exactly with the end of exon 13, whereas the 5' end occurs in a sequence (GG/GTACT; see Figure 3A, nucleotides 2576–2582) matching the splice donor consensus (AG/GTPuAG; Pu = A or G) (Senapathy et al., 1990). This suggests that, in allele b, a cryptic donor is utilized instead of the normal exon 13 splice donor. Therefore, a mutation in the splice donor of exon 13 of allele b was expected. At the level of DNA, however, neither the normal nor the cryptic splice donor was mutated (Figure 7B). The only detectable mutation was a C^{2648} to T transition 30 nt upstream of the normal
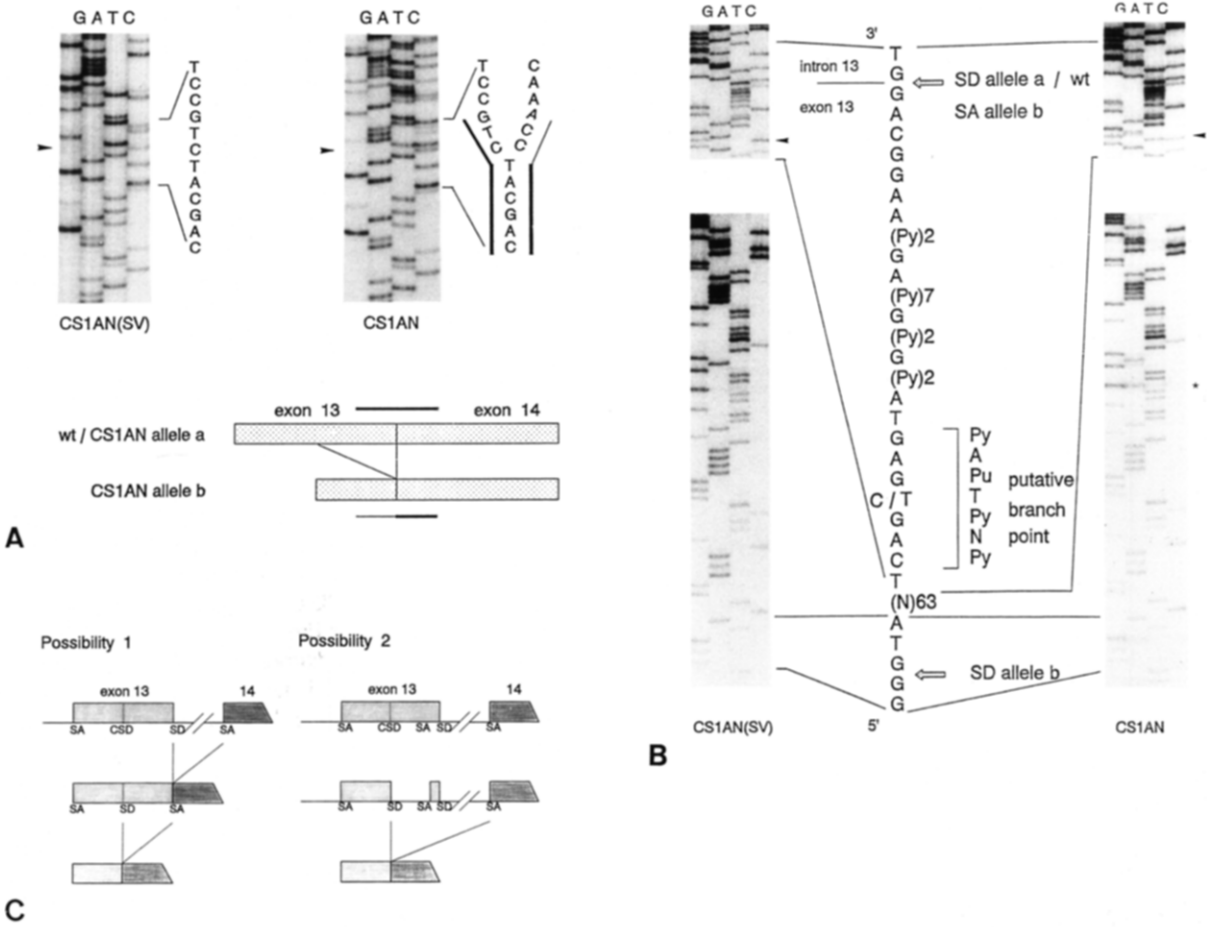
Figure 7. ERCC6 Mutation in Allele b of Patient CS1AN

(A) ERCC6 nucleotide sequence from the exon 14/exon 13 border of both CS1AN.S3.G2 (CS1AN(SV)) and the primary fibroblasts CS1AN. Sequence analysis was performed on total PCR product obtained after amplification of RNA. The start of the double sequence ladder in CS1AN, resulting from a deletion of 100 bp (allele b, schematically depicted underneath the sequences), is indicated by an arrowhead. The sequence obtained with CS1AN.S3.G2 is identical to that using HeLa RNA.

(B) Nucleotide sequence of part of exon 13 at the level of the genome of CS1AN.S3.G2 (CS1AN(SV)) and the primary fibroblasts CS1AN. The C to T transition in CS1AN, responsible for the 100 nt deletion at the mRNA level, is indicated by an arrowhead. The asterisk next to the CS1AN sequence indicates a sequence artifact that was also detected in CS1AN(SV) (and HeLa) on longer exposures.

