The ERCC1–xeroderma pigmentosum group F DNA repair complex

Het ERCC1-xeroderma pigmentosum groep F DNA herstel complex

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. Dr P.W.C. Akkermans M.A.
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op

woensdag 31 maart 1999 om 9.45 uur

door
Anneke Mieke Sijbers
geboren te Venray

Promotiecommissie

Promotoren:	Prof. Dr J.H.J. Hoeijmakers
	Prof. Dr D. Bootsma
Overige leden:	Prof. Dr D. van Lindhout
	Prof. Dr A. Westerveld
	Prof. Dr B.A. Oostra
Co-promotor:	Dr N.G.J. Jaspers

Dit proefschrift kwam tot stand binnen het instituut Celbiologie en Genetica van de Faculteit der Geneeskunde en Gezondheidswetenschappen van de Erasmus Universiteit Rotterdam. Het instituut maakt deel uit van het Medisch Genetisch Centrum Zuid-West Nederland. Het onderzoek is gedeeltelijk financieel ondersteund door de Nederlandse vereniging voor Wetenschappelijk Onderzoek (NWO-GMW-900-501-113).

voor mijn ouders

voor Paul

Contents

CHAPTER 1

DNA damage, repair and human hereditary disease

AN INTRODUCTION

In its long nucleotide-chains, the DNA double helix contains the genetic information for ten thousands of proteins. The DNA molecule, however, is subject to constant change. In order to maintain its integrity, several mechanisms that cope with DNA damage, inflicted by various naturally occurring and man-made agents, have evolved. Individuals with the rare genetic disease xeroderma pigmentosum demonstrate the importance of DNA repair systems. These patients carry a defect in proteins that are involved in the removal of UV-induced damage from the DNA, which results in sun-sensitivity and the frequent occurrence of skin cancers. Hence, studies on DNA repair and human repair disorders such as xeroderma pigmentosum help our understanding of the origin of cancer and processes like aging. This introduction briefly reviews our current knowledge of different types of DNA damage and DNA repair processes, with an emphasis on the nucleotide excision repair pathway. Xeroderma pigmentosum and other human repair disorders are dealt with later in the chapter.

DNA damage and consequences

Numerous endogenous and environmental agents can react with DNA and produce a wide variety of DNA lesions. Major DNA lesions together with the corresponding repair mechanisms are listed in Table 1.

Origin	DNA lesion	Repair pathway
endogenous		
DNA synthesis (incorporation incorrect or damaged base or slippage) DNA recombination	mismatches	mismatch repair
Reactive metabolites (*e.g.* oxygen radicals)	base alterations (oxidation and alkylation), loss of bases and strand breaks	base excision repair recombinational repair end-joining
environmental		
UV-light	cyclobutane pyrimidine dimers, 6-4 photoproducts	nucleotide excision repair photoreactivation
Ionizing radiation	single and double strand breaks, loss of bases, base alterations	recombinational repair end joining base excision repair
Cross-linking agents (*e.g.* mitomycin C, cisplatin)	intra- and inter-strand cross-links	recombinational repair nucleotide excision repair
Alkylating agents	alkylated bases	DNA-alkyltransferases base excision repair nucleotide excision repair
Aromatic compounds	bulky adducts	nucleotide excision repair

Table 1 Major DNA lesions (adapted from Friedberg *et al* 1995)

The most common natural environmental genotoxic agent is the UV component of sunlight. UV links neighboring pyrimidine bases on a DNA strand. As these dimers distort the DNA double helix, they can block essential cellular processes like replication and transcription. DNA lesions can also have altered base-pairing properties. Unless removed, inappropriate or mispaired bases may bring about

permanent alterations in the genetic code. Sequence alterations can consist of base pair substitutions, deletions and insertions. Such mutations can change gene expression or alter the structure and/or function of encoded proteins resulting in disturbed cell growth. Moreover, an accumulation of mutations in somatic cells plays an important role in the development of cancer and aging. When mutations occur in germline cells, they may even give rise to hereditary disease.
It is clear that removal of DNA damage is of vital importance to the cell in order to keep its normal function and to prevent it from malignant transformation. The following section describes different repair processes that give cells the ability to survive DNA damage.

DNA repair mechanisms

DNA repair mechanisms are many and varied. The present section briefly summarizes the different DNA repair pathways known to date. For an extensive overview, the reader is referred to Friedberg *et al* (1995).

Reversal of DNA damage

The simplest way to remove certain types of DNA damage is through direct reversal in a single-enzyme reaction. Alkyltransferases simply extract the alkyl group from alkylated bases that is transferred to an internal cysteine residue, thereby inactivating themselves (Teo *et al* 1984). Photolyases, on the other hand, revert UV-induced dimers in a light-dependent reaction called photo-reactivation (Sancar 1990, Yasui and Eeker 1998). Photolyases specific for cyclobutane pyrimidine dimers are wide-spread among prokaryotes and eukaryotes, from bacteria and fungi up to vertebrates and even aplacental mammals, but they have not been found in placental mammals (Yasui *et al* 1994). Photolyases for 6-4 photoproducts have also been found in various species, including Drosophila, Xenopus and plants (Todo *et al* 1993). Only recently, two genes with homology to the class of 6-4 photolyases and blue-light receptors have been identified in man. However, no overt photolyase activity has been demonstrated for these genes yet (Todo *et al* 1996, van der Spek *et al* 1996, Yasui and Eeker 1998).

Double-strand break repair

More complex processes are required for the repair of double-strand breaks (DSBs) (Fig. 1). Strand breaks result from exposure to ionizing radiation and certain chemicals but are also naturally occurring intermediates in cellular processes like V(D)J-recombination and meiosis. Besides the simple rejoining of nonhomologous ends, that may not be error-free (Chu 1997, Kanaar and Hoeijmakers 1997), DSBs can be restored through homology-dependent recombination that can occur by homologous recombination repair or single-strand annealing (SSA).
Recombination is initiated by a 5'→3'exonuclease that generates long single-stranded 3' tails (White and Haber 1990, Sugawara and Haber 1992). In homologous recombination, one of these tails invades a homologous DNA duplex. The donor

sequence can serve as a template for new DNA synthesis. The daughter strand anneals with another single-stranded tail on the other side of the break. In this way, a recombination intermediate is formed with two crossover points also known as Holliday junctions. Resolution of the Holliday junction restores the exact sequence into two recombinant DNA molecules, with or without crossing-over.

