CLONING AND CHARACTERIZATION OF THE HUMAN DNA-EXCISION REPAIR GENE *ERCC-1*

KLONERING EN KARAKTERISERING VAN HET HUMANE DNA-EXCISIE HERSTELGEN *ERCC-1*

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
AAN DE ERASMUS UNIVERSITEIT ROTTERDAM
OP GEZAG VAN DE RECTOR MAGNIFICUS
PROF. DR. A.H.G. RINNOOY KAN
EN VOLGENS BESLUIT VAN HET COLLEGE VAN DEKANEN.
DE OPENBARE VERDEDIGING ZAL PLAATSVINDEN OP
WOENSDAG 2 NOVEMBER 1988 OM 13.45 UUR

DOOR

MICHEL VAN DUIN

GEBOREN TE ROTTERDAM

PROMOTIECOMMISSIE

PROMOTOR:	PROF. DR. D. BOOTSMA
OVERIGE LEDEN:	PROF. DR. A.J. VAN DER EB
	PROF. DR. P. VAN DE PUTTE
	DR. E.C. ZWARTHOFF
CO-PROMOTOR:	DR. J.H.J. HOEIJMAKERS

Dit proefschrift werd bewerkt binnen de vakgroep Celbiologie en Genetica van de Faculteit der Geneeskunde, Erasmus Universiteit Rotterdam. Het onderzoek is financieel gesteund door MEDIGON, Stichting voor Medisch onderzoek en Gezondheidsonderzoek in Nederland (contract no. 900-501-091) en het 'Radiation Protection Programme' van de Commissie van Europese Gemeenschappen (contract no. BI6-E-141-NL).

Aan mijn ouders
Voor Mike

CONTENTS

ABBREVIATIONS

APRT	adenine phosphoribosyltransferase
agpt	aminoglycosyl phosphotransferase
ASE-1	antisense *ERCC-1*
ASR10	antisense *RAD10*
AT	ataxia telangiectasia
BHK	baby hamster kidney
BS	Bloom's syndrome
CDC	cell division control
cDNA	complementary deoxyribonucleic acid
CHO	Chinese hamster ovary
CS	Cockayne's syndrome
DNA	deoxyribonucleic acid
E.coli	*Escherichia coli*
Ecogpt	*E.coli* guanine phosphoribosyltransferase
ERCC	*E*xcision *R*epair *C*ross *C*omplementing
FA	Fanconi's anemia
gpt	guanine phosphoribosyltransferase
HAT	hypoxanthine, aminopterin, thymidine
HPRT	hypoxanthine phosphoribosyltransferase
J	Joule
kb	kilobase
kD	kilo Dalton
mer-	methylation repair deficient
MMC	mitomycin C
MPA	mycophenolic acid
mRNA	messenger ribonucleic acid
MW	molecular weight
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
S. pombe	*Schizosaccharomyces pombe*
SV40	simian virus 40
UDS	unscheduled DNA synthesis
UV	ultraviolet light
T4	bacteriophage T4
TK	thymidine kinase
XP	xeroderma pigmentosum

I GENERAL INTRODUCTION

I.1 INTRODUCTION

Spontaneously and environmentally induced structural DNA alterations are a constant threat to the genetic integrity and genomic stability of living creatures. A plethora of chemical and physical agents are highly reactive with DNA and have the potential to introduce a large number of different DNA damages in the cellular genome. In order to cope with the inevitable exposure to such inimical elements, evolution has provided most organisms with defence systems with DNA repair functions. Imperfect correction or absence of repair of DNA lesions can lead to mutation, recombination, gene rearrangement or amplification. If such changes occur at specific genomic sites this can evoke activation of proto-oncogenes which are implicated in subsequent carcinogenesis (Hanawalt and Sarasin, 1986; Duesberg, 1987). This is amply illustrated by members of the *ras* proto-oncogene family which have been found to underlie the process of carcinogenesis in a considerable number of tumors (Barbacid, 1986; 1987). Recently, in approximately 40% of colorectal cancers activated Ki-*ras* genes have been detected (Bos et al. 1987; Forrester et al. 1987) whereas human carcinoma's of the exocrine pancreas seem almost exclusively induced by Ki-*ras* activation (Almoguera et al. 1988). In all cases single point-mutations are responsible for the acquired transforming properties of the corresponding *ras* proteins.

The first experimental evidence for a correlation between defective DNA repair and carcinogenesis dates back twenty years, when Cleaver reported that cell lines from patients with the cancer prone hereditary disease xeroderma pigmentosum (XP) were defective in the excision repair of ultraviolet light (UV) induced DNA damage (Cleaver, 1968; 1969). XP is a world wide occurring, rare autosomal recessive disease with a frequency of one to ten patients per million (see Kraemer and Slor, 1985; Kraemer, 1987 for an extensive review on XP). In addition to several clinical abnormalities, XP patients have a high risk for developing skin tumors on sunlight exposed parts of the body whereas also indications were found for an increased frequency of internal neoplasias (Kraemer et al. 1984). The UV hypersensitivity of XP fibroblasts was already notified earlier in the 1960's (Gartler, 1964) at the same time when nucleotide excision repair of UV induced pyrimidine dimers

in *Escherichia coli* (*E.coli*) was discovered (Boyce and Howard- Flanders, 1964). Several years later evidence for excision repair in mammalian cells was found (Regan et al. 1968; Cleaver and Painter, 1968) and soon the relationship between defective DNA excision repair and predisposition to cancer in XP became apparent (Cleaver,1968).

This finding did not disclose the genetic defect in XP, nor did it elucidate the mechanism of DNA excision repair in human cells. It rather appeared to be a prelude to the uncovering of a very complex DNA repair process in which many different gene products operate. Cell fusion experiments with cultured fibroblasts from different patients have provided evidence for ten XP complementation groups, designated XP-A to I and XP-V (variant) (De Weerd-Kastelein et al. 1972; Fischer et al. 1985; Kraemer, 1987). Cell lines from all groups share a hypersensitivity to UV and with the possible exception of XP-V an impaired capacity to efficiently remove UV lesions (see also section I.2.3).

Increased susceptibility to DNA damage appears to be a common denominator for a number of other rare human genetic disorders. Patients with ataxia telangiectasia (AT), Fanconi's anemia (FA), Bloom's syndrome (BS) and Cockayne's syndrome (CS) share an elevated cellular sensitivity to physical and chemical DNA damaging agents and have apart from CS a high risk to develop malignancies at early stages of life (Friedberg et al. 1979; Lehmann, 1982a; Kraemer, 1983; Hanawalt and Sarasin, 1986). Like for XP, the molecular basis of the genetic defect in these abnormalities is completely unknown and furthermore characterized by a considerable genetic heterogeneity. For AT, FA and CS, respectively four, two and three genetically distinct groups have been found (AT:Jaspers et al. 1985; 1988; FA:Zakrzewski and Sperling, 1980; Duckworth-Rysiecki et al. 1985; CS: Lehmann, 1982b). Only in case of BS a single mutation seems to underlie the clinical manifestations of the disease. Recently, evidence was reported for abnormalities in DNA ligase I in several independent BS cell lines which seems in accordance with the high rate of sister chromatid exchange found in BS cells (Willis and Lindahl, 1987; Chan et al. 1987; Willis et al. 1987). BS cell lines from individuals of different populations share aberrant ligase I proteins although the characteristics of the observed aberrations differ. This may suggest that different mutations in the ligase I gene underlie BS. Interestingly, however, Sirover and coworkers have recently reported evidence for an altered uracil DNA glycosylase in five BS cell lines using a set of monoclonal antibodies against this enzyme, which

has a normal activity in BS cells (Vollberg et al. 1987; Seal et al. 1988). In the Askenazic Jewish population the number of BS patients is relatively high and it is believed that BS patients in this population are derived from a single ancestor (German, 1984). This renders it possible that certain protein polymorphisms occur at high frequencies which might have consequences for the variety in antigenic determinants and explain the results found with the monoclonal antibodies. However, the fact that similar data have been obtained with BS cells from an American black (Vollberg et al. 1987) seems to rule out this possibility. Cloning and sequence analysis of the ligase I and uracil DNA glycosylase genes will have to reveal their relation to the genetic defect in BS. At present it is not excluded that the observed ligase I deficiency and altered uracil DNA glycosylase in BS are secondary to the primary defect which has yet to be determined.

For the other DNA repair deficiency syndromes hypersensitivity to DNA damage is an indirect indication for defective DNA repair. In case of CS, which is characterized by an elevated cellular sensitivity to UV, this has recently been substantiated by Mayne et al. (1988a) who reported evidence for impaired repair of lesions in actively transcribed genes (see also section I.3). FA cells exhibit an increased sensitivity to DNA-cross-linking agents which is thought to be due to DNA interstrand cross-links. However, experiments to confirm this have yielded conflicting results (Papadopoulo et al. 1987 and Refs. therein). AT cells are sensitive to ionizing radiation and other DNA strand breaking agents suggesting a defect in DNA repair. However, direct experimental evidence for this assumption is lacking. The absence of inhibition of DNA synthesis in AT cells after exposure to X-rays rather suggests that the primary defect in AT is at the level of cell cycle control (see Jaspers, 1985).

RFLP (Restriction Fragment Length Polymorphism) and genetic linkage studies (White and Lalouel, 1987) are a straightforward approach to map the genetic defect of the above mentioned DNA repair syndromes to a particular genomic site which can subsequently be used to identify and clone the putative DNA repair genes. However, only in case of AT efforts in that direction have been undertaken, which resulted in the exclusion of 15% of the human genome to harbor the defective AT complementation group AB gene (Gatti et al. 1987; Lehmann et al. 1987). In general, RFLP studies on DNA repair syndromes are hampered by the very low number of patients afflicted with these diseases and the considerable genetic heterogeneity. Therefore, linkage

studies seem less applicable to localize defective genes in the various repair syndromes.
Microcell mediated transfer of human chromosomes to XP-A cells has been reported, but did not permit the chromosomal localization of the human XP-A gene (Ishizaki et al. 1981; Schultz et al. 1987). Using the same experimental strategy human chromosome 15 was recently found to correct the defect in XP-F cells (Schultz et al. 1988). In fusion experiments of XP-A fibroblasts with cytoplasts of human/Chinese hamster hybrids a concordance of an XP-A correcting factor with human chromosome 1 was found (Keijzer et al. 1987). However, the authors do not exclude the possibility that the observed XP-A correction is the result of a combined action of human and Chinese hamster proteins rather than the XP-A protein alone.

In some instances an association of complementation groups of different genetic disorders has been observed. The single patients of XP groups B and H also displayed diagnostic criteria for CS (Robbins et al. 1974; Moshell et al. 1983). XP-B fibroblasts have been classified as the third CS complementation group (Lehmann, 1983). Furthermore, some trichothiodystrophy (TTD) patients were recently found to be hypersensitive to UV and corresponding TTD cell lines failed to complement the genetic defect in XP-D cells (Stefanini et al. 1986; Rebora and Crovato, 1987). Although intragenic complementation can not be excluded, the combined occurrence of diagnostic characteristics for different genetic disorders might be the result of large chromosomal aberrations involving more than one locus. However, uptil now cytogenetic studies have not provided evidence for this assumption (W. Keijzer, pers. comm.).

The cancer proneness of human diseases with an impaired resistance against DNA damage indicates that it is very important to investigate the process of DNA repair in mammalian cells. In order to understand the connection between DNA damage and carcinogenesis detailed information is required about the function of the genes and proteins that survey the integrity of the genetic information and attack aberrant structural DNA alterations.
The DNA excision repair pathway is one of the major DNA repair processes in mammalian cells (Friedberg, 1985) and it is evident that cultured fibroblasts from XP patients are valuable tools to study this repair system. In addition, the last decade has witnessed the isolation of a large number of *in vitro* mutagenized UV-sensitive rodent cell lines with XP-like features (Thompson, 1985; Collins and Johnson, 1987 and refs. therein). To date, at least seven

complementation groups of UV sensitive Chinese hamster cells have been identified (Thompson et al. 1981; Thompson and Carrano, 1983; Thompson et al. 1987a; Zdzienicka et al. 1988). With respect to a possible relationship between the 9 excision defective XP and 7 rodent complementation groups limited information is available and evidence for overlap between the human and rodent classifications is lacking (Thompson et al. 1985a). Hence, it is not excluded that the excision of UV-induced DNA adducts in mammalian cells is a highly complex process in which 16 or more gene products are involved. The idea that a large number of gene products operate in DNA excision repair is further supported by recent findings on this mode of DNA correction in the baker's yeast *Saccharomyces cerevisiae* and the recently elucidated mechanism of excision repair in *E.coli* (see below, section I.2).

Currently, efforts to identify the defective excision repair genes or gene products in UV-sensitive cell lines of human and rodent origin involve two main strategies. XP cells are exploited to develop assay systems to purify 'XP-factors'. Using different test systems several laboratories have reported transient correction of the defect in several XP complementation groups with crude and partially purified protein extracts of wild type cells (Yamaizumi et al. 1986; Nishida et al. 1987; Kaufman and Briley, 1987; Wood et al. 1988; Hoeijmakers, 1988; see also section I.2). The second approach to identify mammalian repair genes involves correction of the UV-sensitive phenotype of cultured cells by means of DNA mediated gene transfer. Recently, transfection experiments with UV-sensitive Chinese hamster ovary (CHO) cells have yielded the isolation of several human DNA repair genes (see Chapter II) (Westerveld et al. 1984, appendix paper I; Weber et al. 1988; Hoeijmakers et al. 1988a). As will become apparent throughout this thesis, well characterized CHO DNA repair mutants are of major importance for DNA repair research and invaluable tools to elucidate the intricacies of DNA excision repair in mammalian cells.

I.1.1 Aim of this thesis

It is the aim of the experimental work presented in this thesis to identify and characterize genes and gene products involved in mammalian DNA excision repair. To this end UV-sensitive CHO cells are used in transfection experiments to isolate human genes that complement the defective capacity to repair UV- induced DNA insults.

The next sections of chapter I are devoted to important recent aspects of DNA repair research with the emphasis on the process of excision repair. More comprehensive surveys of the literature on DNA repair can be found in several excellent reviews (Friedberg, 1985; Strauss, 1985; Sedgwick, 1986; Collins et al. 1987).

Chapter II gives the current 'state of the art' with respect to the identification and cloning of human genes involved in DNA excision repair. Particular attention is given to the strategy of DNA mediated gene transfer which is at present the most straightforward approach for the isolation of *ERCC*-genes that confer wild type DNA repair functions to UV-sensitive CHO cells. The experimental work presented in this thesis concerns the cloning and characterization of the first representative of this group which is described in appendix papers I to VII.

I.2 BIOCHEMISTRY OF NUCLEOTIDE EXCISION REPAIR

Nucleotide excision repair can be defined as the process in which a part of the damaged DNA strand including a bulky DNA lesion is excised followed by DNA repair synthesis to fill the single strand gap (Friedberg, 1985). A number of different steps can be discerned in this mode of repair:

1. Recognition of the damage.
2. Endonucleolytic action adjacent to the damage to 'open' the damaged strand.
3. Exonucleolytic action to remove a stretch of nucleotides including the damage.
4. Resynthesis to fill the single strand gap.
5. Ligation to rejoin the newly synthesized strand with the pre-existing DNA.

The multiprotein character of this process is easy to envisage and can also be deduced from the large number of different UV- sensitive mutants among different organisms. In *E.coli* the mechanism of DNA excision repair has been elucidated in great detail whereas for eukaryotes - especially yeast -

information on the genes and proteins involved is gradually emerging. A comparison of the number of *E.coli*, yeast and mammalian mutants with defective excision repair strongly suggests that there has been an evolutionary pressure to generate repair systems with increasing complexity.

I.2.1 Excision repair in *E.coli*

The isolation of UV-sensitive *E.coli uvr* mutants (Howard-Flanders et al. 1966; Ogawa et al. 1968) has provided the basis for the identification and cloning of bacterial excision repair genes and the subsequent unraveling of the excision machinery (Seeberg et al. 1976). The *uvrA*, *B* and *C* genes encode proteins of 103874 (Husain et al. 1986), 76118 (Arikan et al. 1986) and 66038 Dalton (Sancar et al. 1984) which have been purified to homogeneity. This has allowed the *in vitro* reconstitution of the *uvrABC* excision nuclease activity and detailed analysis of the working mechanism of the protein complex (Sancar and Rupp, 1983; Yeung et al. 1983). A schematical model of the *uvrABC* mediated excision of a dimer (Oh and Grossman, 1986; Grossman et al. 1988) is shown in Figure I.1. The *uvrA* protein has an ATP dependent DNA binding activity (Seeberg et al. 1982). ATP binding leads to a conformational change of *uvrA* that favours the formation of protein dimers which have a high affinity for damaged sites in duplex DNA. In this dimer configuration *uvrA* can form a stable complex with *uvrB* and *uvrC*. The *uvrAB* complex has an ATP dependent helicase activity in a 5' to 3' direction (Oh and Grossman, 1987). The recently reported *in vivo* kinetics of repair of UV damage suggest that the *uvrAB* helicase is scanning the DNA at the expense of ATP (Elliot et al. 1988). Due to a higher affinity for DNA helix distortions the complex stops at a damaged site and the *uvrC* protein joins the complex to mediate incision of the damaged strand at 7 nucleotides 5' and 3 or 4 nucleotides 3' of the lesion (Sancar and Rupp, 1983). The release of the excised oligonucleotide and turnover of the *uvrC* protein requires the combined action of the *uvrD* and *polA* gene products: helicase II and DNA polymerase I respectively (Caron et al. 1985; Husain et al. 1985). The resulting single strand gap is subsequently filled in by polymerase I activity and finally closed by a ligation step. Following exposure to DNA damage, in *E.coli* at least 17 genes are induced as part of the SOS response (reviewed by Walker, 1985) including the *uvrA*, *B* and *D* genes. This 'stress' reaction leads to a temporarily high level of 'repairosomes' (i.e. the combined activity of the

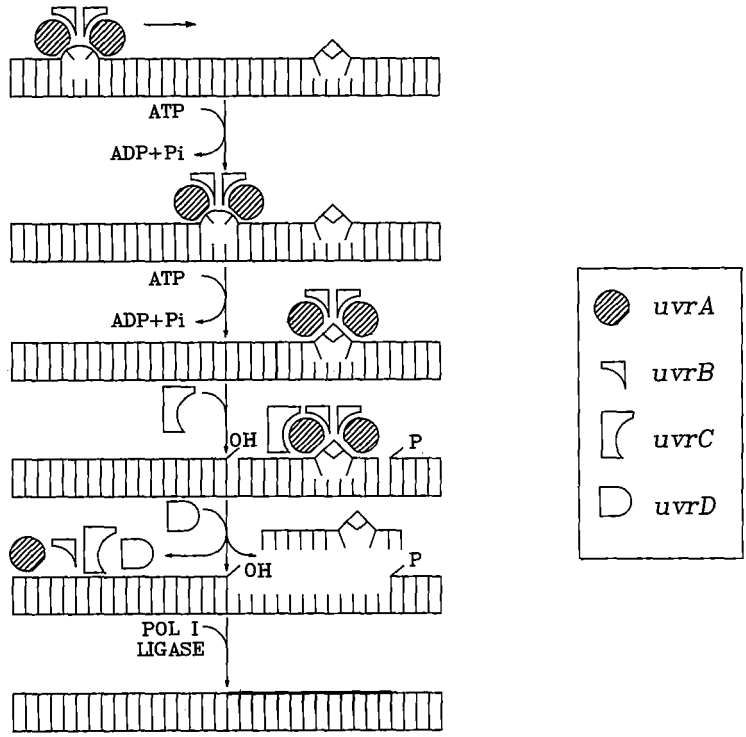

Figure I.1
Possible working mechanism of the *E.coli uvrABC* excision nuclease (see text for further details).

uvr and *polI* gene products) in order to efficiently attack the chromosomal DNA damage.

The *uvrABC* nucleotide excision system acts in a virtually non-specific manner in that it removes a wide variety of structurally unrelated bulky DNA adducts (e.g Sancar et al. 1985; Van Houten and Sancar, 1987). Distortion of the helical structure seems the only trigger for the excision activity.

I.2.1.1 Introduction of *uvrABC* in XP cells

Several laboratories have investigated the possibility whether the bacterial

repair system can correct the repair defect in XP. However, uptil now the introduction of the *E.coli uvr* components either by nuclear microinjection (Zwetsloot et al. 1986) or by DNA transfection (Dickstein, 1988) did not effect measurable changes in the UV response of the treated cells.

I.2.2. Excision repair in yeast

A large number of *Saccharomyces cerevisiae* mutants (>100) with abnormal sensitivity to DNA damage have been isolated (see Haynes and Kunz, 1981 for a review on yeast repair mutants). At least 10 UV-sensitive mutants, belonging to the so-called *RAD3* epistasis group have been identified that are mutated at genetic loci involved in nucleotide excision repair (Friedberg, 1987; Friedberg et al. 1987a). The *rad1*, *rad2*, *rad3*, *rad4*, *rad10* and *mms19* mutants are highly defective in the incision of DNA containing pyrimidine dimers or interstrand cross-links (Reynolds and Friedberg, 1981; Wilcox and Prakash, 1981; Miller et al. 1982) whereas the other excision repair mutants (*rad7*, *rad14*, *rad16*, and *rad23*) display a moderate capacity for strand incision which corresponds with the UV-sensitivity of the different mutants (Friedberg, 1985). The *RAD1*, *RAD2*, *RAD3*, *RAD4*, *RAD7* and *RAD10* genes have been cloned and except for *RAD4* the primary structure of the proteins has been predicted from the nucleotide sequence. Details of the *RAD* proteins are summarized in Table I.1.

The individual characterization of the *RAD* proteins does not allow yet to device a model for yeast DNA excision repair and it is not known whether a concerted action of the *RAD* proteins is responsible for excision repair. The finding that the *RAD23* gene can complement an N-terminally truncated *RAD7* protein lacking 99 amino acids is the first concrete indication for physical interaction between individual gene products involved in eukaryotic DNA repair (Perozzi and Prakash, 1986). Any analogy between yeast excision repair and the bacterial *uvrABC* system is purely speculative although perhaps some common features exist. The molecular weights of the bacterial and yeast proteins show some resemblance and particularly the recent finding that *RAD3* is a DNA dependent ATPase with helicase activity is striking (Sung et al. 1987a;b). It is possible that the helicase action of the *RAD3* and *E.coli uvrD* protein serve a similar function in the excision process. Like the bacterial *uvr* genes, the *RAD* genes are constitutively expressed at very low levels (Friedberg, 1987). Although no evidence exists for the presence of an

Table I.1.
Summary of properties of cloned yeast excision repair proteins.

Protein	Mw (Dalton)	Characteristics	Reference*)
RAD1	126360		1
RAD2	111081	UV-inducible	2,3,4
RAD3	89779	ATPase, Helicase	5,6,7,8
RAD4	87100		9,10
RAD7	63705	interaction with *RAD23*	11
RAD10	24310	homology with human *ERCC-1*	12,13
RAD23	not known	interaction with *RAD7*	11,14

*) 1:Reynolds et al. 1987; 2:Nicolet et al. 1985; 3:Robinson et al. 1986; 4:Madura and Prakash, 1986; 5:Reynolds et al. 1985; 6:Naumovski et al. 1985; 7:Sung et al. 1987a; 8:Sung et al. 1987b; 9:Fleer et al. 1987; 10:Burtscher et al. 1988b; 11:Perozzi and Prakash, 1986; 12:Reynolds et al. 1985; 13: van Duin et al. 1986, appendix paper II; 14:McKnight et al. 1981.

SOS response, for *RAD2* a 2-6 fold induction upon exposure to a variety of DNA damaging agents has been reported (Robinson et al. 1986; Madura and Prakash, 1986). However, the significance of this induction for the efficiency of DNA repair is still obscure.

The *RAD3* protein is at present the best characterized yeast excision repair protein. Interestingly, mutagenesis studies have revealed that the *RAD3* protein has a dual function. In addition to its absolute requirement for the incision of DNA during excision repair, the *RAD3* gene is essential for the viability of haploid yeast cells. Those distinct functions could be specifically eliminated in mutation studies (Naumovski and Friedberg, 1986; 1987).

Significant amino acid homology between *RAD10* and the mammalian excision repair protein *ERCC-1* was discovered (van Duin et al. 1986, appendix paper II, see also section II.4.3.1) suggesting an equivalent function. In support of this assumption, recently partial complementation of the UV-sensitive phenotype of a CHO mutant defective in *ERCC-1* by the yeast *RAD10* gene was reported (Burtscher et al. 1988a;b).

Introduction of several *RAD* genes in *E.coli* mutants had no effect on UV-sensitivity and also transfection of *RAD3* and *RAD10* constructs to XP-A cells were negative in this respect (Friedberg, 1987).

Large scale purification of the *RAD* proteins is a point of common interest in several laboratories in order to attempt *in vitro* reconstitution of a yeast 'repairosome' as has been shown for the *E.coli* excision repair system.

I.2.3 Mammalian nucleotide excision repair

At the single cell level this mode of DNA repair can be measured by incubating UV irradiated cells in medium containing ^{3}H-thymidine. By means of autoradiography the level of UV-induced unscheduled DNA synthesis (UDS) can be visualized as silver grains over the nucleus. The UV-sensitivity of XP cells and CHO mutants is in many cases accompanied by reduced levels of UDS indicating a defective excision of DNA damage.

In mammalian cells two types of excision repair have been identified i.e. short patch and long patch repair (Regan and Setlow, 1974; Hanawalt et al, 1979). The patch size, the length of the newly synthesized DNA strand, is dependent on the type of damage. The short patch repair process is responsible for restoration of 'X-ray like' lesions and damage induced by alkylating agents and requires the insertion of only a few nucleotides. UV-induced DNA damage and other bulky lesions are subject to long patch repair which means that a stretch of 30-100 nucleotides or more are renewed (Edenberg and Hanawalt, 1972; Walker and Th'ng, 1982; Th'ng and Walker, 1983; Smith and Okumoto, 1984).

Many laboratories have addressed the question whether DNA polymerase α or β provide strand renewal and which mechanism accounts for the preceeding excision of the damaged DNA (see Keyse and Tyrrell, 1985; Th'ng and Walker, 1985; Wilson et al. 1988 and refs. therein). Experiments with aphidicoline, cytosine arabinoside which inhibits DNA-polymerase α and dideoxythymidine which specifically blocks polymerase β, gave indications that both enzymes are to some extend involved in excision repair. It remained unclear whether each enzyme alone could fill in the repair gaps (Th'ng and Walker, 1985). Recently, Nishida et al. (1987) have succeeded in isolating DNA polymerase δ from HeLa cells. The protein had a relative molecular weight of 220,000 daltons, which is similar to the calf thymus DNA polymerase δII (Crute et al. 1986). In a system of permeabilized cells purified DNA polymerase δ could in combination with T4 endonuclease V, which specifically cuts at the site of pyrimidine dimers (Nakabeppu and Sekiguchi, 1981), restore UV-induced DNA repair synthesis without the presence of

polymerase α or β. Interestingly, and in contrast to polymerase α and β , the purified polymerase δ also harbors a 3' to 5' exonuclease activity which might be crucial in the initiation of strand removal in the 3' direction so that a proper primer is formed for the polymerase activity. The polymerase δ activity is boosted by the presence of DNA polymerase α (Nishida et al. 1987). Hence, it is conceivable that a concerted action of polymerases subsequent to the initiation by polymerase δ is responsible for strand renewal. On the other hand, it is possible that different repair processes, i.e. preferential repair of active genes (see section I.3) or random DNA damage removal, are mediated by different polymerases. A strand displacement activity as observed for DNA polymerase β (Mosbaugh and Linn, 1983) or a yet to be determined additional 5' to 3' exonuclease activity might prepare the patch for DNA synthesis in 5' to 3' direction.

Damaged strand excision and DNA repair synthesis are late steps in the excision process and triggered by damage recognition and DNA incision. By measuring the number of single strand breaks after UV-irradiation and performing the T4 endo-assay in which cuts are made at the site of dimers, it appears that most UV- sensitive mutants (XP and CHO) that have been identified are defective in the early steps of the excision pathway. However, the biochemistry of these initial steps of eukaryotic repair is still unknown.

I.2.3.1 Excision repair in XP

XP cells of complementation groups A to I are characterized by a reduced level of UV-induced UDS which ranges from 0 to 50 % of the wild type level (Cleaver and Karentz, 1987). In addition also kinetic analysis of the repair reaction revealed differences among the various complementation groups (Keijzer et al. 1982; Giannelli et al. 1982; Squires and Johnson, 1988). It was shown by alkaline elution studies that XP cells have a reduced capacity to incise their DNA following UV exposure (Fornace et al. 1976). Using the prokaryotic T4 endonuclease or *Micrococcus luteus* endonuclease in all nine XP complementation groups (A to I) this defect could be bypassed resulting in enhanced excision levels (Tanaka et al. 1975; 1977; Hayakawa et al. 1981; De Jonge et al. 1985). Also introduction of the T4 *denV* gene coding for the dimer specific T4 endonuclease in immortalized cells of XP-A and D conferred partial UV resistance (Valerie et al. 1987; Arrand et al. 1987). Hence, it was inferred that, although to a different extent, all nine XP complementation groups have a defect prior to or in the incision step of the DNA excision

repair pathway. Therefore, identification of the genetic defect in XP will be an important contribution to the understanding of the mammalian excision repair process.

Recently, several laboratories have succeeded in generating *in vitro* assays for the defective repair in XP. By microinjection of heterologous or wild type protein extracts into XP fibroblasts of various complementation groups the defect could be restored as measured by UDS (De Jonge et al. 1983; Vermeulen et al. 1986; Yamaizumi et al. 1986; Hoeijmakers et al. 1987). This transient microinjection assay is currently employed for purification of the 'correcting factor' which is defective in XP-A and XP-C. It was recently found that calf thymus contains a relatively high XP-A correcting activity. Therefore this bovine tissue promises to be a good source for large scale purification of the XP-A factor which in preliminary experiments was found to have an estimated molecular weight of $\pm$ 45 kilodalton (Hoeijmakers et al. 1988a). Interestingly, correction of the XP-A and G defect could also be achieved after microinjection of poly(A)+ RNA of HeLa cells (Legerski et al. 1984), thereby opening a more direct way to isolate the mRNA and cDNA responsible for the XP correction. By microinjection of size fractionated poly(A)+ RNA fractions it was found that the XP-A correcting protein is encoded by a mRNA of 1200-1400 base pairs (Hoeijmakers et al. 1988a) which is in accordance with the putative size of the correcting bovine protein.

In several laboratories permeabilized human fibroblasts have been used for developing a functional assay for (XP) factors involved in the excision repair process. Kaufman and Briley (1987) showed that saponin treated cells accumulate single strand breaks in the absence of deoxyribonucleoside triphosphate precursors . This was not observed in XP-A cells but could be induced with the *M.luteus* endonuclease or DNase I, indicating that permeabilized XP cells might be used as a 'test tube' to analyse crude and partially purified protein fractions of normal cells. Nishida et al. (1987) have preliminary evidence that the repair activity in cytosol depleted XP cells can be recovered by exogenously supplied crude cell extracts provided that purified polymerase δ is also present. Hence, this looks an attractive system for purifying excision repair proteins.

Recently, Wood et al. (1988) managed to develop a cell free repair assay which monitors the UV-dependent incorporation of ^{32}P-labeled dATP in UV-irradiated plasmid DNA. Extracts from XP-cells of all complementation groups examined were inactive in this repair assay but showed repair syn-

thesis in the presence of the *M.luteus* endonuclease. Moreover, a mixture of heterologous XP extracts had normal repair activity indicating that the *in vitro* repair assay can also serve as an *in vitro* complementation assay (Wood et al. 1988). In agreement with earlier studies (Dressler and Lieberman, 1983; Kaufman and Briley, 1987) a strong ATP dependence for the *in vitro* reaction was observed (Wood et al. 1988). Partially purified extracts that were active in the microinjection assay with XP-A cells were also positive in the cell free repair assay (Eker and Wood, pers. comm.) which indicates that both *in vitro* and *in vivo* repair assays have a comparable specificity. The results obtained with the cell free repair assay are difficult to reconcile with previously reported work of Mortelmans et al. (1976) and Kano and Fujiwara (1983) which indicated that XP cell extracts were capable to excise UV damage from naked DNA or chromatin of normal and heterologous XP cells in an ATP independent fashion . These results suggest that the chromatin structure plays an important role in excision repair and that the repair defect in XP concerns recognition or accessibility of the damage rather than the capacity to perform the incision step. It is clear that further experiments are required to clarify the apparent discrepancy between the conclusions of the different experimental approaches to study the defect in XP.

XP-variant (XP-V) cells are characterized by a moderate UV- sensitivity and near normal UDS levels (Cleaver et al. 1972) and therefore were in contrast to the classical XP-A to I groups not suspected to harbor a defect in excision repair. A deficiency in a repair mechanism termed post-replication or daughter strand repair (Lehmann et al. 1975) was thought to underlie the clinical manifestations of XP-V patients which are not much different from those belonging to other complementation groups. Surprisingly, Wood et al. (1988) recently reported that XP-V cells were as deficient as classical XP cells in the cell free repair assay. These results suggest that XP-V cells are not as exceptional as suspected but rather indicative of a tenth XP complementation group with defective excision repair. The data of Kondo et al. (1987) who reported similar UDS levels for XP-E and XP-V cells and an extensive study of Thielmann et al (1985) on repair parameters of many XP-V cells support this assumption. Also in studies on the UV-induced mutation frequency and transformation XP-V cells behaved as 'normal' XP's (Maher et al. 1976; McCormick et al. 1986). Hence, it appears that the results of a number of independent studies are in accordance with Wood's finding. At present it is too early to explain the fact that residual repair activity of XP-V cells

as measured with the UDS assay is not noticed in the *in vitro* repair assay. Possibly, the latter assay lacks regulatory factors that are responsible for the residual UDS produced by a cripple repair machinery *in vivo*.

I.3. PREFERENTIAL REPAIR OF DNA DAMAGE

UV-induced DNA damage has a terminating effect on transcription (Sauerbier and Hercules, 1979) which is probably underlying the lethal effects of cellular UV exposure. In transfection experiments with XP cells even the presence of one dimer completely abolished the activity of a transfected marker gene (Protic-Sabljic and Kraemer, 1985). UV-irradiation of normal cells inhibits the rate of RNA synthesis which recovers in a few hours to original values although after that period still a considerable number of dimers can be detected (van Zeeland et al. 1981). This relatively rapid recovery of transcription was not observed in UV-sensitive cell strains of CS and XP patients (Mayne and Lehmann, 1982). Using different assays CS cells were found to have a normal excision repair (Mayne, 1982). To explain the inhibition of RNA synthesis recovery in CS cells it was postulated that mammalian cells harbor a mechanism which preferentially repairs DNA lesions in transcriptionally active regions which is disturbed in CS cells (Mayne, 1982; 1984; Mayne and Lehmann, 1982). Several other independent arguments were suggestive for the presence of a heterogeneity in repair of genomic DNA. Firstly, human and CHO cells exhibit a similar UV-survival, however at 24 hrs after irradiation (5-10 J/m^2) in cultured CHO cells still 80-90 % of induced dimers are present whereas in human cells 80 % is removed (Zelle et al. 1980; van Zeeland et al. 1981). Secondly, in monkey cells a difference in excision repair in repetitive DNA (non-transcribed) compared to the bulk DNA was observed (Zolan et al. 1982; 1984; Leadon et al. 1983; Leadon and Hanawalt, 1984) Thirdly, it was found that the residual repair activity of XP-C cells (10-20% of wild type) is localized in specific genomic regions (Mansbridge and Hanawalt, 1983) that are associated with the nuclear matrix (Mullenders et al. 1984). In this respect it is worth-while mentioning that Robinson et al. (1983) reported that transcriptionally active genes are associated with the nuclear matrix.

Hanawalt and coworkers have recently developed a technique that allows the detection of repair of DNA damage at the single gene level (see Hanawalt, 1986; Smith, 1987; Bohr; 1987). Using this procedure it was found that

approximately 70% of UV-induced dimers were removed from an actively transcribed dihydrofolate reductase (DHFR) gene in CHO cells whereas for regions outside the gene DNA repair levels comparable with overall CHO repair levels (i.e 10-15%) were found (Bohr et al. 1985; 1986). Similar results were obtained in experiments with the human DHFR gene (Mellon et al. 1986). Additional experiments showed that the efficiency of dimer removal correlates with the transcriptional activity of the examined DNA since a differential repair was found for active and inactive proto-oncogenes. Within 24 hours after UV-irradiation 85% of dimers were removed from the active mouse c-*abl* gene and only 20% from the silent mouse c-*mos* gene (Madhani et al. 1987). Furthermore, induction of the expression of the CHO metallothionein gene leads to a concomittant increase of the DNA repair of this particular gene (Okumoto and Bohr, 1987).

Accessibility of DNA damage during RNA synthesis might be the explanation for the close correlation between repair and transcription. However, in elegant experiments Mellon et al. (1987) recently reported that only the transcribed strand is subject to preferential repair. Hence, this indicates that accessibility of the DNA during transcription is not the sole factor determining repair efficiency. It is conceivable that there is a close interaction between transcription and repair processes and that the proteins involved are associated in one protein complex.

I.3.1 Selective removal of DNA damage in UV-sensitive cells

In normal cells the essential DHFR gene is found to be preferentially repaired (Mellon, 1986). In contrast no repair of this gene could be detected in XP-C cells (Bohr et al. 1986). Hence, it appears that the nuclear matrix associated repair activity in XP-C does not include all active genes (see section I.3). It was concluded from this observation that the UV hypersensitivity of XP-C cells correlated with the reduced efficiency of repair of essential genes (Bohr et al. 1986). In XP-D cells and normal CHO cells which both have a similar overall repair efficiency compared to XP-C cells, no evidence for nuclear matrix associated repair was found (Mullenders et al. 1986; 1987; Karentz and Cleaver, 1986). The aberrant mode of repair in XP-C suggests that the genetic defect may concern a regulatory mechanism which controls the accessibility of DNA adducts and hence the efficiency of DNA repair. A similar explanation might also be valid for XP-V cells. Only in that case it

can be speculated that the preferential repair of DNA damage is disturbed and that the random repair capacity of XP-V cells functions normally which explains the relatively high residual repair activity and at the same time high UV-induced mutation frequency.

Introduction of the bacteriophage T4 *denV* gene in UV-sensitive CHO mutants of complementation group 1 induces a high overall efficiency of DNA repair which does not seem to depend on transcriptional activity (Bohr and Hanawalt, 1987). This indicates that the accessibility of DNA damage to repair proteins is not restricted to the transcribed regions although it should be realized that the T4 endonuclease is a relatively small protein (19 kd).

Mayne et al. (1988a; pers. comm.) have recently reported that CS cells of complementation groups A and B are indeed defective in the repair of actively transcribed genes as postulated earlier (Mayne and Lehmann, 1982). Similar preliminary results were mentioned by Mellon et al. (1987). This is the first concrete indication that the genetic defect in CS concerns a DNA repair function.

I.4 UV-LIGHT, MUTAGENESIS AND CARCINOGENESIS

Skin cancer is the most common malignancy in man (Penn, 1985; Doll, 1986). It is evident that the major etiologic factor underlying the process of this type of carcinogenesis is exposure to sunlight, especially the UV component of the sunlight spectrum. This is clearly illustrated by the markedly increased frequency of UV-induced mutation frequency with XP cells compared to normal cells (Maher and McCormick, 1976). For a long time, the major biological effects of UV - lethality and mutagenesis - have been attributed to pyrimidine dimers. Recently, however, it has been found that approximately 10% of UV-induced DNA adducts are pyrimidine-pyrimidone (6-4)-photoproducts (Franklin et al. 1982; Haseltine, 1983). Examination of the chemical structure of the (6-4) lesions revealed that the 3' pyrimidone of the compound can no longer basepair with the purine residue of the opposite strand, leaving a premutagenic gap (Franklin et al. 1985). Indeed, several lines of research have indicated that the (6-4) photoproduct plays a dominant role in UV-induced mutagenesis in bacteria (Brash and Haseltine, 1982), phage lambda (Wood et al. 1984; Wood, 1985) and mammalian cells (Lebkowski et al. 1986; Drobetsky et al. 1985). However, using a shuttle vector system in human and monkey cells Kraemer and coworkers showed that *in vitro* photoreactiva-

tion of dimers before transfection of the UV-irradiated shuttle vector reduced the mutation frequency of a target gene in the vector to 10-20%, indicating a considerable mutagenic potential for the major UV lesion in mammalian cells (Protic-Sabljic et al. 1986; Bredberg et al. 1986; Brash et al. 1987).

Several mutation studies with mammalian cells have shown that UV primarily induces single point mutations which are often G.C-A.T transitions (DuBridge and Calos, 1987). The underlying mechanism is present in both normal and XP cells indicating that it is unlikely to be related to excision repair (Bredberg et al. 1986). G.C-A.T mutagenesis probably occurs during replication of the UV damaged strand. It has been found that prokaryotic polymerases preferentially insert adenine residues across from sites of DNA damage (Strauss et al. 1982; Schaaper et al. 1983). It is conceivable that this 'A-rule' model also accounts for the results of mutation studies with mammalian cells. In that case T<>T dimers would not be mutagenic but T<>C and C<>T dimers and (6-4) lesions (at TC sites) would result in an A residue insertion which will end up as a G.C-A.T transition. The controversy on the relative contribution of the two main UV lesions - pyrimidine dimers or (6-4) photoproducts - is not completely solved by this model and requires additional experiments.

Interestingly, it was recently found that a considerable overlap exists between spectra of spontaneous and UV-induced mutations (Drobetsky et al. 1987; De Jong et al. 1988). Moreover, it appeared that the frequency of photoproducts does not always correlate with mutation hot and cold spots (Brash et al. 1987). When these data are taken together one can speculate that DNA context at the site of the lesion plays a significant role in mutation fixation (Drobetsky et al. 1987).

At present it is not known whether the results from the above mentioned *in vitro* studies reflect the *in vivo* situation which accounts for sunlight induced skin tumor formation in XP patients. In this respect it might be important to realize that mutation induction by UV is wavelength dependent (Enninga et al. 1985) and that most *in vitro* studies were performed with UV of 254 nm which is not present in the solar spectrum that reaches the earth's surface and consequently of less biological significance (Parrish et al. 1978).

Recently, Keijzer et al. (submitted for publication) succeeded for the first time in establishing a cell line from a tumor of an XP-C patient and found that an A.T to C.G or T.A transversion in a short pyrimidine stretch, the target for UV- induced DNA lesions, was responsible for the oncogenic

activation of the N-*ras* gene. In addition, a similar mode of N-*ras* activation was found in two independent tumors of another XP patient (A.Sarasin, pers. comm.). *In vitro* studies on UV- induced mutations in the endogenous adenine phosphoribosyltransferase gene of CHO cells revealed that 15% of single base substitutions were A.T to C.G or T.A transversions (Drobetsky et al. 1987). Hence, it is not excluded that UV was the carcinogenic agent in the above mentioned XP cases.

It is conceivable that a specific set of proto-oncogenes is subject to UV-induced carcinogenesis. The involvement of a particular oncogene in cellular transformation will probably depend on the efficiency of its repair following UV-exposure. Since it was recently reported that CS cells appear to be deficient in repair of active chromatin (Mayne et al. 1988a) it is intriguing that in case of CS patients hypersensitivity to sunlight is not accompanied by cancer proneness (Kraemer, 1983). This suggests that the target genes for neoplasia formation are efficiently repaired and that the assumed defect in CS cells does not hold for the entire genome. In this respect it is worth while mentioning that experiments on DNA repair at the single gene level concern only a few genes (see section I.3). It is possible that additional heterogeneities in efficiency of repair of specific genomic regions exist which may explain why in some instances apparent repair defects are not reflected in enhanced carcinogenesis.

II CLONING OF HUMAN REPAIR GENES BY GENOMIC DNA TRANSFECTION.

II.1 INTRODUCTION

In prokaryotic and yeast systems many mutations can be corrected by the introduction of wild type donor DNA. In general, the molecular cloning of the transferred sequences can be readily achieved with standard molecular cloning techniques, provided that this DNA is linked to a selectable marker or some other kind of 'tag'. Important conditions for application of this powerful strategy to the mammalian system have been worked out by Szybalska and Szybalski (1962) who pioneered the transfer of genes to cultured animal cells and by Graham and van der Eb (1973) who considerably improved the efficiency of this process by introducing the calciumphosphate DNA co-precipitation technique. Transfer of mammalian genes is hampered by the large size and the high complexity of mammalian genomes compared to prokaryotes and yeast. This means that much more recipient cells have to be transfected to cover in total the equivalent of one genome. Moreover, the efficiency of transferring intact functional gene copies decreases with increasing gene size. Furthermore, molecular rescue of the transferred gene from the genome of a mammalian transformant requires more sophisticated cloning techniques compared to lower organisms. Notwithstanding these limitations, inherent to mammalian genomic DNA transfection, many laboratories have successfully adopted this approach in the last decade for the isolation of a variety of mammalian genes with different functions. It is evident that genomic DNA transfer to repair deficient cells is - at present - the most straightforward approach to clone human DNA repair genes. In the first part of this chapter general aspects of genomic DNA transfer to mammalian cells are discussed. The subsequent sections focus on results obtained in genomic transfections with human cells of DNA repair deficiency syndromes and *in vitro* mutagenized repair defective rodent cells. Further details on the methodology and strategy of genomic transfections have been described in several recent review papers (Friedberg et al. 1987b; McClelland et al. 1987; Hoeijmakers et al. 1988b).

TABLE II.1

Overview of mammalian genes that have been isolated following DNA mediated gene transfer with genomic DNA.

GENE	DONOR DNA	RECIPIENT CELL	COTRANSFECTED DNA	SELECTION	CLONING STRATEGY('TAG')	GENE SIZE	REFERENCE*
NGF-receptor	Human	Ltk$^-$	pTK	HAT, Rosette	Alu	15-30 kb	1
Transferrin Recept	Human	Ltk$^-$	pTK	HAT, FACS	Alu	31 kb	2,3
Thymidine kinase	Chicken	Ltk$^-$	pBR322	HAT	plasmid rescue	2 kb	4
	Human	Ltk$^-$	-	HAT	Alu	5-15 kb	5,6,7
HPRT	Human	LA9 hprt$^-$	-	HAT	Alu	±30 kb	8
APRT	Human	Ltk$^-$aprt$^-$	pBR322	azaser./aden.	plasmid rescue	<7 kb	9
T8 (Leu-2)	Human	Ltk$^-$	pTK	HAT, Rosette	subtr. cDNA cl.	<13 kb	10
Lyt-2	Mouse	Ltk$^-$	pTK	HAT, FACS	subtr. cDNA cl.	5 kb	11,12,13
T8 (leu-2)	Human	Ltk$^-$	pTK	HAT, FACS	subtr. cDNA cl.	?	14
T4	Human	Ltk$^-$	pTK	HAT, Rosette	subtr. cDNA cl.	30 kb	15
Cell Cycle: G1	Human	BHKts11	-	39.5^{o}C	Alu	?	16
RCC-1	Human	BHKtsBN2	-	39.5^{o}C	Alu	±30 kb	17
CCG-1	Human	BHKtsBN462	-	39.5^{o}C	Alu	25-30 kb	18
Cell Cycle: G1	Human	BHKtsBN51	-	39.5^{o}C	Alu	7-8 kb	19
XRCC-1	Human	CHO EM-9	pSV2gpt	MPA,chlorodUrd.	Alu	?	20,21,22
ERCC-1	Human	CHO 43-3B	pSV3gptH	MPA, UV, MMC	linked to gpt	15-17 kb	23
ERCC-2	Human	CHO UV5	pSV2gpt	MPA, UV	Alu	±20 kb	24
ERCC-3	Human	CHO 27-1	pSV3gptH	MPA, UV	Alu	35-40 kb	25
(Proto)-oncogenes	Human	NIH/3T3	pSVgpt pSVneo	MPA, oncogenic G418,transformation	Alu	5-40 kb	26,27

* 1:Chao et al. 1986; 2:Kuhn et al. 1984; 3:Newman et al. 1983; 4:Perucho et al. 1980; 5:Bradshaw et al. 1983; 6:Lin et al. 1983; 7:Lau et al. 1984; 8:Jolly et al. 1982; 9:Lowly et al. 1980; 10:Littman et al. 1985; 11:Kavathas and Herzenberg, 1983; 12:Kavathas et al. 1984; 13:Hsu et al. 1984; 14:Nakauchi et al. 1985; 15:Maddow et al. 1985; 16:Greco et al. 1987; 17:Kai et al. 1986; 18:Sekiguchi et al. 1987; 19:Ittmann et al. 1987; 20-21:Thompson et al.1985b; 1988; 22:Thompson, 1988; 23:Westerveld et al. 1984; 24:Weber et al. 1988; 25: Weeda, pers. comm.; 26: Cooper and Lane, 1984; 27:Bishop, 1987.

II.2 THE STRATEGY OF GENOMIC DNA TRANSFECTIONS

In Table II.1 a summary is presented of the mammalian genes isolated by genomic DNA transfections as extracted from the literature. The TK, HPRT and APRT genes have been cloned using mutant mouse L-cells whereas normal L-cells were employed to isolate genes encoding human cell surface proteins like CD4, CD8, nerve growth factor receptor and the transferrin receptor (see ref. 1-15 Table II.1). In the latter experiments antibodies against the gene product of interest have permitted the rapid identification of transformants either by the fluorescent activated cell sorter (FACS) or a rosette assay. In case of the T-cell surface proteins a subtractive cDNA cloning procedure was followed to recover the transfected genes from primary transformants. Syrian hamster BHK (baby hamster kidney) ts-mutants appeared to be a suitable tool for the cloning of several cell cycle control genes (see refs. 16-19,Table II.1) and the isolation of human DNA repair genes has been accomplished using UV-sensitive CHO mutants (see refs. 20-25, Table II.1 and also section II.4). In the field of cancer research the pre-neoplastic NIH/3T3 mouse fibroblast has become the standard recipient cell for the transfer of transforming sequences from tumor DNA of many sources (reviewed in Cooper and Lane, 1984). This has facilitated the isolation of a diverse assortment of activated oncogenes. The first to be identified were alleles of the c-*ras* oncogenes which had acquired a transforming potential by single point mutations (Varmus, 1984). An increasing group of oncogenes that are picked up by the NIH/3T3 assay harbors large abnormalities, primarily rearrangements which in a number of cases are also found back in the original tumor DNA but in some instances seem to have arisen during the transfection process (Bishop, 1987). Although differences exist between the strategies followed to clone the genes listed in Table II.1, the basic outline of the approach encompasses a number of common steps:

1. Transfection of high molecular weight genomic DNA to the recipient cells by the calciumphosphate method or alternative techniques.

2. Selection of cells with the desired genetically transformed phenotype.

3. Segregation of irrelevant cotransfected DNA by a second round of transfection using DNA of a primary transformant.

4. Application of recombinant DNA techniques to recover the transfected gene from the DNA of a secondary transformant.

From the available data it can be inferred that the efficiency of genomic DNA transfer depends on a variety of factors which are discussed below.

II.2.1 Recipient cells

Obviously, naturally occurring mutants or *in vitro* mutagenized cells are likely candidates to be used in genomic transfections. However, it is apparent from the cloning of several genes encoding cell surface proteins that also normal cells can be utilized provided that they do not express by themselves the phenotype selected for. Recipient cell lines should be stable with respect to the property selected for or otherwise have a reversion frequency which is considerably lower than the transfection efficiency of the gene of interest. Uptil now only a limited number of recipient cell strains, all of rodent origin have been used successfully in genomic transfections (see Table II.1). As far as known uptil now no reports have claimed the cloning of genes after genomic DNA transfer to human cells. Especially cell lines from patients with DNA repair syndromes have been utilized extensively in many genomic DNA transfections in different laboratories. However, with one notable exception (Tanaka et al. 1988, see also section II.3), unequivocal evidence for the generation of 'bona fide' transformants retaining one specific correcting gene is lacking. One explanation for the inefficient gene transfer experienced with human cell lines is that compared to the rodent cells mentioned above, in general only very small quantities of exogenous sequences are integrated in human cells. (Hoeijmakers et al. 1987; Mayne et al. 1988b). In addition, factors such as transfection frequency and degree of scrambling of exogenous DNA molecules also determine the suitability of a given cell line for genomic transfections. It is of interest to know to what extent these parameters are genetically determined. Recent experiments of Shiomi et al. (1988) suggest that the property of 'transfection frequency' is dominant and can be conferred to a specific cell by cell fusion with an X-irradiated good recipient cell line. It is of considerable importance to investigate this further because it could turn any interesting mutant into one which is also valuable for gene cloning.

II.2.2 The transfected donor DNA

Currently, in most transfection protocols both in the first and second round of transfection genomic DNA is cotransfected with a dominant marker gene

Ecogpt or *agpt* to provide a preselection for the small proportion of cells that are competent to integrate DNA. In order to 'follow' the transfected sequences DNA it is preferable to use genomic DNA of a heterologous species. In case of human DNA the abundant presence of dispersed 'Alu' repeats (Schmid and Jelinek, 1982) provides a natural 'tag' which can be used to detect the presence of transfected DNA in primary and secondary transformants of rodent origin and most importantly it facilitates the recovery of transfected sequences from a recombinant DNA library of secondary transformant DNA. The cloning of transfected genes is seriously hampered when after a second round of transfection no physical linkage of the gene of interest with either repetitive DNA or any other kind of 'tag' can be detected. Absence of detectable transfected sequences on Southern blots of secondary transformants probed with species specific repeats leaves two possibilities open. First, the transfected gene is very poor in repetitive elements and hence difficult to visualize on Southern blots. This possibility was put forward by Albino et al. (1985) and Dulhanty and Whitmore (1988) to explain their transfection results. Second, the secondary 'transformant' does not contain any transfected DNA and is a revertant. A way to get around this problem is to try to establish physical linkage between the gene of interest and a dominant marker copy. In the secondary transfection this requires first a selection on the dominant marker gene copies present in the genomic DNA of a primary transformant. If subsequent co-selection on the gene of interest yields secondary transformants there is compelling evidence for a physical linkage between the dominant marker and the transfected gene. The cloned marker copy can then serve as a 'tag' for the isolation of the transferred gene that must be located in close viscinity of this dominant marker gene (Westerveld et al. 1984, Appendix paper I). Because exogenous DNA is subject to scrambling in the cell during the transfection process, gene size is an important factor determining efficiency of gene transfer in genomic transfections. On the average one double stranded break is induced every 5-15 kb in transfection experiments with various cell lines (Wake et al. 1984; Hoeijmakers et al. 1987). However, in this respect also considerable differences between individual cell strains have been noted (Hoeijmakers et al. 1987). The largest gene isolated thus far by DNA mediated gene transfer is Ki-*ras* (38 kb, McGrath et al. 1983). However, the recently cloned *ERCC-6* gene appears to reside on a fragment of 105 kb (C. Troelstra, pers. comm.) and could well shift the upper limit for a transfectable gene size to the 100 kb range.

