ERCC6, A GENE INVOLVED IN COCKAYNE'S SYNDROME

Christine Troelstra

ERCC6, A GENE INVOLVED IN COCKAYNE'S SYNDROME

ERCC6, EEN GEN BETROKKEN BIJ COCKAYNE'S SYNDROME

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
AAN DE ERASMUS UNIVERSITEIT ROTTERDAM
OP GEZAG VAN DE RECTOR MAGNIFICUS
PROF. DR. C.J. RIJNVOS
EN VOLGENS BESLUIT VAN HET COLLEGE VAN DEKANEN.

DE OPENBARE VERDEDIGING ZAL PLAATSVINDEN OP
WOENSDAG 24 FEBRUARI 1993 OM 15.45 UUR

DOOR
CHRISTINE TROELSTRA
GEBOREN TE AMSTERDAM

PROMOTIECOMMISSIE

Promotor:	Prof. Dr. D. Bootsma
Overige leden:	Prof. Dr. A. Westerveld
	Prof. Dr. Ir. A. A. van Zeeland
	Prof. Dr. M. F. Niermeijer
Co-Promotor:	Dr. J. H. J. Hoeijmakers

Dit proefschrift werd bewerkt binnen de vakgroep Celbiologie en Genetica van de Faculteit der Geneeskunde en Gezondheidswetenschappen van de Erasmus Universiteit Rotterdam. De vakgroep maakt deel uit van het Medisch Genetisch Centrum Zuid-West Nederland.

Gedrukt door: Drukkerij Haveka B.V., Alblasserdam

Contents

Chapter I

General introduction

General introduction

The integrity of the genetic material is continuously threatened by a variety of environmental agents (e.g. UV irradiation, various chemicals), faulty DNA metabolism, and chemical instability of the DNA itself. To protect cells from the deleterious accumulation of DNA damage, several DNA repair mechanisms have evolved (Friedberg, 1985 for a review). One of the most important and intensively studied repair processes is the nucleotide excision repair (NER). NER is capable of removing many different lesions, among which are UV (and sunlight)-induced cyclobutane pyrimidine dimers (CPD), pyrimidine (6-4) pyrimidone photoproducts ((6-4) photoproducts), and thymine glycols, as well as bulky chemical adducts and DNA crosslinks. If unremoved, these lesions will disturb fundamental cellular processes: they will inhibit both DNA-transcription and -replication, which might cause cell death. Moreover, erroneous replication passed a lesion will lead to mutations. These can interfere with the expression or functioning of genes involved in - for instance - regulation of cell proliferation and differentiation, thereby possibly providing a step in the process of carcinogenesis.

The first experimental evidence for the importance of DNA repair in prevention of cancer was obtained in 1968, when Cleaver reported a deficiency of NER in patients with the cancer-prone, hereditary disease xeroderma pigmentosum (XP) (Cleaver, 1968). XP is a rare, autosomal recessive disease, with a frequency of about 1 to 10 patients per million individuals. Clinically, patients are characterised by sun sensitivity of the skin, pigmentation abnormalities, and a highly elevated risk for developing various types of skin tumors. Some patients have, in addition, neurological abnormalities and growth disturbances (Cleaver and Kraemer, 1989 for a review on XP). Cell fusion experiments have identified 8 complementation groups (cg) among cells of different XP patients, 7 of which are NER-deficient (XP-A to -G; De Weerd-Kastelein, et al., 1972; Vermeulen, et al., 1991). The eighth cg (XP-variant) comprises XP patients with an apparently normal NER, but a defect in postreplication repair (Lehmann, et al., 1975). Increased susceptibility for DNA damaging agents has now been identified as a common denominator for a number of other rare, cancer-prone (autosomal recessive) genetic disorders, such as ataxia telangiectasia (AT), Bloom's syndrome (BS), and Fanconi's anemia (FA). In contrast to patients with these cancer-prone repair disorders are individuals suffering from Cockayne's syndrome (CS) and some of the cases with trichothiodystrophy (TTD). Cells from both CS and a fraction of TTD patients carry defects in the NER process, but the disorders are not

associated with an appreciably elevated risk for (skin)tumor formation (Lehmann, 1987). Both syndromes are, as XP, genetically heterogeneous - two complementation groups have been identified among the classical CS patients (Tanaka, et al., 1981; Lehmann, 1982), and at least two in NER-deficient TTD (Stefanini, et al., 1986, and W. Vermeulen, personal communication), one of which is overlapping with XP-D - suggesting mammalian NER to be highly complex.

The molecular mechanism of NER has been a major subject of interest for many years. In *E.coli* the process has been unravelled to considerable detail (Van Houten, 1990); knowledge about the more complex eukaryotic process is now rapidly increasing (Huang, et al., 1992; Shivji, et al., 1992). To understand this intriguing repair system in mammalian cells, as well as its relation to carcinogenesis, extensive information about the function of the genes and gene products involved is required. The NER-deficient cultured cells from XP and CS patients are, evidently, valuable tools. XP cells are being exploited to purify XP-correcting factors, which has resulted in the isolation of at least one NER protein (Robins, et al., 1991; Sugano, et al., 1991; Eker, et al., 1992). In another approach, NER-deficient cell lines serve to isolate correcting genes by means of DNA-mediated gene transfer. For this purpose, the *in vitro* mutagenised UV-sensitive rodent cell lines with XP-like phenotypes have proven to be at least as valuable as XP or CS cells. Among these UV-sensitive rodent cell lines, 12 complementation groups have been identified (Thompson, et al., 1988b; Zdzienicka, et al., 1988; Busch, et al., 1989; Hata, et al., 1991; Stefanini, et al., 1991 and D. Busch, personal communication). Gene transfer to the rodent mutant cell lines has resulted in the isolation of five NER genes (Westerveld, et al., 1984; Weber, et al., 1988; Mudgett and MacInnes, 1990; Weeda, et al., 1990a, and Chapter III), at least three of which are involved in XP or CS syndromes (Weeda, et al., 1990b; Flejter, et al., 1992, and Chapter V). DNA transfections to XP cells have resulted in the isolation of another two XP-correcting genes (Tanaka, et al., 1990; Legerski and Peterson, 1992). Though far from complete, this collection of NER genes and proteins is a powerful source from which we can begin to assemble our understanding of the many different reactions involved in the removal of one single DNA lesion in a mammalian cell.

Scope of the thesis

Knowledge on the molecular mechanism of nucleotide excision repair (NER) in mammalian cells is rapidly increasing, but far from complete. The aim of the experimental work described in this thesis is to enrich our understanding of the NER process, through the identification and characterization of gene(s) involved. Hereto, the UV-sensitivity of a Chinese hamster ovary (CHO) mutant cell line from rodent complementation group 6 (UV61) was restored to normal levels by means of DNA-mediated gene transfer. The correcting human gene, designated *ERCC6*, has been isolated and characterized - as described in Chapters III to VI. Mutations in the gene appeared to be responsible for (the most common form of) the hereditary DNA repair disorder Cockayne's syndrome. A brief overview of (mammalian) NER is given in Chapter II. Finally, in Chapter VII the current data on the molecular mechanism of (mammalian) NER are discussed.

Chapter II

Nucleotide excision repair

Nucleotide excision repair

Nucleotide excision repair (NER) is one of the most versatile DNA repair mechanisms of the cell. It is capable of removing a broad category of DNA lesions having very dissimilar structures such as UV-induced cyclobutane pyrimidine dimers (CPD), pyrimidine (6-4) pyrimidone photoproducts ((6-4) photoproducts), and thymine glycols, as well as bulky chemical adducts. The importance of this repair pathway is apparent from the phenotype of NER-deficient mutants in various organisms. These mutants are being extensively utilized in studies aimed at unravelling the NER process. Most of our current knowledge on the molecular mechanism comes from the bacterium *Escherichia coli* (Van Houten, 1990).

NER in *Escherichia coli*

NER can be schematically divided into five successive steps, being (1) damage recognition, (2) incision of the damaged strand at both sites of the lesion, (3) excision of the damaged strand including the adduct, (4) DNA repair synthesis to fill the single stranded gap, and (5) ligation of the newly synthesized strand. In *E.coli*, at least six proteins are involved in this repair process: UvrA, -B, -C, -D, DNA polymerase I and DNA ligase. The purification of these proteins as well as the isolation of the genes encoding them have been key steps towards the current detailed - but still far from complete - mechanistic model (Van Houten, 1990 for a review).

A simplified scheme of this repair reaction is presented in Figure 1. The UvrA subunit dimerizes, and associates with UvrB, to form a $UvrA_2B$ complex which will bind DNA. The $UvrA_2B$ complex possesses a DNA helicase activity, by which it is able to scan the DNA strand in search for DNA damage (Grossman and Yeung, 1990). The complex is believed to recognize a general structural alteration of the helix, rather than the damaged base(s) itself. This could explain the broad spectrum of damage that can be removed. Upon encountering a damaged site, a stable preincision complex is formed and UvrA probably dissociates from the DNA (Orren and Sancar, 1989). Once the preincision complex is formed UvrC binds. This ensures nicking of the 5th phosphodiester bond 3' to the lesion by UvrB; subsequently UvrC will incise the damaged strand at the eighth phosphodiester bond 5' to the damaged base (Lin, et al., 1992; Lin and Sancar, 1992). Both the resulting oligonucleotide, and the UvrC protein are released through the action of the helicase activity of UvrD, while DNA polymerase I is needed to fill the gap and release the UvrB protein (Orren, et al., 1992). Finally, the newly synthesized strand is ligated to the parental

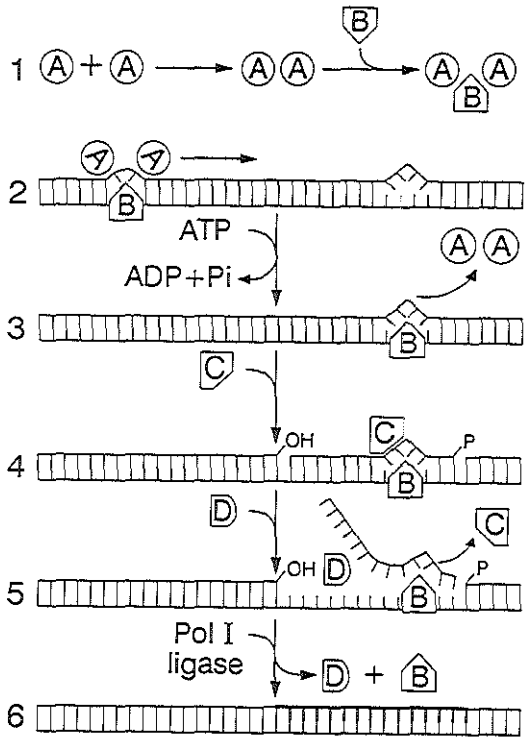


Figure 1. A model for the nucleotide excision repair mechanism in *Escherichia coli*.

DNA by DNA ligase.

Preferential repair of active genes: a NER subpathway

To efficiently remove DNA damage from the genome, NER appears to be divided into two - partly overlapping - subpathways. The presence of lesions in the DNA has important consequences on both DNA replication and transcription, the most urgent effect being the inhibition of actively transcribing RNA polymerase complexes. To get around the lethal effect of this transcription block, a sophisticated NER-subpathway has evolved that is primarily directed towards removal of lesions from the transcribed strand of active genes (Link Jr., et al., 1991). This ensures fast repair of the relatively small - but vital - active fraction of the genome. The second subpathway is dedicated to the removal of damage from both inactive chromatin and the non-transcribed strand of active genes. The rate of this "overall genome" repair is lower

than that of repair of the transcribed strand of active genes, and for some lesions repair is incomplete (Link Jr., et al., 1991 for a review). In human cells, repair of the nontranscribed strand of active genes is faster than repair of an inactive locus, but not as efficient as that of the transcribed strand (Venema, et al., 1992). The delicate subpathway of preferential repair was first recognized in rodent cells (Bohr, et al., 1985). These cells have - in contrast to human fibroblasts - almost no detectable overall genome repair, whereas they are as UV-resistant as human cells (Zelle, et al., 1980; Van Zeeland, et al., 1981). The rodent cells appear to mainly rely on the repair of the active fraction of their genome, whereas human cells have a more extensive overall genome repair (Bohr, et al., 1985; Mellon, et al., 1986; Mellon, et al., 1987). The existence of preferential repair of the transcribed strand of active genes has been documented also in yeast (Smerdon and Thoma, 1990), and *E.coli* (Mellon and Hanawalt, 1989), indicative of its fundamental importance.

In *E.coli*, a direct coupling between transcription and preferential NER seems to exist. If CPD removal is measured in the lactose operon upon UV-irradiation of *E.coli* cells, preferential repair of the transcribed strand is only found when transcription is induced. Repair of the non-transcribed strand is slower, and resembles that of both strands in the non-induced operon (Mellon and Hanawalt, 1989). Preferential repair (of CPD) of a transcribed gene also occurs in an *E. coli in vitro* repair assay. A defined NER system, consisting of purified UvrABC excinuclease, UvrD, DNA polymerase I, and T4 DNA ligase, together with RNA polymerase exhibits strand-selective repair only upon addition of a recently purified factor, called TRCF (Transcription Repair Coupling Factor; Selby and Sancar, 1991). When the factor is omitted, transcription no longer stimulates, but - on the contrary - appears to inhibit preferential repair of the transcribed strand in this defined system (Selby and Sancar, 1990). Based on these results, a model was proposed explaining the role of TRCF in coupling repair to transcription (Selby, et al., 1991). In this model, TRCF recognizes the RNA polymerase complex as soon as it is stalled at a DNA lesion. Either alone or together with the RNA polymerase it would then serve as a high affinity binding site for the $UvrA_2B$ complex, thus ensuring fast repair of the transcribed strand.

An *E.coli* mutant, *mfd* (*m*utation *f*requency *d*ecline), appears to lack a functional TRCF. Cell-free extracts show no enhanced repair of the transcribed strand of an active gene in an *in vitro* repair assay; a phenomenon restored by the addition of purified TRCF (Selby, et al., 1991). The mutant itself, however, shows a (practically) no increased UV-sensitivity, but is reported to have an elevated (5x) mutation frequency (George and Witkin, 1974). This might be explained by the fact that

cultured *E.coli* cells grow - compared to bacteria in non-laboratory conditions - exceptionally fast. Since DNA replication is much faster than transcription, DNA polymerase complexes will frequently encounter and subsequently remove RNA polymerase molecules stalled at a DNA lesion. Translesion DNA synthesis will thereafter be necessary for replication to continue, resulting in the observed elevated mutation frequency. It would, on the other hand, allow fast resumption of RNA synthesis (on the duplicated genome), which could account for the unaffected UV-survival of the mutant cells.

Human NER-deficient disorders

The phenotypic consequence of a NER deficiency in human cells is clearly illustrated by two different, rare, hereditary disorders: xeroderma pigmentosum (XP) and Cockayne's syndrome (CS). Patients suffering from the prototype DNA repair syndrome XP are extremely photosensitive, exhibit pigmentation abnormalities in sun-exposed areas of the skin and have a highly elevated risk for developing skin tumors (≈2000x). Some XP patients (≈20-35%) have, in addition, a progressive neurological degeneration. In CS too, patients are photosensitive, but in contrast to XP this syndrome is not associated with an elevated risk for developing skin tumors. CS is further characterized by cachectic dwarfism, skeletal and retinal abnormalities, as well as progressive neurological degeneration (Lehmann, 1987; Nance and Berry, 1992 for reviews).

XP fibroblasts show enhanced killing by UV light. The UV-dependent DNA repair synthesis seen in normal cells after irradiation is (severely) reduced, indicating that UV-induced DNA damage is not repaired in XP fibroblasts (Cleaver and Kraemer, 1989 for an extensive review on XP). Introduction of the dimer-specific T4-endonuclease V, isolated from T4-infected *E.coli*, or the *Micrococcus luteus* endonuclease into XP cells results in near normal levels of DNA repair synthesis (Tanaka, et al., 1975; Tanaka, et al., 1977; Hayakawa, et al., 1981; De Jonge, et al., 1985). From these experiments it was concluded that XP cells are probably deficient in one of the early steps of NER. Cell fusion experiments have identified seven excision-deficient XP complementation groups (XP-A to -G, summarized in Table I), suggesting a considerable biochemical complexity (De Weerd-Kastelein, et al., 1972; Vermeulen, et al., 1991).

Complementation group A comprises a large number of patients. Generally, NER seems to be blocked completely in XP-A fibroblasts; the residual repair synthesis in most cases is less than 5% (and could be derived from base excision repair of e.g. UV-induced thymine glycols). Consequently, most group A patients manifest a severe clinical phenotype, with clear neurological degeneration (see Table I). Several patients

Table I. Characteristics of NER-deficient XP and CS complementation groups

Group	Clinical characteristics			Repair parameters		Rodent equivalent	Correcting gene	Remarks
	Skin cancer	Neurol. abnorm.	Relative frequency	UV sens.	Residual UDS			
XP-A	+	++	high	+++	<5%	?	*XPAC*	
XP-B	+/-	+++/+	very rare	++	<10%	group 3	*ERCC3*	combined XP/CS
XP-C	+	-	high	+	15-30%	?	*XPCC*	overall genome repair deficient
XP-D	+	++/-	intermediate	++	15-50%	group 2	*ERCC2*	includes TTD and XP/CS
XP-E	+	-	rare	±	>50%	?	?	
XP-F	+	-	rare/ intermediate	+	15-30%	?	?	repair slow but prolonged
XP-G	+/-	++/+	rare	++	<10%	?	?	includes XP/CS
CS-A	-	++	rare	+	wild type	?	?	preferential repair of transcribed genes deficient
CS-B	-	++	high	+	wild type	group 6	*ERCC6*	preferential repair of transcribed genes deficient

have been described that exhibit a milder form of the disease. The gene correcting the repair defect of XP-A cells (the XP-A correcting or *XPAC* gene) has been isolated and characterized (Tanaka, et al., 1990). Mutation analysis of this gene has been performed in some of the patients. The mutations found in the milder form of the disease seem to have a less rigorous effect on protein function than those found in severely affected patients (Satokata, et al., 1990; Satokata, et al., 1992a; Satokata, et al., 1992b).

In the extremely rare group B, only consisting of 3 patients, all individuals exhibit a remarkable combination of clinical characteristics of XP and CS (Robbins, et al., 1974; Cleaver and Kraemer, 1989 and Vermeulen, manuscript in preparation). Also here, excision repair is severely affected, the residual repair synthesis is $<10\%$.

Fibroblasts from complementation group C are less UV-sensitive compared to the severely affected group A and B cells; the residual repair synthesis is between 15 and 30%. Interestingly, these cells harbor a specific deficiency in the repair of the "overall genome", whereas the transcribed strand of active genes is repaired at normal levels (Kantor, et al., 1990; Venema, et al., 1990b; Venema, et al., 1991). This implies the defective protein in XP-C to function mainly (if not solely) in repair of inactive chromatin and the non-transcribed strand of active genes. Phenotypically, XP-C patients differ from group A patients in that they seldomly exhibit neurological abnormalities (Cleaver and Kraemer, 1989).

XP group D is, clinically as well as biochemically, very heterogeneous (Johnson and Squires, 1992 for a recent review). It encompasses classical XP patients, patients suffering from a combined XP/CS phenotype, and patients with trichothiodystrophy (TTD). TTD is a rare, recessive disorder characterized by sulphur-deficient brittle hair, mental and growth retardation, ichthyosis, and - only in 20% of the cases - photosensitivity (Lehmann, 1989; Stefanini, et al., 1992). In several individuals the genetic defect resulting in photosensitivity is demonstrated to be similar to the XP-D defect (Stefanini, et al., 1986; Stefanini, et al., 1992). The patients do not show the pigmentation abnormalities or elevated incidence of skin tumors characteristic for XP. It is possible, though, that the severe ichthyosis masks part of the XP-specific skin defects; it might even (partly) protect cells in the skin from sun(UV)light. Among the XP group D TTD patients, the NER deficiency can either concern removal of both (6-4) photoproducts and CPD or removal of (6-4) photoproducts only (Broughton, et al., 1990).

DNA repair in the rare XP group E is only moderately affected, the residual repair synthesis being $>50\%$. The skin symptoms of patients in this group are mild, and neurological degeneration has not been reported. In 1988, Chu and Chang (Chu and Chang, 1988) reported the apparent absence of a DNA binding activity in cells

from two (consanguineous) XP-E patients (XP2RO and XP3RO). Recently it was reported that in 8 out of 9 other, unrelated, Japanese XP-E patients the binding activity was retained (Keeney, et al., 1992). More conclusive results concerning the XP-E deficiency should await the isolation and characterization of the XP-E correcting gene and protein.

XP-F patients have very mild clinical symptoms. NER in XP-F fibroblasts is very slow but long-lasting, and residual repair levels vary between 15 and 30%.

XP-G patients are more severely affected. As in group B and group D, also in group G patients have been described suffering from the combined XP/CS syndrome (Vermeulen, et al., 1992). The existence of 3 complementation groups that encompass patients suffering from 2 rare genetic syndromes, XP and CS, suggests both diseases to be - at the molecular level - closely related.

CS is, as XP, heterogeneous (see Table I): at least two complementation groups exist among the "classical" form (i.e. not in combination with XP) of the disease (Tanaka, et al., 1981; Lehmann, 1982). In sharp - and unresolved - contrast to XP, patients suffering from CS do not exhibit the pigmentation abnormalities and the increase in skin tumors typical for XP. A distinctive characteristic of CS fibroblasts is that after UV-irradiation they perform repair synthesis at normal rates, but are - in contrast to normal cells - unable to recover their RNA synthesis (Mayne and Lehmann, 1982). The precise molecular defect in this rare disorder has been mysterious for many years. A model postulated by Mayne and Lehmann in 1982 (Mayne and Lehmann, 1982) has only recently been confirmed experimentally, when Venema and colleagues pinpointed the CS defect to the sophisticated, preferential repair of CPD in active genes. Repair of inactive chromatin appears to be unaffected in CS cells (Venema, et al., 1990a). In cells from a patient in CS group A, thus far the smallest of the two complementation groups, strand specificity of repair of active genes is demonstrated to be abolished (Venema, 1991).

Among CS patients (without XP symptoms), three clinical subtypes have been described. The clinical groups mainly differ in the time of onset and severity of the clinical symptoms (Nance and Berry, 1992). Patients in which the clinical symptoms become apparent several months after birth are classified as CS I. These patients are reported to exhibit growth failure (after $\approx$1 year), neurodevelopmental and neurological dysfunction, cutaneous photosensitivity, progressive pigmentary retinopathy with or without cataracts and/or optic disk atrophy, sensorineural hearing loss, dental caries, and cachetic dwarfism. This is the largest group, with approximately 120 cases described. About 20 patients that had clinical characteristics of the disease already at birth have been classified as "severe" or "CS II". The patients have low birth weight, severely reduced growth and poor neurological

development. Their prognosis is generally much worse than that of the "CS I" patient (death occurring at an earlier age). Patients within the third subgroup, referred to as "mild CS", only have some of the clinical features (e.g. patients with normal intelligence or reproductive capacity) or the symptoms are mild and/or late in onset (for a detailed review, see Nance and Berry, 1992). This clinical heterogeneity does not coincide with the genetic heterogeneity as defined by the different complementation groups. Both group A and group B (Tanaka, et al., 1981; Lehmann, 1982), encompass patients classified as CS I, thereby group A also includes cells from a "mild" patient. Cells from patients clinically classified as "severe" have not been included in complementation analysis. It is not excluded, however, that some of these severe CS cases in fact also have an XP-like NER defect, and would display the clinical hallmarks of XP if they would become older.

NER-deficient rodent mutant cell lines

Although the existing human NER-deficient XP and CS cell lines are valuable and indispensable tools to enrich our knowledge of the mammalian NER mechanism, they have technical limitations such as transfectability (Hoeijmakers, et al., 1987). Furthermore, it is likely that certain genes involved in NER will remain unnoticed, since mutations within these genes might be incompatible with embryonic development. Therefore, considerable effort has been devoted to generate, and characterize a collection of UV-sensitive rodent mutant cell lines (Thompson, et al., 1988b; Zdzienicka, et al., 1988; Busch, et al., 1989; Hata, et al., 1991; Stefanini, et al., 1991). Three rodent cell lines - Chinese hamster ovary (CHO) cells, Chinese hamster cell line V79, and mouse lymphoma cell line L5178Y - have been used for this purpose. In CHO cells, partial hemizygosity of the genome has been demonstrated (Siciliano, et al., 1983; Thompson, et al., 1989), implying that only one gene copy has to be mutated (in these regions) for a recessive phenotype to become apparent.

An extensive collection of UV-sensitive rodent mutants is currently available. Among those, at least twelve complementation groups have been identified (see Table II). Representatives of the first five groups are very UV-sensitive, and are thought to be disturbed in the early steps of the NER pathway. Both group 1 and group 4 are, in addition to UV, extremely sensitive to mitomycin C (MMC), a crosslinking agent (Busch, et al., 1989). Groups 5 to 10 and 12 represent mutants with only a moderate UV-sensitivity (Thompson, et al., 1988b; Zdzienicka, et al., 1988; Busch, et al., 1989; Stefanini, et al., 1991 and D. Busch, personal communication). UV61, a group 6 mutant, has a partial deficiency in repair of CPD from the transcribed strand of active genes. Repair of the non-transcribed strand, and the genome overall is - characteristically for rodent cells - low (Lommel and Hanawalt, 1991). The mutant

Table II. Characteristics of rodent NER-deficient complementation groups

Complementation group	Representative mutant	Parental strain	Sensitivity[a] to UV	Sensitivity[a] to MMC	Incision deficiency	XP or CS equivalent	correcting gene cloned
1	UV20, 43-3B	CHO	++	+++	+	not available	*ERCC1*
2	UV5, V-H1	CHO, V79	++	+	+	XP-D	*ERCC2*
3	UV24, 27-1	CHO	++	+	+	XP-B (=CS-C)	*ERCC3*
4	UV41	CHO	++	+++	+	?	
5	UV135, Q31	CHO, L5178Y	+(+)	±	+	?	*ERCC5*
6	UV61, US46	CHO, L5178Y	+	+	partial	CS-B	*ERCC6*
7	VB11	V79	+	±	partial	?	
8	US31	L5178Y	+	+	?	?	
9	CHO7PV	CHO	+	+	partial	?	
10	CHO4PV	CHO	+	+	partial	?	
11	UVS1	CHO	+(+)	+	+[b]	?	
12	UV140[c]	CHO	+	±[d]	+[b]	?	

[a] +, 2-5x wt; ++, 5-10x wt; +++, >10x wt sensitivity
[b] A. Collins, personal communication
[c] D. Busch, personal communication
[d] M. Zdzienicka, personal communication

seems to be completely proficient in repair of (6-4) photoproducts, when determined in the genome overall (Thompson, et al., 1988a). Chapters III, IV, V, and VI describe the isolation and characterization of the human gene complementing the rodent group 6 repair deficiency.

UV-induced mutagenesis

Since the first evidence for the involvement of NER in the prevention of skin cancer -the absence of NER in the cancer-prone disorder XP (Cleaver, 1968) - much attention has been focussed on UV-induced mutagenesis. The two most prevalent UV-induced DNA lesions formed upon irradiation with 254 nm UV light are CPD and (6-4) photoproducts, in a ratio of about 3:1 (Mitchell and Nairn, 1989). "Overall genome" repair of the - more helix disturbing - (6-4) photoproduct in mammalian cells is faster and more complete than repair of CPD (Mitchell and Nairn, 1989 for a review); removal of (6-4) photoproducts from transcribed regions of the genome is only slightly more efficient (Link Jr., et al., 1991; Link Jr., et al., 1992). Although both photoproducts seem to add to the induction of mutations, the relative contribution of each of the two lesions is still not completely clear. The influence of CPD is suggested by several experiments. For example, the hypermutable (5x) rodent cell line UV61 (complementation group 6) is thought to have normal (6-4) photoproduct repair, and is partially deficient only in (preferential) repair of CPD (Thompson, et al., 1988a; Lommel and Hanawalt, 1991). Furthermore, removal of CPD from an irradiated SV40-based shuttle vector (by photoreactivation of the vector) before transfection to mammalian cells results in a reduction of the mutation frequency (Bourre, et al., 1989). There is, nevertheless, also convincing evidence suggesting an important role for (6-4) lesions. A partial revertant, RH1-26, has been isolated from rodent mutant cell line V-H1, complementation group 2 (Zdzienicka, et al., 1992). V-H1 is completely deficient in repair of CPD, and partially defective in (6-4) photoproduct repair (Mitchell, et al., 1989). The revertant has a similarly defective CPD removal (both from active and inactive chromatin), but regained normal (6-4) photoproduct repair. The 7-fold enhanced mutagenicity of V-H1, in contrast to the normal mutagenicity of RH1-26, suggests (6-4) photoproducts to be the main mutagenic lesion (Zdzienicka, et al., 1992). Also results obtained with the bacterium *E.coli* suggest (6-4) lesions to be the main - though not sole - mutagenic lesion (LeClerc, et al., 1991). Taken these data together, it seems very unlikely that mutagenesis is simply the result of only one of the two major photoproducts.

The type of mutations introduced by UV-light is diverse, with both transitions and transversions occurring. Predominance of GC to AT transitions has been reported, however, after UV-irradiation of *E.coli* (*lacI*; Schaaper, et al., 1987), yeast (the tRNA

gene *SUP4-o*; Armstrong and Kunz, 1990), and human cells (the *HPRT* gene; McGregor, et al., 1991). This observation is thought to result from the preferential insertion of an A residue opposite a non-instructive lesion (the "A-rule"; Strauss, 1991). Data concerning hamster cells are contradictory: in both the endogenous and exogenous, stably integrated *APRT* gene GC to AT transitions predominated (Drobetsky, et al., 1987; Drobetsky, et al., 1989), whereas in the *HPRT* gene of V79 and AA8, transversions were the predominant type of mutations present (Vrieling, et al., 1989; Menichini, et al., 1991). Possibly, the nature of mutations in different target genes is influenced by the regional pattern of replication and transcription.

Preferential repair of the transcribed strand of active genes has a profound influence on the strand specificity of UV-induced mutations. In normal human, hamster, and *E.coli* cells, mutations resulting from lesions in the nontranscribed strand of active genes prevail (Vrieling, et al., 1989; Koehler, et al., 1991; McGregor, et al., 1991; Menichini, et al., 1991; Vrieling, et al., 1991). In NER-deficient *E.coli* cells, this strand bias of mutations is strongly reduced, conform the absence of (preferential) repair (Koehler, et al., 1991). In four mammalian NER-deficient cell lines, XP12BE (XP group A) and hamster mutants VH-1, UV5 (rodent group 2), and UV24 (rodent group 3) mutations in the *HPRT* gene are mostly targeted by lesions in the transcribed strand (Vrieling, et al., 1989; McGregor, et al., 1991; Menichini, et al., 1991; Vrieling, et al., 1992). A hypothesis to explain this unexpected reversion of strand specificity has been put forward by Vrieling and coworkers (Vrieling, et al., 1989; Vrieling, et al., 1992). They suggested that the bias found in the absence of repair might be dependent on the fidelity of replication. Different DNA polymerases (polymerase α, δ, and ϵ) are involved in synthesis of the lagging and leading strand. The ability to bypass a lesion might be different for the different polymerases, a fact that could either be intrinsic to the polymerases itself or to the nature of replication of the leading (continuous replication) and lagging (discontinuous replication) strand. The presence of a sequence in the first *HPRT* intron, which is able to function as an autonomously replicating sequence in yeast and is associated with the nuclear matrix (Sykes, et al., 1988), suggests that during replication most of the *HPRT* transcribed strand functions as the template for the leading strand. But whatever might be the mechanism for the strand preference in situations where no repair has occurred, it is clear that if preferential repair actively operates, most mutations will reside from lesions in the nontranscribed strand.

Yeast genes involved in NER

The molecular mechanism of NER in eukaryotes is yet far from understood.

The isolation and characterization of genes and gene products involved will be an important clue to the understanding of this complex DNA repair process. For this purpose, the yeast *Saccharomyces cerevisiae*, amenable to both genetic and molecular manipulation, has proven to be a far-reaching tool. Mutant analysis has demonstrated a high degree of complexity among DNA repair processes in yeast: at least 30 different gene products are involved. The different mutants are divided in three epistasis groups (an epistasis group contains mutants disturbed in the same DNA repair pathway): *RAD3*, *RAD6*, and *RAD52* (Friedberg, 1988). The *RAD6* group is composed of mutants disturbed in so called postreplication repair, whereas the *RAD52* group comprises mutants defective in DNA damage related recombination processes. Mutants from the *RAD3* epistasis group are UV-sensitive and deficient in NER; at least 11 gene products are involved (Hoeijmakers and Bootsma, 1990; Friedberg, 1991 for recent reviews). Of these, RAD1, 2, 3, 4, 10, and 14 are believed to be essential for the incision step of NER. RAD16 is suggested to be important for repair of inactive chromatin. Removal of CPD from the two identical, but differentially transcribed mating type loci *MATα* and *HMLα* was compared within *rad16* mutants. Repair of the transcriptionally inactive *HMLα* locus was completely defective, whereas repair of the transcribed *MATα* locus did occur, although slower than in normal cells - a phenotype resembling that of XP-C cells (Terleth, et al., 1990).