(C) Schematical representation of the two possible splicing reactions that could result in the 100 nt deletion in the mRNA of CS1AN.

splice donor of exon 13. How could a single point mutation at this position have such a profound influence on the apparent functioning of the still normal splice donor sequence 30 nt downstream? A search for dramatic changes in secondary structure caused by the C to T transition failed to reveal any clues to explain this effect. Close inspection of the sequence revealed that the C to T transition might create a putative branch point sequence (Ruskin et al., 1984; Senapathy et al., 1990) at position 2645–2651, which, together with a coincidently present pyrimidine stretch and one of the three subsequent AG dinucleotides, completes a possible splice acceptor site (see Figures 3A and 7B). This newly created acceptor, together with the cryptic splice donor in exon 13, may generate a new intron. The 100 nt deletion could thus be the result of an extra splice event if the third (C)AG is used. The normal splice donor site of exon 13 should then have a dual role: splice donor in the normal splicing of intron 13 and splice acceptor in a subsequent erroneous splice event, removing the last 100 nt of exon 13 (see Possibility 1 in Figure 7C). Alternatively, use of the first AG dinucleotide, which in in vitro and in vivo studies is demonstrated to be strongly preferred above any second AG (Smith et al., 1989; Weeda et al., 1990a), would define a very small exon of 9 nt at the end of exon 13. It is known from studies of several investigators that such small exons may be skipped (both in vitro and in vivo) when the surrounding splice sites are nonoptimal (Black, 1991; Dominski and Kole, 1991, 1992). This would result in the ligation of the cryptic splice donor within exon 13 to exon 14 (see Possibility 2 in Figure 7C). For the ERCC6 gene product, the resulting 100 nt deletion has profound consequences. The protein will only be 857 amino acids long and consequently lacks half of the helicase region (see Figure 3B). Presumably, as with allele a,

Table 1. Comparison of Repair Characteristics of UV61 and CS Cells

Repair Characteristics	UV61[a]	CS
UV sensitivity (at D₁₀)	≈2.7×[b]	≈3–3.5×[c]
RNA synthesis recovery after UV	−[d]	−[e]
Repair synthesis (percentage of wild type)	≈70%[f]	Wild-type range[g]
Preferential repair of CPDs in active genes	37%[h]	−[i]
Overall genome repair	−[b,j]	+[g]
Repair of (6–4) photoproducts	+[b]	+[k]
UV-induced mutagenesis	Elevated[b]	Elevated[l,m]

[a] Mutant, probably leaky (Troelstra et al., 1990).
[b] Thompson et al. (1988a).
[c] CS1AN.S3.G2 (unpublished data).
[d] This study.
[e] Mayne and Lehmann (1982).
[f] Unpublished data.
[g] Andrews et al. (1978).
[h] Lommel and Hanawalt (1991).
[i] Venema et al. (1990a).
[j] Characteristic of rodent cells.
[k] Barrett et al. (1991).
[l] Henderson and Long (1981).
[m] Arlett and Harcourt (1982).

allele b in CS1AN cells will produce a nonfunctional ERCC6 protein.

Discussion

A Rodent Equivalent of CS-B

The ERCC6 gene was previously isolated by virtue of its ability to correct the repair deficiency of UV61, a NER-deficient rodent mutant cell line from complementation group 6 (Troelstra et al., 1990). In this paper, we report the isolation of a functional ERCC6 cDNA molecule that efficiently and specifically corrects both the UV sensitivity of UV61 (and other representatives of rodent group 6) and its inability to restore RNA synthesis after UV exposure. Introduction of this cDNA into cells of several XP, CS, and rodent complementation groups by transfection or microneedle injection demonstrates a specific correction of the repair deficiency of CS-B fibroblasts: both UV survival as well as RNA synthesis recovery are restored to a level within wild-type range. This result implies UV61 and other members of rodent complementation group 6 (Thompson et al., 1988b; Busch et al., 1989) to be the rodent equivalents of the human repair disorder CS-B, the most common form of the disease. In retrospect we note that the repair phenotypes of UV61 and CS-B can be reconciled with each other (Table 1), taking into account the fact that rodent and human excision repair display some principal differences (Zelle et al., 1980; van Zeeland et al., 1981; Mellon et al., 1987) and the notion that UV61 is very likely a leaky mutant (Troelstra et al., 1990).

The ERCC6 Mutations in Patient CS1AN

Sequence analysis of the ERCC6 gene in patient CS1AN revealed deleterious mutations in both alleles. Allele a,

which is the only allele present in CS1AN.S3.G2 cells (see below), carries an A[1088] to T transversion leading to a premature stop at amino acid 337. The encoded protein of 336 amino acids lacks all postulated domains (Figure 3B). Allele b contains a C[2648] to T transition that triggers an erroneous splice event and thereby results in a frameshift. As depicted in Figure 7C, two splice mechanisms can be envisaged that could give rise to the deletion of 100 bp. Although undetected in CS1AN mRNA, we cannot exclude the presence of trace amounts of the partially spliced intermediate depicted in Figure 7C (Possibility 1). Similarly, undetectable amounts of mRNA including the newly defined 9 nt exon (depicted in Figure 7C, Possibility 2) might be produced. However, none of these transcripts will encode a protein longer then 860 amino acids. Presumably, these severely truncated ERCC6 proteins will be nonfunctional. Nevertheless, cells from patient CS1AN seem to be affected in preferential DNA repair only, implying that ERCC6 is specific for this NER subpathway. In addition, these findings suggest ERCC6 to be a nonvital protein, in contrast with the ERCC2 and ERCC3 gene products, of which at least the yeast homologs RAD3 and ERCC3[Sc] (SSL2) are shown to be indispensable for cell survival (Higgins et al., 1983; Naumovski and Friedberg, 1983; Gulyas and Donahue, 1992; J. H. J. H., S. Prakash, and L. Prakash, unpublished data).

The immortalized CS1AN derivative CS1AN.S3.G2 only retained allele a and lacked allele b. Presumably, this is not due to loss of one total copy of chromosome 10, since CS1AN.S3.G2 is polymorphic for at least one locus on 10q26 (EFD75; see Figure 6B). A similar case was reported recently with another repair disorder resulting from mutations in the ligase I gene. The primary fibroblasts of the patient, designated 46BR, carried different mutations in each ligase I gene copy, whereas the SV40-transformed derivative 46BR.1G1 only retained one of the two (mutated) alleles (Barnes et al., 1992). Although the genomic instability of SV40-transformed cell lines is generally known, these findings further emphasize the importance of including primary cell lines as controls in this type of study.