A gene transfer approach which is only feasible for genes smaller than approximately 40 kb involves transfection of total DNA from a genomic cosmid library. The presence of a selectable marker on the cosmid vector provides a physical linkage with the desired gene which will facilitate its molecular cloning from transformant DNA. Using cosmid transfections Pinney et al. (1988) have recently transferred a mouse gene (*myd*) involved in myogenic lineage determination and differentiation into mouse C3H 10T1/2 cells. In two rounds of transfection co-inheritance with the dominant marker was shown, which will aid the isolation of the *myd* gene. Interestingly, it was found that transformants could only be obtained with cosmid DNA prepared in methylase deficient bacteria. Apparently, hypomethylation is an important trigger for activation of the *myd* gene, which was confirmed in additional experiments (Pinney et al. 1988). Furthermore, it can be inferred from this work that genomic transfections with high molecular weight DNA may be influenced by the methylation state of the transfected donor DNA.

Chromosome mediated gene transfer or interspecific cell fusions are alternative means to introduce large pieces of DNA into recipient cells (Cirullo et al. 1983; Pritchard and Goodfellow, 1986; Fallows et al. 1987). However, these procedures have the disadvantage that for gene cloning purposes too much DNA is transferred, as a result of which the use of repetitive DNA as a probe to isolate the sequences of interest becomes virtually impossible.

Transfection of DNA from cDNA expression libraries seems at least in theory a feasible approach to circumvent the relatively poor ability of human cells to integrate large pieces of foreign DNA. Moreover, it should also permit the transfer of cDNA from genes that are too large to be transferred in a genomic transfection. Okayama and Berg (1983; 1985) have designed an efficient cDNA cloning procedure in which the cDNA is inserted in an oriented manner in between the strong SV40 large T promoter and a SV40 derived poly(A) signal. Transfection of a total cDNA library to mouse hypoxanthine phosphoribosyltransferase (HPRT) deficient cells yielded wild type transformants which was shown to be the result of transfected full lenght HPRT cDNA clones (Jolly et al. 1983).

Factors critically determining the rate of success with this cDNA transfection approach are the level of the transcription and the size of the mRNA (which indirectly influence the fraction of functional cDNA copies in the library). Another unknown parameter concerns the possibility that the gene of interest requires a strictly regulated expression, which is obviously not

provided by the regulatory elements on the cDNA expression vector. Finally, when more than one cDNA is necessary, because of alternative splicing of the gene to be cloned, the cDNA approach will become impossible.

II.2.3 Integrity of transformants

In general in genomic transfection experiments a relatively small number of transformants are obtained, which does not allow statistics to be used to discriminate between the frequency of real transformants and of revertants. Moreover, reversion frequencies are often determined in independent experiments, that are usually not of the same scale or carried out under different conditions as used for the transfection experiments. Southern blot analysis of transformant DNA using repetitive probes of the transfected DNA is an efficient means to establish the integrity of transformants. For instance detection of transfected DNA in a secondary transformant is very important in this respect . On the other hand a similar finding with a primary transformant is not conclusive since it is conceivable that many cells have taken up transfected DNA in a primary round of transfection irrespective of the cause of their phenotypical change. The integrity of transformants is considerably substantiated when in Southern blot experiments similar sized fragments of transfected DNA can be detected in independent secondary transformants.

In a great number of reports the successful transfer of 'genomic' genes has been claimed without providing convincing Southern blot data to verify the integrity of transformants (Montisano and Hankinson, 1985; El-Gewely and Oxender, 1985; Albino et al. 1985; Kojima et al. 1987; Ding et al. 1985; Yarosh et al. 1986; Kaina et al. 1987; Johnson et al. 1988). It is not excluded that in a number of cases transfection-induced phenotypical changes have occurred which might be due to the phenotype of the recipient cells used. For instance the *mer*$^-$ phenotype (i.e. sensitivity to alkylating agents due to absence of measurable activity of O^6-methylguanine alkyltransferase) of mammalian cells seems highly unstable (Day et al. 1987) which might have been of influence in transfection studies aimed at the isolation of the human O^6-methylguanine alkyltransferase (Ding et al. 1985; Yarosh et al. 1986; Kaina et al. 1987).

II.3 GENOMIC TRANSFECTIONS TO HUMAN DNA REPAIR DEFICIENT CELLS

A large number of permanent human cell lines from DNA repair deficiency syndromes are available (reviewed by Collins and Johnson, 1987) that can be used for genomic transfections. The hypersensitivity of these cells to DNA damaging treatments provides a relatively easy selection criterium for the isolation of repair proficient transformants. Many efforts to correct the excision repair defect in XP cells have been unsuccessful (reviewed by Lehmann, 1985). However, recently in an impressive series of transfection experiments with high molecular weight mouse DNA Tanaka et al. (1988) managed to get two UV-resistant primary XP-A transformants. In a subsequent round of transfection with genomic DNA of one 'primary' a single secondary transformant was obtained that had retained mouse sequences. Preliminary data have shown that two isolated overlapping lambda clones with mouse DNA from the secondary transformant were able to confer a repair proficient phenotype to the XP2OS-SV40 cells with a high efficiency (Tanaka et al. 1988). These data indicate that fragments of the mouse homologue of the XP-A gene have been isolated. Since apparently the encoded murine protein is functional in human cells the nucleotide sequence conservation between the murine and human XP-A gene probably will be sufficient to use mouse probes for cloning the human gene. In cell fusion studies the mouse chromosome 4 was found to correct the XP-A defect (Lin and Ruddle, 1981). It will be of interest to see whether the sequences isolated by Tanaka et al. (1988) also map to this murine chromosome.

The elegant work of Tanaka and coworkers promises to be the first example of gene cloning following DNA mediated gene transfer to human cells. A further conclusion from the work of the Japanese group may be that the poor transfectability of human cells as reported by Hoeijmakers et al. (1987) and Mayne et al. (1988b) may be overcome at least for the XP-A gene by increasing the number of transfected cells.

As a first step towards cloning of the XP-A gene Karentz et al. (1987) have made repair proficient XP-A/CHO hybrids containing hamster chromosome fragments. In these experiments rescue of 'Alu'-type Chinese hamster sequences is envisaged as a next step. However, it is likely that this approach will be seriously hampered by the presence of too much irrelevant hamster DNA.

Partial complementation of the XP-A and XP-C phenotype was achieved

by transfection of DNA from total human cDNA (Matsumoto and Fujiwara, 1987; Teitz et al. 1987 respectively). In case of the XP-C experiments secondary transformants could be generated which should facilitate the cloning of the XP-C gene by plasmid rescue (Teitz et al. 1987).

Mock transfections with XP group A cells have yielded UV-resistant clones with a frequency of 10^{-7} (Schultz et al. 1985; Royer-Pokora and Haseltine, 1984). In most cases these apparent revertants were obtained after multiple UV irradiations suggesting that the reversion of the XP mutation was the result of UV-induced mutagenesis. The fact that the XP-A mutation is revertible to normal, indicates that the genetic defect in XP-A is most likely a point mutation. Recently, Cleaver et al. (1987) have produced a number of XP-A revertants by treatment of XP12RO-SV40 cells with ethylmethylsulfonate and an accumulated UV dose of 60 J/m^2. It appeared that in this revertant (6-4) photolesions were repaired normally in contrast to pyrimidine dimers. This may suggest that the XP-A protein has a variable affinity for different kinds of DNA adducts.

Attempts to correct the DNA-repair deficient phenotype of other human cell lines from DNA repair deficiency syndromes have been unsuccessful so far or are still in progress. Lehmann and coworkers have performed an extensive series of transfection experiments with an SV40 transformed AT line. Approximately 400.000 gpt^+ transformants were screened for resistance against gamma irradiation. However, no transformants were obtained that could be used to recover the AT gene (Green et al. 1987). FA cells are characterized by an elevated sensitivity to crosslinking agents which can be used as a selective tool in genomic transfections. Recently, Diatloff-Zito et al. (1986) and Shaham et al. (1987) have claimed that a normal response to MMC or diepoxybutane can be conferred to primary FA-fibroblasts after transfer of genomic DNA whereas no 'transformants' were obtained in mock experiments or transfections with FA DNA. These data are remarkable as it has been found by others that human cells and especially primary fibroblasts are less favourable for DNA mediated gene transfer experiments (Schultz et al. 1985; Buchwald et al. 1987; Hoeijmakers et al. 1987; Mayne et al. 1988b). It was found recently that the FA mutation can revert with a frequency of about $2x10^{-7}$ (Buchwald et al. 1987). However, this can not provide a plausible explanation for the transfection results with primary FA fibroblasts since these experiments involved relatively small number of cells and yielded 'transformants' with a much higher frequency. Based on current knowledge of

of DNA mediated gene transfer, particularly into primary fibroblasts it seems unlikely that genomic transfections using these cells will aid disclosure of the genetic defect in FA.

II.4 DNA REPAIR DEFICIENT RODENT CELLS

In recent years a large number of rodent mutant cells with increased sensitivity to DNA damaging agents have been isolated by *in vitro* mutagenesis (reviewed by Thompson, 1985; Collins and Johnson, 1987; Hickson and Harris, 1988). In most cases UV and ionizing radiation have been used as the selective agents after treatment with mutagens like ethyl methanesulfonate and ethylnitrosourea. Many of the laboratory induced mutant cells are phenotypically resembling the cellular hypersensitivity to DNA damage manifested in human DNA repair syndromes. However, uptil now there are no direct indications that these naturally occurring mutations are represented in *in vitro* induced rodent mutant cells. A defective incision step of the excision repair process underlies the UV-sensitivity of a number of CHO mutants which is very similar to XP (Thompson et al. 1982a). Hence, it is conceivable that one or more of the ten XP mutations exists among the UV-sensitive rodent cells. Several MMC-sensitive CHO mutants resemble in some respects the FA phenotype and recently, X-ray sensitive CHO cells have been isolated that show a lack of inhibition of DNA synthesis after X-ray exposure, which is a consistent feature of AT cells (Zdzienicka et al. submitted for publication). Therefore, in general rodent repair mutants are not only valuable tools for studying mammalian DNA repair, but may also facilitate identification of the hereditary repair defects in human cells.

A large number of mutants exhibits a cross-sensitivity to a variety of DNA damaging agents (Hickson and Harris, 1988). This suggests that common steps might be involved in the repair of single strand breaks, monofunctional adducts and DNA cross-links. To shed more light on this apparent complexity it is important that complementation tests are performed to establish the actual number of different mutated genes in the collection of mutants isolated in various laboratories.

The presently isolated UV-sensitive mutants of Chinese hamster origin have been classified in 7 complementation groups (Thompson et al. 1981; Thompson and Carrano, 1983; Thompson et al. 1987a; Zdzienicka et al. 1988). Recent data on the characterization of those complementation groups

and the human genes that correct the defect in corresponding rodent cells are summarized in Table II.2. Four complementation groups have been described for UV-sensitive mouse lymphoma lines (Shiomi et al. 1982) and recently evidence was reported for overlap between the mouse and Chinese hamster classifications. In fusion experiments mouse group I cells failed to complement CHO group 5 (Thompson et al. 1987b). Moreover, the repair defect in both groups could be corrected by a gene on human chromosome 13 (Hori et al. 1983; Thompson et al. 1987b). The US31 mutant which is the sole representative of mouse group III does not complement the first 6 Chinese hamster groups (Thompson et al.1988b). Fusion experiments with the recently described seventh hamster group (Zdzienicka et al. 1988) will have to reveal whether mouse group III represents the eight rodent UV- sensitive complementation group.

With respect to the potential use of these UV-sensitive cell lines in genomic transfections aimed at isolating complementing human genes it is important to note that considerable differences in transfectability exist. The mouse lymphoma mutants have a very low transfection frequency (Shiomi et al. 1988) and are very unsuitable for this approach. V79 Chinese hamster cells were much less efficient DNA 'uptakers' than CHO- 9 and AA8 cells which also makes these cells less favourable for genomic DNA transfer (Hoeijmakers et al. 1987).

II.4.1 Chromosomal localization of *ERCC*-genes

By cell fusions between UV-sensitive rodent mutants and wild type human cells proliferating repair proficient hybrids could be generated. Karyotyping of these hybrids allowed the chromosomal assignment of the complementing human genes that have been designated as *ERCC* genes (*E*xcision *R*epair *C*ross *C*omplementing rodent repair deficiency). In this way the human *ERCC-1* to *ERCC-5* genes, which complement the first 5 CHO complementation groups, have been mapped on human chromosomes 19, 19, 2, 16 and 13 respectively (see Table II.2) (Thompson et al. 1985b; Siciliano et al. 1986a; Thompson et al. 1987b). For *ERCC-1* and *ERCC-3* these data have been confirmed by Southern blot analysis of panels of human/rodent hybrids with probes of the cloned genes (Rubin et al. 1985; Van Duin et al. 1986, Appendix paper I; G. Weeda, pers. comm.). Surprisingly, the human gene *XRCC-1* which corrects the double strand break repair defect

TABLE II.2

Complementation groups of UV-sensitive CHO mutants and the chromosomal localization and other properties of the human *ERCC*-genes that correct the CHO repair defect as found by cell fusion experiments and genomic DNA mediated gene transfer.

Compl. group	Mutant(s)	Parental line	UV^s	MMC^s (>5x)	Ref.[b]	Gene name	Human Chrom.	Ref.[c]	Cloned (y/n)	Gene (kb)	mRNA (kb)	Ref.[d]
1*	43-3B;UV20	CHO9;AA8	+	+	1,2	*ERCC-1***	19	6	y	15-17	1.1	9
2	UV5	AA8	+	-	2	*ERCC-2*	19	7	y	20	2.5-3	10
3	27-1;UV24	CHO9;AA8	+	-	1	*ERCC-3*	2	8	y	35-40	3-3.5	11
4	UV41	AA8	+	+	2	*ERCC-4*	16	7	n	?	?	
5	UV135;Q31	AA8;L5178Y	+	-	2,3	*ERCC-5*	13	8	n	?	?	
6	UV61	AA8	+	?	4	*ERCC-6*	?		n[a]	?	?	12
7	VB11	V79	+	?	5	*ERCC-7*	?		n	?	?	

* At a recent UCLA meeting (Taos, USA, January 1988) it was decided to rename complementation group1 and 2 respectively in order to match the numbering with the complementing human genes.

** Gene nomenclature according to 7th Human Gene Mapping conference (Shows et al. 1984). *ERCC*: *E*xcision *R*epair *C*omplementing defective repair in *C*hinese hamster cells. In order to include mouse mutants the meaning of *ERCC* was recently converted to *E*xcision *R*epair *C*ross *C*omplementing rodent repair deficiency.

a) gene cloning is in advanced state.

b) References in which mutants are described: 1:Wood and Burki, 1982; 2:Thompson et al. 1981; 3:Shiomi et al. 1982; 4:Thompson et al. 1987a; 5:Zdzienicka et al. 1988;

c) References on chromosomal assignment: 6:Thompson et al. 1985b; 7:Siciliano et al. 1986a; 8:Thompson et al. 1987b.

d) References of gene cloning: 9:Westerveld et al. 1984, Appendix paper I; 10: Weber et al. 1988; 11: Weeda, pers. comm.; 12: Troelstra, pers.comm.

in CHO mutant EM9 was also localized on chromosome 19 (Siciliano et al. 1986b), which means that 3 out of 6 CHO mutants are corrected by a gene on the same human chromosome. Several other markers for this chromosome map on Chinese hamster chromosome 9 which is hemizygous in CHO cells (Siciliano et al. 1983). Therefore, the monosomic state of hamster chromosome 9 might explain the apparent nonrandom chromosomal distribution of mutations that account for the different CHO complementation groups.

II.4.2 Isolation of *ERCC*-genes by genomic DNA transfer

In several laboratories genomic DNA transfection experiments have been used or are in progress to isolate human *ERCC*-genes complementing the CHO repair defects. In general UV is used as the selective agent for the isolation of primary transformants. However, mutants belonging to group 1 and 4 display a marked cross-sensitivity to MMC (Hoy et al. 1985) which also could be used as a selective drug.

The human *ERCC-1* gene which complements CHO group 1 is the first human DNA repair gene that was isolated by genomic DNA transfer (Westerveld et al. 1984, Appendix paper I). In a primary transfection total human DNA was partially digested with PstI and after ligation to the selectable marker *Ecogpt* transfected to 43-3B cells. Repair proficient primary transformants were obtained by selection in medium containing MPA and MMC. In a secondary transfection a linked transfer of MPA and MMC resistance could be achieved. This allowed screening of a cosmid library of a secondary transformant, with the *Ecogpt* as a probe, which yielded the isolation of a functional *ERCC-1* gene close to one of the *Ecogpt* copies (Westerveld et al. 1984, Appendix paper I). Using the CHO mutant UV20 of group 1 Rubin et al. (1983) also managed to generate secondary transformants. However, the subsequently followed cloning strategy only resulted in the isolation of fragments of the *ERCC-1* gene (Rubin et al. 1985; Van Duin et al. 1986, Appendix paper II). The UV5 mutant, belonging to group 2, has been successfully employed for the cloning of the *ERCC-2* gene (Weber et al. 1988) and recently also the isolation of human *ERCC-3* was achieved after gene transfer to the 27-1 mutant (Hoeijmakers et al. 1988a). Molecular analysis revealed that tertiairy UV135 (group 5) transformants had retained human sequences (MacInnes et al. 1984; Strniste et al. 1988) and independent primary and secondary UV61 (group 6) transformants were found to harbor

identical, transfected, unique human sequences (C.Troelstra, pers. comm.) indicating that the cloning of the *ERCC-5* and *ERCC-6* genes is in an advanced state. Except for the strategy used for *ERCC-1* in all other cases total human DNA was transfected and human Alu probes were used for analysis of transformants and isolation of transfected sequences from recombinant DNA libraries. The experiments of Rubin et al. (1985) show that this procedure would also have been applicable to the *ERCC-1* gene. In contrast, it seems that efforts to isolate *ERCC-4* are hampered due to the apparent low abundance of human repetitive sequences in the *ERCC-4* gene as a result of which transfected human sequences can not be detected in DNA of secondary transformants (Rubin et al. 1983; Dulhanty and Whitmore, 1988). For the isolation of *ERCC-4* it looks worth-while to apply the protocol followed for *ERCC-1* cloning. Transfection of the isolated human *ERCC-1* and *ERCC-2* genes to representative cell lines of all 6 CHO complementation groups only yielded repair correction in group 1 and 2 mutants respectively (Van Duin et al. 1988a, Appendix paper V; Weber et al. 1988). Furthermore, it was found that introduction of *ERCC- 1* in 43-3B cell gives close to wild type levels of all repair parameters examined (cross-sensitivity to other agents, UV- induced mutagenesis, DNA incision as measured by UV-endo assay and preferential repair) (Zdzienicka et al. 1987; Bohr et al. 1988). Comparable data were obtained for UV5 cells with the cloned *ERCC-2* gene (Weber et al. 1988). Hence it is reasonable to assume that *ERCC-1* and *ERCC-2* are the human homologues of the mutated Chinese hamster repair genes.

With respect to the characterization of the cloned *ERCC*-genes it is of prime interest to investigate whether there is a relationship with the human DNA repair syndromes. Southern blot analysis has revealed that the *ERCC-1* gene is not deleted or grossly rearranged in all nine XP (A to I) and the CS and FA complementation groups A and B. Furthermore, Northern blot experiments revealed correct transcription of *ERCC-1* in all nine XP groups examined (Van Duin et al. 1988c, Appendix paper VI) Introduction of *ERCC-1* into representative cell lines of all XP and CS complementation groups groups did not confer DNA repair proficiency or UV resistance. Taken together these data strongly suggest that *ERCC-1* is most likely not the mutated gene in XP and CS cells examined (Van Duin et al. 1988c, Appendix paper VI). The *ERCC-2* and *ERCC-3* gene have recently been cloned. The uptil now limited data on the molecular characterization of both genes are summarized in Table II.2.

II.4.2.1 The *ERCC-1* gene

A functional human *ERCC-1* cDNA which corrects the CHO group 1 mutation has been isolated and using human probes also a biologically active mouse cDNA was obtained (Van Duin et al. 1986; 1988b, Appendix papers II and IV). The mammalian *ERCC-1* protein has a predicted molecular weight of 32-33 kd and a putative DNA binding domain and nuclear location signal have been tentatively assigned to the deduced protein based on comparison with consensus amino acid sequences of functional domains in other proteins (Hoeijmakers et al. 1986). Currently, experiments are in progress to lay hand on the *ERCC-1* protein which permits *in vitro* analysis of *ERCC-1* functions.

Interestingly, it appeared that the mammalian *ERCC-1* protein exhibits significant homology to the yeast excision repair protein *RAD10* and parts of the E.coli excision repair proteins *uvrA* and *uvrC* (Doolittle et al. 1986; Hoeijmakers et al. 1986; Van Duin et al. 1986; 1988b, Appendix papers II and IV). The similarity between the different repair proteins is schematically depicted in Figure II.1. The homology with the bacterial proteins concerns relatively short regions which may be indicative of a similar domainal function. However, the similarity with *RAD10* includes half of the yeast protein and one third of the mammalian repair protein. Therefore, it seems likely that *ERCC-1* and *RAD10* provide a similar function in DNA excision repair. The postulated *ERCC-1* DNA binding domain coincides with the most homologous region indicating that this protein part may represent the functional centre of *ERCC- 1/RAD10.* It seems likely to assume that *ERCC-1* and *RAD10* have evolved from a similar ancestral gene. Recently, Friedberg and coworkers have tailored the *RAD10* gene behind a mammalian promoter and transfected it to UV20 cells belonging to CHO group 1 (Burtscher et al. 1988a). A small but consistent correction of the UV sensitivity by the yeast gene was reported which represents the first experimental evidence that *ERCC-1* and *RAD10* have functions in common.

The complete genomic organization of the human *ERCC-1* gene and part of the mouse gene have been established (Van Duin et al. 1987; 1988b, Appendix papers III and IV). The architecture of the human *ERCC-1* gene region is shown in Figure II.2. The human *ERCC-1* gene has a length of 15-17 kb and consists of 10 exons of which the first exon is non-coding (open box), the eigth exon (hatched) subject to alternative splicing and the last exon (coding part: black, non-coding: open) to differential polyadenylation

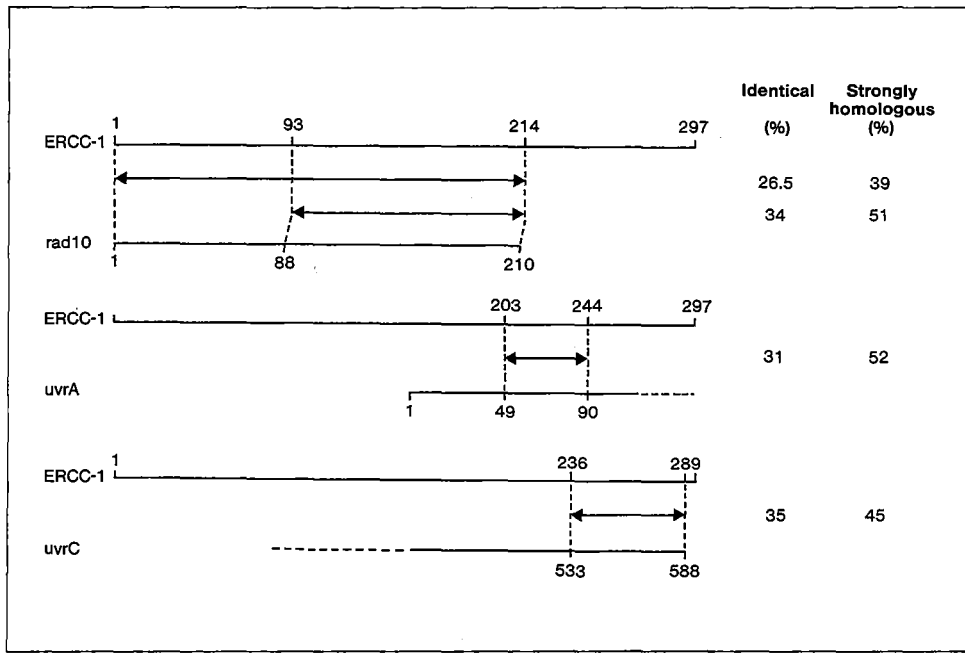

Figure II.1

Schematical representation of the amino acid (AA) homology between *ERCC-1* and *RAD10*, *uvrA* and *uvrC*. Left column, perfect AA homology. Right column, conserved AA homology.

which yields transcripts of 1.0, 1.1, 3.4 and possibly 3.8 kb.
The significance of both the alternative splicing and 3'end processing is unknown. It was found that the larger mRNAs are primarily located in the nucleus (Van Duin et al. submitted, Appendix paper VII) which might suggest that they represent an intermediate RNA molecule during RNA processing. Northern blot experiments have revealed a relatively low *ERCC-1* expression level in a variety of human cells and different mouse organs and stages of development. UV exposure or MMC treatment of HeLa cells did not induce *ERCC-1* transcription (van Duin et al. 1987; Appendix paper III). The *ERCC-1* promoter was found to be devoid of known promoter elements and therefore might represent a class of eukaryotic promoters that are responsible for a low constitutive rate of transcription.

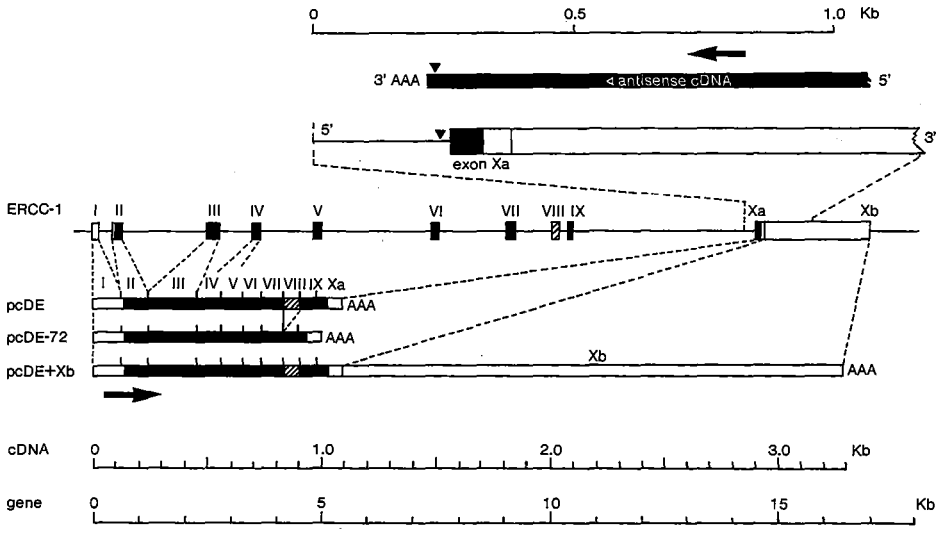

Figure II.2

Architecture of the human *ERCC-1* gene (see text for explanation).

Surprisingly, an opposite transcription unit was found to be located in the 3' *ERCC-1* region of mouse and man (see also Figure II.2). Characterization of a partial cDNA of the human antisense gene (designated *ASE-1*, antisense *ERCC-1*) and genomic sequencing indicated that this opposite oriented transcript of 2.6 kb is overlapping with *ERCC-1* exon X and terminating in the ninth intron (Van Duin et al. submitted, Appendix paper VII). This is the first example of naturally occurring gene overlap in the human genome and might be a fortuitous observation. Strikingly, however, it was found independently by Prakash and coworkers that the yeast *RAD10* locus is also harboring an antisense gene, (designated *ASR10*, antisense *RAD10*), encoding a protein of 525 amino acids (Van Duin et al. submitted, Appendix paper VII). As far as known only two other examples of overlapping transcripts in yeast have been reported in the literature (Hahn et al. 1988; Barker et al. 1985). Hence, it seems likely that evolutionary conservation accounts for the presence of opposite transcripts in the *ERCC-1* and *RAD10* loci which is highly exceptional both in man and yeast. Full length cloning of the human antisense cDNA and sequence analysis will have to reveal whether there is any homology between the yeast and human antisense genes. From the

presently available data it is clear that the antisense gene product is not required for correction of the CHO mutation by *ERCC-1* since *ERCC-1* cDNA constructs were fully competent in this respect. The yeast genetic system allows experimental manipulation of the antisense *RAD10* gene in order to identify its biological significance.

III CONCLUDING REMARKS AND PROSPECTS.

Complementation studies with human and rodent UV-sensitive cell lines have indicated that the mammalian excision repair process is very complex and involves many gene products. This conclusion is based on the large number of XP complementation groups and the still growing collection of laboratory induced mutants of Chinese hamster and mouse origin (Hickson and Harris, 1988). The classifications of XP and UV-sensitive rodent mutants include together more than 18 different complementation groups. It is presently unknown whether there is any overlap between both groups of UV-sensitive cell lines. Phenotypical comparison of XP and rodent mutants renders it possible that similar genes are mutated. However, the cloning of the mutated genes is awaited in order to perform conclusive experiments in this respect. Considering the large number of mutants that have been identified it seems reasonable to assume that some *in vitro* induced rodent mutants carry 'repair syndrome' mutations and consequently may facilitate unraveling the genetic defect in disorders like XP. On the other hand, it can not be excluded that the induced mutations in rodent cells are incompatible with life and therefore not represented in the human repair syndromes.

Correction of the genetic defect in UV-sensitive cells by DNA mediated gene transfer is the most direct route to achieve the cloning of the complementing gene. As it turns out, the isolation of UV-sensitive CHO mutants has been of great importance in this regard. These rodent mutants have far better transfection properties compared to their human XP counterparts (Hoeijmakers et al. 1987). However, as has been reported by Tanaka et al. (1988), genomic transfections with XP cells are not impossible. Based on the data reported by the Japanese group it can be expected that detailed information on the XP-A gene will be available soon.

In several laboratories genomic DNA transfections into UV- sensitive CHO mutants are in progress with the aim to isolate human *ERCC*-genes. At present this has resulted in the cloning of the *ERCC-1* gene (Westerveld et al. 1984; Appendix paper I) while recently the human *ERCC-2* and *ERCC-3* genes have been isolated (Weber et al. 1988; Hoeijmakers et al. 1988a). Genomic transfections with representative cell lines of the other CHO groups are in an advanced state. Therefore, it seems likely that the cloning of the remaining *ERCC*-genes will be achieved in the next few years.

It is evident that further isolation of rodent repair mutants deserves a high

priority. Considering the differences that exist among different cell lines in terms of DNA uptake in gene transfer experiments (Hoeijmakers et al. 1987) it is preferable to isolate new mutants from suitable parental cells. The pattern of hemizygozity of such lines largely determines the spectrum of mutants that can be generated (see II.4.1). In order to avoid repeated isolation of the same mutants it is advisable to perform mutagenesis with rodent cells other than CHO which presumably harbor other patterns of hemizygozity. Using the Chinese hamster lung fibroblast line V79 Zdzienicka et al. (1988) have recently isolated the seventh UV-sensitive rodent complementation group. At present only for the *ERCC-1* protein the amino acid sequence has been deduced from the sequence of the cloned cDNA (Van Duin et al. 1986, Appendix paper II). One of the most remarkable results of the experimental work on *ERCC-1* is its significant homology with the yeast *RAD10* and *E.coli uvrA* and *uvrC* proteins (Doolittle et al. 1986; Hoeijmakers et al. 1986; Van Duin et al. 1988b, Appendix paper IV). The interesting finding that the *ERCC-1* amino acid sequence has significant homology with yeast and bacterial excision repair proteins suggests that functional aspects of DNA excision repair are well conserved during evolution. Assuming that this will also hold for DNA repair mechanisms as a whole it is of interest to investigate whether the other yeast repair genes also have homologous human counterparts. In this respect recently promising results have been obtained for the *RAD6* gene. In hybridization experiments with *RAD6* probes sequence conservation up to mammals was reported (Jentsch et al. 1987; Ozkaynak et al. 1987). Therefore it seems an attractive approach to recrute the other yeast repair genes and use them as probes to clone homologous genes in higher organisms. If no direct homology with mammalian sequences can be detected a step by step approach with intermediate species like *Drosophila* might be followed to gradually ascent the evolutionary ladder.

It was recently reported by Lee and Nurse (1987) that mutations of the fission yeast *Schizosaccharomyces pombe* (*S.pombe*) could be corrected by transfer of a human cDNA library. This resulted in the isolation of the human homologue of the fission yeast cell cycle control gene *CDC2*. A large number of repair deficient *rad* mutants of *S.Pombe* have been described (Phipps et al. 1985). Gene transfer of total human cDNA expression libraries to *S.pombe* might be a straight forward alternative approach to identify and clone human DNA repair genes.

The isolation of the *ERCC*-genes means that detailed information on the

primary structure of human excision repair proteins becomes available. As has been shown for *ERCC-1*, putative functions can be ascribed to the amino acid sequence. However, it is of course essential to verify these presumptions and obtain additional data from experimental work on the purified repair proteins. Uptil now the *ERCC-1* protein has been very reluctant to overexpression in *E.coli* and mammalian cells (unpublished results) which might be inherent to the function of *ERCC-1*. Efforts to overproduce the *ERCC-1* protein in yeast are currently in progress in collaboration with Dr. S. Prakash (Rochester, USA). In this regard it is hoped that the biochemical characterization of *ERCC-1* might benefit from the data already available from *RAD10* protein work and perhaps to a lesser extent the results obtained on the characterization of the *uvrA* and *uvrC* proteins.

At this moment it is completely unknown whether the different *ERCC* polypeptides act in a combined fashion in excision repair. Once purified *ERCC*-proteins are available it will be of great interest to investigate whether *in vitro* excision repair activity can be established by mixing different *ERCC* polypeptides in a fashion comparable to *E.coli uvrABC* repair. In this respect it is important that recently considerable progress has been made in developing an *in vitro* excision repair assay (Wood et al. 1988). This work might be an important contribution to experiments aimed at reconstitution of the excision repair process once the isolated *ERCC*-proteins are available.

The mechanism of mammalian DNA excision will no doubt be much more complex than the *E.coli* system. Not only the large number of repair mutants points in this direction, but also the recent indications of Hanawalt and coworkers for a close interaction between DNA repair processes and transcription (Mellon et al. 1987). Compelling evidence has been obtained for a role of *ERCC-1* in DNA excision repair. However, to further explore the significance of the *in vivo* function of *ERCC-1* it will be interesting to investigate whether 'reverse genetics' can be applied to manipulate *ERCC-1* expression. Considering the low level of transcription and most likely also low translation level, inhibition of *ERCC-1* expression in cultured cells by adding antisense oligonucleotides to the culture medium (Stein and Cohen, 1988) might be feasible. If such results can also be achieved by transfection of 'antisense' gene constructs, it might be of interest to attempt generation of transgenic mice carrying an antisense *ERCC-1* transgene. This will open the way the study the role of *ERCC-1* in processes like differentation, ageing and tumorigenesis.

IV SUMMARY

A number of indications suggest that defects in DNA repair systems are responsible for the cellular hypersensitivity to DNA damage and cancer proneness of a number of hereditary diseases. However, only in case of xeroderma pigmentosum (XP) profound evidence has been obtained for mutations in genes involved in DNA excision repair. Cell fusion experiments with cell lines from different XP patients have provided evidence for ten XP complementation groups. In addition, by *in vitro* mutagenesis a number of ultraviolet light (UV)-sensitive rodent mutant cells with XP-like features have been obtained. At present the rodent classification of UV-sensitive mutants consists of seven complementation groups. Although only limited information is available, thusfar evidence for overlap with the classification of XP groups is still lacking. Hence, it is possible that ten to seventeen different gene products are involved in mammalian DNA excision repair.

It is the aim of the work described in this thesis to isolate and characterize human genes involved DNA excision repair. This will facilitate the understanding of the mechanism of this repair process whereas it also provides an important step to better understand the relationship between DNA damage and carcinogenesis.

Many attempts to complement the genetic defect in XP by genomic DNA transfer have failed which is most likely due to the poor transfectability of human cells. In contrast to XP cells, UV- sensitive Chinese hamster ovary (CHO) mutants are more suitable for genomic transfection which renders it possible to isolate the human excision repair cross complementing (*ERCC*) genes that correct the CHO mutations.

The isolation of these genes opens the possibility to unravel the complex mechanism of DNA excision repair. In addition, it is still hoped that part of the spectra of rodent and human UV- sensitive cells overlap. In that case the cloning of *ERCC*-genes by genomic transfections with CHO mutants might at the same time aid the elucidation of the genetic defects in XP. The experimental work described in this dissertation concerns the cloning and characterization of the first human DNA repair gene, designated *ERCC-1*, which complements the UV and mitomycin-C sensitivity of CHO mutants of complementation group 1. The *ERCC-1* gene was cloned after cotransfection of high molecular weight human DNA with a dominant marker *Ecogpt* to CHO mutant 43-3B. The linked transfer of the *ERCC-1* and the *Ecogpt*

gene in the second round of transfection allowed the isolation of the *ERCC-1* gene from a genomic cosmid library using a probe from the dominant marker. Correction of the repair deficient phenotype of 43-3B cells by transfection of cosmid 43-34 confirmed the cloning of *ERCC-1* (Appendix paper 1). Using single copy probes from cosmid 43-34 the human *ERCC-1* cDNA was isolated and in subsequent experiments human *ERCC-1* cDNA probes allowed the cloning of the mouse *ERCC-1* gene and cDNA. The results on the characterization of the human and murine *ERCC-1* genes are described in Appendix papers II, III and IV.

ERCC-1 is located on human chromosome 19q. The gene consists of 10 exons which span a region of 15-17 kb. The last exon of the human *ERCC-1* gene has a variable size due to alternative polyadenylation which yields 3.4 and 3.8 kb transcripts in addition to the major 1.1 kb *ERCC-1* mRNA. ERCC-1 transcription is not notably induced in UV-irradiated HeLa cells and a low constitutive transcription level is found for *ERCC-1* expression in a several human cell lines and different mouse organs and stages of development examined. Both the human and mouse *ERCC-1* promoter are devoid of known promoter elements and might represent a category of constitutive lowly expressed genes. Evidence has been obtained for alternative splicing of exon VIII of the human *ERCC-1* gene. Transfection of a human *ERCC-1* cDNA clone without exon VIII to 43-3B cells does not correct the DNA repair defect of 43-3B cells. Furthermore, no evidence is found for alternative splicing of mouse *ERCC-1* transcripts. Therefore, the biological significance of the shorter *ERCC-1* transcripts in human cells is unclear. Transfection of full length human and murine *ERCC-1* cDNA to 43-3B cells results in a fully repair proficient phenotype indicating that they encode functional *ERCC-1* proteins. The mammalian *ERCC-1* gene encodes a protein of 32 - 33 kD (297-298 amino acids). A nuclear location signal and DNA binding domain have been tentatively assigned to the protein based on structural amino acid similarity with functional domains in other proteins. Significant amino acid sequence homology was found with the yeast excision repair protein *RAD10*. The pattern of homology showed much less sequence conservation in the N-terminal part of the protein than in the middle and C-terminal portion and was similar to the pattern of homology between human and mouse *ERCC-1*. The possibility that the N-terminus is less relevant for the repair function of *ERCC-1* was confirmed in transfection experiments with a 5' truncated *ERCC-1* cDNA. A construct lacking the sequence coding the first 54 amino

acids was still competent to correct the 43-3B mutation. All introduced mutations in the C- terminal half of *ERCC-1* abolish its repair function.

The *ERCC-1* protein is 87 amino acids longer than *RAD10*. Intriguingly, this C-terminal extension of the protein exhibits significant amino acid similarity to parts of the *E.coli* repair proteins *uvrA* and *uvrC*. This might indicate that during evolution different DNA repair functions have been accommodated in *ERCC-1*. Another important conclusion is that (parts of) the mechanism of DNA excision repair is (are) well conserved throughout evolution. A possible implication of this finding is that functional studies on the *ERCC-1* protein might benefit from the work on the related yeast and bacterial repair proteins. Furthermore, the evolutionary conservation between *ERCC-1* and *RAD10* suggests that the already isolated yeast *RAD* genes can be recruted to clone repair genes of higher organisms based on nucleotide sequence homology.

In order to investigate whether the *ERCC-1* gene specifically corrects CHO mutants of complementation group 1, the cloned gene was transfected to representative lines of the first 6 groups. The transfection experiments consistently showed that *ERCC-1* only corrects group 1 mutants suggesting that these cells carry a mutation in the endogenous Chinese hamster *ERCC-1* gene. Moreover, these results endorse the idea that the classification of rodent repair mutants is based on mutations in different genes (Appendix paper V).

Using similar approaches the possible relationship between *ERCC-1* and human DNA repair syndromes has been investigated (Appendix paper VI). Southern and Northern blot experiments have not revealed abnormalities in the endogenous *ERCC-1* gene and its expression in cell lines from XP group A through I and Cockayne's syndrome (CS) cells of group A and B. Introduction of the cloned *ERCC-1* gene into these cells did not yield an altered cellular response to UV-irradiation. It is inferred from these experiments that *ERCC-1* is not involved in the genetic defect of the XP and CS complementation groups examined.

Appendix paper VII reports that the homology between *ERCC-1* and *RAD10* extends to the genomic organization of both genes. The 3' *ERCC-1* region harbors an antisense transcription unit, yielding a 2.6 kb RNA, which is overlapping with the last *ERCC- 1* exon. Also the *RAD10* gene overlaps in the 3' region with an antisense transcription unit. This is the first example of a naturally occurring overlapping antisense transcription in the human

genome whereas this phenomenon is also exceptional in yeast. In view of its rarity, the fact that antisense transcription is found in evolutionarily related genes suggests an important biological function.

SAMENVATTING

Een aantal aanwijzingen suggereren dat defecten in DNA herstel systemen verantwoordelijk zijn voor cellulaire overgevoeligheid voor DNA-schade en kankerpredispositie bij diverse erfelijke afwijkingen. Echter, alleen in het geval van xeroderma pigmentosum (XP) is er een duidelijke bewijs voor mutaties in genen betrokken bij DNA excisie herstel. Celfusie experimenten met cellijnen van verschillende XP patienten hebben tien XP complementatiegroepen aangetoond. Daarnaast zijn momenteel zeven groepen beschreven van door middel van *in vitro* mutagenese verkregen UV-gevoelige knaagdier cellijnen met XP- achtige eigenschappen. Hoewel slechts een beperkte hoeveelheid informatie beschikbaar is, tot nu toe zijn er geen directe aanwijzingen voor overlap tussen de XP en knaagdier klassificaties, hetgeen betekent dat mogelijk tien tot zeventien verschillende genprodukten betrokken zijn bij DNA excisie herstel in zoogdiercellen.

Het doel van het in dit proefschrift beschreven werk is de isolatie en karakterisering van humane DNA excisie herstel genen. Hierdoor kan meer inzicht worden verkregen in het mechanisme van dit herstelproces, terwijl het tevens een belangrijke stap kan zijn om de relatie tussen DNA schade en het ontstaan van kanker beter te kunnen begrijpen.

Vele pogingen om het XP-defect te corrigeren door transfectie met genomisch DNA van normale cellen zijn mislukt. De meest waarschijnlijke verklaring hiervoor is het geringe vermogen van humane cellen om DNA op te nemen in genomische transfectie experimenten. In tegenstelling tot XP cellijnen zijn Chinese hamster ovarium (CHO) cellen zeer geschikt voor genomische transfecties. Hierdoor is het mogelijk om in combinatie met moleculaire kloneringstechnieken humane 'excision repair cross complementing' (*ERCC*)-genen te isoleren die het defect in UV- gevoelige CHO mutanten corrigeren.

De isolatie van deze humane *ERCC*-genen opent de mogelijkheid om het mechanisme van DNA excisie herstel te ontrafelen. Daarnaast is er de hoop dat er overlap bestaat tussen de spectra van CHO en XP complementatiegroepen waardoor de klonering van DNA herstelgenen via genomische trans-

fecties met CHO mutanten tevens zou kunnen leiden tot de opheldering van het genetisch defect in een of meerdere XP complementatiegroepen.

Het experimentele werk dat in dit proefschrift wordt beschreven betreft de klonering en karakterisering van het eerste humane DNA herstelgen, *ERCC-1* genoemd, dat de UV en mitomycine-C gevoeligheid van CHO mutanten in complementatiegroep 1 corrigeert.

Het *ERCC-1* gen is gekloneerd na cotransfectie van humaan genomisch DNA met de dominante marker *Ecogpt* naar de CHO mutant 43-3B. De gekoppelde overdracht van het *ERCC-1* en *Ecogpt* gen in de tweede ronde van transfectie maakte het mogelijk om het *ERCC-1* gen te isoleren uit een cosmide-bank van een secundaire transformant met een probe van de dominante vector. De correctie van het DNA herstel defect in 43-3B cellen na transfectie van cosmide 43-34 bevestigde vervolgens de klonering van *ERCC-1* (Appendix publicatie I).

Met behulp van unieke probes van cosmide 43-34 kon het humane *ERCC-1* cDNA worden geisoleerd, terwijl vervolgens met humane cDNA probes het homologe *ERCC-1* gen en cDNA van de muis kon worden gekloneerd. De resultaten van de klonering en karakterisering van het humane en het muize *ERCC-1* gen zijn beschreven in Appendix publicaties II, III and IV.*ERCC-1* is gelokaliseerd op de q-arm van het humane chromosoom 19. Het gen bestaat uit tien exons en heeft een lengte van 15 tot 17 kb. Het laatste exon heeft een variabele grootte als gevolg van alternative polyadenylering hetgeen resulteert in *ERCC-1* transcripten van 3.4 kb en 3.8 kb naast het belangrijkste *ERCC-1* transcript van 1.1 kb. In HeLa cellen wordt de *ERCC-1* transcriptie niet geinduceerd door UV bestraling. Een laag constitutief niveau van expressie is gevonden voor *ERCC-1* transcriptie in diverse humane cellijnen en geanalyseerde muize organen en ontwikkelingsstadia. De humane en muize *ERCC-1* promoters bevatten geen bekende promoter elementen en zouden daardoor een klasse van laag geexpresseerde genen kunnen representeren. Een aantal experimentele gegevens tonen aan dat *ERCC-1* exon VIII alternatief gespliced kan worden. Transfectie van het cDNA zonder dit exon leidt niet tot correctie van het DNA herstel defect in 43-3B cellen. Er zijn geen aanwijzingen voor alternatieve splicing van het muize *ERCC-1* transcript en het is daarom nog onduidelijk wat de biologische significantie van het kortere humane *ERCC-1* transcript is.

Transfectie van het volledige humane en muize 1.1 kb cDNA naar 43-3B cellen resulteert in complete correctie van het DNA herstel deficiente phe-

notype hetgeen aangeeft dat beide cDNA's een functioneel *ERCC-1* eiwit koderen. De *ERCC-1* cDNA sequentie kodeert voor een eiwit van 32-33 kD (297-298 aminozuren). Op basis van sequentie overeenkomst met functionele domeinen in andere eiwitten zijn in het *ERCC-1* eiwit een kern lokatie signaal en een DNA bindend domein gepostuleerd. Verder blijkt het *ERCC-1* eiwit een significante aminozuur homologie te hebben met het gist DNA herstel eiwit *RAD10*. Het N-terminale deel van beide eiwitten toont veel minder sequentie konservering dat het midden en C-terminale deel. Dit patroon van homologie wordt ook tussen het muize en humane *ERCC-1* eiwit gevonden. De mogelijke verklaring dat het N-terminale *ERCC-1* deel minder belangrijk is, wordt ondersteund door transfectie-experimenten met een 5' afgebroken *ERCC-1* cDNA. Een construct dat de koderende sequentie voor 54 N-terminale aminozuren mist, is in staat de 43-3B mutatie volledig te corrigeren. Alle mutaties in het C- terminale deel resulteerden in volledige inactivatie van *ERCC-1* DNA herstel functie. Het *ERCC-1* eiwit is 87 aminozuren langer dan *RAD10* en het is intrigerend dat juist dit extra C-terminale deel van *ERCC-1* significante homology vertoont met delen van de *E.coli* DNA excisie herstel eiwitten *uvrA* en *uvrC*. Dit zou kunnen betekenen dat tijdens de evolutie diverse DNA herstel functies in het het *ERCC-1* gen zijn verenigd. Een andere belangrijke conclusie is dat het er op lijkt dat (delen van) het mechanisme van DNA excisie herstel zijn gekonserveerd. Dit heeft als mogelijke implicaties dat resultaten van functionele studies met de gist en bacterie eiwitten van nut kunnen zijn voor het begrijpen van het mechanisme van DNA herstel in hogere organismen. Verder suggereert de evolutionaire conservering van DNA hersteleiwitten dat reeds gekloneerde gist DNA herstelgenen mogelijk gebruikt kunnen worden voor de klonering van homologe genen van hogere eukaryoten.

Om te onderzoeken of het *ERCC-1* gen specifiek het defect in CHO mutanten van komplementatiegroep 1 complementeert, is DNA van het gekloneerde gen op cosmide 43-34 getransfecteerd naar representatieve cellijnen van de eerste 6 CHO groepen. UV overlevings- en UDS- experimenten tonen aan dat *ERCC-1* alleen groep 1 mutanten complementeert. Dit suggereert dat deze CHO cellen gemuteerd zijn in het endogene Chinese hamster *ERCC-1* gen. Bovendien onderschrijft dit resultaat de gedachte dat de klassificatie knaagdier DNA herstel mutanten gebaseerd is op mutaties in verschillende genen (Appendix publicatie V).

Op vergelijkbare wijze is de mogelijke betrokkenheid van *ERCC-1* met

human DNA herstel syndromen onderzocht (Appendix publicatie VI). Door middel van Southern en Northern blot experimenten zijn geen afwijkingen gevonden in het endogene *ERCC-1* gen van XP-cellijnen uit XP groep A tot en met I en Cockayne's syndrome (CS) cellen van groep A en B. Introductie van het gekloneerde *ERCC-1* gen in deze cellen resulteerde niet in een veranderde cellulaire response op UV-bestraling. Op grond van deze bevindingen is geconcludeerd dat het *ERCC-1* gen niet betrokken is bij het genetisch defect in XP en CS.

Tenslotte wordt in Appendix publicatie VII gerapporteerd dat er ook met betrekking tot de genomische organisatie van *ERCC-1* en *RAD10* opvallende overeenkomsten bestaan. Zowel het humane *ERCC- 1* gen als het gist *RAD10* gen overlappen in de 3' regio met een gen dat door de complementaire DNA streng wordt gekodeerd. Voor zover bekend is dit het eerste voorbeeld van overlappende antisense transcriptie in humane cellen terwijl het ook in gist een zeldzaam verschijnsel is. Met het oog op deze uitzonderlijkheid suggereert de konservering van antisense transcriptie in loci van gerelateerde DNA herstelgenen een belangrijke biologische functie.

REFERENCES

Albino, A.P., Graf, L.H., Kantor, R.R.S., McLean, W., Silagi, S. and Old, L.J. (1985). DNA-mediated transfer of human melanoma cell surface glycoprotein gp130: Identification of transfectants by erythrocyte rosetting. Mol.Cell.Biol. 5, 692-697.

Almoguera, C., Shibata, D., Forrester, K., Martin, J., Arnheim, N. and Perucho, M. (1980). Most human carcinomas of the exocrine pancreas contain mutant c-K-*ras* genes. Cell 53, 549-554.

Arikan, E., Kulkarni, S., Thomas, D.C. and Sancar, A. (1986). Sequences of the *E.coli uvrB* gene and protein. Nucl.Acids Res. 14, 2637-2650.

Arrand, J.E., Squires, S., Bone, N.M. and Johnson, R.T. (1987). Restoration of UV-induced excision repair in Xeroderma D cells transfected with the *den V* gene of bacteriophage T4. EMBO J. 6, 3125-3131.

Barbacid, M. (1986). Oncogenes and human cancer: cause or consequence? Carcinogenesis 7, 1037-1042.

Barbacid, M. (1987). *RAS* Genes. Ann.Rev.Biochem. 56, 779-827.

Barker, D.G., White, J.H.M. and Johnston, L.H. (1985). The nucleotide sequence of the DNA ligase gene (CDC9 from *Saccharomyces cerevisiae*: a gene which is cell-cycle regulated and induced in response to DNA damage. Nucl.Acids Res. 23, 8323-8336.

Bishop, J.M. (1987). The molecular genetics of cancer. Science 235, 305-311.

Bohr, V.A., Smith, C.A. Okumoto, D.S. and Hanawalt, P.C. (1985). DNA repair in 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.

Bohr, V.A., Okumoto, D.S., Ho, L. and Hanawalt, P.C. (1986). Characterization of a DNA repair domain containing the dihydrofolate reductase gene in Chinese hamster ovary cells. J. Biol. Chem. 261, 16666-16672.

Bohr, V.A. (1987). Differential DNA repair within the genome. Cancer Rev. 7, 28-55.

Bohr, V.A. and Hanawalt, P.C. (1987). Enhanced repair of pyrimidine dimers in coding and non-coding genomic sequences in CHO cells expressing a prokaryotic repair gene. Carcinogenesis 8, 1333-1336.