A summary of the main properties of genes from the *RAD3* epistasis group that have been isolated and characterized thus far is shown in Table III. Of these, only the *RAD2* gene appears to be inducible by DNA damage (Madura and Prakash, 1986). Many of the genes - *RAD1* (Reynolds, et al., 1987), *RAD3* (Naumovski, et al., 1985; Reynolds, et al., 1985a), *RAD4* (Gietz and Prakash, 1988), *RAD7* (Perozzi and Prakash, 1986), and *SSL2* (other names: *ERCC3*^{Sc} or *RAD25*; Gulyas and Donahue, 1992; Park, et al., 1992) - encode proteins with acidic amino acid stretches, which could be involved in chromatin binding via electrostatic interactions with the basic histones.

Three of the gene products listed in Table III encode (presumed) helicases. The RAD3 gene product is demonstrated to exhibit (5' to 3') DNA helicase activity (Sung, et al., 1987). The second - in this case putative - DNA helicase is encoded by the *SSL2* or *ERCC3*^{Sc} gene (Gulyas and Donahue, 1992; Park, et al., 1992). Both proteins are of vital importance: "complete loss of function" mutations in either of those genes are not compatible with yeast life (Higgins, et al., 1983; Naumovski and Friedberg, 1983; Gulyas and Donahue, 1992). For RAD3, the helicase activity is indispensable for its repair function. Mutations eliminating the helicase function, however, do not affect viability (Sung, et al., 1988). The third, again presumed, helicase is encoded

Table III. Isolated and sequenced yeast genes involved in nucleotide excision repair

Gene	Chromosomal localization	Protein size (aa)	Human homolog	Protein characteristics
RAD1	XVI	1100	?	acidic C-terminus, involved in recombination, can interact with RAD10
RAD2	VII	1031	?	transciption inducible by UV
RAD3	V	778	*ERCC2*	DNA-binding, 5'→3' DNA.DNA and DNA.RNA helicase, acidic C-terminus, vital function
RAD4	V	754	?	helix-turn-helix motif present, acidic C-terminus
RAD7	X	565	?	acidic stretches
RAD10	XIII	210	*ERCC1*	involved in recombination, can interact with RAD1, binds ss-DNA, renaturation of ss-DNA
RAD14	?	246	*XPAC*	Zn^{2+}-finger motifs, no vital function
RAD16	II	790	?	member of a subfamily of postulated helicases (ERCC6 also belongs to this family), Zn^{2+}-finger motifs
SSL2	IX	843	*ERCC3*	also designated *ERCC3*^{Sc} or *RAD25*, helicase signatures present, acidic stretches, a mutant SSL2 protein (*SSL2-1*) is able to promote translation past a strong hairpin structure, vital function

For references, see text.

by the *RAD16* gene (Bang, et al., 1992; Mannhaupt, et al., 1992; Schild, et al., 1992). This postulated helicase is a member of a rapidly expanding subfamily of putative helicases, which share a high degree of homology over the entire region encoding the helicase signatures. Other members of this family are two proteins of unknown function (STH1 and FUN30), transcription regulators (SNF2, MOT1, brm), a protein involved in preservation of chromosome stability (lodestar), and three proteins involved in three important DNA repair pathways (RAD54 in recombination repair, RAD5 in postreplication repair and ERCC6 in preferential NER) (Bang, et al., 1992; Clark, et al., 1992; Davis, et al., 1992; Johnson, et al., 1992; Laurent, et al., 1992, and Chapter V).

Several proteins contain putative DNA binding domains within their predicted amino acid sequence. Four of the *RAD3* group genes, *RAD3*, *4* (Gietz and Prakash, 1988; Cuoto and Friedberg, 1989), *10* (Reynolds, et al., 1985b) and *SSL2* (*ERCC3*^{Sc}), encode peptides with homology to the "helix-turn-helix" DNA binding motif; RAD14 (Bankmann, et al., 1992) and RAD16 (Bang, et al., 1992; Mannhaupt, et al., 1992; Schild, et al., 1992) contain zinc-finger motifs. So far, experimental evidence actually demonstrating DNA binding has only been obtained for RAD3 (Sung, et al., 1987) and RAD10 (Sung, et al., 1992). The RAD1 and RAD10 proteins appear to have, besides their role in NER, a role in mitotic recombination; presumably they are both involved in the same recombination process (Schiestl and Prakash, 1988; Schiestl and Prakash, 1990). The two proteins appear to be complexed with each other, both *in vivo* and *in vitro* (Bailly, et al., 1992; Bardwell, et al., 1992).

Isolation of human genes involved in NER

The most generally used procedure for isolation of human DNA repair genes is transfection of genomic DNA from normal cells to repair-deficient mutants, followed by selection of repair-competent transformants. The correcting gene can subsequently be isolated via standard techniques. This method has resulted in the isolation of several human genes capable of correcting different complementation groups of the *in vitro* mutagenized, UV-sensitive, rodent cell lines (discussed below; see Table IV). Genomic DNA transfections to cell lines from patients with DNA repair syndromes (i.e. XP, CS) have so far - with one notable exception (Tanaka, et al., 1989) - proven unsuccessful. The difficulty to isolate repair-competent transformants after transfection of repair-defective human cells is due to the low amount of DNA that will stably integrate in the genome of these cells. Compared to rodent cells, approximately 30 - $\geq$100 fold less DNA is integrated. This raises the number of cells required to be transfected to generate one genomic transformant

Table IV. Isolated human genes involved in nucleotide excision repair

Gene	Chromosomal localization	Gene size (kb)	Protein size (aa)	Yeast homolog	Protein characteristics[a)]
ERCC1	19q13.2	~17	297	*RAD10*	DNA binding?, involved in recombination?, homology to parts of UvrA and C in C-terminus
ERCC2	19q13.2	~20	760	*RAD3*	helicase?, vital function?
ERCC3	2q21	~45	782	*SSL2*, *RAD25*, *ERCC3*^{Sc}	helicase?, vital function?
ERCC5	13q32-33	~32	?	?	
ERCC6	10q11-21	~85	1493	?	member of a subfamily of postulated helicases, presumably not essential for viability
XPAC	9q34	~25	273	*RAD14*	Zn^{2+} finger, binds to ss and UV-irradiated ds DNA
XPCC	?	?	823	?	(limited) homology to RAD4 in C-terminus

For references see text.

[a)] ? denotes a putative protein characteristic deduced from the amino acid sequence or from the function of a yeast homolog; direct proof is still lacking

containing a specific gene to extremely high levels (depending on gene length to more than 10^9 cells (Hoeijmakers, et al., 1987)).

DNA-mediated gene transfer to human NER-deficient cell lines

The one example of successful isolation of a NER gene via large scale genomic DNA transfection to a XP cell line is the XPAC gene (Tanaka, et al., 1989). This gene, localized on chromosome 9, is approximately 25 kb in length, and encodes a protein of 273 amino acids (aa) (Tanaka, et al., 1990). The XPAC protein has been (partially) purified from HeLa cells and calf thymus, using either microinjection or an *in vitro* repair assay as a test for activity (Robins, et al., 1991; Sugano, et al., 1991; Eker, et al., 1992). The encoded protein has a predicted molecular weight of 31 kD, but has been found to migrate with an apparent molecular weight of 40 - 45 Kd (SDS-PAGE) (Miura, et al., 1991; Robins, et al., 1991; Sugano, et al., 1991; Eker, et al., 1992). The sequence predicts the presence of a zinc-finger, suggestive of DNA-binding (Tanaka, et al., 1990). *In vitro* studies with the purified XPAC protein confirm the DNA-binding capacity (Robins, et al., 1991; Eker, et al., 1992). Mutation analysis among XP-A patients revealed a founder effect in Japan: of the 21 patients analyzed, 16 were homozygous and 4 were heterozygous for the G to C transversion at the 3' splice acceptor site of intron 3 (Satokata, et al., 1990). The clinical heterogeneity among other patients seems to be correlated with the occurrence of different mutations (Satokata, et al., 1992b). A XP-A revertant cell line has been isolated (Cleaver, et al., 1987), in which a nonsense codon at aa 207 (originating from an arginine) is altered such that it encodes again an aa (Jones, et al., 1992; Satokata, et al., 1992a). This revertant is as UV-resistant as normal cells. Repair of (6-4) photoproducts is restored to wild type levels, but dimer removal has remained as deficient as in the original XP-A (Cleaver, et al., 1987), suggesting the XPAC protein to function - directly or indirectly - in damage recognition.

For XP cells, an alternative gene transfer strategy, the transfection of cDNA libraries, has proven successful too (Peterson and Legerski, 1991). The XP-C correcting gene has recently been isolated following this method. The gene, which seems to be specifically involved in repair of inactive chromatin (Venema, et al., 1991), encodes a protein of 823 aa. In its C-terminus, the protein exhibits (limited) homology to the yeast NER protein RAD4 (Legerski and Peterson, 1992). Further research on the XPCC gene might shed more light on its function in repair of inactive chromatin.

DNA-mediated gene transfer to NER-deficient rodent cells

DNA-mediated gene transfer to rodent mutants has - so far - led to the isolation

of five human genes correcting complementation groups 1, 2, 3, 5, and 6. The first human NER gene isolated was the *ERCC1* (*E*xcision *R*epair *C*ross *C*omplementing rodent repair deficiency) gene, correcting cells from rodent complementation group 1 (Westerveld, et al., 1984). The gene seems not to be involved in one of the XP or CS complementation groups (Van Duin, et al., 1989). *ERCC1*, covering a region of 15 - 17 kb on human chromosome 19q13.2 (Van Duin, et al., 1987; Mohrenweiser, et al., 1989; Smeets, et al., 1990), encodes a protein of 297 aa. The predicted aa sequence contains a putative nuclear location signal (NLS) and a postulated DNA binding domain (of the helix-turn-helix type). It shows a significant "overall" aa homology to the gene product of the yeast repair gene RAD10, suggesting both proteins might have similar functions in the excision repair process (Van Duin, et al., 1986). Currently, studies are in progress to assess whether ERCC1 is involved in recombination, as has been demonstrated for the RAD10 protein (Schiestl and Prakash, 1990). The participation of ERCC1 in recombinational repair might explain the extreme sensitivity of rodent group 1 mutants for the crosslinking agent MMC. To completely repair the damage caused by an interstrand crosslink NER alone will probably be insufficient, and additional recombination repair necessary. A role of ERCC1 in recombination is thus an attractive hypothesis. The predicted ERCC1 aa sequence extends at the C-terminus beyond the RAD10 sequence (RAD10: 210 aa, *ERCC1*: 297 aa), where it exhibits homology to parts of the bacterial proteins UvrA and UvrC (Van Duin, et al., 1988). Mutational analysis has shown the C-terminus (homology to the UvrC terminus) of utmost importance for the repair function of the *ERCC1* protein (Van Duin, et al., 1988, P. van der Spek and J. van den Berg, personal communication). In *E.coli*, the UvrC C-terminal half was demonstrated to be important for the incision step of NER (Lin and Sancar, 1991). The UvrC protein is suggested to introduce an incision 5' to a DNA lesion (Lin and Sancar, 1992). Whether the ERCC1 protein has a function in mammalian repair that is comparable to that of uvrC in prokaryotic NER (incision) remains to be elucidated. The four residues demonstrated to be essential for incision by UvrC, however, are outside the region of homology with ERCC1. In contrast to the ERCC1 C-terminus, the - much less conserved - N-terminus is relatively unimportant for its repair function, since deletions up to 93 aa do not affect the functionality of the protein in repair of MMC- and UV-induced lesions (Van Duin, et al., 1988, J. van den Berg, personal communication).

ERCC2, the human gene correcting cells from rodent group 2, is like *ERCC1* located at chromosome 19q13.2, within 250 kb from *ERCC1* (Mohrenweiser, et al., 1989; Smeets, et al., 1990). The gene (about 20 kb in length) encodes a protein of 760

aa (Weber, et al., 1990). Within the aa sequence, seven consecutive domains are present, which are conserved in two families of determined or presumed DNA and RNA helicases. The gene is strongly homologous (approximately 50% identity) to the yeast RAD3 gene product (Weber, et al., 1990), which is demonstrated to have both 5' to 3' DNA-DNA (Sung, et al., 1987) and 5' to 3' DNA-RNA helicase activity (Bailly, et al., 1991; Naegeli, et al., 1992). Recently, the *ERCC2* gene was shown to correct the repair defect of cells from XP complementation group D (Flejter, et al., 1992). Mutational analysis on the *ERCC2* gene might be able to shed light on the molecular background of the diverse phenotypes of patients in this complementation group.

A third NER gene, *ERCC3*, corrects the UV-sensitivity of rodent group 3 cells. The gene is located on chromosome 2q21 (Weeda, et al., 1991), and spans a region of approximately 45 kb. The predicted protein is similar in size to ERCC2: 782 aa. It contains a putative NLS, several acidic regions that might be involved in chromatin binding, a postulated DNA binding domain of the helix-turn-helix type, and as in the ERCC2 protein, the seven consecutive "helicase domains" (Weeda, et al., 1990b). This makes ERCC3 the second - putative - helicase functioning in mammalian NER. A yeast homolog has been isolated, designated *SSL2* (also *RAD25* or *ERCC3*^{Sc}), sharing about 50% identity at the aa level (Gulyas and Donahue, 1992; Park, et al., 1992). The *ERCC3* gene appears to be involved in XP-B, a rare XP complementation group containing only patients that show the clinical symptoms of both XP and CS (Weeda, et al., 1990b). If *ERCC3*, as its yeast homolog, has a vital function (Gulyas and Donahue, 1992; Park, et al., 1992), this might explain the sporadic occurrence of patients in the XP-B group.

ERCC5 has recently been isolated by virtue of its ability to correct rodent cells from complementation group 5. The gene is approximately 32 kb, and is assigned to chromosome 13q33 (Mudgett and MacInnes, 1990). Further research concerning *ERCC5* is in progress.

The experimental work described in this thesis (Chapters III, IV, V, and VI) focusses on the isolation and characterization of the human NER gene *ERCC6*. The gene has been isolated after transfection of genomic DNA to a rodent mutant cell line of cg 6. The mutant, UV61, is (partially) deficient in the removal of CPD, whereas repair of (6-4) photoproducts is seemingly unaffected (Thompson, et al., 1988a). Lommel and Hanawalt recently demonstrated that removal of CPD from the transcribed strand of an active gene was reduced (Lommel and Hanawalt, 1991). The *ERCC6* gene is located on human chromosome 10q11-21, and encodes a predicted helicase of 1493 aa. In addition, the predicted aa sequence contains two possible NLS

and an acidic region that could, as postulated for the ERCC3 protein, be involved in chromatin binding. The gene belongs to a gene family consisting of several presumed helicases, demonstrating an extremely high degree of homology between the helicase regions (the yeast NER protein RAD16 also is a member of this family, see above). *ERCC6* appears to be involved in the human syndrome CS-B (so far, the most common form of the disease), and thus in preferential repair of active genes. Mutations found in the *ERCC6* alleles of a CS-B patient suggest that the gene product is specific for this subpathway of excision repair and has - possibly in contrast to ERCC2 and ERCC3 - no essential function (see Chapter V).

References

Armstrong, J. D., and Kunz, B. A. (1990). Site and strand specificity of UVB mutagenesis in the *SUP4-o* gene of yeast. Proc. Natl. Acad. Sci. USA *87*, 9005-9009.

Bailly, V., Sommers, C. H., Sung, P., Prakash, L., and Prakash, S. (1992). Specific complex formation between proteins encoded by the yeast DNA repair and recombination genes *RAD1* and *RAD10*. Proc. Natl. Acad. Sci. USA *89*, 8273-8277.

Bailly, V., Sung, P., Prakash, L., and Prakash, S. (1991). DNA-RNA helicase activity of RAD3 protein of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA *88*, 9712-9716.

Bang, D. D., Verhage, R., Goosen, N., Brouwer, J., and Putte, P. v. d. (1992). Molecular cloning of *RAD16*, a gene involved in differential repair in *Saccharomyces cerevisiae*. Nucl. Acids Res. *20*, 3925-3931.

Bankmann, M., Prakash, L., and Prakash, S. (1992). Yeast *RAD14* and human xeroderma pigmentosum group A DNA repair genes encode homologous proteins. Nature *355*, 555-558.

Bardwell, L., Cooper, A. J., and Friedberg, E. C. (1992). Stable and specific association between the yeast recombination and DNA repair proteins RAD1 and RAD10 *in vitro*. Mol. Cell. Biol. *12*, 3041-3049.

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.

Bourre, F., Benoit, A., and Sarasin, A. (1989). Respective roles of pyrimidine dimer and pyrimidine (6-4) pyrimidone photoproducts in UV mutagenesis of Simian virus 40 DNA in mammalian cells. J. Virol. *63*, 4520-4524.

Broughton, B. C., Lehmann, A. R., Harcourt, S. A., Arlett, C. F., Sarasin, A., Kleijer, W. J., Beemer, F. A., Nairn, R., and Mitchell, D. L. (1990). Relationship between pyrimidine dimers, 6-4

photoproducts, repair synthesis and cell survival: studies using cells from patients with trichothiodystrophy. Mutat. Res. *235*, 33-40.

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.

Chu, G., and Chang, E. (1988). Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. Science *242*, 564-567.

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. (1968). Defective repair replication in xeroderma pigmentosum. Nature *218*, 652-656.

Cleaver, J. E., Cortes, F., Lutze, L. H., Morgan, W. F., Player, A. N., and Mitchell, D. L. (1987). Unique DNA repair properties of a xeroderma pigmentosum revertant. Mol. Cell. Biol. *7*, 3353-3357.

Cleaver, J. E., and Kraemer, K. H. (1989). Xeroderma pigmentosum. In: The metabolic basis for inherited disease. Vol. II, Scriver, C. R., Beaudet, A. L., Sly, W. S. and Valle, D., eds. (New York: McGraw-Hill Book Co.), 2949-2971.

Cuoto, L. B., and Friedberg, E. C. (1989). Nucleotide sequence of the wild type *RAD4* gene of *Saccharomyces cerevisiae* and characterization of mutant *rad4* alleles. J. Bacteriol. *171*, 1862-1869.

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 Jonge, A. J., Vermeulen, W., Keijzer, W., Hoeijmakers, J. H. J., and Bootsma, D. (1985). Microinjection of Micrococcus luteus UV-endonuclease restores UV-induced unscheduled DNA synthesis in cells of 9 xeroderma pigmentosum complementation groups. Mutat. Res. *150*, 99-105.

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

Drobetsky, E. A., Grosovski, A. J., and Glickman, B. W. (1987). The specificity of UV-induced mutations at an endogenous locus in mammalian cells. Proc. Natl. Acad. Sci. USA *84*, 9103-9107.

Drobetsky, E. A., Grosovski, A. J., Skandalis, A., and Glickman, B. W. (1989). Perspectives on

UV light mutagenesis: investigation of the CHO *APRT* gene carried on a retroviral shuttle vector. Som. Cell Mol. Genet. *15*, 401-409.

Eker, A. P. M., Vermeulen, W., Miura, N., Tanaka, K., Jaspers, N. G. J., Hoeijmakers, J. H. J., and Bootsma, D. (1992). Xeroderma pigmentosum group A correcting protein from calf thymus. Mutat. Res. *274*, 211-224.

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).

Friedberg, E. C. (1988). Deoxyribonucleic acid repair in the yeast *Saccharomyces cerevisiae*. Microbiol. Rev. *52*, 70-102.

Friedberg, E. C. (1991). Eukaryotic DNA repair: glimpses through the yeast *Saccharomyces cerevisiae*. Bioessays *13*, 295-302.

George, D. L., and Witkin, E. M. (1974). Slow excision in an *mfd* mutant of *Escherichia coli* B/r. Mol. Gen. Genet. *133*, 283-291.

Gietz, R. D., and Prakash, S. (1988). Cloning and nucleotide sequence analysis of the *Saccharomyces cerevisiae RAD4* gene required for excision repair of UV-damaged DNA. Gene *74*, 535-541.

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

Gulyas, K. D., and Donahue, T. F. (1992). *SSL2*, a suppressor of a stem-loop mutation in the *HIS4* leader encodes the yeast homologue of human *ERCC3*. Cell *69*, 1031-1042.

Hata, H., Numata, M., Tohda, H., Yasui, A., and Oikawa, A. (1991). Isolation of two chloroethylnitrosourea-sensitive Chinese hamster cell lines. Cancer Res. *51*, 195-198.

Hayakawa, H., Ishizaki, K., Inoue, M., Yagi, T., Sekiguchi, M., and Takebe, H. (1981). Repair of ultraviolet radiation damage in xeroderma pigmentosum cells belonging to complementation group F. Mutat. Res. *80*, 381-388.

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.

Hoeijmakers, J. H. J., and Bootsma, D. (1990). Molecular genetics of eukaryotic DNA excision repair. Cancer Cells 2, 311-320.

Hoeijmakers, J. H. J., Odijk, H., and Westerveld, A. (1987). Differences between rodent and human cell lines in the amount of integrated DNA after transfection. Exp. Cell Res. *169*, 111-9.

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.

Johnson, R. T., and Squires, S. (1992). The XPD complementation group. Insights into xeroderma pigmentosum, Cockayne's syndrome and trichothiodystrophy. Mutat. Res. *273*, 97-118.

Jones, C. J., Cleaver, J. E., and Wood, R. D. (1992). Repair of damaged DNA by extracts from a xeroderma pigmentosum complementation group A revertant and expression of a protein absent in its parental cell line. Nucl. Acids Res. *20*, 991-995.

Kantor, G. J., Barsalou, L. S., and Hanawalt, P. C. (1990). Selective repair of specific chromatin domains in UV-irradiated cells from xeroderma pigmentosum complementation group C. Mutat. Res. *235*, 171-80.

Keeney, S., Wein, H., and Linn, S. (1992). Biochemical heterogeneity in xeroderma pigmentosum complementation group E. Mutat. Res. *273*, 49-56.

Koehler, D. R., Awadallah, S. S., and Glickman, B. W. (1991). Sites of preferential induction of cyclobutane pyrimidine dimers in the non-transcribed strand of *lacI* correspond with sites of UV-induced mutation in *Escherichia coli*. J. Biol. Chem. *266*, 11766-11773.

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.

LeClerc, J. E., Borden, A., and Lawrence, C. W. (1991). The thymine-thymine pyrimidine-pyrimidone(6-4) ultraviolet light photoproduct is highly mutagenic and specifically induces 3' thymine-to-cytosine transitions in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *88*, 9685-9689.

Legerski, R., and Peterson, C. (1992). Expression cloning of a human DNA repair gene involved in xeroderma pigmentosum group C. Nature *359*, 70-73.

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.

Lehmann, A. R. (1989). Trichothiodystrophy and the relationship between DNA repair and cancer. Bioessays *11*, 168-70.

Lehmann, A. R., Kirk-Bell, S., Arlett, C. F., Paterson, M. C., Lohman, P. H. M., de Weerd-Kastelein, E. A., and Bootsma, D. (1975). Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. Proc. Natl. Acad. Sci. USA *72*, 219-223.

Lin, J. J., Phillips, A. M., Hearst, J. E., and Sancar, A. (1992). Active site of (A)BC excinuclease. II. Binding, bending, and catalysis mutants of UvrB reveal a direct role in 3' and an indirect role in 5' incision. J. Biol. Chem. *267*, 17693-17700.

Lin, J. J., and Sancar, A. (1991). The C-terminal half of UvrC protein is sufficient to reconstitute (A)BC excinuclease. Proc. Natl. Acad. Sci. USA *88*, 6824-8.

Lin, J. J., and Sancar, A. (1992). Active site of (A)BC excinuclease. I. Evidence for 5' incision by UvrC through a catalytic site involving Asp^{399}, Asp^{438}, Asp^{466}, and His538 residues. J. Biol. Chem. *267*, 17688-17692.

Link Jr., C. J., Burt, R. K., and Bohr, V. A. (1991). Gene-specific repair of DNA damage induced by UV irradiation and cancer chemotherapeutics. Cancer Cells *3*, 427-436.

Link Jr., C. J., Mitchell, D. L., Nairn, R. S., and Bohr, V. A. (1992). Preferential and strand-specific DNA repair of (6-4) photoproducts detected by a photochemical method in the hamster *DHFR* gene. Carcinogen. *13*, 1975-1980.

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.

Madura, K., and Prakash, S. (1986). Nucleotide sequence, transcript mapping, and regulation of the *RAD2* gene of *Saccharomyces cerevisiae*. J. Bacteriol. *166*, 914-923.

Mannhaupt, G., Stucka, R., Ehnle, S., Vetter, I., and Feldmann, H. (1992). Molecular analysis of yeast chromosome II between *CMD1* and *LYS2:* the excision repair gene *RAD16* located in this region belongs to a novel group of finger proteins. Yeast *8*, 397-408.

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 Research *42*, 1473-1478.

McGregor, W. G., Chen, R. H., Lukash, L., Maher, V. M., and McCormick, J. J. (1991). Cell cycle-dependent strand bias for UV-induced mutations in the transcribed strand of excision repair-proficient human fibroblasts but not in repair-deficient cells. Mol. Cell. Biol. *11*, 1927-1934.

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., 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., 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.

Menichini, P., Vrieling, H., and van Zeeland, A. A. (1991). Strand-specific mutation spectra in repair-proficient and repair-deficient hamster cells. Mutat. Res. *251*, 143-155.

Mitchell, D. L., and Nairn, R. S. (1989). The biology of the (6-4) photoproduct. Photochem. Photobiol. *49*, 805-819.

Mitchell, D. L., Zdzienicka, M. Z., van Zeeland, A. A., and Nairn, R. S. (1989). Intermediate (6-4) photoproduct repair in Chinese hamster V79 mutant V-H1 correlates with intermediate levels of DNA incision and repair replication. Mutat. Res. *226*, 43-47.

Miura, N., Miyamoto, I., Asahina, H., Satokata, I., Tanaka, K., and Okada, Y. (1991). Identification and characterization of XPAC protein, the gene product of the human *XPAC* (xeroderma pigmentosum group A complementing) gene. J. Biol. Chem. *266*, 19786-19789.

Mohrenweiser, H. W., Carrano, A. V., Fertitta, A., Perry, B., Thompson, L. H., Tucker, J. D., and Weber, C. A. (1989). Refined mapping of the three DNA repair genes *ERCC1, ERCC2,* and *XRCC1* on chromosome 19. Cytogenet. Cell Genet. *52*, 11-14.

Mudgett, J. S., and MacInnes, M. A. (1990). Isolation of the functional human excision repair gene *ERCC5* by intercosmid recombination. Genomics *8*, 623-633.

Naegeli, H., Bardwell, L., Harosh, I., and Friedberg, E. C. (1992). Substrate specificity of the Rad3 ATPase/DNA helicase of *Saccharomyces cerevisiae* and binding of Rad3 protein to nucleic acids. J. Biol. Chem. *267*, 7839-7844.

Nance, M. A., and Berry, S. A. (1992). Cockayne syndrome: Review of 140 cases. Am. J. Med.

Genet. *42*, 68-84.

Naumovski, L., Chu, G., Berg, P., and Friedberg, E. C. (1985). RAD3 gene of Saccharomyces cerevisiae: nucleotide sequence of wild-type and mutant alleles, transcript mapping, and aspects of gene regulation. Mol. Cell. Biol. *5*, 17-26.

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.

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.

Park, E., Guzder, S., Koken, M. H. M., Jaspers-Dekker, I., Weeda, G., Hoeijmakers, J. H. J., Prakash, S., and Prakash, L. (1992). *RAD25*, a yeast homolog of human xeroderma pigmentosum group B DNA repair gene is essential for viability. Proc. Natl. Acad. Sci. USA in press.

Perozzi, G., and Prakash, S. (1986). *RAD7* of *Saccharomyces cerevisiae*: Transcripts, nucleotide sequence analysis and functional relationship between the RAD7 and RAD23 gene products. Mol. Cell. Biol. *6*, 1497-1507.

Peterson, C., and Legerski, R. (1991). High-frequency transformation of human repair-deficient cell lines by an Epstein-Barr virus-based CDNA expression vector. Gene *107*, 279-284.

Reynolds, P., Higgins, D. R., Prakash, L., and Prakash, S. (1985a). The nucleotide sequence of the *RAD3* gene of *Saccharomyces cerevisiae*: A potential adenine nucleotide binding amino acid sequence and a nonessential carboxyl terminal domain. Nucl. Acids Res. *13*, 2357-2372.

Reynolds, P., Prakash, L., Dumais, D., Perozzi, G., and Prakash, S. (1985b). Nucleotide sequence of the *RAD10* gene of *Saccharomyces cerevisiae*. EMBO J. *4*, 3549-3552.

Reynolds, P., Prakash, L., and Prakash, S. (1987). Nucleotide sequence and functional analysis of the *RAD1* gene of *Saccharomyces cerevisiae*. Mol. Cell. Biol. *7*, 1012-1020.

Robbins, J. H., Kraemer, K. H., Lutzner, M. A., Festoff, B. W., and Coon, H. G. (1974). Xeroderma pigmentosum. An inherited disease with sun sensitivity, multiple cutaneous neoplasms and abnormal repair. Ann. Intern. Med. *80*, 221-248.

Robins, P., Jones, C. J., Biggerstaff, M., Lindahl, T., and Wood, R. D. (1991). Complementation

of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA. EMBO J. *10*, 3913-3921.

Satokata, I., Tanaka, K., Miura, N., Myamoto, I., Satoh, Y., Kondo, S., and Okada, Y. (1990). Characterization of a splicing mutation in group A xeroderma pigmentosum. Proc. Natl. Acad. Sci. USA *87*, 9908-9912.

Satokata, I., Tanaka, K., Miura, N., Narita, M., Mimaki, T., Satoh, Y., Kondo, S., and Okada, Y. (1992a). Three nonsense mutations responsible for group A xeroderma pigmentosum. Mutat. Res. *273*, 193-202.

Satokata, I., Tanaka, K., Yuba, S., and Okada, Y. (1992b). Identification of splicing mutations of the last nucleotides of exons, a nonsense mutation, and a missense mutation of the *XPAC* gene as causes of group A xeroderma pigmentosum. Mutat. Res. *273*, 203-212.

Schaaper, R. M., Dunn, R. L., and Glickman, B. W. (1987). Mechanisms of ultraviolet-induced mutation. Mutational spectra in the *Escherichia coli lacI* gene for a wild-type and an excision-repair-deficient strain. J. Mol. Biol. *198*, 187-202.

Schiestl, R. H., and Prakash, S. (1988). *RAD1*, an excision repair gene of *Saccharomyces cerevisiae*, is also involved in recombination. Mol. Cell. Biol. *8*, 3619-3626.

Schiestl, R. H., and Prakash, S. (1990). *RAD10*, an excision repair gene of *Saccharomyces cerevisiae*, is involved in the *RAD1* pathway of mitotic recombination. Mol. Cell. Biol. *10*, 2485-2491.

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. (1990). Transcription preferentially inhibits nucleotide excision repair of the template DNA strand in vitro. J. Biol. Chem. *265*, 21330-21336.

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.

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.

Siciliano, M. J., Stallings, R. L., Adair, G. M., Humphrey, R. M., and Siciliano, J. (1983). Provisional assignment of *TPI*, *GPI*, and *PEPD* to Chinese hamster autosomes 8 and 9: a cytogenetic basis for functional haploidy of an autosomal linkage group in CHO cells. Cytogenet. Cell Genet. *35*, 15-20.

Smeets, H., Bachinski, L., Coerwinkel, M., Schepens, J., Hoeijmakers, J. H. J., van Duin, M., Grzeschik, K. H., Weber, C. A., de Jong, P., Siciliano, M. J., and Wieringa, B. W. (1990). A long-range restriction map of the human chromosome 19q13 region: close physical linkage between CKMM and the ERCC1 and ERCC2 genes. Am. J. Hum. Genet. *46*, 492-501.