ERCC6, Member of a Recently Identified Family of Putative Helicases

The predicted amino acid sequence suggests the ERCC6 gene product to be a nuclear DNA (or RNA) unwinding protein. ERCC2 and ERCC3 have also been postulated to encode DNA helicases, as deduced from their amino acid sequence (Weber et al., 1990; Weeda et al., 1990b), and for ERCC2, in addition, based on the demonstrated DNA helicase activity of its yeast homolog RAD3 (Sung et al., 1987). Therefore, the ERCC6 gene may encode the third DNA unwinding function in mammalian NER. In E. coli both the UvrA₂B complex and the UvrD protein have been shown to exhibit DNA unwinding activity (Oh and Grossman, 1987; Orren et al., 1992, and references therein). Similarly, several DNA helicases might be expected to function in mammalian excision repair. The gene product of ERCC6, however, certainly differs from the E. coli

UvrA₂B and UvrD proteins in that it is specific for preferential repair of active genes.

The entire helicase region of the ERCC6 protein appears to be highly homologous to a similar region in proteins from a recently identified, rapidly expanding family of postulated helicases (Bang et al., 1992; Clark et al., 1992; Davis et al., 1992; Johnson et al., 1992; Laurent et al., 1992; Schild et al., 1992; Tamkun et al., 1992). This novel subgroup encompasses proteins implicated in transcription regulation (SNF2, MOT1, and brm), preservation of chromosome stability (lodestar), and DNA repair (RAD16, RAD54, and the very recently identified RAD5 protein). The group also includes two proteins of unknown function (STH1 and FUN30). The degree of homology (around 30% identity and up to 57% similarity) by far exceeds the resemblance normally present among helicases. This suggests a specific type of helicase function common to all members of this gene family.

The transcription regulators SNF2 (Laurent et al., 1991) and brm (Tamkun et al., 1992) are thought to be engaged in transcription activation, SNF2 as a component of a larger functional complex involving SWI1, SWI3, SNF5, and SNF6 (Laurent et al., 1991; Peterson and Herskowitz, 1992). The gene product MOT1, on the other hand, is supposed to be a negative regulator of transcription (Davis et al., 1992). One of the proposed activities for the MOT1 protein is that it strips resident DNA-binding proteins (which would otherwise activate transcription) from their cognate binding sites by melting the DNA helix (Davis et al., 1992).

Mutations in the maternal-effect gene lodestar result in chromatin bridges at anaphase. The precise molecular pathway in which this protein functions is unknown, but one of the processes suggested is completion of incomplete DNA repair or replication prior to mitosis (Girdham and Glover, 1991).

It is of interest to note that the three major DNA repair pathways in yeast are all equipped with at least one member of this subfamily: RAD16 in NER of the inactive HMLα locus, i.e., opposite from the function of ERCC6 (Terleth et al., 1990); RAD54 in recombination repair (Saeki et al., 1981; Budd and Mortimer, 1982); and RAD5 in postreplication repair (Johnson et al., 1992). Our results demonstrate that preferential repair of active genes also involves a member of this helicase subfamily (ERCC6). Common denominators of all proteins comprising this gene family could be complex formation (as described for SNF2 [Laurent et al., 1991; Peterson and Herskowitz, 1992]), complex delivery at the desired site in DNA, and release of DNA-bound proteins (as proposed for MOT1 [Davis et al., 1992]), although other options can be envisaged as well.

Hypothetical Models for ERCC6 Function

The predicted amino acid sequence of ERCC6 advocates the protein to be a nuclear DNA helicase, with strong sequence homology to the subfamily described above. The ability of the ERCC6 cDNA to correct the repair defect of CS-B fibroblasts and the severe mutations in CS1AN cells provides evidence for a specific role of ERCC6 in preferential repair of active genes. Preferential repair has been documented in mammals (Bohr et al., 1985; Mellon et al., 1986, 1987), yeast (Terleth et al., 1989; Smerdon and Thoma, 1990), and E. coli (Mellon and Hanawalt, 1989). The molecular mechanism of the process itself is, nevertheless, completely unknown. A major step toward defining this NER subpathway in E. coli is the recent purification of the transcription repair coupling factor (TRCF) (Selby and Sancar, 1991). This polypeptide, identified as the mfd gene product (Selby et al., 1991), specifically enhances repair of the template strand of an actively transcribed gene in an in vitro repair system. It is speculated that the protein recognizes a RNA polymerase complex that is stalled at a DNA lesion and forms together with this complex a high affinity binding site for other repair proteins (the UvrABC complex) (Selby et al., 1991). With respect to the biochemical reaction that the ERCC6 protein might catalyze within the mammalian preferential repair process, we can envisage several models. First, a transcription complex stalled at a lesion will probably sterically hinder access to the damage by the NER machinery and should somehow be displaced. A possibility would be that the presumed helicase activity of ERCC6 serves to dissociate the RNA polymerase complex from the DNA, e.g., by locally melting the nascent RNA–DNA hybrid region, or that the protein forces the polymerase complex either past the injury or backward (see MOT1 [Davis et al., 1992]). Extrapolating from this idea, it may be that an important function of this subfamily of putative helicases is to strip bound proteins from specific DNA regions either in the context of repair or in transcription regulation. Dissociation of protein complexes from nucleic acids by helicases has already been suggested in the process of splicing (Ruby and Abelson, 1991). Second, or in addition, the ERCC6 helicase could, as part of a repair complex, be involved in scanning the transcribed strand of active genes for a stalled RNA polymerase, thus efficiently guiding the NER machinery to lesions that thwart transcription (see TRCF [Selby et al., 1991]). Which of the presented or other possible hypotheses concerning ERCC6 function corresponds to the actual role of the protein in the preferential repair subpathway remains to be elucidated.

Experimental Procedures

General Procedures

Purification of nucleic acids, restriction enzyme digests, gel electrophoresis, DNA ligation, synthesis of radiolabeled probes using random oligonucleotide primers, and filter hybridization were performed according to established procedures (Sambrook et al., 1989).