Bohr, V.A., Chu, E.H.Y., van Duin, M., Hanawalt, P.C. and Okumoto, D. (1988). Human DNA repair gene restores normal pattern of preferential DNA repair in CHO cells. Nucl. Acids Res. (in press).

Bos, J.L., Fearon, E.R., Hamilton, S.R., Verlaan-de Vries, M., van Boom, J.H., van der A.J. and Vogelstein, B. (1987). Prevalence of *ras* gene mutations in human colorectal can cers. Nature 327, 293-297.

Boyce, R. and Howard-Flanders, P. (1964). Release of ultraviolet light induced thymine dimers from DNA in *E.coli* K-12. Proc.Natl.Acad.Sci. USA, 51, 293- 300.

Bradshaw, H.D. (1983). Molecular cloning and cell cycle- specific regulation of a functional human thymidine kinase gene. Proc.Natl.Acad.Sci.USA 80, 5588-5591.

Brash, D.E. and Haseltine, W.A. (1982). UV-induced mutation hotspots occur at DNA damage hotspots. Nature 298, 189- 192.

Brash, D.E., Seetharam, S., Kraemer, K.H., Seidman, M.M. and Bredberg, A. (1987). Photoproduct frequency is not the major determinant of UV-base substitution hot spots or cold spots in human cells. Proc.Natl.Acad.Sci. USA 84, 3782-2786.

Bredberg, A., Kraemer, K.H. and Seidman, M.M. (1986). Restricted ultraviolet mutational

spectrum in a shuttle vector propagated in xeroderma pigmentosum cells. Proc. Natl. Acad. Sci. USA 83, 8273-8277.

Buchwald, M., Ng, J., Clarke, C. and Duckworth-Rysiecki, G. (1987). Studies of gene transfer and reversion to mitomycin C resistance in Fanconi anemia cells. Mutat. Res. 184, 153-159.

Burtscher, H., Cooper, A.J., Couto, L.B., Harosh, I., Lambert, C., Naumovski, L., Siede, W., Weiss, W.A., Thompson, L.H. and Friedberg, E.C. (1988a). The excision repair *RAD* genes and rad proteins of *S.cerevisiae*: recent progress. J.Cell Biochem. Suppl. 12A, 239.

Burtscher, H.J., Cooper, A.J. and Couto, L.B. (1988b). Cellular response to DNA damage in the yeast *Sacharomyces cerevisiae*. Mutat. Res. 194, 1-9.

Caron, P.R., Kushner, S.R. and Grossman, L. (1985). Involvement of helicase II (*uvrD* gene product) and DNA polymerase I in excision mediated by the uvrABC protein complex. Proc. Natl. Acad. Sci. USA 82, 4925-4929.

Chan, J.Y.H., Becker, F.F., German, J. and Ray, J.H. (1987). Altered DNA ligase I activity in Bloom's syndrome cells. Nature 325, 357-359.

Chao, M.V., Bothwell, M.A., Ross, A.H., Kaprowski, H., Lanahan, A.A., Randall Buck, C. and Sehgal, A. (1986). Gene transfer and molecular cloning of the human NGF receptor. Science 232, 518-521.

Cirullo, R.E., Dana, S. and Wasmuth, J.J. (1983). Efficient procedure for transferring specific human genes into Chinese hamster cell mutants: interspecific transfer of the human genes encoding leucyl and asparaginyl-tRNA synthetases. Mol. Cell. Biol. 3, 892-902.

Cleaver, J.E. (1968). Defective repair replication of DNA in xeroderma pigmentosum. Nature 218, 652-656.

Cleaver, J.E. and Painter, R.B. (1968). Evidence for repair replication of HeLa cell DNA damaged by ultraviolet light. Biochim.Biophys.Acta 161, 552-554.

Cleaver, J.E. (1969). Xeroderma pigmentosum: a human disease in which an initial stage of DNA repair is defective. Proc.Natl.Acad.Sci. USA 63, 428-435.

Cleaver, J.E. (1972). Xeroderma pigmentosum: variants with normal DNA repair and normal sensitivity to ultraviolet light. J.Invest.Dermatol. 58, 124-128.

Cleaver, J.E. and Karentz, D. (1987). DNA repair in man: Regulation by a multigene family and association with human disease. Bio Essays 6, 122-127.

Cleaver, J.E., Corte's, 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.

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

Collins, A., Johnson, R.T. and Boyle, J.M. (1987). Molecular biology of DNA repair. J.Cell Sci.Suppl. 6

Cooper, G.M. and Lane, M.A. (1984). Cellular transforming genes and oncogenesis. Biochim. Biophys. Acta 738, 9-20.

Crute, J.J., Wahl, A.F. and Bambara, R.A. (1986). Purification and characterization of two new high molecular weight forms of DNA polymerase. Biochemistry 25, 26-36.

Day, R.S. III, Babich, M.A., Yarosch, D.B. and Scudiero, D.A. (1987). The role of O^6-methylguanine in human cell killing, sister chromatid exchange induction and mutagenesis: a review. J.Cell Sci. Suppl. 6, 333-353.

Diatloff-Zito, C., Papadopoulo, D., Averbeck, D. and Moustachi, E. (1986). Abnormal response

to DNA cross-linking agents of Fanconi anemia fibroblasts can be corrected by transfection with normal human DNA. Proc.Natl.Acad.Sci.USA 83, 7034-7038.

De Jonge, A.J.R., Vermeulen, W., Keijzer, W., Hoeijmakers, J.H.J. and Bootsma, D. (1983). Microinjection of human cell extracts corrects xeroderma pigmentosum defect. EMBO J. 2, 637-641.

De Jonge, A.J.R., 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 Jong, P.J., Grosovsky, A.J. and Glickman, B.W. (1988). Spectrum of spontaneous mutation at the *APRT* locus of Chinese hamster ovary cells: an analysis of the DNA sequence level. Proc.Natl.Acad.Sci.USA 85, 3499-3503.

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

Dickstein, R., Doll Huh, N., Sandlie, J. and Grossmann, L. (1988). The expression of the *Escherichia coli uvrA* in human cells. Mutat.Res. 193, 75-86.

Ding, R., Ghosh, K., Eastman, A. and Bresnick, E. (1985). DNA- mediated transfer and expression of a human DNA repair gene that demethylates O^6-methylguanine. Mol.Cell.Biol. 5, 3293-3296.

Doll, R. (1986). Cancer: a world-wide perspective. Biochemical and molecular epidemiology of cancer, Liss, New York, pp. 111-125.

Doolittle, R.F., Johnson, M.S., Husain, I., Van Houten, B., Thomas, D.C. and Sancar, A. (1986). Domainal evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. Nature 323, 451-453.

Dresler, S.L. and Lieberman, M.W. (1983). Requirement of ATP for specific incision of ultraviolet-damaged DNA during excision repair in permeable human fibroblasts. J.Biol.Chem. 258, 12269-12273.

Drobetsky, E.A., Grosovsky, 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.

DuBridge, R.B. and Calos, M.P. (1987). Molecular approaches to the study of gene mutation in human cells. Trends in Genet. 3, 293-297.

Duckworth-Rysiecki, G., Cornish, K., Clarke, C. and Buchwald, M. (1985). Identification of two complementation groups in Fanconi anemia. Somat.Cell Mol.Genet. 11, 35-41.

Duesberg, P.H. (1987). Cancer genes: rare recombinations instead of activated oncogenes (a review). Proc.Natl.Acad.Sci.USA 84, 2117-2124.

Dulhanty, A.M. and Whitmore, G.F. (1988). Complementation of the incision repair defect in a group 4 CHO mutant with human DNA that lacks abundant repetitive sequences. J.Cell.Biochem. 12A, 312.

Edenberg, H. and Hanawalt, P. (1972). Size of repair patches in the DNA of ultraviolet-irradiated HeLa cells. Biochim.Biophys.Acta 272, 361-372.

El-Gewely, M.R. and Oxender, D.L. (1985). Gene transfer and cloning of the amino acid transport system L from human cells. Ann.NY Acad.Sci. 456, 417-419.

Elliott, A.G. and Lloyd, R.S. (1988). Molecular analysis of plasmid DNA repair within UV-irradiated *E-coli*: implications of distributive and processive repair pathways. J.Cell. Biochem. 12A, 235.

Enninga, J.C., Groenendijk, R.T.L., Filon, A.R., van Zeeland, A.A. and Simons, J.W.I.M. (1986). The wavelength dependence of UV-induced pyrimidine dimer formation, cell killing and mutation induction in human diploid skin fibroblasts. Carcinogenesis 7, 1829-1836.

Fallows, D., Kent, R.B., Nelson, D.L., Rettig Emanuel, J., Levenson, R. and Housman, D.E. (1987). Chromosome mediated transfer of the murine Na, K-ATPase alpha subunit confers ouabain resistance. Mol.Cell.Biol. 7, 2985-2987.

Fischer, E., Keijzer, W., Thielmann, H.W., Popanda, O., Bohnert, E., Edler, L., Jung, E.G. and Bootsma, D. (1985). A ninth complementation group in xeroderma pigmentosum, XP-I. Mutat.Res. 145, 217-225.

Fleer, R., Nicolet, C.M., Pure, G.A. and Friedberg, E.C. (1987). *RAD4* gene of *Saccharomyces cerevisiae*: Molecular cloning and partial characterization of a gene that is inactivated in *Escherichia coli*. Mol.Cell.Biol. 7, 1180- 1192.

Fornace, A.J., Kohn, K.W. and Kann, H.E. (1976). DNA single strand breaks during repair of UV-damage in human fibroblasts and abnormalities of repair in xeroderma pigmentosum. Proc.Natl.Acad.Sci.USA 73, 39-43.

Forrester, K., Almoguera, C., Han, K., Grizzle, W.E. and Perucho, M. (1987). Detection of high incidence of K-*ras* oncogenes during human colon tumorigenesis.

Franklin, W.A., Lo, K.M. and Haseltine, W.A. (1982). Alkaline lability of novel fluorescent photoproducts produced in ultraviolet light irradiated DNA. J.Biol.Chem. 257, 13535- 13543.

Franklin, W.A., Doetsch, P.W. and Haseltine, W.A. (1985). Structural determination of the ultraviolet light-induced thymine-cytosine pyrimidine-pyrimidone (6-4) photoproduct. Nucl.Acids Res. 14, 5317-5325.

Friedberg, E.C., Ehmann, U.K. and Williams, J.I. (1979). Human diseases associated with defective DNA repair. Adv.Rad.Biol. 8, 85-174.

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

Friedberg, E.C. (1987). The molecular biology of nucleotide excision repair of DNA: recent progress. J.Cell.Sci.Suppl. 6, 1-23.

Friedberg, E.C., Cooper, A.J., Couto, L., Fleer, R., Kalainov, D. Naumovski, L., Nicolet, C., Rehm, J., Robinson, G.W., Ryo, B.H., Sugino, A., Sugino, T., Weiss, W.A.. and Yang, E. (1987a). Recent advances in the characterization of genes and proteins for nucleotide excision repair in the yeast *S.cerevisiae*. Photochem.Photobiophys.Suppl. 333-341.

Friedberg, E.C., Backendorf, C., Burke, J., Collins, A., Grossman, L., Hoeijmakers, J.H.J., Lehmann, A.R., Seeberg, E., van der Schans, G.P. and van Zeeland, A.A. (1987b). Molecular aspects of DNA repair. Mutat.Res. 184, 67-86.

Gartler, S.M. (1964). Inborn errors of metabolism at the cell culture level, in Fishbein, M. (ed): Second International conference on congenital malformations. New York, International Medical Congress, p. 94.

Gatti, R.A., Davis, R.C., Weeks, D.E., Jaspers, N.G.J., Sparkes, R.S. and Lange, K. (1987). Genetic linkage studies of ataxia telangiectasia: phenotypic blood markers. Disease Markers 5, 207-213.

Giannelli, F., Pawsey, S.A. and Avery, J.A. (1982). Differences in patterns of complementation of the more common groups of XP: possible implication. Cell 29, 451-458.

Graham, F.L. and van der Eb, A.J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52, 456-467.

Greco, A., Ittmann, M. and Basilico, C. (1987). Molecular cloning of a gene that is necesary for G1 progression in mammalian cells. Proc. Natl. Acad. Sci. USA 84, 1565-1569.

Green, M.H.L., Lowe, J.E., Arlett, C.F., Harcourt, S.A., Burke, J.F., James, M.R. and Lehmann, A.R. (1987). A gamma-ray- resistant derivative of an Ataxia telangiectasia cell line obtained following DNA mediated gene transfer. J. Cell Sci. Suppl. 6, 127-137.

Grossman, L. Caron, P. Mazur, S. and Oh, E.Y. (1988). The dynamics of the *Escherichia coli uvr* nucleotide excision repair system. J.Cell.Chem.Suppl. 12A, 234.

Hahn, S., Pinkham, J., Wei, R., Miller, R. and Guarente, L. (1988). The *HAP3* regulatory locus of *Saccharomyces cerevisiae* encodes divergent overlapping transcripts. Mol.Cell.Biol. 8, 655-663.

Hanawalt, P.C., Cooper, P.K., Ganesan, A.K. and Smith, C.A. (1979). DNA repair in bacteria and mammalian cells. Ann.Rev.Biochem. 48, 783-836.

Hanawalt, P.C. (1986). Intragenomic heterogeneity in DNA damage processing: potential implications for risk assessment. In: Mechanisms of DNA damage and repair, Grossman, L. and Upton, A.C. (eds.). Plenum Press, pp. 489-498.

Hanawalt, P.C. and Sarasin, A. (1986). Cancer-prone hereditary diseases with DNA processing abnormalities. Trends in Genet. 2, 124-129.

Haseltine, W.A. (1983). Ultraviolet light repair and mutagenesis revisited. Cell 33, 13-17.

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.

Haynes, R.H. and Kunz, B.A. (1981). DNA repair and mutagenesis in yeast. in: The molecular biology of the yeast Saccharomyces: Life cycle and inheritance (Strathern, J., Jones, E, and Broach, J. (eds.) Cold Spring Harbor Laboratories, Cold Spring Harbor, New York pp. 371-414

Hickson, J.D. and Harris, A.L. (1988). Mammalian DNA repair - use of mutants hypersensitive to cytotoxic agents. Trends in Genet. 4, 101-106.

Hoeijmakers, J.H.J., van Duin, M., Westerveld, A., Yasui, A. and Bootsma, D. (1986). Identification of DNA repair genes in the human genome. Cold Spring Harbor Symp. Quant. Biol. Vol. LI, 91-101.

Hoeijmakers, J.H.J. (1987). Characterization of genes and proteins involved in excision repair of human cells. Cell Sci.Suppl. 6, 111-125.

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, 111-119.

Hoeijmakers, J.H.J. (1988). Use of microneedle injection to study DNA repair in mammalian cells. In: DNA repair: A laboratory manual of research procedures. (Friedberg, E.C. and Hanawalt, P.C., eds.). Marcel Dekker Inc., New York pp. 133-150.

Hoeijmakers, J.H.J., van Duin, M., Weeda, G., van der Eb, A.J., Troelstra, C., Eker, A., Jaspers, N.G.J., Westerveld, A. and Bootsma, D. (1988a). Analysis of mammalian excision repair: from mutants to genes and gene products. In: Mechanisms and Consequences of DNA Damage Processing, UCLA Symposium on Molecular Biology, New Series Vol. 83, in press.

Hoeijmakers, J.H.J., Westerveld, A. and Bootsma, D. (1988b). Methods and Strategies for molecular cloning of mammalian DNA repair genes by DNA mediated gene transfer. In: Friedberg, E.C. and Hanawalt, P.C. (eds.), DNA repair, a laboratory manual for research procedures, Marcel Dekker Inc., New York Vol. 3, 181-201.

Hori, T-A., Shiomi, T. and Sato, K. (1983). Human chromosome 13 compensates a DNA repair defect in UV-sensitive mouse cells by mouse-human cell hybridization. Proc.Natl.Acad.Sci.USA

80, 5655-5659.

Howard-Flanders, P., Boyce, R.P. and Theriot, R. (1966). Three loci in *Escherichia coli* K-12 that control the excision of thymine dimers and certain other mutagen products from host or phage DNA. Genetics 53, 1119-1136.

Hoy, C.A., Thompson, L.H., Mooney, C.L. and Salazar, E.P. (1985). Defective DNA cross-links removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. Canc. Res. 45, 1737-1743.

Hsu, C., Kavathas, P. and Herzenberg, L.A. (1984). Cell-surface antigens expressed on L-cells transfected with whole DNA from non-expressing and expressing cells. Nature 312, 68- 69.

Husain, J., van Houten, B., Thomas, D.C., Abdel-Monem, M. and Sancar, A. (1985). Effect of DNA polymerase I and DNA helicase II on the turnover rate of uvrABC excision nuclease. Proc.Natl.Acad.Sci.USA 82, 6774-6778.

Husain, J., van Houten, B., Thomas, D.C. and Sancar, A. (1986). Sequences of *Escherichia coli uvrA* gene and protein reveal two potential ATP binding sites. J.Biol.Chem. 261, 4895- 4901.

Ishizaki, K., Shima, A., Sekiguchi, T. and Tabeke, H. (1981). Microcell-mediated transfer of human DNA repair genes into xeroderma pigmentosum cells. Jpn.J.Genet. 56, 105-115.

Ittmann, M., Greco, A. and Basilico, C. (1987). Isolation of the human gene that complements a temperature- sensitive cell cycle mutation in BHK cells. Mol.Cell.Biol. 7, 3386-3393.

Jaspers, N.G.J., Painter, R.B., Paterson, M.C., Kidson, C. and Inoue, T. (1985). Complementation analysis of ataxia telangiectasia.In: Gatti and Swift, eds. pp. 147-162.

Jaspers, N.G.J. (1985). DNA synthesis in ataxia telangiectasia. Thesis, Erasmus University Rotterdam.

Jaspers, N.G.J., Gatti, R.A., Baan, C., Linssen, P.C.M.L. and Bootsma, D. (1988). Genetic complementation analysis of ataxia telangiectasia and the Nijmegen Breakage Syndrome: a survey of 50 patients. Cytogenet. Cell Genet. (in press).

Jentsch, S., Mc Grath, J.P. abd Varshavsky, A. (1987). The yeast DNA repair gene *RAD6* encodes a ubiquitin- conjugating enzyme. Nature 329, 131-134.

Johnson, R.T., Arrand, J.E. and Bone, N. (1988). A Hamster gene which complements the xeroderma D defect. J.Cell.Biochem. suppl. 12A, p. 315.

Jolly, D.J., Esty, A.C., Bernard, H.U. and Friedmann (1982). Isolation of a genomic clone partially encoding human hypoxanthine phosphorybosyltransferase. Proc. Natl. Acad. Sci. USA 79, 5038-5041.

Jolly, D.J., Okayama, H., Berg, P., Esty, A.C., Filpula, D., Bohlen, P., Johnson, G.G., Shivery, J.E., Hunkapillar, T. and Friedman, T. (1983). Isolation and characterization of a full-length expressible cDNA for human hypoxanthine phosphoribosyltransferase. Proc.Natl.Acad.Sci.USA 80, 477- 481.

Kai, R., Ohtsubo, M., Sekiguchi, M., Nishimoto, T. (1986). Molecular cloning of a human gene that regulates chromosome condensation and in essential for cell proliferation. Mol.Cell Biol. 6, 2027-2032.

Kaina, B., van Zeeland, A.A., Backendorf, C., Thielmann, H.W. and van de Putte, P. (1987). Transfer of human genes conferring resistance to methylating mutagens, but not to UV irradiation and cross-linking agents, into Chinese ovary cells. Mol.Cell.Biol. 7, 2024-2030.

Kano, Y. and Fujiwara, Y. (1983). Defective thymine dimer excision from xeroderma pigmentosum chromatin and its characteristic catalysis by cell free extracts. Carcinogenesis 4, 1419-1424

Karentz, D. and Cleaver, J.E. (1986). Excision repair in xeroderma pigmentosum group C but

not group D is clustered in a small fraction of the total genome. Mutat.Res. 165, 165-174.

Karentz, D., Mitchell, D. and Cleaver, J.E. (1987). Correction of excision repair in xeroderma pigmentosum by hamster chromosome fragments involves both classes of pyrimidine dimers. Somat.Cell Mol.Genet. 13, 621-625.

Kaufmann, W.K. and Briley, L.P. (1987). Reparative strand incision in saponin-permeabilized human fibroblasts. Mutat.Res.184, 237-243.

Kavathas, P. and Herzenberg, L.A. (1983). Stable transformation of mouse L cells for human membrane T-cell differentiation antigens, HLA and {2-microglobulin: selection by fluorescent activated cell sorting. Proc.Natl.Acad.Sci.USA 80, 524-528.

Kavathas, P., Sukhatme, V.P., Herzenberg, L.A. and Parnes, J.R. (1984). Isolation of the gene encoding the human T-lymphocyte differentiation antigen Leu-2 (T8) by gene transfer and cDNA subtraction. Proc.Natl.Acad.Sci.USA 81, 7688- 7692.

Keijzer, W., Verkerk, A. and Bootsma, D. (1982). Phenotypic correction of the defect in xeroderma pigmentosum after fusion with isolated cytoplasts. Exp.Cell Res. 140, 119- 125.

Keijzer, W., Stefanini, M., Bootsma, D., Verkerk, A., Geurts van Kessel, A.H.M., Jongkind, J.F. and Westerveld, A. (1987). Localization of a gene involved in complementation of the defect in xeroderma group A cells on human chromosome 1. Exp. Cell Res. 169, 490-501.

Keijzer, W., Mulder, M.P., Langeveld, J.C.m., Bos, J.L., Bootsma, D. and Hoeijmakers, J.H.J. (1988). Establishment and characterization of a melanoma cell line from a xeroderma pigmentosum patient: activation of N-ras at a potential pyrimidine dimer site (submitted for publication).

Keyse, S.M. and Tyrell, R.M. (1985). Excision repair in permeable arrested human skin fibroblasts damaged by UV (254 nm) radiation: Evidence that α|- and β-polymerases act sequentially at the repolymerisation step. Mutat.Res. 146, 109-119.

Kojima, H., Tsuchiya, S., Sekiguchi, K., Gelinas, R. and Hakomori, S.-I. (1987). Predefined gene transfer for expression of a glycosphingolipid antigen by transfection with a cosmid genomic library prepared from a cell line in which the specific glycosphingolipid is highly expressed. Biochim.Biophys.Res.Comm. 143, 716-722.

Kondo, S., Satoh, Y. and Kuroki, T. (1987). Defect in UV- induced unscheduled DNA synthesis in cultured epidermal keratinocytes from xeroderma pigmentosum. Mutat.Res. 183, 95-101.

Kraemer, K.H. (1983). Heritable disease with increasaed sensitivity to cellular injury. In: Fitzpatrick, T.B., Eisen, A.Z., Wolff, K. Freedberg, I.M. and Austen, K.F. (eds.) Update: Dermatology in general medicine. McGraw- Hill Book Company, New York, pp. 113-142.

Kraemer, K.H., Lee, M.M. and Scotto, J. (1984). DNA repair protects against cutaneous and internal neoplasia: Evidence from xeroderma pigmentosum. Carcinogenesis 5, 511-514.

Kraemer, K.H. and Slor, H. (1985). Xeroderma pigmentosum. Clin. Dermatol. 3, 33-69.

Kraemer, K.H., Lee, M.M. and Scotto, J. (1987). Xeroderma pigmentosum. Arch. Dermatol. 123, 241-250.

Kuhn, L.C., McClelland, A. and Ruddle, F.H. (1984). Gene transfer, expression and molecular cloning of the human transferrin receptor gene. Cell 37, 95-103.

Lau, Y.F. and Kan, Y.W. (1984). Direct isolation of the functional human thymidine kinase gene with a cosmid shuttle vector. Proc.Natl.Acad.Sci.USA 81, 414-418.

Leadon, S.A. Zolan, M.E. and Hanawalt, P.C. (1983). Restricted repair of aflatoxin B1 induced DNA damage in alpha DNA of monkey cells. Nucl.Acids.Res. 11, 5675-5689.

Leadon, S.A. and Hanawalt, P.C. (1984). Ultraviolet irradiation of monkey cells enhances the repair of DNA adducts in alpha DNA. Carcinogenesis 5, 1505-1510.

Lebkowski, J.S., Clancy, S., Miller, J.H. and Calos, M.P. (1985). The *LacI* shuttle: rapid analysis of the mutagenic specificity of ultraviolet light in human cells. Proc.Natl. Acad.Sci.USA 82, 8606-8610.

Lee, M.G. and Nurse, P. (1987). Complementation used to clone a human homologue of the fission yeast cell cycle control gene *cdc2*. Nature 327, 31-35. **Legerski**, R.J., Brown, D.B., Peterson, C.A. and Robberson, D.L. (1984). Transient complementation of xeroderma pigmentosum cells by microinjection of poly(A)$^+$ RNA. Proc. Natl. Acad. Sci. USA 81, 5676-5679.

Lehmann, A.R., Kirk-Bell, S., Arlett, C., 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- irradation. Proc. Natl. Acad. Sci. USA 72, 219-223.

Lehmann, A.R. (1982a). Xeroderma pigmentosum, Cockayne syndrome and ataxia telangiectasia: Disorders relating DNA repair to carcinogenesis. Cancer Surveys 1, 93-118.

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

Lehmann, A.R. (1985). Use of recombinant DNA techniques in cloning DNA repair genes and in the study of mutagenesis in human cells. Mutat.Res. 150, 61-67.

Lin, P.F. and Ruddle, F.H. (1981). Murine DNA repair gene located on chromosome 4. Nature 289, 191-194.

Lin, P.F., Zhao, S.Y. and Ruddle, F.H. (1983). Genomic cloning and preliminary characterization of the human thymidine kinase gene. Proc.Natl.Acad.Sci.USA 80, 6528-6532.

Littman, D.R., Thomas, Y., Maddon, P.J., Chess, L. and Axel, R. (1985). The isolation and sequence of the gene encoding T8: A molecule defining functional classes of T lymphocytes. Cell 40, 237-246.

Lowly, J., Pellicer, A., Jackson, J.F., Sim, G.K., Silverstein, S. and Axel, R. (1980). Isolation of transforming DNA: cloning of the hamster aprt gene. Cell 22, 817-823.

MacInnes, M.A., Bingham, J.D., Thompson, L.H. and Strniste, G.F. (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.

Maddon, P.J., Littman, D.R., Godfrey, M., Maddon, D.E., Chess, L. Axel, R. (1985). The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: A new member of the immunoglobulin gene family. Cell 42, 93-104.

Madhani, H.D., Bohr, V.A. and Hanawalt, P.C. (1986). Differential DNA repair in transcriptionally active and inactive proto-oncogenes: *c-abl* and *c-mos*. Cell 45, 417- 423.

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

Maher, V.M., Ouellette, L.M., Curren, R.D. and McCormick, J.J. (1976). Frequency of ultraviolet light-induced mutations is higher in xeroderma pigmentosum variant cells than in normal human cells. Nature 261, 593-595.

Maher, V.M. and McCormick, J.J. (1976). Effect of DNA repair on the cytotoxicity and mutagenicity of UV-irradiation and of chemical carcinogens in normal and xeroderma pigmentosum cells. In: 'Biology of Radiation Carcinogenesis' (Yuhas, J.M., Tennant, R.W., Regan, J.D., eds.) Raven Press, New York, pp. 129-145.

Mansbridge, J.N. and Hanawalt, P.C. (1983). Domain-limited repair of DNA in ultraviolet irradiated fibroblasts from xeroderma pigmentosum complementation group C, in: Friedberg,

E.C. and Bridges, B.R. (eds.): Cellular responses to DNA damage, Liss, New York pp. 195-207

Matsumoto, A. and Fujiwara, Y. (1987). Human DNA transfection confers a partial UV-resistance on xeroderma pigmentosum cells. Eight International Congress of Radiation Research, Edinburgh, Abstr. C12-7V, p.110.

Mayne, L.V. (1982). Excision repair in Cockayne syndrome. Mutat.Res. 106, 179-189.

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. (1984). Inhibitors of DNA synthesis (aphidicolon and araC/HU) prevent the recovery of RNA synthesis after UV-irradiation. Mutat.Res. 131, 187-191.

Mayne, L.V., Mullender, L.H.F. and van Zeeland, A.A. (1988a). Cockayne's syndrome: A UV-sensitive disorder with a defect in the repair of transcribing DNA but normal overall, excision repair. In: Mechanisms and Consequences of DNA Damage Processing (Friedberg, E.C. and Hanawalt, P. eds.). UCLA Symposia on Molecular and Cellular Biology, New Series vol. 83 (in press).

Mayne, L.V., Jones, T., Dean, S.W., Harcourt, S.A., Lowe, J.E., Priestley, A., Steingrimsdottir, H., Sykes, H., Green, M.H.L. and Lehmann, A.R. (1988b). SV40-transformed normal and DNA-repair-deficient human fibroblasts can be transfected with high frequency but retain only limited amounts of integrated DNA. Gene 66, 65-76.

McClelland, A., Kamarck, M.E. and Ruddle, F.H. (1987). Molecular cloning of receptor genes by transfection. Meth.Enzymol. 147, 280-291.

McCormick, J.J., Kateley-Kohler, S., Watanabe, M. and Maher, V.M. (1986). Cells from xeroderma pigmentosum variants are abnormally sensitive to transformation by UV, buth their excision repair is 'error-free'. Biochem.Molec.Epidemiol. Canc., Liss, New York, pp. 419-426.

McGrath, J.P., Capon, D.J., Smith, D.H., Chen, E.Y., Seeburg, P.H., Goeddel, D.V. and Levinson, A.D. (1983). Structure and organization of the human Ki-*ras* proto-oncogene and a related processed pseudogene. Nature 304, 501-506.

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.

Miller, R.D., Prakash, L., and Prakash, S. (1982). Genetic control of excision of *Saccharomyces cerevisiae* interstrand DNA cross-links induced by psoralen plus near UV light. Mol.Cell.Biol. 2, 939-948.

Montisano, D. and Hankinson, O. (1985). Transfection by genomic DNA of cytochrome P_1-450 enzymatic activity and inducibility. Mol.Cell.Biol. 5, 698-704.

Mortelmans, K., Friedberg, E.C., Slor, H., Thomas, G. and Cleaver, J.E. (1976). Defective thymine dimer excision by cell free extracts of xeroderm pigmentosum cells. Proc. Natl. Acad. Sci. USA 73, 2757-2761.

Mosbaugh, D.W. and Linn, S. (1983). Excision repair and DNA synthesis with a combination of HeLA DNA polymerase { and DNase V. J.Biol.Chem. 258, 108-118.

Moshell, W.N., Ganges, M.B., Lutzner, M.A., Coon, H.G., Barett, S.F., Dupuy, J.M. and Robbins, J.H. (1983). A new patient with both xeroderma pigmentosum and Cockayne's syndrome establishes the new xeroderma pigmentosum complementation group H. In: Friedberg, E.C., Bridges, B. (eds.): Cellular response to DNA damage, Liss, New York pp. 209-213.

Mullenders, L.H.F., van Kesteren, A.C., Bussmann, C.J.M., van Zeeland, A.A. and Natarajan,

A.T. (1984). Preferential repair of nuclear matrix associated DNA in xeroderma pigmentosum complementation group C. Mutat.Res. 141, 75- 82.

Mullenders, L.H.F., van Kesteren, A.C., Bussmann, C.J.M., van Zeeland, A.A. and Natarajan, A.T. (1986). Distribution of UV-induced repair events in higher order chromatin loops in human and hamster fibroblasts. Carcinogenesis 7, 995- 1002.

Mullenders, L.H.F., van Zeeland, A.A. and Natarajan, A.T. (1987). The localization of ultraviolet induced excision repair in the nucleus and the distribution of repair events in higher order chromatin loops in mammalian cells. J. Cell Sci.Suppl. 6, 243-262.

Nakabeppu, Y. and Sekiguchi, M. (1981). Physical association of pyrimidine dimer DNA glycosylase and apurinic-apyrimidinic DNA endonuclease essential for repair of ultraviolet- damaged DNA. Proc.Natl.Acad.sci. USA 78, 2742-2746.

Nakauchi, H., Nolan, G.P., Hsu, C., Huang, H.S., Kavathas, P. and Herzenberg, L.A. (1985). Molecular cloning of Lyt-2, a membrane glycoprotein marking a subset of mouse T lymphocytes: Molecular homology to its human counterpart, Leu-2/T8, and to immunoglobulin variable regions. Proc.Natl.Acad.Sci.USA 82, 5126-5130.

Naumovski, L. Chu, G., Berg, P. and Friedberg, E.C. (1985). *RAD3* gene of *Sacchoromyces 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. (1986). Analysis of the essential and excision repair functions of the *RAD3* gene of *Saccharomyces cerevisiae* by mutagenesis. Mol. Cell. Biol. 6, 1218-1227.

Naumovski, L. and Friedberg, E.C. (1987). The *RAD3* gene of *Saccharomyces cerevisia*: isolation and characterization of a temperature-sensitive mutant in the essential function and of extragenic suppressors of this mutant. Mol.Gen.Genet. 209, 458-466.

Newman, R., Domingo, D., Trotter, J. and Trowbridge, J. (1983). Selection and properties of a mouse L-cell transformant expressing human transferrin receptor. Nature 304, 643- 645.

Nicolet, C.M. Chenevert, J.M., Friedberg, E.C. (1985). The *RAD2* gene of *Saccharomyces cerevisiae*: nucleotide sequence and transcript mapping. Gene 36, 225-234.

Nishida, C., Reinhard, P. and Linn, S. (1988). DNA repair synthesis in human fibroblasts requires DNA polymerase. J.Biol.Chem. 263, 501-510.

Ogawa, H., Shimada, K. and Tomizawa, J. (1968). Studies on radiation-sensitive mutants of *E.coli*. I. Mutants defective in the repair synthesis. Mol. Gen. Genet. 101, 227-244.

Oh, E.Y. and Grossman, L. (1986). The effect of *Escherichia coli Uvr* protein binding on the topology of supercoiled DNA. Nucl.Acids Res. 14, 8557-8571.

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

Okayama, H. and Berg, P. (1983). A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. Mol.Cell Biol. 3, 280-289.

Okayama, H. and Berg, P. (1985). Bacteriophage lambda vector for transducing a cDNA clone library into mammalian cells. Mol.Cell.Biol. 5, 1136-1142.

Okumoto, D.S. and Bohr, V.A. (1987). DNA repair in the metallo thionein gene increases with transcriptional activation. Nucl.Acids Res. 15, 10021-10030.

Ozkaynak, E., Finley, D., Solomon, M.J. and Varsharsky, A. (1987). The yeast ubiquitin genes: a family of natural gene fusions. EMBO J. 6, 1429-1439.

Papadopoulo, D., Averbeck, D. and Moustachi, E. (1987). The fate of 8-methoxypsoralen-

photoinduced DNA interstrand crosslinks in Fanconi's anemia cells of defined genetic complementation groups. Mutat.Res. 184, 271-280.

Parrish, J.A., Rox Anderson, R., Urbach, F. and Pitts, D. (1978). Sources of UV-A, in: UV-A, biological effects of ultraviolet radiation with emphasis on human responses to longwave ultraviolet, Plenum, New York, pp. 7-36.

Penn, J. (1985). Ultraviolet light and skin cancer. Immun. Today 6, 206-208.

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

Perucho, M., Hanahan, D., Lipsich, L and Wigler, M. (1980). Isolation of the chicken thymidine kinase gene by plasmid rescue. Nature 285, 207-210.

Phipps, J., Nasim, A. and Miller, D.R. (1985). Recovery repair and mutagenesis in Schizosaccharomyces pombe. Adv.in Genetics 23, 1-72.

Pinney, D.F., Pearson-White, S.H., Konieczny, S.F., Latham, K.E. and Emerson jr., C.P. (1988). Myogenic lineage determination and differentiation: Evidence for a regulatory gene pathway. Cell 53, 781-793.

Pritchard, C. and Goodfellow, P.N. (1986). Development of new methods in human gene mapping: selection for fragments of the human Y chromosome after chromosome-mediated gene transfer. EMBO J. 5, 979-985.

Protic-Sabljic, M. and Kraemer, K.H. (1985). One pyrimidine dimer inactivates expression of a transfected gene in xeroderma pigmentosum cells. Proc.Natl.Acad.Sci.USA 82, 6622-6626 .

Protic-Sabljic, M., Tuteja, N., Munson, P.J., Hauser, J., Kraemer, K.H. and Dixon, K. (1986). UV-light induced cyclo butane pyrimidine dimers are mutagenic in mammalian cells. Mol.Cell.Biol. 6, 3349-3356.

Rebora, A. and Crovato, F. (1987). Trichothiodysthrophy with xeroderma pigmentosum (group D) mutation. J.Am.Acad. Dermatol. 16, 940-947.

Regan, J. and Setlow, R. (1974). Two forms of repair in the DNA of human cells damaged by chemical carcinogens and mutagens. Cancer Res. 34, 3318-3325.

Regan, J.D., Trosko, J.E., Carrier, W.L. (1968). Evidence for excision of ultraviolet-induced pyrimidine dimers from the DNA of human cells in vitro. Biophys. J. 8, 319-325.

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

Reynolds, R.J. and Friedberg, E.C. (1981). Molecular mechanism of pyrimidine dimer excision in *Saccharomyces cerevisiae*: incision of ultraviolet-irradiated deoxyribonucleic acid in vivo. J. Bacteriol. 146, 692-704.

Reynolds, P., Prakash, L., Dumais, D., Perozzi, G. and Prakash, S. (1985). Nucleotide sequence of the *RAD10* gene of *Saccha romyces 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 malignant neoplasms and abnormal DNA repair. Ann.Intern.Med. 80, 221-248.

Robinson, G.W., Nicolet, C.M., Kalainov and Friedberg, E.C. (1986). A yeast excision-repair

gene is inducible by DNA damaging agents. Proc.Natl.Acad.Sci. USA 83, 1842-1846.

Royer-Pokora, B. and Haseltine, W.A. (1984). Isolation of UV resistant revertants from a xeroderma pigmentosum group A cell line. Nature 311, 390-392.

Rubin, J.S., Joyner, A.L., Bernstein, A. and Whitmore, G.F. (1983). Molecular identification of a human DNA repair gene following DNA mediated gene transfer. Nature 306, 206-208.

Rubin, J.S., Prideaux, V.R., Huntington, F.W., Dulhanty, A.M., Whitmore, G.F. and Bernstein, A. (1985). Molecular cloning and chromosomal localization of DNA sequences associated with a human DNA repair gene. Mol.Cell.Biol. 5, 398-405.

Sancar, A. and Rupp, W.D. (1983). A novel repair enzyme: uvrABC excision nuclease of *Escherichia coli* cuts a DNA strand on both sides of the damaged region. Cell 33, 249-260.

Sancar, G.B., Sancar, A. and Rupp, W.D. (1984). Sequences of the *E.coli uvrC* gene and protein. Nucl.Acids.Res. 12, 4593-4608.

Sancar, A., Franklin, A. and Sancar, G. (1985). Repair of psoralen and acetylaminofluorene DNA adducts by ABC excinuclease. J.Mol.Biol. 184, 725-734.

Sauerbier, W. and Hercules, K. (1978). Gene and transcription unit mapping by radiation effects. Ann.Rev.Genet. 12, 329- 363.

Schaaper, R.M., Kunkel, T.A. and Loeb, L.A. (1983). Infidelity of DNA synthesis associated with bypasss of apurinic sites. Proc.Natl.Acad.Sci. 80, 487-491.

Schmid, C.W. and Jelinek, W.R. (1982). The Alu family of dispersed repetitive sequences. Science 216, 1065-1070.

Schultz, R.A., Barbis, D.P. and Friedberg, E.C. (1985). Studies on gene transfer and revision to UV resistance in xeroderma pigmentosum cells. Somat.Cell Mol.Genet. 11, 617-624.

Schultz, R.A., Saxon, P.J., Glover, T.W. and Friedberg, E.C. (1987). Microcell-mediated transfer of a single human chromosome complements xeroderma pigmentosum group A fibroblasts. Proc.Natl.Acad.Sci.USA 84, 4176-4179.

Schultz, R.A., Saxon, P.J., Chao, C.K. and Friedberg, E.C. (1988). Phenotypic complementation of xeroderma pigmentosum cells following microcell-mediated chromosome transfer. J. Cell.Biochem.Suppl. 12A, 245.

Seal, G., Brech, K., Karp, S.J., Cool, B.L. and Sirover, M.A. (1988). Immunological lesions in human uracil DNA glycosylase: association with Bloom syndrome. Proc.Natl. Acad.Sci. USA, 2339-2343.

Sedgwick, S.G. (1986). Stability and change through DNA repair. In: Galas, D.J., Kirkwood, T.B.L. and Rosenberger, R.F. (eds).: Accuracy in molecular processes, Chapman and Hall, London.

Seeberg, E., Nissen-Meijer, J. and Strike, P. (1976). Incision of ultraviolet-irradiated DNA by extracts of *E.coli* requires three different gene product. Nature 263, 524- 526.

Seeberg, E. and Steinum, A.L. (1982). Purification and properties of the *uvrA* protein from *Escherichia coli.* Proc.Natl.Acad.Sci.USA 79, 988-992.

Seeberg, E., Steinum, A.L., and Blingsmo, O.R. (1982). Two separable protein species which restore uvrABC endonuclease activity in extracts from *uvrC* mutated cells. Biochimie 64, 825-828

Sekiguchi, T., Yoshida, M.C., Sekiguchi, M. and Nishimoto, T. (1987). Isolation of a human X chromosome-linked gene essential for progression from G1 to S phase of the cell cycle. Exp.Cell Res. 169, 395-407.

Shaham, M., Adler, B., Gonguly, S. and Chaganti, R.S.K. (1987). Transfection of normal hu-

man and Chinese hamster DNA corrects diepoxybutane-induced chromosomal hypersensitivity of Fanconi anemia fibroblasts. Proc. Natl. Acad. Sci. USA 84, 5853-5857.

Shiomi, T., Hieda-Shiomi, N. and Sato, K. (1982). Isolation of UV-sensitive mutants of mouse L5178Y cells by a cell suspesnsion spotting method. Somat.Cell Genet. 8, 329-345.

Shiomi, T., Hieda-Shiomi, N., Sato, K., Yoshizumi, T. and Nakazawa, T. (1988). Cell fusion-mediated improvements in transfection competence for repair deficient mutant of mouse T cell line. Somat.Cell Molec.Genet. 14, 195-203.

Shows, T.B., McAlpine, P.J. and Miller, R.J. (1984). Proceedings 7th Human Gene Mapping Conference (Sparker, R.S., Berg, K., Evans, H.J. and Klinger, H.P. (eds.) Karger, Basel pp. 340-390.

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.

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

Siciliano, M.J., Carrano, A.V. and Thompson, L.H. (1986b). Assignment of a human DNA-repair gene associated with sister-chromatid exchange to chromosome 19. Mutat.Res. 174, 303-308.

Smith, C.A. and Okumoto, D.S. (1984). Nature of DNA repair synthesis to inhibitors of polymerase | in human cells. Biochemistry 23, 1383-1391.

Smith, C.A. (1987). DNA repair in specific sequences in mammalian cells. J.Cell.Sci.Suppl. 6, 225-241.

Squires, S. and Johnson, R.T. (1988). Kinetic analysis of UV- induced incision discriminates between fibroblasts from different xeroderma pigmentosum complementation groups, XPA heterozygotes and normal individuals. Mutat.Res. 193, 181-192.

Stefanini, M.,, Lagomarsini, P. Arlett, C.F., Borrone, C, Crovato F., Trevisan, G., Gordone, 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.

Stein, C.A. and Cohen, J.S. (1988). Oligodeoxynucleotides as inhibitors of gene expression: a review. Cancer Res. 48, 2659-2668.

Strauss, B., Rabkin, S., Sagher, D. and Moore, P. (1982). The role of DNA polymerase in base substitution mutagenesis on non-instructional templates. Biochemie 64, 829-838

Strauss, B.S. (1985). Cellular aspects of DNA repair. Adv.in Cancer Res. 45, 45-105.

Strniste, G.F., Chen, D.J., de Bruin, D., McCoy, L.S., Luke, J.A., Mudgett, J.S., Nickols, J.W., Okinaka, R.T., Tesmer, J.G. and MacInnes, M.A. (1988). Restoration of Chinese hamster cell radiation resistance by the human repair gene *ERCC-5* and progress in molecular cloning of this gene. In: Mechanisms and Consequences of DNA damage processing, UCLA Symposia on Molecular Biology New Series vol. 83, in press.

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

Sung, P., Prakash, L., Weber, S. and Prakash, S. (1987b). The *RAD3* gene of *Saccharomyces cerevisiae* encodes a DNA dependent ATPase. Proc.Natl.Acad.Sci.USA 84, 6045-6049.

Szybalska, E.H. and Szybalski, W. (1962). Genetics of human cells, IV. DNA mediated heritable transformation of a bio chemical trait. Proc.Natl.Acad.Sci.USA 48, 2026-2034.

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

Tanaka, K., Hayakawa, M., 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., Satokata, I., Ogita, Z. and Okada, Y. (1988). Toward the molecular cloning of the gene for xeroderma pigmentosum complementation group A by DNA transfection method. J.Cell. Biochem. Suppl. 12A, p. 246.

Teitz, T., Naiman, T., Avissar, S.S., Bar, S., Okayama, H. and Canaani, D. (1987). Complementation of the UV-sensitive phenotype of a xeroderma pigmentosum human cell line by transfection with a cDNA clone library. Proc.Natl.Acad. Sci.USA 84, 8801-8804.

Thielmann, H.W., Edler, L., Popanda, O. and Friemel, S. (1985). Xeroderma pigmentosum patients from the federal republic of Germany: Decrease in post-UV colony forming ability in 30 xeroderma pigmentosum fibroblast strains is quantitatively correlated with a decrease in DNA - incising capacity. J.Canc.Res.Clin.Oncol. 109, 227-240.

Th'ng, J.P.H. and Walker, I.G. (1983). DNA repair patch size measurements with nucleosomal DNA. Carcinogenesis 4, 975- 978.

Th'ng, J.P.H. and Walker, I.G. (1985). Excision repair of DNA in the presence of aphidicolin. Mutat.Res. 165, 139-150.

Thompson, L.H., Busch, O.B., Brookman, K., Mooney, C.I. and Glaser, D.A. (1981). Genetic diversity of UV-sensitive DNA repair mutants of Chinese hamster ovary cells. Proc.Natl. Acad.Sci.USA 78, 3734-3737.

Thompson, L.H., Brookman, K.W., Dillehay, L.E., Mooney, C.L. and Carrano, A.V. (1982a). Hypersensitivity to mutation and sister-chromatid exchange induction in CHO cell mutants defective in incising DNA containing UV lesions. Somat.Cell Genet. 8, 759-773.

Thompson, L.H., Carrano, A.V. (1983). Analysis of mammalian cell mutagenesis and DNA repair using *in vitro* selected CHO cell mutants. In: Cellular Responses to DNA Damage, Liss, New York pp. 125-143.

Thompson, L.H. (1985). DNA repair mutants. In: Molecular cell Genetics, M.M. Gottesman (eds.) Wiley and Sons, New York pp. 641-667.

Thompson, L.H., Mooney, C.L. and Brookman, K.W. (1985a). Genetic complementation between UV-sensitive CHO mutants and xero derma pigmentosum fibroblasts. Mutat.Res. 150, 423-429.

Thompson, L.H., Brookman, K.W., Minkler, J.L., Fuscoe, J.C., Henning, K.A. and Carrano, A.V. (1985b). DNA-mediated gene transfer of a human DNA repair gene that controls sister chromatid exchange. Mol.Cell.Biol. 5, 881-884.

Thompson, L.H., Mooney, C.L., Burkhart-Schultz, K., Carrano, A.V. and Siciliano, M.J. (1985c). Correction of a nucleotide-excision repair mutation by human chromosome 19 in hamster-human hybrid cells. Somat.Cell Molec.Genet. 11, 87-92.

Thompson, L.H., Salazar, E.P., Brookman, K.W., Collins, C.C., Stewart, S.A., Busch, D.B. and Weber, C.A. (1987a). Recent progress with DNA repair mutants of Chinese hamster ovary cells. J.Cell Sci.Suppl. 6, 97-110.

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

Molec.Genet. 13, 539-551.

Thompson, L.H. (1988). Use of Chinese hamster ovary cell mutants to study human DNA repair genes. In: Friedberg, E.C. and Hanawalt, P.C. (eds), DNA repair, a laboratory manual of research procedures, Dekker, New York, Vol. 3, 115-132.

Thompson, L.H., Weber, C.A. and Carrano, A.V. (1988). Human DNA repair genes. In: Friedberg, E. and Hanawalt, P. (eds). Mechanisms and consequences of DNA damage processing. UCLA Symposia on Molecular Biology New Series Vol. 83 (in press).

Valerie, K., Green, A.P., de Riel, J.K. and Henderson, E.E. (1987). Transient and stable complementations of ultra violet repair in xeroderma pigmentosum cells by the *denV* gene of bacteriophage T4. Cancer Res. 47, 2967-2971.

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- 9213.

Van Duin, M., Janssen, J.H. de Wit, J., Hoeijmakers, J.H.J., Thompson, L.H., Bootsma, D. and Westerveld, A. (1988a). Transfection of the cloned human excision repair gene *ERCC-1* to UV-sensitive CHO mutants only corrects the repair defect in complementation group 2 mutants. Mutat.Res. 193, 123-130.

Van Duin, M., van den Tol, J., Warmerdam, P., Odijk, H., Meijer, D., Westerveld, A., Bootsma, D. and Hoeijmakers, J.H.J. (1988b). 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., Weeda, G., Klein, B., Hoeijmakers, J.H.J., Bootsma, D. and Westerveld, A. (1988c) The cloned human DNA excision repair gene *ERCC-1* fails to correct xeroderma pigmentosum complementation groups A through I. Mutat. Res. (in press).

Van Houten, B. and Sancar, A. (1987). Repair of N-methyl-N' nitro-N-nitrosoguanidine-induced DNA damage by ABC excinuclease. J.Bacteriol. 169, 540-545.

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.

Varmus, H.E. (1984). The molecular genetics of cellular onco genes. Ann.Rev.Genet. 18, 553-612.

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

Vollberg, T.M., Seal, G. and Sirover, M.A. (1987). Monoclonal antibodies detect conformational abnormality of uracil DNA glycosylase in Bloom's Syndrome cells. Carcinogenesis 8, 1725-1729

Vuksanovic, L. and Cleaver, I.E. (1987). Unique cross-link and monoadduct repair characteristics of a xeroderma pigmentosum revertant cell line. Mutat.Res., 184, 255-263.

Wake, C.T., Gudewitz, T., Porter, T., White, A. and Wilson, J.H. (1984). How damaged is the biologically active subpopulation of transfected DNA. Mol.Cell.Biol. 4, 387- 398.

Walker, I.G. and Th'ng, J.P.H. (1982). Excision-repair patch size in DNA from human KB cells treated with UV-light, or methyl methanesulfonate. Mutat.Res. 105, 277-285.

Walker, G.C. (1985). Inducible DNA repair systems. Annu.Rev.Biochem. 54, 425-457.

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.

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

White, R. and Lalouel, J.-M. (1987). Investigation of genetic linkage in human families. In: Harris, H. and Hirschhorn, K. (Eds.), Adv. in human genetics, Plenum, New York, Vol. 16, 121-228.

Wilcox, D.R. and Prakash, L. (1981). Incision and post-incision steps of pyrimidine dimer removal in excision defective mutants of *Saccharomyces cerevisiae*. J.Bacteriol. 148, 618-623.

Willis, A.E. and Lindahl, T. (1987). DNA ligase I deficiency in Bloom's Syndrome, Nature 325, 355-357.

Willis, A.E., Weksberg, R., Tomlinson, S. and Lindahl, T. (1987). Structural alterations of DNA ligase I in Bloom syndrome. Proc.Natl.Acad.Sci.USA 84, 8016-8022.

Wilson, S., Abbotts, J. and Widen, S. (1988). Progress toward molecular biology of DNA polymerase {. Biochim.Biophys.Acta 949, 149-157.

Wood, R.D. and Burk, H.J. (1982). Repair capability and the cellular age response for killing and mutation induction after UV. Mutat. Res. 95, 505-514.

Wood, R.D., Skopek, T.R. and Hutchinson, F. (1984). Changes in DNA base sequence induced by targeted mutagenesis of lambda phage by ultraviolet light. J.Mol.Biol. 173, 273-291.

Wood, R.D. (1985). Pyrimidine dimers are not the principle pre-mutagenic lesion induced in lambda phage DNA by ultraviolet light. J.Mol.Biol. 184, 577-585.

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.

Yamaizumi, M., Sugano, T., Asahina, H., Okada, Y. and Uchida, T. (1986). Micro-injection of partially purified protein factor restores DNA damage specifically in group A of xeroderma pigmentosum cells. Proc.Natl.Acad.Sci.USA 83, 1476-1479.

Yarosh, D.B., Barnes, D. and Erickson, L.C. (1986). Transfection of DNA from a chloroethylnitrosourea resistant tumor cell line (MER^+) to a sensitive tumor cell line (MER^-) results in a tumor cell line resistant to MNNG and CNU that has increased O^6-methylguanine-DNA methyl transferase levels and reduced levels of DNA interstrand cross-linking Carcinogenesis 7, 1603-1606.

Yeung, A.T., Mattes, W.B., Oh, E.Y. and Grossman, L. (1983). Enzymatic properties of purified *Escherichia coli* uvrABC proteins. Proc.Natl.Acad.Sci.USA 80, 6157-6161.

Zakrzewski, S. and Sperling, K. (1980). Genetic heterogeneity in Fanconi's anemia demonstrated by somatic cell hybrids. Hum.Genet. 56, 81-84.

Zdzienicka, M.Z., Rosa, L., Westerveld, A., Bootsma, D. and Simons, J.W.I.M. (1987). Biological and biochemical consequences of the human *ERCC-1* repair gene after transfection into a repair deficient CHO cell line. Mutat.Res. 183, 69-74

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. (in press).