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

Stefanini, M., Collins, A. R., Riboni, R., Klaude, M., Botta, E., Mitchell, D. L., and Nuzzo, F. (1991). Novel Chinese hamster ultraviolet-sensitive mutants for excision repair form complementation groups 9 and 10. Cancer Res *51*, 3965-3971.

Stefanini, M., Giliani, S., Nardo, T., Marinoni, S., Nazzaro, V., Rizzo, R., and Trevisan, G. (1992). DNA repair investigations in nine Italian patients affected by trichothiodystrophy. Mutat. Res. *273*, 119-125.

Stefanini, M., Lagomarsini, P., Arlett, C. F., Marinoni, S., Borrone, C., Crovato, F., Trevisan, G., Cordone, G., and Nuzzo, F. (1986). Xeroderma pigmentosum (complementation group D) mutation is present in patients affected by trichothiodystrophy with photosensitivity. Hum. Genet. *74*, 107-112.

Strauss, B. S. (1991). The "A rule" of mutagen specificity: a consequence of DNA polymerase bypass of non-instructional lesions? BioEssays *13*, 79-84.

Sugano, T., Uchida, T., and Yamaizumi, M. (1991). Identification of a specific protein factor defective in group A xeroderma pigmentosum cells. J. Biochem. *110*, 667-664.

Sung, P., Higgins, D., Prakash, L., and Prakash, S. (1988). Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. EMBO J. *7*, 3263-3269.

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, L., and Prakash, S. (1992). Renaturation of DNA catalysed by yeast DNA repair and recombination protein RAD10. Nature *355*, 743-745.

Sykes, R. C., Lin, D., Hwang, S. J., Framson, P. E., and Chinault, A. C. (1988). Yeast ARS function and nuclear matrix association coincide in a short sequence from the human *HPRT* locus.

Mol. Gen. Genet. *212*, 301-309.

Tanaka, K., Hayakawa, H., Sekiguchi, M., and Okada, Y. (1977). Specific action of T4 endonuclease V on damaged DNA in xeroderma pigmentosum cells *in vivo*. Proc. Natl. Acad. Sci. USA *74*, 2958-2962.

Tanaka, K., Kawai, K., Kumahara, Y., Ikenaga, M., and Okada, Y. (1981). Genetic complementation groups in Cockayne syndrome. Somat. 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.

Tanaka, K., Satokata, I., Ogita, Z., Uchida, T., and Okada, Y. (1989). Molecular cloning of a mouse DNA repair gene that complements the defect of group A xeroderma pigmentosum. Proc. Natl. Acad. Sci. USA *86*, 5512-5516.

Tanaka, K., Sekiguchi, M., and Okada, Y. (1975). Restoration of ultraviolet-induced unscheduled DNA synthesis of xeroderma pigmentosum cells by the concomitant treatment with bacteriophage T4 endonuclease V and HVJ (Sendai virus). Proc. Natl. Acad. Sci. USA *72*, 4071-4075.

Terleth, C., Schenk, P., Poot, R., Brouwer, J., and Putte, P. v. d. (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., Bachinski, L. L., Stallings, R. L., Dolf, G., Weber, C. A., Westerveld, A., and Siciliano, M. J. (1989). Complementation of repair gene mutations on the hemizygous chromosome 9 in CHO: a third repair gene on human chromosome 19. Genomics *5*, 670-679.

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. Somat. Cell. Mol. Genet. *14*, 605-612.

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., Koken, M. H. M., van den Tol, J., ten Dijke, P., Odijk, H., Westerveld, A., Bootsma, D., and Hoeijmakers, J. H. J. (1987). Genomic characterization of the human DNA

excision repair gene ERCC-1. Nucl. Acids Res. *15*, 9195-213.

Van Duin, M., van den Tol, J., Warmerdam, P., Odijk, H., Meijer, D., Westerveld, A., Bootsma, D., and Hoeijmakers, J. H. J. (1988). Evolution and mutagenesis of the mammalian excision repair gene ERCC-1. Nucl. Acids Res. *16*, 5305-5322.

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. Ph. D. Thesis. State University of Leiden.

Venema, J., Bartosova, Z., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. F. (1992). Transcription affects the rate but not the extent of repair of cyclobutane pyrimidine dimers in the human adenosine deaminase gene. J. Biol. Chem. *267*, 8852-8856.

Venema, J., Mullenders, L. H. F., 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., Karcagi, V., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. F. (1991). Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. Mol. Cell. Biol. *11*, 4128-34.

Venema, J., Van Hoffen, A., Natarajan, A. T., Van Zeeland, A. A., and Mullenders, L. H. F. (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., Jaeken, J., Jaspers, N. G. J., Bootsma, D., and Hoeijmakers, J. H. J. (1992). Xeroderma pigmentosum complementation group G associated with Cockayne's syndrome. submitted.

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., van Rooijen, M. L., Groen, N. A., Zdzienicka, M. Z., Simons, J. W. I. M., Lohman,

P. H. M., and van Zeeland, A. A. (1989). DNA strand specificity for UV-induced mutations in mammalian cells. Mol. Cell. Biol. *9*, 1277-1283.

Vrieling, H., Venema, J., Van Rooyen, M. L., Van Hoffen, A., Menichini, P., Zdzienicka, M. Z., Simons, J. W. I. M., Mullenders, L. H. F., 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.

Vrieling, H., Zhang, L. H., van Zeeland, A. A., and Zdzienicka, M. Z. (1992). UV-induced *hprt* mutations in a UV-sensitive hamster cell line from complementation group 3 are biased towards the transcribed strand. Mutat. Res. *274*, 147-155.

Weber, C. A., Salazar, E. P., Stewart, S. A., and Thompson, L. H. (1988). Molecular cloning and biological characterization of a human gene, *ERCC2*, that corrects the nucleotide excision repair defect in CHO UV5 cells. Mol. Cell. Biol. *8*, 1137-1146.

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.

Weeda, G., Wiegant, J., van der Ploeg, M., Geurts van Kessel, A. H. M., van der Eb, A. J., and Hoeijmakers, J. H. J. (1991). Localization of the xeroderma pigmentosum group B-correcting gene *ERCC3* to human chromosome 2q21. Genomics *10*, 1035-1040.

Westerveld, A., Hoeijmakers, J. H. J., van Duin, M., de Wit, J., Odijk, H., Pastink, A., Wood, R. D., and Bootsma, D. (1984). Molecular cloning of a human DNA repair gene. Nature *310*, 425-429.

Zdzienicka, M. Z., Schans, G. P. v. d., 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.

Zdzienicka, M. Z., Venema, J., Mitchell, D. L., van Hoffen, A., van Zeeland, A. A., Vrieling, H., Mullenders, L. H. F., Lohman, P. H. M., and Simons, J. W. I. M. (1992). (6-4) Photoproducts and not cyclobutane pyrimidine dimers are the main UV-induced mutagenic lesions in Chinese hamster cells. Mutat. Res. *273*, 73-83.

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.

Chapter III

Molecular cloning of the human excision repair gene ERCC6

MOLECULAR AND CELLULAR BIOLOGY, Nov. 1990, p. 5806–5813
0270-7306/90/115806-08$02.00/0

Vol. 10, No. 11

Molecular Cloning of the Human DNA Excision Repair Gene *ERCC-6*

C. TROELSTRA,[1] H. ODIJK,[1] J. DE WIT,[1] A. WESTERVELD,[2] L. H. THOMPSON,[3] D. BOOTSMA,[1] AND J. H. J. HOEIJMAKERS[1]*

Department of Cell Biology and Genetics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam,[1] and Institute of Antropogenetics, Academic Medical Centre, 1105 AZ Amsterdam,[2] The Netherlands, and Biomedical Sciences Division, Lawrence Livermore National Laboratory, Livermore, California 94550[3]

Received 30 May 1990/Accepted 7 August 1990

The UV-sensitive, nucleotide excision repair-deficient Chinese hamster mutant cell line UV61 was used to identify and clone a correcting human gene, *ERCC-6*. UV61, belonging to rodent complementation group 6, is only moderately UV sensitive in comparison with mutant lines in groups 1 to 5. It harbors a deficiency in the repair of UV-induced cyclobutane pyrimidine dimers but permits apparently normal repair of (6–4) photoproducts. Genomic (HeLa) DNA transfections of UV61 resulted, with a very low efficiency, in six primary and four secondary UV-resistant transformants having regained wild-type UV survival. Southern blot analysis revealed that five primary and only one secondary transformant retained human sequences. The latter line was used to clone the entire 115-kb human insert. Coinheritance analysis demonstrated that five of the other transformants harbored a 100-kb segment of the cloned human insert. Since it is extremely unlikely that six transformants all retain the same stretch of human DNA by coincidence, we conclude that the *ERCC-6* gene resides within this region and probably covers most of it. The large size of the gene explains the extremely low transfection frequency and makes the gene one of the largest cloned by genomic DNA transfection. Four transformants did not retain the correcting *ERCC-6* gene and presumably have reverted to the UV-resistant phenotype. One of these appeared to have amplified an endogenous, mutated CHO *ERCC-6* allele, indicating that the UV61 mutation is leaky and can be overcome by gene amplification.

To counteract the deleterious consequences of DNA injury, an intricate network of DNA repair systems has evolved (for a review of DNA repair in general, see reference 10). One of the major, universal repair processes is the nucleotide excision repair pathway, which is molecularly best defined in *Escherichia coli* (10, 13, 21, 24). This system removes a broad category of DNA lesions having a very dissimilar structure, such as UV-induced cyclobutane pyrimidine dimers and (6–4) photoproducts, as well as bulky chemical adducts and DNA cross-links. To elucidate the molecular mechanism of the mammalian excision repair pathway, several nucleotide excision repair-deficient mutants are available. Cell lines from patients with the human hereditary disease xeroderma pigmentosum (XP) are one example (for a review, see reference 5). Extensive genetic heterogeneity has been demonstrated among XP patients. Cell fusion experiments have identified at least seven excision-deficient XP complementation groups (2, 9, 17). In addition to XP cells, a set of UV-sensitive, nucleotide excision repair-deficient rodent (mainly Chinese hamster) mutant cell lines have been generated in the laboratory. Eight genetic complementation groups have been described (32, 37, 45; for recent reviews, see references 3, 6, and 30). The cell fusion experiments performed between some of the CHO mutants and XP fibroblasts did not reveal any overlap between these two classes of repair mutants (28, 35). This suggests that a considerable biochemical complexity underlies the nucleotide excision repair process in mammalian cells, for which the molecular mechanism is largely unknown.

CHO mutants of various complementation groups have been successfully used to isolate the correcting human *ERCC* (excision repair cross-complementing rodent repair deficiency) genes, *ERCC-1*, *ERCC-2*, and *ERCC-3*, complementing the excision defects of mutants of groups 1 (44), 2 (41), and 3 (43), respectively. The first five groups display a similar, high degree of UV sensitivity and are deficient in the incision step of the excision repair process (31–33). In contrast, the two mutants making up group 6, CHO mutant UV61 (3) and mouse lymphoma mutant US46 (27, 37), are only moderately sensitive to UV exposure, and UV61 is partially deficient in the incision of damaged DNA (27, 36). Furthermore, mutant UV61 is remarkable in harboring a specific deficiency in the repair of cyclobutane pyrimidine dimers and bulky chemical adducts, but permitting apparently normal repair of (6–4) photoproducts (34). This phenotype suggests that the gene product affected in this mutant is involved in the repair of cyclobutane dimers and bulky adducts but not in the repair of (6–4) lesions. Alternatively, it is possible that the mutation in the UV61 protein alters the affinity of the repair complex for different types of damage. The alteration would then be such that the rate of removal of cyclobutane dimers and bulky adducts was notably diminished, whereas repair of (6–4) photoproducts was not significantly affected. In both hypotheses the UV61 (*ERCC-6*) polypeptide plays an important role in repair of specific types of DNA injury.

To obtain insight into the function of this particular protein in the DNA excision repair process, we have cloned and

* Corresponding author.

partially characterized the human gene correcting the mutation in UV61.

MATERIALS AND METHODS

Cell culture. The CHO mutant cell line UV61 used in this study was isolated by Busch et al. (3) and assigned to complementation group 6 (36). Its repair-competent, parental cell line is the CHO line AA8 (3). All cells were grown in 1:1 F10-Dulbecco minimal essential medium supplemented with antibiotics and 8% fetal calf serum.

DNA transfection. High-molecular-size HeLa DNA (ca. 200 to 300 kb), isolated as described previously (16), was either partially cleaved to an average fragment size of 50 to 60 kb and ligated to a dominant marker molecule before transfection or directly used for cotransfection with a dominant marker. For transfection of ligated DNA, HeLa DNA partially digested with *PstI* or *MboI* was ligated to *PstI*-digested pSV3gptH (44) or *Bam*HI-digested pMCS (14), respectively, in a molar ratio of 1:3, by using T4 DNA ligase (GIBCO/Bethesda Research Laboratories, Gaithersburg, Md.). Ligations were tested on agarose gels.

For transfections, 5×10^5 UV61 cells were seeded in 100-mm Petri dishes 1 day before DNA transfection. The high-molecular-size DNA, together with one of the two dominant marker molecules mentioned, was transfected into UV61 cells by a modification of the calcium phosphate precipitation technique of Graham and Van der Eb (12). A 20-μg sample of DNA (20 μg of ligated DNA or 15 μg of HeLa DNA plus 5 μg of circular pSV3gptH DNA) was applied to each petri dish. Following a 16-h exposure to the DNA, cells were treated for 30 min with 10% dimethyl sulfoxide and grown for 24 h on nonselective medium to allow expression of the transfected marker. Thereafter, the medium was changed to selective medium. When selecting for expression of the pMCS dominant marker, G418 (GIBCO Ltd., Paisley, Scotland) was added to the F10-Dulbecco minimal essential medium described above (concentration, 750 μg/ml during the first 3 days of selection and 800 μg/ml thereafter). Selection for pSV3gptH was done by adding aminopterin (0.2 μg/ml), thymidine (5 μg/ml), xanthine (10 μg/ml), hypoxanthine (15 μg/ml), mycophenolic acid (20 μg/ml), and deoxycytidine (2.3 μg/ml) to F10-Dulbecco minimal essential medium. The selection medium was refreshed every 3 to 4 days.

The same procedure was followed for secondary transfections. DNA of a primary transformant (PT) was used with or without addition of dominant marker pRSVneo (11) (20 μg of PT DNA plus 2 μg of pRSVneo in cotransfection).

UV selection and UV survival. After the appearance of G418- or mycophenolic acid-resistant colonies (within 10 to 12 days after transfection), the cells were trypsinized. UV selection was started 16 to 20 h after trypsinization. Cells were exposed three times to 8.4 J/m^2 (Philips TUV low-pressure mercury tube, 15 W, 0.6 J/m^2/s; predominantly 254 nm), with 24-h intervals. UV-resistant colonies were isolated, grown in selective medium, and characterized with respect to UV resistance and human DNA content by Southern blot hybridization.

For determination of UV sensitivity, cells were plated at densities varying from 2×10^2 to 2×10^4 cells per 60-mm petri dish, depending on the cell line and the UV dose. Cells were irradiated 1 day after plating. A series of dishes was irradiated for each cell line, each dish receiving a single dose (three dishes per UV dose). Clones were counted 6 to 7 days after UV irradiation, and relative cloning efficiencies were determined.

TABLE 1. Repair characteristics of mutant UV61

Characteristic	Result for UV61[a]
Sensitivity to DNA-damaging agents[b]	
UV	ca. 2.7×[c]
7-BrMeBA[d]	ca. 3×[c]
Rate of incision	Normal[e]
Removal of (6–4) photoproducts	Normal
Removal of dimers	Deficient
Removal of bulky adducts	Deficient
UDS[f]	ca. 70%[e]
UV-induced mutagenesis	ca. 5×[e]
Postreplication recovery	Normal

[a] Data from reference 33.
[b] At D_{10}.
[c] Compared with that of the wild type.
[d] 7-BrMeBA, 7-Bromomethylbenz[*a*]anthracene.
[e] Measured first at 2 h after UV irradiation.
[f] Own unpublished results.

Southern blot hybridization. Digestion of DNA with restriction enzymes, gel electrophoresis, and hybridizations were performed by using routine procedures as described by Maniatis et al. (19). Southern blotting to Zeta probe blotting membranes (Bio-Rad, Richmond, Calif.) was performed by using alkaline transfer (22), as described by the manufacturer.

Library construction and screening. High-molecular-size DNA was partially digested with *MboI* and size fractionated on a sucrose gradient. The 15- to 25-kb fraction was ligated to the *Bam*HI sites of lambda EMBL-3 phage arms (Stratagene, La Jolla, Calif.), packaged as described previously (25), and then used to infect *E. coli* LE392 cells. A total of 3×10^6 phage were plated, and replica filters were prepared as described previously (19).

The filters were hybridized with radioactively labeled human Cot-1 DNA. Hybridizing plaques were plated again for a second round of screening. Single plaques were grown in liquid culture, and DNA was purified as described previously (19).

Overlapping phages were identified by using restriction enzyme site mapping and the restriction fingerprinting technique as described by Coulson et al. (7). Enzymes used in the latter procedure were *Sau*3A and *Hin*dIII.

RESULTS

Generation of repair-proficient PTs. The main characteristics of mutant UV61 are summarized in Table 1. In view of the moderate UV sensitivity of UV61 cells, it was necessary to develop an optimal selection protocol for repair-proficient transformants. In reconstruction experiments wild-type and UV61 cells were mixed in various ratios and different selection protocols were tested. The deduced UV selection procedure is described in Materials and Methods.

Two strategies were followed simultaneously to generate PTs. One approach, transfection of ligated DNA (partially digested HeLa DNA ligated in vitro to dominant marker molecules), increases the chance that a dominant marker copy is near the repair gene. This permits secondary transfection of the gene linked to a dominant marker copy and facilitates gene identification and cloning. The cotransfection, on the other hand, is more efficient for large genes.

Transfection of ligated DNA to more than 3.6×10^8 cells resulted in two PTs that survived the UV selection (PT-3 and PT-4). Cotransfection of high-molecular-size HeLa DNA

TABLE 2. Transfection efficiencies of UV-resistant PTs and STs of UV61

Transfection	Approach[a]	No. of UV^r transformants[b]	Transfection efficiency[c]
Primary	Ligation	2 (0)	1:150,000
Primary	Cotransfection	4 (4)	1:100,000 (1:100,000)
Secondary	Linking	0 (0)	
Secondary	Cotransfection	4 (2)	1:20,000 (1:40,000)

[a] For a detailed description of procedures, see Materials and Methods.
[b] Number of UV-resistant, dominant marker-containing transformants isolated; the number after subtraction of revertants is given in parentheses.
[c] Transfection efficiencies given as UV-resistant transformants per dominant marker-expressing transformant. Efficiencies after correction for the presence of revertants are given in parentheses.

and pSV3gptH yielded four PTs (PT-1, PT-2, PT-5, and PT-6) in experiments involving 6×10^8 UV61 cells. The transfection frequencies (i.e., number of UV-surviving transformants per dominant marker-containing transformants) were extremely low (Table 2).

Characterization of PTs. To assess the degree of repair proficiency of the transformants, we determined their UV survival. PT-1 had regained wild-type (AA8) UV resistance, and the other PTs were in the wild-type range too (Fig. 1).

The PTs were analyzed for the presence of human sequences by using human Cot-1 DNA or a cloned human Alu repeat (8) as probes. As expected for PTs, Southern blot analysis (Fig. 2) revealed the presence of considerable amounts of human DNA in PT-1, PT-2, PT-3, PT-5, and PT-6 (results for PT-5 and PT-6 not shown). Surprisingly, PT-4 did not contain any detectable human sequences. This could mean that PT-4 does not contain a human gene, but is in fact a dominant marker-containing revertant. In view of the large scale of the genomic DNA transfections, this possibility cannot be ruled out.

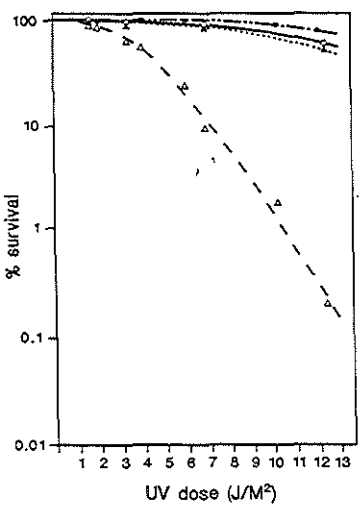


FIG. 1. UV survival curves of wild-type cell line AA8 (○), mutant UV61 (△), and UV61 transformants PT-1 (●) and ST-1 (▲).

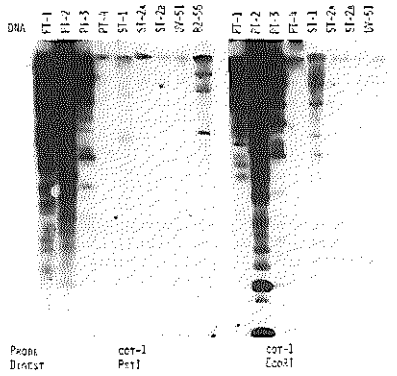


FIG. 2. Southern blot analysis of *Pst*I- (left) or *Eco*RI (right)-digested DNA (20 μg) from mutant UV61 and from UV61 PTs and STs after hybridization with ^{32}P-labeled human Cot-1 DNA as the probe. B2-56 (lane 9, left panel) is a secondary transformant of CHO mutant 27-1, containing the human *ERCC-3* gene (43), and served as a positive control.

In our selection protocol each petri dish with colonies of dominant marker-containing UV61 transformants (>100 cells per colony) was trypsinized prior to UV irradiation. When a transformant colony stably retains the correcting human *ERCC-6* gene, one expects that respreading the cells will result in many (>10 to 50) UV-resistant clones, assuming normal cloning efficiency and growth rate. For a revertant arising during dominant marker selection or, more likely, during UV selection, this number is expected to be much smaller. Table 3 summarizes our findings with respect to the number of UV-resistant clones after trypsinization and UV irradiation. It is striking that PT-4 falls in the second class (i.e., only one or a few UV-surviving colonies). This further suggests that PT-4 is a UV-resistant revertant.

Generation of repair-proficient STs. In secondary transfection experiments with DNA of PTs without addition of extra dominant-marker molecules, no clones with combined MPA and UV resistance or G418 and UV resistance were isolated from a total of about 9×10^8 transfected UV61 cells. This

TABLE 3. Characteristics of UV-resistant transformants of UV61

Transformant	No. of UV^r clones/dish	Presence of human DNA
PT-1	10–50	++
PT-2	<3	++++
PT-3	10–50	++
PT-4	<3	–[a]
PT-5	10–50	++
PT-6	10–50	++++
ST-1	10–50	+
ST-2a	<3	–
ST-2b	<3	–
ST-5	5–10	–

[a] The minus sign indicates that no human DNA is detectable on Southern blots with human repeat (Cot-1, Alu) as the probe.

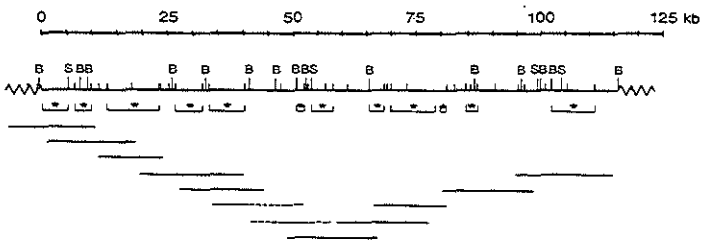


FIG. 3. Physical map of the human insert cloned from ST-1. Fragments of <300 bp have not been mapped. Indicated below the map are various overlapping lambda clones isolated from the library by screening with human Cot-1 DNA (for the sake of simplicity, only part of the lambda clones isolated and mapped are indicated). Symbols: *, repeat-containing fragments; ⊥, *Eco*RI; B, *Bam*HI; S, *Sal*I; —, human insert; ~, flanking UV61 DNA; - - -, phages isolated after chromosomal walking.

result suggests that the distance between an integrated dominant-marker copy and the human gene conferring UV resistance, in the DNA of the PTs tested, is too large to be transferred simultaneously. Therefore, new dominant-marker molecules (pRSVneo) were added to PT DNAs for subsequent secondary transfection. This approach yielded four UV-resistant secondary transformants (STs) (after transfection of 4.0×10^8 UV61 cells), designated ST-1, ST-2a, ST-2b, and ST-5 (the numbers refer to the parental PTs; ST-2a and ST-2b were isolated from two independent dishes in one experiment with PT-2 as the donor DNA). Again, transfection efficiencies were very low (Table 2).

Characterization of STs. The UV survival of all STs was determined. ST-1 was corrected up to wild-type (AA8) and PT-1 UV resistance (Fig. 1). Similar results were obtained for the other STs.

Southern blot analysis with human Cot-1 DNA as the probe revealed the presence of human sequences only in the DNA of ST-1 (Fig. 2). For ST-2a and ST-2b (and ST-5 [not shown]) we were unable to detect any human repetitive DNA.

The absence of detectable human DNA in three of the four STs and in PT-4 can be explained in several ways. The transfected repair gene may contain two few repetitive sequences to be visualized by Southern blot hybridization with human Cot-1 DNA as the probe. Alternatively, the three empty STs (ST-2a, ST-2b, and ST-5) may simply be dominant marker-containing revertants. Two of these three STs (ST-2a and ST-2b) were from dishes with one or a few UV-resistant clones after trypsinization and UV selection (Table 3), as was PT-4 (see above). These results further suggest that ST-2a and ST-2b may be revertants.

Molecular cloning of the human insert. ST-1, which contains detectable human material, is from a dish with multiple UV-resistant clones after trypsinization and UV selection. Furthermore, ST-1 is derived from a PT (PT-1), also belonging to this category (Table 3). This information suggests that ST-1 is a bona fide transformant. Therefore, ST-1 DNA was chosen to construct a genomic library in lambda EMBL-3. The library (2.5×10^6 plaques; total complexity, 16 times the haploid genome) was screened with human Cot-1 DNA as the probe. Two hundred plaques containing human DNA sequences were identified; after a second round of screening pure cultures were grown from 50 plaques. DNA of these lambda recombinants was isolated and characterized by restriction mapping. The human and hamster parts were identified by hybridization with species-specific repeat probes.

Overlapping lambda clones were identified as described in Materials and Methods. A restriction site map of the inserts resulted in the identification of two nonoverlapping segments. Both segments harbored UV61 sequences at one end. Chromosomal walking was performed by using unique sequences from the human termini of both segments. Subsequently, phages were isolated that joined all contiguous parts, yielding a human insert with a total size of approximately 115 kb. A physical map is shown in Fig. 3. The phages isolated via chromosomal walking are indicated. These cover a part which contains very few repetitive sequences, explaining why these lambda recombinants had escaped detection in the first Cot-1 screening. In fact, the cloned region overall has relatively few repeats (Fig. 3).

Coinheritance analysis of independent transformants. A systematic search for coinheritance of the same human DNA segment in independent UV61 transformants was performed. This search was based on the notion that the human gene specifically conferring UV resistance to UV61 should be present in each genuine transformant. To this aim, 14 unique probes spread over the entire cloned human region were isolated and hybridized to genomic DNA digests of all UV-resistant PTs and STs obtained. The results are shown in Fig. 4A and summarized in Fig. 4B. Probe IV, for instance, does recognize the expected 6.4-kb *Eco*RI fragment in ST-1, from which the library was made. In addition, it hybridizes to a band of the same size in PT-1, PT-2, PT-5, PT-6, ST-5, and HeLa DNA. PT-3, PT-4, ST-2a, and ST-2b do not contain this human fragment. The same holds true for probes III through XIV, with the exception of probe VIII (Fig. 4). This indicates that the same contiguous region of at least 100 kb is present in 6 of 10 UV61 transformants.

Probes I and II are near the left UV61 border in ST-1 (Fig. 4). Probe I recognizes a rearranged *Eco*RI fragment in DNA of ST-1 compared with HeLa DNA, or, in the case of PT-2, PT-6, and ST-5, no fragment at all (Fig. 4). Probe II recognizes a 1.8-kb *Eco*RI fragment in ST-1 DNA, as in HeLa DNA; a rearranged fragment is detected in PT-2 (Fig. 4B). At the right-hand end the DNA of ST-1 is the first to diverge from the human genomic sequence and that of the other transformants (Fig. 4B). From these results, we deduce that the length of the human segment common to all transformants is approximately 100 kb.

The signals in the lanes with PT-1 and ST-1 are stronger than those in the other lanes (Fig. 4A and 5). Because both probe VII (Fig. 4A) and probe XV (Fig. 5) weakly cross-hybridize with hamster sequences, the endogenous bands (2.5 kb for probe VII in Fig. 4A [arrowhead]; the lower two

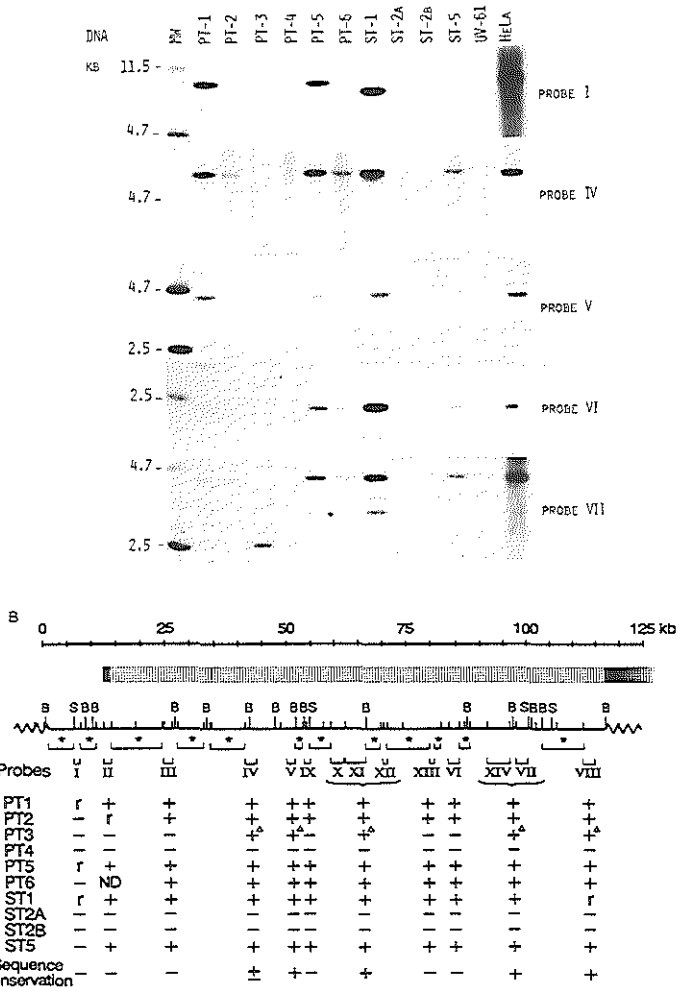


FIG. 4. Coinheritance analysis of all transformants for the retention of human sequences present in ST-1. (A) The same Southern blot of *Eco*RI-digested DNA from all primary and secondary UV61 transformants, mutant UV61, and HeLa cells was hybridized with different probes from the human region present in ST-1. Probes are indicated below the map in panel B. (B) Summary of coinheritance analysis. Symbols: +, same fragment hybridizing as in HeLa DNA; −, no fragment hybridizing; r, hybridizing fragment rearranged compared with HeLa DNA; +△, endogenous UV61 fragment hybridizing more strongly than in other transformants and UV61; ND, not determined. See the legend to Fig. 3 for explanations of other symbols.

hybridizing bands for probe XV in Fig. 5) serve as a convenient internal control for the amount of DNA loaded. Apparently, both PT-1 and ST-1 have a 5- to 10-fold amplification of the transfected, human region.

The amplification of this region in ST-1 is probably the reason that we were able to detect it by using human Cot-1 DNA as the probe on Southern blots of genomic DNA. ST-5, on the other hand, has only one copy (Fig. 4A and 5), which was not visible on our blots. Since this region contains relatively few repeats, the Southern blot hybridizations were clearly not sensitive enough to detect it as a single copy.

Presence of UV61 revertants. The coinheritance analysis revealed that PT-3, PT-4, ST-2a, and ST-2b all lack the entire human locus (Fig. 4B), which strongly suggests that these

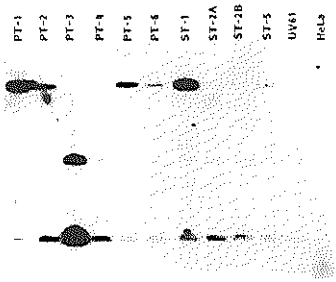


FIG. 5. Southern blot analysis of *Hin*dIII-digested DNA (20 μg) from HeLa cells, mutant UV61, and all PTs and STs with ^{32}P-labeled unique sequences from the region between probes VII and VIII on the map in Fig. 4B (probe XV).