Isolation of cDNA Clones and Sequence Analysis

Unique genomic fragments were used as a probe to hybridize a λgt11 human testis cDNA library (Clonetech). cDNA fragments isolated were subsequently used as probes to screen cDNA libraries from human testis, placenta (Ye et al., 1987), brain (cortex), amygdala, ovary (Clonetech), adenovirus-transformed retina cells, or pre-B cell line Nalm-6 (provided by Dr. A. Bernards), and a cDNA library provided by Dr. H. Okayama. Sequence analysis was performed on denatured double-stranded plasmid by the dideoxy chain termination method (Sanger et al., 1977) using T7 DNA polymerase (Pharmacia). Sequences were determined on both strands either by construction of plasmids containing progressive unidirectional deletions with the Erase-a-Base System (Promega) or by sequence-specific primers. Se-

quences have been compared with the protein and nucleotide data bases using TFASTA and BLAST (Altschul et al., 1990).

Plasmids

Construction of the cDNA expression plasmid was done as follows (see Figure 1): HTL1B and HTL1C (nucleotides 1462–3746 and ±2700–4714), containing several internal EcoRI sites, were isolated from the λ vector (cloning site EcoRI) via PCR and joined at the unique BglII site (nucleotide 2991). The 3.2 kb cDNA thus obtained was cloned into the SalI site of pTZ19R containing a more 5′ but overlapping fragment (HTL1J(0.8E)) of 700 bp, with 205 bp overlap in the following orientation: EcoRI–700 bp fragment 5′ to 3′–EcoRI–SalI). Transformation into E. coli strain DH5αF′ resulted in a plasmid that had recombined the two ERCC6 fragments somewhere in the 205 bp overlapping region. The ERCC6 sequence of this plasmid was verified, and one mistake originating from the PCR was removed by exchanging the NsiI fragment containing the error with the original NsiI fragment. The total ORF was then obtained by ligation of the most 5′ fragment to the PflMI site at position 1361. This almost full-length cDNA molecule was then ligated into a modified pSVL eukaryotic expression vector (Pharmacia; M. van Duin, unpublished data), yielding pSLME6(+).

Cell Lines, Transfection Experiments, UV Selection, and UV Survival

The CHO cell lines used were wild-type AA8, mutants UV41, UV135, UV61, and VB11 from CG-4, CG-5, CG-6, and CG-7, respectively (Zdzienicka et al., 1988; Busch et al., 1989), and two new and unpublished mutants from group 6 provided by Dr. D. Busch. The immortalized human fibroblasts are SV40-A′dam (wild type), CW12 (XP-C from XP7CA), CW3 (XP-E from XP2RO) (Wood et al., 1987), XP2YOSV (XP-F) (Yagi and Takebe, 1983), CS3BE.S3.G1 (CS-A) (Mayne et al., 1986), and CS1AN.S3.G2 (CS-B) (Mayne et al., 1986). Primary fibroblasts used are XP1BI (XP-G) (Keijzer et al., 1979), CS3BE (CS-A), CS1AN (CS-B), and CS1BE (CS-B) (Lehmann, 1982). All CHO cells and immortalized human fibroblasts used were grown in 1:1 F10–Dulbecco's minimal essential medium supplemented with antibiotics and 8%–10% fetal calf serum (F10–DMEM). Primary fibroblasts were cultured in F10 supplemented with antibiotics and 10%–15% fetal calf serum. Transformed cells harboring the E. coli neo dominant marker gene (plasmid pSV2neo) were selected and cultured in medium containing G418 (concentration 300–850 μg/ml, depending on the cell line).

Transfection of CHO cells and immortalized fibroblasts was performed using lipofectin (Bethesda Research Laboratories), essentially as described by the manufacturer. All cells were transfected at approximately 80% confluence (on 60 mm petri dishes), with 2 μg of pSV2neo and 5 μg of pSLME6(+). DNA (a total of 7 μg) and 14 μg of lipofectin were separately added to Opti-MEM (Bethesda Research Laboratories); thereafter, the two were mixed and left at room temperature for 15 min. Cells were washed with phosphate-buffered saline (PBS) before addition of the Opti-MEM–DNA–lipofectin mixture. CHO cells were incubated for 6–7 hr; thereafter 10% fetal calf serum was added. Medium was refreshed with F10–DMEM after approximately 16 hr. Immortalized human fibroblasts were incubated for a similar period, but medium was subsequently refreshed with F10–DMEM. Cells were trypsinized (and, depending on cell density, divided over more dishes) and placed on selection medium 48 hr after transfection. Selection for repair-proficient transformants was started after the appearance of G418-resistant colonies. Cells were trypsinized, and UV selection was started 16–20 hr after trypsinization for CHO cells. 20–48 hr after trypsinization for immortalized human fibroblasts, cells were irradiated three times, CHO cells at 1 day intervals and the SV40-immortalized fibroblasts at 2 day intervals. The UV dose given (Philips TUV low pressure mercury tube, 15 W, 0.45 J/m²/s; predominantly 254 nm) varied with the cell line: 2 J/m² (UV135); 4 J/m² (UV41, XP2YOSV); 6 J/m² (CW12, CS1AN.S3.G2); 8 J/m² (CS3BE.S3.G1); 8.4 J/m² (UV61); and 11 J/m² (VB11).

For determination of UV sensitivity of G418-resistant mass populations (CS3BE.S3.G1 and CS1AN.S3.G2 transformants), G418-resistant plus UV-resistant mass populations (61E6a), or isolated UV-resistant colonies (E61ANa and E61ANd), cells were plated at densities varying from 2 × 10² to 5 × 10⁵ cells per 60 mm petri dish, depending on cell line and UV dose. Cells were rinsed with PBS and exposed to UV light about 1 day

after plating. A series of dishes was irradiated for each cell line, each receiving a single UV dose (three dishes per UV dose). Clones were fixed and stained with Coomassie brilliant blue 6–10 days after UV irradiation, and relative cloning efficiencies were determined.