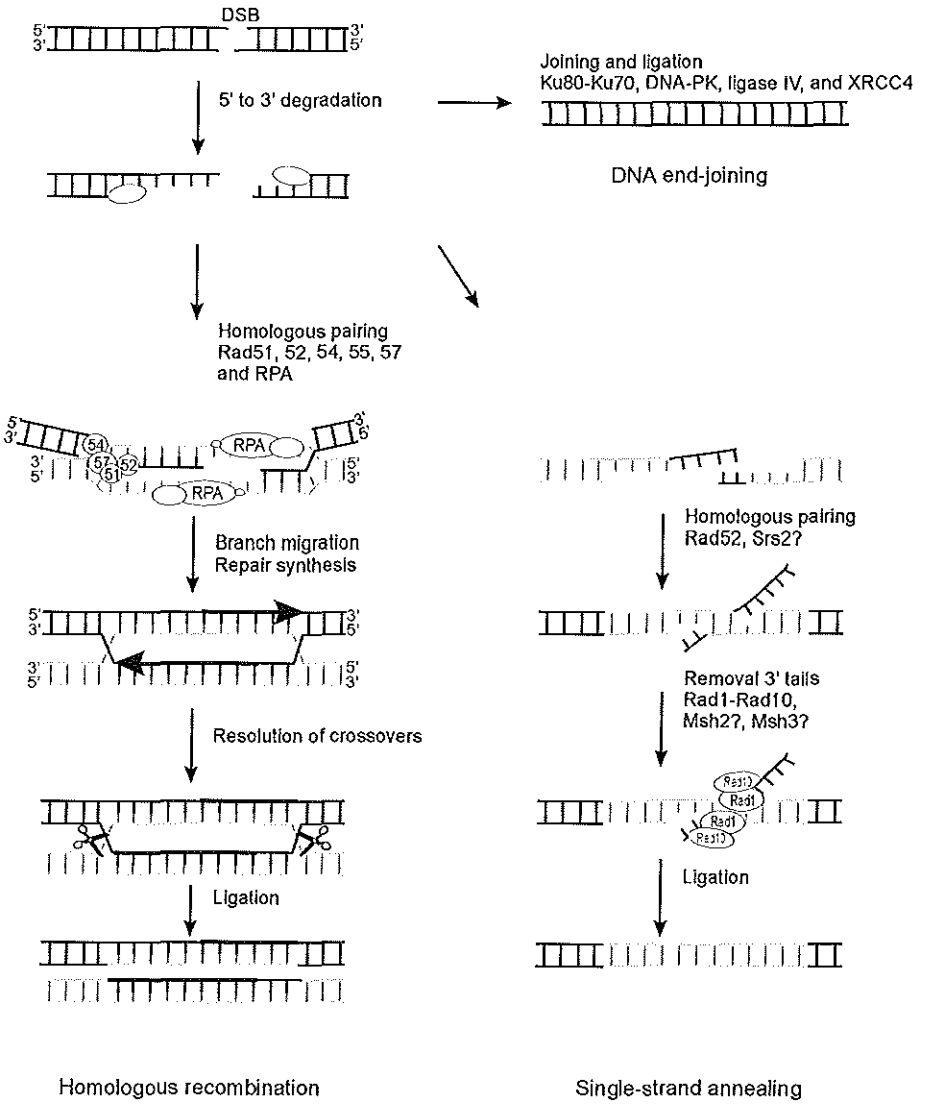


Fig. 1 Model for double-strand break repair in yeast. In the case of homologous recombination, the homologous DNA molecule is indicated in gray. For the single-strand annealing pathway, these gray areas represent repeats within the same DNA double helix. Arrows indicate new DNA synthesis. For further details see Chapter 6.

In the alternative SSA pathway, degradation from the breakpoint continues until a relatively short region of homology is found (*e.g.* within repeats) (Sugawara and Haber 1992). Subsequent annealing of the single-stranded complementary strands produces an intermediate with non-homologous 3' ends that must be removed before new DNA synthesis and ligation can take place (Fishman-Lobell and Haber 1992, Yao *et al* 1997). Repair of a DSB via single-strand annealing is accompanied by a deletion and can also involve a translocation.

Postreplication repair

In case the DNA replication machinery encounters a lesion, DNA synthesis can restart downstream of the lesion, leaving a gap. These single-strand gaps opposite lesions in the newly replicated DNA strand can be filled in by postreplicative repair. The process, also known as daughter strand gap repair, involves homologous pairing and strand exchange with the undamaged sister molecule (Friedberg *et al* 1995). At present, little is known about the proteins involved and their role in this process. Another way to pass by a lesion in the DNA template is translesion synthesis, which is less accurate and therefore error-prone (Paz-Elizur *et al* 1996). The remaining damage will be repaired via one of the other repair processes mentioned here.

Mismatch repair

Mismatch repair corrects misincorporation errors after DNA synthesis and acts on mispaired bases that occur in recombination intermediates. In *Escherichia coli*, the MutS dimer specifically binds to mispairs (Grilley *et al* 1989) (Fig. 2). Subsequent binding of MutL activates MutH that makes a single-strand incision, which can be on either side of the mismatch (Au *et al* 1992). The repair proteins select the newly synthesized DNA strand, as it remains temporarily unmethylated. Subsequent excision initiates at the nick and proceeds towards the mismatch depending on the concerted action of MutS, MutL, the UvrD helicase and an exonuclease activity (Grilley *et al* 1993). The reaction is completed by resynthesis over several hundred bases. A similar process takes place in eukaryotes utilizing proteins highly homologous to MutS and MutL. However, in eukaryotes, mismatch recognition involves heterodimeric protein complexes rather than the MutS and MutL homodimers employed by *E. coli*, *e.g.* MSH2 forms a complex with either MSH3 or MSH6 (all MutS homologs) (Drummond *et al* 1995, Palombo *et al* 1995, Palombo *et al* 1996). These complexes have different but overlapping substrate specificities. Whereas the human MSH2–MSH6 dimer binds to single base mispairs and loops of one or two bases, the MSH2–MSH3 dimer prefers larger looped structures (Modrich 1994, Alani 1996, Johnson *et al* 1996, Marsischky *et al* 1996). The mechanism that discriminates between the newly synthesized strand and the template strand is unknown, but does not involve hemi-methylation as in *E. coli*. Mutations in the encoding human genes are associated with hereditary nonpolyposis colorectal cancer (HNPCC) (Modrich 1994, Fishel and Kolodner 1995).

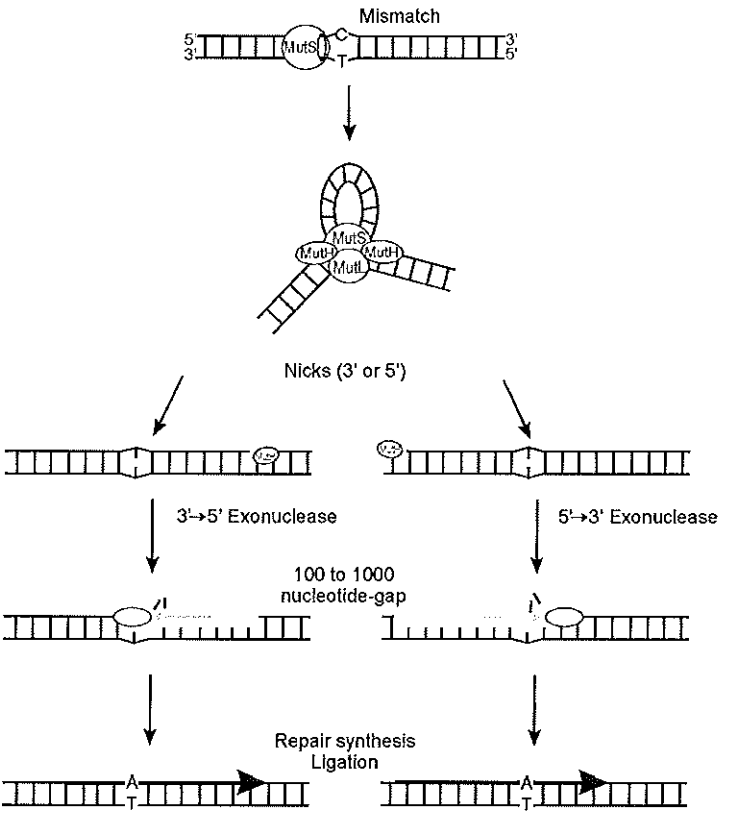


Fig. 2 Model for mismatch repair in *E. coli*.

Base excision repair

Base excision repair (BER) protects the cell against the effects of various types of endogenously generated small types of DNA base damage. BER is also important to eliminate similar lesions produced by ionizing radiation and alkylating agents (see Table 1). Different DNA glycosylases with partially overlapping substrate specificity remove modified bases by cleavage of the base-sugar bond leaving an abasic site (Fig. 3). Subsequently, the sugar residue is removed by an AP-endonuclease or AP-lyase activity and the one-nucleotide gap is filled-in by DNA polymerase β and DNA ligase IV in eukaryotes. Occasionally, longer repair patches of 2 to 10 residues are observed (Matsumoto *et al* 1994, Frosina *et al* 1996). It is likely that these longer patches result from repair synthesis associated with strand displacement in the 5'→3' direction. The produced 'flap' structure would be a substrate for the flap endonuclease FEN-1 (Harrington and Lieber 1994a, Klungland and Lindahl 1997).

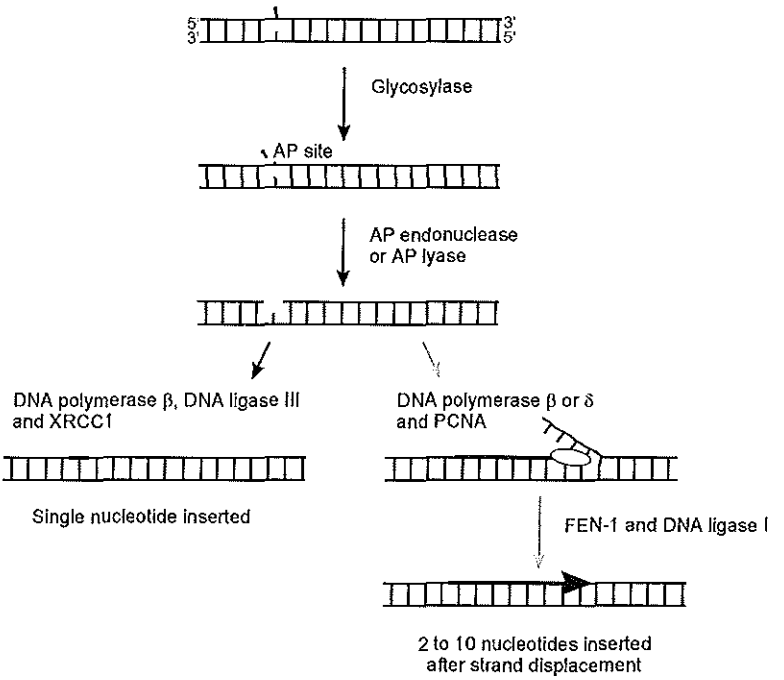


Fig. 3 Model for base excision repair.