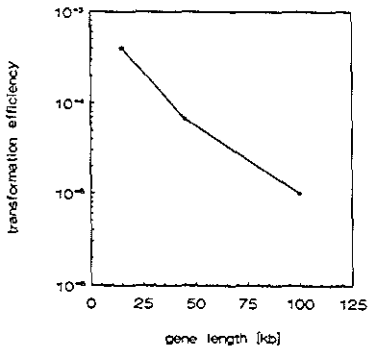


FIG. 6. Relationship between primary transfection efficiencies and gene size of *ERCC-1*, *ERCC-3*, and *ERCC-6*. Transfection efficiency is given as UV-resistant transformants per dominant marker-containing transformants. For *ERCC-1*, ligated genomic DNA (50 to 60 kb in size) was used.

are revertants. This agrees well with the fact that three of them (PT-4, ST-2a, and ST-2b) lack detectable human sequences (Fig. 2) and originated from a dish with a small number of UV-resistant clones after UV selection (Table 3). (One additional transformant, PT-2, also yielded a small number of UV-resistant clones, but nevertheless appeared to contain the common human segment (compare Table 3 and Fig. 4). This may be explained by initial instability of the transfected sequences, low cloning efficiency, and/or slow growth of this transformant.)

Various human genomic probes contained conserved sequences and cross-hybridized under normal stringency to Chinese hamster genomic sequences representing the hamster *ERCC-6* homolog (for example, probe VII in Fig. 4A [arrowhead] and probe XV in Fig. 5). The cross-hybridization to Chinese hamster DNA in the lane of PT-3 is considerably stronger than that in the other lanes, irrespective of the probe or digest. This indicates that PT-3 has amplified an endogenous, mutated allele at the CHO *ERCC-6* locus.

DISCUSSION

The evidence that we have indeed cloned *ERCC-6* is based on the coinheritance analysis of independent transformants for the retention of human sequences present in ST-1. This study showed that six transformants harbored the same human segment of 100 kb. The chance that this is due to coincidence is extremely small. When one assumes that every transformant integrates on the average 10^4 kb of exogenous sequences, as we have shown for several CHO lines (15), the likelihood that four independent PTs contain the same human fragment by coincidence is approximately 10^{-10}. Taking into account that two STs also possessed this region, this chance becomes vanishingly small. Two additional findings also strengthen the significance of the correlation between the cloned human integrate and the repair-proficient phenotype of the UV61 transformants: (i) the exceptionally large size of the common genomic fragment for transfection experiments and (ii) the observation that one of the transformants (PT-3) appeared to have amplified the CHO sequence cross-hybridizing to the human integrate. Although the gene is spread over too many lambda recombinants to permit direct transfection (of pooled clones) to UV61 cells, we conclude, on the basis of the arguments summarized above, that *ERCC-6* must be retained on the 100-kb common region. Hence this gene was cloned notwithstanding the moderate sensitivity of the mutant to UV, the extremely low transfection frequency, the low content of repetitive elements in the gene, and the occurrence of revertants.

The *ERCC-6* gene probably spans most of the cloned common area because during transfection double-strand breaks have been shown to occur on the average every 5 to 15 kb in various cell lines (15). Preliminary mapping of partial *ERCC-6* cDNA clones on the isolated genomic region confirms this proposition. A cDNA clone representing approximately 75% of the smallest *ERCC-6* mRNA already covers 80 kb of the cloned locus. As far as we know, *ERCC-6* is the largest gene to be cloned by using the genomic transfection approach. The frequent occurrence of double-strand breaks in transfected DNA molecules during the transformation process (15) strongly selects against large genes. This observation provides a reasonable explanation for the extremely low transfection efficiency encountered in the cloning of *ERCC-6* (Table 2), which is even more pronounced when we correct for the revertants among the UV61 transformants isolated. The transfection frequency (primary cotransfections) for *ERCC-6* is 1 in 100,000 dominant marker-containing UV61 transformants. This is 40-fold lower than the frequency of *ERCC-1* transformants (44) (gene size, 15 to 17 kb [39]) and 7-fold below that of the *ERCC-3* gene (approximate size, 45 kb [43]). These genes were isolated by us under comparable conditions. Hence a relationship exists between gene size and transfection frequency in the CHO lines used for these experiments (Fig. 6).

An interesting aspect concerns the identification of revertants. The occurrence of revertants has been described before in attempts to clone other repair genes (see, e.g., references 1, 4, 18, 23, and 26). For two simian virus 40-transformed fibroblasts isolated from XP patients and belonging to complementation group A, which have been used extensively as recipients for genomic DNA transfections, the recovery of partial and complete revertants has been reported by several groups (1, 23). One partial revertant has been characterized in detail and found to be proficient in (6–4) photoproduct excision, but affected in dimer

removal (4), a remarkable phenotype that resembles UV61 and a Chinese hamster mutant, VH-1, belonging to complementation group 2 (20, 46). The frequency of revertants seems to be enhanced when long-lasting UV selection protocols are used. This is one of the reasons why we have tried to minimize the number of UV exposures of UV61.

In none of the instances mentioned above has the nature of reversion been elucidated. The analysis of PT-3 in this study revealed that amplification of an endogenous, mutated Chinese hamster *ERCC-6* allele underlies the acquired UV resistance in this UV61 revertant. We have found recently that the amplification is also reflected at the mRNA level. These data suggest that the mutation in UV61 is leaky and that its effect can be overcome by increasing the copy number of the gene and mRNA, thereby raising the amount of (partly functional) gene product. This explanation would be compatible with a model in which the affinity of the *ERCC-6* protein for dimer lesions is lowered by the UV61 mutation and can be compensated for by larger quantities of the defective polypeptide. The molecular basis of the reversion in the other three UV61 revertants is unknown, but must involve other mechanisms than gene amplification.

ERCC-6 is the fourth of a set of human repair genes that have been cloned by using UV-sensitive, laboratory-induced CHO mutants. One of the most interesting features to emerge from the analysis of these genes is their high interspecies sequence conservation. The *ERCC-1* protein is homologous to the yeast RAD10 excision repair gene product and to parts of the *E. coli* UvrA and UvrC proteins (38, 40). The protein encoded by the *ERCC-2* gene, cloned by Weber et al. (41), possesses a high level of identity with the yeast RAD3 protein (42), which was shown by Sung et al. to specify a 5'-to-3' DNA helicase (29). Finally, the *ERCC-3* gene also is very strongly conserved, and a homologous yeast gene has been identified (M. H. M. Koken, G. Weeda, and J. H. J. Hoeijmakers, unpublished data). This gene was recently found to be involved in the human repair diseases XP and Cockayne's syndrome (G. Weeda, R. C. A. van Ham, W. Vermeulen, D. Bootsma, A. J. van der Eb, and J. H. J. Hoeijmakers, Cell, in press). Future analysis of *ERCC-6* should reveal the level of sequence conservation of this gene, its specific role in dimer repair, and whether it is implicated in one of the human repair disorders.

ACKNOWLEDGMENTS

We are grateful to N. G. J. Jaspers and C. W. Weber for critical reading of the manuscript. We also thank M. Kuit for photography and G. Weeda for providing genomic DNA from a transformant of CHO mutant 27-1.

This work was supported in part by the Commission of the European Community contract BI6-141-NL and Medigon, Foundation of Medical and Health Research in the Netherlands, contract 900-501-091.

LITERATURE CITED

1. **Blum, M., I. Baumann, H. Lohrer, H. J. Rahmsdorf, and P. Herrlich.** 1989. A promising genomic transfectant into xeroderma pigmentosum group A with highly amplified mouse DNA and intermediate UV resistance turns revertant. Biochem. Biophys. Res. Commun. 160:647–655.
2. **Bootsma, D., W. Keijzer, E. G. Jung, and E. Bohnert.** 1989. Xeroderma pigmentosum complementation group XP-I withdrawn. Mutat. Res. 218:149–151.
3. **Busch, D., C. Greiner, K. Lewis, R. Ford, G. Adair, and L. Thompson.** 1989. Summary of complementation groups of UV-sensitive CHO mutants isolated by large-scale screening. Mutagenesis 4:349–354.
4. **Cleaver, J. E., F. Cortés, L. H. Lutze, W. F. Morgan, A. N. Player, and D. L. Mitchell.** 1987. Unique DNA repair properties of a xeroderma pigmentosum revertant. Mol. Cell. Biol. 7:3353–3357.
5. **Cleaver, J. E., and K. H. Kraemer.** 1989. Xeroderma pigmentosum, p. 2949–2971. *In* C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle (ed.), The metabolic basis for inherited disease, vol. 2. McGraw-Hill Book Co., New York.
6. **Collins, A., and R. T. Johnson.** 1987. DNA repair mutants in higher eukaryotes. J. Cell Sci. Suppl. 6:61–82.
7. **Coulson, A., J. Sulston, S. Brenner, and J. Karn.** 1986. Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 83:7821–7825.
8. **Deininger, P. L., D. J. Jolly, C. M. Rubin, T. Friedmann, and C. W. Schmid.** 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. J. Mol. Biol. 151:17–33.
9. **De Weerd-Kastelein, E. A., W. Keijzer, and D. Bootsma.** 1972. Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization. Nature (London) New Biol. 238:80–83.
10. **Friedberg, E. C.** 1985. DNA repair. W. H. Freeman & Co., San Francisco.
11. **Gorman, C., R. Padmanabhan, and B. H. Howard.** 1983. High efficiency DNA-mediated transformation of primate cells. Science 221:551–553.
12. **Graham, F. L., and A. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52:456–467.
13. **Grossman, L., P. R. Caron, S. J. Mazur, and E. Y. Oh.** 1988. Repair of DNA-containing pyrimidine dimers. FASEB J. 2:2696–2701.
14. **Grosveld, F. G., T. Lund, E. J. Murray, A. L. Mellor, H. H. M. Dahl, and R. A. Flavell.** 1982. The construction of cosmid libraries which can be used to transform eukaryotic cells. Nucleic Acids Res. 10:6715–6732.
15. **Hoeijmakers, J. H. J., H. Odijk, and A. Westerveld.** 1987. Differences between rodent and human cell lines in the amount of integrated DNA after transfection. Exp. Cell Res. 169:111–119.
16. **Hoeijmakers, J. H. J., A. Westerveld, and D. Bootsma.** 1988. Methods and strategies for molecular cloning of mammalian DNA reair genes by DNA-mediated gene transfer, p. 181–201. *In* E. C. Friedberg and P. C. Hanawalt (ed.), DNA repair, vol. 3. Marcel Dekker, Inc., New York.
17. **Johnson, R. T., G. C. Elliot, S. Squires, and V. C. Joysey.** 1989. Lack of complementation between xeroderma complementation groups D and H. Hum. Genet. 81:203–210.
18. **MacInnes, M. A., J. D. Bingham, L. H. Thompson, and G. F. Strniste.** 1984. DNA-mediated cotransfer of excision repair capacity and drug resistance into Chinese hamster ovary cell line UV-135. Mol. Cell. Biol. 4:1152–1158.
19. **Maniatis T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Mitchell, D. L., M. Z. Zdzienicka, A. A. van Zeeland, and R. S. Nairn.** 1989. Intermediate (6–4) photoproduct repair in Chinese hamster V79 mutant V-H1 correlates with intermediate levels of DNA incision and repair replication. Mutat. Res. 226:43–47.
21. **Orren, D. K., and A. Sancar.** 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.
22. **Reed, K. C., and D. A. Mann.** 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. 13:7207–7221.
23. **Royer-Pokora, B., and W. A. Haseltine.** 1984. Isolation of UV-resistant revertants from a xeroderma pigmentosum group A cell line. Nature (London) 311:390–392.
24. **Sancar, A., and G. B. Sancar.** 1988. DNA repair enzymes. Annu. Rev. Biochem. 57:29–67.
25. **Scalenghe, F., E. Turco, J. E. Edström, V. Pirrotta, and M. Melli.** 1981. Microdissection and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes. Chromosoma 82:205–216.

26. **Schultz, R. A., D. P. Barbis, and E. C. Friedberg.** 1985. Studies on gene transfer and reversion to UV resistance in xeroderma pigmentosum cells. Somatic Cell Mol. Genet. 11:617–624.
27. **Shiomi, T., N. Hieda-Shiomi, and K. Sato.** 1982. Isolation of UV-sensitive mutants of mouse L5178Y cells by a cell suspension spotting method. Somatic Cell Genet. 8:329–345.
28. **Stefanini, M., W. Keijzer, A. Westerveld, and D. Bootsma.** 1985. Interspecies complementation analysis of xeroderma pigmentosum and UV-sensitive Chinese hamster cells. Exp. Cell Res. 161:373–380.
29. **Sung, P., L. Prakash, S. W. Matson, and S. Prakash.** 1987. RAD3 protein of *Saccharomyces cerevisiae* is a DNA helicase. Proc. Natl. Acad. Sci. USA 84:8951–8955.
30. **Thompson, L. H.** 1989. Somatic cell genetics approach to dissecting mammalian DNA repair. Environ. Mol. Mutagen. 14:264–281.
31. **Thompson, L. H., K. W. Brookman, L. E. Dillehay, C. L. Mooney, and A. V. Carrano.** 1982. Hypersensitivity to mutation and sister-chromatid exchange induction in CHO cell mutants defective in incising DNA containing UV-lesions. Somatic Cell Genet. 8:759–773.
32. **Thompson, L. H., D. B. Busch, K. Brookman, C. I. Mooney, and D. A. Glaser.** 1981. Genetic diversity of UV-sensitive DNA repair mutants of Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA 78:3734–3737.
33. **Thompson, L. H., and A. V. Carrano.** 1983. Analysis of mammalian cell mutagenesis and DNA repair using *in vitro* selected CHO cell mutants, p. 125–143. *In* E. C. Friedberg and B. A. Bridges (ed.), Cellular responses to DNA damage. Alan R. Liss, Inc., New York.
34. **Thompson, L. H., D. L. Mitchell, J. D. Regan, S. D. Bouffler, S. A. Stewart, W. L. Carrier, R. S. Nairn, and R. T. Johnson.** 1988. CHO mutant UV61 removes (6–4)photoproducts but not cyclobutane dimers. Mutagenesis 4:140–146.
35. **Thompson, L. H., C. L. Mooney, and K. Brookman.** 1985. Genetic complementation between UV-sensitive CHO mutants and xeroderma pigmentosum fibroblasts. Mutat. Res. 150:423–429.
36. **Thompson, L. H., E. P. Salazar, K. W. Brookman, C. C. Collins, S. A. Stewart, D. B. Busch, and C. A. Weber.** 1987. Recent progress with DNA repair mutants of Chinese hamster ovary cells. J. Cell Sci. Suppl. 6:97–110.
37. **Thompson, L. H., T. Shiomi, E. P. Salazar, and S. A. Stewart.** 1988. An eighth complementation group of rodent cells hypersensitive to ultraviolet radiation. Somatic Cell Mol. Genet. 14:605–612.
38. **Van Duin, M., J. de Wit, H. Odijk, A. Westerveld, A. Yasui, M. H. M. Koken, J. H. J. Hoeijmakers, and D. Bootsma.** 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.
39. **Van Duin, M., M. H. M. Koken, J. van den Tol, P. ten Dijke, H. Odijk, A. Westerveld, D. Bootsma, and J. H. J. Hoeijmakers.** 1987. Genomic characterization of the human DNA excision repair gene *ERCC-1*. Nucleic Acids Res. 15:9195–9213.
40. **Van Duin, M., J. van den Tol, P. Warmerdam, H. Odijk, D. Meijer, A. Westerveld, D. Bootsma, and J. H. J. Hoeijmakers.** 1988. Evolution and mutagenesis of the mammalian excision repair gene *ERCC-1*. Nucleic Acids Res. 16:5305–5322.
41. **Weber, C. A., E. P. Salazar, S. A. Stewart, and L. H. Thompson.** 1988. Molecular cloning and biological characterization of a human gene, *ERCC2*, that corrects the nucleotide excision repair defect in CHO UV5 cells. Mol. Cell. Biol. 8:1137–1146.
42. **Weber, C. A., E. P. Salazar, S. A. Stewart, and L. H. Thompson.** 1990. *ERCC-2*: cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast *RAD3*. EMBO J. 9:1437–1447.
43. **Weeda, G., R. C. A. van Ham, R. Masurel, A. Westerveld, H. Odijk, J. de Wit, D. Bootsma, A. J. van der Eb, and J. H. J. Hoeijmakers.** 1990. Molecular cloning and biological characterization of the human excision repair gene *ERCC-3*. Mol. Cell. Biol. 10:2570–2581.
44. **Westerveld, A., J. H. J. Hoeijmakers, M. van Duin, J. de Wit, H. Odijk, A. Pastink, R. D. Wood, and D. Bootsma.** 1984. Molecular cloning of a human DNA repair gene. Nature (London) 310:425–429.
45. **Zdzienicka, M. Z., G. P. van der Schans, and J. W. I. M. Simons.** 1988. Identification of a new seventh complementation group of UV-sensitive mutants in Chinese hamster cells. Mutat. Res. 194:165–170.
46. **Zdzienicka, M. Z., G. P. van der Schans, A. Westerveld, A. A. van Zeeland, and J. W. I. M. Simons.** 1988. Phenotypic heterogeneity within the first complementation group of UV-sensitive mutants of Chinese hamster cell lines. Mutat. Res. 193:31–41.

Chapter IV

Localization of the nucleotide excision repair gene ERCC6 to human chromosome 10q11-q21

GENOMICS 12, 745–749 (1992)

Localization of the Nucleotide Excision Repair Gene *ERCC6* to Human Chromosome 10q11–q21

C. TROELSTRA,* R. M. LANDSVATER,†,[1] J. WIEGANT,‡ M. VAN DER PLOEG,‡ G. VIEL,† C. H. C. M. BUYS,† AND J. H. J. HOEIJMAKERS*,[2]

*Medical Genetic Centre, Department of Cell Biology and Genetics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; †Department of Medical Genetics, University of Groningen, Groningen, The Netherlands; and ‡Department of Cytochemistry and Cytometry, Sylvius Laboratory, Leiden, The Netherlands

Received September 12, 1991; revised November 29, 1991

We have cloned the human DNA excision repair gene *ERCC6* by virtue of its ability to correct the uv sensitivity of Chinese hamster ovary cell mutant UV61. This mutant is a member of complementation group 6 of the nucleotide excision repair-deficient rodent mutants. By means of *in situ* hybridization and Southern blot analysis of mouse × human somatic cell hybrids, the gene was localized to human chromosome 10q11–q21. An RFLP detected within the *ERCC6* locus can be helpful in linkage analysis. © 1992 Academic Press, Inc.

INTRODUCTION

To protect DNA from deleterious accumulation of damage and permanent mutations, an intricate network of DNA repair systems has evolved (reviewed by Friedberg, 1985). One of the major DNA repair processes is the nucleotide excision repair pathway. This system removes a broad range of DNA lesions, such as uv-induced cyclobutane pyrimidine dimers and (6–4) photoproducts, bulky chemical adducts, and DNA crosslinks.

Two human genetic diseases in which the excision repair mechanism is defective are known: xeroderma pigmentosum (XP) and Cockayne syndrome (CS). The autosomal, recessive disorder XP is clinically characterized by extreme sensitivity of the skin to sunlight (uv), pigmentation abnormalities, predisposition to skin cancer, and frequently neurological complications (Cleaver and Kraemer, 1989). Cell fusion experiments have identified at least seven excision-deficient XP complementation groups (De Weerd-Kastelein *et al.*, 1972; Vermeulen *et al.*, 1991, and references therein). CS patients exhibit sun sensitivity, dwarfism, microcephaly, wizened appearance, deafness, and severe mental retardation. CS is, unlike XP, not associated with an elevated risk for skin tumor formation (reviewed by Lehmann, 1987). Recently, CS cells were found to be disturbed in a subpathway of the nucleotide excision repair: the preferential repair of actively transcribed genes (Venema *et al.*, 1990). Also, CS is genetically heterogeneous; at least three complementation groups have been identified (Lehmann, 1982).

A second class of mammalian excision repair-deficient cell lines consists of laboratory-induced, uv-sensitive, rodent cell lines. Eight complementation groups have been identified (reviewed by Collins and Johnson, 1987; Busch *et al.*, 1989). So far, only one overlap between these two classes has been detected: the gene correcting XP complementation group B is identical to the gene correcting complementation group 3 of the rodent mutant cell lines (Weeda *et al.*, 1990).

Recently, we have cloned the human *ERCC6* gene (excision repair cross complementing rodent repair deficiency). This gene is capable of correcting a uv-sensitive CHO mutant belonging to group 6: UV61 (Troelstra *et al.*, 1990). Mutant UV61 is remarkable in harboring a specific deficiency in the repair of cyclobutane pyrimidine dimers and bulky chemical adducts, but permitting apparently normal repair of (6–4) photoproducts (Thompson *et al.*, 1988). This suggests that the gene product affected in UV61 is, directly or indirectly, involved in damage recognition. Whether *ERCC6* is implicated in one of the XP or CS complementation groups remains to be elucidated.

Here we report the chromosomal localization of the *ERCC6* gene by *in situ* hybridization with biotinylated cDNA probes and Southern blot hybridization of *ERCC6* probes to DNA of somatic cell hybrids. Furthermore, an RFLP that can be used for linkage analysis has been identified within the *ERCC6* locus.

MATERIALS AND METHODS

In situ hybridization. Treatment of human lymphocyte metaphase spreads prior to hybridization was as described by Weeda *et al.* (1991).

[1] Present address: Department of Internal Medicine, AZU, Utrecht, The Netherlands.

[2] To whom correspondence should be addressed.

0888-7543/92 $3.00

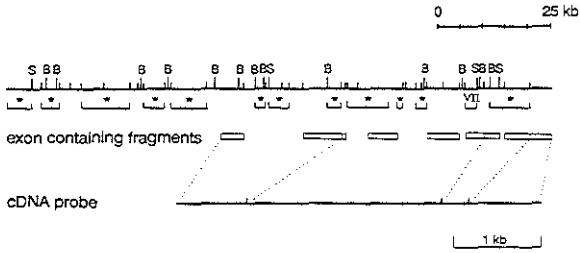


FIG. 1. Restriction map of the human *ERCC6* locus. The cDNA probe used for *in situ* hybridizations and the genomic probe VII (see Troelstra *et al.*, 1990) used on Southern blots are shown. Exon containing fragments recognized by the cDNA probe are indicated. Short bars represent *Eco*RI restriction sites. Symbols: B, *Bam*HI; S, *Sal*I; *, repeat containing fragment; VII, genomic probe VII.

In situ hybridization experiments using the *ERCC6* cDNA fragment shown in Fig. 1 or cDNAs extending more 5', a chromosome 8-specific marker (thyroglobulin gene), and the centromere-specific marker D10Z1 as biotin-labeled probes were performed as described elsewhere (Pinkel *et al.*, 1986). Both probe and target DNA were denatured at 80°C for 5 min. Hybridization was 16 h at 37°C in 50% formamide, 2× SSC, 40 m*M* sodium phosphate (pH 7.0), 10% dextran sulfate, 100 ng sonicated salmon sperm DNA, 100 ng yeast tRNA, and 20 ng labeled probe. The slides were washed with 50% formamide, 2× SSC, pH 7, at 45°C followed by 4× SSC plus 0.05% Tween 20 at room temperature. Slides were incubated with 5 μg/ml avidin D–FITC (Vector, U.S.A.), and the fluorescent signal was amplified with biotinylated goat anti-avidin D, washed, dehydrated with ethanol, and air-dried. The slides were either counterstained with propidium iodide in antifade medium or banded with 4',6'-diamidino-2-phenylindole (DAPI) and actinomycin D.

Cell lines. Somatic cell hybrids CY18 and CY5 are mouse–human hybrid cell lines containing a single human chromosome 16 and a der 10t(10;16)(q26;q22) translocation chromosome, respectively, as their only human material (Callen, 1986). The original mouse cell line used for fusion was GM346A, a HPRT- and APRT-deficient mouse L-cell.

Southern blot analysis. Digestion of DNA with the indicated restriction enzymes, gel electrophoresis, labeling of DNA probes, and hybridizations were performed using routine procedures as described (Maniatis *et al.*, 1982). Southern blotting to Zeta probe blotting membranes was performed by alkaline transfer, as described by the manufacturer (Bio-Rad, Richmond, CA).

RESULTS

In Situ Hybridization

To map the *ERCC6* locus, *in situ* hybridization was carried out using biotinylated *ERCC6* cDNA probes. The 3.6-kb probe (Fig. 1) represents the 3' half of the smallest of the two *ERCC6* transcripts of about 5 and 7.5 kb detectable on Northern blots (unpublished results). This 3.6-kb cDNA fragment covers 75 kb of the *ERCC6* locus. Also, longer cDNA probes, extending more 5', have been used in *in situ* hybridizations. Two representative *in situ* hybridizations are shown in Fig. 2. Although the chromosome banding is not as detailed as that in a routine G-banding procedure, the hybridizing chromosome can be identified as human chromosome 8 or 10, in the area close to the centromere. *In situ* hybridization, simultaneously using the *ERCC6* cDNA probe and a probe specific for either chromosome 8 (a single-copy probe from the thyroglobulin gene) or chromosome 10 (an alpha satellite DNA probe, D10Z1), clearly demonstrated the *ERCC6* locus to be on chromosome 10q11–q21 (results not shown).

Hybridization to DNA from Somatic Cell Hybrids

To confirm the assignment of the *ERCC6* gene to chromosome 10, two human × mouse somatic cell hybrids (CY5 and CY18) were used. Hybrid CY5 contains almost the complete chromosome 10 and part of chromosome 16 as the only human chromosomes; in hybrid CY18 only human chromosome 16 is present (Callen, 1986). The probe used in these experiments was a unique 2.2-kb human genomic DNA fragment (probe VII, see Fig. 1) that recognizes a *Taq*I fragment of 3.2 kb in HeLa DNA (Fig. 3). In DNAs of both hybrids and mouse cells, a cross-hybridizing fragment (4.3 kb) from the mouse *ERCC6* homologue is detected, indicating that the probe contains a conserved sequence. In CY5 the human *ERCC6* *Taq*I fragment (3.2 kb) is present (Fig. 3). In hybrid CY18 no human fragment is detected, meaning that the *ERCC6* locus segregates with human chromosome 10, completely in accordance with the results from the *in situ* hybridization.

Restriction Fragment Length Polymorphism in the ERCC6 Gene

The genetic defect in multiple endocrine neoplasia type 2 (MEN2) has been mapped to the pericentromeric region of chromosome 10 (Mathew *et al.*, 1987; Simpson *et al.*, 1987; Norum *et al.*, 1990; Wu *et al.*, 1990; Lairmore *et al.*, 1991). The *ERCC6* gene might therefore reside in the vicinity of the MEN2 locus. This prompted us to look for RFLPs within the *ERCC6* region. DNAs of 30

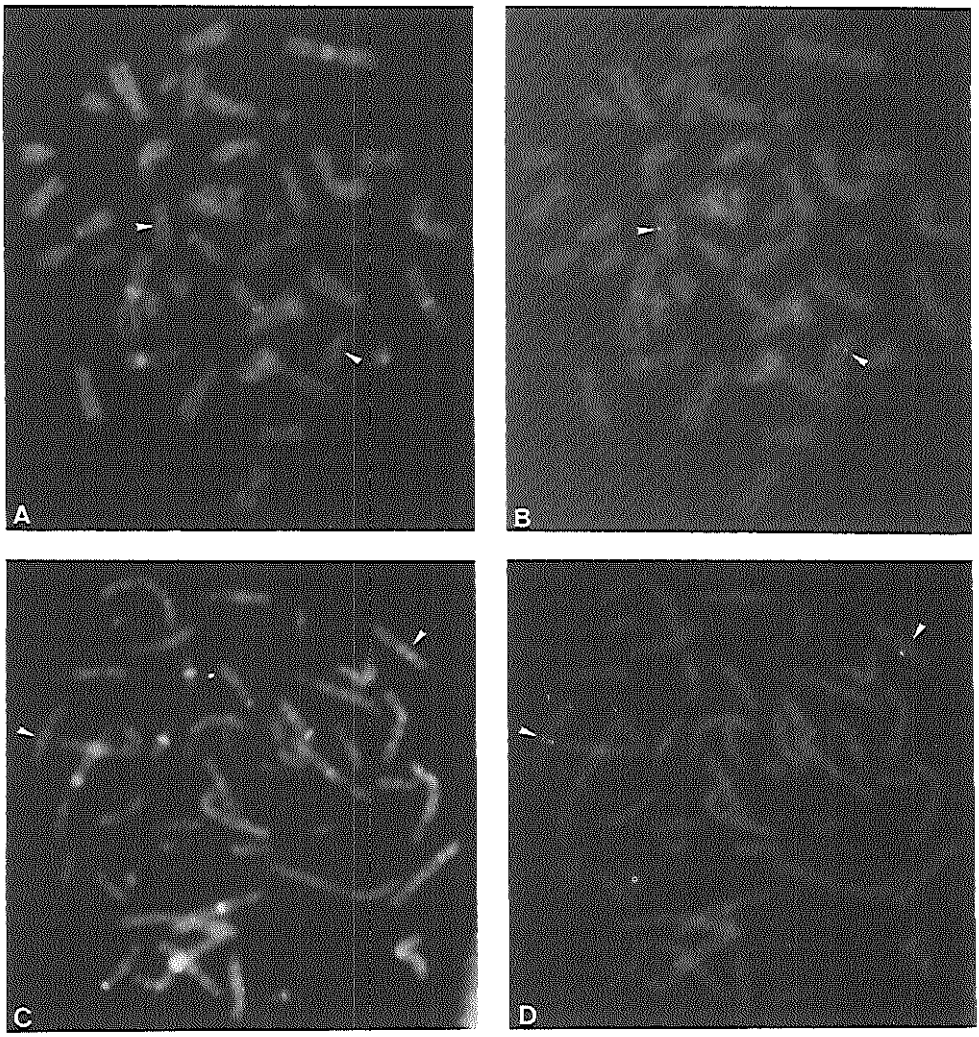


FIG. 2. Localization of *ERCC6* by *in situ* hybridization. (A, C) Karyotypes of human lymphocyte metaphase spreads, stained with DAPI; arrowheads indicate chromosome 10. (B, D) Photographs showing the corresponding fluorescent *in situ* hybridization with the human *ERCC6* cDNA fragment as a probe. Arrowheads indicate the fluorescent label on chromosome 10.

unrelated European Caucasians were digested with *Ban*I, *Bgl*II, *Eco*RI, *Hind*III, *Msp*I, *Pst*I, *Taq*I, and *Xba*I. Southern blots were hybridized to various *ERCC6* probes and checked for the presence of polymorphisms. Using probe 4J.ES3 (probe VII in Fig. 1, results shown from analysis of a family) *Taq*I identified a two-allele polymorphism (A1 = 10.3 kb, A2 = 3.2 kb) without a constant band (Fig. 4, individuals homozygous for A1 not shown). Based on the analysis of DNA from these 30 unrelated European Caucasians, an allelic frequency of

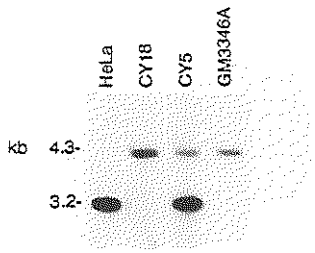


FIG. 3. Localization of *ERCC6* by hybridization to somatic cell hybrids. Southern blot analysis of genomic DNA (20 μg) from HeLa, somatic cell hybrids CY18 and CY5, and mouse cell line GM346A, digested with *TaqI* and hybridized with genomic probe VII (Fig. 1).

0.42 for A1 and 0.58 for A2 can be calculated. No RFLPs were detected with this probe in the other digests tested.

DISCUSSION

By means of *in situ* hybridization and analysis of somatic cell hybrids, we have assigned the *ERCC6* locus to human chromosome 10, region q11–q21. The *ERCC6* gene product participates in the nucleotide excision repair pathway, which removes a wide range of potentially mutagenic and carcinogenic lesions in the DNA. As such, the *ERCC6* protein is involved in the prevention of carcinogenesis. Whether mutations or deletions in the *ERCC6* locus are of any importance for development of specific tumors is unknown. It is, however, interesting to note that during progression from astrocytoma to glioblastoma (part of) chromosome 10 is frequently found to be lost (James *et al.*, 1988).