RNA Synthesis Recovery

The method used was as described by Mayne and Lehmann (1982), with several modifications. Cells were seeded in 30 mm petri dishes in 1.5 ml of medium. One day after seeding, the medium was changed for F10, supplemented with antibiotics, 15% fetal calf serum, 20 mM HEPES, and 0.025 μCi/ml [¹⁴C]thymidine (50 mCi/mmol). The cells were grown for 16 hr, and then the medium was removed; cells were rinsed with PBS and UV irradiated (at 254 nm, fluence rate 0.7 J/m²/s, three petri dishes for each time point or UV dose). At various times after UV irradiation, the medium was replaced by medium containing 8 μCi/ml [³H]uridine (45 Ci/mmol), and the cells were incubated for 1 hr. At the end of the labeling period, cells were washed once with ice-cold PBS containing >1 mM uridine and once with ice-cold PBS. Cells were scraped from the dishes and spotted onto GF/C Whatman glass filters soaked in 10% trichloroacetic acid. These were then washed once in 10% trichloroacetic acid and 96% ethanol, dried, and counted in a liquid scintillation counter. ¹⁴C counts served as an internal control for the number of cells present.

Microinjection

ERCC6 construct pSLME6(+) was injected into nuclei of XP or CS homopolykaryons as described by van Duin et al. (1989). Cells were incubated 24–48 hr to allow expression of the injected DNA. To determine the influence of the ERCC6 gene on DNA repair, injected XP cells were subjected to UV-induced repair synthesis (unscheduled DNA synthesis), which was measured by autoradiography as described previously (Vermeulen et al., 1986). Injected CS fibroblasts were UV irradiated and subsequently incubated with [³H]uridine to measure recovery from UV-induced RNA synthesis inhibition.

DNA Amplification and Mutation Determination

Total RNA (10 μg) was used for preparation of cDNA with ERCC6-specific primers. RNA was dissolved in 5 μl of water, heated for 3 min at 85°C, and cooled on ice. Subsequently, the RNA was added to 15 μl containing 50 pmol of 3′ primer, 18 U of RNAguard (Boehringer), 12.5 U of avian myeloblastosis virus reverse transcriptase (Boehringer), 1 mM dNTPs, 50 mM KCl, 20 mM Tris HCl (pH 8.4 at 20°C), 2.5 mM MgCl₂, and 100 μg/ml bovine serum albumin; incubation was for 1 hr at 37°C. The cDNA (4 μl of the above reaction mixture) was added to 96 μl containing 10 pmol of 5′ primer, 0.2 mM dNTPs, 50 mM KCl, 20 mM Tris HCl (pH 8.4 at 20°C), 2.5 mM MgCl₂, 100 μg/ml BSA, and 2.5 U of Taq polymerase (Bethesda Research Laboratories); the mixture was overlayed with oil. In some experiments, Taq polymerase was added only after the samples reached the denaturing temperature of 94°C (hot start). Amplification was done in 35 cycles: 1 (or 2) cycles of 5 min denaturation at 94°C, 5 min annealing (temperature depending on primer combination used), and 5 min extension at 72°C, followed by 34 (or 33) cycles of 2 min at 94°C, 2 min annealing, and 3 min at 72°C. Chromosomal DNA was amplified as above, starting from 200 ng, in 25 cycles. Asymmetric PCR was performed on 1 μl of PCR product, using the conditions mentioned above (35 cycles), with a primer ratio of 1:100 (0.5 pmol of one and 50 pmol of the other primer). If available, internal primers were used. Products were extracted once with phenol–chloroform–isoamylalcohol and once with chloroform, separated from oligonucleotides and dNTPs on a Centricon-100 column (Amicon), and dried under vacuum in a SpeedVac concentrator (New Brunswick Scientific). Pellets were dissolved in 15 μl, 7 μl of which was used for each sequencing reaction.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Andrews, A. D., Barrett, S. F., Yoder, F. W., and Robbins, J. H. (1978). Cockayne's syndrome fibroblasts have increased sensitivity to ultraviolet light but normal rates of unscheduled DNA synthesis. J. Invest. Dermatol. 70, 237–239.

Arlett, C. F., and Harcourt, S. A. (1982). Variation in response to mutagens amongst normal and repair-defective human cells. In Molecular Mechanisms and Their Implications for Environmental Protection, C. W. Lawrence, ed. (New York: Plenum Press), pp. 249–266.

Bang, D. D., Verhage, R., Goosen, N., Brouwer, J., and van der Putte, P. (1992). Molecular cloning of RAD16, a gene involved in differential repair in Saccharomyces cerevisiae. Nucl. Acids Res., in press.

Barnes, D. E., Tomkinson, A. E., Lehmann, A. R., Webster, A. D. B., and Lindahl, T. (1992). Mutations in the DNA ligase I gene of an individual with immunodeficienies and cellular hypersensitivity to DNA-damaging agents. Cell 69, 495–503.

Barrett, S. F., Robbins, J. H., Tarone, R. E., and Kraemer, K. H. (1991). Evidence for defective repair of cyclobutane pyrimidine dimers with normal repair of other DNA photoproducts in a transcriptionally active gene transfected into Cockayne syndrome cells. Mutat. Res. 255, 281–291.

Black, D. L. (1991). Does steric interference between splice sites block the splicing of a short c-src neuron-specific exon in non-neuronal cells? Genes Dev. 5, 389–402.

Bohr, V. A., Smith, C. A., Okumoto, D. S., and Hanawalt, P. C. (1985). DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. Cell 40, 359–369.

Budd, M., and Mortimer, R. K. (1982). Repair of double-strand breaks in a temperature conditional radiation-sensitive mutant of Saccharomyces cerevisiae. Mutat. Res. 103, 19–24.