Zdzienicka, M.Z., Jaspers, N.G.J., van der Schans, G.P., Natarajan, A.T. and Simons, J.W.I.M. Ataxia telangiectasia- like Chinese hamster V79 cell mutants with radioresistant DNA synthesis, chromosomal instability and normal DNA strand break repair. Submitted for publication.

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.

Zolan, M.E., Cortopassi, G.A., Smith, C.A. and Hanawalt, P.C. (1982). Deficient repair of chemical adducts in alpha DNA of monkey cells. Cell 28, 613-619.

Zolan, M.E., Smith, C.A. and Hanawalt, P.C. (1984). Formation and repair of furocoumarin adducts in alpha deoxyribonucleic acid and bulk deoxyribonucleic acid of monkey cells. Biochemistry 23, 63-69.

Zwetsloot, J.C.M., Barbeiro, A.P., Vermeulen, W., Arthur, H.M., Hoeijmakers, J.H.J. and Backendorf, C. (1986). Micro- injection of *Echerichia coli* Uvr A,B,C and C proteins into fibroblasts of xeroderma pigmentosum complementation groups A and C does not result in restoration of UV- induced unscheduled DNA synthesis. Mutat.Res. 166, 89-98.

DANKWOORD

Mede dankzij de steun en toewijding van een groot aantal mensen is dit proefschrift tot stand gekomen. Een ieder die een steen(tje) heeft bijgedragen aan deze klus ben ik daarvoor bijzonder dankbaar. Een aantal van hen wil ik met name noemen.

Mijn promotor, Prof. Dr. D. Bootsma, wil ik bedanken voor het in mij gestelde vertrouwen en voor de mogelijkheden die hij mij heeft geboden om met de vele facetten van wetenschappelijk onderzoek kennis te maken. Dick, ik heb de afgelopen jaren grote bewondering gekregen voor jouw stuurmanskunst en de wijze waarop jij de koers van het onderzoek bewaakt. Ik beschouw het als een voorrecht dat ik in jouw 'DNA-repair' groep heb mogen werken.

Jan Hoeijmakers, mijn co-promotor, wil ik bedanken voor de wijze waarop hij mij heeft begeleid bij de experimenten en het schrijven van de 'papers'. Jan, jouw ongekende inzet en motivatie werken aanstekelijk en zijn van grote betekenis geweest voor het in dit proefschrift beschreven werk.

Andries Westerveld, Jan de Wit en Hanny Odijk ben ik zeer erkentelijk voor hun voortreffelijke transfectiewerk. Het kan niet vaak genoeg gezegd worden dat jullie werk de basis heeft gevormd voor het huidige onderzoek.

Jacolien van den Tol en Marcel Koken wil ik bedanken voor hun belangrijke aandeel in het praktische werk. 'Jaco en Koko', het doet mij een groot plezier dat jullie mij op de dag van de promotie als paranimfen terzijde willen staan.

Rita Boucke dank ik voor haar immer stimulerende goede humeur en voortreffelijke hulp bij het bereiden van manuscripten.

Cor van Dijk ben ik veel dank verschuldigd voor de deskundige en snelle hulp bij het vervaardigen van foto's en tekeningen.

'Onze computerdeskundige', Sjozèf van Baal, komt een bijzonder woord van dank toe voor de fantastische hulp bij het printen van dit proefschrift.

De zorg en stimulans van mijn ouders en de kansen die zij mij geboden hebben, zijn de basis voor het bereiken van de huidige mijlpaal. Ik ben hen daarvoor zeer dankbaar.

Bovenal gaat mijn dank en bewondering uit naar jou, Mike. Jouw steun in de afgelopen jaren en je begrip voor de abnormale werktijden van een promovendus zijn van grote waarde geweest voor het tot stand komen van dit boekje. Het is voor onze toekomst geschreven.

CURRICULUM VITAE

De auteur van dit proefschrift werd op 26 juli 1956 te Rotterdam geboren. In 1974 werd aan de Chr. Scholengemeenschap 'Maarten Luther' het HAVO diploma behaald. De hierna gekozen opleiding tot scheepswerktuigkundige aan de Hogere Zeevaartschool te Rotterdam moest voortijdig worden beeindigd. Daarop trad hij in 1975 alsnog in de voetsporen van zijn biologieleraar. In 1980 werd de lerarenopleiding biologie en scheikunde aan de Stichting Opleiding Leraren (Utrecht) met succes afgerond, waarna hij 'doorstroomde' naar de Rijksuniversiteit te Utrecht. Het doctoraalexamen Biologie werd in oktober 1983 afgelegd (cum laude) met als hoofdvakken Scheikundige Dierfysiologie (prof. Dr. D.I. Zandee, Utrecht) en Moleculaire Neurobiologie (Prof. Dr. W.H. Gispen, Utrecht) en als bijvak Genetica (Prof. Dr. D. Bootsma, Erasmus universiteit Rotterdam). Tevens werd de 1^e graads onderwijsbevoegdheid biologie behaald.

Vanaf 1 november 1983 is hij in dienst van de Erasmus universiteit te Rotterdam en verricht hij wetenschappelijk onderzoek op de afdeling celbiologie en genetica onder leiding van Prof. Dr. D. Bootsma. Dit onderzoek wordt vanaf 1 januari 1986 voor een periode van vier jaar financieel gesteund door MEDIGON, Stichting voor Medisch Onderzoek en Gezondheidsonderzoek in Nederland. De resultaten van dit onderzoek hebben geleid tot het schrijven van dit proefschrift.

APPENDIX PAPER I

Nature 310, 425-429 (1984)

Molecular cloning of a human DNA repair gene

A. Westerveld*, J. H. J. Hoeijmakers, M. van Duin, J. de Wit, H. Odijk, A. Pastink†, R. D. Wood‡§ & D. Bootsma

Department of Cell Biology and Genetics, Erasmus University, Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands
† Department of Medical Biochemistry, State University of Leiden, 2333 AL Leiden, The Netherlands
‡ Lawrence Berkeley Laboratory, Biomedical Division and Graduate Group in Biophysics, University of California, Berkeley, California 94720, USA

Cell strains derived from patients having a hereditary disorder associated with defects in repair of DNA damage such as xeroderma pigmentosum[1,2] and mutants isolated from established rodent cell lines[3,4] provide the tools for genetic and biochemical analysis of DNA repair pathways in mammalian cells. Complementation studies using these cells have illustrated the genetic and biochemical complexity of these pathways[3,5,6]. The precise nature of the genes and gene products involved in these mutants has not yet been resolved. Isolation of repair genes by recombinant DNA technology would open up new approaches to the elucidation of repair mechanisms in mammalian cells. Here we report the molecular cloning of a human repair gene (*ERCC1*) that complements the repair defect in a Chinese hamster ovary (CHO) mutant cell line.

The CHO mutant cell line 43-3B (ref. 7) used in this study is sensitive to UV light and mitomycin C and falls into complementation group 2 of the classification described by Thompson *et al.*[3,7]. The mitomycin C sensitivity of these cells allows efficient selection of repair-competent transformants after transfection with human DNA[8].

The strategy for the isolation of the human repair gene is depicted in Fig. 1. High-molecular weight DNA isolated from HeLa cells is partially cleaved to an average fragment size of 50–60 kilobase pairs (kbp), ligated to linearized plasmid pSV_3gptH and transfected to the 43-3B CHO mutant cells. pSV_3gptH contains the *Escherichia coli gpt* (*Ecogpt*) gene that renders transformed cells resistant to medium containing mycophenolic acid (MPA medium[9]), which provides a means of preselecting the fraction of transfection-competent recipient cells[10]. The physical linkage between pSV_3gptH and the repair gene can be used to identify in an *E. coli* recombinant library of a (secondary) transformant, clones that contain the two genes, using colony hybridization with a labelled *Ecogpt* probe[11].

The transfection frequency of 43-3B cells with the human–pSV_3gptH hybrid DNA molecules for MPA resistance (MPA^r)

* To whom reprint requests should be addressed.
§ Present address: Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, PO Box 5666, New Haven, Connecticut 06511, USA.

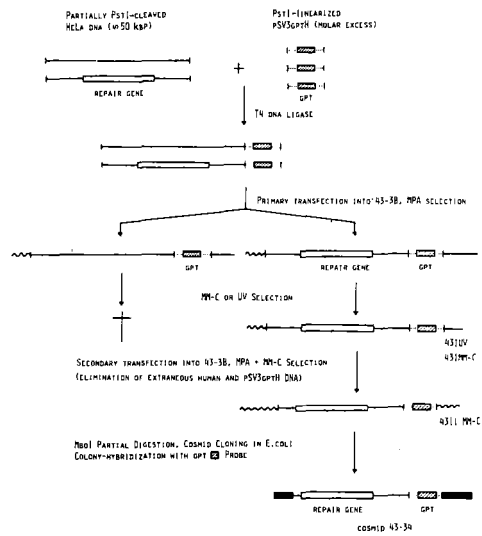

Fig. 1 Experimental strategy for the isolation of the human repair gene *ERCC1*.
Methods: For transfections, $\sim 5 \times 10^5$ 43-3B cells were seeded in 100 mm Petri dishes 1 day before DNA transformation. The cells were grown in F10/Dulbecco's minimal essential medium (DMEM) 1:1 supplemented with antibiotics and 3% fetal and 7% newborn calf serum. High-molecular weight HeLa DNA, isolated as described previously[17], was partially cleaved to an average fragment size of 50–60 kbp with restriction endonuclease *Pst*I. Provided that the repair gene is not exceptionally large (>40–50 kbp), sufficient gene copies in the DNA should remain intact. The restricted DNA was subsequently ligated to a twofold molar excess of *Pst*I-linearized dominant marker pSV_3gptH using T4 DNA ligase (Biolabs). Vector pSV_3gptH is a derivative of pSV_3gpt (ref. 9), from which the simian virus 40 (SV40) early region and a 120-bp *Hin*dIII–*Bgl*II fragment in front of the *gpt* sequence (which contains false start codons[18]) were removed. After testing the ligation on agarose gels the hybrid molecules were transfected into 43-3B CHO repair mutant cells by the calcium phosphate precipitation method[19]. 20 μg of DNA (vector + HeLa DNA) were applied to each Petri dish. Following overnight exposition of the DNA, cells were treated with dimethyl sulphoxide (10% for 30 min) and grown for 24–48 h on non-selecting culture medium to allow expression of the transfected markers. Selection for the dominant marker was in modified MPA medium containing: F10/DMEM 1:1, antibiotics (3% fetal, 7% newborn calf serum), aminopterin (0.2 μg ml^{-1}), thymidine (5 μg ml^{-1}), xanthine (10 μg ml^{-1}), hypoxanthine (15 μg ml^{-1}), mycophenolic acid (25 μg ml^{-1}) and deoxycytidine (2.3 μg ml^{-1}). The selection medium was refreshed every 3–4 days. After the appearance of MPA-resistant (MPA^r) colonies (within 2 weeks) the cells of each Petri dish were reseeded on two dishes, one of which was UV-irradiated (3×5 J m^{-2} at 1-day intervals), the other exposed to mitomycin C (10^{-8} M)[8] in MPA medium. These treatments were lethal to the 43-3B mutant, but not to repair-competent CHO cells. The UV^r and $MM\text{-}C^r$ colonies appearing on the duplicates of one parental Petri dish were expanded into mass culture (10^8–10^9 cells) in selective conditions (except for UV) and characterized with respect to UV and mitomycin C sensitivity (Fig. 3), UV-induced unscheduled DNA synthesis (Table 1) and Southern blot hybridization (Fig. 2). Secondary and tertiary transformation and selection were done as above except that 43IUV and 43IIMM-C DNA, respectively, were used as donor DNAs and that selection for both markers (MPA^r and $MM\text{-}C^r$) was done simultaneously. To construct a cosmid library of 43IIMM-C DNA, a partial *Mbo*I size-fractionated digest of 43IIMM-C DNA with an average fragment size of 40–50 kbp was ligated to *Bam*H1-cleaved pTCF-cosmid vector arms, packaged *in vitro* and transduced into bacterial host ED8767 essentially as described earlier[20]. This library, which consisted of 0.8×10^6 independent recombinants (equivalent to five times the diploid Chinese hamster genome), was screened by colony filter hybridization[20,21] using an *Ecogpt* probe (the isolated[22] small *Sph*I fragment of pSV_3gpt containing the *Ecogpt* gene and part of the SV40 early region[9]). The *gpt* probe was labelled *in vitro* by nick-translation[23], using [α-^{32}P]-labelled dATP and dCTP (Amersham, 3,000 Ci $mmol^{-1}$) to a specific activity of $2\text{–}5 \times 10^8$ c.p.m. μg^{-1}. During hybridization excess unlabelled pTCF plasmid was present as competitor. Hybridized and washed filters (Millipore, HA nitrocellulose, 0.45 μm pore size) were exposed to X-ray film (Fuji Rx) at −70 °C with intensifying screen (Fuji). Seven hybridizing *E. coli* clones were picked from the master filter, rescreened with the same probe and grown in 1 l culture medium (L-broth + ampicillin, 50 μg ml^{-1}). Cosmid DNAs were isolated according to ref. 20. All cosmids transfected into 43-3B as described above using 1–5 μg of cosmid DNA gave MPA^r colonies; only cosmid 43-34 (see Figs 2–4) induced mitomycin C-resistance as well.

was 50 clones per μg DNA per 10^6 cells and for combined $MPA^r + UV^r$ or MPA^r + mitomycin C resistance ($MM\text{-}C^r$) 0.02 clones per μg DNA per 10^6 cells. The latter transformation efficiency is comparable to that found for another CHO mutant of the same complementation group[8] and is in the order of that expected for unique genes[12,13]. One UV^r and one $MM\text{-}C^r$ transformant (designated 43IUV and 43IMM-C respectively) were analysed for the presence of pSV_3gptH and human sequences by blot hybridization and characterized with respect to various repair parameters. Southern blot analysis revealed the presence in both genomes of multiple copies (probably ≥100) of the dominant marker as well as a considerable amount of human DNA (see, for 43IUV, Fig. 2*a*, *e*). As a first end point of repair, we investigated UV-induced unscheduled DNA synthesis in the transformants. The level of unscheduled DNA synthesis in mutant 43-3B is only a fraction of that performed by repair-proficient CHO cells, while the level in transformants 43IUV and 43IMM-C equals that of the CHO wild type (Table 1). As in other rodent cells, the level of unscheduled DNA synthesis in CHO wild-type cells is lower than that in human cells[14]. Survival curves show that both transformants have regained the repair-proficient constitution for UV and MM-C (Fig. 3). The finding that these curves do not coincide with those of CHO wild-type cells might be explained by the heterologous (human) nature of the gene and its product. As repair-competent revertants were never observed in parallel mock-transfection experiments, we conclude that the repair-proficient character of 43IUV and 43IMM-C is due to the stable incorporation and expression of a complementing human repair gene.

After transformation of 43-3B cells with 43IUV DNA (secondary transfection), coupled transfer of MPA^r and $MM\text{-}C^r$ was found to be only slightly lower than that of either marker alone (about 0.02 clones per μg DNA per 10^6 cells), indicating a close physical linkage between the two in the DNA of 43IUV. One secondary transformant, 43IIMM-C, was selected for further characterization. The levels of its unscheduled DNA synthesis and its MM-C- and UV-survival curves were not significantly different from those of the primary transformant (Fig. 3, Table 1). Southern blot analysis of the DNA of 43IIMM-C revealed that three copies of the dominant marker and a specific subset of human DNA fragments had been retained (Fig. 2*a*, *e*). Restriction enzyme mapping data indicated that two of the three pSV_3gptH copies were very closely linked, whereas the third copy was located at some distance from the other two (see below). In a tertiary transformation, combined transfer of MPA^r and $MM\text{-}C^r$ was found with a similar frequency as for the secondary transformation, indicating that no rearrangement had disconnected the two genes. Therefore DNA of the secondary transformant was used to construct a cosmid recombinant library in *E. coli* (Fig. 1).

Restriction endonuclease analysis of cosmids of seven *E. coli* clones hybridizing with the *Ecogpt* probe revealed that all had overlapping sequences (for a representative trio covering the entire 70-kbp cloned region, see Fig. 2*b–d*). Some harboured all three pSV_3gptH copies present in the 43IIMM-C genome (cosmid 43-31), others contained only two (cosmid 43–12) or one copy (cosmid 43-34, Fig. 2*b*). The remaining part of the inserts consisted of only human DNA. From the hybridization with human repeats we conclude that most of the human sequences integrated in 43IIMM-C are covered by these clones (compare Fig. 2*d* and *e*). A physical map of the cosmid inserts (Fig. 4, upper part) appeared to be consistent with the Southern blot data obtained from the genomic DNA of 43IIMM-C. Apparently this transformant has stably integrated a continuous 43IUV DNA fragment longer than 70 kbp.

After transfection to 43-3B all cosmids induced MPA resistance with a high frequency (70 clones per μg DNA per 10^6 cells); however, only cosmid 43–34 transferred resistance to MM-C and UV (Fig. 3) and restoration of unscheduled DNA synthesis (Table 1) with an efficiency comparable to that of the dominant marker. This is more than 3×10^3 higher than that of genomic DNA. Cosmid 43–34 extends the furthest to the right in the map of Fig. 4, and contains only one pSV_3gptH copy, indicating that the UV/MM-C-resistance gene must reside to the right of pSV_3gptH and extends at least in part beyond the overlap with cosmid 43-23 (Fig. 4). This is consistent with the finding that a tertiary transformant, 43IIIMM-C, has retained only the single pSV_3gptH copy and the same fragment of human DNA (Fig. 2*a*, *e*). A more precise localization of the gene was obtained by digesting 43–34 cosmid DNA with various restriction endonucleases and scoring for the intactness of the gene by transfection into the 43-3B cells. Of the enzymes tested, only *Sal*I and *Cla*I left the gene intact; other endonucleases (*Bam*HI,

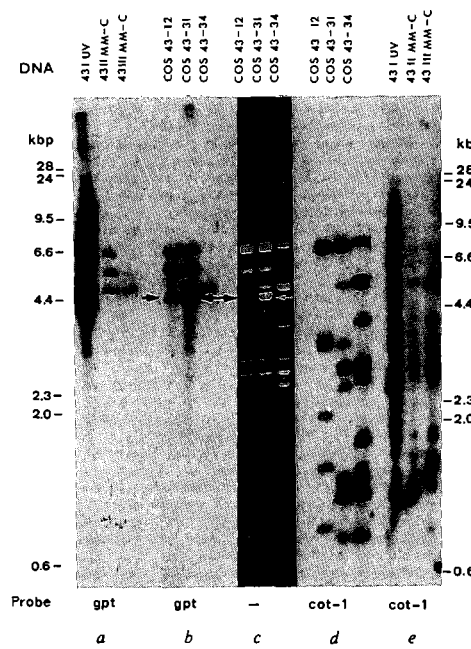

Fig. 2 Fifteen μg of high-molecular weight DNA of 43IUV (primary transformant), 43IIMM-C (secondary transformant) and 43IIIMM-C (tertiary transformant) were digested with *Pst*I. Three μg of cosmids 43-12, -31 and -34 (see Fig. 4) were cleaved with *Sal*I to liberate the insert (Fig. 4; ref. 21) and further digested with *Pst*I. The DNAs were electrophoresed and hybridized with the probes as indicated. Arrows indicate non-equimolar bands, hybridizing with the *Ecogpt* probe, that are derived from contaminating free pSV_3gptH molecules in the cosmid DNA preparation. As these bands were only found in preparations of cosmids harbouring the tandem copies of pSV_3gptH, we presume they have arisen during growth of the bacteria for preparation of cosmid DNA as a result of a recombination event, which might be due to the presence of three to four pBR and at least two *Ecogpt* sequences in these cosmid molecules.
Methods: DNAs were prepared as detailed elsewhere[17,21]. Genomic DNAs were ethanol-precipitated before layering onto a 0.5% agarose gel and electrophoresis. After transfer to a nitrocellulose filter[24], one set of blots was hybridized with the ^{32}P-labelled *Ecogpt* probe (Fig. 1 legend), the other with a ^{32}P-labelled human *cot-1* DNA probe, ligated using T4 DNA ligase before nick-translation to increase its size. Nick-translation, hybridization, washing (0.3 × SSC at 65 °C) and autoradiography were as described in Fig. 1.

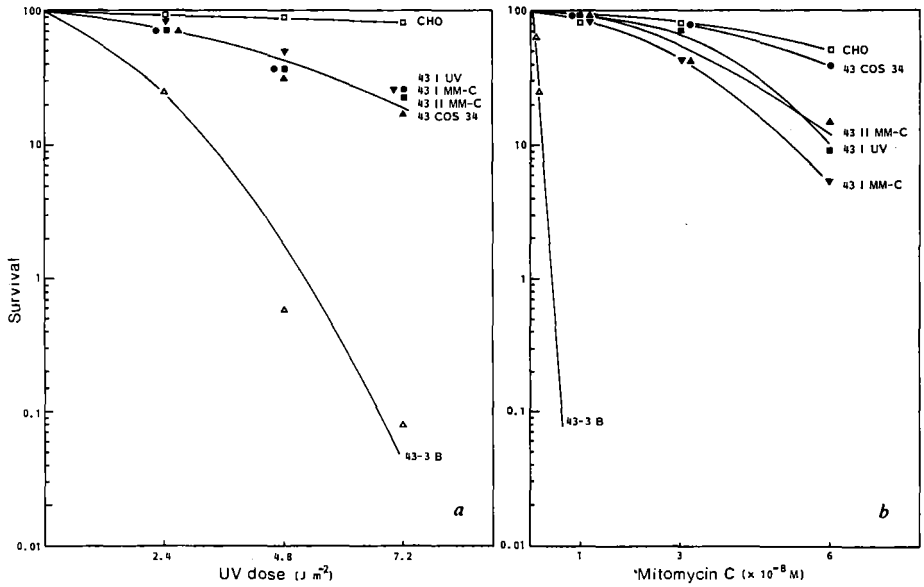

Fig. 3 UV (*a*) and mitomycin C (*b*) sensitivities of CHO (□), 43-3B (△) and the primary (43IIMM-C, 43IUV), secondary (43IMM-C) and cosmid (43cos34) transformants (▼, ■, ▲ and ●, respectively). Cells in normal growth medium were inoculated into 60 or 100 mm Petri dishes at densities of 2×10^2–10^4 cells per dish. One day after seeding the cells were either irradiated with 254 nm UV from a low-pressure mercury, germicidal lamp (Philips TUV lamp) at a dose rate of 0.6 J m^{-2} s^{-1}, or grown in medium with mitomycin C. The cells were cultured for 7 days after which period the clones were fixed, stained and scored.

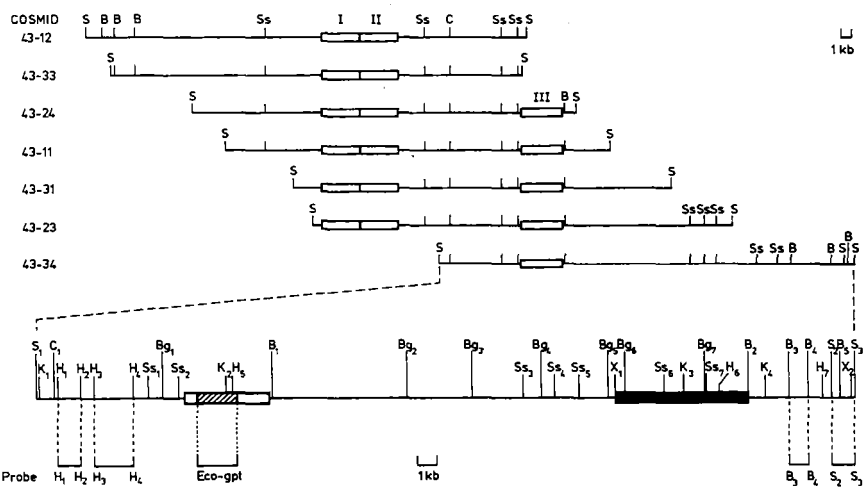

Fig. 4 Upper part, physical maps of seven cosmids covering the transfected *Ecogpt* region of 43IIMM-C. The pSV_3gptH copies, designated I–III, are indicated by open bars. Lower part, detailed physical map of cosmid 43-34. The pSV_3gptH copy is indicated as an open bar (the hatched region is the *Ecogpt* probe used). The minimal region covered by the repair gene *ERCC1* is drawn as a solid bar. Probes used for hybridization of 43-3B transformants are indicated separately. Restriction enzymes include: *Bam*HI (B), *Bgl*II (Bg), *Cla*I (C), *Hin*dIII (H), *Kpn*I (K), *Sal*I (S), *Sst*I (Ss) and *Xho*I (X).

Methods: Restriction enzyme maps are based on single, double and triple endonuclease digestion, on hybridization data with the *Ecogpt* and human *cot-1* probes and individual fragments isolated from low-melting point agarose gels[22], and on subclones from insert regions. To obtain information on the location of the repair gene, 10 μg of cosmid 43-34 DNA were cleaved with various restriction endonucleases (see text) and co-transfected with pSV_3gptH to 43-3B cells. To prevent reconstitution of the gene by religation of the fragments in the cell after transfection, the restricted cosmid DNA was treated before transfection with alkaline phosphatase and, when appropriate, a molar excess of human DNA fragments with the same (cohesive) ends was added as carrier DNA. One μg of digested cosmid DNA was applied to each Petri dish, containing 0.4–1.0×10^6 43-3B cells. Selection was based on MPA and mitomycin C-resistance simultaneously.

*Bgl*II, *Eco*RI, *Hin*dIII, *Kpn*I, *Nun*II, *Pst*I, *Pvu*II, *Sma*I, *Sst*I, *Xho*I) largely or completely abolished MM-C resistance. From these data we conclude that at least sites X_1, K_3, Bg_7, H_6 and B_2 are within essential parts of the DNA repair gene, leaving a minimal gene size of 7 kbp (Fig. 4, lower part).

Consistent with this conclusion is the finding that the largely unique 1.0-kbp B_3–B_4 fragment (Fig. 4, lower part) was present in all of the 13 independent primary transformants checked, and some of them had retained the more distant 1.1-kbp S_2–S_3 fragment also. As expected, none of these transformants contained two unique *Hin*dIII fragments (H_1–H_2 and H_3–H_4; Fig. 4, lower part) from the other side of pSV_3gptH, that is, from a region unrelated to the repair gene (data not shown). Following the nomenclature proposed at the 17th Human Gene Mapping Conference[15], we term the repair gene described here *ERCC1* (excision repair complementing defective repair in Chinese hamster cells).

Other groups have reported the transfer and identification of human repair genes. In one case the transfer of human DNA conferring UV resistance to xeroderma pigmentosum cells of complementation group A has been claimed[16] but this was not followed by molecular identification or cloning of the responsible gene. Extensive efforts to reproduce these results were unsuccessful in our hands (unpublished observations). Recently, a human repair gene was identified[8] following gene transfer into a Chinese hamster mutant of the same complementation group as mutant 43-3B used in this study. The molecular cloning of the repair gene reported here makes possible an investigation of its function in repair, its relation to human repair-deficiency syndromes such as XP and general characteristics of mammalian repair processes.

We thank Drs G. C. Grosveld, A. de Klein, N. G. J. Jaspers and M. James for stimulating discussions and helpful suggestions; J. A. Bruyn, H. Heyboer and R. Kanaar for help in the experiments; and Mrs R. J. Boucke for typing the manuscript and Mr T. van Os for photography. This work was supported by Euratom contract no. BIO-E-404-NL and Fungo, Foundation of Medical Scientific Research in the Netherlands, contract 13–23–35.

Table 1 UV-induced unscheduled DNA synthesis of 43-3B transformants

Cell line	No. of grains
HeLa	37.0±0.9
CHO	21.6±0.8
43-3B	4.5±0.2
43IUV*	17.6±1.7
43IMM-C	25.0±0.7
43IIMM-C	21.9±0.5
43IIIMM-C	31.8±0.6
43Cos34	31.5±1.2

Two days after seeding the cells were exposed to UV light (16 J m^{-2}) and incubated for 2 h in F10 medium without TdR containing ^{3}H-thymidine (10 μCi ml^{-1}; specific activity 46 Ci $mmol^{-1}$) and 8% dialysed fetal calf serum. After fixation of the cells the preparations were processed for autoradiography (Kodak AR10 stripping film), exposed for 1 week at 4 °C, developed, fixed and stained with Giemsa solution. The table shows the results of one typical experiment. For each cell line at least 50 nuclei were counted. Number of grains is given as mean ±s.e.m. per fixed square of non-S-phase nucleus.

* Note that about 20% of the 43IUV population did exhibit 43-3B unscheduled DNA synthesis levels; this might be due to segregation of the repair marker.

Received 2 April; accepted 15 June 1984.

1. Kraemer, K. H. in *Update: Dermatology in General Medicine* (eds Fitzpatrick, T. B., Eisen, A. Z., Wolff, K., Freedberg, I. M. & Austen, K. F.) 113–141 (McGraw-Hill, New York, 1983).
2. Lehmann, A. R. & Karran, P. *Int. Rev. Cytol.* **72**, 101–146 (1981).
3. Thompson, L. H., Busch, D. B., Brookman, K., Mooney, C. L. & Glaser, D. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3734–3737 (1981).
4. Shiomi, T., Hieda-Shiomi, N. & Sato, K. *Somat. Cell Genet.* **8**, 329–345 (1980).
5. de Weerd-Kastelein, E. A., Keijzer, W. & Bootsma, D. *Nature new Biol.* **238**, 80–83 (1972).
6. Moshell, A. N. *et al.* in *Cellular Responses to DNA Damage* (eds Friedberg, E. C. & Bridges, B.) 209–213 (Liss, New York, 1983).
7. Wood, R. D. & Burki, H. J. *Mutat. Res.* **95**, 505–514 (1982).
8. Rubin, J. S., Joyner, A. L., Bernstein, A. & Whitmore, G. F. *Nature* **306**, 206–208 (1983).
9. Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072–2076 (1981).
10. Wigler, M. *et al. Cell* **16**, 777–785 (1979).
11. Lowy, I. *et al. Cell* **22**, 817–823 (1980).
12. Pellicer, A. *et al. Science* **209**, 1414–1422 (1980).
13. Scangos, G. & Ruddle, F. H. *Gene* **14**, 1–10 (1981).
14. Hori, T.-A., Shiomi, T. & Sato, K. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5655–5659 (1983).
15. Shows, T. B., McAlpine, P. J. & Miller, R. J. in *Proc. 7th Human Gene Mapping Conference* (eds Sparkes, R. S., Berg, K., Evans, H. J. & Klinger, H. P.) 340–390 (Karger, Basel, 1984).
16. Takano, T., Noda, M. & Tamura, T. *Nature* **296**, 269–270 (1982).
17. Jeffreys, A. J. & Flavell, R. A. *Cell* **12**, 429–439 (1977).
18. Mulligan, R. E. & Berg, P. *Molec. cell. Biol.* **1**, 449–459 (1981).
19. Graham, F. & Van der Eb, A. *Virology* **52**, 456–467 (1973).
20. Grosveld, F. G., Dahl, H. H. M., de Boer, E. & Flavell, R. A. *Gene* **13**, 227–237 (1981).
21. Grosveld, F. G. *et al. Nucleic Acids Res.* **10**, 6715–6732 (1982).
22. Groffen, J. *et al. Cell* **36**, 93–99 (1984).
23. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
24. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).

APPENDIX PAPER II

Cell 44, 913-923 (1986)

Molecular Characterization of the Human Excision Repair Gene *ERCC-1*: cDNA Cloning and Amino Acid Homology with the Yeast DNA Repair Gene *RAD10*

Marcel van Duin, Jan de Wit, Hanny Odijk, Andries Westerveld, Akira Yasui, Marcel H. M. Koken, Jan H. J. Hoeijmakers, and Dirk Bootsma
Department of Cell Biology and Genetics
Erasmus University Rotterdam
P. O. Box 1738
3000DR Rotterdam, The Netherlands

Summary

The human excision repair gene *ERCC-1* was cloned after DNA mediated gene transfer to the CHO mutant 43-3B, which is sensitive to ultraviolet light and mitomycin-C. We describe the cloning and sequence analysis of the *ERCC-1* cDNA and partial characterization of the gene. *ERCC-1* has a size of 15 kb and is located on human chromosome 19. The *ERCC-1* precursor RNA is subject to alternative splicing of an internal 72 bp coding exon. Only the cDNA of the larger 1.1 kb transcript, encoding a protein of 297 amino acids, was able to confer resistance to ultraviolet light and mitomycin-C on 43-3B cells. Significant amino acid sequence homology was found between the *ERCC-1* gene product and the yeast excision repair protein *RAD10*. The most homologous region displayed structural homology with DNA binding domains of various polypeptides.

Introduction

Nucleotide excision repair, which removes DNA lesions like pyrimidine dimers induced by ultraviolet (UV) light and bulky chemical adducts, is the major DNA repair pathway in mammalian cells (Friedberg, 1985). In the hereditary disease xeroderma pigmentosum (XP) defective DNA excision repair is believed to underlie the extreme sensitivity of patients to sunlight and their predisposition to develop tumors on exposed parts of the skin (for review see Kraemer, 1983). XP displays a considerable genetic heterogeneity; cell fusion experiments have demonstrated the presence of at least nine complementation groups (de Weerd-Kastelein et al., 1972; Fischer et al., 1985). The genes or gene products that are mutated in this cancer prone disorder are unknown. As an approach to the elucidation of these mutations and the understanding of mammalian DNA repair, a number of Chinese hamster ovary (CHO) cell lines that are sensitive to UV light have been isolated (Wood and Burki, 1982; Thompson et al., 1981; Thompson and Carrano, 1983). Genetic complementation revealed that these mutants constitute at least five different complementation groups (Thompson et al., 1981; Thompson and Carrano, 1983), which are, like XP, all defective in the incision step of the excision repair pathway (Thompson et al., 1982).

With the aid of DNA mediated gene transfer, we recently cloned a human excision repair gene designated *ERCC-1* (Westerveld et al., 1984). This gene was cloned by virtue of its ability to correct the excision repair defect in CHO mutant 43-3B, which belongs to complementation group two in the classification of CHO mutants sensitive to UV light (Wood and Burki, 1982) and is also sensitive to mitomycin-C (MM-C). To isolate this gene, human genomic DNA was partially digested with PstI and ligated to the dominant marker pSV3gptH. In a primary transfection of this DNA to 43-3B cells, transformants resistant to mycophenolic acid, which selects for the presence of pSV3gptH, and to UV light or MM-C were isolated. Using genomic DNA of these primary transformants in a secondary transfection, linked transfer of pSV3gptH and the correcting human gene to 43-3B cells could be achieved. This made it possible to isolate *ERCC-1* from a cosmid library of a secondary transformant using pSV3gptH probes (Westerveld et al., 1984). The extensive use of XP cells in this approach did not result in the generation of repair proficient transformants (for review see Lehmann, 1985). In contrast, a number of other successful transfections of CHO mutants using human genomic DNA have been reported (Rubin et al., 1983; MacInnes et al., 1984; Thompson et al., 1985a).

We report the cDNA cloning and partial genomic characterization of *ERCC-1*. Significant homology was found at the amino acid level between *ERCC-1* and the Saccharomyces cerevisiae excision repair gene *RAD10*, suggesting the evolutionary conservation of DNA excision repair. Part of the homologous region has structural homology with DNA binding domains of other polypeptides.

Results

Localization of *ERCC-1* on Cos43-34

The cloning strategy of *ERCC-1* involved the screening of a cosmid library of a repair proficient secondary 43-3B transformant (Westerveld et al., 1984). Of the seven overlapping cosmids isolated from this library, one (cos43-34) contained a functional *ERCC-1* gene. The overlap of cos43-34 with the six other cosmids concerned the left-hand region of the insert, indicating that relevant *ERCC-1* sequences are on the right-hand part (Westerveld et al., 1984). To narrow down the location of *ERCC-1*, cos43-34 DNA was partially digested with Sau3A and size fractionated fragments of 15–20 kb were cloned in a λEMBL-3 replacement vector. A suitable set of λ-recombinants covering the putative *ERCC-1* region was selected for transfection to 43-3B cells in order to screen for a functional *ERCC-1* gene. The results of these experiments are shown in Figure 1. Two recombinant clones extending the farthest to the left did not give repair proficient 43-3B transformants. All other phages were positive in this assay. These results limit the position of *ERCC-1* on cos43-34 to the 15–17 kb region depicted in Figure 1. The higher number of MM-C resistant transformants generated by

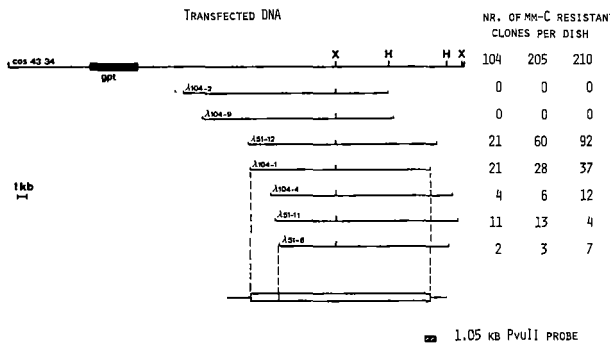

Figure 1. Localization of *ERCC-1* on Cos43-34
Cosmid DNA was partially digested with Sau3A. Fragments of 15–20 kb were cloned in λEMBL3. A suitable set of recombinants was selected for transfection to 43-3B cells. For each of three petri dishes from one transfection experiment the number of MM-C resistant colonies is given. A 1.05 kb genomic PvuII probe was found to be free from repeats and was used for the isolation of cDNA clones.

λ51-12 and λ104-1 than by λ104-4, λ51-11, and λ51-6 might be due to the fact that the latter three lack nonessential portions of *ERCC-1*.

Cloning of *ERCC-1* cDNA

The isolation of unique probes from the genomic *ERCC-1* region was hampered by the abundance of repetitive elements. However, a 1.05 kb PvuII fragment situated close to the right end of *ERCC-1* (Figure 1) was found to be free of repeats and was used as a probe for screening the human expression cDNA library generously provided by Dr. H. Okayama (Okayama and Berg, 1983). This resulted in a number of hybridizing clones, three of which (pcD3A, pcD3C, and pcD3B7), varying in size from 800 to 1000 bp, will be described in more detail. Northern blot analysis of poly(A)$^+$ RNA of HeLa cells revealed that the cDNA clones hybridized mainly to an mRNA of 1.0–1.1 kb. Faint hybridization with a 3.0 kb RNA species was also observed (Figure 2). Identical results were obtained in a Northern blot analysis of the human chronic myelogenous leukemia cell line K562 and an SV40 transformed human fibroblast line. The simplest interpretation of these results is that the 1.0–1.1 kb mRNA is the mature *ERCC-1* transcript. The 3.0 kb band may represent a precursor RNA species.

The aligned physical maps of the three cDNA clones are shown in Figure 3. Sequence analysis of these clones (shown below) revealed that the largest clone, pcD3B7, lacked 104 bp of an internal cDNA region and that clone pcD3A lacked a stretch of 72 bp. However, by substitution of the internal SmaI fragment of pcD3C in the corresponding sites of pcD3B7 a complete *ERCC-1* cDNA could be constructed. This clone, designated pcDE, combined all sequences present in the three cDNA clones.

Surprisingly, the 72 bp region that is absent in pcD3A appeared to correspond exactly to a single exon at the genome level. Sequence analysis of genomic *ERCC-1* DNA at this position revealed the sequence 5′-caccctttccag-GTGAC. . . .TTGGAgtaaggaatggct-3′, which showed that the 72 bp region (capitals) is flanked by expected splice donor and acceptor sequences (underlined). Since the chance that an artificial deletion coincides precisely with

Figure 2. Northern Blot Analysis of Poly(A)$^+$ RNA from HeLa Cells
Twenty micrograms of poly(A)$^+$ RNA was separated in a 1% agarose gel containing formaldehyde and after transfer to nitrocellulose was hybridized to a ^{32}P-labeled *ERCC-1* cDNA probe from pcD3A.

a single exon is extremely low, this finding rendered it very likely that clone pcD3A was derived from an alternatively spliced *ERCC-1* mRNA lacking this 72 bp exon. To obtain additional evidence for differential processing of the *ERCC-1* transcript, S1 nuclease analysis was performed. A BamHI-PvuII fragment from cDNA clones pcD3A and pcD3B7 labeled with γ-^{32}P-ATP at the 3′ PvuII site was hybridized to human poly(A)$^+$ RNA and subsequently treated with nuclease S1. The results of these experi-

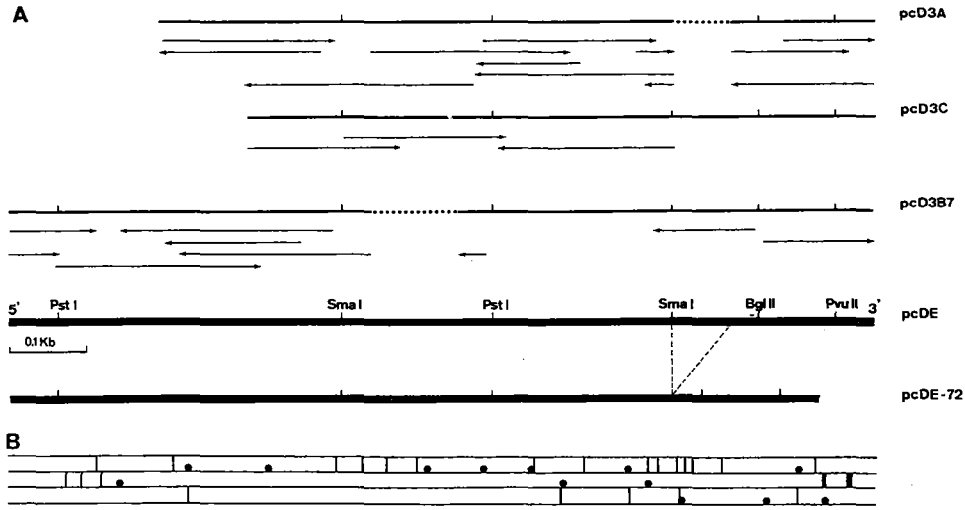

Figure 3. Cloning of the *ERCC-1* cDNA

(A) The screening of the human cDNA expression library with a genomic *ERCC-1* probe yielded three overlapping clones: pcD3A, pcD3C, and pcD3B7. These clones were sequenced according to the indicated strategy (arrows). The dotted parts in the aligned physical maps represent deletions. A complete *ERCC-1* cDNA (pcDE) was constructed by cloning the internal SmaI fragment of pcD3C in the corresponding sites of pcD3B7. Clone pcDE-72 was obtained by ligation of the SmaI-BglII fragment of pcD3A to the corresponding sites of pcDE.

(B) All stop codons and ATGs in the three reading frames of the *ERCC-1* sequence of pcDE (shown in Figure 5) are indicated. Dots: ATG triplets. Vertical bars: stop codons.

ments are shown in Figure 4. After incubation of poly(A)$^+$ RNA from HeLa cells with the pcD3A probe two protected bands of 129 bp and 856 bp were found. These bands can be explained by hybridization of two mRNAs, one completely homologous to the probe and the other differing at a distance of 129 bp from the labeled PvuII site. This position corresponds exactly with the 3' border of the 72 bp deletion found in cDNA clone pcD3A. S1 analysis of poly(A)$^+$ RNA from HeLa and K562 cells with a 3'-labeled probe from pcD3B7 (which includes the 72 bp exon) also yielded a protected band of 129 bp, indicating the presence of *ERCC-1* transcripts without the 72 bases. These data indicate that the *ERCC-1* precursor RNA is subject to alternative splicing. To obtain complete cDNA clones from both transcripts, in addition to clone pcDE, a cDNA clone (pcDE-72) lacking the 72 bp fragment was constructed by replacing the internal SmaI-BglII fragment of pcDE with the corresponding fragment of pcD3A (see Figure 3).

The construction of the cDNA library by the method developed by Okayama and Berg (1983) enables the expression of full length cDNAs in mammalian cells due to the presence of a strong SV40 promoter. This promoter functions optimally in primate cells, but it also displays considerable activity in CHO cells (Simonsen and Levinson, 1983; Scott McIvor et al., 1985; Gorman et al., 1983). In order to investigate the integrity of the cloned *ERCC-1* cDNAs, these cDNAs were transfected to the UV light and MM-C sensitive 43-3B cells. The results of these experiments are summarized in Table 1. In contrast to clone pcD3B7, the reconstructed cDNA clone pcDE conferred resistance to UV light and MM-C after transfection to 43-3B cells. This suggests that the 104 bp deletion in pcD3B7 has inactivated the *ERCC-1* gene and is most likely a cloning artifact. Reconstructed clone pcDE-72 and clone pcD3A did not compensate for the repair defect in 43-3B cells, indicating that the 72 bp region which is absent in the 3' part of these clones is essential for *ERCC-1* functioning in these mutant cells. Surprisingly, clone pcD3C was positive in three independent experiments of this transfection assay. However, sequence analysis (discussed below) revealed that this clone lacks 302 bp of the 5' part including the translational start and 54 N-terminal encoded amino acids. Inspection of the sequence of the pcD3C insert did not show any other potential start codons followed by a long open reading frame (see Figure 3B). Apparently, in 43-3B cells transfected with pcD3C the repair defect is corrected by a truncated *ERCC-1* protein that is translated from an ATG present in the 5' region of the cDNA expression vector (Okayama and Berg, 1983).

Sequence of *ERCC-1* cDNA

Following the strategy depicted in Figure 3A the nucleotide sequence of the *ERCC-1* cDNA clones was determined by the chemical modification method developed by Maxam and Gilbert (1980). The nucleotide sequence and deduced amino acid sequence of the 1097 bp insert of

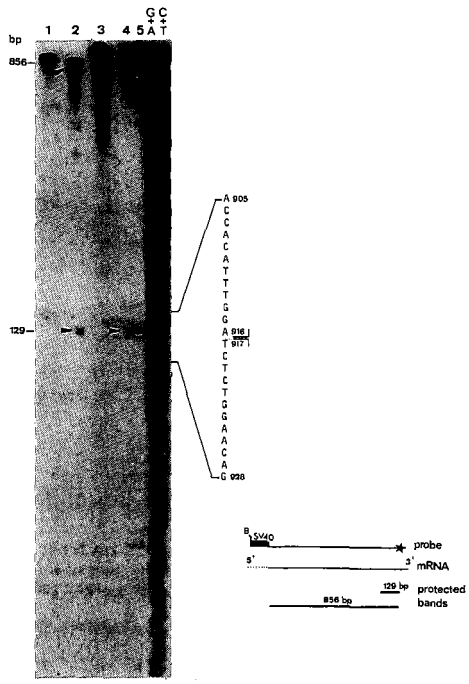

Figure 4. S1 Nuclease Analysis of *ERCC-1* mRNA

The probes used were BamHI-PvuII fragments, labeled at the PvuII sites of pcD3A and pcD3B7, which span from the BamHI site (B) in the SV40 part of the vector to the 3′ PvuII site (asterisk) in the cDNA. After hybridization with RNA and nuclease S1 treatment the samples were separated on a 6% polyacrylamide gel next to a sequence ladder starting from the PvuII site in pcD3B7. The pcD3A probe was incubated with yeast tRNA (lane 1) and poly(A)$^+$ RNA (lane 2) from HeLa cells. The pcD3B7 probe was incubated with yeast tRNA (lane 3) and poly(A)$^+$ RNA from HeLa cells (lane 4) and K562 cells (lane 5). The arrowheads indicate S1 protected bands. The sequence on the right with numbering according to Figure 5 shows the position at which the two *ERCC-1* mRNAs deviate. The diagram shows the protection of two bands of 129 and 856 bp after hybridization of human poly(A)$^+$ RNA with the pcD3A probe.

clone pcDE is shown in Figure 5. The position of all possible start and stop codons is depicted in Figure 3B. The first ATG of the sequence is followed by the longest open reading frame, of 891 bases. A computer search for protein coding regions based on codon preference (Staden and McLachlan, 1982) gave a strong bias in favor of this reading frame. Moreover, since the other reading frames encode much smaller polypeptides (maximum of 60 amino acids) we conclude that the reading frame translated in Figure 5 specifies the *ERCC-1* protein. This corresponds well with the finding that in 95% of all reported cases the 5′ proximal ATG serves as the start codon for translation (Kozak, 1984). The open reading frame is preceded by an untranslated region of 142 bp containing three in-frame termination codons. However, it is worth

Table 1. Complementation of the Sensitivity of 43-3B Cells to UV Light and MM-C by Transfection with *ERCC-1* cDNA

cDNA Clone	Correction
pcD3A	–
pcD3C	+
pcD3B7	–
pcDE	+
pcDE-72	–

noting that the purine residue found at the −3 position of most eukaryotic ATG start codons (Kozak, 1984) is not present in front of the *ERCC-1* translational start codon. The genomic DNA sequence at this point confirmed the cDNA sequence (not shown), ruling out the possibility that this deviation is due to cDNA cloning artifacts. However, the G residue frequently found at position +4 (Kozak, 1984) is present in *ERCC-1*. In conclusion, the *ERCC-1* cDNA clone pcDE encodes a protein of 297 amino acids with a calculated molecular weight of 32,562. The cDNA clone lacking the alternatively spliced 72 bp region might encode a protein of 29,993 daltons. It is remarkable that the region of 24 amino acids lacking in this putative protein is exceptionally rich in threonine residues (about 30%).

Inspection of the 3′ region revealed the sequence of the common polyadenylation signal AATAAA (Wickens and Stephenson, 1984) varying in the different cDNA clones from 19 to 21 bases upstream of the poly(A) tail. However, the pentanucleotide sequence CAYTG, which is found adjacent to the polyadenylation site in many eukaryotic mRNAs (Berget, 1984), is not found in the 3′ part of the *ERCC-1* sequence.

ERCC-1 and Yeast *RAD10* Protein Show Significant Homology

To determine whether *ERCC-1* is partially homologous to other prokaryotic and yeast DNA repair genes a computer analysis was performed using the DIAGON software developed by Staden (1982). No significant homologies were found with published sequences of the genes encoding the E. coli *uvrC*, *Phr*, *Alk A*, bacteriophage T4 *den V*, and yeast *RAD1*, *RAD3*, *RAD6*, and *RAD52* proteins. However, at the protein level significant similarity was found between *ERCC-1* and the recently cloned yeast *RAD10* DNA repair gene (Weiss and Friedberg, 1985; Prakash et al., 1985; Reynolds et al., 1985). The *RAD10* gene encodes a protein of 210 amino acids (Reynolds et al., 1985), which is approximately 90 amino acids smaller than *ERCC-1*. A comparison of the C-terminal half of *RAD10* with the middle part of *ERCC-1* is shown in Figures 6A and 6B. Over a region of approximately 110 amino acids 35% homology exists. At the center of this region a stretch of 25 amino acids shows 56% homology. If the amino acids are classified into four groups (Schwartz and Dayhoff, 1978), the group homology in the 110 amino acid region is 63%.