Several other human DNA repair genes have been assigned to various chromosomes (Mohrenweiser *et al.*, 1989; Smeets *et al.*, 1990; Weeda *et al.*, 1991; Siciliano *et al.*, 1986; Thompson *et al.*, 1987; Kaur and Athwal, 1989; Ishizaki *et al.*, 1990). A remarkable clustering of repair genes occurs on chromosome 19, in the q13.2–q13.3 area. This chromosome contains *ERCC1*, *ERCC2*, *XRCC1* (*X*-ray-repair cross complementing rodent repair deficiency gene) and ligase I. *ERCC1* and *ERCC2* have been shown to be separated by less than 250 kb (Mohrenweiser *et al.*, 1989; Smeets *et al.*, 1990). *ERCC6* is the second DNA repair gene assigned to chromosome 10, the other being a gene encoding a human methyltransferase (Rydberg *et al.*, 1990).

Previous reports have mapped the genetic defect in the MEN2 syndrome to the pericentromeric region of chromosome 10, possibly in the vicinity of the *ERCC6* locus. Two clinical types of MEN2 exist: MEN2A and MEN2B. By linkage analysis, both have been located more precisely: to a small region (~11 cM) around the centromere, bordered by FNRB (at 10p11.2) and RBP3 (at 10q11.2), tightly linked to D10Z1 (Norum *et al.*, 1990; Wu *et al.*, 1990, Lairmore *et al.*, 1991). Polymorphic DNA markers that are closely linked to the gene for a genetic disease have proved to be of great value for determining the genotypes at the disease locus of members of afflicted families. In addition, such markers are a valuable tool in the molecular cloning of the disease gene. Although it remains to be determined how closely linked the *TaqI* RFLP reported here and the MEN2 locus are, the RFLP might be of importance in future MEN2 studies.

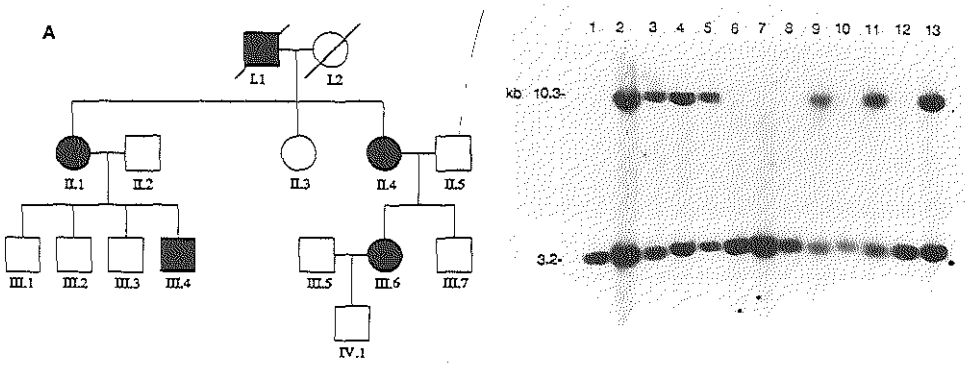


FIG. 4. Detection of an RFLP within the *ERCC6* gene. (A) Pedigree of the MEN2A family used in B. (B) Southern blot analysis of genomic DNA (20 μg) of 13 family members digested with *TaqI* and hybridized with the genomic probe VII (Fig. 1) detecting the A1 and A2 alleles indicated. Lane 1, II.1; lane 2, II.2; lane 3, III.1; lane 4, III.2; lane 5, III.3; lane 6, III.4; lane 7, II.3; lane 8, II.4; lane 9, II.5; lane 10, III.6; lane 11, III.5; lane 12, IV.1; lane 13, III.7.

ACKNOWLEDGMENTS

We are grateful to Dr. D. Bootsma for helpful suggestions and critical reading of the manuscript. We thank M. Kuit for photography. This study was supported in part (R.M.L., G.V., and C.H.C.M.B.) by the Dutch Prevention Fund.

Note added in proof. After the acceptance of this manuscript a clinical report appeared (Fryns *et al.* (1991) *Am. J. Med. Genet.* **40:** 343–344) presenting a Cockayne syndrome patient with an interstitial deletion of the long arm of chromosome 10 (del(10)(q11.23q21.2)). By Southern blot analysis of DNA from this patient, we have obtained indications that one copy of the *ERCC6* gene is missing in DNA of this patient. This makes *ERCC6* a candidate gene for the excision repair disorder Cockayne syndrome.

REFERENCES

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

Callen, D. F. (1986). A mouse–human hybrid cell panel for mapping human chromosome 16. *Ann. Genet. (Paris)* **29:** 235–239.

Cleaver, J. E., and Kraemer, K. H. (1989). Xeroderma pigmentosum. *In* "The Metabolic Basis for Inherited Disease" (A. L. Beaudet, W. S. Sly, and D. Valle, Eds.), Vol. 2, pp. 2949–2971, McGraw-Hill, New York.

Collins, A., and Johnson, R. T. (1987). DNA repair mutants in higher eukaryotes. *J. Cell Sci. Suppl.* **6:** 61–82.

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.

Friedberg, E. C. (1985). "DNA Repair," Freeman, San Francisco.

Ishizaki, K., Oshimura, M., Sasaki, M. S., Nakamura, Y., and Ikenaga, M. (1990). Human chromosome 9 complements UV sensitivity of xeroderma pigmentosum group A cells. *Mutat. Res.* **235:** 209–215.

James, C. D., Carlbom, E., Dumanski, J. P., Hansen, M., Nordenskjold, M., Collins, V. P., and Cavenee, W. K. (1988). Clonal genomic alterations in glioma malignancy stages. *Cancer Res.* **48:** 5546–5551.

Kauer, G. P., and Athwal, R. S. (1989). Complementation of a DNA repair defect in xeroderma pigmentosum cells by transfer of human chromosome 9. *Proc. Natl. Acad. Sci. USA* **86:** 8872–8876.

Lairmore, T. C., Howe, J. R., Korte, J. A., Dilley, W. G., Aine, L., Aine, E., Wells, S. A., Jr., and Donis-Keller, H. (1991). Familial medullary thyroid carcinoma and multiple endocrine neoplasia type 2B map to the same region of chromosome 10 as multiple endocrine neoplasia type 2A. *Genomics* **9:** 181–192.

Lehmann, A. R. (1982). Three complementation groups of Cockayne's syndrome. *Mutat. Res.* **106:** 347–356.

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

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mathew, C. G. P., Chin, K. S., Easton, D. F., Thorpe, K., Carter, C., Liou, G. I., Fong, S. L., Bridges, C. D. B., Haak, H., Niewenhuizen Kruseman, A. C., Schifter, S., Hansen, H. H., Telenius, H., Telenius-Berg, M., and Ponder, B. A. J. (1987). A linked genetic marker for multiple endocrine neoplasia type 2A on chromosome 10. *Nature* **328:** 527–528.

Mohrenweiser, H. W., Carrano, A. V., Fertitta, A., Perry, B., Thompson, L. H., Tucker, J. D., and Weber, C. A. (1989). Refined mapping of the three DNA repair genes *ERCC1*, *ERCC2*, and *XRCC1* on chromosome 19. *Cytogenet. Cell Genet.* **52:** 11–14.

Norum, R. A., Lafreniere, R. G., O'Neal, L. W., Nikolai, T. F., Delaney, J. P., Sisson, J. C., Sobol, H., Lenoir, G. M., Ponder, B. A. J., Willard, H. F., and Jackson, C. E. (1990). Linkage of the multiple endocrine neoplasia type 2B gene (MEN2B) to chromosome 10 markers linked to MEN2A. *Genomics* **8:** 313–317.

Pinkel, D., Straume, T., and Gray, J. (1986). Cytogenetic analysis using quantitative highly sensitive, fluorescence hybridization. *Proc. Natl. Acad. Sci. USA* **83:** 2934–2938.

Rydberg, B., Spurr, N., and Karran, P. (1990). cDNA cloning and chromosomal assignment of the human O^6-methylguanine-DNA methyltransferase: cDNA expression in *Escherichia coli* and gene expression in human cells. *J. Biol. Chem.* **265:** 9563–9569.

Siciliano, M. J., Carrano, A. V., and Thompson, L. H. (1986). DNA repair gene on chromosome 16 and a third repair gene on chromosome 19. *Am. J. Hum. Genet.* **39:** A42.

Simpson, N. E., Kidd, K. K., Goodfellow, P. J., McDermid, H., Myers, S., Kidd, J. R., Jackson, C. E., Duncan, A. M. V., Farrer, L. A., Brasch, K., Castiglione, C., Genel, M., Gertner, J., Greenberg, C. R., Gusella, J. F., Holden, J. J. A., and White, B. N. (1987). Assignment of multiple endocrine neoplasia type 2A to chromosome 10 by linkage. *Nature* **328:** 528–530.

Smeets, H., Bachinsky, L., Coerwinkel, M., Schepens, J., Hoeijmakers, J. H. J., van Duin, M., Grzeschik, K-H., Weber, C. A., de Jong, P., Siciliano, M. J., and Wieringa, B. (1990). A long-range restriction map of the human chromosome 19q13 region: Close physical linkage between CKMM and the *ERCC1* and *ERCC2* genes. *Am. J. Hum. Genet.* **46:** 492–501.

Thompson, L. H., Carrano, A. V., Sato, K., Salazar, E. P., White, B. F., Stewart, S. A., Minkler, J. L., and Siciliano, M. J. (1987). Identification of nucleotide-excision repair genes on human chromosomes 2 and 13 by functional complementation in hamster–human hybrids. *Somat. Cell Mol. Genet.* **13:** 539–551.

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. (1988). CHO mutant UV61 removes (6–4) photoproducts but not cyclobutane dimers. *Mutagenesis* **4:** 140–146.

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 *ERCC6*. *Mol. Cell. Biol.* **10:** 5806–5813.

Venema, J., Mullenders, L. H. F., Natarajan, A. T., van Zeeland, A. A., and Mayne, L. (1990). The genetic defect in Cockayne's 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.

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

Weeda, G., van Ham, R., C., A., Vermeulen, W., Bootsma, D., van der Eb, A. J., and Hoeijmakers, J. H. J. (1990). A presumed DNA helicase, encoded by the excision repair gene *ERCC3* is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. *Cell* **62:** 777–791.

Weeda, G., Wiegant, J., van der Ploeg, M., Geurts van Kessel, A. H. M., van der Eb, A. J., and Hoeijmakers, J. H. J. (1991). Localization of the xeroderma pigmentosum group B-correcting gene *ERCC3* to human chromosome 2q21. *Genomics* **10:** 1035–1040.

Wu, J., Carson, N. L., Myers, S., Pakstis, A. J., Kidd, J. R., Castiglione, C. M., Anderson, L., Hoyle, L. S., Genel, M., Verdy, M., Jackson, C. E., Simpson, N. E., and Kidd, K. K. (1990). The genetic defect in multiple endocrine neoplasia type 2A maps next to the centromere of chromosome 10. *Am. J. Hum. Genet.* **46:** 624–630.

Chapter V

ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes

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 $UvrA_2B$ 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-

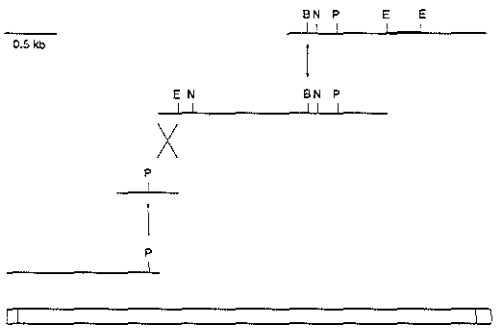


Figure 1. Construction of a Functional *ERCC6* cDNA

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

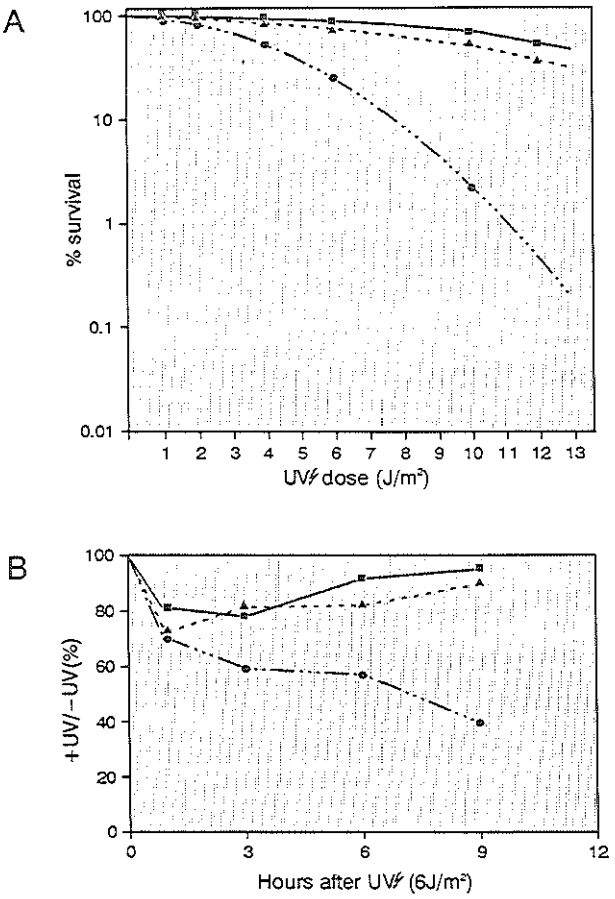


Figure 2. Correction of the UV61 Repair Defect by *ERCC6* cDNA

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

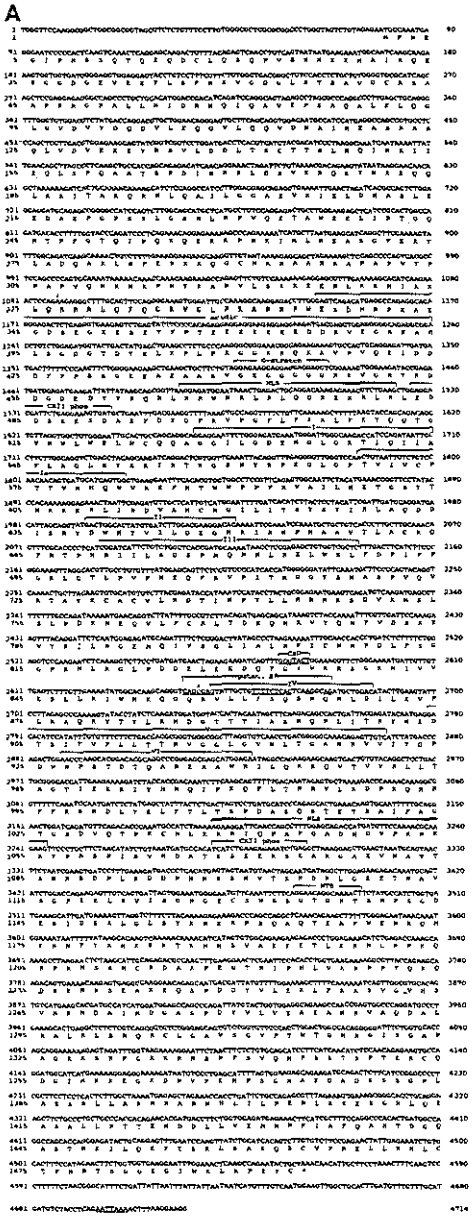
(B) Recovery of RNA synthesis in wild-type cell line AA8, mutant UV61, and *ERCC6* cDNA transformant 61E6a. Cells unirradiated or UV irradiated with 6 J/m² were pulse labeled with [³H]uridine at different times after irradiation; the incorporated radioactivity was determined by liquid scintillation counting (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 [³H]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 correcting 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-



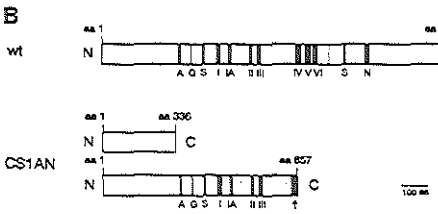


Figure 3. Nucleotide and Predicted Amino Acid Sequence of the *ERCC6* cDNA

(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.

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 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

(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/m^2, 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*

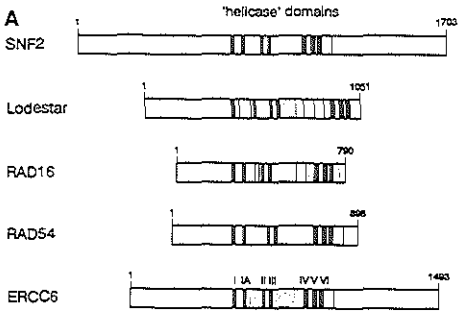
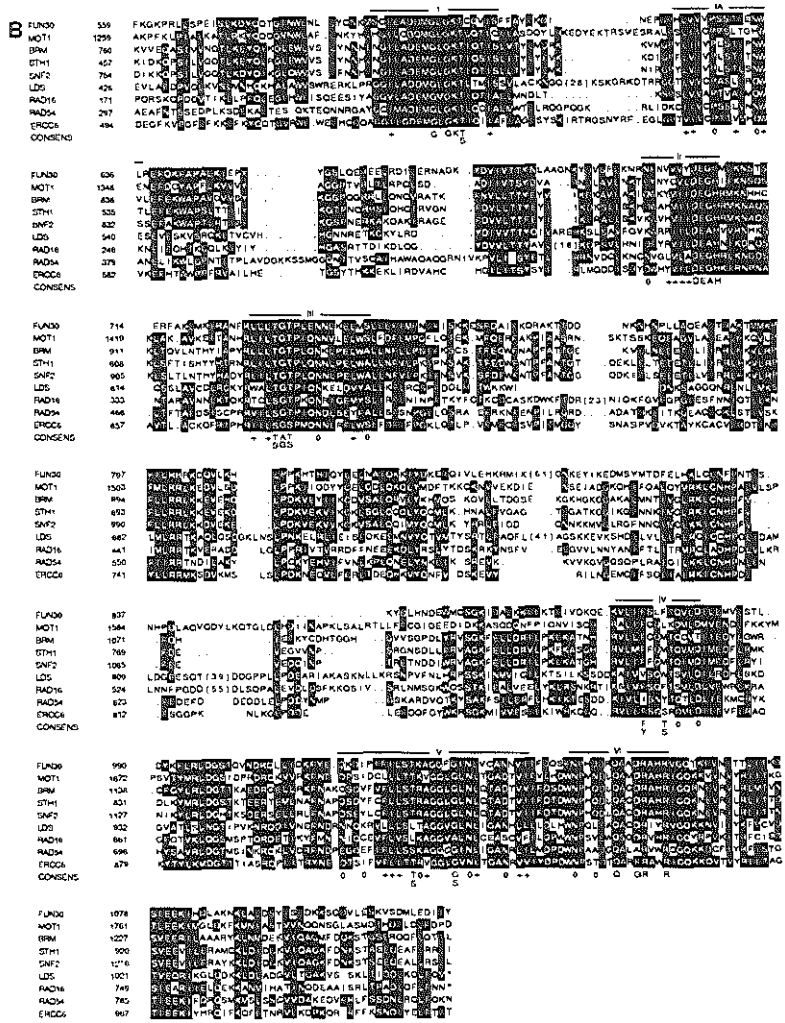


Figure 4. Homology of the ERCC6 Helicase Region to the Similar Region of RAD54, RAD16, lodestar, SNF2, STH1, brm, MOT1, and FUN30

(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 superfamilies 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).



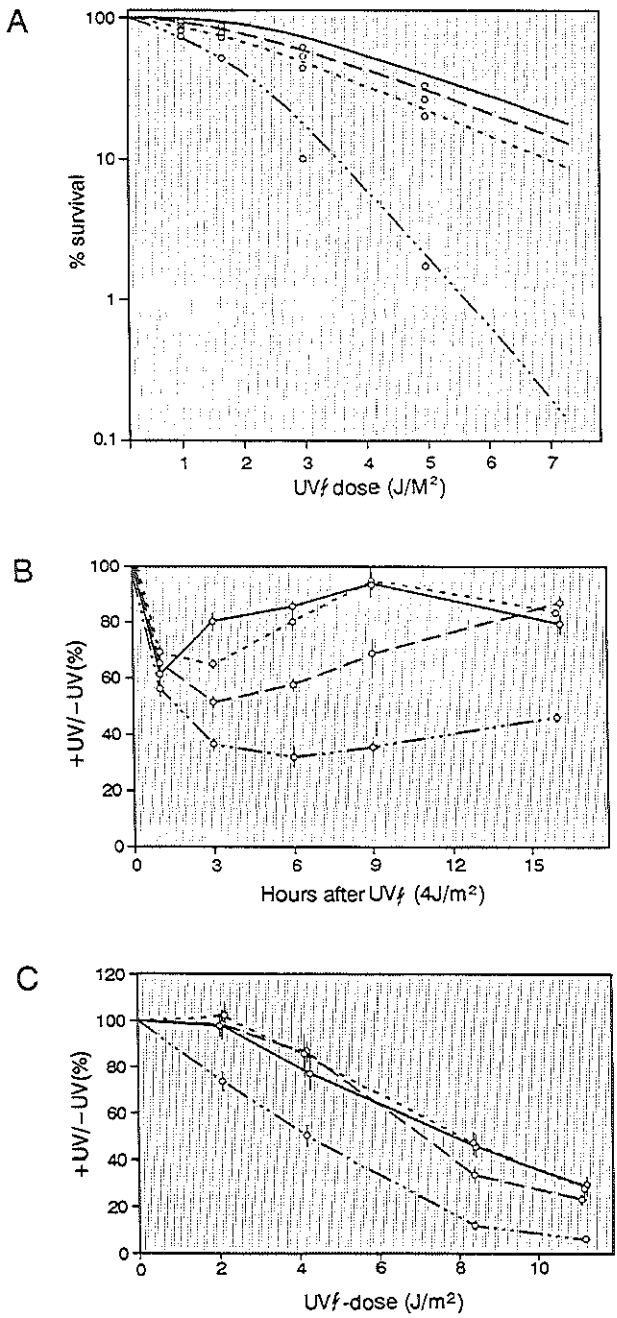


Figure 5. Correction of the CS1AN.S3.G2 Repair Defect by *ERCC6* cDNA

(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).

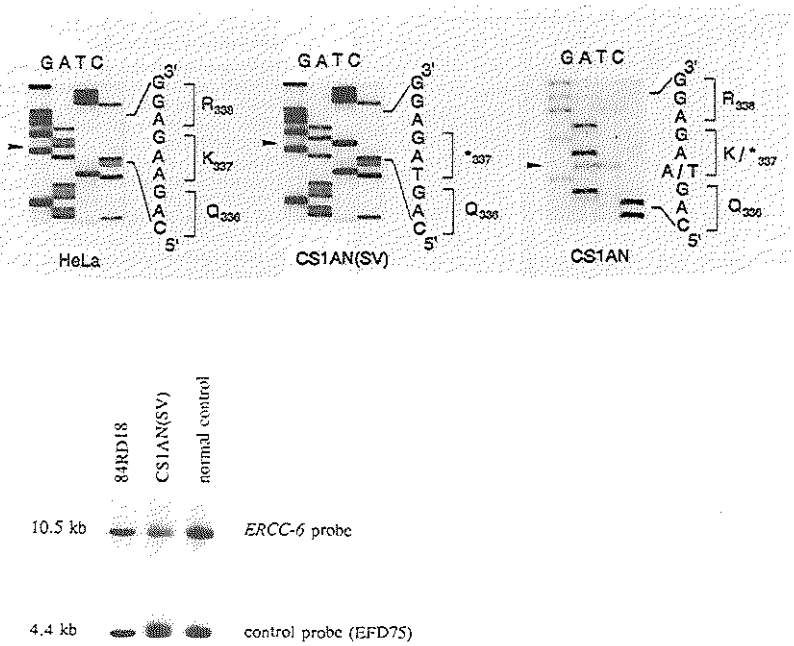


Figure 6. *ERCC6* Mutation in CS1AN.S3.G2 and Allele a of Patient CS1AN

(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^{2646} to T transition 30 nt upstream of the normal

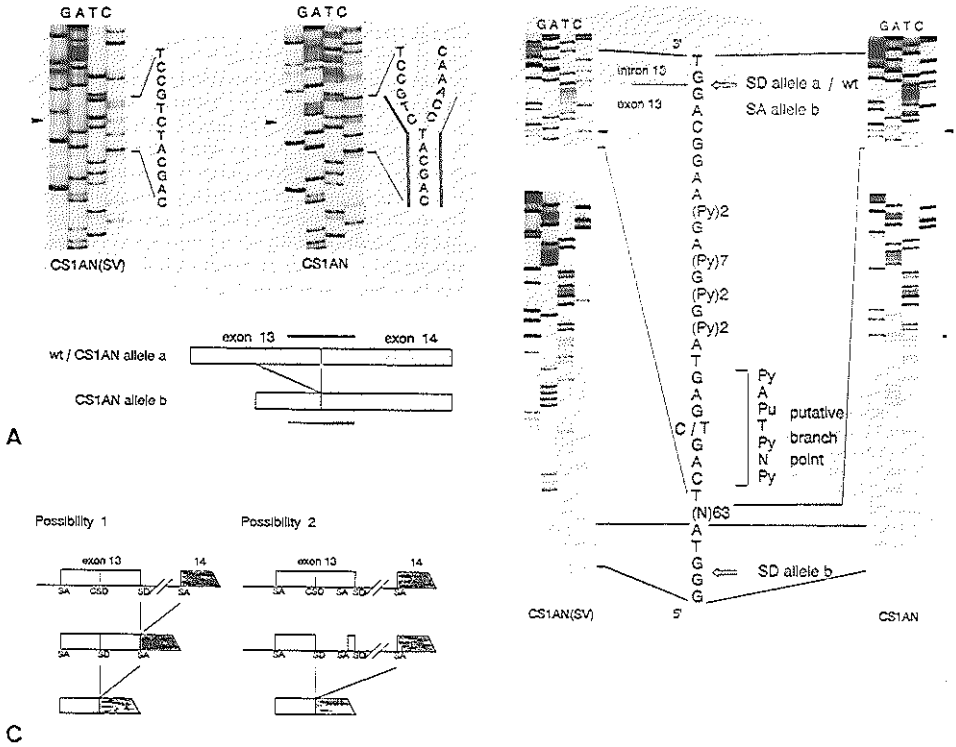


Figure 7. *ERCC6* Mutation in Allele b of Patient CS1AN

(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 D_{10})	≈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	−[h,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 $UvrA_2B$ 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

$UvrA_2B$ 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 × 10^2 to 5 × 10^5 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 [^{14}C]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 [^{3}H]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. ^{14}C 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 [^{3}H]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 $MgCl_2$, 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 $MgCl_2$, 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.

Acknowledgments

We are very grateful to Dr. A. Bernards and Dr. H. Okayama for providing us with cDNA libraries, to Dr. D. Busch for two new and unpublished group 6 mutants, and to Dr. A. R. Lehmann and Dr. L. V. Mayne for some of the CS lines. We thank Dr. D. Schild for sharing with us unpublished information. Furthermore, we greatly acknowledge the help of G. Moolhuizen in the construction of the cDNA expression plasmid and the assistance of Dr. N. G. J. Jaspers in RNA synthesis

recovery experiments. We thank S. van Baal for his assistance in computer work, P. van der Spek for the valuable use of his computer program that pointed out the presence of the helicase domains, and M. Kuit for photography. This research was supported in part by Euratom (contract number BJ6-141-NL).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received July 20, 1992; revised September 25, 1992.

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.

Fleijter, 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 *Iodestar*, 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

The accession number for the sequence reported in this paper is L04791.

Chapter VI

Structure and expression of the excision repair gene ERCC6, involved in the human disorder Cockayne's syndrome group B

Structure and expression of the excision repair gene *ERCC6*, involved in the human disorder Cockayne's syndrome group B

Christine Troelstra, Wouter Hesen, Dirk Bootsma, and Jan H. J. Hoeijmakers
MGC, Department of Cell Biology and Genetics,
Erasmus University, Rotterdam, The Netherlands

(Accepted for publication in *Nucleic Acids Research*)

SUMMARY

The human repair gene *ERCC6* - a presumed DNA (or RNA) helicase - has recently been found to function specifically in preferential nucleotide excision repair (NER). This NER subpathway is primarily directed towards repair of (the transcribed strand of) active genes. Mutations in the *ERCC6* gene are responsible for the human hereditary repair disorder Cockayne's syndrome complementation group B, the most common form of the disease. In this report, the genomic organization and expression of this gene are described. It consists of at least 21 exons, together with the promoter covering a region of 82 - 90 kb on the genome. Postulated functional domains deduced from the predicted amino acid sequence, including 7 distinct helicase signatures, are - with one exception - encoded on separate exons. Consensus splice donor and acceptor sequences are present at all exon borders with the exception of the unusual splice donor at the end of exon VII. The "invariable" GT dinucleotide in the consensus (C,A)AG / GTPuAGT is replaced by the exceptional GC. Based on 42 GC splice donor sequences identified by an extensive literature search we found a statistically highly significant better "overall" match of the surrounding nucleotides to the consensus sequence compared to normal GT-sites. This confirms and extends the observation made recently by Jackson (*Nucl. Acids Res.*, 19, 3795-3798 (1991)) derived from analysis of 26 cases. Analysis of *ERCC6* cDNA clones revealed the occurrence of alternative polyadenylation, resulting in the (differential) expression of two mRNA molecules (which are barely detectable on Northern blots) of 5 and 7 kb in length.

INTRODUCTION

Nucleotide excision repair (NER) is one of the major DNA repair systems functioning in mammalian cells (Friedberg, 1985). It is able to remove a broad variety of DNA lesions, such as UV-induced pyrimidine dimers and (6-4) photoproducts, as well as bulky chemical lesions and DNA cross links. After induction of DNA damage, the NER machinery appears to be directed first primarily towards the transcribed regions in the genome. This enables the cells to rapidly resume the vital process of transcription, that otherwise would remain blocked by lesions in the DNA template (Bohr, et al., 1985). This preferential repair of active genes is even specific for the transcribed strand (Mellon, et al., 1987). Removal of lesions from the non-transcribed strand, as well as from inactive chromatin proceeds more slowly and appears to be incomplete for several lesions (Link Jr., et al., 1991 for a recent review). In human cells, repair of the non-transcribed strand is more rapid than that of inactive genes (Venema, et al., 1992).

In the rare hereditary disorders xeroderma pigmentosum (XP) and Cockayne's syndrome (CS), deficiencies in NER are thought to underlie the clinical symptoms. XP patients are clinically characterized by severe sun(UV)sensitivity of the skin, pigmentation abnormalities, a highly elevated risk for developing skin tumors, and often neurological degeneration (Cleaver and Kraemer, 1989). Cell fusion experiments with cells from these patients have distinguished 7 NER-deficient complementation groups (cg; 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 (Shivji, et al., 1992). For XP-C cells the defect appears to be specific for repair of inactive chromatin; these cells have retained the ability to remove damage from the transcribed strand of active genes (Kantor, et al., 1990; Venema, et al., 1990b; Venema, et al., 1991). CS patients have, in sharp contrast to XP, no pigmentation abnormalities and no elevated risk for skin tumor formation. They are clinically characterised by skeletal and retinal abnormalities, growth retardation, progressive neurological degeneration, and a sunsensitive skin (Lehmann, 1987; Nance and Berry, 1992 for reviews). Like XP, CS is heterogeneous: 2 cg have been identified (CS-A, -B; Tanaka, et al., 1981; Lehmann, 1982). The molecular defect in CS-A and -B was shown to affect preferential repair of active genes only, whereas repair of the inactive chromatin regions proceeds normally (Venema, et al., 1990a). A third complementation group was defined (CS-C), containing one patient that exhibited a combination of characteristics of XP and CS. This patient is also classified as XP-B (Robbins, et al.,

1974; Lehmann, 1982). Several patients exhibiting this combined XP/CS phenotype are reported; one of these has been assigned to XP group D (Vermeulen, et al., 1991).

Another class of mammalian NER-deficient mutants consists of laboratory-induced, UV-sensitive, rodent (mainly chinese hamster) mutant cell lines. Among these, at least 10 cg have been defined (Thompson, et al., 1988; Zdzienicka, et al., 1988; Busch, et al., 1989; Stefanini, et al., 1991). A number of rodent mutants have been successfully used for isolation of human genes capable of correcting the rodent repair deficiency (Hoeijmakers and Bootsma, 1990 for a recent review). Subsequent introduction of these genes into XP and CS cells has demonstrated an overlap between the hamster and human cg: rodent cg 2 and 3, corrected by *ERCC2* and *-3*, are the rodent equivalents of XP-D and XP-B, respectively (Weeda, et al., 1990; Flejter, et al., 1992). Recently, we have isolated the human repair gene *ERCC6* by virtue of its ability to correct the UV-sensitivity of the Chinese hamster ovary (CHO) mutant cell line UV61, cg 6 (Troelstra, et al., 1990). The gene encodes a protein of 1493 amino acids; the predicted amino acid sequence suggests that the *ERCC6* gene product is a nuclear DNA helicase. Mutations in the *ERCC6* gene appear to be responsible for the repair deficiency of CS-B cells, implying that the *ERCC6* gene product is specifically involved in the process of preferential repair (Troelstra, et al., 1992b). This paper describes the genomic architecture and expression of the *ERCC6* gene.