Busch, D., Greiner, C., Lewis, K., Ford, R., Adair, G., and Thompson, L. (1989). Summary of complementation groups of UV-sensitive CHO mutants isolated by large-scale screening. Mutagenesis 4, 349–354.

Clark, M. W., Zhong, W. W., Keng, T., Storms, R. K., Barton, A., Kaback, D. B., and Bussey, H. (1992). Identification of a Saccharomyces cerevisiae homolog of the SNF2 transcriptional regulator in the DNA sequence of an 8.6 kb region in the LTE1–CYS1 interval on the left arm of chromosome I. Yeast 8, 133–145.

Cleaver, J. E., and Kraemer, K. H. (1989). Xeroderma pigmentosum. In The Metabolic Basis of Inherited Disease, Vol. 2, C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, eds. (New York: McGraw-Hill), pp. 2949–2971.

Davis, J. L., Kunisawa, R., and Thorner, J. (1992). A presumptive helicase (MOT1 gene product) affects gene expression and is required for viability in the yeast Saccharomyces cerevisiae. Mol. Cell. Biol. 12, 1879–1892.

de Weerd-Kastelein, E. A., Keijzer, W., and Bootsma, D. (1972). Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization. Nature New Biol. 238, 80–83.

Dominski, Z., and Kole, R. (1991). Selection of splice sites in pre-mRNAs with short internal exons. Mol. Cell. Biol. 11, 6075–6083.

Dominski, Z., and Kole, R. (1992). Cooperation of pre-mRNA sequence elements in splice site selection. Mol. Cell. Biol. 12, 2108–2114.

Emery, H. S., Schild, D. S., Kellogg, D. E., and Mortimer, R. K. (1991).

Sequence of RAD54, a Saccharomyces cerevisiae gene involved in recombination and repair. Gene 104, 103–106.

Flejter, W. L., McDaniel, L. D., Johns, D., Friedberg, E. C., and Schultz, R. A. (1992). Correction of xeroderma pigmentosum complementation group D mutant cell phenotypes by chromosome and gene transfer: involvement of the human ERCC2 DNA repair gene. Proc. Natl. Acad. Sci. USA 89, 261–265.

Friedberg, E. C. (1985). DNA Repair (San Francisco: W. H. Freeman and Company).

Fryns, J. P., Bulcke, J., Verdu, P., Carton, H., Kleczkowska, A., and Van der Berghe, H. (1991). Apparent late-onset Cockayne syndrome and interstitial deletion of the long arm of chromosome 10 (del(10)(q11.-23q21.2)). Am. J. Med. Genet. 40, 343–344.

Girdham, C. H., and Glover, D. M. (1991). Chromosome tangling and breakage at anaphase result from mutations in lodestar, a Drosophila gene encoding a putative nucleoside triphosphate-binding protein. Genes Dev. 5, 1786–1799.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucl. Acids Res. 17, 4713–4730.

Gulyas, K. D., and Donahue, T. F. (1992). SSL2, a suppressor of a stem–loop mutation in the HIS4 leader encodes the yeast homolog of human ERCC-3. Cell 69, 1031–1042.

Henderson, E. E., and Long, W. K. (1981). Host cell reactivation of UV- and X-ray-damaged herpes simplex virus by Epstein–Barr virus (EBV)-transformed lymphoblastoid cell lines. Virology 115, 237–248.

Higgins, D. R., Prakash, S., Reynolds, P., Polakowska, R., Weber, S., and Prakash, L. (1983). Isolation and characterization of the RAD3 gene of Saccharomyces cerevisiae and inviability of RAD3 deletion mutants. Proc. Natl. Acad. Sci. USA 80, 5680–5684.

Huang, J. C., Svoboda, D. L., Reardon, J. T., and Sancar, A. (1992). Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. Proc. Natl. Acad. Sci. USA 89, 3664–3668.

Johnson, R. E., Henderson, S. T., Petes, T. D., Prakash, S., Bankmann, M., and Prakash, L. (1992). Saccharomyces cerevisiae RAD5-encoded DNA repair protein contains DNA helicase and zinc-binding sequence motifs, and affects the stability of simple repetitive sequences in the genome. Mol. Cell. Biol. 12, 3807–3818.

Keijzer, W., Jaspers, N. G. J., Abrahams, P. J., Taylor, A. M. R., Arlett, C. F., Zelle, B., Takebe, H., Kinmont, P. D. S., and Bootsma, D. (1979). A seventh complementation group in excision deficient xeroderma pigmentosum. Mutat. Res. 62, 183–190.

Kemp, B. E., and Pearson, R. B. (1990). Protein kinase recognition sequence motifs. Trends Biochem. Sci. 15, 342–346.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. 266, 19867–19870.

Lapeyre, B., Bourbon, H., and Amalric, F. (1987). Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. Proc. Natl. Acad. Sci. USA 84, 1472–1476.

Laurent, B. C., Treitel, M. A., and Carlson, M. (1991). Functional interdependence of the yeast SNF2, SNF5, and SNF6 proteins in transcriptional activation. Proc. Natl. Acad. Sci. USA 88, 2687–2691.

Laurent, B. C., Yang, X., and Carlson, M. (1992). An essential Saccharomyces cerevisiae gene homologous to SNF2 encodes a helicase-related protein in a new family. Mol. Cell. Biol. 12, 1893–1902.

Lehmann, A. R. (1982). Three complementation groups in Cockayne syndrome. Mutat. Res. 106, 347–356.

Lehmann, A. R. (1987). Cockayne's syndrome and trichothiodystrophy: defective repair without cancer. Cancer Rev. 7, 82–103.

Lommel, L., and Hanawalt, P. C. (1991). The genetic defect in the Chinese hamster ovary cell mutant UV61 permits moderate selective repair of cyclobutane pyrimidine dimers in an expressed gene. Mutat. Res. 255, 183–191.

Mayne, L. V., and Lehmann, A. R. (1982). Failure of RNA synthesis to recover after UV irradiation: an early defect in cells from individuals

with Cockayne's syndrome and xeroderma pigmentosum. Cancer Res. 42, 1473–1478.