Nucleotide excision repair

Another excision repair process is the nucleotide excision repair pathway (NER). The major difference between NER and BER is the way the damage is removed. Basically, NER cuts out the damage as part of an oligonucleotide fragment, while BER excises only one nucleotide. NER will be described in more detail in the next section as the XPF and ERCC1 proteins engaged in NER are the main subject of the present thesis.

An alternative DNA excision repair pathway for UV-lesions

In most organisms NER is thought to be a major pathway for the repair of UV-induced damage. In addition to NER, photo-reactivation and in some species BER are also involved in UV-damage repair (*i.e.* CPD-specific DNA glycosylases have been identified in bacteriophage T4 and bacterium *Micrococcus luteus*) (Dodson *et al* 1994). Recently, an alternative nucleotide excision repair pathway for the removal of UV photoproducts has been identified in *Schizosaccharomyces pombe* and the filamentous fungus *Neurospora crassa* (Doetsch 1995, Freyer *et al* 1995, Yajima *et al* 1995). It involves a new type of nicking activity, named the *S. pombe* DNA endonuclease (SPDE). The enzyme makes an incision immediately 5' to the dimer (Bowman *et al* 1994, Freyer *et al* 1995). Following incision, either a 3'endonuclease

or a 5'→3'exonuclease (perhaps the yeast FEN-1 homolog) may proceed with the repair reaction. The latter would explain the reduced repair of UV-induced dimers found in an UV-sensitive *S. pombe fen-1* mutant (McCready *et al* 1993, Murray *et al* 1994). Subsequent resynthesis and ligation may involve proteins shared with other repair systems (Murray *et al* 1994).

For a long time it was thought that the different repair pathways would operate independent from one another. However, a number of repair proteins appear to be involved in multiple processes, providing a way to regulate the different ways of repair. This issue is addressed in Chapter 6.

Nucleotide excision repair

Nucleotide excision repair is currently the most intensively studied repair process. Although NER has a very broad substrate-specificity, it is primarily involved in the removal of UV-induced damage from the DNA. The repair reaction sequentially involves: recognition of the damage, DNA unwinding, incision of the damaged strand on both sides of the lesion, excision of an oligonucleotide containing the damage, and gap-filling DNA synthesis. NER has been maintained during evolution. The process is the most simple in bacteria. In *E. coli* only three proteins -UvrA, UvrB, and UvrC- are required for the incision reaction. UvrA forms a dimer that in a complex with UvrB is responsible for damage recognition (Mazur and Grossman 1991, Myles and Sancar 1991). Subsequent binding of UvrB to the damaged site induces a conformational change in the DNA helix. As a result, the UvrA molecules dissociate allowing UvrC to bind (Moolenaar *et al* 1995). Both incisions occur in this UvrBC–DNA complex (Lin and Sancar 1992, Moolenaar *et al* 1995). Then the UvrD helicase releases the approximately 13 bases long oligonucleotide containing the lesion, DNA polymerase I fills the gap, and finally DNA ligase seals the repair patch. The NER processes in yeast and mammalian cells are very much alike. Therefore, the present chapter, as most of the thesis, will focus on NER in mammalian cells.

The mechanism of mammalian NER

The process in mammals is far more complex than in bacteria. No less than 15 proteins are required for dual incision. Most of the encoding genes have been isolated through transfection of repair-deficient rodent mutant cell lines and recovery of the excision repair cross complementing (*ERCC*) human genes. Nearly all ERCC proteins subsequently appeared to be involved in the human repair disorders xeroderma pigmentosum (XP) and Cockayne syndrome (CS), and are therefore called XP and CS factors, respectively. Two genes have been isolated using the XP cells themselves: *XPA* and *XPC*. Others have been cloned using the protein sequences of homologs in yeast. Interesting features of the XP and ERCC factors (Table 2) that take part in the different steps of NER (Fig. 4) are summarized in the following section.

Lesion recognition

Removal of the DNA lesion requires its recognition first. Three proteins preferentially bind damaged DNA *in vitro* and are probably required for the initiation of the repair reaction. The XPC protein strongly binds (UV-damaged) DNA (Drapkin *et al* 1994, Masutani *et al* 1994, Reardon *et al* 1996). XPA recognizes various lesions, including UV-induced 6-4 photoproducts (Robins *et al* 1991, Jones and Wood 1993, Asahina *et al* 1994), whereas the UV-DDB protein (for UV-damaged DNA binding) specifically binds to 6-4 photoproducts (Treiber *et al* 1992, Protic and Levine 1993, Reardon *et al* 1993a), but also cyclobutane pyrimidine dimers and cisplatin adducts (Chu and Chang 1988, Chu and Chang 1990).

The XPC protein exists in a complex with the human homolog of the *Saccharomyces cerevisiae* repair protein Rad23, HHR23B, that enhances XPC-dependent excision repair *in vitro* (Sugasawa *et al* 1996). There is strong evidence that XPC binds the damaged site first (Sugasawa *et al* 1998).

The XPA protein can interact with several repair factors that modulate its binding properties. For instance, XPA binds to the Replication protein A, RPA (Matsuda *et al* 1995, Saijo *et al* 1996). This trimeric single-stranded DNA-binding protein is required for both the incision reaction as well as repair synthesis (Coverley *et al* 1992, Aboussekhra *et al* 1995, Mu *et al* 1995). RPA stimulates the interaction of XPA with yet another repair protein, ERCC1 (Li *et al* 1994, Park and Sancar 1994, Li *et al* 1995b, Saijo *et al* 1996). ERCC1 resides in a complex with the XPF protein. Together they form an endonuclease that is responsible for one of the incisions around the lesion (Chapter 3). Like the homologous proteins in *S. cerevisiae*, XPA may bind the ERCC1–XPF complex with even higher affinity than either subunit alone (Guzder *et al* 1996). Both RPA and ERCC1 enhance the preference of XPA for damaged DNA (He *et al* 1995, Li *et al* 1995a, Nagai *et al* 1995), whereas its binding to damaged DNA, in turn, stimulates the interaction with the basal transcription factor TFIIH (Park *et al* 1995b, Nocentini *et al* 1997). RPA, on the other hand, can bind to the endonuclease XPG (He *et al* 1995), possibly XPF (Bessho *et al* 1997b), and UV-DDB (Otrin *et al* 1997). Small amounts of UV-DDB stimulate NER *in vitro* (Aboussekhra *et al* 1995, Mu *et al* 1995), possibly via this specific interaction with RPA.

Taken together, data strongly suggest that the XPC–HHR23B and pre-assembled XPA–RPA complexes bind to the damaged site and recruit the factors required for dual incision. However, it is also possible that a complex consisting of TFIIH, XPA and RPA acts as one entity after initial damage detection by XPC–HHR23B.

DNA unwinding

During NER the DNA is opened over a region of at least 25 bases with help of TFIIH, a large protein complex that participates both in initiation of basal transcription and in NER (Evans *et al* 1997a, Wood 1997). The XPB and XPD proteins are subunits of the TFIIH factor and their helicase activities (with opposite polarity) are probably required for local unwinding of the DNA double helix (Schaeffer *et al* 1993, Drapkin *et al* 1994, Humbert *et al* 1994, Schaeffer *et al* 1994, van Vuuren *et al* 1994). XPC–HHR23B may assist this strand separation by interacting with

TFIIH (Drapkin *et al* 1994). However, the need for XPC depends on the DNA structure, as for the removal of lesions in a partially unwound DNA substrate the XPC–HHR23B complex is not required (Mu *et al* 1996, Mu and Sancar 1997, Mu *et al* 1997b). Binding of RPA to the non-damaged strand may prevent incision of the strand (that serves as a template for repair synthesis) and stabilize the open structure (Kim *et al* 1992, Seroussi and Lavi 1993).