Since *ERCC-1* is involved in DNA repair, we have investigated whether DNA binding properties of the protein could be deduced from the amino acid sequence. Based

```
                                                                    5'  AAGTGCTGCGAGCCCTGGGCCAC

GCTGGCCGTGCTGGCAGTGGGCCGCCTCGATCCCTCTGCAGTCTTTCCCTTGAGGCTCCAAGACCAGCAGGTGAGGCCTCGCGGCGCTGAAACCGTGAGGCCCGGACCACAGGCTCCAG
                        50                                                                  100

  1
Met Asp Pro Gly Lys Asp Lys Glu Gly Val Pro Gln Pro Ser Gly Pro Pro Ala Arg Lys Lys Phe Val Ile Pro Leu Asp Glu Asp Glu
ATG GAC CCT GGG AAG GAC AAA GAG GGG GTG CCC CAG CCC TCA GGG CCG CCA GCA AGG AAG AAA TTT GTG ATA CCC CTC GAC GAG GAT GAG
       150                                                                  200
 31
Val Pro Pro Gly Val Ala Lys Pro Leu Phe Arg Ser Thr Gln Ser Leu Pro Thr Val Asp Thr Ser Ala Gln Ala Ala Pro Gln Thr Tyr
GTC CCT CCT GGA GTG GCC AAG CCC TTA TTC CGA TCT ACA CAG AGC CTT CCC ACT GTG GAC ACC TCG GCC CAG GCG GCC CCT CAG ACC TAC
                  250                                                                 300
 61
Ala Glu Tyr Ala Ile Ser Gln Pro Leu Glu Gly Ala Gly Ala Thr Cys Pro Thr Gly Ser Glu Pro Leu Ala Gly Glu Thr Pro Asn Gln
GCC GAA TAT GCC ATC TCA CAG CCT CTG GAA GGG GCT GGG GCC ACG TGC CCC ACA GGG TCA GAG CCC CTG GCA GGA GAG ACG CCC AAC CAG
                                  350                                                 400
 91
Ala Leu Lys Pro Gly Ala Lys Ser Asn Ser Ile Ile Val Ser Pro Arg Gln Arg Gly Asn Pro Val Leu Lys Phe Val Arg Asn Val Pro
GCC CTG AAA CCC GGG GCA AAA TCC AAC AGC ATC ATT GTG AGC CCT CGG CAG AGG GGC AAT CCC GTA CTG AAG TTC GTG CGC AAC GTG CCC
                                        450                                           500
121
Trp Glu Phe Gly Asp Val Ile Pro Asp Tyr Val Leu Gly Gln Ser Thr Cys Ala Leu Phe Leu Ser Leu Arg Tyr His Asn Leu His Pro
TGG GAA TTT GGC GAC GTA ATT CCC GAC TAT GTG CTG GGC CAG AGC ACC TGT GCC CTG TTC CTC AGC CTC CGC TAC CAC AAC CTG CAC CCA
                                              550
151
Asp Tyr Ile His Gly Arg Leu Gln Ser Leu Gly Lys Asn Phe Ala Leu Arg Val Leu Leu Val Gln Val Asp Val Lys Asp Pro Gln Gln
GAC TAC ATC CAT GGG CGG CTG CAG AGC CTG GGG AAG AAC TTC GCC TTG CGG GTC CTG CTT GTC CAG GTG GAT GTG AAA GAT CCC CAG CAG
       600                                                                  650
181
Ala Leu Lys Glu Leu Ala Lys Met Cys Ile Leu Ala Asp Cys Thr Leu Ile Leu Ala Trp Ser Pro Glu Glu Ala Gly Arg Tyr Leu Glu
GCC CTC AAG GAG CTG GCT AAG ATG TGT ATC CTG GCC GAC TGC ACA TTG ATC CTC GCC TGG AGC CCC GAG GAA GCT GGG CGG TAC CTG GAG
                  700                                                                 750
211
Thr Tyr Lys Ala Tyr Glu Gln Lys Pro Ala Asp Leu Leu Met Glu Lys Leu Glu Gln Asp Phe Val Ser Arg Val Thr Glu Cys Leu Thr
ACC TAC AAG GCC TAT GAG CAG AAA CCA GCG GAC CTC CTG ATG GAG AAG CTA GAG CAG GAC TTC GTC TCC CGG GTG ACT GAA TGT CTG ACC
                                  800                                                 850
241
Thr Val Lys Ser Val Asn Lys Thr Asp Ser Gln Thr Leu Leu Thr Thr Phe Gly Ser Leu Glu Gln Leu Ile Ala Ala Ser Arg Glu Asp
ACC GTG AAG TCA GTC AAC AAA ACG GAC AGT CAG ACC CTC CTG ACC ACA TTT GGA TCT CTG GAA CAG CTC ATC GCC GCA TCA AGA GAA GAT
                                        900                                           950
271
Leu Ala Leu Cys Pro Gly Leu Gly Pro Gln Lys Ala Arg Arg Leu Phe Asp Val Leu His Glu Pro Phe Leu Lys Val Pro  *  *
CTG GCC TTA TGC CCA GGC CTG GGC CCT CAG AAA GCC CGG AGG CTG TTT GAT GTC CTG CAC GAG CCC TTC TTG AAA GTA CCC TGATGACCCCAG
                                              1000

CTGCCAAGGAAACCCCCAGTGTAATAATAAATCGTCCTCCCAGGCCAGGCTCA  3'
   1050
```

Figure 5. Composite Nucleotide and Encoded Amino Acid Sequence of Human *ERCC-1* cDNA

Amino acids are numbered on the left and nucleotides below each line. The sequence of 1097 bp is derived from the *ERCC-1* cDNA clones as shown in Figure 3 and represents the insert of pcDE. The alternatively spliced exon and polyadenylation signal are underlined. The asterisks mark the translational stop codons.

on amino acid homology with a number of well characterized prokaryotic DNA binding proteins, DNA binding properties have been proposed for the homeo-box proteins of various eukaryotes and yeast mating type regulatory (*MAT*) proteins (Laughon and Scott, 1984; Shepherd et al., 1984). In case of the yeast *MATα2* gene product, DNA binding capacity has been demonstrated (Johnson and Herskowitz, 1985). A comparison of *ERCC-1* and *RAD10* with a number of these protein domains is shown in Figure 6C. Specific amino acids at positions 5, 8–10, and 15 are considered to be important for the formation of two adjacent α-helical structures (α2 and α3) characteristic of DNA binding domains (Pabo and Sauer, 1984). The residues at positions 11–13, 16, 17, and 20 (Wharton and Ptashne, 1985), and possibly those at 14 and 19 (Laughon and Scott, 1984), are thought to play a role in determining the DNA sequence specificity of DNA–protein interaction. It appeared that both *RAD10* and *ERCC-1* contained identical or related amino acids at the positions, which are crucial for the configuration of the two helices. This region coincides with the 56% homologous part of *ERCC-1* and *RAD10*. Moreover, similarities between the compared protein domains were also found in the flanking positions (e.g. positions 22–24). From these observations it is tempting to speculate that the most conserved region of *ERCC-1* and *RAD10* comprises a DNA binding domain which might be essential for its DNA repair function.

Chromosomal Localization of *ERCC-1*

Using ^{32}P-labeled *ERCC-1* cDNA probes the DNA from a panel of 45 human × rodent hybrids was screened for human *ERCC-1* sequences. A representative Southern blot analysis is shown in Figure 7. The overall results of this screening are presented in Table 2. The highest correlation is found with chromosome 19. No hybrids were found in the important category in which chromosome 19 is present and *ERCC-1* is absent (+/− column, Table 2), in contrast to the findings with all other chromosomes. However, three hybrids were found in which no chromosome 19 could be detected cytologically, but in which *ERCC-1* by

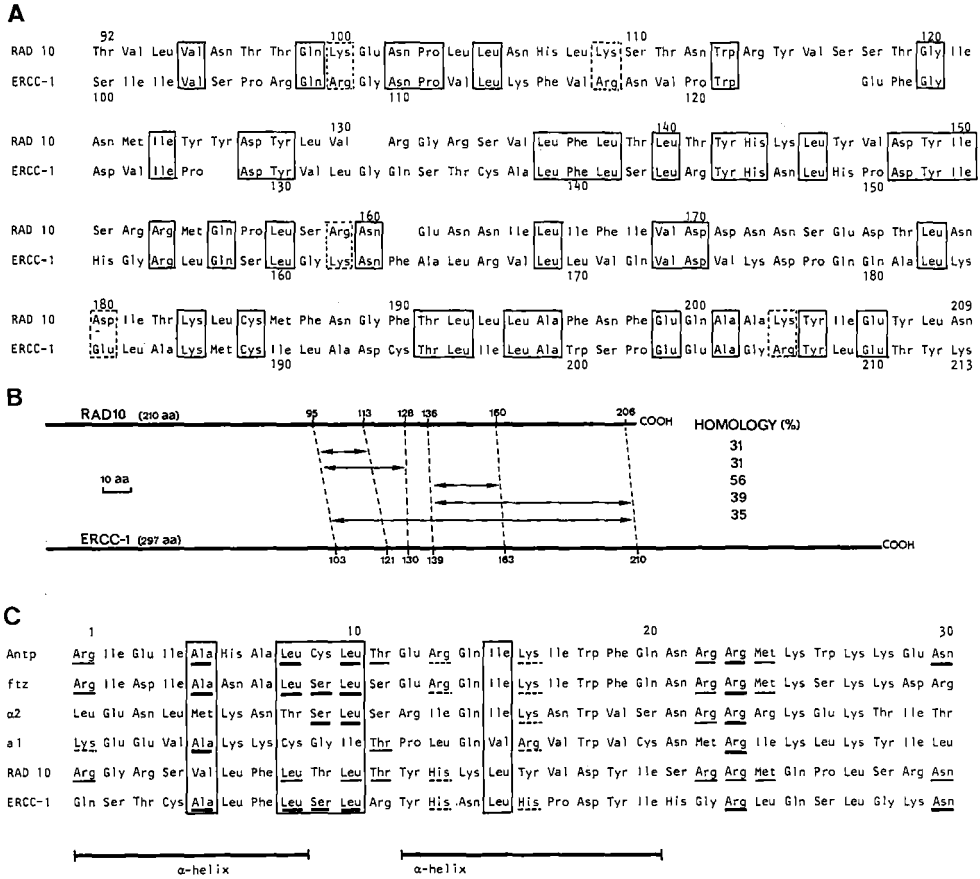

Figure 6. Comparison of *ERCC-1* with Other Proteins

(A) *ERCC-1* amino acids 100–213 aligned with the yeast *RAD10* C-terminal part starting from amino acid 92 to 209. The *RAD10* sequence is from Reynolds et al. (1985). The earlier published *RAD10* nucleotide and deduced amino acid sequence of Weiss and Friedberg (1985) deviates in the 3′ region (from codon 170) from that of Reynolds et al. (1985). Only with the latter sequence, a substantial homology is found in this region. There is now agreement that the sequence determined by Reynolds et al. (1985) is correct (Friedberg, personal communication). Closed boxes and open boxes indicate homologous and related basic or acidic amino acid residues respectively.

(B) Schematic diagram showing *ERCC-1* and *RAD10* proteins. The table on the right gives the percentage of homology of the regions indicated by arrows between the dashed lines. The numbering corresponds to the amino acid sequence of both proteins.

(C) Amino acid comparison of *ERCC-1* and *RAD10* and DNA binding domains in Drosophila *fushi tarazu* (ftz) and *Antennapedia* (Antp) homeo box proteins and yeast mating type regulatory proteins a1 and α2 as compiled by Shepherd et al. (1984). The positions of the two α-helices of the DNA binding domain are shown at the bottom. The amino acids at positions 5, 8–10, and 15, which are important for the formation of the α-helical structures, are boxed. Amino acids homologous to *ERCC-1* are underlined by a solid bar. In addition, thin lines refer to amino acids homologous to *RAD10*. Dashed lines indicate related basic and acidic amino acid residues.

Southern blot hybridization was shown to be present. When these exceptional hybrids were analyzed by enzyme electrophoresis, it appeared that the chromosome 19 marker glucose phosphate isomerase (EC 5.3.1.9) was present, indicating that these hybrids had retained cytologically undetectable fragments of chromosome 19. From these data we conclude that *ERCC-1* is on human chromosome 19. Additional screening of a number of hybrids with translocations in chromosome 19 (Worwood et al., 1985) revealed that *ERCC-1* is most likely located on 19q13.2-q13.3 (Brook et al., 1985).

Discussion

We have presented the molecular characterization of the human excision repair gene *ERCC-1*. Transfection of mu-

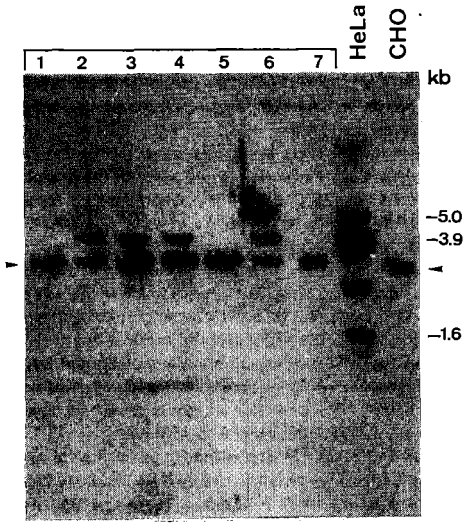

Figure 7. Southern Blot Analysis of PstI Digested DNA (15 μg) from Seven Human × Chinese Hamster Somatic Cell Hybrids and Control HeLa and CHO Cells with a ^{32}P-Labeled *ERCC-1* cDNA Probe

The arrow indicates hybridization with the Chinese hamster *ERCC-1* gene. Hybrids 2, 3, 4, and 6 retained the human *ERCC-1* sequences whereas the others did not.

tant 43-3B cells with *ERCC-1* complements the repair defect in these cells as measured by survival of exposure to UV light and MM-C, unscheduled DNA synthesis (Westerveld et al., 1984), and dimer removal (unpublished results). The 43-3B mutant belongs to complementation group two (Wood and Burki, 1982) in the classification of CHO mutants sensitive to UV light as described by Thompson et al. (1981) and Thompson and Carrano (1983). Recently, the repair defect in CHO mutant UV20, which also falls in complementation group two, was found to be corrected by a human gene on chromosome 19 (Rubin et al., 1985; Thompson et al., 1985b). Two EcoRI fragments of 4.6 and 8.6 kb associated with this repair gene have been cloned by Rubin et al. (1985). Comparison of the physical maps of these fragments with genomic *ERCC-1* pieces indicated that similar EcoRI fragments are present in *ERCC-1* (unpublished data). This finding, together with the identical chromosomal assignment and the fact that UV20 and 43-3B cells belong to the same complementation group, implies that the EcoRI fragments cloned by Rubin et al. (1985) are parts of the *ERCC-1* gene.

The stretch of 72 bp absent in cDNA clone pcD3A appeared to coincide precisely with a single exon at the genome level. Subsequent S1 analysis confirmed the presence of two types of *ERCC-1* mRNAs that are the result of alternative splicing of the 24 amino acid coding exon. Detailed analysis of human genomic DNA by Southern hybridization did not reveal any pseudogene or *ERCC-1*

Table 2. Relationship between the Human *ERCC-1* Gene and Human Chromosomes in 45 Human-Rodent Somatic Cell Hybrid Clones

Chromosome	No. Hybrid Clones with Chromosome/*ERCC-1* Gene Retention[a]				% Concordance
	+/+	+/−	−/+	−/−	
1	14	5	18	8	49
2	5	2	27	11	36
3	17	4	15	9	58
4	10	3	22	10	44
5	17	4	15	9	58
6	19	5	13	8	60
7	15	3	17	10	56
8	18	5	14	8	58
9	13	3	19	10	51
10	14	3	18	10	53
11	20	5	12	8	62
12	21	2	11	11	71
13	13	3	19	10	51
14	16	4	16	9	56
15	13	2	19	11	53
16	18	3	14	10	62
17	13	4	19	9	49
18	13	3	19	10	51
19	29	0	3	13	93←
20	15	3	17	10	56
21	25	4	7	9	76
22[b]	19	5	7	9	66
X	23	5	9	8	69

[a] +/. . and −/. . indicate the presence and absence of the human chromosome. . . /+ and . . /− refer to the presence and absence of human *ERCC-1* sequences as detected by Southern hybridization.

[b] The arrow indicates the chromosome with the highest concordance. Four clones containing chromosome 22 translocations were excluded.

gene family. This rules out the possibility that cDNA clone pcD3A is derived from a transcribed related gene.

Alternative splicing has been found in a number of other gene systems. As a result of this mode of RNA processing single genes are able to meet different functional demands of the cell. In most cases the reported differences between alternatively spliced mRNAs concern the 5′ and 3′ ends of the transcript. In these cases the use of alternative promoters and polyadenylation sites provokes the generation of different mRNAs. The inclusion or exclusion of a separate internal coding exon, as occurs during the processing of *ERCC-1* mRNA, has been reported in a few other cases such as the Drosophila myosin (Rozek and Davidson, 1983), the rodent αA-crystallin (King and Piatigorsky, 1983), and the bovine preprotachikinin (Nawa et al., 1984) genes. However, in the first two cases it is not known whether the alternative splicing serves any function. Only for the preprotachikinin gene a tissue specific splicing of a single exon, yielding two functional mRNAs, has been reported (Nawa et al., 1984).

In the case of *ERCC-1*, our transfection experiments indicate that only the cDNA derived from the larger transcript is able to complement fully the excision defect in 43-3B cells. This rules out the possibility that one of the *ERCC-1* gene products is involved in the repair of lesions such as those caused by UV light and the other in the removal of damages caused by cross-linking agents such as MM-C. The significance of the smaller transcript of *ERCC-1* is still uncertain. First, the possibilities that it has no function and that it is the result of an artifact in the splicing system have not been excluded. Second, it is possible that the smaller gene product is involved in the removal of lesions other than those induced by UV light or MM-C. Finally, the mutation of 43-3B cells might inactivate only the larger gene product. Consequently, one cannot expect complementation with the smaller *ERCC-1* cDNA, although this might be essential for the repair process as well. Further studies are required to discriminate between these possibilities.

Comparison of the *ERCC-1* encoded amino acid sequence with the sequences of isolated DNA repair genes from prokaryotes and yeast revealed striking homology with the S. cerevisiae excision repair gene *RAD10*. The sequence of this gene has recently been determined by Reynolds et al. (1985). The putative *RAD10* and *ERCC-1* proteins consist of 210 and 297 amino acids respectively. The C-terminal half of *RAD10* displays significant homology with the central 110 amino acids of *ERCC-1*. Furthermore, from amino acid sequence homology we have tentatively identified a DNA binding domain in the most homologous part of this region. A general feature of DNA binding protein domains is the presence of an α-helix–turn–α-helix motif, which is involved in the DNA–protein interaction (for review see Pabo and Sauer, 1984). With respect to the amino acids that are important for the configuration of these α-helices, structural homology was found between a number of prokaryotic DNA binding proteins, eukaryotic homeo-box domains, and yeast *MAT* proteins (Laughon and Scott, 1984; Shepherd et al., 1984). A comparison of *ERCC-1* and *RAD10* with the eukaryotic proteins revealed that the most conserved region of *ERCC-1* and *RAD10* shows a structural homology with the putative DNA binding domains of these polypeptides. Concerning *ERCC-1*, the α-helical propensities of the individual amino acids of the putative DNA binding region (calculated according to Finkelstein and Ptitsyn, 1976) are not incompatible with the assumed α-helical structures that comprise the DNA binding domain. However, the presence of the proline at position 17 in the middle of the C-terminal α3 helix of *ERCC-1* (Figure 6) is worth noting. In general, prolines are considered strong helix breakers, introducing kinks in the secondary structure of a protein (Chou and Fasman, 1974). Although proline residues are found in the α3 helix of some of the prokaryotic DNA binding proteins they are located at the N-terminal part (and rarely at the C terminus) of the helix but not in the center (Pabo and Sauer, 1984). This helix is believed to be in direct contact with the major groove of the DNA; therefore it is tempting to speculate that the deformation of the putative α3 helix in *ERCC-1* caused by proline might be related to the structural deformation in the DNA helix caused by DNA lesions induced by UV light or MM-C. However, it is evident that X-ray diffraction and two-dimensional NMR studies on the purified protein–DNA complexes are required to test this hypothesis. In conclusion, we consider the structural homologies between the highly conserved region of *ERCC-1* and *RAD10* and the DNA binding domains of various eukaryotic DNA binding proteins strong enough to suggest that the role of *ERCC-1* and *RAD10* in the removal of DNA damage is mediated through a DNA–protein interaction.

Since the homology covers only a part of the *RAD10* and *ERCC-1* proteins one can question whether the homology only reflects a common DNA binding property or whether both genes are evolutionary related and serve a common DNA repair function. Several observations favor the last option. First, the homology in the DNA binding domains of unrelated peptides predominantly concerns amino acids at fixed positions (Pabo and Sauer, 1984) and spans a region of 20–25 amino acids, which is much smaller than the homologous region of *ERCC-1* and *RAD10*. Second, the positive transfection experiments with the truncated *ERCC-1* cDNA clone pcD3C showed that the absence of 54 N-terminal amino acids does not inactivate the *ERCC-1* gene product in 43-3B cells. If this region of the protein is not essential for its function it is conceivable that it is less subject to evolutionary conservation. Third, the mutant phenotypes of the 43-3B cell line and the yeast rad10 strain have a number of characteristics in common. Similar to 43-3B, the rad10 strain is sensitive to UV light, associated with a defective incision step of the excision repair pathway. In addition, 43-3B and rad10 both are sensitive to 4-nitroquinoline-1-oxide (4NQO) and methylmethanesulphonate and both have enhanced mutagenesis induced by UV light and 4NQO (Prakash, 1976; Zdzienicka and Simons, 1986; Haynes and Kunz, 1981). To confirm the evolutionary relationship between *ERCC-1* and *RAD10* it will be of interest to establish whether rad10 is also sensitive to cross-linking agents like MM-C, and whether the repair defects in rad10 and 43-3B can be corrected by introduction of the human *ERCC-1*

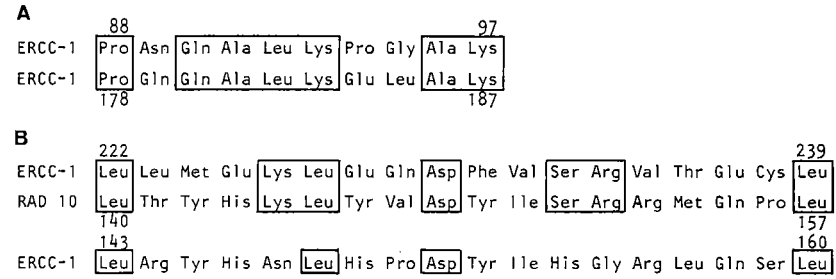

Figure 8. Alignment of Internal *ERCC-1* and *RAD10* Regions

Numbers correspond to amino acid residues in *ERCC-1* as indicated in Figure 5 and *RAD10* as published by Friedberg et al. (1985) and Reynolds et al. (1985). Homologous amino acids are boxed.

and the yeast *RAD10* genes respectively. If such studies support the idea that *ERCC-1* and *RAD10* are derived from the same ancestor gene, it is evident that evolutionary changes have resulted in a large difference in size between both proteins. *ERCC-1* has approximately 90 more C-terminal amino acids than *RAD10*. The alternatively spliced coding exon from *ERCC-1* is located in this region and was essential for *ERCC-1* activity in 43-3B cells. If *ERCC-1* and *RAD10* provide a similar function in DNA repair this would mean either that the functional domains of the two proteins are organized in a different fashion or that additional yeast proteins serve the function of the C-terminal part of the *ERCC-1* gene product.

With respect to the first possibility it is interesting that computer analysis also revealed homology between the C-terminal "extra" part of *ERCC-1* and the region of *RAD10* that already exhibited extensive homology with the middle portion of *ERCC-1*. This concerns a stretch of 18 amino acids (residues 222–229 in *ERCC-1* and 140–157 in *RAD10*) that has a homology of 38% (Figure 8B). The finding that within *ERCC-1* there are two regions which share homology with the same part of *RAD10* suggests the existence of internal homology within the *ERCC-1* protein. Indeed, the indicated 18 amino acids have partial homology with a more central region of *ERCC-1* (residues 143–160, Figure 8B). In addition, a very homologous region of 10 amino acids occurs twice in *ERCC-1* at positions 88–97 and 178–187 (Figure 8A). Taken together, these data provide evidence suggestive of a duplication of part of the *ERCC-1* gene in the course of evolution, which might explain the difference in size between *ERCC-1* and *RAD10*.

Genetic analysis of a large number of yeast and higher eukaryotic mutant cells deficient in excision repair led to the identification of different complementation groups. The interspecies relationship of these complementation groups is unknown. Our finding that *ERCC-1* and *RAD10* are in part homologous is the first indication for evolutionarily conserved DNA repair systems. This emphasizes the importance of yeast and rodent DNA repair deficient mutants in studies which are aimed at the identification of the primary defect in human hereditary diseases associated with defects in DNA repair.

Experimental Procedures

General Procedures

Purification of nucleic acids, restriction enzyme digestion, gel electrophoresis, transfer of DNA and RNA to nitrocellulose, nick translation, and filter hybridization were performed according to established procedures as described by Maniatis et al. (1982).

DNA Transfection

In order to screen for a functional *ERCC-1* gene, DNA from λEMBL3 recombinants containing fragments of cos43-34 or *ERCC-1* cDNA clones were cotransfected with the dominant marker pSV3gptH to CHO 43-3B cells. Usually 1–3 μg of DNA was cotransfected with 2 μg of pSV3gptH DNA. Transfection and selection procedures were essentially as previously described (Westerveld et al., 1984). One day before transfection $\pm 5 \times 10^5$ cells were seeded in 100 mm petri dishes. After 10–14 days of selection the cells were fixed and clones were counted.

Molecular Cloning in λEMBL3

Cos43-34 was partially digested with Sau3A and separated in a low melting agarose gel. Fragments of 15–20 kb were isolated and cloned in the BamHI site of a λEMBL3 replacement vector (Frischauf et al., 1983).

Isolation of cDNA Clones

A cDNA library made from a human SV40 transformed fibroblast was generously provided by Dr. H. Okayama (Okayama and Berg, 1983). In this cDNA expression library the cloned cDNA insert is under the direction of the SV40 early region promoter sequences, allowing a direct use of the cDNA clones in transfection experiments with mammalian cells. The library was screened by colony filter hybridization with a ^{32}P-labeled nick-translated genomic 1.05 kb PvuII probe.

Subcloning and DNA Sequencing

For detailed mapping of restriction sites, parts of cos43-34 were subcloned in pUC-vectors by standard procedures (Maniatis et al., 1982). The nucleotide sequence of the *ERCC-1* cDNA clones was determined following the chemical modification procedure developed by Maxam and Gilbert (1980).

S1 Mapping Procedure

The nuclease S1 protection experiments were carried out according to the procedure described by Grosveld et al. (1981). For the preparation of the probes, the 3′ PvuII sites of cDNA clones pcD3A and pcD3B7 were labeled with polynucleotide kinase and γ-^{32}P-ATP.

Chromosomal Localization

Southern blot analysis was done on DNA from a panel of 45 human × rodent (mouse or Chinese hamster) somatic cell hybrids. Three different Chinese hamster cell lines, designated A3, E36, and CHO,

and two mouse cell lines, Pg19 and Wehi-3B, have been used for the construction of the hybrids according to previously described procedures (de Wit et al., 1979; Geurts van Kessel et al., 1981). Prior to DNA isolation the chromosome content of the hybrids was determined. At least 16 metaphases per hybrid were analyzed.

Acknowledgments

We are very grateful to Dr. H. Okayama for providing the cDNA library, Drs. L. Prakash and S. Prakash for giving permission to use the *RAD10* sequence before publication, Dr. A. H. M. Geurts van Kessel for the hybrid panel used for the chromosomal localization, Drs. D. Shaw and T. Hulsebos for the regional chromosomal mapping, P. ten Dijke and R. Kanaar for the excellent help in the experiments, T. Kartasova for contribution in the cDNA sequencing, and Dr. G. C. Grosveld for helpful suggestions. Furthermore, we thank Mrs. R. Boucke for typing the manuscript and Mr. T. van Os for photography. This work was supported by Euratom contract no. BJ6-141-NL and FUNGO, Foundation of Medical Scientific Research in the Netherlands.

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 U.S.C. Section 1734 solely to indicate this fact.

Received October 7, 1985; revised December 16, 1985.

References

Berget, S. M. (1984). Are U4 small nuclear ribonucleoproteins involved in polyadenylation? Nature *309*, 179–181.

Brook, J. D., Shaw, D. J., Meredith, A. L., Worwood, M., Cowell, J., Scott, J., Knott, T. J., Litt, M., Bufton, L., and Harper, P. S. (1985). A somatic cell hybrid panel for chromosome 19 localization of known genes and RFLP's and orientation of the linkage group. Cytogen. Cell Gen. *40*, 590–591.

Chou, P. Y., and Fasman, G. D. (1974). Prediction of protein formation. Biochemistry *13*, 222–245.

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.

de Wit, J., Hoeksema, H. L., Bootsma, D., and Westerveld, A. (1979). Assignment of structural β-galactosidase loci to human chromosome 3 and 22. Hum. Genet. *51*, 259–267.

Finkelstein, A. V., and Ptitsyn, O. B. (1976). A theory of protein molecule organization. IV. Helical and irregular local structures of unfolded protein chains. J. Mol. Biol. *103*, 15–24.

Fischer, E., Keijzer, W., Thielman, H. W., Popanda, O., Bohnert, E., Edler, E. G., Jung, E. G., and Bootsma, D. (1985). A ninth complementation group in xeroderma pigmentosum, XP-I. Mutation Res. *145*, 217–225.

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

Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. (1983). Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. *170*, 827–842.

Geurts van Kessel, A. H. M., ten Brinke, H., Boere, W. A. M., den Boer, W. C., de Groot, P. G., Hagemeijer, A., Meera Khan, P., and Pearson, P. L. (1981). Characterization of the Philadelphia chromosome by gene mapping. Cytogenet. Cell Genet. *30*, 83–91.

Gorman, C., Padmanablan, R., and Howard, B. H. (1983). High efficiency DNA-mediated transformation of primate cells. Science *221*, 551–553.

Grosveld, G. C., Koster, A., and Flavell, R. A. (1981). A transcription map for the rabbit β-globin gene. Cell *23*, 573–584.

Haynes, R. H., and Kunz, B. A. (1981). DNA repair and mutagenesis in yeast. In The Molecular Biology of the Yeast *Saccharomyces*: Life Cycle and Inheritance, J. Strathern, E. Jones, and J. Broach, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 371–414.

Johnson, A. D., and Herskowitz, I. (1985). A repressor (*MATα2* product) and its operator control expression of a set of cell type specific genes in yeast. Cell *42*, 237–247.

King, C. R., and Piatigorsky, J. (1983). Alternative RNA splicing of the murine αA-crystallin gene: protein-coding information within an intron. Cell *32*, 707–712.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucl. Acids Res. *12*, 857–872.

Kraemer, K. H. (1983). Heritable diseases with increased sensitivity to cellular injury. In Update: Dermatology in General Medicine, T. B. Fitzpatrick, A. Z. Eisen, K. Wolff, I. M. Freedberg, and K. F. Austen, eds. (New York: McGraw-Hill Book Company), pp. 113–141.

Laughon, A., and Scott, M. P. (1984). Sequence of a Drosophila segmentation gene: protein structure homology with DNA-binding proteins. Nature *310*, 25–30.

Lehmann, A. R. (1985). Use of recombinant DNA techniques in cloning DNA repair genes and in the study of mutagenesis in human cells. Mutation Res. *150*, 61–67.

MacInnes, M. A., Bingham, J. D., Thompson, L. H., and Strniste, G. F. (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.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular cloning. A Laboratory Manual. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65*, 499–560.

Nawa, H., Kotani, H., and Nakanishi, S. (1984). Tissue-specific generation of two preprotachikinin mRNAs from one gene by alternative RNA splicing. Nature *312*, 729–734.

Okayama, H., and Berg, P. (1983). A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. Mol. Cell. Biol. *3*, 280–289.

Pabo, C. O., and Sauer, R. T. (1984). Protein–DNA recognition. Ann. Rev. Biochem. *53*, 293–321.

Prakash, L. (1976). Effect of genes controling radiation sensitivity on chemically induced mutations in *Saccharomyces cerevisiae*. Genetics *83*, 285–301.

Prakash, L., Dumais, D., Polakowska, R., Peruzzi, G., and Prakash, S. (1985). Molecular cloning of the *RAD10* gene of *Saccharomyces cerevisiae*. Gene *34*, 55–61.

Reynolds, P., Prakash, L., Dumais, D., Perruzi, G., and Prakash, S. (1985). The nucleotide sequence of the *RAD10* gene of *Saccharomyces cerevisiae*. EMBO J. *4*, 3549–3552.

Rozek, C. E., and Davidson, N. (1983). Drosophila has one myosin heavy-chain gene with three developmentally regulated transcripts. Cell *32*, 23–34.

Rubin, J. S., Joyner, A. L., Bernstein, A., and Whitmore, G. F. (1983). Molecular identification of a human DNA repair gene following DNA-mediated gene transfer. Nature *306*, 206–208.

Rubin, J. S., Prideaux, V. R., Huntington, F. W., Dulhanty, A. M., Whitmore, G. F., and Bernstein, A. (1985). Molecular cloning and chromosomal localization of DNA sequences associated with a human DNA repair gene. Mol. Cell. Biol. *5*, 398–405.

Schwartz, R. M., and Dayhoff, M. D. (1978). Matrices for detecting distant relationships. In Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 3, M. D. Dayhoff, ed. (Washington, D.C.: National Biomedical Research Foundation), pp. 353–358.

Scott McIvor, R., Goddard, J. M., Simonsen, C. C., and Martin, D. W., Jr. (1985). Expression of a cDNA sequence encoding human purine nucleoside phosphorylase in rodent and human cells. Mol. Cell. Biol. *5*, 1349–1357.

Shepherd, J. C. W., McGinnes, W., Carrasco, A. E., De Robertis, E. M., and Gehring, W. J. (1984). Fly and frog homoeo domains show homologies with yeast mating type regulatory proteins. Nature *310*, 70–71.

Simonsen, C. C., and Levinson, A. D. (1983). Isolation and expression of an altered mouse dihydrofolate reductase cDNA. Proc. Natl. Acad. Sci. USA *80*, 2495–2499.

Staden, R. (1982). An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucl. Acids Res. *10*, 2951–2961.

Staden, R., and McLachlan, A. D. (1982). Codon preference and its use in identifying protein coding regions in long DNA sequences. Nucl. Acids Res. *10*, 141–156.

Thompson, L. H., and Carrano, A. V. (1983). Analysis of mammalian cell mutagenesis and DNA repair using in vitro selected CHO cell mutants. In Cellular Responses to DNA Damage, U.C.L.A. Symposium on Molecular and Cellular Biology New Series, Vol. 11, E. C. Friedberg and B. R. Bridges, eds. (New York: Alan R. Liss), pp. 125–143.

Thompson, L. H., Busch, D. B., Brookman, K. W., Mooney, C. L., and Glaser, P. A. (1981). Genetic diversity of UV-sensitive DNA-repair mutants of Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA *78*, 3734–3737.

Thompson, L. H., Brookman, K. W., Dillehay, L. E., Mooney, C. L., and Carrano, A. V. (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.

Thompson, L. H., Brookman, K. W., Minkler, J. L., Fuscoe, J. C., Henning, K. A., and Carrano, A. V. (1985a). DNA-mediated transfer of a human DNA repair gene that controls sister chromatid exchange. Mol. Cell. Biol. *5*, 881–884.

Thompson, L. H., Mooney, C. L., Burkhart-Schultz, K., Carrano, A. V., and Siliciano, M. J. (1985b). Correction of a nucleotide excision repair mutation by human chromosome 19 in hamster-human hybrid cells. Somat. Cell Mol. Genet. *11*, 87–92.

Weiss, W. A., and Friedberg, E. C. (1985). Molecular cloning and characterization of the yeast *RAD10* gene and expression of *RAD10* protein in *E. coli*. EMBO J. *4*, 1575–1582.

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

Wickens, M., and Stephenson, P. (1984). Role of a conserved AAUAAA sequence: four AAUAAA point mutations prevent messenger 3′ end formation. Science *226*, 1045–1051.

Wharton, R. P., and Ptashne, M. (1985). Changing the binding specificity of a repressor by redesigning an α-helix. Nature *316*, 601–605.

Wood, R. D., and Burki, H. J. (1982). Repair capability and the cellular age response for killing and mutation induction after UV. Mutation Res. *95*, 505–514.

Worwood, M., Brook, J. D., Cragg, S. J., Hellkuhl, B., Jones, B. M., Perera, P., Roberts, S. H., and Shaw, D. J. (1985). Assignment of human ferritine genes to chromosomes 11 and 19q13.3-19qter. Hum. Genet. *69*, 371–374.

Zdzienicka, M. Z., and Simons, J. W. I. M. (1986). Analysis of repair processes by the determination of the induction of cell killing and mutations in two repair deficient Chinese hamster ovary cell lines. Mutation Res., in press.

Note Added in Proof

We have recently identified a region in the deduced *ERCC-1* protein (amino acids 12–23) that shows a striking homology with the nuclear location signal of the SV40 large T antigen (Kalderon et al., Cell *39*, 499–509, 1984; Kalderon et al., Nature *311*, 33–38, 1984).

APPENDIX PAPER III

Nucleic Acids Research 15, 9195-9213 (1987)

Volume 15 Number 22 1987 Nucleic Acids Research

Genomic characterization of the human DNA excision repair gene *ERCC-1*

Marcel van Duin, Marcel H.M.Koken, Jacolien van den Tol, Peter ten Dijke, Hanny Odijk, Andries Westerveld, Dirk Bootsma and Jan H.J.Hoeijmakers

Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

Received September 11, 1987; Accepted October 22, 1987

ABSTRACT

In this report the genomic characterization of the human excision repair gene ERCC-1 is presented. The gene consists of 10 exons spread over approximately 15 kb. By means of a transfection assay the ERCC-1 promoter was confined to a region of ± 170 bp upstream of the transcriptional start site. Classical promoter elements like CAAT, TATA and GC-boxes are absent from this region. Furthermore, ERCC-1 transcription is not UV-inducible. A possible explanation is provided for the previously reported alternative splicing of exon VIII. Analysis of ERCC-1 cDNA clones revealed the occurrence of differential polyadenylation which gives ERCC-1 transcripts of 3.4 and 3.8 kb in addition to the major 1.1 kb mRNA. Apparent evolutionary conservation of differential polyadenylation of ERCC-1 transcripts suggests a possible role for this mode of RNA processing in the ERCC-1 repair function.

INTRODUCTION

In order to cope with DNA-lesions induced by mutagenic and carcinogenic agents, living organisms have developed a variety of DNA repair systems. One of the most important and best studied repair processes is the excision repair pathway (see 1 for a review). In Escherichia coli this system constitutes part of the damage inducible SOS-response. Lesions causing a relatively strong deformation of the normal helix structure (such as bulky adducts or ultraviolet light (UV) induced pyrimidine dimers) are recognized and removed by the concerted action of at least 4 excision repair proteins: uvr A,B,C and D (2). In yeast, mutant analysis and gene cloning have demonstrated the existance of more than 10 genetic loci implicated in excision repair (1). In higher eukaryotes cell hybridization experiments have revealed the presence of at least 9 complementation groups in cells from excision deficient xeroderma pigmentosum patients (3,4) and of at

least 5 groups in laboratory induced CHO repair mutant cells (5,6). This extensive genetic heterogeneity suggests a considerable biochemical complexity underlying the excision repair pathway. However, the mechanism of this system in eukaryotes is poorly understood.

We have recently cloned the first human repair gene, designated ERCC-1, that corrects the excision defect in CHO mutants of complementation group 2 (7). The major ERCC-1 mRNA is 1.1 kb and specifies a protein of 32.5 kD. Evidence was obtained for the occurrence of alternative splicing yielding in addition a transcript of 1.0 kb, lacking a 72 bp coding exon (8). However, only the cDNA corresponding to the 1.1 kb mRNA, inserted in a mammalian expression vector was able to correct the repair defect in CHO mutants of group 2 (8). At the amino acid level a significant homology was found with the yeast repair protein RAD10 and parts of the E.coli uvrA and C gene products (8,9,10) indicating strong evolutionary conservation of DNA repair systems.

In this report we present the genomic organization of the human ERCC-1 gene. We have identified the ERCC-1 promoter region and provide evidence that the ERCC-1 mRNA is subject to differential polyadenylation during processing.

MATERIALS AND METHODS

Cell culture and transfections

UV-sensitive Chinese hamster ovary (CHO) cell line 43-3B (11) and HeLa cells were grown in DMEM/F10$^+$ (1:1) medium supplemented with 10% fetal calf serum and antibiotics (penicillin, 100 U/ml and streptomycin 0.1 mg/ml). ERCC-1 DNA constructs were cotransfected with 0.5 - 3.0 µg pSV3gptH to 43-3B cells as described previously (8). After 10-14 days of selection with mitomycin-C (MM-C) and mycophenolic acid resistant clones were fixed and counted.

Identification of intron-exon borders

Using general procedures (12) all BglII fragments of cosmid43-34 that hybridized to ^{32}-labeled ERCC-1 cDNA probes were subcloned in pUC-vectors. Hybridizing parts of these clones were

sequenced by the chemical degradation method of Maxam and Gilbert (13).

Isolation of cDNAs

A human cDNA expression library made of poly(A)$^+$ RNA from a SV40 transformed fibroblast cell line was kindly provided by Dr. H. Okayama and screened with a 1.05 kb PvuII fragment of the 3′ region of ERCC-1 (8).

Plasmid construction of the ERCC-1 promoter

The genomic 3.7 kb BglII fragment Bg3-Bg4 of cosmid43-34 (7) was subcloned in pUC9 yielding p56-4. The ERCC-1 cDNA clone pcDE has been described previously (8). The plasmids pROM-1 and pUCPROMH-1 containing the putative ERCC-1 promoter region linked to the ERCC-1 cDNA and pERCC-SVP were constructed as follows: pcDE was cleaved with HindIII, treated with Klenow enzyme, and further digested with StuI which removes the SV40 promoter region (HindIII-StuI) and an ERCC-1 StuI fragment. The resulting blunt ended vector part of pcDE was ligated to this StuI fragment yielding the promoterless construct pERCC-SVP. The pROM-1 construct was obtained by the consecutive cloning of a 1.3 kb genomic PvuII-StuI fragment of 56-4 harboring the 5′ half of exon 1 at its 3′ end and the ERCC-1 StuI fragment in the HindIII (Klenow)-StuI digested pcDE vector. A 1.3 kb Klenow treated HinfI-BamHI fragment of pROM-1 carrying the complete ERCC-1 cDNA and 170 bp of 5′ genomic DNA was ligated to HindII linearized pUC18 resulting in pUCPROMH-1. Plasmid pROM-2 is similar to pROM-1 but has the ERCC-1 StuI fragment in the anti-sense orientation.

Nuclease S1 analysis

S1-mapping was performed as described by Grosveld et al. (14). A uniformly ^{32}P-labeled single stranded probe was synthesized from M13-templates using the method of Burke (15) and annealed with poly(A)$^+$RNA (10 µg) of K562 cells. After treatment with nuclease S1 the protected fragments were loaded on 6% polyacrylamide sequence gels next to a sequence ladder as molecular weight marker.

Northern blotting

Total RNA was prepared from cell cultures and isolated nuclei as described by Auffray and Rougeon (16). The mRNA

fraction was purified by two passages over oligo dT cellulose. For Northern blot analysis poly(A)$^+$ RNA (± 20 µg) was size fractionated by electrophoresis through 1% agarose/formaldehyde gels, transferred to nitrocellulose, and hybridized to ^{32}P-labeled nick-translated probes.

UV-induction experiments

HeLa cells were grown to near confluency on 15 cm Petri dishes and irradiated with 254 nm UV (0.6 J/m^2). Subsequently the cells were rinsed with PBS and incubated with medium. After several time intervals HeLa cells of 6 Petri dishes were harvested by polishman and poly(A)$^+$ RNA was isolated.

RESULTS

Gene Structure

Cosmid43-34, harboring the human ERCC-1 gene, was isolated from a cosmid library of a repair proficient CHO 43-3B transformant (7). The size of the ERCC-1 gene and its location on cosmid 43-34 (cos43-34) was determined in 3 ways: i. By transfection of restriction enzyme digested cos43-34 DNA to 43-3B cells to test for intactness of the gene (7). ii. By subcloning of Sau3A fragments of cos43-34 in EMBL-3 and screening for recombinant phages carrying a functional gene (8). iii. By Southern blot analysis of independent genomic 43-3B transformants for the presence or absence of specific probes from the ERCC-1 flanking regions (unpublished results). From these results it appeared that ERCC-1 covered a region of 15 kb. A detailed physical map of the deduced ERCC-1 region on cos43-34 is presented in Figure 1A. To determine the structural organization of the gene all BglII fragments in this region were subcloned in pUC-vectors and hybridized with ^{32}P-labeled probes of ERCC-1 cDNA clone pcDE (8). Subsequently, appropriate hybridizing fragments were further subcloned to sequence all exons and intron-exon junctions. The sequence strategy, deduced transcriptional orientation and the genomic organization of ERCC-1 are depicted in Figure 1B. The human ERCC-1 gene appears to consist of 10 exons ranging in size between 60 and 216 bp, spread over a region of 15 kb. The nucleotide sequence of the exons was completely in accordance with the previously reported ERCC-1 cDNA sequence (8). As shown

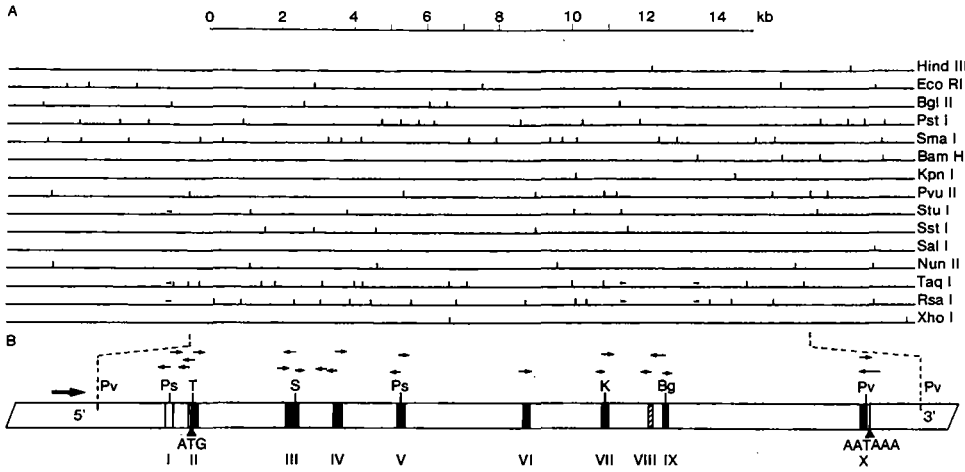

Figure 1. Genomic organization of ERCC-1.
Detailed physical map (A) and intron-exon structure (B) of the human excision repair gene ERCC-1. Arrow heads (in A) define regions for which not all restriction enzyme cutting sites have been mapped. Transcriptional orientation is indicated by the large arrow. Exons (depicted as boxes) are indicated by Roman numbers. The sequence strategy to determine intron-exon junctions is indicated by small arrows. Filled boxes represent coding sequences beginning with the indicated ATG. The differentially spliced exon VIII is shown as a shaded box. Abbreviations: Ps : PstI; T : TaqI; S : SmaI; K : KpnI; Bg : BglII; Pv : PvuII.

in Figure 2, the sequences around the intron-exon borders are consistent with the consensus donor and acceptor splicing signals (17) The much less conserved sequence PyNPyTPuAPy which is involved in branchpoint formation during the splicing process (18,19) could tentatively be identified in most introns at the appropriate distance (20 to 40 nucleotides) proximal to the splice acceptor sites (Figure 2, see also Discussion).

The ERCC-1 promoter region

The nucleotide sequence upstream from the first ERCC-1 exon is shown in Figure 3. A perfect alignment of 142 bp with the 5' terminus of the previously reported cDNA sequence (8) was found. The precise start site of the ERCC-1 mRNA was determined by nuclease S1 analysis as demonstrated in Figure 4. A 2.0 kb genomic fragment spanning from the PstI site in exon I (position 60 in cDNA clone pcDE, (8)) to an upstream BglII site was

INTRON SPLICE ACCEPTOR	EXON				INTRON SPLICE DONOR	
	I	(153)	ACCACAG	34	GT GCGGGAGGCG	(0.3)
TGAGAACGTAGACCCCAGTGG CTCACTCGCTCCTGGCACCTTCCCTTTCAG	GCTCCAG 35	II	(112)	GLY VAL GGA GTG 107	GT AGGACAA	(1.9)
CTCAGATGCCTCGCTCACCCCAACCA GCCCCTGCAAACTCCCTTTTCTCCCCACAG	ALA LYS GCC AAG 108	III	(216)	ARG GLN CGG CAG 142	GT GAGGAGGGAG	(0.7)
AT TGATGGCTTCTGCCCTTCGTCCCTCCCCAG	ARG GLY AGG GGC 143	IV	(104)	LEU SE CTC AG 175	GT GAGCTCTGC	(1.2)
CTTTGGGGGTCTTGAAGAGACTGAGAC CTCTCAACTCCCGCTCCCAGCCCCCACCAG	R LEU C CTC 176	V	(100)	ASP VAL GAT GTG 201	GT AAGCAGGGGG	(2.3)
GCTAATTTTTGTAATTCATGGCTTCT AGGGTTTCTAATTCTGATTTTCTCCTCCAG	LYS ASP AAA GAT 202	VI	(77)	TRP SE TGG AG 234	GT GAGATGAGGG	(1.8)
AATCTGATACCTCACC TCCCGGCTGCCCTGTATCCTGTTATCCCAG	R PRO C CCC 235	VII	(100)	SER ARG TCC CGG 258	GT GAGGCCACC	(0.9)
GCTCAGTCTTGCTGGCCTTTCTTCT CCGACACACTCCTGCCTCCACCCTTTCCAG	VAL THR GTG ACT 259	VIII	(72)	PHE GLY TTT GGA 281	GT AAGGAATGGCT	(0.2)
GGACAGGGGAGGCTTTTGTGCTCAA CTGCCTCGACCCCTCGCTTTCACCTTTCAG	SER LEU TCT CTG 282	IX	(60)	GLN LYS CAG AAA	GT AAGAGCTCT	(4.1)
CCAGAAAACAGCTCCTTTAATGCT GGTTCCTCATTTTTTTCTCCCTCCATCCAG	ALA ARG GCC CGG	X	(112)	3' POLY A TAIL		

Figure 2. Structural details of the ERCC-1 intron-exon organization. The nucleotide sequence of each intron-exon junction is shown. Vertical lines represent intron-exon borders. Splice acceptor and donor sequences are in accordance with reported consensus sequence $(C)_nNCAG/G$ and AAG/GTA respectively (17). The corresponding amino acids at the beginning and end of each exon are shown in the three letter code and numbered as reported previously (8). Exon numbering is as used in Fig. 1. The size of each intron and exon (respectively kb and bp) is given in parentheses. The length of exon I is based on the results presented in Fig. 3 and 4. Heptamers which display the best match with the splicing branchpoint consensus sequence PyNPyTPuAPy are underlined.

subcloned in M13-mp8. Homogeneously ^{32}P-labeled single stranded DNA of this fragment was annealed to poly(A)$^+$ RNA of K562 cells and treated with nuclease S1. Figure 4 shows a protected band of 73-75 bp. We conclude from this result that the transcription most likely initiates at the A residues indicated in Figure 3 since these are generally preferred as 'capping'-site (17). From these results it appears that cDNA clone pcDE is almost complete (lacking only 9-11 bp at the 5' end, Figure 4) and consequently, that exon 1 has a size of 151-153 bp. The ERCC-1 5' flanking sequence is lacking the TATA and CAAT promoter signals. However at position -100 a sequence with a low level of homology with the

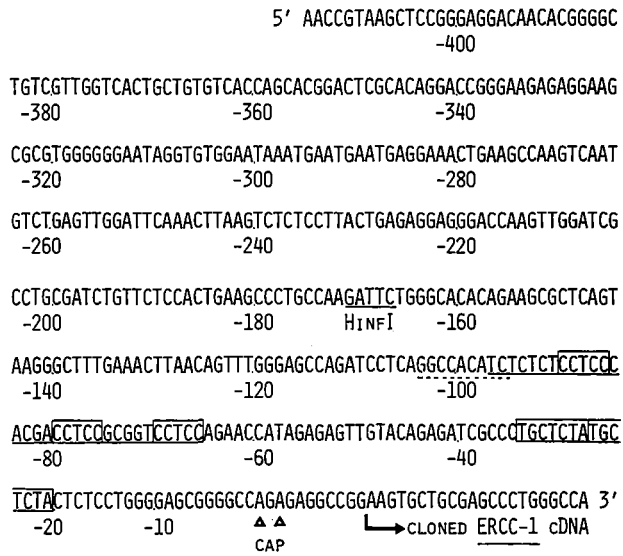

Figure 3. Genomic sequence upstream of ERCC-1 exon I. The 5' border of the cloned ERCC-1 cDNA (8) is indicated. Transcriptional start site(s) (as determined in Fig. 4) are shown by triangles. Tandem sequence TGCTCTA and the 3 times repeated sequence CCTCC are boxed. An 'abbreviated' CAAT box at -100 is indicated by a dashed line. The pyrimidine rich stretch of nucleotides and the HinfI site used for promoter construct pUCPROMHI (Fig. 4) are underlined.

CAAT-box consensus sequence GGTCAATCT can be found (see Figure 3). At position -30 the sequence TCTATGC very remotely resembles the TATA$^{T}_{A}A^{T}_{A}$ consensus. Several other sequence motifs are present in the ERCC-1 promoter region. At position -20 to -30 a tandem repeat of TGCTCTA is located. Further downstream the pentanucleotide CCTCC occurs three times with intervals of 5 nucleotides (see Figure 3). This region (-65 to -100) is very rich in pyrimidine residues (78%). The most upstream CCTCC motif is part of a stretch of 13 pyrimidines.

To verify whether this region can drive ERCC-1 transcription two genomic fragments (a 1.3 kb PvuII-StuI and 180 bp HinfI-StuI fragment) were constructed in front of the ERCC-1 cDNA yielding PROM-1 and pUCPROMHI respectively (see Figure 4 and Materials and Methods). These DNA constructs were transfected to 43-3B cells in coprecipitate with the dominant vector pSV3gptH. Transformants

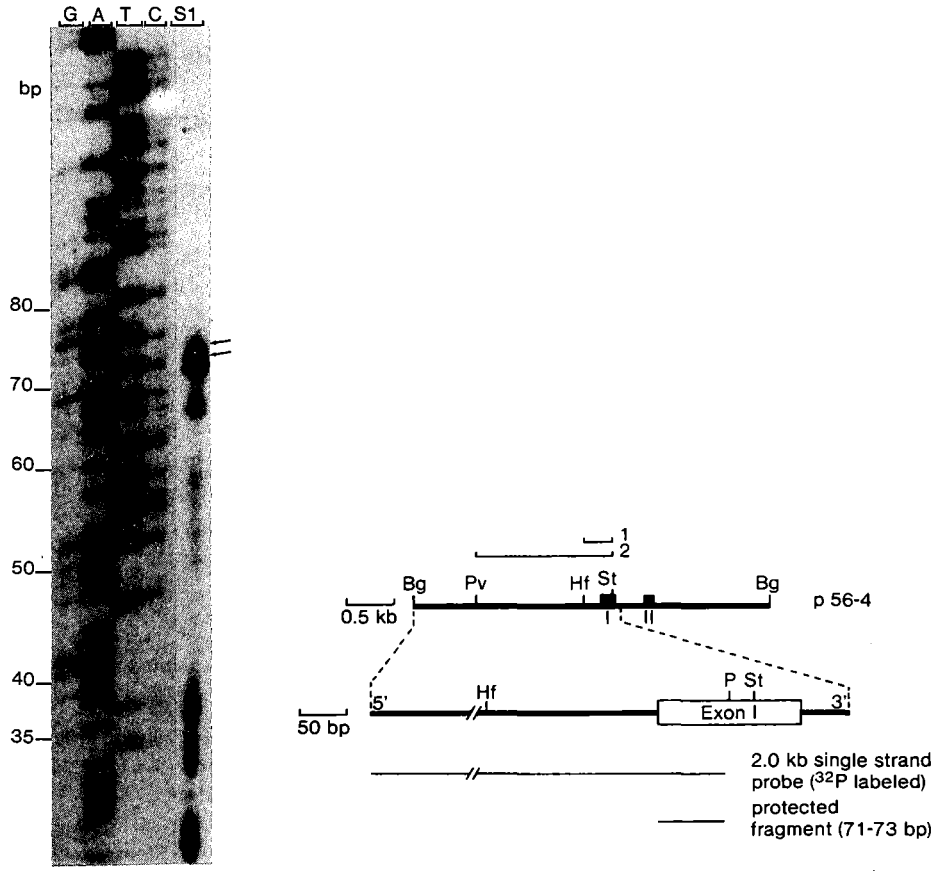

Figure 4. Analysis of ERCC-1 transcriptional start site. A two kb BglII-Pst fragment of genomic ERCC-1 subclone p56-4 (harboring exon I and II) was subcloned in M13-mp8. Single stranded ^{32}P-labeled DNA of this fragment was annealed to K562 poly(A)$^+$RNA and treated with nuclease S1 (4000U, 1 hr, 16^{o}C). Protected fragments (shown by arrows) were visualized by polyacrylamide gelelectrophoresis and autoradiography. A Maxam-Gilbert sequence ladder of an exon I fragment labeled at the StuI site was used as a molecular weight marker. After incubation of the probe with yeast tRNA no protected fragments were observed (not shown). DNA fragments 1 and 2 are genomic PvuII-StuI and HinfI-StuI fragments that were used for constructing pROM-1 and pUCPROMH-1 respectively (see Table 1).

were selected with mycophenolic acid for expression of the E.coli gpt gene and with MM-C for ERCC-1 expression.