Mayne, L. V., Priestley, A., James, M. R., and Burke, J. F. (1986). Efficient immortalization and morphological transformation of human fibroblasts by transfection with SV40 DNA linked to a dominant marker. Exp. Cell Res. 162, 530–538.

Mellon, I., and Hanawalt, P. C. (1989). Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand. Nature 342, 95–98.

Mellon, I., Bohr, V. A., Smith, C. A., and Hanawalt, P. C. (1986). Preferential DNA repair of an active gene in human cells. Proc. Natl. Acad. Sci. USA 83, 8878–8882.

Mellon, I., Spivak, G., and Hanawalt, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. Cell 51, 241–249.

Nance, M. A., and Berry, S. A. (1992). Cockayne syndrome: review of 140 cases. Am. J. Med. Genet. 42, 68–84.

Naumovski, L., and Friedberg, E. C. (1983). A DNA repair gene required for the incision of damaged DNA is essential for viability in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA 80, 4818–4821.

Oh, E. Y., and Grossman, L. (1987). Helicase properties of the Escherichia coli UvrAB protein complex. Proc. Natl. Acad. Sci. USA 84, 3638–3642.

Orren, D. K., and Sancar, A. (1989). The (A)BC excinuclease of Escherichia coli has only the UvrB and UvrC subunits in the incision complex. Proc. Natl. Acad. Sci. USA 86, 5237–5241.

Orren, D. K., Selby, C. P., Hearst, J. E., and Sancar, A. (1992). Post-incision steps of nucleotide excision repair in Escherichia coli: disassembly of the UvrBC–DNA complex by helicase II and DNA polymerase I. J. Biol. Chem. 267, 780–788.

Peterson, C. L., and Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell 68, 573–583.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature 335, 683–689.

Rihs, H. P., Jans, D. A., Fan, H., and Peters, R. (1991). The rate of nuclear cytoplasmic protein transport is determined by the casein kinase II site flanking the nuclear localization sequence of the SV40 T-antigen. EMBO J. 10, 633–639.

Robbins, J., Dilworth, S. M., Laskey, R. A., and Dingwall, C. (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell 64, 615–623.

Ruby, S. W., and Abelson, J. (1991). Pre-mRNA splicing in yeast. Trends Genet. 7, 79–85.

Ruskin, B., Krainer, A. R., Maniatis, T., and Green, M. R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38, 317–331.

Saeki, T., Machida, I., and Nakai, S. (1981). Genetic control of diploid recovery after gamma-irradiation in the yeast Saccharomyces cerevisiae. Mutat. Res. 73, 251–265.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990). The P-loop: a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. 15, 430–434.

Schild, D., Glassner, B. J., Mortimer, R. K., Carlson, M., and Laurent, B. C. (1992). Identification of RAD16, a yeast excision repair gene homologous to the recombinational repair gene RAD54 and to the SNF2 gene involved in transcriptional activation. Yeast 8, 385–395.

Selby, C. P., and Sancar, A. (1991). Gene- and strand-specific repair in vitro: partial purification of a transcription-repair coupling factor. Proc. Natl. Acad. Sci. USA 88, 8232–8236.

Selby, C. P., Witkin, E. M., and Sancar, A. (1991). Escherichia coli mfd

mutant deficient in "mutation frequency decline" lacks strand-specific repair: in vitro complementation with purified coupling factor. Proc. Natl. Acad. Sci. USA 88, 11574–11578.

Senapathy, P., Shapiro, M. B., and Harris, N. L. (1990). Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. Meth. Enzymol. 183, 252–278.

Shivji, M. K. K., Kenny, M. K., and Wood, R. D. (1992). Proliferating cell nuclear antigen is required for DNA excision repair. Cell 69, 367–374.

Simpson, N. E., and Cann, H. (1990). Report of the committee on the genetic constitution of chromosome 10. Cytogenet. Cell Genet. 55, 142–152.

Smerdon, M. J., and Thoma, F. (1990). Site-specific DNA repair at the nucleosome level in a yeast minichromosome. Cell 61, 675–684.

Smith, C. W. J., Porro, E. B., Patton, J. G., and Nadal-Ginard, B. (1989). Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. Nature 342, 243–247.

Sung, P., Prakash, L., Matson, S. W., and Prakash, S. (1987). RAD3 protein of Saccharomyces cerevisiae is a DNA helicase. Proc. Natl. Acad. Sci. USA 84, 8951–8955.

Sung, P., Prakash, S., and Prakash, L. (1988). The RAD6 protein of Saccharomyces cerevisiae polyubiquitinates histones, and its acidic domain mediates this activity. Genes Dev. 2, 1476–1485.

Tamkun, J. W., Deuring, R., Scott, M. P., Kissinger, M., Pattatucci, A. M., Kaufman, T. C., and Kennison, J. A. (1992). brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell 68, 561–572.

Tanaka, K., Kawai, K., Kumahara, Y., Ikenaga, M., and Okada, Y. (1981). Genetic complementation groups in Cockayne syndrome. Som. Cell Genet. 7, 445–455.

Tanaka, K., Miura, N., Satokata, I., Miyamoto, I., Yoshida, M. C., Satoh, Y., Kondo, S., Yasui, A., Okayama, H., and Okada, Y. (1990). Analysis of a human DNA excision repair gene involved in group A xeroderma pigmentosum and containing a zinc-finger domain. Nature 348, 73–76.

Terleth, C., van Sluis, C. A., and van der Putte, P. (1989). Differential repair of UV damage in Saccharomyces cerevisiae. Nucl. Acids Res. 17, 4433–4439.

Terleth, C., Schenk, P., Poot, R., Brouwer, J., and van der Putte, P. (1990). Differential repair of UV damage in rad mutants of Saccharomyces cerevisiae: a possible function of G_2 arrest upon UV irradiation. Mol. Cell. Biol. 10, 4678–4684.