MATERIALS AND METHODS

General procedures

Purification of nucleic acids, restriction enzyme digests, gel electrophoresis, DNA ligation, synthesis of radiolabeled probes using random oligonucleotide primers, the polymerase chain reaction (pcr), sequence analysis (dideoxy-mediated chain-termination), and filter hybridization were performed according to established procedures (Sambrook, et al., 1989).

Construction of the cDNA plasmid

Construction of the (almost) full length *ERCC6* cDNA was as described (Troelstra, et al., 1992b). The cDNA insert was subcloned into both the vector pTZ19R (Pharmacia) yielding pTZE6total, and the mammalian expression vector pSLM, a derivative of pSVL (Pharmacia; Van Duin, unpublished results). In the mammalian expression vector the *ERCC6* cDNA is under the control of the SV40 late promoter, whereas in pTZE6total the cDNA is not preceded by a eukaryotic promoter.

Cell culture, transfection, and selection

UV-sensitive CHO cell line UV61 and wt CHO cell line AA8 were grown in 1:1 F10-Dulbecco minimal essential medium supplemented with antibiotics and 8% fetal calf serum. The *ERCC6* cDNA construct pTZE6total (3 μg) was cotransfected with 2 μg pSV2neo and 20 μg lambda phage 6B (see Figure 1) to UV61 cells. In order to release the lambda arms and plasmid vector from the inserts prior to transfection, both pTZE6total and lambda phage 6B were digested with *Sal* I; an enzyme which does not cut within the insert. Transfection and selection were performed as described before (Troelstra, et al., 1992b).

Identification of intron-exon borders

All genomic fragments hybridizing to the *ERCC6* cDNA were subcloned in pTZ19R (Pharmacia) or pBluescript II KS (Stratagene), and sequenced with *ERCC6*-specific primers. All sequence reactions were performed on double-stranded templates by the dideoxy chain termination method, using T7 DNA polymerase (Pharmacia). Intron length was determined either by restriction enzyme digestions and subsequent Southern blot analysis or by pcr with exon-specific primers on subcloned genomic fragments.

Northern blot analysis

RNA samples were separated on an agarose gel and transferred to a nylon membrane (Zeta probe from Bio-Rad) as described by Fourney et al. (Fourney, et al., 1988). The filters were hybridized at 65°C, in 3xSSC, 10xDenhardt's reagent, 0.1% SDS, 90 μg/ml dextran sulfate, and 50 μg/ml denatured salm sperm DNA. Filters were washed 2 times (10 min.) in 3 x SSC, 0.1% SDS and once (10 min.) in 1 x SSC, 0.1% SDS, at 65°C.

Synthesis of strand-specific probes

Strand-specific probes were synthesized according to the method described by Espelund et al. (Espelund, et al., 1990), with several modifications. The template was generated by pcr, with one normal and one biotinylated primer. The biotinylated product was then bound to streptavidin-coated magnetic beads (Dynabeads M-280, Dynal) through a 30 min. incubation at 37°C in 5xSSPE (20xSSPE: 3.6M NaCl, 200mM NaH_2PO_4 pH7.4, 20mM EDTA pH7.4). The beads were washed 4x with a solution containing 0.17% (w/v) Triton X-100, 100 mM NaCl, 10 mM Tris.HCl pH7.5, and 1 mM EDTA; subsequently the non-biotinylated strand was removed by 2 cycles (7 min. each) of denaturation in 125 mM NaOH, 100 mM NaCl. After washing twice with 100 mM Tris.HCl pH7.6, 150 mM NaCl, the biotinylated strand

is radioactively labelled via primer extension. The beads were washed 3x, and the DNA was denatured by two incubations (7 min. each) in 125 mM NaOH, 100 mM NaCl. The supernatant was neutralized with an equal volume of 1M Tris.HCl (pH7.5), and used in hybridizations.

RESULTS

Architecture of the *ERCC6* gene

ERCC6 promoter region

The *ERCC6* gene was isolated from a lambda EMBL3 library originating from a repair-proficient secondary UV61 transformant. The genomic region coinherited by independent, repair-proficient primary and secondary UV61 transformants, as judged by Southern blot analysis, was approximately 100 kb, and consequently separated over several lambda clones (Troelstra, et al., 1990). A physical map of this region is shown

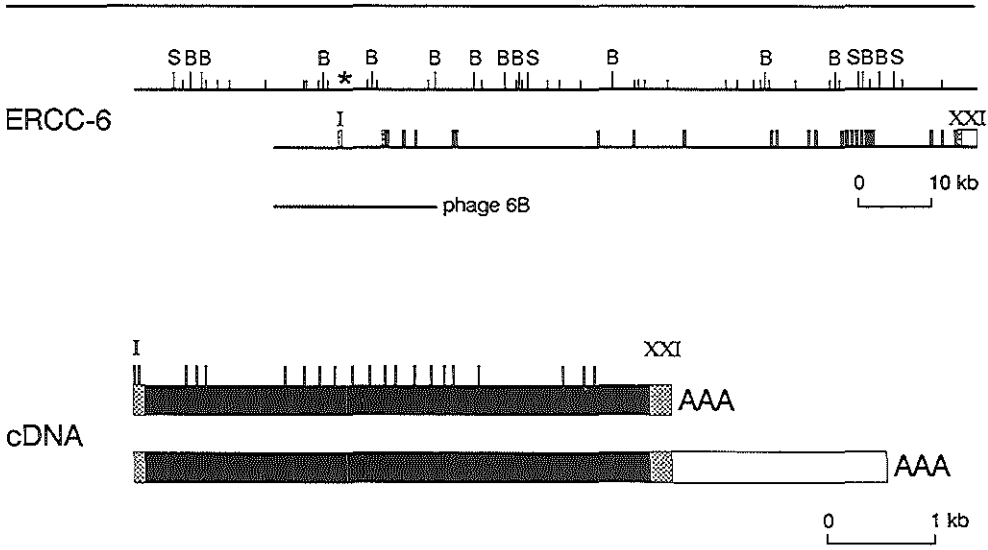


Figure 1. Genomic organization of *ERCC6*. A physical map of the *ERCC6* locus is shown. Exons (I to XXI) are indicated by boxes on the bar underneath the map. cDNAs from the two encoded, differentially poly-adenylated mRNAs are shown; white and hatched areas represent untranslated regions, the black area represents the ORF. Exon borders in the cDNA are indicated by vertical lines on top of the shortest cDNA. Phage 6B: a λ-phage containing the promoter region of the *ERCC6* gene. Symbols: ●, a CpG-island encompassing a *Bss* HI, *Not* I, and a *Sac* II restriction enzyme site located just 3' of exon I; S, *Sal* I; B, *Bam* HI; short bars in the map represent *Eco* RI restriction sites.

in Figure 1. The size of the *ERCC6* locus was determined by hybridization of an (almost) full length, functional cDNA of 4.7 kb, encoding at least the total open reading frame and the 3' end (of the shortest mRNA, see below), to the different lambda clones. The cDNA, previously localized on chromosome 10q11-21 (Troelstra, et al., 1992a), was shown to encompass 82 kb at the genome level (Figure 1) and to reside on a 430 kb *Not* I fragment (data not shown). One of the *Not* I sites is part of a CpG-island present in the 5' region of the *ERCC6* gene (Figure 1). To pinpoint the region containing the promoter and the 5' untranslated end of the mRNA (part of which might be lacking in the cDNA), a promoterless cDNA construct (pTZE6total, see Materials and Methods) was cotransfected with lambda clone 6B, containing the first 4 exons present in the cDNA, and approximately 8 kb more upstream sequences (Figure 1). To be expressed, the promoterless cDNA should recombine within the cell with the cotransfected lambda phage, thus producing a "mini gene". After transfection of the two DNAs together UV-resistant clones were obtained, with an efficiency that was $\approx$ 10x lower than after transfection of a eukaryotic expression vector carrying the *ERCC6* cDNA behind the SV40 late promoter. The UV-sensitivity was corrected to a level within wt range (Figure 2), as determined by survival experiments with pooled

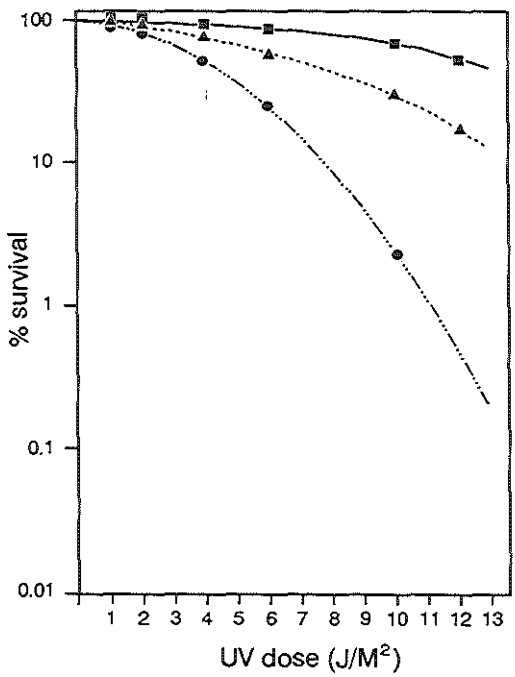


Figure 2. Correction of the UV-sensitivity of mutant UV61 by the *ERCC6* gene. UV survival curves of wild type CHO cell line AA8 (■——■), mutant UV61 (●..-..●), and UV61 cells cotransformed with λ-phage 6B (see Fig.1) and a promoterless *ERCC6* cDNA (▲....▲).

clones. To rule out the possibility that the UV-resistant transformants obtained in the cotransfection experiments were caused by fortuitous integration behind an endogenous promoter, instead of reconstitution of a functional gene by recombination between the 5' *ERCC6* gene part with the rest of the cDNA, we also transfected the cDNA without a functional promoter at its 5' end. In this experiment, in which the cDNA was placed in the inverted orientation in vector pSLM, no UV-resistant clones were obtained. These findings strongly suggest that a functional gene has indeed been generated via homologous recombination (between the lambda and cDNA clones) within the cell. The promoter should then be situated within the ± 8 kb genomic region 5' of the first exon present in the cDNA clone. The *ERCC6* locus thus covers a region of 82 - 90 kb.

Intron exon structure

Appropriate fragments were subcloned into plasmid vectors in order to determine intron-exon borders. Sequence analysis demonstrated the *ERCC6* cDNA to be dispersed over 21 exons (Figure 1 and 3). The CpG-island present in the 5' region of the *ERCC6* gene is situated within the first intron of *ERCC6* (Figure 1). The first exon present in the cDNA does not contain any coding information, analogous to the first exon in another repair gene, *ERCC1* (Van Duin, et al., 1987). Since we have not determined the transcriptional start site, we cannot exclude the presence of an additional exon more 5' on the genome, which is absent in the cDNA. The presence of a CpG-island in the first intron, however, argues against the existence of an additional 5' exon located far upstream. The tentatively identified functional domains are - with the exception of helicase signature VI - encoded by separate exons (Figure 4).

As shown in Figure 3, all sequences around intron-exon and exon-intron borders are consistent with the consensus splice acceptor and donor signals (Senapathy, et al., 1990), with the exception of the splice donor of the intron VII. The canonical GT dinucleotide at the beginning of the intron is replaced by GC (see below). In all introns, the weakly defined branchpoint sequence (PyNPyTPuAPy) could be tentatively identified at the appropriate distance (10 - 50 nt) from the splice acceptor site (underlined in Figure 3; Ruskin, et al., 1984; Senapathy, et al., 1990). One alternatively spliced cDNA molecule has been isolated, in which exon VIII was found to be skipped. The sizes of internal exons range from 79 to 745 nt. Three of the internal exons (exon II, V, and XVIII) have an exceptional size: 436, 745 and 708 nt, respectively; the average exon length for vertebrate internal exons being 137 nt (Hawkins, 1988). Intron lengths range from 85 to 19,000 nt; the average size of vertebrate introns being 1,127 nt (Hawkins, 1988).

INTRON SPLICE ACCEPTOR		EXON			INTRON SPLICE DONOR	
		I	(≥65)	CCCTGG	?	(6 kb)
TACAATGAATCCTATATAACGAAAA	66			501		
GGCTTATAAATTTTTTTCCTTTTAG	GTAGTC	II	(436)	CCTCAC	GTGAGGTCCAGCC	(1.8 kb)
TATAAAAATGTTATCAGAGTGCAAA	502			622		
ATAGACAATTTATCTTATTTTTCAG	GTCATG	III	(121)	AATAAG	GTGATTCAGAATA	(2.1 kb)
CAATAAATCACTTCATTGCTGTTTT	623			731		
CTGTCTTCTATATTTACCATTCAAG	GAACAA	IV	(109)	ATGCAG	GTGAGGATGCTAG	(4 kb)
CCCTCGGAAAGTTTCATGCTAGTGG	732			1476		
CAAAGCATTGCTATTGTTCTTTCAG	AGCCGG	V	(745)	GTTAAG	GTCGGTCTGTGGG	(19 kb)
TCTGTCTTGTGATCAAAATAATGGA	1477			1605		
AATGTGATTTTTATTTTCATGGTAG	GAGATG	VI	(129)	TTTTAA	GTATGTACCATAT	(5 kb)
TATCCACCATTTGCCATTTTCTCTT	1606			1764		
TTCTTGTTGGTGTTTGTTGTCATAG	GTACCA	VII	(159)	TTACAG	GCAAGTGCTCCTC	(7.5 kb)
TCTTTCATGGTT	1765			1900		
GTTTTCTTCTTGCGTTTGGATGCAG	GTTTGA	VIII	(136)	AAAAAG	GTAACACAATATT	(9 kb)
GGTGTCACTTCTTATTTAAAACTGA	1901			2071		
CTTTACCATTTTATTGTGGTCTCAG	GAGAAA	IX	(171)	AAACAG	GTATGACCTCTTT	(0.5 kb)
GATTGCTAAAAGGTAAATTTAATAA	2072			2248		
AAAGGTGCTTCCTTTCCTCAAATAG	TTTCGC	X	(177)	GTACAG	GTAAAATATTAGG	(3.5 kb)
TTCCTTAAACATATATAGTAAT	2249			2365		
GTTCCCTTCTCTGCTCTTATTAAAG	GTCAAA	XI	(117)	GAACAG	GTCTGTAAATCCA	(2 kb)
GTTAATTTTTTTTTTGAAATTATAG	2366			2461		
TTTCCTTGTTTTTCCCGGTTCGTAG	GTCTTA	XII	(96)	ATGCAG	GTCAGCTAAAAAA	(2 kb)
TTACTTTACATGGGGTCATCTGAGT	2462			2677		
GTACATGTACTCTTTCTTACGACAG	ATTTTC	XIII	(216)	AGGCAG	GTGAGTGCACAGA	(0.5 kb)
CTGGGAATGTGTATTTGCTTTGCAA	2678			2788		
ACTCCTATCCCCCACCTCCAAACAG	ATGCTG	XIV	(111)	AATGAG	GTAACATGTGAAT	(0.4 kb)
TGTAACTGGTCTTAAGTGTGTGTGC	2789			2908		
TCAGTGTTGTGTGTCTTACCTCTAG	GACACA	XV	(120)	ACGCAG	GTTTGTTTTTATT	(0.4 kb)
GTCATTGGGAAGGATTCTCGTTG	2909			3003		
AGAGGTCTCTCTCTCTGTTGCAG	GCCCGG	XVI	(95)	CCACCG	GTCAGTGCACACA	(1.1 kb)
TAGGTAGAGCTACACATTGTTTTAT	3004			3149		
ACCAGCTTATCTTTTATTTTTTTAG	ACAAAT	XVII	(146)	TTGCAG	GTATTACATAAAA	(85 bp)
TTGCACAAGATGATACAATATAGTA	3150			3857		
TTAGTGGTGTTTTTCCTCTTTACAG	GAACTG	XVIII	(708)	AATCAG	GTAACCATTTGAT	(8.5 kb)
TTTCTTGCATACAGAGTGAAATATC	3858			4062		
ACTTTGCTATTTCTTTTCTTGCTAG	TTGGCG	XIX	(205)	AAAAAA	GTAAGAGATTGCT	(0.8 kb)
AAGTCTCAAAAGCAAACATTTAATC	4063			4141		
TACTGTGATGCTTTTCCTTTTTTAG	GAGTAG	XX	(79)	TGCCAG	GTAATATAGATAA	(1.2 kb)
GGCATAAACTAGAAATTAAATATAT	4142					
CAGTATAGTGGTCTTCTTTTTATAG	GATGGC	XXI	(573 / 2.5 kb)		3'poly(A)	

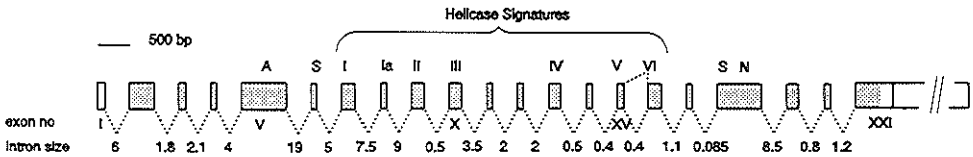


Figure 4. Schematic depiction of the postulated functional domains as encoded by different exons. The shaded area represents the ORF; encoded domains are indicated above the exons. Intron length is given in kb. The postulated functional domains (Troelstra, et al., 1992b) encoded by the different exons are indicated: A, acidic amino acid stretch; S, putative nuclear location signal followed by a postulated casein kinase II phosphorylation site; the helicase signatures are numbered I to VI; N, putative nucleotide binding fold. The interrupted box added to exon XXI indicates the 3' end of the longest *ERCC6* mRNA, produced by alternative poly-adenylation.

ERCC6 gene expression and alternative poly-adenylation

Northern blot analysis of human poly(A)$^+$ RNA demonstrated the presence of two *ERCC6* transcripts of approximately 5 and 7 kb in length, both expressed at a very low level (Figure 5). To examine the structure of the long *ERCC6* mRNA, various cDNA libraries were screened for the presence of clones with extra sequences not found in the 4.7 kb cDNA. One of the isolated (partial) cDNAs had a 3' untranslated region extending 2 kb more 3' as the mentioned 4.7 kb construct. It contained a poly(A)tail, preceded by the common polyadenylation signal AATAAA (Birnstiel, et al., 1985; Wickens, 1990). The total length of the isolated cDNA then becomes $\pm$6.7 kb, probably reflecting the longest mRNA. Sequence analysis and restriction enzyme digests showed the 3' untranslated region to be colinear with the genome. This suggests the 2 mRNA products to be the result of alternative polyadenylation (Figure 1).

Analysis of mouse tissues revealed the presence of elevated (though still very low) levels of *ERCC6* mRNA in brain and testis; in most other tissues transcription was below detection level (Figure 5). Also in mouse two transcripts are identified,

Figure 3. Structural organization of the *ERCC6* gene. The nucleotide sequence of each intron-exon junction is shown, with the exception of the splice donor of exon I. The vertical lines represent intron-exon borders. The splice donor of exon VII has a GC dinucleotide (indicated in bold) at the position of the canonical GT. All other acceptor and donor sites are in reasonable accordance with the consensus sequence $(Py)_nNCAG/G$ and (C,A)AG/GTPuAGT (Senapathy, et al., 1990). The nucleotides at the borders of each exon are numbered as reported previously (Troelstra, et al., 1992b). The size of introns and exons are given between parenthesis (in bp, if not indicated otherwise).

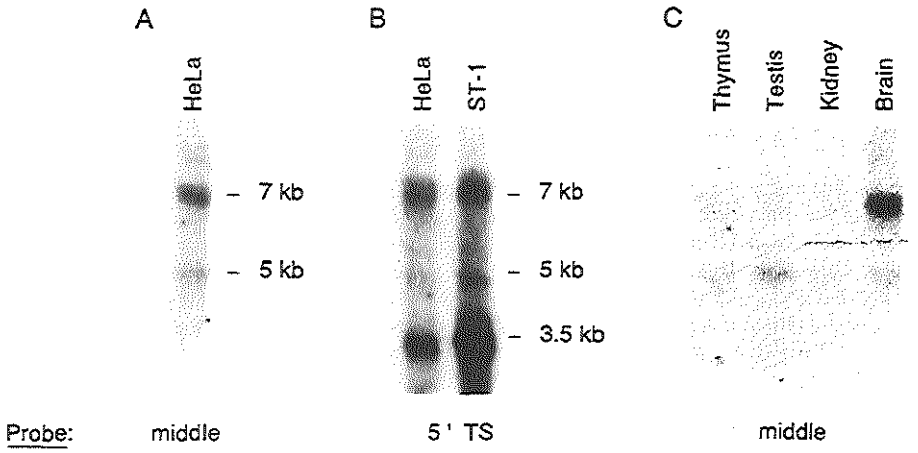


Figure 5. Expression of the *ERCC6* gene.
A: Northern blot analysis of 5 μg poly(A)$^+$ mRNA from HeLa cells, hybridized with a ^{32}P-labeled *ERCC6* cDNA fragment extending from nt 1662 - 3746 (middle probe). *B*: Northern blot analysis of 5 μg poly(A)$^+$ mRNA from HeLa cells and ST1, a secondary repair proficient transformant of UV61, containing the human *ERCC6* gene (Troelstra, et al., 1990). The blot is hybridized with a strand-specific probe (^{32}P-labeled transcribed strand, TS) spanning nt 16 to 1667 from the *ERCC6* cDNA. A radioactively labeled non-transcribed strand did not recognize any transcript (result not shown). *C*: Northern blot analysis of 20 μg of total RNA from different mouse tissues. The probe used is similar to that in *A*.

suggesting that the alternative polyadenylation is conserved. Remarkably, the larger transcript is the most abundant in mouse brain, whereas in testis the smaller mRNA is the more frequently occurring one (Figure 5).

In human RNA, a third transcript of 3.5 kb is detected, with probes in the 5' 1.7 kb of the gene. A probe from nt 1662 to 3746 did not detect this short transcript. Both a probe covering the first 350 bp of the cDNA (exon I and part of exon II) and one spanning nt 1020 to 1667 (part of exon V, exon VI and 60 nt of exon VII) recognize this short transcript, suggesting that at least sequences from exon I or II as well as exon V or VI (and maybe the 60 nt of exon VII) should be present. The transcript is detected in both total and poly A$^+$ RNA. Hybridizations with strand-specific probes, as shown in Figure 5, have excluded the possibility of an overlapping antisense gene: the transcript appears to be produced from the same strand as the two longer *ERCC6* mRNAs. Since we have been unable to identify *ERCC6* transcripts in

mouse tissues with a 5' probe, it is unclear whether the 3.5 kb transcript is conserved through evolution.

DISCUSSION

The *ERCC6* gene product, which is predicted to be a nuclear DNA helicase, has recently been shown to be involved in the human DNA repair disorder Cockayne's syndrome (Troelstra, et al., 1992b). In this paper, the expression and structural organization of the *ERCC6* gene are reported.

Two lowly expressed *ERCC6* mRNA molecules of 5 and 7 kb have been identified by Northern blot analysis of human poly(A)$^+$ RNA (Figure 5). Two types of cDNAs have been isolated, varying in their 3' end by differential polyadenylation. The second polyadenylation site, accompanied by the common polyadenylation signal AATAAA, is present 2 kb downstream of the first one that contains the less preferred signal ATTAAA (Birnstiel, et al., 1985; Wickens, 1990). Alternative polyadenylation occurs in many genes, including another human NER gene *ERCC1* (Van Duin, et al., 1987). It is unclear though, whether this has any regulatory function. In this respect, it is interesting to note that often the alternative polyadenylation is evolutionarily conserved, as it is - presumably - for *ERCC6*. We noted differential expression of the two transcripts in mouse brain and testis (Figure 5). It is unknown whether the elevated expression of the longest *ERCC6* mRNA in brain is related to the (severe) progressive neurological degeneration in CS patients. Also the function of the (human) 3.5 kb transcript encompassing only the 5' region of the *ERCC6* cDNA is unresolved. Hybridization with strand-specific probes ruled out the possibility of an overlapping antisense gene. An overlapping gene in sense direction seems unlikely, since it should then have (parts) of at least 2 exons (exons I or II and V, VI or VII) in common with the *ERCC6* gene. A third possibility is that the 3.5 kb transcript is the result of a (strong) transcriptional stop due to the presence of an attenuation site. The attenuation site should then be within or 3' of exon V, which is approximately 15 kb (or more) from the transcriptional start site. Normally, attenuation takes place within a few hundred bp from the transcription initiation site, although there is one example of a termination site $\approx$2 kb from the start (c-*myb*; Watson, 1988). If, for the *ERCC6* gene, the stop site is 3' of exon V, it could be within the 19 kb long intron V. Cross-hybridization of unique human sequences from this intron to rodent DNA suggested the presence of 2 conserved regions within intron XIX (unpublished observations). It is possible that a second gene is located within this large intron, with transcription

termination sequences that can be recognized by RNA polymerase complexes transcribing the *ERCC6* gene. Splicing of this prematurely terminated transcript might result in the observed 3.5 kb mRNA. A more conclusive statement concerning the role of this short (truncated) *ERCC6* transcript should, however, await further research, particularly by the isolation of corresponding cDNAs. So far, no such cDNAs have been isolated, despite the screening of several cDNA libraries.

The *ERCC6* locus is 82 - 90 kb in length, and harbors at least 21 exons. The presence of a CpG-island within intron I (Figure 1) suggests that there is no unidentified exon more 5' (containing extra 5' untranslated sequences not present in the cDNA), although we cannot completely exclude this possibility. It is interesting to note that also *ERCC2* and *ERCC3* have CpG-islands within the first intron (Weber, et al., 1990; Weeda, et al., 1991). The putatively identified functional domains in the *ERCC6* cDNA sequence are - with one exception - dispersed over separate exons. It has been proposed that modern eukaryotic genes may have been assembled from a limited number of ancestral exons encoding separate functional domains (Dorit, et al., 1990). Within *ERCC6*, only helicase domain VI is split by intron XV (see Figure 4). Interruption of helicase domains by introns has been noted before: yeast RNA helicase p68 contains one intron in domain V (Iggo, et al., 1991), which is positionally conserved (between human, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*); within the *ERCC3* gene (encoding a putative DNA helicase) both domain I and V are interrupted (Weeda, et al., 1991).

In one of the isolated *ERCC6* cDNAs exon VIII was found to be skipped. Although we cannot completely exclude the possibility that a mRNA molecule without exon VIII has any functional potential, it seems unlikely. The exon skipping results in a frameshift; consequently, the mRNA only encodes a very short protein. Presumably, this unconventionally spliced mRNA is the result of an erroneous splice event, rather than the product of a mechanism that creates an alternative, functional *ERCC6* protein.

All splice sites, with the exception of the splice donor at the 3' end of exon VII, are in reasonable accordance with the consensus sequences $(Py)_nCAG$ / G and (C,A)AG / GTPuAGT for donor and acceptor, respectively (Senapathy, et al., 1990, see Figure 3). In the splice donor of intron VII, the "invariable" GT dinucleotide is replaced by GC. "GC" instead of "GT" splice donors are very rare. Recently, Jackson (Jackson, 1991) compared 26 "GC" splice donors from the literature, and noted that as a group they have a better "overall' match to the (C/A)AG/GTAAGT consensus than the regular "GT" donors. To assess the validity of this finding and to make conclusions statistically more significant, we have screened various data bases for

Table I. Summary of GC-splice donor sequences.

Gene	Sequence		Reference
	Exon	Intron	
Human cytochrome P-450	CACTAAG	GCAAGCCCACA	(Jackson, 1991 and references therein)
Hamster αA-crystallin	CATCAAG	GCAAGTTACAT	(Jackson, 1991 and references therein)
Mouse αA-crystallin	CGTCAAG	GCAAGTTACAT	(Jackson, 1991 and references therein)
Mole rat αA-crystallin	CATCAAG	GCAAGTTTCGT	(Jackson, 1991 and references therein)
Chicken αD-globin	TTTCAAG	GCAAGCAAAGG	(Jackson, 1991 and references therein)
Duck αD-globin	CTTCAAG	GCAAGCGGGGA	(Jackson, 1991 and references therein)
Chicken myosin heavy chain	GCTGCAG	GCAAGTGTCTG	(Jackson, 1991 and references therein)
Rat heme oxygenase	TCGACAG	GCAAGCGACTA	(Jackson, 1991 and references therein)
Soybean nodulin-24	AAAGAGG	GCAAGTTAATT	(Jackson, 1991 and references therein)
Human factor XII	AGGACCG	GCGAGTACCCG	(Jackson, 1991 and references therein)
Pig growth hormone	GCTGCAG	GCAAGTGCCCC	(Jackson, 1991 and references therein)
Human acetylcholine receptor	CCGCAAG	GCAAGGACCCT	(Jackson, 1991 and references therein)
Rat pyruvate kinase	CACCCAG	GCATGTGCTAT	(Takenaka, et al., 1989)
Human superoxide dismutase-1	GCAGAAG	GCAAGGGCTGG	(Jackson, 1991 and references therein)
Mouse superoxide dismutase-1	GCAGAAG	GCAAGGCCCGG	(Jackson, 1991 and references therein)
Human prothrombin	TGTGCTG	GCAAGTCTGTG	(Jackson, 1991 and references therein)
Human factor VII	GCTGCAG	GCGGGGTCGCT	(Jackson, 1991 and references therein)
Human erythrocyte α-spectrin	CATTCAG	GCAAGTTCAA	(Kotula, et al., 1991)
Rat D2a receptor	AATACAG	GCAAGTCTGGC	(O'Malley, et al., 1990)
Earthworm hemoglobin chain c	TCACCAA	GCAAGTCTCCC	(Jackson, 1991 and references therein)
Minute virus of mice	TTTACAG	GCCTGAAATC	(Jongeneel, et al., 1986)
Human α-glucosidase	CACCAAG	GCAAGA	(Hoefsloot, et al., 1990)
Bovine aspartyl protease	TGGCGAG	GCAAGTCCAG	(Jackson, 1991 and references therein)
Mouse APRT	ATCGCAG	GCGAGTGGCCT	(Jackson, 1991 and references therein)
Hamster APRT	ATCGCAG	GCGAGTGGCCA	(Jackson, 1991 and references therein)
Human APRT	ATCGCAG	GCGAGTGCCAG	(Jackson, 1991 and references therein)
Neurospora qa repressor	CCTCTAG	GCACGTCTGTA	(Jackson, 1991 and references therein)
Human argininosuccinate lyase	GCTGCAG	GCAAGACATCA	(Abramson, et al., 1991)
Mouse TRP-1	GTGGAAG	GCAAGTAA	(Jackson, 1991 and references therein)
Murine fifth complement (C5)	ACGGGCT	GCAAGTGGT	(Haviland, et al., 1991)
Human platelet GP IIb (exon V)	CACTCAG	GCGAGTAGGGA	(Heidenreich, et al., 1990)
Human platelet GP IIb (exon VIII)	ACTACAG	GCAAGAAATCC	(Heidenreich, et al., 1990)
Human keratin 18	CGAAGAG	GCAAGCAGGGG	(Jackson, 1991 and references therein)
Mouse RNA polymerase (exon VII)	CTATAAG	GCATGTAAATA	(Jackson, 1991 and references therein)
Mouse RNA polymerase (exon XIII)	CTCCCAG	GCAAGATGCTT	(Jackson, 1991 and references therein)
E.typhina TUB-8	CAACGAG	GCAAGTCTTCA	(Byrd, et al., 1990)
Rat ESP-1	CTATCAG	GCAAGAATGCT	(Girotti, et al., 1992)
HSV-1 LAT	CAAGAAG	GCATGTGTCCC	(Spivack, et al., 1991)
Human C3	TACCCAG	GCAAGT	(Barnum, et al., 1989)
Human DNA ligase I	ACGCAAG	GCAAGT	(Noguiez, et al., 1992)
Human ERCC3	TATTAAG	GCAAGTGACAG	(Weeda, et al., 1991)
Human ERCC6	ATTACAG	GCAAGTGCTCC	(This paper)
Sequence complementary to U1 snRNA	CAG	GTAAGT	(Rosbash and Seraphin, 1991)

additional "GC" splice donors, and extended the list to a total of 42 (Table I). The comparison presented in Table II clearly confirms and strengthens the observation made by Jackson: "GC" splice sequences display as a whole at all (remaining) positions a substantially higher level of complementarity to the U1 snRNA than the group of "GT" donors. Particularly the G residue at position +5 seems invariant, whereas the "G" at -1 is almost invariant. Most sites have 6-8 bases matching U1 snRNA (6 sites have 8 matches, 21 sites have 7 matches, 12 sites have 6 matches, the remaining 3 sites only have 5 bases matching to U1 snRNA). Apparently, the mismatch in base pairing with U1 snRNA caused by the GT to GC alteration needs to be compensated by a better match of the rest of the sequence, in order to permit the assembly of a functional spliceosome.