The results summarized in Table 1 show that both ERCC-1 plasmids can confer MM-C resistance to 43-3B cells with an efficiency comparable to the genomic ERCC-1 gene on cosmid 43-34. This

Table 1.
Transfected DNAs are described in Materials and Methods.

Transfected DNA		Number of mpa and MM-C resistant clones/µg transfected DNA[a)]
cos43-34	(genomic ERCC-1)	25
pcDE	(cDNA + SV40 early promoter)	5
pROM-1	(cDNA + 1.3 kb promoter fragment)	22
pUCPROMH-1	(cDNA + 170 bp promoter fragment)	32
pROM-2	(inactivated ERCC-1 cDNA)	0
pERCC-SVP	(cDNA without promoter)	2

a) average of three dishes
Identification of the ERCC-1 promoter region. ERCC-1 promoter activity was determined by scoring the number of mitomycin-C and mycophenolic acid (mpa) resistant clones after cotransfection of ERCC-1 DNA and pSV3gptH to 43-3B cells.

indicates that the two genomic regions tested display promoter activity and confines the ERCC-1 promoter to a stretch of 170 bp upstream of the transcriptional start site. Surprisingly, with pcDE, harboring ERCC-1 cDNA under the direction of the strong SV40 early promoter, always a lower transfection frequency is observed. No repair-proficient transformants are obtained with a non-functional ERCC-1 construct (pROM-2). However, a very low but consistent transfection frequency is found with pERCC-SVP which is lacking promoter sequences. This may be explained by insertion of the SV40 early promoter derived from cotransfected pSV3gptH or by integration of pERCC-SVP DNA in the vicinity of endogenous promoters in the genome of 43-3B cells.

Northern blot analysis of different human cell lines revealed a low ERCC-1 transcription level (8, data not shown). To investigate whether the ERCC-1 promoter is DNA-damage inducible, Northern blot analysis was performed on poly(A)$^+$ RNA from Hela cells at various periods of time after UV irradiation. The results are presented in Figure 5. Hybridization with the probe for the glyceraldeyde-3-phosphate dehydrogenase gene (GAPDH; 20) included as control, indicates that approximately equal amounts of poly(A)$^+$ RNA are loaded in each lane. It is evident that UV irradiation with a UV dose of 1 J/m^2 did not result in detectable changes of ERCC-1 transcription within a

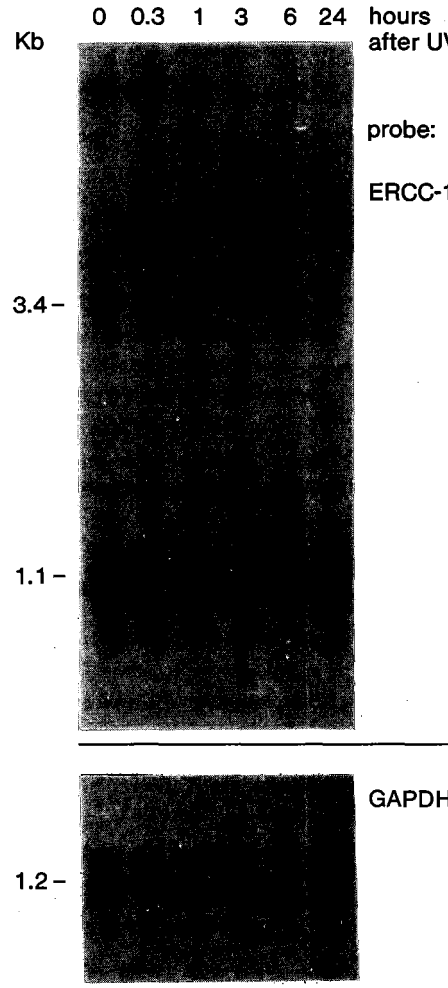

Figure 5. Effect of UV-irradiation on ERCC-1 transcription. Poly(A)$^+$RNA (20 µg) of exponentially growing UV-irradiated (1 J/m^2) HeLa cells was size fractionated on 1% agarose gel and after blotting to nitrocellulose hybridized to a nick-translated ^{32}P-labeled ERCC-1 cDNA probe. To verify that equal amounts of RNA were loaded on the gel the filter was rehybridized with a GAPDH probe (20).

period of 24 hours. Nor did a UV dose of 10 J/m^2 or treatment of HeLa cells with MM-C (10^{-8}M) affect the level of ERCC-1 transcription (data not shown).

The 3´ region

The ERCC-1 cDNA was isolated from a human expression cDNA

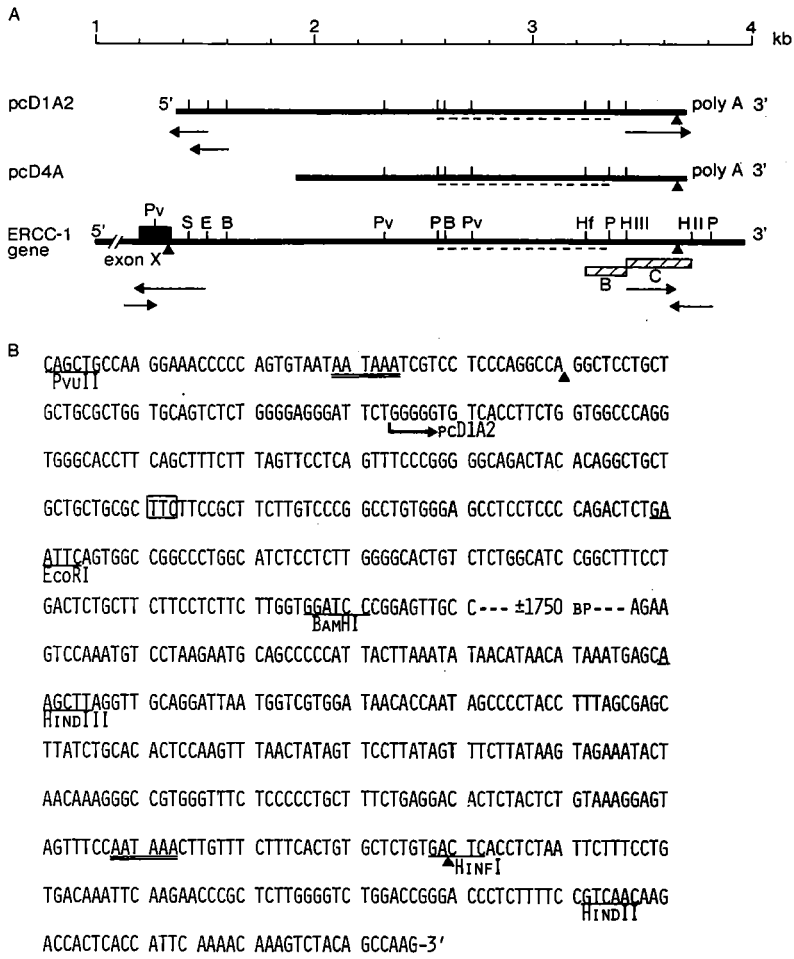

Figure 6. Detailed analysis of the 3′ ERCC-1 flanking region. A. Alignment of cDNA clones pcD4A and pcD1A2 with the genomic region flanking the 3′ site of ERCC-1 exon X. Triangles (▲) indicate AATAAA polyadenylation signals. The dashed line specifies fragments containing repetitive sequences as found by Southern hybridization using human repetitive DNA as a probe (not shown). Arrows indicate the sequence strategy and the shaded boxes B and C are probes used for the Northern blot analysis (see Fig. 7). Not all restriction enzyme sites are shown. Abbreviations: Pv : PvuII; S : SmaI; E : EcoRI; B : BamHI; P : PstI; Hf : HinfI; HIII : HindIII; HII : HindII. B. Composition of genomic and cDNA sequence data. Appropriate restriction enzyme sites shown in Fig.6A are indicated. The polyadenylation signals AATAAA are indicated by double underlining. The boxed sequence TTC might represent a polymorphism as it was not found in the genomic sequence. Triangles (▲) indicate polyadenylation sites as found by sequence analysis.

library (21) using a 3′ 1.05 kb genomic PvuII fragment as a probe. In addition to the earlier described ERCC-1 cDNA clones (8) several other clones (pcD1A2 and pcD4A) were isolated with this probe. As illustrated in Figure 6A the physical maps of these clones are colinear with the genomic map of the 3′ ERCC-1 flanking region. The longest clone pcD1A2 has an insert size of approximately 2.2 kb. Furthermore, Southern blot hybridization with human repeats as a probe revealed the presence of repetitive sequences in both cDNA clones as well as in the corresponding genomic fragments (not shown). The nucleotide sequence of the 5′ and 3′ termini of pcD1A2 and the corresponding genomic regions of cosmid 43-34 (7) were determined by the chemical cleavage method of Maxam and Gilbert (13). A composition of both sequences is presented in Figure 6B. The 3′ end of pcD1A2 harbors a poly(A) signal followed by a poly(A) tail indicating a transcriptional orientation which is similar to that of ERCC-1. In the 5′ direction clone pcD1A2 extends to 40 bp downstream of the polyadenylation site of the 1.1 kb ERCC-1 mRNA (see Figure 6B). This renders it conceivable that pcD1A2 is a partial cDNA clone from a longer ERCC-1 mRNA with an alternative polyadenylation site. To investigate this hypothesis poly(A)$^+$ RNA of HeLa nuclei and human primary fibroblasts was analysed by Northern blotting using a coding ERCC-1 cDNA probe and two genomic probes of the pcD1A2 region (see Figure 6A). The autoradiogram of this experiment is shown in Figure 7. ERCC-1 probe 3A3 which encodes the N-terminal half of the ERCC-1 protein recognizes the mature 1.0 - 1.1 kb ERCC-1 mRNA, a 3.4 kb transcript in fibroblasts and a 3.4 and 3.8 kb mRNAs in HeLa nuclei. These 3.4 and 3.8 kb RNA species are also recognized by two probes (B and C) flanking the HindIII site in the 3′ pcD1A2 region. The 3.8 kb transcript seems to be absent in primary fibroblasts. However, although with a much lower hybridization signal than the 3.4 band, the 3.8 kb could be detected in poly(A)$^+$ RNAs of a number of other cell lines (e.g. Figure 5, data not shown). We conclude from these results that pcD4A and pcD1A2 are partial cDNA clones derived from the 3.4 kb ERCC-1 transcript. The 3.8 kb molecule most likely terminates at a yet unidentified polyadenylation site further downstream.

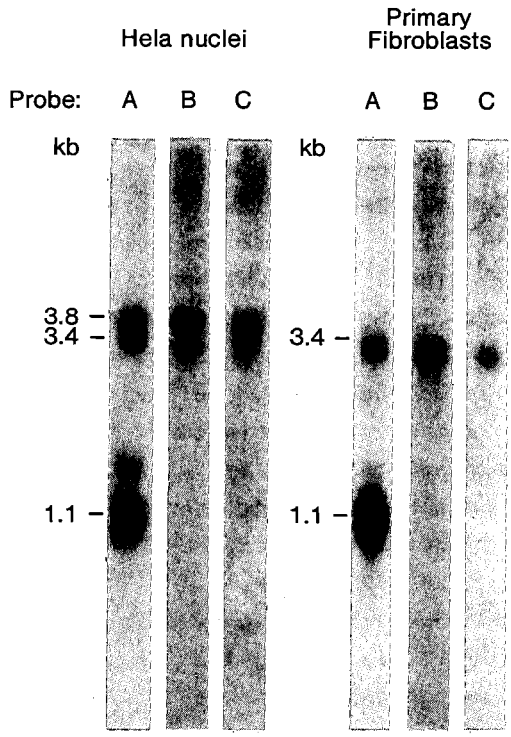

Figure 7. Northern blot analysis of ERCC-1 transcripts. Poly(A)$^+$RNA (± 20 µg) of HeLa nuclei and primary fibroblasts was size fractionated by agarose gelelectrophoresis (1% agarose) and after blotting to nitrocellulose filters hybridized to three ^{32}P-labeled DNA probes. A: 3A3, a 5′ PstI-fragment of ERCC-1 cDNA clone pcDE (8) B and C are genomic DNA fragments (HinfI-HindIII and HindIII-HindII respectively) from the 3′ terminal region of the ERCC-1 gene shown in Fig. 6A.

DISCUSSION

The excision repair gene ERCC-1 has been assigned to human chromosome 19 (8,22,23) and regionally mapped to 19q13.2-13.3 (24). Here we report the structural organization of the ERCC-1 gene. The gene is about 15 kb in length and harbors 10 exons. We have recently shown that the coding exon VIII of 72 bp is subject to alternative splicing. Only the cDNA of the larger ERCC-1 transcript (which includes exon VIII) inserted in a mammalian expression vector can correct the excision repair mutation in 43-3B cells (8). During pre-mRNA processing excision of introns requires branchpoint formation at a distance of -40 to -20 bp

from the 5′ exon borders. The consensus sequence PyNPyTPuAPy has been proposed for this branchpoint region (19,25,26). In contrast to yeast where a strict requirement for the heptamer TACTAAC has been observed (27) the sequence constraints for higher eukaryotes seem to be much more relaxed. Deletion or mutation of the branchpoint region can lead to branch formation at other criptic positions close by (19). It should be noted that except for exon VIII all ERCC-1 3′ intron junctions contain putative branchpoint motifs matching with the consensus PyNPyTPuAPy (Figure 2). The sequence proximal to exon VIII that most closely resembles the branchpoint consensus has a C-residue at the conserved T-position. Only in a few cases the presence of a C-residue at this position of the predicted branchpoint has been observed (18). Therefore, it is possible that the exon VIII branchpoint sequence is less efficiently utilized during RNA processing. Since the 3′ intron junction with ERCC-1 exon IX is located ,only 0.3 kb downstream it is conceivable that occasionally the branchpoint signal of exon VIII is skipped as a result of which the alternative splicing of this exon can occur. Analysis of a number of mouse ERCC-1 cDNA clones (manuscript in preparation) have not given any evidence for differential RNA procesing in mouse brain and testes tissue. It is therefore not excluded that the observed alternative splicing of ERCC-1 mRNA is the result of an inaccurate splicing event in human cells rather than a means to provide additional biological functions to the ERCC-1 protein.

Using a functional assay we have shown that the ERCC-1 promoter is located in a stretch of 170 bp proximal to the transcriptional start site. In this region there is no clear cut TATA box or AT-rich region which is typically positioned between -20 and -30 from the CAP-site (17). An ′abbreviated′ version of a CAAT-box might be located at -100. The apparent absence of these promoter motifs is frequently found in housekeeping genes that are driven by promoters containing the transcription factor Spl binding box GGGCGG (28). The ERCC-1 promoter is also devoid of such GC-boxes. The absence of classical transcription elements like CAAT, TATA and GC-boxes has recently been reported for the weakly expressed c-mos proto-oncogene (29) and is exceptional for eukaryotic genes. Therefore, it is possible that the ERCC-1

promoter represents a specific class of promoters. Several other sequence motifs, of which the significance is still unknown, have been notified in the ERCC-1 promoter region (Figure 3). Three CCTCC repeats are located in a pyrimidine rich region from -65 to -100. The promoter of the EGF-receptor gene harbors pyrimidine rich stretches of nucleotides with 4 times repeated TCC elements (30). These regions might have a regulatory function in EGF-receptor expression since nuclear proteins were found to bind to the TCC-containing regions (31). Possibly, ERCC-1 expression is mediated through a similar type of interaction between regulatory proteins and the CT-rich promoter elements.

We have not quantitatively determined the transcription activity of the ERCC-1 gene. However, qualitative observations from Northern blot analysis of poly(A)$^+$ RNA of different cell lines indicate low levels of ERCC-1 mRNA. Low constitutive expression of repair proteins has also been observed in E.coli and yeast (1,32). In bacteria an efficient removal of DNA lesions is provided by an excision repair system that is part of the DNA-damage inducible SOS-response. In yeast, the agents that elicit this response induce the excision repair gene RAD2 whereas transcription of RAD1, RAD3 and RAD10 is not affected by DNA-damage (33). From our experiments it can be inferred that ERCC-1 expression is not influenced by UV or MM-C induced DNA damage.

Several observations suggest very low levels of intracellular ERCC-1 protein. Firstly, as mentioned before, ERCC-1 transcripts belong to a class of low abundant mRNAs in a variety of human cell lines (8, results not shown). Secondly, the AUG start codon of the ERCC-1 protein lies in a very unfavourable sequence context (TCCAGATGG) for translation initiation (34) because of the presence of a pyrimidine at position -3. More than 95% of eukaryotic mRNAs examined have an A (75%) or a G-residue (20%) at that position and the consensus sequence CCACCATGG that emerged from a compilation of the mRNA initiation sites used by higher eukaryotes completely matched the most optimal translational start context as found by mutation analysis (35,36). In those studies the presence of a T at -3, as found in ERCC-1 mRNA, caused a 10-20 fold drop in the efficiency of translation initiation. Hence, the presence of a pyrimidine at

position -3 of the ERCC-1 mRNA suggests a low constitutive level of ERCC-1 translation.

The transfection efficiency of ERCC-1 constructs seems to be promoter dependent (Table 1), since the strong SV40 promoter yields less transformants than the weak ERCC-1 promoter. Further experiments are required to study the effect of the ERCC-1 expression level, however the preliminary data presented in Table 1 suggest that overexpression of ERCC-1 can be deleterious to the cell and that mimimum amounts of ERCC-1 protein are required for repair of DNA damages.

ERCC-1 mRNA molecules of 1.1, 3.4 and 3.8 kb have been identified by Northern blot analysis using coding ERCC-1 cDNA probes. The 1.1 kb transcript which is the major mRNA species encodes the protein that corrects the 43-3B mutation and harbors all exons shown in Figure 1. Although no full length cDNA clones have been isolated of the longer transcripts, there is strong evidence that differential polyadenylation gives rise to the 3.4 and 3.8 kb mRNAs: i. Both transcripts are recognized by 5′ ERCC-1 cDNA probes. ii. Genomic probes located approximately 2 kb 3′ of the 1.1 kb mRNA polyadenylation site hybridize with the 3.4 and 3.8 kb molecules. iii. cDNA clones of the 3.4 kb mRNA have the ERCC-1 transcriptional orientation and have a poly(A) tail preceeded by the polyadenylation signal AATAAA. These cDNA clones contain human repetitive sequences and are completely colinear with the genomic DNA as shown by restriction enzyme and partial sequence analysis. iv. Recently, functional mouse ERCC-1 cDNA clones have been isolated with different 3′ untranslated regions due to alternative polyadenylation (van Duin, in preparation).

In addition to the hexanucleotide AATAAA several other loosely conserved sequence elements have been suggested to play a role in mRNA 3′ end formation (see 37 for a review). The pentanucleotide CAYTG is found 3′ of the AATAAA before or after the poly(A) site in many genes and is suggested to mediate poly(A) formation through interaction with RNA of the U4 small ribonucleoprotein (38). Furthermore, T-rich and GT-rich sequences motifs are located downstream of poly(A) sites in many transcripts (37). Recent studies of Gil and Proudfoot (39) and McDevitt et al. (40) have demonstrated that these motifs function

synergistically and are important for efficient and accurate 3′ end formation. Inspection of the 3′ ERCC-1 sequence revealed the presence of CAYTG box and a T-rich element near the second AATAAA suggesting that this can be a better target for 3′ end processing than the more proximal AATAAA.

Multiple polyadenylation sites occurring in 3′ untranslated regions have been reported for a number of transcripts (41,42,43). It is unknown whether the selection of poly(A) sites within one transcriptional unit is a regulatory event. In this respect it is interesting to note that 3′ untranslated sequences of several transiently expressed genes have been implicated in mRNA stability and that recently AU-rich sequences with AUUUA motifs have been identified defining mRNA instability (44,45). Deletion of such sequences from the c-fos proto-oncogene abolishes the rapid turnover of c-fos transcripts (46) and confers a transforming potential to the gene (47). Such elements have not been found in the partially sequenced 3′ ERCC-1 flanking region. However, the apparent evolutionary conservation of differential polyadenylation of ERCC-1 transcripts could imply that mRNA 3′ end processing plays a regulatory role in ERCC-1 expression.

ACKNOWLEDGEMENTS

We thank Anneke Graus and Petra Warmerdam for technical assistance with the promoter studies. Dr. H. Okayama is kindly acknowledged for providing the cDNA library and Dr. F.J. Benham for giving the GAPDH probe. We are indebted to Rita Boucke, Tar van Os and Mirco Kuit for preparing the manuscript.
This work was financially supported by EURATOM (contract no. B16-141-NL) and MEDIGON, Foundation of Medical Scientific Research in the Netherlands (contract no. 900-501-091).

REFERENCES

1. Friedberg, E.C. (1985) DNA repair, Freeman & Company, San Francisco
2. Walker, G.C. (1985) Annu. Rev. Biochem. 54, 425-457.
3. De Weerd-Kastelein, E.A., Keijzer, W. and Bootsma, D. (1972) Nature 238, 80-83.
4. Fischer, E., Keijzer, W., Thielman, H.W., Popanda, O., Bohnert, E., Edler, E.G., Jung, E.G. and Bootsma, D. (1985) Mutation Res. 145, 217-225.

5. Thompson, L.H., Busch, D.B., Brookman, K.W., Mooney, C.L. and Glaser, P.A. (1981) Proc.Natl.Acad.Sci.USA 78, 3734-3737.
6. Thompson, L.H. and Carrano, A.V. (1983) UCLA Symposium on Molecular and Cell Biology New Series (Friedberg, E.C. and Bridges, B.R., eds.) vol. 11, 125-143.
7. Westerveld, A., Hoeijmakers, J.H.J., van Duin, M., de Wit, J., Odijk, H., Pastink, A., Wood, R.D. and Bootsma, D. (1984) Nature 310, 425-429.
8. Van Duin, M., de Wit., J., Odijk, H. Westerveld, A., Yasui, A., Koken, M.H.M., Hoeijmakers, J.H.J. and Bootsma, D. (1986) Cell 44, 913-923.
9. Doolittle, R.F., Johnson, M.S., Husain, I., van Houten, B., Thomas, D.C., Sancar, A. (1986) Nature 323, 451-453.
10. Hoeijmakers, J.H.J., van Duin, M., Westerveld, A., Yasui, A. and Bootsma, D. (1986) Cold Spring Harbor Symposia on Quantitative Biology vol. LI, 91-101
11. Wood, R.D. and Burki, H.J. (1982) Mutation Res. 95, 505-514.
12. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning. Cold Spring Harbor, New York.
13. Maxam, A.M. and Gilbert, W. (1980) Meth.Enzymol. 65, 499-560.
14. Grosveld, G.C., Koster, A and Flavell, R.A. (1981) Cell 23, 573-584.
15. Burke, J. (1984) Gene 30, 63-68.
16. Auffray C. and Rougeon, F. (1980) Eur.J.Biochem. 107, 303-314.
17. Breatnach, R. and Chambon, P. (1981) Annu.Rev.Biochem. 50, 349-384.
18. Keller, E.B. and Noon, W.A. (1984) Proc.Natl.Acad.Sci.USA 81, 7417-7420.
19. Reed, R. and Maniatis, T. (1985) Cell 41, 95-105.
20. Benham, F.J., Hodgkinson, S., Davies, K.E. (1984) EMBO J. 3, 2635-2640
21. Okayama, H. and Berg, P. (1983) Mol.Cell.Biol. 3, 280-289
22. Rubin, J.S., Prideaux, V.R., Huntington, F.W., Dulhanty, A.M., Whitmore, G.F. and Bernstein, A. (1985) Mol.Cell.Biol. 5, 398-405.
23. Thompson, L.H., Mooney, C.L., Burkhart-Schultz, K., Carrano, A.V. and Siciliano, M.J. (1985) Mol.Cell.Biol. 5, 881-884.
24. Brook, J.D., Shaw, D.J. Meredith, A.L. Worwood, M., Cowell, J., Scott, J., Knott, T.J., Litt, M., Bufton, L. and Harper, P.S. (1985) Cytogen.Cell Genet. 40, 590-591.
25. Keller, W. (1984) Cell 39, 423-425.
26. Aebi, M. and Weissmann, C. (1987) Trends in Genet. 3, 102-107.
27. Langford, C.J., Klin, F.J., Donath, C. and Gallwitz, C. (1984) Cell 36, 645-653.
28. Kim, S.H., Moores, J.C., David, D., Respess, J.G., Jolly, J. and Friedman, T. (1986) Nucleic Acids Res. 14, 3108-3118.
29. Propst, F.,Rosenberg, M.P., Lyer, A., Kaul, K. and Van de Woude, G.F. (1987) Mol. Cell. Biol. 7,1629-1637.
30. Ishii, S., Xu, Y.-H., Stratton, R.H., Roe, B.A. Merlino, G.T. and Pastan, I. (1985) Proc.Natl.Acad.Sci.USA 82, 4920-4924.
31. Ishii, S., Imamato, F., Yamanashi, Y., Toyoshima, K. and Yamamoto, T. (1987) Proc.Natl.Acad.Sci.USA 84, 4373-4378.

32. Friedberg, E.C. (1987) J.Cell Sci.Suppl. 6, 1-23.
33. Robinson, G.W., Nicolet, C.M., Kalainov, D. and Friedberg, E.C. (1986) Proc.Natl.Acad.Sci.USA 83, 1842-18466.
34. Kozak, M. (1986) Cell 47, 481-483.
35 Kozak, M. (1984) Nucl.Acids Res. 12, 857-872.
36. Kozak, M. (1986) Cell 44, 283-392.
37. Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Cell 41, 349-359.
38. Berget, S.M. (1984) Nature 309, 179-181.
39. Gil, A. and Proudfoot, N.J. (1987) Cell 49, 399-406.
40. McDevitt, M.A., Hart, R.P., Wong, W.W. and Nevins, J.R. (1986) EMBO J. 5, 2907-2913.
41. LeMeur, M.A., Galliot, B. and Gerlinger, P. (1984) EMBO J. 3, 2779-2786.
42. Hagenbüchle, O., Wellauer, P.K., Cribbs, D.L. and Schibler, U. (1984) Cell 38, 737-744.
43. Chen, M.-J., Shimada, T., Moulton, A.D., Cline, A., Humphries, R.K., Maizel, J. and Nienhuis, A.W. (1984) J.Biol.Chem. 259, 3933-3943.
44. Shaw, G. and Kamen, R. (1986) Cell 46, 659-667.
45. Brawerman, G. (1987) Cell 48, 5-6.
46. Treisman, R. (1985) Cell 42, 889-902.
47. Meylink, F., Curran, T., Miller, A.D. and Verma, I.M. (1985) Proc.Natl.Acad.Sci. USA 82, 4987-4991.

APPENDIX PAPER IV

Nucleic Acids Research 16, 5305-5322 (1988)

Volume 16 Number 12 1988 Nucleic Acids Research

Evolution and mutagenesis of the mammalian excision repair gene *ERCC*-1

M.van Duin, J.van den Tol, P.Warmerdam, H.Odijk, D.Meijer, A.Westerveld, D.Bootsma and J.H.J.Hoeijmakers

Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

Received April 7, 1988; Revised and Accepted May 19, 1988 Accession nos X07413, X07414

ABSTRACT

The human DNA excision repair protein ERCC-1 exhibits homology to the yeast RAD10 repair protein and its longer C-terminus displays similarity to parts of the E.coli repair proteins uvrA and uvrC. To study the evolution of this 'mosaic' ERCC-1 gene we have isolated the mouse homologue. Mouse ERCC-1 harbors the same pattern of homology with RAD10 and has a comparable C-terminal extension as its human equivalent. Mutation studies show that the strongly conserved C-terminus is essential in contrast to the less conserved N-terminus which is even dispensible. The mouse ERCC-1 amino acid sequence is compatible with a previously postulated nuclear location signal and DNA-binding domain. The ERCC-1 promoter harbors a region which is highly conserved in mouse and man. Since the ERCC-1 promoter is devoid of all classical promoter elements this region may be responsible for the low constitutive level of expression in all mouse tissues and stages of embryogenesis examined .

INTRODUCTION

Cell hybridization experiments have identified 6 complementation groups within DNA excision repair deficient UV-sensitive Chinese hamster ovary (CHO) cell lines (1-3). In many respects these mutants resemble cells from the genetic disorder xeroderma pigmentosum (XP) in which defects in at least nine genes (4,5) underly the extreme sensitivity of XP patients to sun exposure (UV light) and a predisposition to skin cancer. Cell fusion studies - although incomplete - have thusfar not revealed any overlap between these two classes of mutants (6). Hence, it is possible that 15 or more genes play a role in the excision of UV induced DNA damage in mammalian cells. Also in yeast and prokaryotes many loci have been found to be implicated in excision repair (see 7 for a review). In E.coli the uvrABC and D

gene products play a key role in the excision repair machinery and it is likely that comparable multiprotein complexes are operative in higher organisms.

By applying genomic DNA transfer we have isolated the human ERCC-1 gene which corrects the repair defect of UV and mitomycin-C (MMC) sensitive CHO 43-3B cells belonging to complementation group 1*) (8). Recently, also the ERCC-2 gene complementing group 2*) mutants has been isolated (9). The ERCC-1 gene only corrects CHO mutants of complementation group 1 (10) and restores all impaired repair characteristics to wild type levels (11) suggesting that ERCC-1 is the human homologue of the mutated gene in CHO group 1 mutants.

Based on similarity of the predicted ERCC-1 amino acid sequence with functional domains of other proteins, a putative nuclear location signal (NLS), DNA binding domain and ADP-monoribosylation site have been assigned to the ERCC-1 protein (12). Furthermore, a significant homology with yeast RAD10 and parts of the E.coli uvrA and uvrC repair proteins has been found (12-14). This suggests that DNA repair systems are well conserved during evolution. In this respect it is worth noting that the RAD6 protein which is involved in cellular response to DNA damage in yeast and recently found to be a ubiquitin conjugating enzyme, is also very strongly conserved up to mammals (15,16). The extent of homology between ERCC-1 and RAD10 makes it tempting to speculate that both proteins are functionally equivalent although ERCC-1 has a C-terminal extension of 83 amino acids which is absent in RAD10 (12). It is intriguing that this extra ERCC-1 part displays similarity with bacterial excision repair proteins. It is possible that evolution has endowed ERCC-1 with functional domains of distinct repair proteins in prokaryotes or that in the course of evolution the tail of RAD10 was lost. To investigate these possibilities and to further establish the significance of the postulated functional domains in the ERCC-1 protein, it is of interest to characterize the ERCC-1 gene of other organisms. Here we report the characterization of the mouse ERCC-1 gene and present mutation studies on the human ERCC-1 cDNA.

MATERIALS AND METHODS

Cell culture and transfection.

CHO 43-3B cells (17) were routinely grown in DMEM/F10 (1:1) medium with 5% fetal calf serum and antibiotics. To test for a functional ERCC-1 gene, DNA constructs (5-10 µg) were cotransfected with pSV3gptH (2-5 µg) to 5.10^5 43-3B cells in three 9 cm dishes as described previously (8). After 10-14 days of selection on mycophenolic acid (MPA) and MMC the cells were fixed, stained and clones were counted.

Unscheduled DNA Synthesis.

Two days after seeding in medium without MPA and MMC the cells were exposed to UV light (16 J/m^2) and incubated in thymidine-free, Ham's F10 medium supplemented with ^{3}H-thymidine (10 µCi/ml; specific activity 50 Ci/mmol) and 5% dialyzed fetal calf serum After Bouin fixation the preparations were processed for autoradiography (Kodak AR10 stripping film), exposed for 1 week at 4°C, developed and stained with Giemsa solution. For each preparation the number of grains per fixed square of 50 nuclei was counted.

RNA preparation and Northern blotting.

Total RNA was isolated from adult mice (Balb C), size fractionated on 1% agarose gels and after transfer to nitrocellulose filters hybridized to mouse ERCC-1 cDNA probes following standard procedures (18). Probes were labeled using the random primer technique (19).

cDNA cloning

A mouse brain cDNA library was prepared in phage λgt10 and screened with a human ERCC-1 cDNA probe as reported earlier (20).

Plasmid Constructions and Sequencing.

Routine protocols were used for plasmid isolation, subcloning and ligation (18). Nucleotide sequences were determined by the chemical cleavage (21) or chain termination method (22). Oligonucleotides for specific priming and Tab-linker mutagenesis were made with an Applied Biosystems DNA synthesizer. Human and mouse ERCC-1 cDNA plasmids were constructed as follows:

pTZME. Both EcoRI inserts of λgt10 mouse ERCC-1 cDNA were subcloned in pTZ19R (Pharmacia) yielding pM4a-2, harboring the 5'

cDNA part and pM4a-1 containing the remaining 3′ half. The artificial 5′ EcoRI site of pM4a-2 was removed by very short Bal31 treatment starting from the adjacent SstI site in the polylinker. The retained HindIII and PstI site of the polylinker were used to insert the HindIII-PstI fragment of pcDX (23) harboring the SV40 early promoter. Finally the EcoRI fragment of pM4a-1 was subcloned behind the 5′ cDNA part yielding pTZME.

pcDEMP1 and pcDEMP2. The 0.42 kb SmaI fragment of human ERCC-1 cDNA clone pcDE (14) was subcloned in pSP65 in which the unique PstI site was deleted giving pSPSma. This plasmid was linearized with PstI or KpnI (unique ERCC-1 cDNA sites) and 0.5 µg ligated overnight at 4°C to 4 µg of either a PstI-Tab linker (5′ GCTGCA 3′) or a Kpn Tab-linker (5′ GCGTAC 3′) in 10 µl (24). After ethanol precipitation (to eliminate excess unligated linker) the DNA was kinased and ligated to generate circular molecules. The insertion of a single Tab-linker in PstI and KpnI site was confirmed by sequence analysis. The PstI and KpnI mutated inserts of pSPSma were subsequently recloned in pcDE yielding pcDEMP-1 and pcDEMP-2 respectively.

pcDEBgl. pcDE (14) was digested with BglII, treated with Klenow DNA polymerase to create blunt ends and religated.

pcDEASt. The StuI fragment of pcDE was deleted by StuI digestion and religation yielded pcDEΔStu with a single StuI site and ERCC-1 sequences 3′ of codon 214 (14). Construct pcDEΔStu was linearized by StuI/AvaI double digestion which releases a ± 60 bp StuI-AvaI fragment. After klenow treatment to fill in the AvaI site the earlier deleted StuI fragment was inserted again yielding pcDEASt.

RESULTS

Characterization of mouse ERCC-1 cDNA.

Southern blot hybridization with DNA digests of various vertebrates indicated that the ERCC-1 cDNA was strongly conserved in evolution. Under reduced stringency conditions specific hybridization was found with mammalian, reptile, avian and fish DNA and very weakly with DNA of Drosophila (data not shown) whereas no hybridization was found with DNA from Trypanosomes, yeast and E.coli . This indicated that it should be possible to

isolate the mouse ERCC-1 cDNA from a mouse library using human ERCC-1 probes. A full length mouse ERCC-1 cDNA clone (designated λcDME) was isolated from a brain λgt10 cDNA library using human ERCC-1 cDNA as a probe. To establish whether this clone encoded a functional ERCC-1 protein the cDNA insert of λcDME was released with EcoRI, subcloned in a SV40 based mammalian expression cartridge (see Materials and Methods) and transfected to CHO 43-3B cells. The results of this experiment, in which human ERCC-1 cDNA construct pcDE (14) served as a positive control, are shown in Table 1. Using the mouse ERCC-1 cDNA construct pTZME, stable MMC resistant transformants could be generated with a transfection frequency (not shown) similar to pcDE. Furthermore, mouse cDNA transformants displayed levels of unscheduled DNA synthesis (UDS) in the wild type range (see Table 1), indicating that the isolated mouse cDNA harbors a functional ERCC-1 gene.
The complete nucleotide sequence of λcDME and predicted amino acid sequence are given in Figure 1. The mouse ERCC-1 cDNA appears to encode a protein of 298 amino acids and deduced molecular weight of 32970 Dalton. The alignment of the mouse ERCC-1 protein with its human homologue, the postulated functional domains (12) and the homology with the yeast RAD10 and

TABLE 1
Test for functional mouse ERCC-1 gene.

Cell line	Transfected DNA	MMC resistant clones	UDS*)
43-3B	pcDE	+	25 ± 1
43-3B	pTZME	+	27 ± 1
43-3B	-DNA	-	6 ± 1
CHO-9			19 ± 1**)
HeLa			33 ± 1**)

*) expressed as average number of grains (±SEM) per fixed square in 50 nuclei.
**) UDS of untransfected cells.
Mouse ERCC-1 cDNA was constructed behind the SV40 early promoter (yielding pTZME) and cotransfected with pSV3gptH to 43-3B cells followed by selection on mycophenolic acid and mitomycin-C (MMC). To substantiate repair proficiency of transformants UV-induced unscheduled DNA synthesis (UDS) was determined.

Figure. 1.
Nucleotide sequence and translated amino acids (in three letter code) of the mouse ERCC-1 cDNA. Amino acids are numbered on the left and nucleotides are numbered below the sequence. The polyadenylation signal AATAAA is underlined.

E.coli uvrA and uvrC repair proteins is shown in Figure 2A. Despite the differences between both mammalian proteins their extent of homology with the yeast and bacterial proteins is comparable. The positions of the conserved and non-conserved amino acid changes between the mouse and human protein are schematically depicted in Figure 2B. The overall homology between both proteins is 85%. However, it is striking that the majority (>70%) of the amino acid and nucleotide substitutions are concentrated in the N-terminal part of ERCC-1. Of the first 100 amino acids 70% are homologous whereas the region from 100 to 200 and 200 to 298 have similarities of 97% and 89% respectively. Two amino acid changes are found in the postulated NLS domain, 3 in the suggested ADP-monoribosylation site and none in the potential DNA-binding domain. The 2 conservative substitutions in the NLS (Ala→Thr and Lys→Arg) are at positions which have been shown not to be critical in the SV40 T-antigen NLS (25). Therefore these changes are not expected to abolish a potential NLS function of this domain. The high degree of sequence conservation in the putative DNA-binding domain supports the idea that this part of the protein is very important for its function and cannot

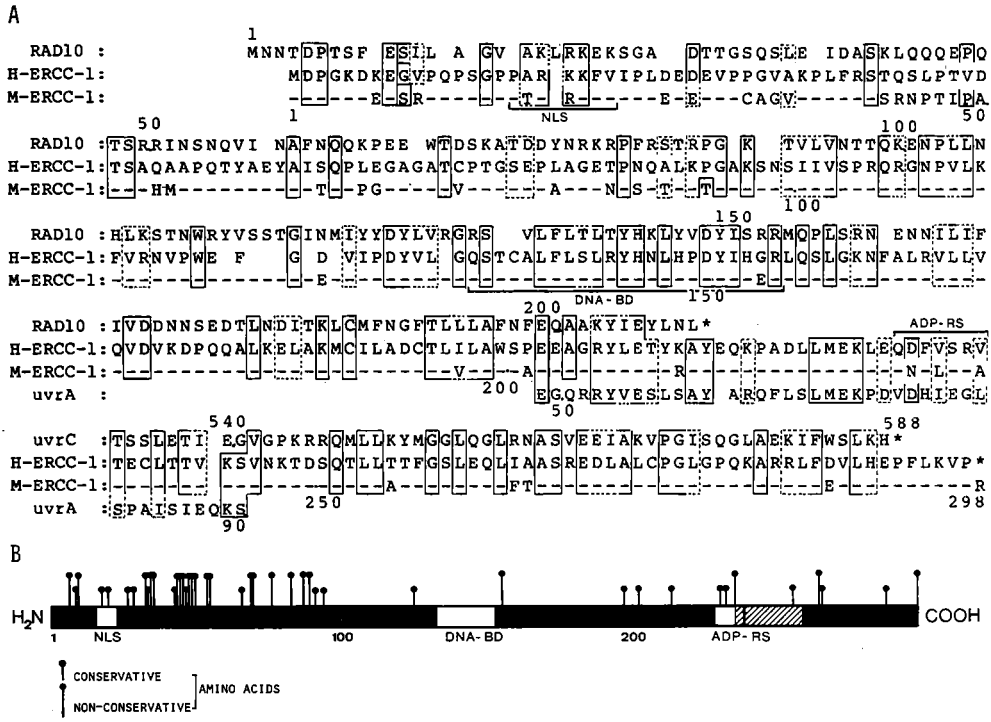

Figure. 2.
A: Alignment of mouse (**M**) and human (**H**) ERCC-1 amino acids (in one letter code) with yeast RAD10 and parts of E.coli uvrA and uvrC proteins. Of the mouse protein (horizontal bars) only differences with human ERCC-1 are depicted. Identical amino acids are boxed with solid lines and physico-chemically related residues (K, R; D, E and V, I, L) are indicated by open boxes between the microbial and mammalian proteins. Previously postulated domains (12) are shown by black lines: NLS, nuclear location signal; DNA-BD, DNA binding domain; ADP-RS, ADP-monoribosylation site. Asterisks indicate stopcodons. The RAD10, H-ERCC-1, uvrA and uvrC amino acid sequence and numbering are as described (14, 41, 42, 43 respectively).
B: Schematical presentation of conserved and non-conserved amino acid changes in mouse and human ERCC-1 protein. Conserved amino acids are grouped as follows: A, S, T, P and G; N, D, E and Q; H, R and K; M, L, I and V; F, Y and W. The postulated NLS, DNA-BD and ADP-RS are depicted as boxes (see also Fig.2A). The shaded region corresponds to human exon VIII which was previously found to be subject to alternative splicing (14).

tolerate many amino acid alterations. The 2 conservative amino acid changes in the region to which we have tentatively assigned an ADP-monoribosylation site are at positions not incompatible

with such a function. However, the non-conservative substitution at position 235 (Val→Ala) neighbours the arginine residue that is the actual site of ADP-ribosylation by cholera toxin in the consensus sequence of this domain (26). Hence, the difference between the mouse and human protein at this position makes it less likely that ADP-monoribosylation plays an important role in ERCC-1 protein processing.

At the nucleotide level the similarity of the coding regions of mouse and human ERCC-1 is 82% and of all base changes 70% are at 'wobble' base positions. With respect to the 5' untranslated region it is worth noting that at position -3 upstream of the translation initiation site a conserved C-residue is located which is highly exceptional for eukaryotic mRNAs (27).

Partial characterization of the mouse ERCC-1 promoter region.

The mouse ERCC-1 gene was isolated from a genomic EMBL-3 library using mouse cDNA as a probe. Four overlapping EMBL-3 clones, shown in Figure 3, hybridized to 5' and 3' cDNA probes (not shown) indicating that the mouse ERCC-1 gene has a maximum size of 16-17 kb which is comparable to the previously reported size of the human ERCC-1 gene (28). In order to determine the mouse ERCC-1 promoter sequence a 0.9 kb HindIII-BamHI fragment hybridizing with a 5' cDNA probe was subcloned in pTZ19R, yielding pMHB5. The sequence strategy of this clone, depicted in Figure 3, revealed genomic sequences that were completely identical to the cDNA sequence (Figure 1). Moreover, it was found that the genomic organization of mouse ERCC-1 exon 2 is similar to that of the human gene (not shown). A comparison of the mouse 5' genomic and untranslated cDNA sequence with the corresponding human sequence (28) is presented in Figure 4. A long stretch of homologous nucleotides is found around the transcriptional start site of the human gene (box A, Figure 4), which makes it likely that transcription of the mouse ERCC-1 gene initiates at the corresponding position in box A. This means that deletions or insertions rather than nucleotide substitutions have mainly contributed to the differences in the mouse (113 nucl.) and human (153 nucl.) 5' untranslated regions. We previously showed that the ERCC-1 promoter is located within 170 bp upstream of the transcriptional start site (28). Like the human ERCC-1 promoter

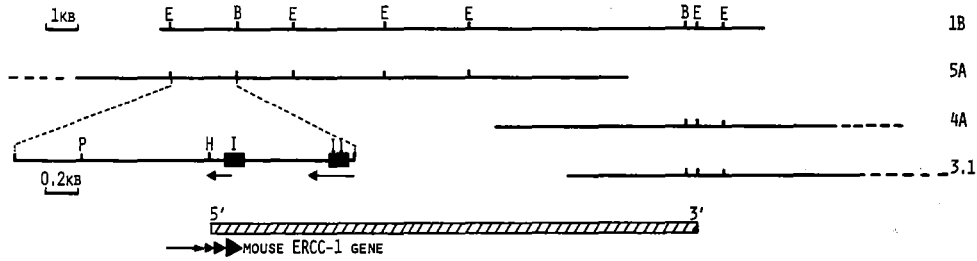

Figure. 3.
Isolation of the mouse ERCC-1 gene. Physical maps are shown of 4 overlapping clones that were isolated from a genomic EMBL-3 library. The position of the ERCC-1 gene (shaded bar) was deduced from hybridization with cDNA probes. Not all restriction sites are given. Non-characterized flanking parts are shown as dashed lines. The 5′ EcoRI-BamHI fragment was studied in detail and sequenced as shown by arrows. Exon I and II are indicated as black boxes. Restriction sites: E, EcoRI; B, BamHI; H, HindIII; P, PstI.

the mouse promoter is lacking clearly identifyable CAAT and TATA boxes as well as GC-rich regions. However, at 60 to 70 bp upstream of the 'cap'-site a domain of 30-40 bp (box B, Figure 4) is coinciding with a region in the human promoter which harbors three CCTCC pentanucleotides that are each spaced by 5 nucleotides. The conservation of this region suggests that it has

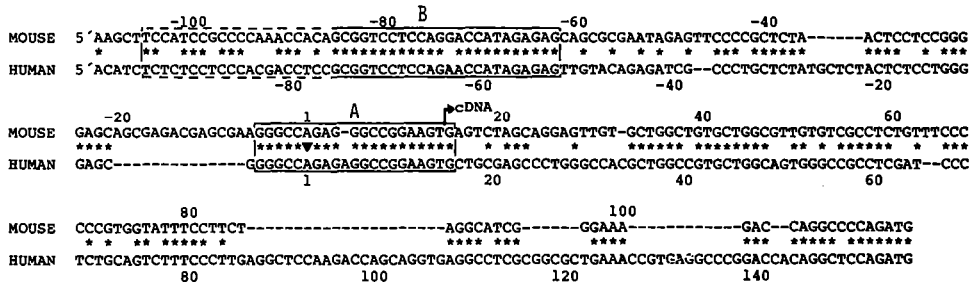

Figure. 4.
Alignment of 5′ mouse and human ERCC-1 sequences upstream of the translational start site (ATG). The mouse sequence is a compilation of genomic and cDNA (Fig. 1) sequence. The human sequence is from Van Duin et al. (14, 28). Boxes A and B show the most conserved regions in the ERCC-1 promoter. Nucleotide numbering is based on transcriptional start site (▲) of the human gene. Gaps are introduced to improve alignment of both sequences.

a regulatory function for ERCC-1 transcription. We have screened the EMBL sequence data base for nucleotide sequences homologous to box B, however, no apparent homology turned up.

Expression of ERCC-1 in mouse organs and developmental stages.

To examine possible regulatory aspects of the ERCC-1 promoter, ERCC-1 expression was investigated by Northern blot analysis of various mouse organs and stages of development. As a control for differences in the amount of RNA the blot was rehybridized with with a probe for the GAPDH mRNA which is considered to be present in relatively constant amounts and has a size very close to that of ERCC-1. As is shown in Figure 5 ERCC-1 transcripts could be detected in all organs (including also liver, stomach, bonemarrow and thymus, not shown) and stages of development examined. The differences in hybridization signal in the different lanes are to a large extent also observed with the GAPDH probe and therefore reflect mainly variation in the quantity of RNA on the filter. From comparison of the hybridization intensities of ERCC-1 and those of other genes (e.g. GAPDH and actin) we infer that ERCC-1 RNA falls in the class of low abundant messengers. Furthermore, in situ hybridization experiments on sections of mouse embryo's did not reveal tissues with a specific high level of ERCC-1 expression (not shown). These results support the conclusion that ERCC-1 is equally expressed at low levels throughout the whole body and at various stages of embryogenesis. The ERCC-1 promoter may thus represent a member of a novel class of promoters specifying a constitutive low expression level.

Mutagenesis of the human ERCC-1 cDNA.

A number of mutant human ERCC-1 cDNAs were constructed (see Materials and Methods) and transfected to CHO 43-3B cells to test for functionality. The results of these experiments are summarized in Figure 6. Construct pcDE-72 is included as negative control, since this clone lacks the alternatively spliced exon VIII which is essential for phenotypic complementation of the mutation in CHO 43-3B cells (14). Similarly all other alterations induced in the C-terminal portion of ERCC-1 appear to be deleterious as well. Construct pcDEASt, that encodes a 'RAD10-like' ERCC-1 protein terminating exactly at the point where (the homology with) the RAD10 gene product stops (see Figure 2A) does

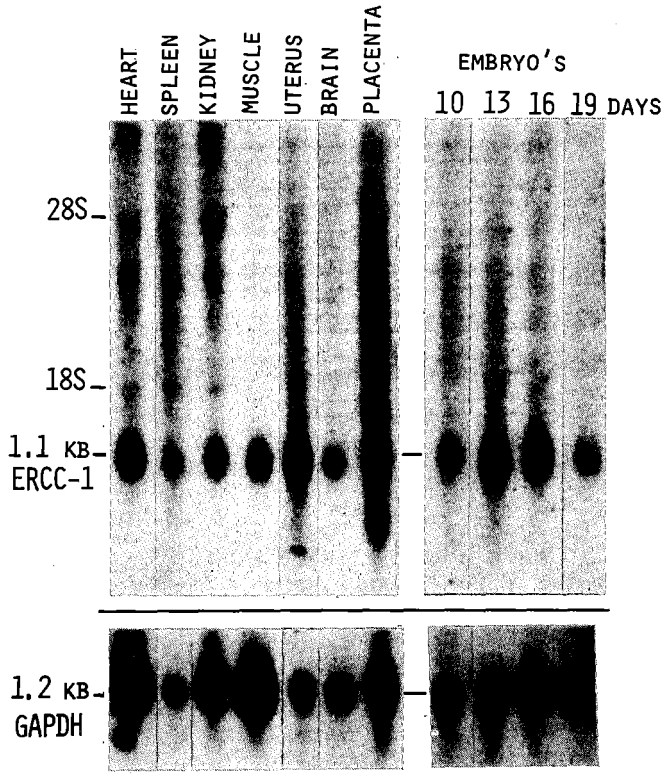

Figure. 5.
Northern blot analysis of mouse organs. Equal amounts (20 µg) of total RNA was size fractionated on a 1% agarose gel and after transfer to nitrocellulose hybridized to ^{32}P-labeled mouse ERCC-1 cDNA (upper part) and a probe for the glyceraldehydephosphate dehydrogenase (GAPDH, 44; lower part). The position of ribosomal subunits and ERCC-1 and GAPDH transcripts are indicated on the left.

not display detectable correcting activity. The same holds for construct pcDEBgl which encodes a truncated protein of 287 amino acids with 17 unrelated C-terminal residues due to a frameshift mutation. These findings are in striking contrast to deletion of the N-terminal ERCC-1 region. We reported previously that a truncated protein lacking the first 54 amino acids (encoded by construct pcD3C) is still able to confer MMC resistance to CHO mutant cells (14). Tab-linker mutagenesis (24) was applied to insert 6 nucleotides in the unique PstI and KpnI site in the coding part of the human ERCC-1 cDNA. The resulting introduction of a leucine and glutamine behind the putative DNA-binding domain (residue 158; construct pcDEMP1) did not affect the repair

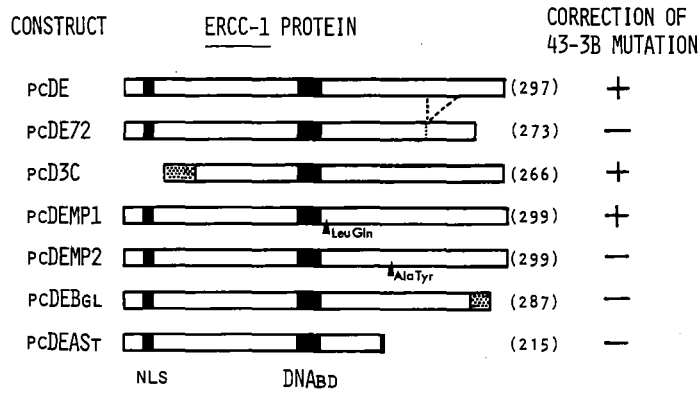

Figure. 6.
Mutagenesis of the ERCC-1 gene. SV40 early promoter driven ERCC-1 cDNA constructs were cotransfected with pSV3gptH to CHO 43-3B cells, to test for a functional gene. Stable transformants were selected on mycophenolic acid and MMC. Construct pcDE, harboring the complete human ERCC-1 cDNA, served as a positive control. pcDE-72 and pcD3C have been described previously (14) and encode proteins that are lacking 24 amino acids encoded by exon VIII and the first 162 coding nucleotides respectively. The other constructs are described in Materials and Methods. Putative functional domains (nuclear location signal, NLS; DNA-binding domain, DNA-BD) that emerged from human and mouse ERCC-1 amino acid comparison are depicted as black boxes. Unrelated amino acid sequences due to frame shifts are shown as dotted areas. Amino acid number is given in parentheses.

function of ERCC-1 in CHO mutants whereas an extra alanine and tyrosine residue distal from amino acid 208 (construct pcDEMP2) inactivated the protein. In conclusion the data presented here are consistent with the notion that the C-terminal part is crucial for ERCC-1 function in contrast to the N-terminus.