Thompson, L. H., Mitchell, D. L., Regan, J. D., Bouffler, S. D., Stewart, S. A., Carrier, W. L., Nairn, R. S., and Johnson, R. T. (1988a). CHO mutant UV61 removes (6–4) photoproducts but not cyclobutane dimers. Mutagenesis 4, 140–146.

Thompson, L. H., Shiomi, T., Salazar, E. P., and Stewart, S. A. (1988b). An eight complementation group of rodent cells hypersensitive to ultraviolet radiation. Som. Cell. Mol. Genet. 14, 605–612.

Troelstra, C., Odijk, H., de Wit, J., Westerveld, A., Thompson, L. H., Bootsma, D., and Hoeijmakers, J. H. J. (1990). Molecular cloning of the human DNA excision repair gene ERCC-6. Mol. Cell. Biol. 10, 5806–5813.

Troelstra, C., Landsvater, R. M., Wiegant, J., van der Ploeg, M., Viel, G., Buys, C. H. C. M., and Hoeijmakers, J. H. J. (1992). Localization of the nucleotide excision repair gene ERCC6 to human chromosome 10q11–q21. Genomics 12, 745–749.

van Duin, M., de Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M. H. M., Hoeijmakers, J. H. J., and Bootsma, D. (1986). Molecular characterization of the human excision repair gene ERCC-1: cDNA cloning and amino acid homology with the yeast DNA repair gene RAD10. Cell 44, 913–923.

van Duin, M., Vredeveldt, G., Mayne, L. V., Odijk, H., Vermeulen, W., Klein, B., Weeda, G., Hoeijmakers, J. H. J., Bootsma, D., and Westerveld, A. (1989). The cloned human DNA excision repair gene ERCC-1 fails to correct xeroderma pigmentosum complementation groups A through I. Mutat. Res. 217, 83–92.

van Houten, B. (1990). Nucleotide excision repair in Escherichia coli. Microbiol. Rev. 54, 18–51.

van Zeeland, A. A., Smith, C. A., and Hanawalt, P. C. (1981). Sensitive determination of pyrimidine dimers in DNA of UV-irradiated mammalian cells. Mutat. Res. *82*, 173–189.

Venema, J. (1991). Gene specific repair of UV-induced DNA damage in mammalian cells. PhD thesis, State University of Leiden, Leiden, The Netherlands.

Venema, J., Mullenders, L. H., Natarajan, A. T., van Zeeland, A. A., and Mayne, L. V. (1990a). The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. Proc. Natl. Acad. Sci. USA *87*, 4707–4711.

Venema, J., van Hoffen, A., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. (1990b). The residual repair capacity of xeroderma pigmentosum complementation group C fibroblasts is highly specific for transcriptionally active DNA. Nucl. Acids Res. *18*, 443–448.

Vermeulen, W., Osseweijer, P., de Jonge, A. J., and Hoeijmakers, J. H. J. (1986). Transient correction of excision repair defects in fibroblasts of 9 xeroderma pigmentosum complementation groups by microinjection of crude human cell extracts. Mutat. Res. *165*, 199–206.

Vermeulen, W., Stefanini, M., Giliani, S., Hoeijmakers, J. H. J., and Bootsma, D. (1991). Xeroderma pigmentosum complementation group H falls into complementation group D. Mutat. Res. *255*, 201–208.

Vrieling, H., Venema, J., van Rooyen, M. L., van Hoffen, A., Menichini, P., Zdzienicka, M. Z., Simons, J. W., Mullenders, L. H., and van Zeeland, A. A. (1991). Strand specificity for UV-induced DNA repair and mutations in the Chinese hamster *HPRT* gene. Nucl. Acids Res. *19*, 2411–2415.

Weber, C. A., Salazar, E. P., Stewart, S. A., and Thompson, L. H. (1990). *ERCC2*: cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast RAD3. EMBO J. *9*, 1437–1447.

Weeda, G., van Ham, R. C. A., Masurel, R., Westerveld, A., Odijk, H., de Wit, J., Bootsma, D., van der Eb, A. J., and Hoeijmakers, J. H. J. (1990a). Molecular cloning and biological characterization of the human excision repair gene *ERCC-3*. Mol. Cell. Biol. *10*, 2570–2581.

Weeda, G., van Ham, R. C. A., Vermeulen, W., Bootsma, D., van der Eb, A. J., and Hoeijmakers, J. H. J. (1990b). A presumed DNA helicase encoded by *ERCC-3* is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. Cell *62*, 777–791.

Wen, L., Huang, J., Johnson, B., and Reeck, G. R. (1989). A human placental cDNA clone that encodes nonhistone chromosomal protein HMG-1. Nucl. Acids Res. *17*, 1197–1214.

Wood, C. M., Timme, T. L., Hurt, M. M., Brinkley, B. R., Ledbetter, D. H., and Moses, R. E. (1987). Transformation of DNA repair-deficient human diploid fibroblasts with a simian virus 40 plasmid. Exp. Cell Res. *169*, 543–553.

Yagi, T., and Takebe, H. (1983). Establishment by SV40 transformation and characteristics of a cell line of xeroderma pigmentosum belonging to complementation group F. Mutat. Res. *112*, 59–66.

Ye, R. D., Wun, T. C., and Sadler, J. E. (1987). cDNA cloning and expression in *Escherichia coli* of a plasminogen activator inhibitor from human placenta. J. Biol. Chem. *262*, 3718–3725.

Zdzienicka, M. Z., van der Schans, G. P., and Simons, J. W. I. M. (1988). Identification of a new seventh complementation group of UV-sensitive mutants in Chinese hamster cells. Mutat. Res. *194*, 165–170.

Zelle, B., Reynolds, R. J., Kottenhagen, M. J., Schutte, A., and Lohman, P. H. M. (1980). The influence of the wavelength radiation on survival, mutation induction and DNA repair in irradiated Chinese hamster cells. Mutat. Res. *72*, 491–509.

GenBank Accession Number