Three of the internal *ERCC6* exons (exons without either a cap site or a polyadenylation signal) are exceptionally large: 436, 708, and 745 nt (exons II, XVIII,

Table II. Nucleotide percentages at GC-type compared to GT-type splice donors.

Consensus GT splice[a]	-1 C/A	-2 A	-3 G	1 G	2 T	3 A/G	4 A	5 G	6 T
G	18	12	79	100	-	35	11	82	18
A	32	60	9	-	-	59	71	7	16
T	13	15	7	-	100	3	9	6	50
C	37	13	5	-	-	3	9	6	16

Consensus GC splice[b]	C	A	G	G	C	A	A	G	T
G	9.5	2.4	95.2	100	-	14.3	2.4	100	9.5
A	38.1	90.4	2.4	-	-	83.3	85.7	-	14.3
T	2.4	2.4	2.4	-		-	9.5	-	64.3
C	50	4.8	-	-	100	2.4	2.4	-	11.9

Significance of increased complementarity to U1 snRNA in GC splice donor sites (shaded boxes)								
p-values	<0.1 (ns)[c]	<0.001	-	-	<0.001	<0.05	<0.01	<0.1 (ns)[c]

[a] Data taken from reference (Senapathy, et al., 1990).
Percentages are based on 3724 different splice donors.
[b] Data from this work and reference (Jackson, 1991).
Percentages are based on 42 different splice donors.
[c] ns = not significantly different.

---

and V respectively). Out of 1305 internal vertebrate exons examined by Hawkins, only 7 exceed 550 nt; the average length being 137 nt (Hawkins, 1988). A conceivable explanation for the scarcity of large internal exons, as proposed by Robberson and colleagues (Robberson, et al., 1990) is based on the idea that, in order to stably form a spliceosome, factors bound at the 5' and 3' border of an exon need to "communicate". The interaction between the different factors might become less stable if the exon length exceeds a certain limit. *In vitro* splicing experiments demonstrated splicing of an intron followed by a large exon (>300 nt) to be less efficient than splicing of the same intron preceding a small exon

(<300 nt; Robberson, et al., 1990). Whether this is true *in vivo* too, is unknown. If so, it might (partly) explain the low expression level of the *ERCC6* gene.

Acknowledgements

We thank Mirko Kuit for photography, and Bart Janssen (Dept. Clin. Genet.) for help in some of the experiments.

REFERENCES

Abramson, R. D., Barbosa, P., Kalumuck, K., and O'Brien, W. E. (1991). Characterization of the argininosuccinate lyase gene and analysis of exon skipping. Genomics *10*, 126-132.

Barnum, S. R., Amiguet, P., Amiguet-Barras, F., Fey, G., and Tack, B. F. (1989). Complete intron/exon organization of DNA encoding the α' chain of human C3. J. Biol. Chem. *264*, 8471-8474.

Birnstiel, M. L., Busslinger, M., and Strub, K. (1985). Transcription termination and 3' processing: the end is in site! Cell *41*, 349-359.

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.

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.

Byrd, A. D., Schardl, C. L., Songlin, P. J., Mogen, K. L., and Siegel, M. R. (1990). The ß-tubulin gene of *Epichloe typhina* from perennial ryegrass (Lolium perenne). Curr. Genet. *18*, 347-354.

Cleaver, J. E., and Kraemer, K. H. (1989). Xeroderma pigmentosum. In: The metabolic basis for inherited disease. Vol. II, Scriver, C. R., Beaudet, A. L., Sly, W. S. and Valle, D., eds. (New York: McGraw-Hill Book Co.), 2949-2971.

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

Dorit, R. L., Schoenbacher, L., and Gilbert, W. (1990). How big is the universe of exons? Science *250*, 1377-1382.

Espelund, M., Prentice Stacy, R. A., and Jakobsen, K. S. (1990). A simple method for generating single-stranded DNA probes labeled to high activities. Nucl. Acids Res. *18*, 6157-6158.

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.

Fourney, R. M., Miyakoshi, J., III, R. S. D., and Paterson, M. C. (1988). Northern blotting: efficient RNA staining and transfer. Focus *10*, 5-7.

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

Girotti, M., Jones, R., Emery, D. C., Chia, W., and Hall, L. (1992). Structure and expression of the rat epididymal secretory protein I gene. Biochem. J. *281*, 203-210.

Haviland, D. L., Haviland, J. C., Fleischer, D. T., and Wetsel, R. A. (1991). Structure of the murine fifth complement component (C5) gene. J. Biol. Chem. *266*, 11818-11825.

Hawkins, J. D. (1988). A survey on intron and exon lengths. Nucl. Acids Res. *16*, 9893-9908.

Heidenreich, R., Eisman, R., Surrey, S., Delgrosso, K., Bennett, J. S., Schwartz, E., and Poncz, M. (1990). Organization of the gene for platelet glycoprotein IIb. Biochem. *29*, 1232-1244.

Hoefsloot, L. H., Hoogeveen-Westerveld, M., Reuser, A. J. J., and Oostra, B. A. (1990). Characterisation of the human lysosomal α-glucosidase gene. Biochem. J. *272*, 493-497.

Hoeijmakers, J. H. J., and Bootsma, D. (1990). Molecular genetics of eukaryotic DNA excision repair. Cancer Cells *2*, 311-320.

Iggo, R. D., Jamieson, D. J., MacNeill, S. A., Southgate, J., McPheat, J., and Lane, D. P. (1991). p68 RNA helicase: identification of a nucleolar form and cloning of related genes containing a conserved region in yeasts. Mol. Cell. Biol. *11*, 1326-1333.

Jackson, I. J. (1991). A reappraisal of non-consensus mRNA splice sites. Nucl. Acids Res. *19*, 3795-3798.

Jongeneel, C. V., Sahli, R., McMaster, G. K., and Hirt, B. (1986). A precise map of splice junctions in the mRNAs of minute virus of mice, an autonomous parvovirus. J. Virol. *59*, 564-573.

Kantor, G. J., Barsalou, L. S., and Hanawalt, P. C. (1990). Selective repair of specific chromatin domains in UV-irradiated cells from xeroderma pigmentosum complementation group C. Mutat. Res. *235*, 171-80.

Kotula, L., Laury-Kleintop, L. D., Showe, L., Sahr, K., Linnenbach, A. J., Forget, B., and Curtis, P. J. (1991). The exon-intron organization of the human erythrocyte α-spectrin gene. Genomics *9*, 131-140.

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.

Link Jr., C. J., Burt, R. K., and Bohr, V. A. (1991). Gene-specific repair of DNA damage induced by UV irradiation and cancer chemotherapeutics. Cancer Cells *3*, 427-436.

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.

Noguiez, P., Barnes, D. E., Mohrenweiser, H. W., and Lindahl, T. (1992). Structure of the human DNA ligase I gene. Nucl. Acids Res. *20*, 3845-3850.

O'Malley, K. L., Mack, K. J., Gandelman, K. Y., and Todd, R. D. (1990). Organization and expression of the rat $D2_A$ receptor gene: identification of alternative transcripts and a variant donor splice site. Biochem. *29*, 1367-1371.

Robberson, B. L., Cote, G. J., and Berget, S. M. (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. *10*, 84-94.

Robbins, J. H., Kraemer, K. H., Lutzner, M. A., Festoff, B. W., and Coon, H. G. (1974). Xeroderma pigmentosum. An inherited disease with sun sensitivity, multiple cutaneous neoplasms and abnormal repair. Ann. Intern. Med. *80*, 221-248.

Rosbash, M., and Seraphin, B. (1991). Who's on first? The U1 snRNP-5' splice site interaction and splicing. Trends Biochem. Sci. *16*, 187-190.

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.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual. (Cold Spring Harbor: Cold Spring Harbor Laboratory Press).

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.

Spivack, J. G., Woods, G. M., and Fraser, N. W. (1991). Identification of a novel latency-specific splice donor signal within the Herpes Simplex Virus type I 2.0-kilobase latency-associated transcript (LAT): translation inhibition of LAT open reading frames by the intron within the 2.0-kilobase LAT. J. Virol. *65*, 6800-6810.

Stefanini, M., Collins, A. R., Riboni, R., Klaude, M., Botta, E., Mitchell, D. L., and Nuzzo, F. (1991). Novel Chinese hamster ultraviolet-sensitive mutants for excision repair form complementation groups 9 and 10. Cancer Res *51*, 3965-3971.

Takenaka, M., Noguchi, T., Inoue, H., Yamada, K., Matsuda, T., and Tanaka, T. (1989). Rat pyruvate kinase M gene. J. Biol. Chem. *264*, 2363-2367.

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

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

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

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., van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D., and Hoeijmakers, J. H. J. (1992b). *ERCC6*, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. Cell *71*, 939-953.

Van Duin, M., Koken, M. H. M., van den Tol, J., ten Dijke, P., Odijk, H., Westerveld, A., Bootsma, D., and Hoeijmakers, J. H. J. (1987). Genomic characterization of the human DNA excision repair gene ERCC-1. Nucl. Acids Res. *15*, 9195-213.

Venema, J., Bartosova, Z., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. F. (1992). Transcription affects the rate but not the extent of repair of cyclobutane pyrimidine dimers in the human adenosine deaminase gene. J. Biol. Chem. *267*, 8852-8856.

Venema, J., Mullenders, L. H. F., 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., Karcagi, V., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. F. (1991). Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. Mol. Cell. Biol. *11*, 4128-34.

Venema, J., Van Hoffen, A., Natarajan, A. T., Van Zeeland, A. A., and Mullenders, L. H. F. (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., 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.

Watson, R. J. (1988). A transcriptional arrest mechanism involved in controlling constitutive levels of mouse c-*myb* mRNA. Oncogene *2*, 267-272.

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., Ma, L. B., van Ham, R. C., van der Eb, A. J., and Hoeijmakers, J. H. (1991). Structure and expression of the human XPBC/ERCC-3 gene involved in DNA repair disorders xeroderma pigmentosum and Cockayne's syndrome. Nucl. Acids Res. *19*, 6301-8.

Weeda, G., Van Ham, R. C. A., Vermeulen, W., Bootsma, D., Van der Eb, A. J., and Hoeijmakers, J. H. J. (1990). 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.

Wickens, M. (1990). How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem. Sci. *15*, 277-281.

Zdzienicka, M. Z., Schans, G. P. v. d., 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.

Chapter VII

General discussion

General considerations

Over the past few years, many eukaryotic nucleotide excision repair (NER) deficient mutants have been characterized, and a number of genes (and gene products) have been isolated and studied to considerable detail (Friedberg, 1988; Hoeijmakers and Bootsma, 1990; Hoeijmakers, 1991, for reviews). These studies have provided many answers, but simultaneously raised new questions. The work described in this thesis focusses on the isolation and characterization of the *ERCC6* gene. Mutations in that gene were found to be responsible for the most common form of the autosomal recessive disease Cockayne's syndrome (CS): CS-B. The defect in preferential repair of active DNA sequences in CS cells implies that the ERCC6 protein is essential for this subpathway of NER. It is as yet unknown, though, how the protein couples repair to transcription, how a deficiency in this gene product results in the observed neurological degeneration specific for CS and why the absence of a functional ERCC6 protein does not lead to enhanced carcinogenesis within CS-B patients. In this chapter these - and other - issues are discussed, and translated into putative molecular models.

Cockayne's syndrome

CS is a rare, autosomal recessive DNA repair disorder, clinically characterised by a growth defect, progressive neurological degeneration, calcification in the brain, ocular lesions (e.g. cataracts, progressive retinal pigmentary degeneration), and sun (UV) sensitivity of the skin. Three clinical subtypes of Cockayne's syndrome have been described, which differ in the time of onset and severity of the disease (see also Chapter II). The subtypes are being referred to as mild (late onset form), CS I (clinical symptoms become apparent within the first year after birth), and CS II, the most severe form (Nance and Berry, 1992). At the cellular level, two complementation groups are distinguished within "classical" (i.e. without XP symptoms) CS patients (Tanaka, et al., 1981; Lehmann, 1982). CS group A so far contains two patients, one defined clinically as CS I, the other having the mild form of the disease. CS-B, the most common subtype of the disease, comprises patients who are clinically defined as CS I. Apparently, the clinical heterogeneity does not coincide with the genetic heterogeneity as defined by the complementation groups. A hallmark of CS cells from group A as well as group B is the absence of preferential (fast) repair of UV-induced CPD in active genes, whereas the slow and less efficient repair of inactive chromatin (overall genome repair) is unaffected (Venema, et al., 1990a). This implies the defective protein in each group to function (specifically) in repair of active

genes.

The molecular defect underlying the repair deficiency of CS group B has been unravelled; patients within this group carry a defective *ERCC6* gene. The gene defect in CS-A patients is as yet unknown. Chapter V of this thesis describes the mutations present in the *ERCC6* alleles of patient CS1AN (group B, a clinical report of this female patient is found in (Schmickel, et al., 1977)). Both alleles carry deleterious mutations, resulting in two - presumably - non-functional proteins (see Figure 3B in Chapter V). This finding suggests cells from unassigned, more severely affected CS cases reported to be unlikely candidates for group B. The assignment of these severely affected patients to any of the existing - or possibly new - complementation groups will have to await further complementation analysis. Individuals with a mild form of the disease - such as patient GM2965 in CS group A - most likely have mutations that only partially affect the function of the responsible gene. A patient with such a mild, late-onset form of the disease that should possibly be assigned to group B is described by Fryns *et al.* (1991). This patient carries an interstitial deletion of chromosome region 10q11-21, including the *ERCC6* gene (Fryns, et al., 1991, B. Smit and A. Hagemeijer, personal communication). Presumably, the other *ERCC6* allele harbours a less severe mutation.

Several patients have been described that exhibit a combination of clinical features of both CS and XP (Robbins, et al., 1974; Lafforet and Dupuy, 1978; Jaeken, et al., 1989). These patients have been assigned to XP group B (also designated CS-C; now containing 3 individuals), XP-D (1 XP/CS case), and XP-G (2 XP/CS individuals) (Robbins, et al., 1974; Lehmann, 1982; Vermeulen, et al., 1991; Vermeulen, et al., 1992, and Vermeulen, manuscript in preparation). They suffer from the neurological abnormalities characteristic for CS as well as the pigmentation abnormalities normally seen in the skin of XP patients. The skin tumors often found in XP patients were present in two cases only (XP11BE from XP group B and XP-CS-2 from XP-D). The absence of tumors in the two XP/CS individuals in XP group G might be related to the young age of both patients (who died at age 2.5 and 6, respectively). The absence of tumor formation in the two other - much older - XP/CS patients (assigned to XP-B) is as yet unexplained. Lehmann and Norris (Lehmann and Norris, 1989) have proposed a model in which the XP repair defect by itself is thought to be insufficient to trigger tumorigenesis and an additional deficiency in immune response is essential to result in the cancer prone XP phenotype. So far, however, there is no definite proof for this idea. The XP- and XP/CS-mouse models that are currently being developed (by gene replacement in mouse embryonic stem cells, G. Weeda, I. Donker, M. Hoogeveen, and J. de Wit, personal communication) might shed light on the difference in tumor susceptibility. At the cellular level, XP/CS

patients have - in contrast to "classical" CS - a severely reduced UV-induced DNA repair synthesis (unscheduled DNA synthesis) (Robbins, et al., 1974; Jaeken, et al., 1989; Vermeulen, et al., 1991; Vermeulen, et al., 1992 and Vermeulen, manuscript in preparation). In several XP/CS cell lines UV-induced RNA synthesis recovery has been examined, and found to be inhibited (Lehmann, 1982; Jaeken, et al., 1989; Vermeulen, et al., 1992 and Vermeulen, manuscript in preparation), indicating that besides a defect in overall genome repair (reduction of UDS) also preferential repair of active genes is deficient (so similar to what has been found for e.g. XP group A).

Cockayne's syndrome: mutation induction and tumorigenesis

Although CS patients are - like XP - photosensitive, other cutaneous symptoms associated with XP (pigmentation abnormalities and the elevated risk for developing skin tumors) are absent (Lehmann, 1987; Nance and Berry, 1992). Studies on UV-induced mutability in CS cells appeared to be difficult because of the premature senescence intrinsic to CS cells (Lehmann, 1987). The few data available, however, indicate an elevated mutation induction in cultured CS cells, albeit less pronounced than in XP (Arlett and Harcourt, 1982). Although preferential repair has a profound effect on cell survival, it is - in human cells - expected to contribute only slightly to the process of mutagenesis. First of all, preferential repair is directed towards the transcribed strand (Mellon, et al., 1987), thus of little influence on mutations originating from lesions in the nontranscribed strand. Secondly, the removal of (6-4) photoproducts (besides CPD the most frequently UV-induced lesion) is fast (Mitchell and Nairn, 1989 for a review), and only slightly influenced by preferential repair of active genes (Thomas, et al., 1989; Link Jr., et al., 1992). Finally the slower, in CS seemingly unaffected, "overall genome" repair might in the end remove most of the lesions from the nontranscribed as well as many from the transcribed strand. Especially in this respect use of *in vitro* cultured cells (fastly growing) could differ from the *in vivo* situation in the patient (where many cells are growing more slowly). The contribution of preferential repair to mutagenesis could be exaggerated in the relatively fast growing cultured cells. The inefficient "overall genome" repair, which *in vivo* might be enough to - finally - remove many lesions from both the nontranscribed and the transcribed strand of active genes, could be too slow to achieve the same in the fastly growing cultured cells. This hypothesis would explain the lack of tumorigenesis in CS patients on the one hand as well as the (slightly) elevated level of mutagenesis reported in cultured CS cells (Arlett and Harcourt, 1982) on the other hand.

Neurological degeneration in XP, CS, and XP/CS patients

Progressive neurological degeneration is a characteristic of all CS, XP/CS, and some XP patients. XP patients with neurological disease have been reported in groups A, B(=XP/CS), D, and G (Robbins, 1988). Within the relatively large group C, one patient was reported to have neurological complications (Robbins, 1989); generally, XP-C patients do not suffer from neurological degeneration. The finding of Venema and coworkers that XP-C cells are specifically disturbed in repair of inactive sequences, while retaining the ability to repair the transcribed strand of active genes (Venema, et al., 1990b; Venema, et al., 1991) - thus being the "biochemical antipole" of CS-A and CS-B - prompted them to speculate neurological disease to be correlated with the absence of preferential repair (Venema, et al., 1990a). This assumption is sustained by the fact that neuronal cells are transcriptionally active, and transcribe a larger fraction of their genome compared to other cells. It does not, however, explain the mysterious phenomenon that the histopathological findings with respect to the progressive degeneration is reported to differ between CS on one, and XP on the other hand. In CS patients, primary demyelination is said to lead to the severe neurological symptoms (Grunnet, et al., 1983; Ohnishi, et al., 1987; Patton, et al., 1989; Sasaki, et al., 1992), whereas in XP they seem to be due to an axonal degeneration (Robbins, 1988 for a review; Robbins, et al., 1991). This suggests that in CS patients the myelin forming glia cells are the primary target, whereas in XP the axons are the more sensitive cell type. Furthermore, CS patients differ - neurologically - from those suffering from XP in that multiple calcifications are found in brains of CS patients. XP/CS patients appear to exhibit the neurological disease of CS rather than that of XP (Robbins, et al., 1974; Lafforet and Dupuy, 1978; Jaeken, et al., 1989). In XP, the (slow) accumulation of lesions throughout the genome will lead to a general deterioration of cell function and finally cell death, possibly resulting in the XP-like axonal degeneration. In CS cells the - still functional - "overall genome" repair pathway would be able to repair part of the lesions, thereby postponing cell death. This might explain why CS patients do not have the neuronal degeneration specific for XP patients. The mechanism responsible for the CS-like demyelination, however, remains unexplained.

In CS cells, preferential repair of (the transcribed strand of) active genes is strongly affected, whereas "overall genome" repair appears normal (Venema, et al., 1990a; Venema, 1991). The defective proteins (ERCC6 for CS-B, still unknown for CS-A) are therefore hypothesized to somehow couple repair to the process of transcription. One of the ways in which this could be achieved is by recognition or displacement of a RNA polymerase complex stalled at a DNA lesion - as hypothesized for the ERCC6 protein in Chapter V. If not recognized and removed, the RNA

polymerase and all pursuing transcription complexes might remain bound at the template strand. "Trapping" of RNA polymerase complexes could - hypothetically - disturb repair and transcription of this particular gene, as well as possibly overall transcription levels. Especially non-dividing cells such as present in the nervous system - in which there is no replication complex that might efficiently remove blocked RNA polymerase complexes from their template strand - would then be prone to deterioration of overall transcription levels.

Myelinating glia cells (Schwann cells and oligodendrocytes) have received considerable attention during the past decade, since demyelination is the primary cause of diseases like Multiple Sclerosis, and several immune-mediated demyelinating diseases. In mouse models, evidence has been obtained that the production of the myelin sheath is strongly determined by the amount of mRNA encoding myelin basic protein (MBP). The correction of the phenotype of mice carrying different mutations in the *MBP* gene by introduction of the normal *MBP* gene in fertilized oocytes was dependent on the transcription level of the transgene: only mice homozygous for the introduced *MBP* gene showed a normal phenotype (Popko, et al., 1987; Readhead, et al., 1987). This suggests the myelination process to be - at least in part - critically dependent on transcription levels. It is unlikely, although not excluded, that the dependence on transcription levels is due to an inhibiting effect of the mutated protein present. The results were obtained in two different mouse strains, *shiverer* (Readhead, et al., 1987) and *mld* (Popko, et al., 1987). *Shiverer* mice are homozygous for a deletion removing the entire *MBP* gene with the exception of the promotor and the first exon (Roach, et al., 1985). In *mld* mice, a complicated duplication and inversion of part of the gene - resulting in the production of normal as well as antisense transcripts - is thought to be the cause. The strongly reduced amount of MBP protein detected in *mld* mice is ascribed to the formation of duplex MRNA molecules (Okano, et al., 1991). It seems therefore possible that if in CS the presence of DNA lesions in transcribed genes has a severe effect on overall transcription levels, this could specifically effect myelin synthesis by glia cells (and thereby result in the progressive demyelination of CS patients).

In XP group A - in which many patients have severe neurological abnormalities - cells are thought to be defective in one of the (early) excision repair steps common to both preferential and overall genome repair (Shivji, et al., 1992). Although this obviously leads to accumulation of DNA lesions in the genome - possibly resulting in the XP-like neuronal degeneration - the (still functional) proteins specific for the preferential repair pathway might be able to recycle the stalled RNA polymerase complexes without inducing DNA repair. This would at least allow transcription of

undamaged genes, ensuring production and maintenance of e.g. the myelin sheath.

XP/CS patients have the cutaneous abnormalities typical for XP, but a CS-like neurodegeneration. At the cellular level, both a reduced level of UDS (unscheduled DNA synthesis) and a defect in recovery of UV-induced RNA synthesis inhibition are present, suggestive of a defect in preferential as well as "overall genome" repair (similar to e.g. XP-A). In the above described hypothesis, the proteins affected in the different XP/CS complementation groups (XP-B/CS-C, XP-D, and XP-G) might perform a dual role: besides their functioning in the actual repair of the lesion, they could be involved in recognition and/or removal of stalled RNA polymerase complexes. If such a protein is defective, the neurological consequences might be determined mainly by the most direct effect: the negative influence of the stalled RNA polymerase complexes on overall transcription levels; thereby "masking" the XP-like neuronal degeneration due to accumulation of lesions throughout the genome. In XP-D patients that have been described having a XP-like neurological degeneration (Johnson and Squires, 1992), the affected protein (ERCC2) should then be mutated such that the ability to recognize and/or remove (together with other proteins) stalled RNA polymerase complexes is retained, but the actual repair function is lost. Alternatively, in this hypothesis, the protein affected in each XP/CS complementation group could have a role outside the NER, in the RNA transcription process. Subtle mutations in these proteins might then lead to a deterioration of transcription levels, leading to the CS-like neurological impairment. For *ERCC2* and *ERCC3* (XP-D and XP-B resp.) vital functions outside the excision repair process have been suggested (on the basis of - unknown - vital functions of their yeast homologs; Higgins, et al., 1983; Naumovski and Friedberg, 1983; Gulyas and Donahue, 1992; Park, et al., 1992).

It seems reasonable to assume that there is a common defect in CS and XP/CS patients, which results in the demyelination and discriminates them - clinically - from XP. Whether this common defect is indeed, as hypothesized, a decrease in transcriptional activity remains to be determined. At present, the above model is highly speculative; there is no evidence to sustain the proposed reduction in overall transcription levels. Other possibilities, such as differences in the nature of DNA lesions in different cell types (due to differences in cell metabolism), or disturbance of cell-cell interactions (between the neuron and the surrounding glia cells) can be envisaged, too.

Characterization of genes and gene products involved in (preferential) NER

The isolation and characterization of genes and gene products involved in the eukaryotic excision repair process has contributed considerably to our insight into the molecular mechanisms of NER. We are now able to assign putative functions to

several individual proteins, based on either *in vitro* activities or the predicted amino acid sequence of the protein in question. Nevertheless the issue as to how the different proteins cooperate and function in the recognition and removal of DNA damage is far from understood and remains subject to speculation.

Genes involved specifically in preferential NER

The NER system is divided in two subpathways: the fast, preferential repair of active genes, and the slower "overall genome" repair (Link Jr., et al., 1991 for a review). Although many gene products are shared by both repair pathways, at least one gene - *ERCC6* - can be pinpointed to be specifically involved in preferential repair only (see Chapter V). The predicted amino acid sequence of this protein (described in Chapter V) suggests it to be a nuclear DNA (or RNA) helicase. In Chapter V several possible functions for the ERCC6 protein are described: *1)* Probably, a RNA polymerase complex that is stalled at a DNA lesion within an actively transcribed gene will sterically hinder access of repair enzymes to the lesion, and should therefore be removed before repair can be accomplished (Selby and Sancar, 1990). ERCC6 might function herein - either by forcing the polymerase complex passed the lesion, pushing it backwards, or completely dissociating it from the DNA template. *2)* Alternatively, the protein could be involved in recognition of a stalled RNA polymerase complex, thereby directing the excision complex to actively transcribed genes. At present, we have no clues as to discriminate between these or even other possible functions of the ERCC6 protein.

A second gene product functioning in preferential gene repair is encoded by the CS-A correcting (*CSAC*) gene: the phenotype of CS-A patients as well as the repair characteristics of CS-A cells are very much alike CS-B (Venema, et al., 1990a). A more conclusive statement concerning the function of the CSAC protein has to await isolation and characterization of the CSAC gene and gene product.

Genes involved in preferential and "overall genome" repair

Mutations in either the *ERCC1*, *ERCC2*, *ERCC3*, *ERCC4*, *ERCC5*, *XPAC* or *XPGC* gene seem to abolish repair of inactive as well as active genes. Rodent mutants from complementation groups 1, 2, 3, 4 and 5 as well as XP-A cells are deficient in the incision step of the NER reaction. Cells from XP-G patients have a severely reduced repair DNA synthesis (UDS). Within the human counterparts of rodent groups 2 and 3 - XP-D and XP-B resp. - patients are found that exhibit the neurological degeneration normally found in CS, but the cellular characteristics specific for XP (Vermeulen, et al., 1991 and Vermeulen, manuscript in preparation). Similar patients

have been assigned to XP group G (Vermeulen, et al., 1992). Above, models are hypothesized, in which the proteins affected (ERCC2, ERCC3, and the not yet isolated XPGC) have a dual role; one in repair of a lesion irrespective of its location in the genome, the other in retaining normal levels of transcription (possibly through "turn-over" of stalled RNA polymerase complexes). For both *ERCC2* and *ERCC3* yeast homologs have been isolated and characterized (Weber, et al., 1990; Gulyas and Donahue, 1992; Park, et al., 1992). The gene product of the *ERCC2* yeast homolog *RAD3* has been investigated to considerable detail. The protein is demonstrated to be a DNA-DNA and DNA-RNA helicase (Sung, et al., 1987; Bailly, et al., 1991; Naegeli, et al., 1992). RAD3, and by implication its human homologue ERCC2, has been postulated to function in scanning the DNA for the presence of damage. It was suggested that the protein scans both inactive and actively transcribed sequences for the presence of DNA damage, by use of its DNA-DNA and DNA-RNA unwinding capacity respectively (Bailly, et al., 1991; Naegeli, et al., 1992). The protein would then encounter RNA polymerase complexes that are stalled at a DNA lesion. It might thus be involved in the recognition of damage as well as the recognition and displacement of a stalled RNA polymerase complex. Based on the phenotype of a specific *rad3* mutant: rad3 Arg-48, an alternative model has been proposed. The rad3 Arg-48 mutant carries a defective *RAD3* gene, encoding a protein in which the invariable lysine within the ATP-binding fold (helicase signature I) is replaced by arginine. Rad3 Arg-48 cells retain the ability to incise DNA upon UV-irradiation (although less efficient then wt), but are disturbed in the removal of pyrimidine dimers. Therefore, the RAD3 protein was suggested to function in a post-incision step, and to be essential for strand displacement and/or turn-over of the incision complex (cf. UvrD; Sung, et al., 1988). Conform with this postulated strand displacement or complex turn-over, the RAD3 (and by implication ERCC2) protein might utilize its DNA-RNA unwinding activity for turn-over of a stalled RNA polymerase complex. If so, the protein might act together or in complex with the ERCC6 gene product. Among the rodent group 2 mutants (which carry mutations in *ERCC2*, the human homolog of *RAD3*) a cell line (V-H1) has been reported that has similar characteristics as the yeast rad3 Arg-48 mutant: no CPD removal, but intermediate levels of incision. This mutant appeared to exhibit intermediate repair of (6-4) photoproducts, thus explaining the intermediate levels of incision (Mitchell, et al., 1989). For the rad3 Arg-48 yeast mutant repair of (6-4) photoproducts has not been investigated. For both the rodent and the yeast mutant, the protein might be affected such that CPD can not be recognized anymore, whereas the more helix disturbing (6-4) photoproducts still are recognized. This would suggest that the proteins (RAD3 and its human homolog

ERCC2) are directly or indirectly involved in damage recognition. Whether the protein in the *in vivo* NER process actually has one (or all) of the above mentioned or other functions awaits further research. Besides the important role in excision repair, RAD3 has another - as yet unidentified - function that is essential for yeast viability (Higgins, et al., 1983; Naumovski and Friedberg, 1983). The high degree of homology between RAD3 and its human homologue prompted the idea that probably also ERCC2 has such a vital function distinct from its involvement in repair. If so, this might provide an explanation for the surprising TTD phenotype seen in some XP group D patients (described in Chapter II). TTD might be the result of mutations affecting, but not abolishing, the vital function of the protein.

The sequence of the *ERCC3* gene revealed the presence of seven motifs conserved between different helicases, suggesting that also this protein possesses unwinding activity (Weeda, et al., 1990b). Several lines of evidence suggest the ERCC3 protein to have a role equivalent or complementary to that of ERCC2 (Weeda, et al., 1990b). The rodent mutants from cg 2 and 3 have comparable sensitivities; both XP group B (ERCC3) and XP-D (ERCC2) encompass XP/CS patients; both genes have yeast homologues with a similar (high) degree of identity and similarity ($\approx$ 50 and 70 % resp.) and both yeast proteins have a (presumed) helicase function and are essential for yeast viability. It has been suggested that ERCC3 might act in the same process as ERCC2 (whatever that process might be), but exhibiting a helicase activity of a different polarity (RAD3 possesses a 5' to 3' unwinding capacity). Alternatively, the proteins might function as a hetero-dimer (or multimer). The vital function of the yeast ERCC3 protein is, as for the RAD3 protein, unknown. Recently, a search for suppressor mutations relieving translation blockage from a MRNA containing an artificial hairpin in the 5' untranslated region (UTR) resulted in the isolation of four genes, designated *SSL1-4*. Surprisingly, sequence analysis of one of these genes - *SSL2* - revealed it to be the *Saccharomyces cerevisiae* equivalent of *ERCC3* (Gulyas and Donahue, 1992). Apparently, the mutated SSL2 protein permits translation from a MRNA containing a hairpin that normally blocks protein synthesis. Possibly, this is achieved through unwinding of the stable hairpin, or by stripping bound proteins from the 5' UTR containing this unusually stable hairpin. The vital function of SSL2 (and presumably its human homologue ERCC3) might therefore be related to the translation (initiation) process. Alternatively, the protein might actually be involved in transcription. As a result of the mutation it might for instance remain bound to the hairpin-region, thus preventing its formation and permitting translation. The involvement in transcription would be more conform the idea postulated above concerning the demyelination reported in both CS and XP/CS.