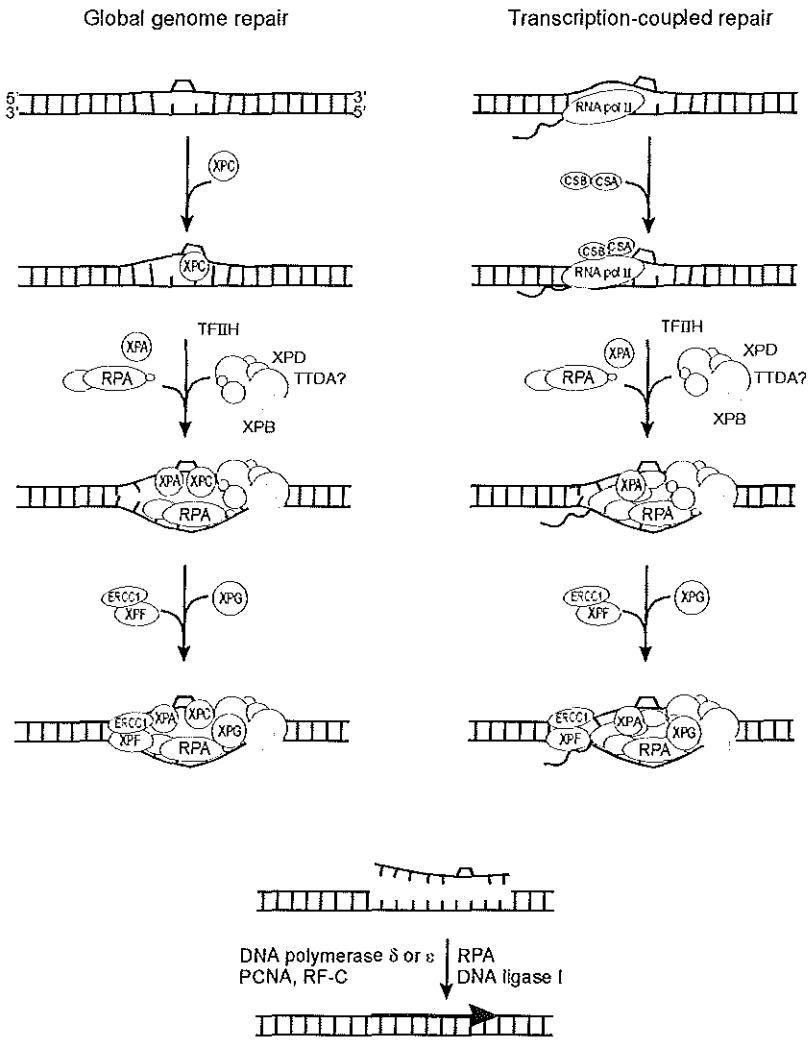


Fig. 4 Model for nucleotide excision repair in mammalian cells.

Dual incision

TFIIH and RPA can both bind to XPG, the protein that probably proceeds with the repair reaction (Mu *et al* 1995, Habraken *et al* 1996, Iyer *et al* 1996, de Laat *et al* 1998b). XPG makes the first incision while it cuts the damaged DNA strand on the 3' side of the lesion (O'Donovan *et al* 1994a, Matsunaga *et al* 1995, Mu *et al* 1996). Next, the ERCC1–XPF complex makes the incision on the opposite side (Mu *et al* 1996). Both endonucleases specifically cut near junctions of unpaired and duplex DNA in a structure-dependent way (Chapter 3, O'Donovan *et al* 1994a). Recent results indicate that RPA helps to position the 3' and 5' incisions (de Laat *et al* 1998b), which could explain the finding that with certain substrates, RPA stimulates the structure-specific incision of ERCC1–XPF (Matsunaga *et al* 1996, Bessho *et al* 1997b).

Excision

Whereas the cleavage positions vary with the type of adduct, the excised oligomers are more or less constant in length. Incision 2 to 9 phosphodiester bonds away from the lesion on the 3' side and 16 to 25 bonds away on the 5' side releases a 24 to 32-nucleotide-long fragment *in vitro* (Huang *et al* 1992, Svoboda *et al* 1993, Matsunaga *et al* 1995, Moggs *et al* 1996). It is not known which protein(s) displace the lesion-containing strand, but both the excised fragment and the gapped DNA molecule remain bound to one or more components of the repair complex (Mu *et al* 1995, Mu *et al* 1996). XPG can bind to the proliferating cell nuclear antigen PCNA (Gary *et al* 1997). This interaction stimulates repair, though is not essential to the incision reaction. It probably stimulates the release of the protein assembly and/or links excision with subsequent repair synthesis (Nichols and Sancar 1992).

Gap-filling

DNA polymerase δ and/or ε fills in the resulting single-stranded gap using the complementary strand as a template. The polymerase is supported by proteins known to be involved in DNA replication: RPA, PCNA and Replication factor C (RF-C) (Miura *et al* 1992, Nichols and Sancar 1992, Shivji *et al* 1992, Aboussekhra *et al* 1995, Wood and Shivji 1997). RF-C is needed to load PCNA onto the DNA (Podust *et al* 1994). PCNA may be required for the initiation of repair synthesis and/or subsequent turnover of the repair complex. It may also stimulate polymerase activity. The single-strand breaks are rejoined by DNA ligase I sealing the repair patch.

Human protein (complex)		ERCC protein	Homologs		
			S. cerevisiae	*S. pombe*	other
XPA			Rad14	[1]	$Dxpa^{Dme}$
RPA	p70		RPA[2]		
	p32				
	p14				
XPC			Rad4		XPC^{Dme}
HHR23B			Rad23		
HHR23A			Rad23		
XPE	UV-DDB1				
	UV-DDB2				
TFIIH	XPB	3	Rad25[3]	$ERCC3^{Spo}$	XPB^{Dme}
(core)	XPD	2	Rad3	Rad15	XPD^{Dme}
	p52[4]		Tfb2		
	p62		Tfb1		
	p44		Ssl1		
	p34		Tfb4		
	TTDA				
XPG		5	Rad2	Rad13	
ERCC1		1	Rad10	Swi10[5]	
XPF		4	Rad1	Rad16[6]	$Mei\text{-}9^{Dme}$
					$Mus38^{Ncr}$
(XPF)		11	Rad1	Rad16	$Mei\text{-}9^{Dme}$
					$Mus38^{Ncr}$
CSA		8	Rad28		
CSB		6	Rad26		
		7			
		9			
		10			
			Rad7		
			Rad16		
			MMS19		

[1] identified in database (our unpublished observation)
[2] complex of three subunits
[3] allelic with SSL2
[4] previously known as p41
[5] allelic with Rad23 (Hang *et al* 1996)
[6] allelic with Swi9, Rad10, and Rad20 (Schmidt *et al* 1989)

Table 2 Eukaryotic NER factors

Protein activity	Sensitivity human/rodent mutant		Additional function
	UV	mitomycin C	
damage recognition	++	+/-	
ssDNA binding			DNA replication recombination
damage recognition/ ssDNA binding	+		
damage recognition/ ssDNA binding	+		
3'→5'helicase	++	+/-	basal transcription
5'→3'helicase	++	+/-	
3' structure-specific endonuclease	++	±	
5' structure-specific endonuclease	++	+++	single-strand annealing
	++	+++	
	+	+	
	+		
ds DNA-dependent ATPase	+		
	+	±	
	+		
	+	+	
ATP-dependent DNA damage sensor			

[Spo] *S. pombe*
[Dme] *D. melanogaster*
[Ncr] *Neurospora crassa*

Transcription-coupled repair

The basal transcription factor TFIIH is involved in both transcription and NER. A number of lesions, when present in actively transcribed genes, are repaired more efficiently than lesions in the rest of the genome via a process known as transcription-coupled repair (TC-NER). Repair is probably initiated by a stalled RNA polymerase II (Hanawalt 1994) (Fig. 4). The CS factors, CSA and CSB, are specifically involved in TC-NER. They probably support the assembly of the repair complex at the site of arrested transcription. Several interactions have been reported that may be relevant to this process, including CSB–RNA polymerase II (van Gool *et al* 1997), CSB–CSA and CSB–XPG (Iyer *et al* 1996), CSB–CSA–TFIIH (Henning *et al* 1995) and CSB with DNA and XPA (Selby and Sancar 1997).