DISCUSSION

The complex evolution of the human ERCC-1 gene, resulting in a 'mosaic' type of homology to repair proteins of lower organisms has prompted us to study the gene in other species. Here we describe the isolation and characterization of cDNA and genomic clones of mouse ERCC-1. Transfection of the mouse ERCC-1 cDNA to 43-3B cells conferred UV- and MMC resistance and restored UDS. The induction of a UDS level inbetween that of CHO9 wild type and HeLa cells was also found after transfection of the human ERCC-1

gene and cDNA (8). The finding that UDS is higher in corrected 43-3B cells compared to wild type CHO cells can be explained by differences in nucleotide poolsizes, cell morphology and other factors influencing UDS. Alternatively, it is not excluded that transfected ERCC-1 genes from human and mouse induce a higher UDS level then the endogenous CHO gene.

Southern blot and sequence analysis of mouse genomic clones indicate that the mouse and human gene are similar in size. Preliminary data also suggest a very similar gene organization. The ERCC-1 gene of both mouse and man is driven by an exceptional promoter lacking 'classical' promoter elements. Comparison of the mouse and human promoter region revealed a highly homologous sequence at 50 to 90 bp upstream of the transcriptional start site. We have not noticed this putative promoter motif in other genes The high sequence conservation suggests that ERCC-1 expression is mediated through interaction of transcription factors with this region. However, additional experiments, including 'footprinting' are required to verify this assumption.

In HeLa cells ERCC-1 expression appears to be constitutive and not inducible by UV and MMC (28). We present here that ERCC-1 is transcribed at low levels in all mouse organs and stages of development investigated. Also Northern blot analysis of a number of different human cell lines revealed low constitutive levels of ERCC-1 expression (not shown). Therefore ERCC-1 is probably not only operative in repair of environmentally induced DNA-damage (e.g. UV photo products) but perhaps more importantly in the removal of DNA injuries that are induced at a constitutive rate in all cells and tissues by various intracellular processes.

With respect to ERCC-1 expression at the protein level it is worth noting that rodent and human ERCC-1 mRNAs have a C-residue at position -3 proximal to the translation initiation site. More than 95% of eukaryotic messengers harbor a purine at that position (27) and mutation studies have demonstrated that this dramatically enhances translation efficiency (29, 30). Although the presence of a G-residue at position +4 in the ERCC-1 transcripts might partially compensate for the lack of a purine at -3 (27) it seems likely that ERCC-1 has a conserved low translational efficiency. Yeast DNA excision repair genes also

belong to a category of lowly expressed genes (both at RNA and protein level) as deduced from codon usage and translation initiation consensus (31, 32). It will be of interest to see whether the recently isolated human repair genes XRCC-1 (L.Thompson, pers. comm.) and ERCC-2 (9) are subject to the same mode of translational control.

The cloning and sequence analysis of the mouse ERCC-1 gene has yielded instructive information that can be used to elucidate the function of the protein. Comparison of mouse and human amino acid sequences shows that the N-terminal protein part is much less conserved than the rest of the protein. This is in accordance with the pattern of similarity with the yeast RAD10 protein, which showed a high level of homology in the middle part of ERCC-1 (corresponding with the C-terminal half of RAD10) but only barely detectable similarity between the N-termini of both proteins. The apparent reduced evolutionary pressure for sequence conservation of the 5' portion of ERCC-1 fits also nicely with the transfection results of the 'decapitated' cDNA construct pcD3C (Figure 6) which demonstrated that a large N-terminal segment can be omitted without affecting the correcting potential of the ERCC-1 protein. This underlines the idea that ERCC-1 and the yeast RAD10 protein operate in a related step in the intricate excision repair process.

Notwithstanding this extensive homology, one important difference remains between the mammalian and yeast protein: a C-terminal extension of ERCC-1, which appears to be essential for proper functioning of the protein in CHO cells. The finding of this region in the mouse ERCC-1 protein indicates that it was present before the evolutionary lines to mouse and man diverged (65-80 million years ago).

Which evolutionary events may have caused the remarkable difference with the yeast gene? It is not excluded that a primordial ERCC-1 gene has lost its C-terminus to generate a RAD10-like version. Alternatively, C-terminal sequences might have been added to an ancestral RAD10-like gene yielding the ERCC-1-like gene structure. The fact that the extra region of ERCC-1 displays homology with prokaryotic repair proteins may be the result of convergent evolution. Another interesting

possibility is that the ERCC-1 gene has acquired functional domains from prokaryotic genes that originally resided in mitochondrial DNA but have migrated to the nucleus in the course of evolution. We are currently trying to isolate the Drosophila ERCC-1/RAD10 homologue which will hopefully shed more light on the evolution of ERCC-1.

Evolution has provided mouse ERCC-1 with one extra C-terminal amino acid compared to the human protein. Despite this minor difference our mutation studies indicate that the C-terminal part of ERCC-1 seems to be very important for its repair function. This region of ERCC-1 displays significant homology with the C-terminus of the E.coli uvrC protein (13). Interestingly, it appears that a mutation in this part of uvrC also leads to inactivation (33). Taken together these data support the idea that ERCC-1 and uvrC share a similar important function.

Several nuclear proteins are provided with positively charged domains that direct active transport through the nuclear membrane (34-38). A consensus sequence for this nuclear location signal (NLS) has emerged from detailed studies of Kalderon et al. (24, 38) and Colledge et al. (39). Amino acids 16 to 23 are conserved between mouse and man and show structural homology with a SV40 T-antigen NLS (12). It is therefore somewhat unexpected that the N-terminal ERCC-1 part appears to be non-essential. However, several explanations can be put forward for this observation. First, the size of the truncated ERCC-1 protein should allow it to enter the nucleus also by passive diffusion (40). Although probably less efficient, this process might still ensure sufficient ERCC-1 levels in the nucleus to permit correction of the CHO 43-3B mutation. In this context it should be realized that cDNA transformants in general have integrated multiple copies of the ERCC-1 cDNA constructs. Moreover, the cDNA inserts are driven by a SV40 promoter which is expected to accomplish a much higher expression level than the promoter of the unique endogenous ERCC-1 gene. Secondly, it is possible that N-terminal truncated ERCC-1 proteins can only be functional in rapidly dividing cells in culture in which the nuclear membrane is frequently absent, allowing the decapitated ERCC-1 protein to 'sneak in'. Therefore one could speculate that a deletion like in

cDNA construct pcD3C would be more serious in non-proliferating tissues *in vivo*. Finally, it cannot be excluded that ERCC-1 is harboring a second NLS of another type as has been found for polyoma virus large T (34) and the glucocorticoid receptor (35). Currently experiments are underway to investigate the hypothesized NLS in more detail.

Purification of the ERCC-1 protein will be an important prerequisite for studying the function of ERCC-1. The results presented in this paper demonstrate that such studies benefit from detailed analysis of evolutionary related genes.

ACKNOWLEDGEMENTS

We thank Dr. A. Yasui for making construct pcDEASt, Dr. G.C. Grosveld for generously providing the mouse genomic EMBL-library, Dr. F.J.Benham for the GAPDH probe, Dr. M.J.Vaessen for *in situ* hybridization experiments, Drs. M. von Lindern and M. Schilham for mouse organ RNAs and Dr. N.G.J. Jaspers for critically reading the manuscript. Mrs. R.J. Boucke and Mr. J. Fengler are acknowledged for preparing the manuscript. This work was financially supported by EURATOM (contract nr. BI6-141-NL) and MEDIGON, Foundation of Medical Scientific Research in The Netherlands.

* At the recent UCLA meeting on 'Mechanisms and Consequences of DNA damage processing ' in Taos (january 1988) it was decided to renumber the CHO excision deficient complementation groups 1 and 2 to match with the number of the human ERCC-genes correcting their respective defects: ERCC-1, group 1 (formerly group 2); ERCC-2, group 2 (formerly group 1).

REFERENCES

1. Thompson, L.H., Busch, D.B., Brookman, K.W., Mooney, C.L. and Glaser, P.A. (1981) Proc.Natl.Acad.Sci.USA 78, 3734-3737.
2. Thompson, L.H. and Carrano, A.V. (1983) UCLA symposium on Molecular and Cellular Biology New Series vol. 11, 125-143.
3. Thompson, L.H., Salazar, E.P., Brookman, K.W., Collins, C.C., Stewart, S.A., Busch, D.B. and Weber, C.A. (1987) J.Cell Sci.Suppl. 6, 97- 110.
4. De Weerd-Kastelein, E.A., Keijzer, W. and Bootsma, D. (1972) Nature 238, 80-83.
5. Fischer, E., Keijzer, W., Thielmann, H.W., Popanda, O., Bohnert, E., Edler, L., Jung, E.G. and Bootsma, D. (1985) Mutat.Res. 145, 217-225.

6. Thompson, L.H., Mooney, C.L. and Brookman, K.W. (1985) Mutat. Res. 150, 423-429.
7. Friedberg, E.C. (1985) DNA repair, Freeman & Company, San Francisco.
8. Westerveld, A., Hoeijmakers, J.H.J., van Duin, M., de Wit,J., Odijk, H., Pastink, A., Wood, R.D. and Bootsma, D (1984) Nature 310, 425-429.
9. Weber, C.A., Salazar, E.P., Stewart, S.A. and Thompson, L.I (1988) Mol.Cell.Biol. 8, 1137-1146.
10. Van Duin, M., Janssen, J.H., de Wit, J., Hoeijmakers, J.H.J., Thompson, L.H., Bootsma, D. and Westerveld, A. (1988) Mutat.Res. 193, 123-130.
11. Zdzienicka, M.Z., Roza, L., Westerveld, A., Bootsma, D. and Simons, J.W.I.M. (1987) Mutat.Res. 183, 69-74.
12. Hoeijmakers, J.H.J., Van Duin, M., Westerveld, A., Yasui, A. and Bootsma, D. (1986). Cold Spring Harbor Symposia on Quantitative Biology vol. LI, 91-101.
13. Doolittle, R.F., Johnson, M.S., Husain, I., Van Houten, B., Thomas, D.C. and Sancar, A. (1986) Nature 323, 451-453.
14. Van Duin, M., de Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M.H.M., Hoeijmakers, J.H.J. and Bootsma, D. (1986) Cell 44, 913-923.
15. Jentsch, S., McGrath, J.P. and Varshavsky, A. (1987) Nature 329, 131-134.
16. Ozkaynak, E., Finley, D., Solomon, M.J. and Varshavsky, A. (1987) EMBO J. 6, 1429-1439.
17. Wood, R.D. and Burki, H.J. (1982) Mutat.Res. 95, 505-514.
18. Maniatis, T., Fritsch, E.F. and Sambrook, J.(1982). Molecular cloning, A laboratory Manual (Cold Spring Harbor, New York).
19. Feinberg, A.P. and Vogelstein , B. (1983) Anal.Biochem. 132, 6-13.
20. Meijer, D., Hermans, A., von Lindern, M., van Agthoven, T.,de Klein, A., Mackenbach, P., Grootegoed, A., Talarico, D. Della Valle, G. and Grosveld, G. (1987) EMBO J. 6, 4041-4048.
21. Maxam, A.M. and Gilbert, W. (1980) Meth.Enzymol. 65, 499-560.
22. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J.Mol.Biol. 143, 161-178.
23. Okayama, H. and Berg, P. (1983) Mol.Cell Biol. 3, 280-289.
24. Barany, F. (1985) Gene 37, 111-123.
25. Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984) Cell 39, 499-509.
26. Van Dop, C., Tsubokawa, M., Bourne, R. and Ramachandran, J. (1984) J.Biol.Chem. 259, 696-698.
27. Kozak, M. (1984) Nucl.Acids Res. 12, 857-872.
28. 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) Nucl.Acids Res. 15, 9195-9213.
29. Kozak, M. (1986) Cell 44, 283-292.
30. Kozak, M. (1987) Nucl.Acids Res. 15, 8115-8148.
31. Sharp, P.M., Tuohy, T.M.F., Mosurski, K.R. (1986) Nucl. Acids Res. 14, 5125-5143.
32. Hamilton, R., Watanabe, C.K., De Boer, H.A. (1987) Nucl. Acids Res. 15, 3581-3593.
33. Van Sluis, C. and Dubbeld, D. (1983) in: DNA Repair, A Laboratory Manual of Research Procedures vol. 2 (eds. E.C

Friedberg and P.C. Hanawalt) 267-283.
34. Richardson, W.D., Roberts, B.L. and Smith, A.E. (1986) Cell 44, 77-85.
35. Picard, D. and Yamamoto, K.R. (1987) EMBO J. 6, 3333-3340.
36. Burglin, T.R. and De Robertis, E.M. (1987) EMBO J. 6,2617-2625.
37. Stone, J., de Lange, T., Ramsay, G., Jakobovits, E., Bishop, M., Varmus, H. and Lee, W. (1987) Mol.Cell Biol. 7, 1697-1709.
38. Kalderon, D., Richardson, W.D., Markham, A.F. and Smith, A.E. (1984) Nature 311, 33-38.
39. Colledge, W.H., Richardson, W.D., Edge, M.D. and Smith, A.E. (1986) Mol.Cell.Biol. 6, 4136-4139.
40. Paine, P.L., Moore, L.C. and Horowitz, S.B. (1975) Nature 254, 109-114.
41. Reynolds, P., Prakash, L., Dumais, D., Perozzi, G. and Prakash, S. (1985) EMBO J. 4, 3549-3552.
42. Husain, I., Van Houten, B., Thomas, D.C. and Sancar, A. (1986) J. Biol. Chem. 261,4895-4901.
43. Sancar, G.B., Sancar, A. and Rupp, D. (1984) Nucl. Acids Res. 12, 4593-4608.
44. Benham, F.J., Hodgkinson, S. and Davies, K.E. (1984) EMBO J. 3, 2635-2640.

APPENDIX PAPER V

Mutation Research 193, 123-130 (1987)

Mutation Research, 193 (1988) 123–130
DNA Repair Reports
Elsevier

MTR 06267

Transfection of the cloned human excision repair gene *ERCC-1* to UV-sensitive CHO mutants only corrects the repair defect in complementation group-2 mutants

M. van Duin, J.H. Janssen, J. de Wit, J.H.J. Hoeijmakers, L.H. Thompson [1], D. Bootsma and A. Westerveld

Department of Cell Biology and Genetics, Erasmus University, Rotterdam (The Netherlands) and [1] Biomedical Sciences Division L-452, Lawrence Livermore National Laboratory, Livermore, CA 94550 (U.S.A.)

(Received 17 August 1987)
(Revision received 22 September 1987)
(Accepted 28 September 1987)

Keywords: DNA excision repair; *ERCC-1*; CHO mutants; DNA transfection.

Summary

The human DNA-excision repair gene *ERCC-1* is cloned by its ability to correct the excision-repair defect of the ultraviolet light- and mitomycin-C-sensitive CHO mutant cell line 43–3B. This mutant is assigned to complementation group 2 of the excision-repair-deficient CHO mutants. In order to establish whether the correction by *ERCC-1* is confined to CHO mutants of one complementation group, the cloned repair gene, present on cosmid 43–34, was transfected to representative cell lines of the 6 complementation groups that have been identified to date. Following transfection, mycophenolic acid was used to select for transferants expressing the dominant marker gene *Ecogpt*, also present on cosmid 43–34. Cotransfer of the *ERCC-1* gene was shown by Southern blot analysis of DNA from pooled (500–2000 independent colonies) transformants of each mutant. UV survival and UV-induced UDS showed that only mutants belonging to complementation group 2 and no mutants of other groups were corrected by the *ERCC-1* gene. This demonstrates that *ERCC-1* does not provide an aspecific bypass of excision-repair defects in CHO mutants and supports the assumption that the complementation analysis is based on mutations in different repair genes.

During the last decade a large number of ultraviolet light (UV)-sensitive Chinese hamster ovary (CHO) cell lines has been isolated that are deficient in DNA-excision repair (Thompson et al., 1981; Wood and Burki, 1982; Stefanini et al., 1982). Genetic characterization performed by fusing cells from different mutant cell lines revealed the presence of at least 6 complementation groups (Thompson et al., 1981, 1987a; Thompson and Carrano, 1983). These mutants are phenotypically comparable with cell strains isolated from DNA-excision-repair-deficient xeroderma pigmentosum (XP) patients. These XP cells have been classified in 9 complementation groups (De Weerd-Kastelein et al., 1972; Fischer et al., 1985). Although the cell fusion data are far from complete, none of the

Correspondence: A. Westerveld, Department of Cell Biology and Genetics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam (The Netherlands).

tested XP and Chinese hamster excision-repair mutants belong to the same group (Stefanini et al., 1985; Thompson et al., 1985a).

The UV sensitivity of CHO mutants belonging to complementation groups 1, 2 and 5 has been corrected to nearly wild-type levels in genomic transfections with human DNA (Rubin et al., 1983; MacInnes et al., 1984; Westerveld et al., 1984; Weber et al., 1987). The human *ERCC-1* gene that corrects the repair defect of the complementation group-2 43–3B cell line has been molecularly cloned (Westerveld et al., 1984; Van Duin et al., 1986). Complementation group-2 mutants display an extreme sensitivity to the drug mitomycin-C (MMC) and have also an increased sensitivity to 4-nitroquinoline-1-oxide and *N*-acetoxy-2-acetyl-aminofluorene and alkylating agents. *ERCC-1* converts all these characteristics of group-2 mutants to wild-type levels. Moreover, the UV-induced mutation frequency at the Na^+/K^+ ATPase locus and the rate of cyclobutane pyrimidine dimer removal were also returned to normal (Zdzienicka et al., 1987).

It is generally assumed that intergenic complementation accounts for the classification of the UV-sensitive CHO mutants in different complementation groups. However, it should be noted that intragenic complementation or aspecific bypass of repair defects after cell fusion in the complementation assay cannot be excluded. To investigate the specificity of the cloned *ERCC-1* gene for group-2 mutants, we have examined whether *ERCC-1* can correct the repair defect in the other CHO complementation groups. Therefore, the cloned *ERCC-1* gene was transfected to representative cell lines of all 6 CHO complementation groups. UV survival and UV-induced repair synthesis data of the *ERCC-1*-containing transformants revealed that only CHO mutants belonging to complementation group 2 are transformed to wild-type.

Materials and methods

Cell lines and culture conditions

The wild-type CHO cell lines CHO-9 and AA8 have been described by Burki et al. (1980) and Thompson et al. (1980), respectively. The UV-sensitive mutants of CHO-9 and AA8 and corresponding complementation-group assignments are listed in Table 1. The cells were routinely cultured in a 1:1 mixture of Ham's F10 and Dulbecco's ME medium (DMEM) supplemented with antibiotics and 3% fetal and 7% newborn calf serum.

TABLE 1

UV-SENSITIVE CHO MUTANTS AND CORRESPONDING COMPLEMENTATION GROUPS THAT WERE USED FOR cos 43–34 TRANSFECTION EXPERIMENTS

Mutant	Complementation group	Parental line	Reference [a]
UV5	1	AA8	1
UV4, UV20	2	AA8	1
43–3B	2	CHO-9	2
UV24	3	AA8	1
UV41, UV47	4	AA8	1
UV135	5	AA8	3
UV61	6	AA8	4

[a] 1 = Thompson et al., 1981; 2 = Wood and Burki, 1982; 3 = Thompson et al., 1982; 4 = Thompson, 1987a.

DNA transfection and selection of transferants

One day before DNA transfection 5×10^5–1×10^6 recipient cells were seeded in 100-mm Petri dishes and cultured in a humidified incubator (37 °C, 5% CO_2). The transfection was carried out using the calcium-phosphate precipitation method (Graham and Van der Eb, 1973). Cosmid 43–34 (Fig. 1) (Westerveld et al., 1984) used in the transfection, harbors the 15-kb human excision-repair gene *ERCC-1* and in addition the dominant marker genes *agpt*, encoding G418 resistance, and *Ecogpt*, which allows selection with mycophenolic acid (MPA). To each Petri dish 1 μg cosmid 43–34 + 10 μg 43–3B carrier DNA was supplied. Following overnight exposure to the precipitate, the cells were treated with dimethylsulfoxide (10% for 30 min) and cultured for 24 h in non-selective medium to allow expression of the transfected DNA. Selection for the *Ecogpt* marker was performed in modified MPA medium containing F10/DMEM 1:1 and antibiotics, 10% fetal calf serum, aminopterin (0.2 μg/ml), thymidine (5 μg/ml), xanthine (10 μg/ml), hypoxanthine (15 μg/ml), mycophenolic acid (25 μg/ml) and deoxycytidine (2.3 μg/ml) (Mulligan and Berg, 1980;

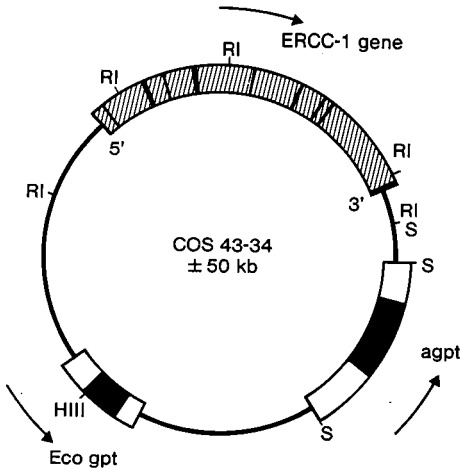

Fig. 1. Schematical presentation of cosmid 43–34 harboring approximately 42 kb of genomic DNA of a secondary 43–3B transformant (Westerveld et al., 1984). The cosmid vector pTCF (Grosveld et al., 1982) is flanked by SalI sites and contains the *agpt* gene. Furthermore the *Ecogpt* gene (on plasmid pSV3*gpt*H) and the human *ERCC-1* gene are located on the cosmid insert. Shaded boxes and small black boxes represent introns and exons of the *ERCC-1* gene, respectively (van Duin, in preparation). Arrows indicate transcriptional orientation. Abbreviations: S: SalI; RI: EcoRI; HIII: HindIII (not all restriction sites are shown).

Westerveld et al., 1984). The selection medium was refreshed every 3–4 days. After 2 weeks MPA-resistant colonies were trypsinized and grown into mass cultures. Within 1 week unscheduled DNA synthesis and UV survival were determined. The transformed cells were cultured for another 2 weeks before genomic DNA was isolated.

UV survival

Exponentially growing cultures were trypsinized and 200–4000 cells were plated onto 60-mm dishes and left to attach for 12 h. Subsequently cells were rinsed with phosphate-buffered saline (PBS) and exposed to UV light with a 254-nm low-pressure mercury, germicidal Philips TUV lamp with a fluence rate of 0.6 J/m^2. After cultivation in non-selective medium for 7 days clones were fixed and stained with Coomassie brilliant blue. For each dose 3–6 dishes were used. The relative survival was plotted versus the UV dose.

Unscheduled DNA synthesis (UDS)

Two days after seeding in medium without MPA, the cells were exposed to UV light (16 J/m^2) and incubated in thymidine-free, Ham's F10 medium supplemented with [^{3}H]thymidine (10 μCi/ml; specific activity 50 Ci/mmole) and 5% dialyzed fetal calf serum. After Bouin fixation the preparations were processed for autoradiography (Kodak AR10 stripping film), exposed for 1 week at 4°C, developed and stained with Giemsa solution. For each preparation the number of grains per fixed square of 25–50 nuclei was counted.

Southern blot analysis

Genomic DNA was isolated as described by Jeffreys et al. (1977), digested to completion with *EcoRI* and electrophoresed in 0.7% agarose gels. After transfer of the size-fractionated DNA to nitrocellulose (Southern, 1975) the filter was hybridized with ^{32}P-labelled nick-translated probes (Rigby et al., 1977). The hybridization was visualized by autoradiography.

Results

Cosmid 43–34, harboring the *ERCC-1* gene and the dominant markers *Ecogpt* and *agpt* (Fig. 1), was transfected to representative mutants of the various UV-sensitive CHO complementation groups. After transfection, a selection was carried out in MPA-containing medium for transferants with a functional *Ecogpt*. The mutants had a transfection frequency of approximately 100 MPA-resistant clones per μg cosmid 43–34 DNA per 10^6 cells. The MPA-resistant clones of each transfected mutant were pooled and grown into mass populations consisting of 500–2000 independent transferants. DNA was isolated from the mass populations and characterized by Southern blot analysis using *Ecogpt* and *ERCC-1* probes. The various transfected populations and the original CHO mutants were analysed by UV survival and UDS.

Southern blot analysis

DNA from the pooled transformants was analyzed by Southern blotting. *EcoRI*-digested

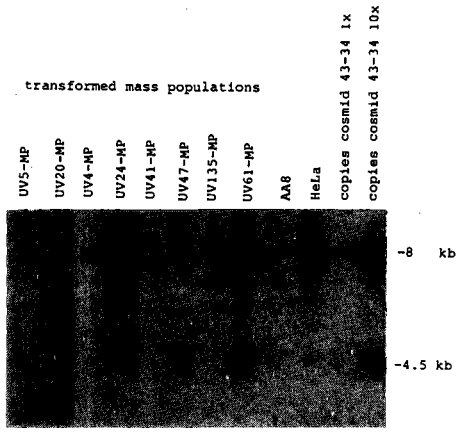

Fig. 2. Southern blot analysis of EcoRI-digested DNA of cos 43–34-transformed mass populations (MP), HeLa (human) and CHO AA8 cells probed with complete human *ERCC-1* cDNA. The last 2 lanes on the right harbor 1 and 10 copies of EcoRI-digested cosmid 43–34 per diploid genome, respectively. EcoRI digestion of cosmid 43–34 results in 3 fragments that hybridize with a complete *ERCC-1* cDNA probe: two 4.5-kb fragments and one 8-kb fragment (see also Fig. 1).

DNA was size-fractionated by agarose-gel electrophoresis, transferred to nitrocellulose and the blot was hybridized with a ^{32}P-labeled *Ecogpt* and a complete *ERCC-1* cDNA probe (Van Duin et al., 1986). The results of the *ERCC-1* hybridization are presented in Fig. 2. The 1.0-kb cDNA probe recognizes an 8-kb and 2 4.5-kb fragments of the human *ERCC-1* gene in control HeLa DNA, whereas the hybridization conditions did not allow detection of the Chinese hamster *ERCC-1* gene in AA8 cells (see also Fig. 1). The number of cosmid 43–34 copies that was integrated by transformants was estimated from the hybridization signal with HeLa DNA and the titration of 1 and 10 cosmid copies per diploid genome. On average the transformed mass populations harbored 1–5 cosmid molecules per cell. UV5 and UV20 transformants displayed the lowest and the UV61-MP the highest hybridization signal. UV61-MP contains in the order of 5–10 copies per cell. A similar conclusion could be drawn from the hybridization experiments with the *Ecogpt* probe (data not shown).

UV survival

The UV sensitivities of the various transformants were analyzed to examine whether the transfected *ERCC-1* can correct the impaired UV survival of one or more CHO complementation groups. The UV-survival plots of the wild-type AA8 cell line, the mutants and the cos 43–34 transformants are presented in Fig. 3. The mutants of complementation groups 1, 3, 4 and 5 showed a relative D37 (D37 mutant/D37 AA8) of 0.1 and this D37 was found for the transfected populations as well. The moderately UV-sensitive UV61 mutant (group 6) and the transformed UV61-MP had an intermediate survival (D37 = 0.3) as compared with the other mutants and AA8 cells. Thus, introduction of *ERCC-1* in the mutants of complementation groups 1, 3, 4, 5 and 6 had no effect on their response to UV irradiation. However, for both representatives of complementation group 2 (UV4 and UV20), the UV sensitivities of the transformed populations varied from those of the corresponding mutants. The survival curves of both populations were biphasic (see Fig. 3). About 65% of the UV4-MP and 10% of the UV20-MP cells behaved AA8-like with a relative D37 (D37 transformants/D37 AA8) value of 1. The UV survival of the sensitive fractions, about 35% of the UV4-MP and 90% of the UV20-MP cells, was similar to that of the mutant cells. For both fractions a relative D37 of 0.1 was calculated.

Unscheduled DNA synthesis

The capacity of cos 43–34 transformants to remove UV-induced DNA damage was measured by the incorporation of [^{3}H]thymidine as the result of UDS. The results, expressed as grains per fixed square of 25–50 nuclei, are presented in Table 2. In 5 of the mutant cell lines, UV4, 5, 20, 24 and 41, the UDS levels range from 5 to 10% of that found in wild-type cells. UV47 (group 4) and UV135 (group 5) exhibited about 30% of the wild-type AA8 activity, while the moderate UV-sensitive UV61 (group 6) cell line showed a UDS level of 60% compared to the AA8. The transformants of complementation groups 1, 3, 4, 5 and 6 did show UDS activities comparable with the original mutant cell line. Restoration to wild-type levels was found in transformants of complementation group 2. However, not all cells displayed

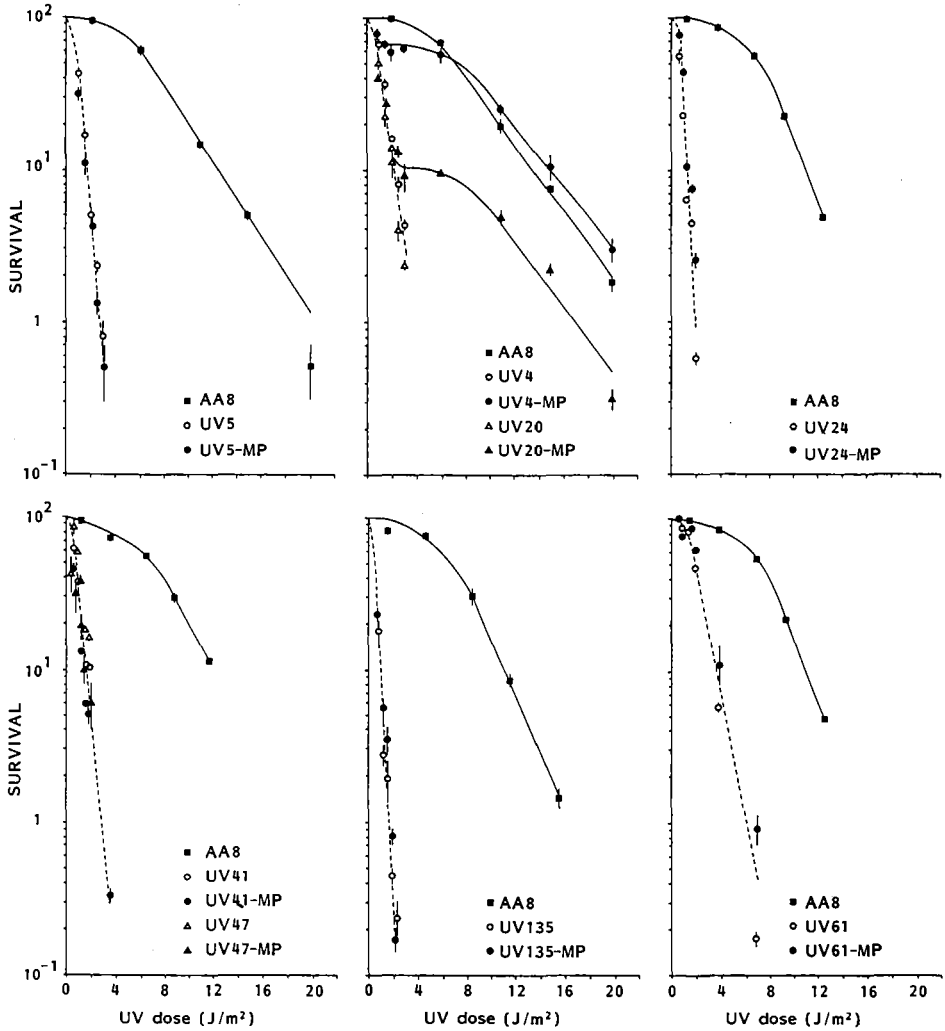

Fig. 3. UV-survival curves of wild-type (AA8), mutants (UV5; UV4; UV20; UV24, UV41; UV47; UV135; UV61) and transfected mass populations (MP): UV5-MP; UV4-MP; UV20-MP; UV24-MP; UV41-MP; UV47-MP; UV135-MP; UV61-MP. Error bars in the data points represent the SEM (standard error of the mean).

UDS activities identical to the wild-type level. In UV20-MP about 10% of the cells exhibited wild-type UDS levels and 90% of the cells had a mutant phenotype. Also UV4 transferants could be divided in 2 classes. About 65% of the UV4-MP cells showed wild-type UDS levels and about 35% were of the mutant type. These results are in accordance with the UV-survival data and ap-

TABLE 2

UV-INDUCED UNSCHEDULED DNA SYNTHESIS

Complementation group	Cell line			Grain counts [a] (mean ± SEM)			
	Mutant	Mass population	Wild type	Exp. 1	Exp. 2	Exp. 3	Exp. 4
			AA8	26.4±0.6			28 ±1.1
			CHO9	19.1±0.5	18.6±0.6	29.0±2.0	
			HeLa	37.9±0.7	33.0±0.6		43 ±2.2
1	UV5				2.4±0.3		
		UV5-MP			2.4±0.3		
2	UV4				1.8±0.2		
		UV4-MP (35%) [b]			1.9±0.3		
		UV4-MP (65%) [b]			25.7±0.6		
	UV20			0.9±0.2			
		UV20-MP (92%) [b]		1.6±0.3			
		UV20-MP (8%) [b]		25.1±0.6			
3	UV24					1.5±0.3	
		UV24-MP				2.6±0.4	
4	UV41						2.1±0.4
		UV41-MP					1.9±0.5
	UV47						7.1±0.8
		UV47-MP					8 ±1.3
5	UV135						11 ±2.2
		UV135-MP					8.2±0.8
6	UV61					19 ±1.1	
		UV61-MP				18 ±1.1	

[a] The number of grains is given as mean ± SEM per fixed square of non-S-phase nucleus and for each preparation 25–50 nuclei were characterized.

[b] Percentage is determined by screening 200 non-S-phase nuclei.

parently due to the fact that not all MPA-resistant transformants harbor an active *ERCC-1* gene.

Discussion

The results presented here demonstrate that *ERCC-1* corrects the DNA-excision-repair defect of mutant cell lines assigned to complementation group 2. The defect in mutants of the other 5 complementation groups is not restored, although Southern blot analyses revealed the presence of *ERCC-1*. The transfected cells were continuously grown in MPA-selection medium, so each cell should contain at least one functional *Ecogpt* copy. The mean copy number per cell for *Ecogpt* as well as *ERCC-1* sequences varied from 5–10 in UV61-MP to 1–5 in the other transformants. Recently we showed that the average amount of stably integrated DNA in the genome of transfected cells is species- and cell-line-specific (Hoeijmakers et al., 1987; unpublished results). In general human cells incorporate 20- to 100-fold less transfected DNA than Chinese hamster cells. The mutants derived from the CHO AA8 subline studied in the present paper integrate considerably less than mutants isolated from CHO-9, underlining the conclusion that the amount of stably integrated DNA is cell-line-specific.

In case of UV20 transformants an inefficient transfer of a functional *ERCC-1* gene was achieved since only 8% of the cells in the transformed population were repair-proficient. The possibility that potential complementation of the other mutants by the *ERCC-1* gene is masked due to inefficient gene transfer seems unlikely, since Southern blot analysis revealed that all cell lines had integrated 1–5 *ERCC-1* copies (Fig. 2) and that UV61 had incorporated even more transfected DNA than UV4 and UV20.

The *ERCC-1* gene has been mapped on the long arm of human chromosome 19 (Rubin et al., 1985; Van Duin et al., 1986). Using cell-fusion

experiments Thompson et al. (1987b, personal communication) have recently assigned *ERCC* genes complementing CHO group-1, -3, -4 and -5 mutants to human chromosomes 19, 2, 16 and 13, respectively. The chromosomal assignment of the UV61-complementing gene is still unknown. Hence, it appears that CHO complementation groups 1 and 2 are corrected by a human gene on chromosome 19. Weber et al. (1987) have recently succeeded in the molecular cloning of a gene that confers UV resistance to UV-5 mutants (group 1). Molecular characterization of this gene, designated *ERCC-2*, revealed that it is clearly different from *ERCC-1* (Weber et al., 1987) indicating that 2 genes involved in excision repair are located on human chromosome 19. In this respect it is interesting to note that a DNA-repair gene associated with sister-chromatid exchange has also been assigned to human chromosome 19 (Siciliano et al., 1986). The conclusion of our transfection results, that *ERCC-1* can only correct the repair defect in CHO group-2 mutants is completely in accord with the chromosomal assignment data obtained by Thompson et al. (1987b).

Aspecific bypass of DNA-excision-repair defects has been observed after introduction of the T4 or *Micrococcus luteus* endonuclease in cell lines of all XP-complementation groups (Tanaka et al., 1977; De Jonge et al., 1985) and CHO mutants of group 1 (Valerie et al., 1985). The specificity of *ERCC-1* for group-2 mutants and the assignments of other *ERCC* genes to different human chromosomes render it unlikely that bypass activities account for the complementation analysis of UV-sensitive CHO mutants. Our results rather support the assumption that mutations in different repair genes underlie the classification of UV-sensitive CHO cell lines in different complementation groups and that *ERCC-1* is the mutated gene in CHO group-2 mutants.

Acknowledgements

We are grateful to Dr. N.G.J. Jaspers for helpful suggestions and a critical reading of the manuscript. Furthermore we thank Mrs. R.J. Boucke for typing the manuscript and Mr. M. Kuit for photography.

This work was supported by Euratom Contract No. BJ6-141-NL and MEDIGON, Foundation of Medical and Health Research in The Netherlands, Contract No. 900-501-091 and it was carried out under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract W-7405-ENG-48.

References

Burki, H.J., C.K. Lam and R.D. Wood (1980) UV-light induced mutations in synchronous CHO-cells, Mutation Res., 69, 347–356.

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

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

Fischer, E., W. Keijzer, H.W. Thielmann, O. Popanda, E. Bohnert, L. Edler, E.G. Jung and D. Bootsma (1985) A ninth complementation group in xeroderma pigmentosum, XP-I, Mutation Res., 145, 217–225.

Graham, F.L., and A.J. Van der Eb (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA, Virology, 52, 456–467.

Grosveld, F.G., T. Lund, E.J. Murray, A.L. Mellor, M.H.M. Dahl and R.A. Flavell (1982) The construction of cosmid libraries which can be used to transform eukaryotic cells, Nucl. Acids Res., 10, 6715–6732.

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.

Jeffreys, A.J., and R.A. Flavell (1977) A physical map of the DNA regions flanking the rabbit β-globin gene, Cell, 12, 429–439.

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.

Mulligan, R.C., and P. Berg (1981) Selection for animal cells that express the *Escherichia coli* gene coding for xanthine-guanine phosphoribosyltransferase, Proc. Natl. Acad. Sci. (U.S.A.), 78, 2072–2076.

Rigby, P.W.J., M. Dreckmann, C. Rhodes and P. Berg (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I, J. Mol. Biol., 113, 237–251.

Rubin, J.S., A.L. Joyner, A. Bernstein and G.F. Whitmore (1983) Molecular identification of a human DNA repair gene following DNA-mediated gene transfer, Nature (London), 306, 206–208.

Rubin, J.S., V.R. Prideaux, F.W. Huntington, A.M. Dulhanty, G.F. Whitmore and A. Bernstein (1985) Molecular cloning and chromosomal localization of DNA sequences associated with a human DNA repair gene, Mol. Cell. Biol., 5, 398–405.

Siciliano, M.J., A.V. Carrano and L.H. Thompson (1986) Assignment of a human DNA-repair gene associated with sister-chromatid exchange to chromosome 19, Mutation Res., 174, 303–308.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis, J. Mol. Biol., 98, 503–517.

Stefanini, M., A. Reuser and D. Bootsma (1982) Isolation of Chinese hamster ovary cells with reduced unscheduled DNA synthesis after UV-irradiation, Somatic Cell Genet., 8, 635–642.

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.

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

Thompson, L.H., and A.V. Carrano (1983) Analysis of mammalian cell mutagenesis and DNA repair using in vitro selected CHO cell mutants, in: E.C. Friedberg and B.R. Bridges (Eds.), Cellular Response to DNA Damage, UCLA Symposium on Molecular and Cellular Biology New Series, Vol. 11, Alan R. Liss, New York, pp. 125–143.

Thompson, L.H., S. Fong and K. Brookman (1980) Validation of conditions for efficient detection of HPRT and APRT mutations in suspension-cultured Chinese hamster ovary cells, Mutation Res., 74, 21–36.

Thompson, L.H., D.B. Busch, K.W. Brookman, C.L. Mooney and P.A. Glaser (1981) Genetic diversity of UV-sensitive DNA repair mutants of Chinese hamster ovary cells, Proc. Natl. Acad. Sci. (U.S.A.) 78, 3734–3737.

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.

Thompson, L.H., C.L. Mooney and K.W. Brookman (1985a) Genetic complementation between UV-sensitive CHO mutants and xeroderma pigmentosum fibroblasts, Mutation Res., 150, 423–429.

Thompson, L.H., C.L. Mooney, K. Burkhart-Schultz, A.V. Carrano and M.J. Siciliano (1985b) Correction of a nucleotide excision repair mutation by human chromosome 19 in hamster–human hybrid cells, Somat. Cell. Mol. Genet., 11, 87–92.

Thompson, L.H., E.P. Salazar, K.W. Brookman, C.C. Collins, S.A. Stewart, D.B. Busch and C.A. Weber (1987a) Recent progress with the DNA repair mutants of Chinese hamster ovary cells, J. Cell Sci., Suppl. 6, 97–110.

Thompson, L.H., A.V. Carrano, K. Sato, E.P. Salazar, B.F. White, S.A. Stewart, J.L. Minkler and M.J. Siciliano (1987b) Identification of nucleotide excision repair genes on human chromosomes 2 and 13 by functional complementation in hamster–human hybrids, Submitted for publication.

Valerie, K., J.K. de Riel and E.E. Henderson (1985) Genetic complementation of UV-induced DNA repair in Chinese hamster ovary cells by the *den*V gene of phage T4, Proc. Natl. Acad. Sci. (U.S.A.), 82, 7656–7660.

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.

Weber, C.A., E.P. Salazar, S.A. Stewart and L.H. Thompson (1987) Molecular cloning and biological characterization of a human gene, *ERCC-2*, which corrects the nucleotide excision repair defect in CHO UV5 cells, Mol. Cell. Biol., in press.

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.

Wood, R.D., and H.J. Burki (1982) Repair capability and the cellular age response for killing and mutation induction after UV, Mutation Res., 95, 505–514.

Zdzienicka, M.Z., L. Roza, A. Westerveld, D. Bootsma and J.W.I.M. Simons (1987) Biological and biochemical consequences of the human *ERCC-1* repair gene after transfection into a repair deficient CHO cell line, Mutation Res., 183, 69–74.

APPENDIX PAPER VI

Mutation Research (in press)

THE CLONED HUMAN DNA EXCISION REPAIR GENE *ERCC-1* FAILS TO CORRECT XERODERMA PIGMENTOSUM COMPLEMENTATION GROUPS A THROUGH I

M. van Duin, G. Vredeveldt, L.V. Mayne*, H. Odijk, W. Vermeulen, B. Klein**, G. Weeda**, J.H.J. Hoeijmakers, D.Bootsma and A. Westerveld.

Department of Cell Biology and Genetics, Erasmus University Rotterdam, PO box 1738, 3000 DR Rotterdam, The Netherlands.

* MRC Cell Mutation Unit (present address:Sussex Centre for Medical Research), University of Sussex, Falmer, Brighton BN1 9RR, UK.

** Department of Medical Biochemistry, University of Leiden, 2333 AL Leiden, The Netherlands.

ABSTRACT

The human DNA excision repair gene *ERCC-1* complements the ultraviolet light (UV) and mitomycin C (MMC) sensitivity of CHO mutants of complementation group 1. We have investigated whether *ERCC-1* is the mutated gene in cell lines from xeroderma pigmentosum (XP) complementation groups A through I by analyzing the endogenous gene in XP cells and by introduction of the gene followed by repair assays. Our studies show that *ERCC-1* is not deleted or grossly rearranged in representative cell lines of nine XP groups. Furthermore, Northern blot analysis revealed correct transcription of *ERCC-1* in all groups. The cloned human *ERCC-1* gene was introduced into immortalized XP cells by DNA transfection (groups A,C,D,E and F). The presence of the integrated transfected sequences was verified on Southern blots and by selection for two dominant marker genes that flank the *ERCC-1* gene on the transfected cos43-34 DNA. *ERCC-1* failed to confer a normal UV-survival and UV-induced unscheduled DNA synthesis (UDS) to transfected populations. In the case of the remaining XP complementation groups (B, G, H and I), nuclear microïnjection was used to introduce an *ERCC-1* cDNA construct driven by an SV40 promoter into primary fibroblasts. Coinjection of the SV40 large T gene and analysis of its expression served as a control for the injection. The *ERCC-1* cDNA failed to induce increased levels of UDS in the microinjected fibroblasts. We infer from these experiments that *ERCC-1* is not the mutated gene in the nine XP complementation groups examined. From similar type of experiments we conclude that *ERCC-1* is not the defective gene in UV-sensitive Cockayne's Syndrome cells.

INTRODUCTION

Cell fusion experiments have so far identified 7 complementation groups of UV-sensitive CHO-mutants (Thompson et al. 1981; 1987; Zdzienicka et al. 1988). Cell lines of groups 1

and 2 have been exploited successfully in DNA mediated gene transfer experiments aimed at the isolation of complementing human DNA repair genes. This has resulted in the cloning of human *ERCC-1* and *ERCC-2* genes (Westerveld et al. 1984; Van Duin et al. 1986; Weber et al. 1988). *ERCC-1* and *ERCC-2* specifically correct the mutant phenotype of CHO group 1 and group 2 [1] mutants respectively, indicating that both genes are the human homologues of the mutated loci in the corresponding CHO complementation groups (Van Duin et al. 1988a; Weber et al. 1988). In addition to an enhanced sensitivity to UV, CHO group 1 mutants display an extreme sensitivity to MMC and other bifunctional alkylating agents (Hoy et al. 1985). A defective incision step of the excision of DNA damage is thought to underlie the phenotype of CHO group 1 mutants (Thompson et al. 1982). The cloned *ERCC-1* gene corrects the impaired excision repair of both UV and MMC induced DNA-adducts (Westerveld et al. 1984). Furthermore, cross-sensitivities of group 1 mutants to other chemical agents are corrected by the cloned gene (Zdzienicka et al. 1987). The phenotype of CHO group 1 mutants is reminiscent of the rare human autosomal recessive disorder xeroderma pigmentosum (XP). Patients suffering from this trait are hypersensitive to sunlight exposure and have an elevated risk for skin tumor development (Kraemer, 1983). Nine XP complementation groups have been identified (Fischer et al. 1985) that are all characterized by a defective incision step of the DNA excision repair pathway. Interspecies complementation studies with UV-sensitive human and rodent DNA repair mutants, to determine whether they belong to identical complementation groups are incomplete and have not yet identified overlap. Stefanini et al. (1987) recently reported transient complementation of the repair defect in heterokaryons of 6 XP groups after fusion with UV sensitive CHO12RO cells. However, this CHO mutant has not been assigned to any of the known CHO complementation groups. Thompson et al. (1985) have attempted complementation analysis between two CHO groups (1 and 4) and XP fibroblasts based on recovery of stable hybrid cells resistant to DNA damage. The low frequency of hybrid formation did not permit conclusions with respect to all XP groups and crosses of both CHO groups with XP-A, C, D, F, and G yielded only partial complementation (Thompson et al. 1985). Therefore the question whether identical genes are affected in CHO and human repair mutants can probably be more efficiently addressed when the isolated genes are available. The phenotypic similarity between CHO group 1 and human XP cells renders *ERCC-1* a potential candidate gene which could account for one of the genetic defects in this repair syndrome. In this report we have investigated this possibility. We conclude from our experiments that *ERCC-1* can most likely be excluded as the mutated gene in the XP complementation groups examined.

MATERIALS AND METHODS

Cell lines and media

The cell lines used in the experiments are listed in Table 1. The immortalized cells were cultured in DMEM/F10 medium supplemented with fetal calf serum (FCS) and antibiotics

[1] At the recent UCLA meeting on 'Mechanisms and consequences of DNA damage processing' in Taos (January 1988) it was decided to rename CHO complementation groups 1 and 2 to 2 and 1 respectively in order to match the numbering with the complementing *ERCC* genes.

Table 1

Cell lines used in the different experiments of this study.

Cell line*)	Experiments**)				References/Source
HeLa	S	N			
XP12RO-SV (A)	S	N	T		Veldhuizen (unpubl.)
XP2OS-SV (A)			T		Takebe et al. 1974
XP12BE-SV (A)	S				Camden cell bank
XP11BE (B)	S	N	M		Kreamer et al. 1975
XP4PA-SV (C)	S				Daya-Grosjean et al.1987
XP8CA-SV (C)		N			Klein (unpubl.)
CW12 (C)			T		Wood et al. 1987
MH3XP (D)		N			Wood et al. 1987
XP1BR (D)	S				Keijzer (unpubl.)
HD2 (D)	S		T		Johnson et al. 1985
CW3 (E)		N	T		Wood et al. 1987
XP2RO (E)	S				De weerd-Kastelein etal. 1973
XP2YO-SV (F)		N	T		Yagi and Takebe, 1983
XP2BI (G)				M	Keijzer et al. 1979
XP3BR (G)	S				Arlett et al. 1980
XP3BR-SV (G)		N	T		Klein (unpubl.)
XP2CS (H)	S	N		M	Moshell et al. 1983
XP3MA (I)	S			M	Fisher et al. 1985
XP20MA (I)		N			Keijzer (unpubl.)

*) Between parentheses the XP complementation group is given.

**) The different cell lines were used in different experiments: Southern analysis (S), Northern analysis (N), Transfection (T) and/or Microinjection (M).

(penicillin, streptomycin). Primary fibroblasts were cultured in F10 medium with FCS and antibiotics. All cells were grown as monolayers in a 37 ^{o}C incubator.

DNA transfection and cloning efficiency of transformants.

On cosmid 43-34 the *ERCC-1* is flanked by the selectable markers genes aminoglycosyl phosphotransferase (*agpt*) and the *E.coli* guanine phosphoribosyl transferase (*Ecogpt*) which allow selection for G418 and mycophenolic acid (MPA) respectively. Cosmid 43-34 DNA was transfected as a calcium phosphate precipitate into the different cell lines as previously described (Westerveld et al. 1984). One to two days after transfection selection was started by adding G418 to the medium. The concentration of G418 was dependent on the transfected cell line and ranged from 400 μg/ml to 800 μg/ml medium). After 10-14 days clones were counted and either isolated or grown into mass populations (MP's). In order to establish the cointegration of *Ecogpt* sequences in G418 resistant mp's, for each MP the cloning efficiency in medium with or without G418 and/or mycophenolic acid (MPA) was determined. For each cell line three independent MP's were tested in threefold experiments. For XP3BR-SVmp 2000 cells per dish were plated and for the other

tested lines 400-500 cells per dish were plated. In Table 2 the average relative cloning efficiency is given.

Southern blotting
Using routine procedures (Maniatis et al. 1982) genomic DNA from cell lines was digested with restriction enzymes and after size fractionation on agarose gels blotted to nitrocellulose filters which were subsequently hybridized to ^{32}P-labeled probes.

Northern blotting
Total RNA was isolated from cultured cells by the LiCl/Urea procedure (Auffray and Rougeon, 1980) and size fractionated in 1.2% agarose gels containing formaldehyde. RNA was then transferred to nitrocellulose and hybridized in 50% formamide to *ERCC-1* probe EPv12 using standard procedures (Maniatis et al. 1982).

UV-survival and unscheduled DNA synthesis (UDS)
The response of G418 resistant XP-transformants to UV was measured by UV survival and UV-induced UDS as reported previously (Van Duin et al. 1988a).

Microinjection of ERCC-1 cDNA into XP primary fibroblasts.
ERCC-1 cDNA construct pcDE (van Duin et al. 1986) and pSV3gpt (Mulligan and Berg, 1981) were mixed in a 1 : 1 molar weight ratio (0.15 $\mu g/\mu l$, final concentration) and coinjected into nuclei of XP homopolykaryons using procedures previously described (De Jonge et al. 1983). At 24 and 48 hrs post- injection one part of the injected cells was stained for the expression of SV40 large T antigen encoded by pSV3gpt using a monoclonal antibody and an FITC containing goat anti mouse conjugate. In order to determine the effect on the repair potential the other half of the injected cells was at 24 and 48 hours post-injection irradiated with UV (15 J/m^2) and incubated for two hours in medium containing ^{3}H-thymidine. Subsequently, UV-induced DNA synthesis was determined by autoradiography as described previously (Vermeulen et al. 1986).

RESULTS

Analysis of the endogenous ERCC-1 gene in XP.
To examine whether XP cells from all complementation groups harbor intact *ERCC-1* gene copies, the endogenous *ERCC-1* gene was analysed by Southern blotting experiments. High molecular weight DNA was isolated from each of the cell lines and after digestion with various restriction endonucleases size fractionated on agarose gels and transferred to nitrocellulose filters that were subsequently hybridized to a^{32}P-labeled *ERCC-1* cDNA probe. DNA from HeLa cells was used as a control since the *ERCC-1* gene was originally isolated from this cell line (Westerveld et al. 1984).

The autoradiogram of a blot with TaqI digested XP DNAs is shown in Fig.1. All cell strains examined showed hybridization patterns similar to the HeLa control. Similar results were obtained with PvuII digested DNAs (not shown). Therefore, as far as can be detected by this method, we conclude that *ERCC-1* is not deleted or grossly rearranged in the representative cell strains of the nine XP complementation groups examined.