Future directions

In the past decade, considerable effort has been devoted to the isolation and characterization of genes involved in mammalian excision repair. The results have provided several key steps towards unravelling the molecular mechanism of mammalian nucleotide excision repair. At least seven genes are cloned (Westerveld, et al., 1984; Weber, et al., 1988; Tanaka, et al., 1989; Mudgett and MacInnes, 1990; Weeda, et al., 1990a; Legerski and Peterson, 1992 and Chapter III), and characterized to some extent. The primary structure of the encoded proteins as well as the phenotype of corrected rodent mutant cell lines or an associated human repair disorder have provided clues to putative functions. Undoubtedly, in the near future several of the postulated activities will be sustained by biochemical evidence. Moreover, insight into interactions between different components of the excision repair machinery can now be obtained (Bailly, et al., 1992; Bardwell, et al., 1992). Even specific steps in which single proteins are involved can be deduced with help of the - recently developed - *in vitro* repair assay based on cell-free extracts (Wood, et al., 1988; Shivji, et al., 1992).

The isolation and characterization of the *ERCC6* gene has provided a primary sequence that strongly suggests the protein to be a nuclear DNA (or RNA) helicase. The work has now come to a stage that will require protein analysis. Overexpression and purification of (parts of) the ERCC6 protein might provide biochemical information sustaining the prediction of a DNA (or RNA) unwinding activity. Thereby, antibodies raised against ERCC6 will be essential in providing insight into possible complexes in which the protein might exist *in vivo*.

The region spanning the different helicase signatures in ERCC6 is highly homologous to similar parts of several other putative helicases functioning in DNA repair (RAD5, RAD16, RAD54), transcriptional regulation (SNF2, MOT1, brahma), preservation of chromosome stability (lodestar), or unknown processes (FUN30, STH1) (see Chapter V). The functional significance of the homology between this group of presumed DNA helicases could be assessed by mutational analysis of the *ERCC6* gene and subsequent transfection to either rodent mutant cell line UV61 or a cell line of CS group B. Alternatively, the helicase regions could be interchanged between one of the yeast genes in this group and a yeast homologue of *ERCC6* could be performed; to test the repair capacity of the hybrid protein, though, an *ERCC6* deficient yeast mutant should be available. Recently, a new member of the above mentioned "helicase" subfamily has been isolated from yeast (A. van Gool, personal communication). The entire helicase region exhibits an extremely high degree of homology to the similar region of the human ERCC6 protein (the homology to human ERCC6 exceeds that to other family members). Whether this gene is indeed "the"

yeast homolog of ERCC6 remains to be elucidated. It will be interesting to determine the degree of UV-sensitivity of an *ERCC6* (and thus preferential repair) deficient mutant. The *E.coli* mutant *mfd*, deficient in strand-specific repair of active genes, has a near normal UV-sensitivity (George and Witkin, 1974). The fact that no *ERCC6* deficient yeast mutant has been isolated yet suggests it to be only marginally UV-sensitive, although other options (e.g. lethality of mutations in the gene) cannot be excluded.

Mutations in the *ERCC6* gene have been demonstrated to be responsible for the repair defect in cells from CS patients. In one clinically severely affected patient deleterious mutations are present in both *ERCC6* alleles. A CS patient with relatively mild clinical symptoms has been described, carrying an interstitial deletion of chromosome 10, region q11-21 including the *ERCC6* gene (Fryns, et al., 1991, B. Smit and A. Hagemeijer, personal communication). Mutational analysis of the second, still present *ERCC6* allele of this patient will provide information on a mutation that most likely only partially inhibits the normal function of the protein.

In this thesis, several models concerning ERCC6 activity in preferential repair of (the transcribed strand of) active genes are discussed. There is, however, no clue as to confirm or reject one of the hypotheses. Whether ERCC6 is involved in the recognition and/or replacement of a stalled RNA polymerase - one of the models proposed - is clearly an important point. An *in vitro* assay, that could measure repair in an actively transcribed plasmid - which is as yet not available - might shed some light on this issue. If a stalled RNA polymerase is to be recognized and/or removed before repair of the lesion can be accomplished, repair of these lesions in an extract that is deficient in this recognition/replacement (possibly CS) would be negatively influenced by transcription. A similar assay has been described for preferential repair in *E.coli* (Selby and Sancar, 1990; Selby and Sancar, 1991).

To study in more detail the *in vivo* function of ERCC6 in processes such as mutagenesis, tumorigenesis, and neurodegeneration, a mouse model would be very informative. By *ERCC6* gene replacement in mouse embryonic stem cells, a CS-B mouse can be generated; the results obtained by mutational analysis of a CS-B patient (see Chapter V) suggest that construction of a mouse model in which both copies of the *ERCC6* gene are deprived of any function is feasible. The construction of different repair-deficient mouse models will undoubtedly result in valuable information on various issues. The effect of the absence of preferential repair on mutagenesis and tumorigenesis can be investigated in more detail, e.g. after the administration of carcinogens, after UV exposure or in CS mice that carry an activated oncogene in their genome. A mouse model for CS will permit accurate determination of damage

in neuronal tissues. The results with respect to mutagenesis, tumorigenesis and neurodegeneration can be compared between XP, XP/CS, and CS mice. Moreover, XP/CS mice could be compared to "double mutant" mice carrying mutations in both alleles of a XP gene (e.g. *XPAC*) as well as in both alleles of a CS gene (e.g. *ERCC6*).

References

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, Lawrence, C. W., eds. (New York: Plenum Press), 249-266.

Bailly, V., Sommers, C. H., Sung, P., Prakash, L., and Prakash, S. (1992). Specific complex formation between proteins encoded by the yeast DNA repair and recombination genes *RAD1* and *RAD10*. Proc. Natl. Acad. Sci. USA *89*, 8273-8277.

Bailly, V., Sung, P., Prakash, L., and Prakash, S. (1991). DNA-RNA helicase activity of RAD3 protein of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA *88*, 9712-9716.

Bardwell, L., Cooper, A. J., and Friedberg, E. C. (1992). Stable and specific association between the yeast recombination and DNA repair proteins RAD1 and RAD10 *in vitro*. Mol. Cell. Biol. *12*, 3041-3049.

Friedberg, E. C. (1988). Deoxyribonucleic acid repair in the yeast *Saccharomyces cerevisiae*. Microbiol. Rev. *52*, 70-102.

Fryns, J. P., Bulcke, J., Verdu, P., Carton, H., Kleczkowska, A., and van den 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.

George, D. L., and Witkin, E. M. (1974). Slow excision in an *mfd* mutant of *Escherichia coli* B/r. Mol. Gen. Genet. *133*, 283-291.

Grunnet, M. L., Zimmerman, A. W., and Lewis, R. A. (1983). Ultrastructure and electrodiagnosis of peripheral neuropathy in Cockayne's syndrome. Neurology *33*, 1606-1609.

Gulyas, K. D., and Donahue, T. F. (1992). *SSL2*, a suppressor of a stem-loop mutation in the *HIS4* leader encodes the yeast homologue of human *ERCC3*. Cell *69*, 1031-1042.

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.

Hoeijmakers, J. H. J. (1991). How relevant is the Escherichia coli UvrABC model for excision repair in eukaryotes? J. Cell Sci. *100*, 687-691.

Hoeijmakers, J. H. J., and Bootsma, D. (1990). Molecular genetics of eukaryotic DNA excision repair. Cancer Cells 2, 311-320.

Jaeken, J., Klocker, H., Schwaiger, H., Bellmann, R., Hirsh-Kauffmann, M., and Schweiger, M. (1989). Clinical and biochemical studies in three patients with severe early infantile Cockayne syndrome. Hum. Genet. *83*, 339-346.

Johnson, R. T., and Squires, S. (1992). The XPD complementation group. Insights into xeroderma pigmentosum, Cockayne's syndrome and trichothiodystrophy. Mutat. Res. *273*, 97-118.

Lafforet, D., and Dupuy, J. M. (1978). Photosensibilite et reparation de l'ADN. Possibilite d'une parente nostologique entre xeroderma pigmentosum et syndrome de Cockayne. Arch. franc. Pediat. (Supplt) *35*, 65-74.

Legerski, R., and Peterson, C. (1992). Expression cloning of a human DNA repair gene involved in xeroderma pigmentosum group C. Nature *359*, 70-73.

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.

Lehmann, A. R., and Norris, P. G. (1989). DNA repair and cancer: speculations based on studies with xeroderma pigmentosum, Cockayne's syndrome and trichothiodystrophy. Carcinogenesis *10*, 1353-1356.

Link Jr., C. J., Burt, R. K., and Bohr, V. A. (1991). Gene-specific repair of DNA damage induced by UV irradiation and cancer chemotherapeutics. Cancer Cells *3*, 427-436.

Link Jr., C. J., Mitchell, D. L., Nairn, R. S., and Bohr, V. A. (1992). Preferential and strand-specific DNA repair of (6-4) photoproducts detected by a photochemical method in the hamster *DHFR* gene. Carcinogen. *13*, 1975-1980.

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.

Mitchell, D. L., and Nairn, R. S. (1989). The biology of the (6-4) photoproduct. Photochem. Photobiol. *49*, 805-819.

Mitchell, D. L., Zdzienicka, M. Z., van Zeeland, A. A., and Nairn, R. S. (1989). Intermediate (6-4) photoproduct repair in Chinese hamster V79 mutant V-H1 correlates with intermediate levels of DNA incision and repair replication. Mutat. Res. *226*, 43-47.

Mudgett, J. S., and MacInnes, M. A. (1990). Isolation of the functional human excision repair gene *ERCC5* by intercosmid recombination. Genomics *8*, 623-633.

Naegeli, H., Bardwell, L., Harosh, I., and Friedberg, E. C. (1992). Substrate specificity of the Rad3 ATPase/DNA helicase of *Saccharomyces cerevisiae* and binding of Rad3 protein to nucleic acids. J. Biol. Chem. *267*, 7839-7844.

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.

Ohnishi, A., Mitsudome, A., and Murai, Y. (1987). Primary segmental demyelination in the sural nerve in Cockayne's syndrome. Muscle & Nerve *10*, 163-167.

Okano, H., Aruga, J., Nakagawa, T., Shiota, C., and Mikoshiba, K. (1991). Myelin basic protein gene and the function of antisense RNA in its repression in myelin-deficient mutant mouse. J. Neurochem. *56*, 560-567.

Park, E., Guzder, S., Koken, M. H. M., Jaspers-Dekker, I., Weeda, G., Hoeijmakers, J. H. J., Prakash, S., and Prakash, L. (1992). *RAD25*, a yeast homolog of human xeroderma pigmentosum group B DNA repair gene is essential for viability. Proc. Natl. Acad. Sci. USA in press.

Patton, M. A., Giannelli, F., Francis, A. J., Baraitser, M., Harding, B., and Williams, A. J. (1989). Early onset Cockayne's syndrome; case reports with neuropathological and fibroblast studies. J. Med. Genet. *26*, 154-159.

Popko, B., Puckett, C., Lai, E., Shine, H. D., Readhead, C., Takahashi, N., Hunt III, S. W., Sidman, R. L., and Hood, L. (1987). Myelin deficient mice: expression of myelin basic protein and generation of mice with varying levels of myelin. Cell *48*, 713-721.

Readhead, C., Popko, B., Takahashi, N., Shine, H. D., Saavedra, R. A., Sidman, R. L., and Hood, L. (1987). Expression of a myelin basic protein gene in transgenic shiverer mice: correction of the dysmyelinating phenotype. Cell *48*, 703-712.

Roach, A., Takahashi, N., Pravtcheva, D., Ruddle, F., and Hood, L. (1985). Chromosomal mapping of mouse myelin basic protein gene and structure and transcription of the partially deleted gene in

shiverer mutant mice. Cell *42*, 149-155.

Robbins, J. H. (1988). Defective DNA repair in xeroderma pigmentosum and other neurologic diseases. Curr. Opin. Neurol. Neurosurg. *1*, 1077-1083.

Robbins, J. H. (1989). A childhood neurodegeneration due to defective DNA repair: a novel concept of disease based on studies of xeroderma pigmentosum. J. Child Neurol. *4*, 143-146.

Robbins, J. H., Brumback, R. A., Mendiones, M., Barrett, S. F., Carl, J. R., Cho, S., Denckla, M. B., Ganges, M. B., Gerber, L. H., Guthrie, R. A., Meer, J., Moshell, A. N., Polinsky, R. J., Ravin, P. D., Sonies, B. C., and Tarone, R. E. (1991). Neurological disease in xeroderma pigmentosum. Documentation of a late onset type of the juvenile onset form. Brain *114*, 1335-1361.

Robbins, J. H., Kraemer, K. H., Lutzner, M. A., Festoff, B. W., and Coon, H. G. (1974). Xeroderma pigmentosum. An inherited disease with sun sensitivity, multiple cutaneous neoplasms and abnormal repair. Ann. Intern. Med. *80*, 221-248.

Sasaki, K., Tachi, N., Shinoda, M., Satoh, N., Minami, R., and Ohnishi, A. (1992). Demyelinating peripheral neuropathy in Cockayne syndrome: a histopathologic and morphometric study. Brain Dev. *14*, 114-117.

Schmickel, R. D., Chu, E. H. Y., Trosko, J. E., and Chang, C. C. (1977). Cockayne syndrome: a cellular sensitivity to ultraviolet light. Pediatrics *60*, 135-139.

Selby, C. P., and Sancar, A. (1990). Transcription preferentially inhibits nucleotide excision repair of the template DNA strand in vitro. J. Biol. Chem. *265*, 21330-21336.

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.

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.

Sung, P., Higgins, D., Prakash, L., and Prakash, S. (1988). Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. EMBO J. *7*, 3263-3269.

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.

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

Tanaka, K., Satokata, I., Ogita, Z., Uchida, T., and Okada, Y. (1989). Molecular cloning of a mouse DNA repair gene that complements the defect of group A xeroderma pigmentosum. Proc. Natl. Acad. Sci. USA *86*, 5512-5516.

Thomas, D. C., Okumoto, D. S., Sancar, A., and Bohr, V. A. (1989). Preferential DNA repair of (6-4) photoproducts in the dihydrofolate reductase gene of Chinese hamster ovary cells. J. Biol. Chem. *264*, 18005-180010.

Venema, J. (1991). Gene specific repair of UV-induced DNA damage in mammalian cells. Ph. D. Thesis. State University of Leiden.

Venema, J., Mullenders, L. H. F., 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., Karcagi, V., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. F. (1991). Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. Mol. Cell. Biol. *11*, 4128-34.

Venema, J., Van Hoffen, A., Natarajan, A. T., Van Zeeland, A. A., and Mullenders, L. H. F. (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., Jaeken, J., Jaspers, N. G. J., Bootsma, D., and Hoeijmakers, J. H. J. (1992). Xeroderma pigmentosum complementation group G associated with Cockayne's syndrome. submitted.

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.

Weber, C. A., Salazar, E. P., Stewart, S. A., and Thompson, L. H. (1988). Molecular cloning and biological characterization of a human gene, *ERCC2*, that corrects the nucleotide excision repair defect in CHO UV5 cells. Mol. Cell. Biol. *8*, 1137-1146.

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.

Westerveld, A., Hoeijmakers, J. H. J., van Duin, M., de Wit, J., Odijk, H., Pastink, A., Wood, R. D., and Bootsma, D. (1984). Molecular cloning of a human DNA repair gene. Nature *310*, 425-429.

Wood, R. D., Robins, P., and Lindahl, T. (1988). Complementation of the xeroderma pigmentosum DNA repair defect in cell-free extracts. Cell *53*, 97-106.

Summary
Samenvatting

Summary

The genetic information of an organism is constantly subject to the introduction of mutations by environmental agents, faulty DNA metabolism or instability intrinsic to the DNA itself. To protect cells from a deleterious accumulation of alterations, various DNA repair mechanisms have evolved. The nucleotide excision repair (NER) process is one of the most versatile repair systems. It is able to remove a broad range of different DNA lesions, such as bulky chemical adducts, DNA cross links, and UV-induced lesions (cyclobutane pyrimidine dimers, (6-4) photoproducts, thymine glycols).

The importance of the NER process is illustrated by the classical DNA repair disorder xeroderma pigmentosum (XP). Due to a NER defect XP patients are extremely sun(UV)sensitive, they have pigmentation abnormalities in sunexposed areas and have a highly elevated (± 2000x) risk for the formation of skin tumors. Sometimes patients also suffer from progressive neurological degeneration. Patients with the hereditary repair disorder Cockayne's syndrome (CS) are, as XP patients, sun(UV)sensitive, but have no elevated risk for the formation of skin tumors. Other clinical characteristics of CS are cachectic dwarfism, retinal abnormalities and progressive neurological degeneration. CS cells perform - in contrast to XP cells - apparently normal repair of UV-induced pyrimidine dimers from the genome overall. They are, however, unable to recover their RNA synthesis after UV exposure. Recently, the repair defect in CS has been pinpointed to a subpathway of NER that primarily focusses on the preferential repair of lesions from (the transcribed strand of) active genes. This ingenious repair system normally permits cells to quickly resume transcription that was blocked by lesions in the template. Consequently, it has a major impact on the cellular resistance to UV and some other genotoxic agents.

To obtain a more detailed understanding of the molecular mechanism of the eukaryotic NER process, isolation and characterization of genes and gene products involved is indispensable. The work described in this thesis focusses on the isolation and characterization of the gene that corrects the NER defect of a UV-sensitive Chinese hamster ovary (CHO) mutant. The mutant, UV61, is a representative of complementation group 6 of the laboratory induced, NER-deficient rodent cell lines. It is only moderately UV-sensitive in comparison with mutant cell lines from groups 1 to 5, and harbors a partial deficiency in the repair of cyclobutane pyrimidine dimers but permits apparently normal repair of (6-4) photoproducts. The correcting gene, designated *ERCC6*, has been isolated through DNA-mediated transfection of UV61.

As described in *Chapter III*, the gene appeared to have a - for transfection cloning - respectable size of close to 100 kb, explaining the extremely low yield at which genomic repair-corrected transformants were obtained. The isolation of a revertant that has amplified an endogenous, mutated CHO *ERCC6* allele suggests that the UV61 mutation is leaky and can be overcome by gene amplification. This hypothesis is sustained by the recent observation of Lommel and Hanawalt (Mutat. Res. **255** (1991), 183-191) that UV61 is only partly disturbed in repair of cyclobutane pyrimidine dimers from the transcribed strand of the active *DHFR* locus.

In *Chapter IV*, we localize the *ERCC6* gene on chromosome 10, region q11-21. Almost simultaneously, Fryns *et al.* (Am. J. Med. Genet. **40** (1991), 343-344) reported a CS patient with a late onset form of the disease, carrying an interstitial deletion of the long arm of chromosome 10 (del(10)(q11.23-q21.2)). This prompted us to transfect the *ERCC6* gene to CS cells from both complementation groups A and B. As reported in *Chapter V*, the repair defect of CS group B cells is corrected upon introduction of the *ERCC6* gene, implying the ERCC6 protein to function in preferential repair of (the transcribed strand of) active genes. Thereby it indicates UV61 and other representatives of rodent group 6 to be the rodent equivalent of CS group B. The gene is spread over 21 exons and produces two lowly expressed, differentially polyadenylated mRNAs (*Chapter VI*). It encodes a protein of 1493 amino acids that contains seven motifs conserved among different DNA and RNA helicases. The region encompassing these helicase signatures exhibits a high degree of homology ($\pm$ 30% identity and 50-60% similarity at the amino acid level) to a number of other postulated helicases functioning in DNA repair, preservation of chromosome stability, transcription regulation, and other yet unknown processes, constituting a gene family. Mutation analysis of RNA and DNA of a CS-B patient revealed the presence of deleterious mutations in both alleles, presumably resulting in the production of non-functional proteins. This suggests that the ERCC6 protein is not essential for cell viability and specific for the preferential repair of active genes.

Samenvatting

De erfelijke informatie van een organisme wordt voortdurend bloot gesteld aan de mutagene effecten van omgevings factoren, onnauwkeurigheden in het DNA metabolisme of instabiliteit van het DNA zelf. Om de cel te beschermen tegen de desastreuze accumulatie van veranderingen, zijn verschillende DNA herstel mechanismen ontstaan. Het nucleotide excisie reparatie (NER) mechanisme is een van de meest veelzijdige herstel systemen. Het is in staat een grote verscheidenheid aan lesies te verwijderen, zoals lesies veroorzaakt door diverse chemische agentia en UV(zonlicht)-geïnduceerde lesies (cyclobutaan pyrimidine dimeren, (6-4) fotoprodukten, thymine glycolen).

Het belang van het NER proces blijkt uit de fenotypische gevolgen van de erfelijke afwijking xeroderma pigmentosum (XP), het klassieke voorbeeld van een DNA herstel defect. Als gevolg van een niet of slecht functionerend NER zijn XP patiënten extreem gevoelig voor zon(UV)licht. Aan de zon blootgestelde delen van de huid vertonen een abnormale pigmentatie, en de patiënten hebben een sterk verhoogde ($\pm$ 2000x) kans op het ontstaan van huid tumoren. De ziekte gaat soms gepaard met progressieve neurologische degeneratie. Patiënten met de erfelijke ziekte Cockayne's syndroom (CS) zijn - evenals XP patiënten - overgevoelig voor zonlicht, maar hebben geen verhoogde kans op het ontstaan van huid tumoren. De ziekte wordt verder gekenmerkt door ernstige groeistoornissen en een progressieve neurologische degeneratie. Cellen van deze patiënten zijn UV(zon)-gevoelig. De verwijdering van UV-geïnduceerde pyrimidine dimeren, wanneer gemeten in het "totale genoom", is - in tegenstelling tot XP - normaal. Recentelijk is gebleken dat het herstel defect in CS zich beperkt tot een onderdeel van het NER systeem dat zich primair richt op de preferentiële verwijdering van lesies uit (de getranscribeerde streng van) actieve genen. Dit ingenieuze herstel mechanisme stelt de cel in staat om het transcriptie proces, dat geblokkeerd was door lesies in de getranscribeerde streng, snel te hervatten. Als gevolg daarvan draagt dit systeem in belangrijke mate bij aan de resistentie van cellen voor verschillende genotoxische agentia.

Een beter inzicht in het moleculaire mechanisme van het eukaryote NER proces is mogelijk door isolatie en karakterisering van betrokken genen en gen produkten. Dit proefschrift beschrijft de isolatie en karakterisering van een humaan gen, dat het NER defect van een UV-gevoelige chinese hamster ovarium (CHO) mutant cellijn corrigeert. De mutant cellijn, UV61, behoort tot complementatie groep 6 van de NER-deficiënte knaagdier cellijnen. UV61 is slechts matig UV-gevoelig en heeft een partieel

defect in het herstel van cyclobutaan pyrimidine dimeren. Herstel van (6-4) fotoprodukten, daarentegen, verloopt normaal. Het corrigerende gen, *ERCC6*, is geïsoleerd door middel van DNA transfectie naar UV61. Het gen heeft een - voor transfectie clonering respectabele - grootte van bijna 100 kb; dit verklaart de lage frequentie waarmee herstel-proficiënte transformanten werden geïsoleerd (*Hoofdstuk III*). De isolatie van een revertant met een geamplificeerd, endogeen gemuteerd CHO *ERCC6* allel suggereert dat het geproduceerde eiwit nog gedeeltelijk kan functioneren. Deze hypothese wordt ondersteund door de recente bevinding van Lommel en Hanawalt (Mutat. Res. **255** (1991), 183-191) dat UV61 slechts gedeeltelijk gestoord is in het herstel van cyclobutaan pyrimidine dimeren in de getranscribeerde streng van het actieve *DHFR* gen.

ERCC6 is gelokaliseerd op chromosoom 10, regio q11-21 (*Hoofdstuk IV*). In een publikatie van Fryns *et al.* (Am. J. Med. Genet. **40** (1991), 343-344) wordt een CS patiënt gerapporteerd met een interstitiële deletie in de lange arm van chromosoom 10 (del(10)(q11.23-q21.2)). Dit heeft ons ertoe gebracht het *ERCC6* gen te transfecteren naar CS cellen van de beide complementatie groepen A en B. Zoals beschreven in *Hoofdstuk V*, kan *ERCC6* inderdaad het herstel defect van CS groep B cellen corrigeren. Dit suggereert dat het ERCC6 eiwit betrokken is bij het preferentiële herstel van (de getranscribeerde streng van) actieve genen. Daarbij betekent het dat UV61 en andere knaagdier groep 6 mutanten het knaagdier equivalent vormen van CS groep B. Het *ERCC6* gen is verdeeld in 21 exonen en expresseert twee, alternatief gepolyadenyleerde, mRNAs (*Hoofdstuk VI*). Beide - niet frequent voorkomende - mRNAs coderen voor een eiwit van 1493 amino zuren (*Hoofdstuk V*). Dit eiwit bevat zeven sequentie motieven die geconserveerd zijn tussen verschillende DNA en RNA helicases. De "helicase-regio" vertoont een hoge mate van homologie met een aantal andere veronderstelde helicases ($\pm$30% identiek en 50-60% vergelijkbaar op amino zuur niveau). De eiwitten gecodeerd door deze genfamilie zijn betrokken in DNA herstel, behoud van chromosoom stabiliteit, transcriptie regulatie en andere nog onbekende processen. Mutatie analyse bij een CS-B patiënt heeft de aanwezigheid van ernstige mutaties in beide *ERCC6* allelen aangetoond; het is aannemelijk dat beide allelen coderen voor niet-functionele eiwitten. Dit suggereert dat het ERCC6 eiwit niet essentieel is voor levensvatbaarheid en waarschijnlijk specifiek betrokken is bij preferentieel herstel van getranscribeerde genen.

Curriculum vitae
Nawoord

Curriculum vitae

August 1, 1964	Born in Amsterdam, The Netherlands
Sept. 1976 - June 1982	Education at the "Gemeentelijke Scholengemeenschap Woensel", Eindhoven, The Netherlands
Sept. 1982 - March 1987	Student at the Agricultural University in Wageningen, The Netherlands
Sept. 1985 - May 1986	Genetics: Department of Genetics, Agricultural University, Wageningen (Prof. Dr. Ir. J. H. van der Veen) Supervisor: Dr. Ir. P. de Boer
May 1986 - March 1987	Molecular biology: Department of Cell Biology and Genetics, Erasmus University, Rotterdam (Prof. Dr. D. Bootsma) Supervisor: Dr. G. C. Grosveld
March 1987	Doctoral degree in Molecular Science (cum laude)
April 1987	Ph. D. student (A.I.O.), Department of Cell Biology and Genetics, Erasmus University, Rotterdam Promotor : Prof. Dr. D. Bootsma Co-promotor : Dr. J.H.J. Hoeijmakers
December 1987	Cambridge Certificate of proficiency in English
September 1992 -	Research associate, Department of Cell Biology and Genetics, Erasmus University and Department of Radiotherapy, Subdivision of Clinical Radiobiology, Dr. Daniel den Hoed Cancer Center, Rotterdam

Nawoord

Het moge duidelijk zijn dat het tot stand komen van dit boekje geen verdienste is van mij alleen. Een woord van dank aan al diegenen die in de afgelopen jaren hebben meegewerkt, meegedacht of hun belangstelling hebben getoond is dan ook op zijn plaats. Enkelen wil ik met name noemen.

Allereerst mijn promotor Professor Bootsma, zonder wiens vertrouwen in mij dit proefschrift nooit verschenen was. Het is verbluffend hoe U het "managen" van een - nog steeds groeiende - groep weet te combineren met een grote interesse voor wat er achter de lab-tafel gebeurt. Een betere promotor had ik me niet kunnen wensen!

Zonder de bezielende leiding van Jan Hoeijmakers zou "Repair" niet zijn wat het nu is. Beste Jan, ik heb in de afgelopen jaren ontzettend veel van je geleerd. Jouw onvoorstelbare en nooit afnemende enthousiasme voor alles wat naar wetenschap ruikt is, vooral in "mindere" tijden, een belangrijke steun geweest. Ik ben er trots op dat ik ook na het voltooien van dit boekwerk met jouw club kan blijven samenwerken.

De clonering van *ERCC6* is begonnen met een ongelooflijke hoeveelheid transfectie werk, hetgeen voor het belangrijkste deel *op conto* kwam van Andries (Professor) Westerveld, Hanny Odijk en Jan de Wit. Zonder jullie geen *ERCC6*! Andries, bedankt dat je mijn manuscript hebt willen doorworstelen en hebt voorzien van waardevol commentaar. Ook Professor van Zeeland en Professor Niermeyer hebben die taak op zich willen nemen: veel dank voor alle goede suggesties. Jan en Hanny, jullie stonden mij aan het begin van mijn AIO-carrière terzijde in het kweekhok; ik ben blij dat jullie me aan het einde van mijn AIO-periode als paranimfen terzijde staan.

In mijn eerste "repair jaren" heb ik vele vragen op jou afgevuurd, Marcel (van Duin). Het heeft me altijd verbaast dat je voor vrijwel alles een oplossing had.

Kokie (Marcel Koken), ik mis je aansporende, veel gehoorde kreet "is het nou nog niet af" nu al (wanneer is dat boekje van jou eigenlijk klaar?). Gelukkig is Parijs niet het andere eind van de wereld, voor goede adviezen weet ik je nog steeds te vinden.

Koos, ook jouw hulp - op velerlei fronten (computer, artikelen, verhalen, etc.) - heb ik enorm gewaardeerd. Ik zal in mijn (radiobiologische) toekomst nog wel vaker een beroep op jouw kennis doen!

Sjozèf, jij bent voor een ieder die af en toe een computer aanraakt van onschatbare waarde! Er is niemand die zoveel domme vragen met zoveel geduld "à la minute" oplost. Behalve het oplossen van vragen deed je overigens veel meer: figuur 4B in hoofdstuk V zal ik niet licht vergeten!

Bij het praktische werk heb ik veel hulp gekregen van studenten. Patricia, Martha,

Emmy, Gerben, en Wout, bedankt voor alle inzet èn - niet te vergeten - de resultaten. Alain, ik ben blij dat je ons op het juiste moment bent komen versterken. Ik wacht met spanning op het volgende *ERCC6* boekje!
Wim, bedankt voor de vele discussies - zowel over de proeven (lang leve het agressie-papier!) als anderszins. Zonder jouw aanwezigheid zou de groep (voor mij) een heel andere zijn.
Ruud en Peer, ik heb dankbaar gebruik gemaakt van jullie computer programma. Zonder dat had het vinden van de *ERCC6* helicase domeinen stukken meer tijd in beslag genomen. Peer, de stress ijsjes en andere "afkoelende middelen" hebben het lab-leven in elk geval levendig gehouden!
Dan zijn er nog al die mensen wiens naam nooit boven een van onze artikelen prijkt, maar die desalniettemin van groot belang zijn voor de afdeling. Ons lab zou niet draaien zonder de afdeling fotografie. Mirko, zeker jij hebt een boel tijd aan dit boekje gespendeerd. Reuze bedankt! Ook de "meiden van de keuken" zijn niet weg te denken. Jopie, Elly en Joke jullie goede humeur is benijdenswaardig en aanstekelijk! Rita en Jeanette ("het" secretariaat), veel dank voor alle gezelligheid en al die "regel-dingetjes". Ook Rein, Mieke en Piet, heel hartelijk dank: bestellingen, reparaties, etc. etc., het gaat allemaal schijnbaar vanzelf...!
Zoals velen heb ook ik een boel aan (schoon)familie te danken. Bedankt voor alle belangstelling en sorry voor alle keren dat ik "moest" werken. Pappa en mamma, het is prettig om te weten dat er mensen zijn die ècht altijd voor je klaar staan. Bedankt voor alle goede raad, de "peptalk", en nog veel meer. Mam, zonder jou waren (o.a.) de ramen nu zò vies geweest dat we geen luxaflex meer nodig zouden hebben.
En tenslotte het thuisfront. Wil(bert), jouw begrip voor mijn werktijden, je schouder bij (o.a.) mislukte proeven, je oprechte interesse in mijn "werk", en al die andere "(kleine) dingen die het doen" waren - en zijn - heel belangrijk!

Christine