The XPC protein, on the other hand, is specifically involved in the repair of the remaining of the genome. Because of the condense chromatin structure, these regions are normally inaccessible. It is likely that XPC increases the accessibility of the damaged site for other repair proteins as XPC seems not to be required for repair of already partially unwound DNA structures (Mu *et al* 1996, Mu and Sancar 1997, Mu *et al* 1997b). XPC may take the role of RNA polymerase II during global genome repair (Sugasawa *et al* 1998).

Complexity

The exact composition of the preincision complex is not known. According to Mu *et al* (1997b) the smallest DNA–protein complex detected *in vitro* is formed in the presence of the purified XPA, RPA, XPC–HHR23B, and TFIIH factors. Recent studies, however, revealed the presence of a DNA–XPC–HHR23B pre-incision factor, which appeared to be the first factor to bind to a lesion (Sugasawa *et al* 1998). The isolation of a factor that consists of the yeast XPA, XPF and ERCC1 homologs suggests that the preincision complex is formed at the damaged site by sequential assembly of the different repair proteins (Guzder *et al* 1995, Guzder *et al* 1996, Habraken *et al* 1996). Other results, however, suggest the existence of a preformed 'repairosome'. Two forms of yeast TFIIH have been isolated as large complexes. The factor probably active in transcription is associated with a kinase activity, whereas the factor involved in repair seems to include repair proteins instead of the kinase (Svejstrup *et al* 1995, Mu *et al* 1996, He and Ingles 1997).

Defects in NER can result in three genetic disorders: xeroderma pigmentosum, Cockayne syndrome and PIBIDS, the photosensitive form of trichothiodystrophy (TTD), which are shortly described in the next section. For a recent overview, the reader is referred to Bootsma *et al* (1998).

Human NER-deficient disorders

Xeroderma pigmentosum

XP is a rare, autosomal recessive disease. In the Netherlands, one in 200,000 individuals suffers from XP. Characteristics of XP are sun-sensitivity, dry appearance

of the skin (*xeroderma*), pigmentation abnormalities (*pigmentosum*), a more than two thousand fold increased incidence of skin cancer, and in some cases neurological disease due to accelerated neurodegeneration. Cell fusion experiments established seven NER-deficient complementation groups, XP-A through XP-G, and a so-called variant group (XP-V) that is thought to be defective in postreplication repair (Table 3) (Vermeulen *et al* 1991).
While XP has a worldwide distribution, XP-A is most common in Japan and in the Mediterranean countries. Mutations in the *XPA* gene usually give rise to a severe phenotype associated with neurological abnormalities, though leaky mutations have been reported (Mimaki *et al* 1996, States and Myrand 1996). Among the Japanese group A patients three different mutations prevail. Two are nonsense mutations that produce truncated proteins. These give rise to a less severe phenotype than the splice mutation responsible for a complete lack of XPA protein. XP-B is extremely rare. At present, only five patients (from three independent families) have been assigned to this complementation group. Two mutations in the *XPB* gene originate in an XP phenotype combined with the severe characteristics of the CS repair disorder, whereas a third mutation gives rise to features of the third NER syndrome, TTD (Weeda *et al* 1997b). The *XPB* gene encodes a subunit of the basal transcription factor TFIIH.
The XPC protein is the damage recognition factor (Sugasawa *et al* 1998). Interestingly, XP-C cells are only defective in genome-overall repair, suggesting that XPC might be specifically involved in repair of transcriptionally inactive DNA. XP-C patients normally have no neurological complications (Kraemer *et al* 1987, Cleaver and Kraemer 1989), although one exceptional case has been reported with neurodegeneration in later life (Robbins *et al* 1993).
Unlike XP-C, XP-D is frequently associated with neurological abnormalities. The *XPD* gene product -like XPB- is a subunit of the TFIIH factor. Mutations in *XPD* result in classical XP, combined XP-CS or TTD. The molecular defect causing XP-E is still unclear. Some XP-E individuals lack the DNA-binding activity of one of the two subunits of the UV-DDB complex. The protein can stimulate, but is not essential to the NER reaction *in vitro*. This might explain the relatively mild clinical symptoms and a high residual cellular repair activity characteristic of XP-E (Wood 1997).
Likewise, XP-F patients have a relatively mild clinical phenotype. XP-F patients hardly experience sensitivity to sunlight. Fifty percent of the XP patients assigned to a complementation group belong to group A or C, contrasting the only eleven cases (five percent) with XP-F (Cleaver and Kraemer 1989). In the past ten years, we have come across at least fourteen additional cases with XP-F. In only two out of the eighteen reported XP-F individuals, neurological symptoms developed later in life (Chapter 4, Moriwaki *et al* 1993).
Patients belonging to complementation group G are in general more severely affected and like in XP groups B and D there is extensive clinical heterogeneity. Low levels of repair and a XP phenotype are the consequence of subtle mutations that inactivate the XPG endonuclease. In those cases where mutations result in a truncated XPG protein, additional CS features are found (Nouspikel *et al* 1997). Some of the features of the different complementation groups are summarized in Table 3.

Group	Freq	UV sens	UDS[1] (%)	Skin cancer	Neurological abnormalities	Features
XP-A	high	+++	2-5	+	present	
XP-B	low[2]	++	10-40	+/–	variable	XP-CS or TTD
XP-C	high	+	15-30	+	absent	genome overall repair-deficient
XP-D	int	++	15-50	+/–	variable	XP, XP-CS, or TTD
XP-E	low	±	40-50	+/–	absent	
XP-F	low	+	15-30	+/–	mild	repair low but long-lasting
XP-G	low	++	2-25	+/–	variable	XP or XP-CS
XP-V	high	+/±	normal	+/–	absent	postreplication repair deficient
CS-A	int	+	normal	–	severe	TC-NER-deficient
CS-B	high	+	normal	–	severe	TC-NER-deficient
TTD-A	low[3]	+	10	–	mild	
UV^s	low[4]	+	normal	–	absent	CS-like

[1] Unscheduled DNA synthesis as percentage of repair synthesis in normal cells
[2] 3 families
[3] 1 family
[4] 2 families

Table 3 NER-deficient complementation groups

Trichothiodystrophy

PIBIDS is the acronym for the photosensitive form of TTD, which is further characterized by ichthyosis (scaling of the skin), (sulfur-deficient) brittle hair, impaired intelligence, decreased fertility, and short stature. While acute sun-sensitivity is associated with PIBIDS, other skin problems typical of XP are not, for instance skin cancer has never been reported in TTD patients. Molecular defects have been assigned to the *XPD* and *XPB* genes and to a still unidentified gene, *TTDA* (Stefanini *et al* 1993a, 1993b, Vermeulen *et al* 1994b). The XPB, XPD, and probably TTDA proteins are associated with the basal transcription factor TFIIH (Schaeffer *et al*

1993, Drapkin *et al* 1994, Schaeffer *et al* 1994, Vermeulen *et al* 1994b). Since features like brittle hair are not easily explained by a NER defect alone, they are assumed to result from a subtle defect in the transcription function of TFIIH (Vermeulen *et al* 1994b).

Cockayne syndrome

Cockayne syndrome patients have an aged appearance due to a dry and sometimes scaly skin, thin hair, and diminished subcutaneous tissue. The neurological and developmental complaints associated with CS are similar to those found in TTD. Growth retardation and skeletal abnormalities generally begin within the first year of life (reviewed by Nance and Berry 1992). The mean age of death in reported patients is about 12 years, although a few affected individuals have lived into their late teens and twenties. Cells from CS patients (without additional XP) are specifically defective in the transcription-coupled repair process (Venema *et al* 1990, van Hoffen *et al* 1993). CS can result from mutations in at least five genes (Lehmann 1995). Classical CS is associated with mutations in the *CSB* or *CSA* genes (Troelstra *et al* 1992b, Henning *et al* 1995). Individuals with an exceptional combined XP-CS phenotype carry mutations in the *XPD*, *XPB*, or *XPG* gene (Vermeulen *et al* 1993, Vermeulen *et al* 1994a, Broughton *et al* 1995). Mutations in these genes can affect the repair and/or transcription initiation function of TFIIH. Features typical of CS, like cataract, hearing loss, dental caries, and poor development, are thought to be the result of a subtle transcription defect.

Notably, several of the CS and TTD features are also found in individuals with COFS and CAMFAK syndrome (CS-like), Tay syndrome and Amish brittle hair syndrome (TTD-like), and other candidate 'transcription syndromes' (Bootsma and Hoeijmakers 1993, Vermeulen *et al* 1994b).