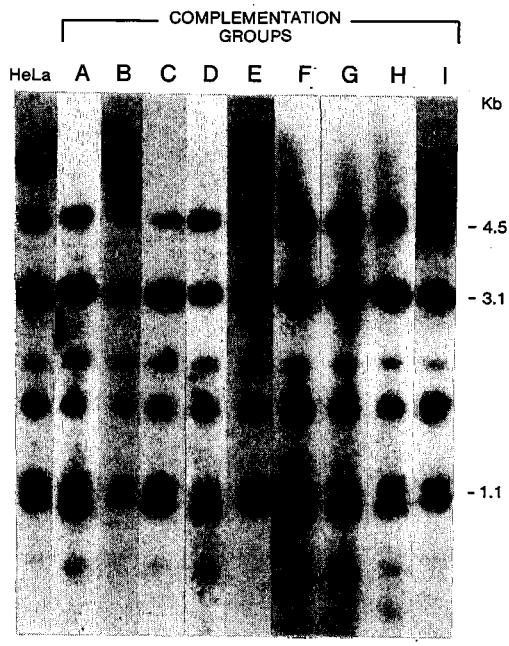

Figure 1
Southern blot analysis of Taq1 digested DNA of all nine XP complementation groups (A to I). HeLa DNA served as a control in this experiment. The blot was hybridized to *ERCC-1* cDNA probe EPv12 which recognizes all 10 *ERCC-1* exons. The molecular weight of the hybridizing bands is indicated on the right.

To investigate the expression of the endogenous *ERCC-1* gene in XP cells, total RNA was isolated and analyzed by Northern blot experiments using *ERCC-1* cDNA as a probe.

The autoradiogram of this experiment is shown in Fig. 2. All nine XP groups examined produced a correct *ERCC-1* transcript similar to that of HeLa cells.

Transfection of ERCC-1 into immortalized XP cells.
The Southern and Northern blot experiments presented above do not rule out the possibility that the *ERCC-1* gene is inactivated by point mutations or other small alterations. Therefore the cloned functional *ERCC-1* gene was examined for its ability to correct the repair defect in XP cells. To that aim the cloned *ERCC-1* gene located on cosmid 43-34 (cos43-34), was transfected into established XP cells of 6 different complementation groups (see Materials and Methods). Before transfection cos43-34 DNA was linearized with ClaI which generates 50 kb molecules in which the 5' and 3' *ERCC-1* ends are flanked by the dominant marker genes *Ecogpt* and *agpt* at a distance of approximately 10 and 4 kb of genomic DNA respectively (see Fig. 3). The transfected cells were selected in medium

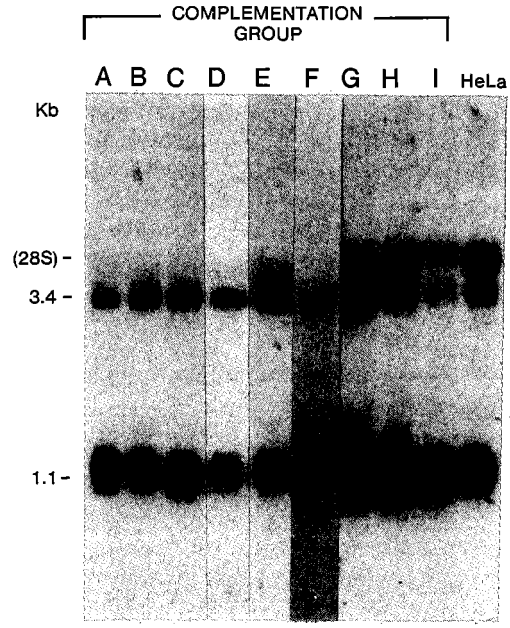

Figure 2
Northern blot analysis of representative cell lines of all nine XP complementation groups (A to I) and HeLa as a control. Per lane 15-30 μg total RNA was loaded. The EPv12 probe described in Fig.1 was used for hybridization in 50 % formamide at 42 ^{o}C. Molecular weight of hybridizing bands is indicated on the left. In some lanes background labeling of ribosomal RNA (28S) is visible due to longer exposure times of the autoradiogram.

Figure 3
Structure of cosmid 43-34. The position of the *ERCC-1* gene and *Ecogpt* and *agpt* dominant markers are shown in black. The open parts of the left and right hand boxes are pBR322 sequences. Only restriction sites are shown that are relevant for the Southern blot analysis in Fig. 4. Details of cosmid 43-34 have been described previously (Westerveld et al. 1984; Van Duin et al. 1987).

containing G418 and after two weeks mass populations (mp's) were generated which consisted of 20-200 independent transformants. In general, transfection efficiencies ranged from 10^{-4} to 10^{-5}. The G418 resistant transformants were subsequently examined for the presence of transfected *ERCC-1* DNA. Since on cos43-34 *ERCC-1* is flanked by the *agpt* and *Ecogpt* markers, cross resistance of G418 resistant cells to mycophenolic acid (MPA), provided by the *Ecogpt* gene, would be indicative of the presence of *ERCC-1* sequences. The presence of the *Ecogpt* marker was determined by measuring the cloning efficiency of G418 resistant cos43-34 transformants in medium containing G418 and MPA. The results presented in Table 2 indicate that XP-C, D and E transformants had integrated both bacterial markers with a high efficiency. No MPA resistant XP-G transformants were obtained. In an earlier report we showed that most of the XP-A and XP-F transformants (60%-90%) had cointegrated *Ecogpt* sequences which was confirmed by Southern blot analysis (Hoeijmakers et al. 1987). The results of Southern blot analysis of the XP-C, D and E transformants are presented in Fig. 4.

Table 2.
Cloning efficiency of G418 resistant cos43-34 transformants (mp's) in normal medium and medium supplemented with G418, MPA or G418 and MPA expressed as percentage of seeded cells which formed colonies (see Materials and Methods).Each experiment was carried out in threefold. The average values for three independent mp's are given.

	Cloning efficiency (%) in different media			
Cell line*	normal	+G418	+MPA	+G418/MPA
CW12mp (C)	35	34	28	27
HD2mp (D)	44	36	31	26
CW3mp (E)	27	24	25	20
XP3BR-SVmp (G)	5	5	0	0

* XP complementation group is given in parentheses.

To discriminate between cos43-34 derived *ERCC-1* sequences and the endogenous gene the DNAs were digested with HindIII/XhoI and BglII which generates cos 43-34 specific *ERCC-1* bands when hybridized to 5' and 3' *ERCC- 1* probes respectively (see also Fig.1). From a comparison between the signals of the endogenous and transfected *ERCC-1* gene it can be inferred that all transformed mass population had integrated multiple copies of *ERCC-1* into their genome. The HD2 mass population has integrated considerably more DNA than CW12 cells in which much more degradation of the transfected DNA seems to have taken place (see Fig. 4). In contrast, intergrated *agpt* sequences but not *ERCC-1* DNA or *Ecogpt* dominant marker could be detected in G418 resistant XP-G transformants (not shown) which is in agreement with the finding that these cells did not show cross resistance to MPA.

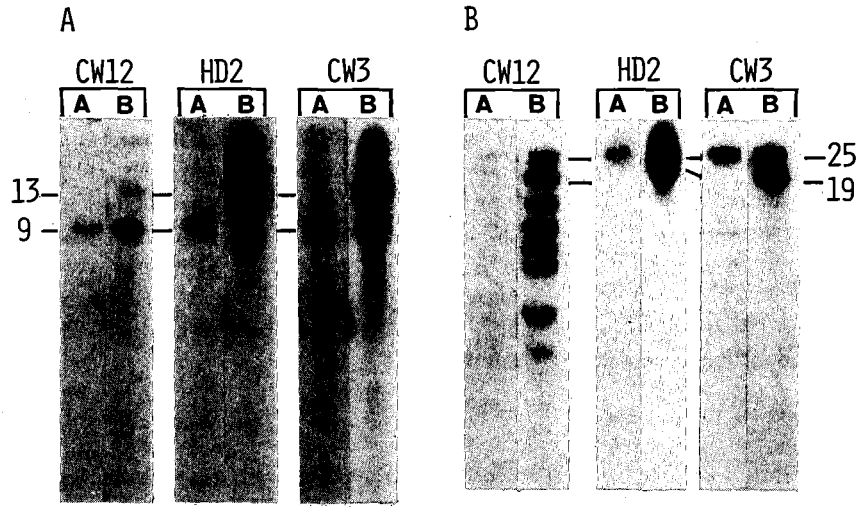

Figure 4

Southern blot analysis of G418 resistant CW12 (XP-C), HD2 (XP- D) and CW3 (XP-E) cells that have been transfected with cos43- 34. In panel A DNA of recipient (lanes A) and transfected cells (lanes B) were digested with BglII and hybridized to probe 4A4 located 3' of *ERCC-1* (see Fig. 1). Panel B shows HindIII/XhoI double digested DNA probed with *ERCC-1* cDNA probe EPs12 harboring the first 5 exons. On left and right side fragment sizes (kb) are indicated.

Response of XP cos43-34 transformants to DNA damage.

To test whether *ERCC-1* restores normal DNA repair characteristics in XP cells we determined UV-survival and UV- induced unscheduled DNA synthesis (UDS) of the G418 resistant mass populations. UV-survival plots are shown in Fig. 5 and UDS data are summarized in Table 3.

All transformed XP cultures (XP-A and XP-F not shown) harboring transfected *ERCC-1* DNA displayed mutant UV survival phenotypes. The UDS levels of the transformants were comparable with those found in the various mutants cell lines and were in agreement with the survival data. With respect to the relatively high UDS values of the XP- E parental cells and transformants it should be noted that the residual UDS level of primary XP-E fibroblasts is up to 50% of that of repair proficient cells. (De Weerd-Kastelein et al. 1973; Fujiwara et al. 1985). Also HD2 and XP2YO-SV (XP-D and F respectively) cells display higher UDS values than the original non transformed diploid XP fibroblast cultures (Johnson et al. 1985; Yagi and Takebe, 1983).

Table 3
UV induced UDS of immortalized XP cell lines and G418 resistant mass populations (mp).

	UDS	
Cell line*	grains **	% of HeLa#
XP12RO-SV (A)	5±2	13
XP12RO-SVmp (A)	5±1	12
CW12 (C)	7±1	23
CW12mp (C)	7±1	23
HD2 (D)	16±1	47
HD2mp (D)	15±1	43
CW3 (E)	26±1	88
CW3mp (E)	25±1	86
XP2YO-SV (F)	5±1	13
XP2YO-SVmp (F)	5±1	13

* XP complementation group is given in parentheses.

** No. of grains (±SEM) per fixed square of non S-phase nuclei.

\# Percentage of UDS of HeLa, which was determined in each independent experiment.

Microinjection of ERCC-1 cDNA.
XP complementation groups of which no immortalized cell lines suitable for transfection are available, were examined for *ERCC-1* correction by means of a microinjection assay. *ERCC-1* cDNA construct pcDE (van Duin et al. 1986), harboring a functional *ERCC-1* cDNA driven by the SV40 early promoter, was injected into nuclei of XP homopolykaryons. After 24 and 48 hours the levels of UV induced DNA synthesis in injected compared to non-injected cells were measured. In previous reports we have shown that fibroblasts of the examined complementation groups can be corrected by microinjection of crude extracts of repair proficient cells or heterologous XP cells. As a positive internal control for the nuclear injection pcDE was coinjected with pSV3gpt carrying a functional SV40 large T antigen gene which is under the same promoter and termination sequences as the *ERCC-1* cDNA in pcDE. Large T expression was determined by immunofluorescent staining. It was found that 30-50% of injected cells exhibited high level expression of the injected large T gene. It is therefore reasonable to assume that the coinjected *ERCC-1* cDNA is correctly expressed in approximately the same fraction of cells. XP-B, G, H and I homopolykaryons were subjected to this analysis, however the introduction of the *ERCC- 1* cDNA did not have a significant positive effect on UDS levels in any of these complementation groups (data not shown).

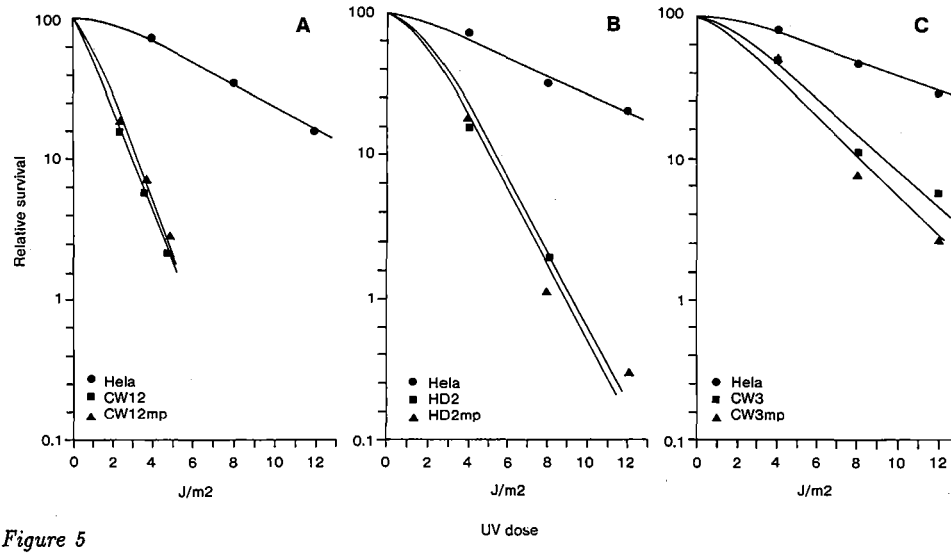

Figure 5

UV-survival plots of XP-C (CW12), XP-D (HD2) and XP-E (CW3) cells and cos43-34 transformed G418 resistant mass populations. HeLa cells served as wild type controls.

DISCUSSION

The human *ERCC-1* gene was isolated by virtue of its ability to complement the DNA repair defect in CHO complementation group 1 cells. These CHO cells are sensitive to UV and MMC, indicating a role for *ERCC-1* in the excision repair of UV and MMC induced DNA damage. We have investigated the possibility that *ERCC-1* is the mutated gene in UV-sensitive XP cells of complementation groups A through I.

Cross-sensitivity to UV and MMC might be used to identify the human counterpart of CHO group 1 mutants. However, for many cell lines of human DNA repair syndromes the response to both agents is poorly documented and often limited to specific complementation groups. Fujiwara and coworkers have reported a normal response of XP-A cells to MMC (Fujiwara et al. 1977; Fujiwara, 1982) and also XP-C (Nagasawa and Little, 1983) and XP-F were reported to have wild type repair of MMC induced DNA damage (Plooy et al. 1985). As far as we know for the other complementation groups cross-sensitivity to MMC is not reported in the literature. The biochemistry of repair of UV induced DNA damage - primarily pyrimidine dimers - and MMC induced DNA cross-links is unknown. The phenotype of CHO group 1 cells suggest that *ERCC-1* is involved in a common (pre)incision step of the removal of both types of DNA lesions. On the other hand, it can not be excluded that *ERCC-1* has a dual function and that different mutations in *ERCC-1* can affect the response to either UV or MMC. The *ERCC-1* protein was found to have a 'mosaic' structure displaying homology to the yeast *RAD10* and parts of the E.coli *uvrA* and *uvrC* repair proteins (Van Duin et al. 1986; Doolittle et al. 1986; Hoeijmakers et al. 1986; Van Duin et al. 1988b). This could imply that during evolution multiple

repair functions have been accommodated in the *ERCC-1* protein. Consequently, it is possible that some mutations in the *ERCC-1* gene might yield only a UV and not a MMC sensitive phenotype.

Southern and Northern blot analysis with *ERCC-1* probes did not give indications of altered *ERCC-1* genes in cell lines from all nine XP complementation groups. These experiments, however, do not exclude point mutations or minor deletions in genomic *ERCC- 1* sequences and do not allow any conclusion on the functionality of the *ERCC-1* protein in the examined XP cells. Introduction of the cloned *ERCC-1* gene by DNA transfection is limited to immortalized cell lines. The experiments with XP3BR- SV (XP-G) cells show that even SV40 transformed lines are not always suitable for efficient DNA transfer. Although G418 resistant clones were obtained after transfection of cos43-34 no cross-resistance to MPA was found indicating that these cells had integrated very small amounts of DNA, which was confirmed by Southern blot analysis (not shown). These results are in accordance with our earlier observations that the quantity of integrated exogenous DNA is cell line specific and that human cells in general take up small amounts of DNA compared to some rodent cell lines (Hoeijmakers et al. 1987; Mayne et al. 1988a).

The *ERCC-1* gene was successfully transfected to XP-A, C, D, E and F cells whereas nuclear microinjection was used to introduce the *ERCC-1* cDNA in XP-B, G, H and I fibroblasts. The presence of transfected *ERCC-1* sequences was confirmed by Southern blot analysis and double selection on G418 and MPA. Simultaneous injection of plasmid pSV3gpt expressing SV40 large T antigen indicated that in 30-50% of the injected cells the introduced *ERCC-1* cDNA should be properly expressed. *ERCC-1* had no effect on UV survival and UDS of mass populations of 20-200 independent G418 resistant transformants. Likewise in the transient microinjection assay *ERCC-1* cDNA did not elevate UV induced DNA repair synthesis.

Although we have not demonstrated that the transfected and microinjected *ERCC-1* gene is correctly expressed into functional proteins, we believe that our negative results are significant. Firstly, the *ERCC-1* constructs used in the experiments are able to correct the defect in CHO group 1 mutants with a very high efficiency (Westerveld et al. 1984; Van Duin et al. 1986) indicating that they encode functional *ERCC-1* protein and are correctly expressed when introduced into mammalian cells. Secondly, control experiments (i.e cross- resistance to mpa, Southern blot analysis of the transfected human gene and coinjection with plasmid pSV3gpt) verified that intact *ERCC-1* copies were present in at least a detectable fraction of the various XP cells into which *ERCC-1* was introduced. Therefore, we anticipate that correction of any of the XP groups by *ERCC-1* when present would not have escaped our attention. Hence, it is unlikely that mutations in the *ERCC-1* gene will account for this complex human genetic disorder. Thompson and coworkers have reported that fusion of CHO group 1 cells, harboring a mutated *ERCC-1* gene, with XP cells of various complementation groups yields (partial) correction of the mutual repair defect, strongly suggesting that these XP groups are not defective in *ERCC-1* (Thompson et al. 1985). This is in agreement with our conclusion. Cell lines from patients with Cockayne's syndrome and Fanconi's anemia display a cellular hypersensitivity to UV and MMC respectively (Kraemer, 1983). Nothing is known about the DNA repair defect in FA. For CS complementation groups A and B it has recently been demonstrated that these cells are deficient in the preferential repair of active genes(Mayne et al. 1988b), which

explains their characteristic absence of RNA synthesis recovery after UV exposure (Mayne and Lehmann, 1982). Southern blot analysis indicated normal *ERCC-1* copies in all three CS complementation groups and both FA groups (not shown). Furthermore, transfection of *ERCC-1* to CS-A and CS-B cells did not correct the defect in RNA synthesis recovery and survival after UV irradiation (not shown).

The experiments thus far show that a defective *ERCC-1* gene is most likely not represented in patients from nine XP excision deficient complementation groups. This implies that at least 10 genes (9 XP-loci and *ERCC-1*) are involved in the incision step of the excision repair mechanism further stressing the biochemical complexity of this repair process. Furthermore, the fact that an *ERCC-1* defect is not found in XP and possibly also not in CS and FA suggest that mutations in the *ERCC-1* gene may be incompatible with life. Alternatively it is not excluded that patients with a defect in *ERCC-1* exist but have not been identified yet.

ACKNOWLEDGEMENTS

We thank Drs. R.T. Johnson, R.E. Moses, A.Sarasin and H. Takebe for generously providing us with the different immortalized XP and FA cells, N. van Damme, J. de Geus, J. van den Tol and R.C.A van Ham for technical assistance and C.van Dijk and M. Kuit for help in preparing the manuscript. This work was financially supported by MEDIGON, Foundation of Medical and Health Research in the Netherlands, contract no. 900-501-091 and EURATOM, contract no. BJ6-141-NL.

REFERENCES

Arlett, C.F. (1980) Mutagenesis in repair deficient human cell strains. In: M. Alacevic (ed.), Progress in environmental mutagenesis, Amsterdam, Elsevier, 162-174.

Auffray, C. and Rougeon, F. (1980) Purification of mouse immunoglobuline heavy-chain messenger RNA's from total myeloma tumor RNA. Eur. J. Biochem. 107, 303-314.

Daya-Grosjean, L., M.R. James, C. Drougard and A. Sarasin (1987) An immortalized xeroderma pigmentosum, group C, cell line which replicates SV40 shuttle vectors. Mutat. Res. 183, 185-196.

De Jonge, A.J.R., W. Vermeulen, B. Klein and J.H.J. Hoeijmakers (1983) Microinjection of human cell extracts corrects xeroderma pigmentosum defect. EMBO J. 2, 637-641.

De Weerd-Kastelein, E.A., W. Keijzer and D. Bootsma (1973) A third complementation group in xeroderma pigmentosum. Mutat. Res. 22, 87-91.

Doolittle, R.F., M.S. Johnson, J. Husain, B. Van Houten, D.C. Thomas and A. Sancar (1986). Domainal evolution of a prokaryotic DNA repair protein and its relationship to active transport proteins. Nature 323, 451-453.

Fischer, E., W. Keijzer, H.W. Thielmann, O. Popanda, E. Bohnert, L. Edler, E.G. Jung and D. Bootsma (1985). A ninth complementation group in xeroderma pigmentosum, XP-I. Mutat.Res. 145, 217-225.

Fujiwara, Y., M. Tatsumi and M.S. Sasaki (1977) Cross-link repair in human cells and its possible defect in Fanconi's anemia cells. J. Mol. Biol. 113, 635-649.

Fujiwara, Y. (1982) Defective repair of mitomycin-C crosslinks in Fanconi's anemia and loss in confluent normal human and xeroderma pigmentosum cells. Biochim. Biophys. Acta 699, 217- 225.

Fujiwara, Y., Y. Uehara, M. Ishihashi, Y. Yamamoto and K. Nishioka (1985) Assignment of 2 patients with xeroderma pigmentosum to complementation group E. Mutat. Res. 145, 55- 61.

Hoeijmakers, J.H.J., M. Van Duin, A. Westerveld, A. Yasui and D. Bootsma (1986). Identification of DNA repair genes in the human genome. Cold Spring Harbor Symposia on Quantitative Biology vol. LI, 91-101.

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.

Hoy, C.A., L.H. Thompson, C.L. Mooney and E.P. Salazar (1985) Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. Cancer Res. 45, 1737-1743.

Johnson, R.T., S. Squires, G.C. Elliott, G.L.E. Koch and A.J. Rainbow (1985) Xeroderma pigmentosum D - HeLa hybrids with low and high ultraviolet sensitivity associated with normal and diminished DNA repair ability, respectively. J. Cell. Sci. 76, 115-133.

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

Kraemer, K.H., E.A. de Weerd-Kastelein, J.H. Robbins., W. Keijzer, S.F. Barrett, R.A. Pettinga and D. Bootsma (1975) Five complementation groups in xeroderma pigmentosum. Mutat. Res. 33, 327-340.

Kraemer, K.H. (1983) Heritable diseases with increased sensitivity to cellular injury, In: Fitzpatrick,T.B, Eisen, A.Z., Wolff, K., Freedberg, I. M. and Austen, K.F. (Eds.), Update: Dermatology in General Medicine, McGraw-Hill, New York, 113-141. Maniatis, T., E.F. Fritsch and J. Sambrook (1982). In: Molecular cloning, A Laboratory Manual (Cold Spring Harbor Laboratory) Cold Spring Harbor, New York.

Mayne, L.V. and A.R. Lehmann (1982). Failure of RNA synthesis to recover after UV-irradiation; An aerly defect in cells from individuals with Cockayne's syndrome and xeroderma pigmentosum. Cancer Res. 42, 1473-1478.

Mayne, L.V., T. Jones, S.W. Dean, S.A. Harcourt, J.E. Lowe, A. Priestley, H. Steingrimsdottir, Sykes, H., Green, M.H.L. and Lehmann, A.R. (1988a) SV40-transformed normal and DNA-repair- deficient human fibroblasts can be transfected with high frequency but

retain only limited amounts of integrated DNA. Gene 66, 65-76.

Mayne, L.V., L.H.F. Mullenders and A.A.van Zeeland, A. (1988b) Cockayne's Syndrome: A UV-sensitive disorder with a defect in the repair of transcribing DNA but normal overall excision repair. In: Mechanisms and Consequences of DNA damage processing. (Friedberg, E.C. and Hanawalt, P., eds.) UCLA symposia on Molecular and Cellular Biology, New Series Vol. 83 (in press).

Moshell, W.N.,M.B. Ganges, M.A. Lutzner, H.G. Coon, S.F. Barett, J.M. Dupuy and J.H. Robbins (1983) A new patient with both xeroderma pigmentosum and Cockayne's syndrome establishes the new xeroderma pigmentosum complementation group H. In: Friedberg, E., Bridges, B. (Eds): Cellular response to DNA damage, Liss, New York , 209-213.

Mulligan, R.C. and P. Berg (1981) Selection for animal cells that express the *Escherichia coli* gene coding for xanthine- guanine phosphoribosyltransferase. Proc. Natl. Acad. Sci. USA 78, 2072-2076.

Nagasawa, H. and B. Little (1983) Suppression of cytotoxic effect of mitomycin-C by superoxide dismutase in Fanconi's anemia and dyskeratosis congenita fibroblasts. Carcinogen. 4, 795-799.

Plooy, A.C.M., M. Van Dijk, F. Berends and P.H.M Lohman (1985) Formation and repair of DNA interstrand cross-links in relation to cytotoxicity and unscheduled DNA synthesis induced in control and mutant human cells treated with *cis*- diamminedichloroplatinum(II). Cancer Res. 45, 4178-4184.

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.

Thompson, L.H., D.B. Busch, K.W. Brookman, C.L. Mooney and P.A Glaser (1981). Genetic diversity of UV-sensitive DNA repair mutants of Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA 78, 3734-3737.

Thompson, L.H., K.W. Brookman, L.E. Dillehay, C.L.Mooney and A.V. Carrano (1982) Hypersensitivity to mutation and sister- chromatide-exchange induction in CHO cell mutants defective in incising DNA containing UV lesions. Somat. Cell Genet. 8, 759- 773.

Thompson, L.H., C.L. Mooney and K.W. Brookman (1985). Genetic complementation between UV-sensitive CHO mutants and xeroderma pigmentosum fibroblasts. Mutat. Res. 150, 423-429.

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 the DNA repair mutants of Chinese hamster ovary cells. J. Cell Sci. Suppl. 6, 97-110.

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.

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 excision repair gene *ERCC- 1*. Nucl. Acids Res. 22, 9195-9213.

Van Duin, M., J.H Janssen, J. de Wit, J.H.J. Hoeijmakers, L.H. Thompson, D. Bootsma and A. Westerveld (1988a) Transfection of the cloned human excision repair gene *ERCC-1* to UV- sensitive CHO mutants only corrects the repair defect in complementation group 2 mutants. Mutat. Res. 193, 123-130.

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

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

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.

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

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

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

Zdzienicka, M.Z., L. Rosa, A. Westerveld, D. Bootsma and J.W.I.M. Simons (1987) Biological and biocemical consequences of the human *ERCC-1* repair gene after transfection into a repair deficient CHO cell line. Mutat. Res. 183, 69-74.

Zdienicka, 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. (in press).

APPENDIX PAPER VII

Submitted for publication

CONSERVED PATTERN OF ANTISENSE OVERLAPPING TRANSCRIPTION IN THE HOMOLOGOUS HUMAN *ERCC-1* AND YEAST *RAD10* DNA REPAIR GENE REGIONS.

M. van Duin, J. van den Tol, J.H.J. Hoeijmakers, D. Bootsma, I.P. Rupp*, P. Reynolds*, L. Prakash* and S. Prakash#.

Dept. of Cell Biology and Genetics, Erasmus University, PO box 1738, 3000 DR Rotterdam, The Netherlands.

* Dept. of Biophysics, University of Rochester, School of Medicine, Rochester, NY 14642, USA.

\# Dept. of Biology, University of Rochester, Rochester, NY 14627, USA

Abstract

Here we report that the genes for the homologous yeast *RAD10* and human *ERCC-1* DNA excision repair proteins harbor overlapping antisense transcription units in their 3' regions. Since naturally occurring antisense transcription is rare in yeast and man, this being the first example in human cells, our observation indicates that antisense transcription in *ERCC- 1/RAD10* gene regions represents an evolutionarily conserved feature.

The DNA excision repair pathway is one of the major repair systems in the cell. It counteracts the mutagenic and carcinogenic effects of DNA lesions. We have recently reported that the first isolated human repair gene *ERCC-1* displays significant homology to the yeast *RAD10* protein whereas in addition also similarity with parts of the *E.coli* proteins *uvrA* and *uvrC* was found (8,30,32). These findings strongly suggest that throughout evolution excision repair functions have been conserved. We now report that antisense transcription units in the *ERCC-1* and *RAD10* loci might be part of the evolutionarily conserved features of DNA repair.

The architecture of the 3' region of the *ERCC-1* gene is schematically depicted in Fig.1A. Alternative polyadenylation of *ERCC-1* transcripts yields mRNAs of 1.1 and 3.4 kb (31). These RNA species are visualised on Northern blots of HeLa total poly(A)$^+$ RNA when hybridized with a randomly primed (9) ^{32}P-labeled 5' *ERCC-1* cDNA probe as shown in Fig. 2 (lane 1). However, a genomic probe from the middle part of the extended version of exon X (the hatched BamHI-PstI fragment in Fig. 1A), expected to hybridize only with the 3.4 kb alternatively polyadenylated *ERCC- 1* transcript, detected in addition a 2.6 kb RNA (Fig.2, lane 2).

The 2.6 kb transcript was not recognized by 5' *ERCC-1* probes. Southern blots of human genomic DNA probed with the unique BamHI-PstI fragment yielded hybridization patterns that were fully consistent with the physical map of this cloned region (not shown), excluding the possibility that cross-hybridization with other sequences accounts for the detection of the 2.6 kb RNA. To investigate the possibility that the 2.6 kb RNA was

transcribed from the opposite strand, HeLa poly(A)$^+$ RNA was hybridized to ^{32}P-labeled strand specific RNA probes of the BamHI-PstI fragment. To that aim the 0.9 kb BamHI-PstI fragment shown in Fig. 1A was subcloned in both orientations downstream of the T7-promoter in pTZ18/19R (Pharmacia). With the aid of T7 RNA polymerase, ^{32}P-labelled single strand probes were synthesized using previously described conditions (20) and hybridized to Northern blots containing size-fractionated HeLa total poly A$^+$ RNA. The autoradiograms of this experiment (Fig.2A, lane 3 and 4) show that the 3.4 kb *ERCC-1* RNA and the 2.6 kb RNA are recognized by different probes indicating that these RNAs are transcribed from opposite DNA strands and that the 2.6 kb RNA represents an antisense transcript.

Screening of an 'Okayama' cDNA library (23) with the 1.0 kb PvuII fragment which includes ERCC-1 exon X (Fig.1A), yielded two partial cDNAs (designated pcD2.1 and pcD3B) of the antisense RNA. (see Fig.1A). The largest clone (pcD3B) had an insert size of 1 kb. Sequence analysis of both clones and the genomic DNA 3' of exon X revealed that pcD3B was completely colinear with the genomic DNA. A compilation of the genomic sequences around exon X and the sequence of the 3' half of pcD3B is presented in Fig. 1B. Both cDNA clones completely overlap with *ERCC-1* exon X and terminate in intron 9 at approximately 60 bp downstream of the splice acceptor of *ERCC-1* exon X. A polyadenylation signal ATTAAA is located at approximately 23 bp upstream of the start of the poly(A) tail, indicating that both cDNAs were derived from an authentic poly(A)$^+$ mRNA which has a 5' to 3' orientation opposite to *ERCC-1*. It can be deduced from these data that the 2.6 kb transcript has an overlap of 170 bp with the *ERCC-1* transcription unit, yielding the major 1.1 kb mRNA, whereas it is completely complementary to the 3.4 kb *ERCC-1* transcript for at least 1 kb. We have provisionally designated this antisense gene *ASE- 1* (antisense *ERCC-1*). The *ASE-1* cDNA clone pcD3B harbors a 5' truncated potential open reading frame encoding 183 amino acids (not shown) which is suggestive of a coding function of the antisense transcript. However, at this stage, we cannot rule out the possibility that the cloned portions of the cDNA represent part of the 3' untranslated region of the 2.6 kb *ASE-1* RNA. To substantiate a coding function of the antisense transcript, poly(A)$^+$ RNA was prepared from HeLa nuclei and polysomes following described protocols (6) and hybridized with *ASE-1* probes (Fig. 2B). Using the XbaI-PstI fragment covering the region of *ERCC-1* exon X as a probe (Fig. 1A), the 3.4 kb *ERCC-1* and 2.6 kb anti-sense transcripts could be detected in both RNA samples, and the 1.1 kb *ERCC-1* transcript showed weak hybridization due to the presence of exon X sequence information in the probe. The relative amounts of the 2.6 and 3.4 kb transcripts differ in nuclei and polysomes. The alternatively polyadenylated 3.4 kb *ERCC-1* RNA appears to be mainly of nuclear origin, whereas the antisense transcript is found predominantly in the polysome fraction, suggesting that *ASE-1* mRNA is translated into a protein.

The human *ASE-1* DNA specifically cross-hybridizes to cloned DNA of the region downstream of mouse *ERCC-1* exon X and strand specific RNA probes identify an antisense RNA of 2.3 kb in this region (data not shown). Hence it appears that the genomic organization of *ERCC- 1* and *ASE-1* is conserved between man and mouse.

The *ERCC-1* homologue in yeast, the *RAD10* gene, and its transcripts are shown schematically in Fig. 3A. The indicated restriction fragments were cloned in both orientations into M13 to investigate transcripts encoded by the *RAD10* gene and to determine

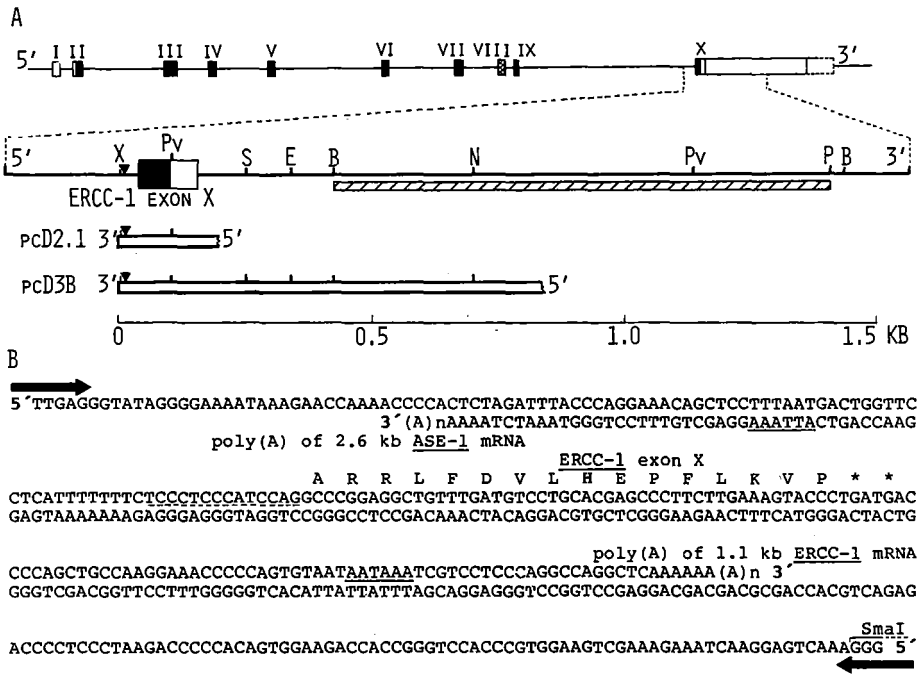

Figure 1

A. Schematic representation of the *ERCC-1* gene and its 3' flanking region. Boxes with roman numerals refer to *ERCC-1* exons. Black parts represent *ERCC-1* coding sequences. The alternatively spliced exon VIII is hatched. The variable size of exon X is due to alternative polyadenylation yielding longer transcripts of 3.4 kb and 3.8 kb in addition the major mRNA of 1.1 kb. (31). The pcD2.1 and pcD3B antisense cDNA clones were isolated from a cDNA library of SV40 transformed fibroblasts (23). An inverted black triangle marks the polyadenylation signal ATTAAA of the antisense transcript. Restriction sites: B,BamHI; E,EcoRI; N,NunII; P,PstI; Pv,PvuII; S,SmaI; X,XbaI. The 0.9 kb BamHI-PstI fragment used as a probe in Fig. 2 is shown by the hatched bar. **B.** Nucleotide sequence of the *ERCC-1* exon X region and cDNAs of the *ASE-1* gene. The arrows indicate the orientation of *ERCC- 1* (upper) and the antisense transcription unit (lower). The double stranded region indicates the overlap between the 1.1 kb *ERCC-1* mRNA and the 2.6 kb *ASE-1* mRNA. The overlap with the longer 3.4 kb *ERCC-1* RNA (not indicated) continues over the entire *ASE-1* sequence shown. The sequence of *ERCC-1* exon X and part of the flanking DNA has been reported earlier (30,31). Amino acids encoded by *ERCC-1* exon X are given in the one letter code. The *ERCC-1* exon X splice acceptor is shown by a dashed line.

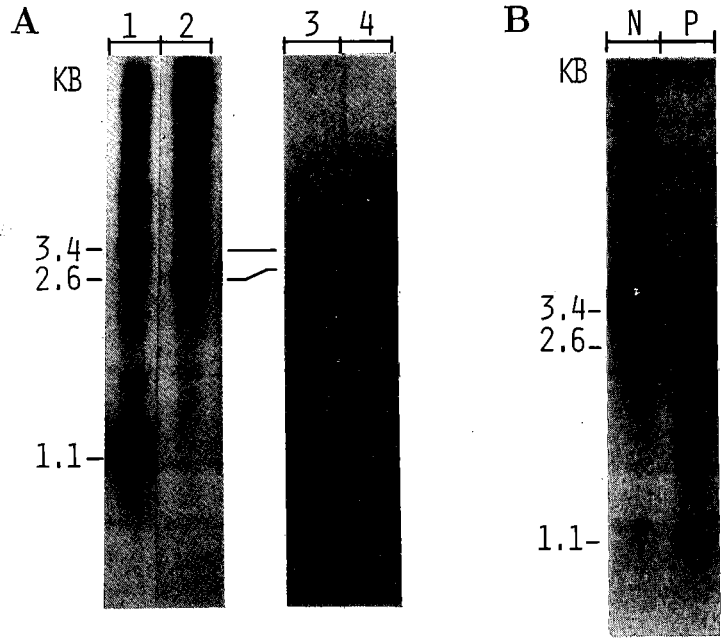

Figure 2

Northern blot analysis of HeLa poly(A)$^+$ RNA. **A**) 20 μg (lane 1 and 2) or 10 μg (lane 3 and 4) poly(A)$^+$ RNA isolated by the LiCl/urea method (2) and passaged twice over oligod T-cellulose using routine procedures (19) was size fractionated and hybridized to the following probes: lane 1, *ERCC- 1* cDNA probe harboring exon I to V; lane 2, the BamHI-PstI fragment shown in Fig. 1A; Lane 3 and 4 show hybridization with ^{32}P-labeled single strand RNA probes of opposite orientations of this BamHI-PstI fragment respectively. The filters hybridized with the RNA probes were after hybridization washed at 68^{o}C in 0.1 x SSC and subsequently treated with RNAse A (1 μg/ml in 2 x SSC, 15 min) to reduce non-specific background labeling. **B** Poly(A)$^+$ RNA (6 μg) of HeLa nuclei (N) and polysomes (P) hybridized with the ^{32}P-labeled XbaI-BamHI fragment of the genomic *ERCC-1* 3' region (see Fig. 1A).

the gene organization 3' of *RAD10*. Isolation of poly(A)$^+$ RNA from the Rad$^+$ strain DBY747, *MAT* **a** *his3-1 leu2-3 leu2-112 trp-289 ura3-52*, RNA blotting and hybridization were as described (18). Poly(A)$^+$ RNA was dissolved in 1 M glyoxal, fractionated on a 0.8% agarose gel in 10 mM $NaPO_4$ (pH 6.5), transferred to genescreen and hybridized to ^{32}P-labeled M13 derived single strand probes which were synthesized by published procedures (27). Northern blot analysis of yeast RNA with ^{32}P-labeled single strand M13 derived DNA probes of the various fragments is shown in Fig.4. A probe corresponding to the transcribed *RAD10* DNA strand of the PvuII-XbaI fragment hybridized to three transcripts of 1.0, 1.5 and 1.8 kb. Strand specific probes of the adjacent XbaI-EcoRV and EcoRV-EcoRV fragments hybridized to the larger two bands, whereas the more downstream EcoRV-BglII fragment recognized only the 1.8 kb *RAD10* transcript (Fig. 4A). The *RAD10* origin of these transcripts was confirmed by analysis of a *RAD10* disruption strain (data not shown). Since the 5' ends of the *RAD10* transcripts map at positions -17 and -32 (25), we conclude that the size heterogeneity in the *RAD10* transcripts arises mainly from differences at the 3' end and that as in the case of *ERCC-1 RAD10* displays alternative polyadenylation.

Surprisingly, hybridization of yeast poly(A)$^+$ RNA with radiolabeled probes corresponding to the non-coding *RAD10* DNA strand revealed a 1.9 kb transcript. As shown in Fig.4B, this transcript was recognized by 5 different strand specific probes of the entire region between the XbaI and NruI sites indicated in Fig. 3A. Hence, the opposite DNA strand in the 3' *RAD10* region encodes a transcript that overlaps the 1.8 kb *RAD10* transcript by at least 600 nucleotides (Fig. 3A). We have provisionally designated the yeast antisense gene as *ASR10* (antisense *RAD10*).

To examine whether the *ASR10* gene encodes a protein, the nucleotide sequence of 3' *RAD10* DNA region was determined. The 1.9 kb antisense RNA contains an open reading frame (ORF) encoding 525 amino acids with a calculated M_r of 59,505 (results not shown). As is illustrated in Fig. 3B the last two codons of the *ASR10* ORF overlap with those of the *RAD10* ORF.

Screening of the EMBL and NBRF data banks did not show any significant similarity of *ASR10* with other genes or proteins. The predicted amino acid sequence of the yeast *ASR10* protein did not show any homology with the disrupted potential open reading frame present in *ASE-1* cDNA clone pcD3B. It will be of interest to examine whether the missing portion of the *ASE-1* cDNA encodes a protein which shows similarity to the *ASR10* amino acid sequence. From these findings we conclude that the phenomenon of overlapping antisense gene organisation is shared between the homologous mammalian *ERCC-1* and yeast *RAD10* gene.

Recently, several examples of antisense trancription in higher eukaryotes have been reported (1,5,13,29,33). However, in all cases, the biological significance or conservation of this phenomenon is unknown. To our knowledge the *ERCC-1/ASE-1* locus represents the first example of naturally occuring antisense transcription in the human genome. Also in yeast antisense transcription is unusual, though not without precedent. Recently, convergent overlapping transcription for two yeast *CDC* genes was described and divergent overlapping RNAs for the *HAP3* locus were reported (3,12). Although the occurrence of overlapping gene transcription in the related mammalian *ERCC-1* and yeast *RAD10* regions can be purely coincidental, we consider this unlikely in view of the rarity of this

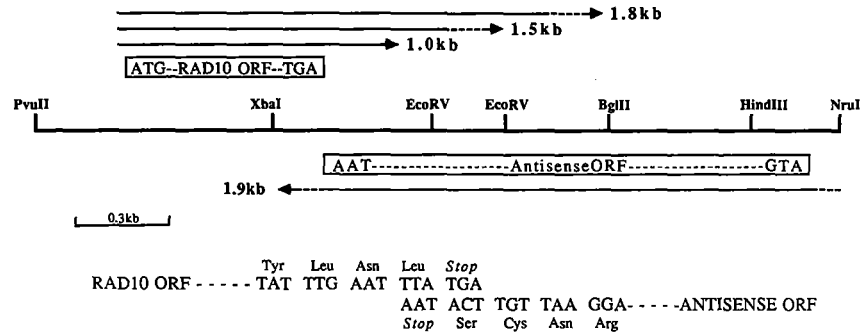

Figure 3

A. Transcript map of the *RAD10* gene and antisense gene region. The horizontal line in the middle gives the restriction map of the *RAD10* region DNA. Only the relevant restriction enzymes sites with six base recognition are given. The position of the *RAD10* and *ASR10* open reading frames (antisense ORF)is indicated by open bars. The straight lines with arrowheads above the *RAD10* ORF and below the antisense ORF represent the direction and approximate size of *RAD10* transcripts and the antisense transcript respectively. The dashed lines at the 5' or 3' ends of the transcripts indicate that the position of these ends was not accurately determined by S1 nuclease mapping. The 5' ends of *RAD10* transcripts were mapped to positions -17 and -32 by S1 mapping (25). The 3' ends of the *RAD10* 1.0 kb transcript map at positions +869 and +938, as determined by S1 mapping (unpublished results). **B.** Nucleotide and amino acid sequence of overlapping coding region of *ASR10* and *RAD10*.

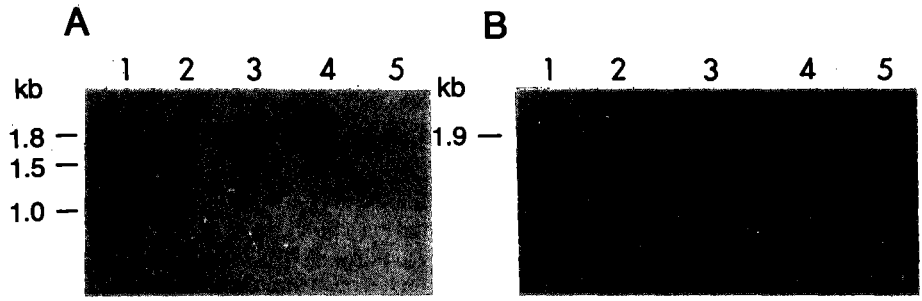

Figure 4

Northern blot analysis of yeast poly(A)$^+$ RNA with strand specific DNA probes of the *RAD10* region. **A**) Transcripts hybridizing to strand specific probes from the *RAD10* gene. Ten μg poly(A)$^+$ RNA was loaded in each lane. The following probes were used after cloning in M13mp18 or 19 (see Fig.3; given in parentheses): 1,PvuII-XbaI (mp19); 2, XbaI-EcoRV (mp18); 3, EcoRV-EcoRV (mp19); 4,EcoRv-BglII (mp18); 5, BglII- HindIII (mp9). **B**) Transcripts recognized by strand specific probes from the antisense gene. 5.6 μg poly(A)$^+$ RNA was loaded per lane. Probes: 1, XbaI-EcoRV (mp19); 2,EcoRv-EcoRv (mp19, in opposite orientation of the clone used in Fig.2A); 3,EcoRV- BglII (mp19); 4, BglII-HindIII (mp8); 5, HindIII-NruI (mp9).

phenomenon.

At present one can only speculate about the function of the antisense transcripts. The presumed coding capacity of both yeast and human antisense RNAs makes it unlikely that the antisense transcripts are solely regulatory RNAs with e.g. no other function than regulating translation as is observed in prokaryotes (11) and as applied in 'reverse genetics' experiments in eukaryotes (4,7,15-17,21,22,26). Several advantages might accrue from partially complementary transcripts derived from overlapping genes. A tail-tail duplex configuration of the 2 mRNAs involved could mediate their transport to a common location in the cytoplasm. Furthermore, the stability of such a tandem transcript could be affected by the duplex state and serve to protect against degradation. At the translational level several options are open. Translation could be inhibited or slowed down by the hybridized 3' tail. Furthermore, unwinding of the 3' termini could offer the possibility of induction of gene expression at the translational level. In these respects it is worth noting that transcription levels, translation initiation, codon usage and protein level (10,28,32) suggest that *RAD10* and *ERCC-1* are low abundance proteins. Another intriguing possibility at the protein level, particularly for lowly expressed genes, is that tail-tail association of sense and antisense transcripts could provide an opportunity for complex formation between the two nascent proteins that will be in close vicinity of each other when the ribosomes on the opposite transcripts encounter each other. The possible regulatory effect of the conserved overlapping antisense transcription on the expression of *ERCC- 1/RAD10* is being examined.

We thank Dr. Okayama for the cDNA library, H. Odijk for technical assistance, M. Koken, A. Westerveld and K. Jaspers for stimulating discussions and J. Fengler for help in preparing the manuscript. This work was financially supported by EURATOM, MEDIGON, Foundation of Medical and Health Research in the Netherlands and by grants from the U.S National Institute of Health and Department of Energy.

LITERATURE CITED

1. Adelman, J.P., C.T. Bond, J. Douglass and E. Herbert. 1987. Two mammalian genes transcribed from opposite strands of the same locus. Science 235: 1514-1517.

2. Auffray, C. and F. Rougeon. 1980. Purification of mouse immunoglobuline heavy chain messenger. Eur.J.Biochem. 107: 303- 314.

3. Barker, D.G., J.H.M. White and L.H. Johnston. 1985. The nucleotide sequence of the DNA ligase gene (*CDC9*) from *Saccharomyces cerevisiae*: a gene which is cell-cycle regulated and induced in response to DNA damage. Nucl. Acids Res. 13: 8323- 8337.

4. Cabrera, M.C., P. Alonso, P. Johnston, R.G. Phillips and P.A. Lawrence. 1987. Phenocopies induced with antisense RNA identify the *wingless* gene. Cell 50: 659-663.

5. Chen, C., T. Malone, S.K. Beckendorf and R.L. Davis. 1987. At least two genes reside within a large intron of the *dunce* gene of *Drosophila.* Nature 329: 721-724.

6. Clemens, M.J. 1987. Purification of eukaryotic messenger RNA. *In* B.D. Hames and S.J. Higgins (eds.) Transcription and translation, a practical approach. IRL press, Oxford, 211-230.

7. Crowley, T.E., W. Nellen, R.H. Gomer and R.A. Firtel. 1985. Phenocopy of Discoidin I-minus mutants by antisense transformation in Dictyostelium. Cell 43: 633-641.

8. Doolittle, R.F., M.S. Johnson, I. Husain, B. van Houten, D.C. Thomas, and A. Sancar. 1986.Domainal evolution of a prokaryotic DNA repair protein and its relationship to active transport proteins. Nature 323: 351-453.

9. Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132: 6-13.

10. Friedberg, E.C. 1987. The molecular biology of nucleotide excision repair of DNA: recent progress. J. Cell Sci. suppl. 6: 1- 23.

11. Green, P.J., O. Pines and M. Inouye. 1986. The role of antisense RNA in gene regulation. Ann. Rev. Biochem. 55: 569-597.

12. Hahn, S., J. Pinkham, R. Wei, R. Miller and L. Guarente. 1988. The *HAP3* regulatory locus of *Saccharomyces cerevisiae* encodes divergent overlapping transcripts. Mol. Cell. Biol. 8: 655-663.

13. Henikoff, S., M. Keene, K.Fechtel and J. Fristrom. 1986. Gene within a gene: nested *Drosophila* genes encode unrelated proteins on opposite strands. Cell 44: 33-42.

14. Hu, N.-T. and J. Messing. 1982. The making of strand-specific M13 probes. Gene 17: 271-277.

15. Izant, J.G. and H. Weintraub. 1984. Inhibition of thymidine kinase gene expression by antisense RNA: a molecular approach to genetic analysis. Cell 36: 1007-1015.

16. Izant, J.G. and H. Weintraub. 1985. Constitutive and conditional suppression of exogenous and endogenous gene by anti-sense RNA. Science 229: 345-352.

17. Kim, S.K. and B.J. Wold. 1985 Stable reduction of thymidine kinase activity in cells expressing high levels of anti-sense RNA. Cell 42: 129-138.

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

19. Maniatis, T., E.F. Fritsch and J. Sambrook. 1982. *In:* Molecular cloning, A Laboratory Manual (Cold Spring Harbor Laboratory) Cold Spring Harbor, New York.

20. Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn and M.R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acids Res. 12: 7035-7056.

21. Melton, D.A. 1985. Injected anti-sense RNAs specifically block messenger RNA translation *in vivo*. Proc.Natl.Acad.Sci.USA 82: 144-148.

22. Nishikura, K. and J.M. Murray. 1987. Antisense RNA of proto-oncogene c-*fos* blocks renewed growth of quiescent 3T3 cells. Mol.Cell.Biol. 7: 639-649.

23. Okayama, H. and P. Berg. 1983. A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. Mol. Cell. Biol. 3: 280-289.

24. Prakash, L., D. Dumais, R. Polakowska, G. Perozzi and S. Prakash. 1985. Molecular cloning of the *RAD10* gen of *Saccharomyces cerevisiae*. Gene 34: 55-61.

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

26. Rosenberg, U.B., A. Preiss, E. Seivert, H. Jackle and D.C. Knipple. 1985. Production of phenocopies by kruppel antisense RNA injection into *Drosophila* embryos. Nature 313: 703-706.

27. Sanger, F., A.R. Coulson, B.G. Barrell, A.J.H. Smith and B.A.Roe. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. 143: 161-178.

28. Sharp, P.M., T.M.F. Tuohy and K.R. Mosurski. 1986. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucl. Acids Res. 14: 5125-5143.

29. Spencer, C., D. Gietz and R. Hodgetts. 1986. Overlapping transcription units in the dopa decarboxylase region of *Drosophila*. Nature 322: 279-281.

30. Van Duin, M., J.de Wit, H. Odijk, A. Westerveld, A. Yasui, M.H.M. Koken, J.H.J. Hoeijmakers and D. 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.

31. 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 excision repair gene *ERCC- 1*. Nucl.Acids Res. 22: 9195-9213.

32. 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*. Nucl. Acids Res. 16: 5305-5322.

33. Williams, T. and M. Fried. 1986. A mouse locus at which transcription from both DNA strands produces mRNAs complementary at their 3' ends. Nature 322: 275-277.