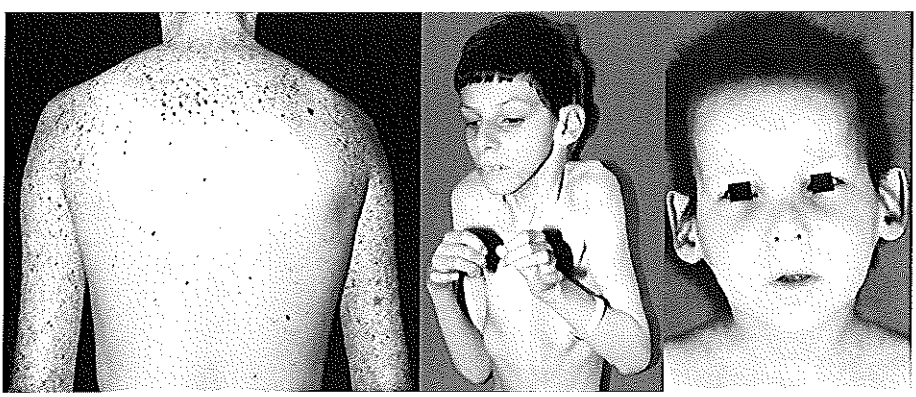


Fig. 5 Xeroderma pigmentosum, Cockayne syndrome, and PIBIDS patient (from left to right). Courtesy of Lehmann and McCuaig.

Outline of the thesis

The aim of the study presented in the present thesis was to unravel the identity and role of the ERCC1-containing protein complex in DNA repair. Chapter 2 presents the results of a mutational analysis revealing the different domains important for ERCC1 activity. Chapter 3 reports on the isolation and identification of the complex partner of ERCC1 as the protein involved in xeroderma pigmentosum group F. The proteins together form a structure-specific endonuclease. The identification of XP-F mutations in a Dutch patient (Chapter 4) and the mutual interaction domains in ERCC1 and XPF (Chapter 5) provided a molecular basis for the clinical phenotype. Finally, the role of the ERCC1–xeroderma pigmentosum group F complex in DNA repair is discussed in Chapter 6.

CHAPTER 2

Mutational analysis of the human nucleotide excision repair gene ERCC1

Anneke M. Sijbers
Peter J. van der Spek
Hanny Odijk
Joke van den Berg
Marcel van Duin
Andries Westerveld
Nicolaas G.J. Jaspers
Dirk Bootsma
Jan H.J. Hoeijmakers

DEPARTMENT OF CELL BIOLOGY AND GENETICS,
MEDICAL GENETICS CENTER,
ERASMUS UNIVERSITY ROTTERDAM, THE NETHERLANDS

Nucleic Acids Res 24: 3370-3380 (1996)

The human DNA repair protein ERCC1 resides in a complex together with the ERCC4, ERCC11, and XP-F correcting activities, thought to perform the 5' strand incision during NER. Its yeast counterpart, Rad1–Rad10, has an additional engagement in a mitotic recombination pathway, probably required for repair of DNA cross-links. Mutational analysis revealed that the poorly conserved N-terminal 91 amino acids of ERCC1 are dispensable for both repair functions in contrast to a deletion of only four residues from the C-terminus. A database search revealed a strongly conserved motif in this C-terminus sharing sequence homology with many DNA break-processing proteins, indicating that this part is primarily required for the presumed structure-specific endonuclease activity of ERCC1. Most missense mutations in the central region give rise to an unstable protein (complex). Accordingly, we found that free ERCC1 is very rapidly degraded, suggesting that protein–protein interactions provide stability. Survival experiments show that the removal of cross-links requires less ERCC1 than UV repair. This suggests that the ERCC1-dependent step in cross-link repair occurs outside the context of NER and provides an explanation for the phenotype of the human repair syndrome xeroderma pigmentosum group F.

Introduction

Repair of damaged DNA prevents accumulation of lesions that give rise to mutations, chromosomal instability, carcinogenesis or cell death. A wide variety of DNA lesions caused by exposure to UV-light and numerous chemical agents is removed via the NER pathway. This repair process involves specific damage recognition, dual incision of the damaged strand, followed by lesion removal, gap-filling and finally strand ligation (for a recent review see Hoeijmakers 1994). Most of the proteins engaged in NER have been identified by making use of UV-sensitive mutant rodent cells (ERCC1 through ERCC11) and cells derived from patients suffering from xeroderma pigmentosum (XP-A through XP-G), Cockayne syndrome (CS-A and CS-B) or trichothiodystrophy (TTD-A). XPA protein is thought to play an important role in the damage recognition step as it specifically binds to damaged DNA (Robins *et al* 1991, Jones and Wood 1993, Asahina *et al* 1994) and interacts with several other repair proteins including the RPA heterotrimer (He *et al* 1995, Matsuda *et al* 1995), XPG (He *et al* 1995), the basal transcription factor TFIIH (Park *et al* 1995b), and the ERCC1 complex (Li *et al* 1994, Park and Sancar 1994, Li *et al* 1995b, Nagai *et al* 1995). Following damage recognition the helicase activities of XPB and XPD (Sung *et al* 1993a, Roy *et al* 1994), present in the TFIIH complex (Schaeffer *et al* 1993, Drapkin *et al* 1994, Schaeffer *et al* 1994), are thought to convert the damaged site into a substrate for XPG and the ERCC1 complex, likely to be responsible for dual strand incision around the lesion. Further action of RPA, PCNA, RF-C, DNA polymerase δ and/or ε, and ligase are needed to complete the full NER reaction (Shivji *et al* 1995).
Although *ERCC1* was the first human NER gene cloned (Westerveld *et al* 1984), information on its enzymatic function is still very limited. The protein exists in a complex together with the ERCC4, ERCC11 and XP-F correcting activities (Biggerstaff *et al* 1993, Reardon *et al* 1993b, van Vuuren *et al* 1993). Largely due to the difficulty of purifying it to homogeneity (van Vuuren *et al* 1995), the exact composition of the complex has not yet been fully resolved, although recently a heterodimeric ERCC1 complex was reported (Park *et al* 1995a). By homology with its *S. cerevisiae* counterpart Rad10 (van Duin *et al* 1986), which associates with the Rad1 protein (Bailly *et al* 1992, Bardwell *et al* 1992), the ERCC1 complex is expected to mediate endonucleolytic incision at the 5' side of the lesion (Sung *et al* 1993b, Tomkinson *et al* 1993, Bardwell *et al* 1994, Tomkinson *et al* 1994, Matsunaga *et al* 1995). The nature of this putative activity, however, remains to be established. The domain of ERCC1 involved in the transient interaction with XPA extends from residues 93 to 120 (Li *et al* 1994), in a region that is strongly conserved in Rad10 (van Duin *et al* 1986). Further, on the basis of this conservation, the area could be involved in the association with the human homolog of Rad1 (Bardwell *et al* 1992), ERCC4 and 11 and/or XPF.

Beyond the central region, towards the C-terminus, significant homology with the C-terminus of the *E. coli* NER protein UvrC is observed (see Fig. 9). This domain is conserved in the *S. pombe* ERCC1 homolog Swi10 (Rödel *et al* 1992), but absent in Rad10 from *S. cerevisiae* (van Duin *et al* 1988). Both homologs have an additional function in a mitotic recombination pathway. In *S. cerevisiae* this pathway involves recombination between direct repeats (Klein 1988, Schiestl and Prakash 1988, Schiestl and Prakash 1990) in which *Rad1* is required for removal of nonhomologous sequences from the 3' ends of recombining DNA (Fishman-Lobell and Haber 1992, Saffran *et al* 1994, Ivanov and Haber 1995). In *S. pombe* this pathway entails mating-type switching (Gutz and Schmidt 1985). The mammalian ERCC1 complex may have such a function as well. This idea is supported by the extreme hypersensitivity to DNA cross-linking agents that is unique to ERCC1- and ERCC4-deficient rodent mutants (Busch *et al* 1989). Interstrand cross-links probably require recombination for their elimination. In order to obtain more information on the significance of various ERCC1 domains for both the NER and recombination functions, we have constructed *ERCC1* cDNAs with specific mutations and measured their ability to correct the mutagen-hypersensitivity of rodent ERCC1-mutant 43-3B.

Results

To identify the regions in ERCC1 essential for its function in NER and cross-link repair, mutated *ERCC1* cDNAs were assayed for correction of the rodent group 1 mutant 43-3B. Like other mutants in this complementation group and in group 4, this UV-sensitive cell line also exhibits an extreme sensitivity to cross-linking agents such as mitomycin C (MMC) and cisplatin. The latter feature is not displayed by other UV-sensitive NER-deficient complementation groups and probably reflects the role of ERCC1 in recombination needed for elimination of interstrand cross-links. The requirement of ERCC1 for UV resistance corresponds with its function in NER. Stably transfected neomycin-resistant mass populations were examined for their responses to UV irradiation and MMC. To validate the findings two separate cDNAs for each mutation were tested. Since a negative result can have trivial reasons we studied in addition, when indicated, individual clones which were verified to contain one or more copies of intact mutated or wildtype *ERCC1* cDNA. Transfection to Chinese hamster 43-3B cells of a wildtype human *ERCC1* cDNA (encoding 297 residues) almost fully complements both repair defects of these cells (see Fig. 1 and 2).

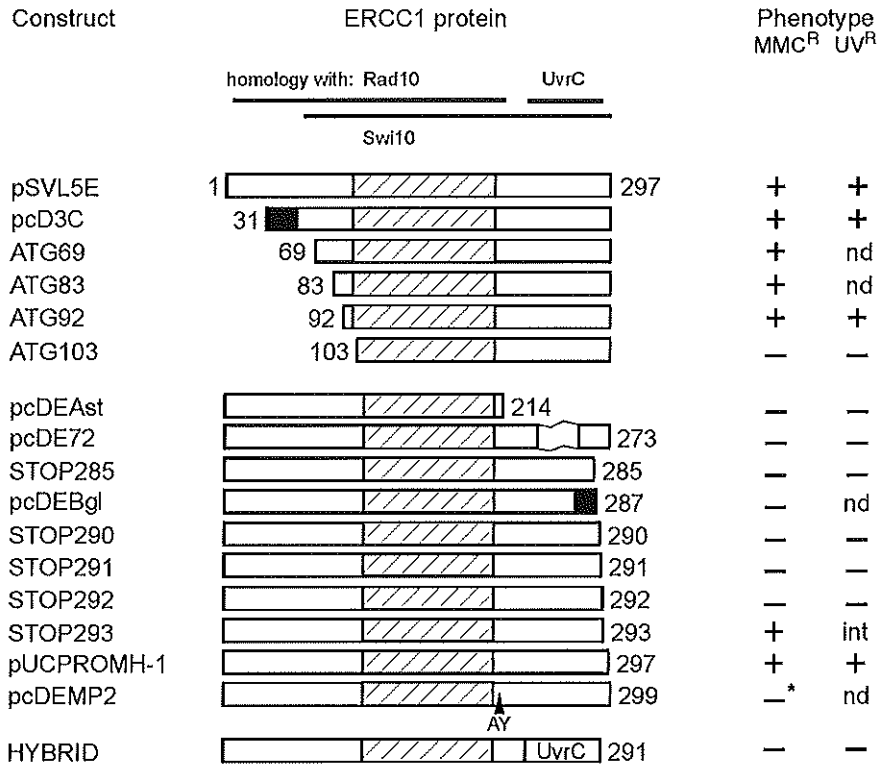


Fig. 1 Schematic representation of *ERCC1* deletion mutants and their correcting abilities. +, correction; int, intermediate correction; —, no correction of the 43-3B mutant phenotype; nd, not determined; *, dominant negative when overexpressed in wildtype cells. Numbers in brackets, number of amino acid residues. The hatched region illustrates the most strongly conserved sequence between ERCC1 and Rad10 and the black regions indicate nonsense sequences. In pcDE72, ERCC1 is lacking exon VIII, whereas in pcDEMP2 an extra alanine and tyrosine (AY) are inserted distal from residue 208. For several mutants the CFA was determined as described by Westerveld *et al* (1984).

N-terminal deletion mutants of ERCC1

Construct pcD3C encoding a truncated ERCC1 protein lacking the first 54 amino acids (Fig. 1) has been shown to confer MMC resistance on 43-3B cells (van Duin *et al* 1986). We have further shortened the protein by constructing *ERCC1* cDNAs containing the start codon at amino acid positions 69, 83, 92, and 103 (Fig. 1), preceded by an optimal translation initiation sequence.

Whereas the constructs ATG69, ATG83, and ATG92 all corrected both the UV and MMC sensitivity of recipient cells, ATG103 could not do so (see Fig. 2A and B for the ATG92 and ATG103 mutants, for others data not shown). We conclude that an N-terminal deletion of 91 residues comprising almost one third of the protein does not interfere with its repair functions.

It was not possible to verify the effect of the ATG92 and ATG103 mutations at the protein level by immunoblot analysis as our affinity-purified ERCC1 antiserum mainly recognizes epitopes in the N-terminus of the protein. Therefore, it remains uncertain whether a deletion of 102 amino acids results in an unstable protein or interferes with the protein activity itself.

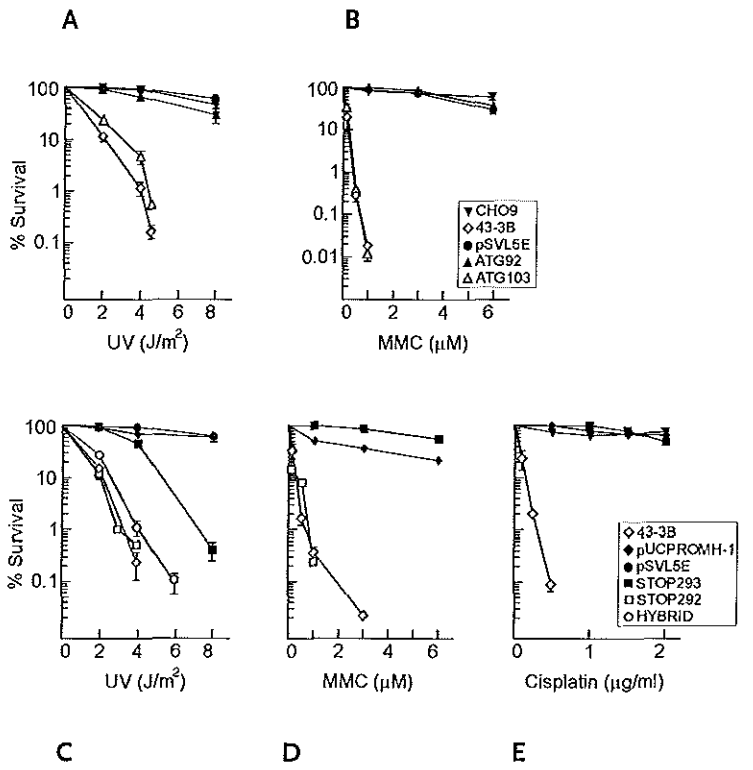


Fig. 2 Survival of 43-3B transfectants containing N-terminal deletion mutants after treatment with A. UV, B. MMC, and C-terminal deletion mutants after treatment with C. UV, D. MMC, and E. cisplatin. ◆ pUCPROMH-1 (mass population) and ● pSVL5E (single clone) both containing the wildtype *ERCC1* cDNA; ▲ ATG92 (mp); △ ATG103 (mp); ■ STOP293 (sc); □ STOP292 (sc); ○ HYBRID (mp). ▼ the parental wildtype cell line CHO9 (mp) and ◇ the mutant 43-3B (mp). The number of proliferating cells was measured as ^{3}H-thymidine incorporation. Points are average values for duplicate wells (or four for the untreated cells) and the error bars represent standard errors of means.

C-terminal deletion mutants of ERCC1

Previous studies have suggested that the strongly conserved C-terminal part of ERCC1 is crucial for its function. pcDEAst in which the *ERCC1* cDNA contains a premature stop codon at amino acid position 214 coding for a 'Rad10-like' ERCC1 protein (Fig. 1) could not correct MMC sensitivity. Neither could pcDEBgl, encoding a truncated protein of 287 amino acids with 17 unrelated C-terminal residues due to a frameshift mutation (van Duin *et al* 1988) nor pcDE72, a splice mutant lacking exon VIII (van Duin *et al* 1986) (Fig. 1).

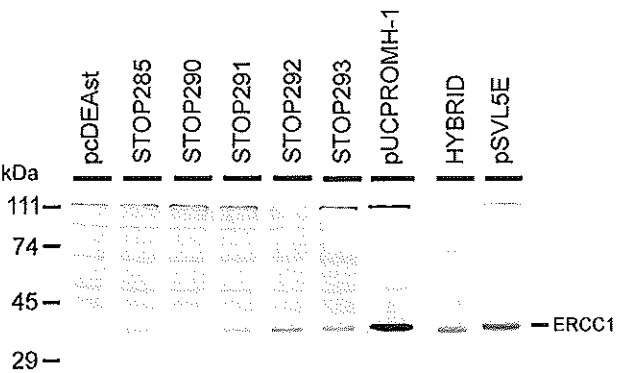


Fig. 3 Expression of wildtype and ERCC1 C-terminal deletion mutant proteins. Equal amounts (90 µg) of whole cell extracts were loaded (pcDEAst, STOP285, 290, 291, and 292, pSVL5E and HYBRID mass populations; STOP293 and pUCPROMH-1 single clones). Blots were incubated with affinity-purified anti-ERCC1 antiserum. ERCC1 protein migrates at 39 kDa. Untransfected 43-3B is shown in Fig. 6. Note that this antiserum is specific for human ERCC1 and does not recognize the endogenous Chinese hamster protein.

To more precisely determine the extent of the C-terminal functional area, a premature stop codon was introduced at amino acid position 286 resulting in a C-terminal deletion of 12 residues (STOP285). In addition, STOP290, STOP291, STOP292, and STOP293 were constructed (Fig. 1). Of these C-terminal truncations only STOP293 was able to correct the sensitivity of 43-3B cells to MMC, whereas for UV only a partial correction was found (Fig. 2C and D). The normal survival of STOP293 transfected cells after exposure to cisplatin (Fig. 2E) showed that the observed resistance to MMC reflects correction of cross-link sensitivity in general and rules out the possibility of deficient drug uptake or metabolism. Thus, only four residues can be deleted from the C-terminus without losing the cross-link repair function, although the UV damage repair function is slightly affected already. The various truncated proteins in whole cell extracts from stable transfectants were analyzed by immunoblotting (Fig. 3). STOP291, STOP292,

STOP293, and STOP290 (weakly) proteins could be visualized, indicating that they are stable *in vivo* though, with the exception of STOP293, non-functional. Further shortening of the protein apparently induces instability as no ERCC1 could be detected in STOP285 and pcDEAst (Fig. 3, note that the endogenous Chinese hamster ERCC1 protein is not recognized by our affinity-purified anti-ERCC1 antiserum, see also Fig. 6).
To assess whether the sequence homology of the C-terminus with the *E. coli* UvrC repair protein extends to the functional level, a hybrid construct was generated, with the human part substituted for its bacterial equivalent (Fig. 1). This ERCC1-UvrC hybrid protein is properly expressed (HYBRID in Fig. 3), however, failed to correct the sensitivity to UV and MMC (Fig. 1 and 2C).

Missense mutations in ERCC1

To further examine the presence of functional domains involved in one or both repair functions of ERCC1, specific amino acids were substituted in the region most strongly conserved between human ERCC1, *S. cerevisiae* Rad10, and *S. pombe* Swi10 (see Fig. 4). This part may contain the binding site for the human

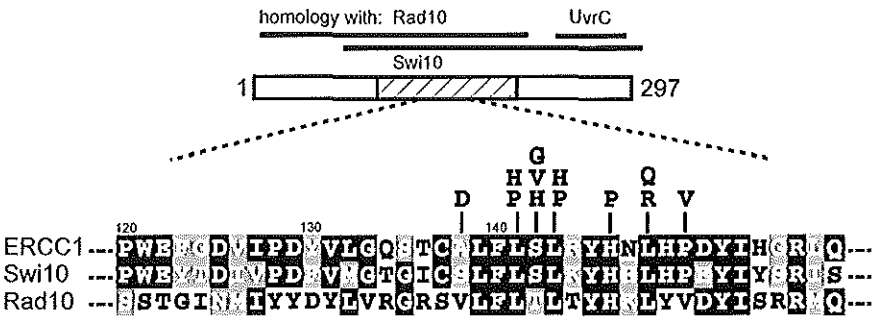


Fig. 4 Homology between ERCC1, Swi10 and Rad10 proteins in their most strongly conserved region. Identical amino acids are in black boxes, physicochemically related residues are in gray boxes, and missense mutations introduced are indicated. Displayed part of ERCC1 (297 amino acids) residues 120-159, Swi10 (252 amino acids) residues 62-101 and Rad10 (210 amino acids) residues 117-156 (numbers according to van Duin *et al* 1986, Rödel *et al* 1992, and Reynolds *et al* 1985).

homolog of Rad1 (Bardwell *et al* 1992) and/or it may harbor a DNA binding site (van Duin *et al* 1986). The types of changes made and the effects on UV and MMC survival after transfection to 43-3B cells are summarized in Table 1. For a number of ERCC1 mutants we found a considerable difference in correction of the UV and MMC sensitivity. This was also observed in individual clones containing intact mutated cDNA. For instance, in clone P150V(I) the UV sensitivity was complemented only partially, whereas the extreme sensitivity to both MMC and cisplatin

was almost fully restored (Fig. 5). In contrast, other mutant *ERCC1* transfectants retained the sensitivity to UV and showed partial or no correction of the MMC sensitivity. Immunoblot analysis revealed that most mutations gave rise to no or hardly detectable ERCC1 protein (Fig. 6), suggesting that they cause protein instability. In those cases where protein was detected a partially corrected phenotype was seen (Table 1), as shown for the P150V and L141H substitutions (see also Fig. 6). These observations suggest that a reduced amount of (mutated) ERCC1 is sufficient for the repair of MMC damage but not for the repair of UV damage. This interpretation is strongly supported by the isolation of two clones carrying the same mutated *ERCC1* cDNA (P150V), but which were found to express the en-

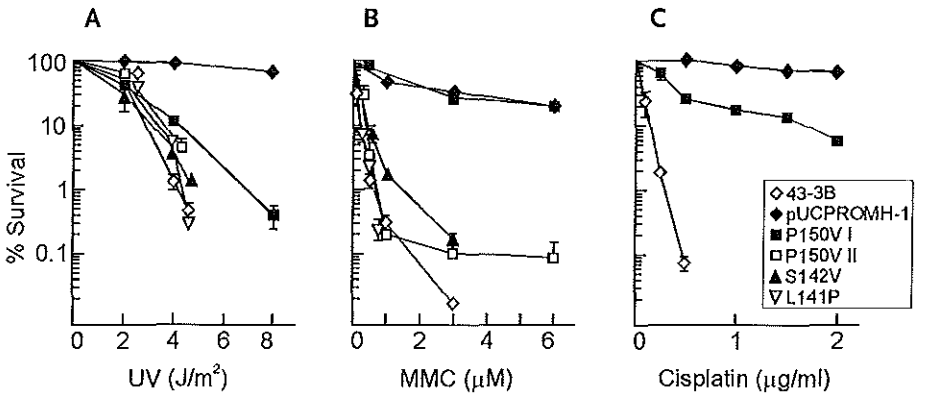


Fig. 5 Survival of 43-3B transfectants containing missense mutants following A. UV, B. MMC, and C. cisplatin treatment. ◇ the mutant cell line 43-3B (mass population) and 43-3B transfected with ▽ L141P (single clone); ▲, S142V (sc); ■, P150V(I) (sc); □, P150V(II) (sc); ◆, pUCPROMH-1 containing the wildtype *ERCC1* cDNA (sc). The number of proliferating cells was measured as either CFA or overall ^{3}H-thymidine incorporation. Points are average values for duplicate wells (or four for the untreated cells) and the error bars represent standard errors of means.

coded protein to a different level in repeated experiments. Clone P150V(II) exhibits only partial correction of the MMC sensitivity and no correction of UV sensitivity in contrast to P150V(I) (Fig. 5B and A). Immunoblot analysis revealed that the level of correction correlated with the amount of mutated ERCC1 protein detectable (Fig. 6, determined by copy number and site of integration, which differs in each transfectant).

ERCC1 expression was also analyzed at the single cell level by immunofluorescence. In rodent cells transfected with pUCPROMH-1, a human wildtype *ERCC1* construct, ERCC1 is expressed in the nucleus of every cell, but the expression level seems somewhat lower than in HeLa cells (Fig. 7). Although variation is seen in transfectants expressing mutated ERCC1 protein (even in cells derived from one sin-

Missense mutation	Checked in	UV^R	MMC^R	Protein expression level
C76W	mp		—	
Q107R	mp		—	
A138D	mp	—	—	not detectable
L141H	mp	int	++	reduced
L141P	sc	—	—	not detectable
S142G	mp		—	
S142H	mp		++	
S142V	sc	—	int	strongly reduced
L143H	mp	int	+	strongly reduced
L143P	mp	—	int	
H146P	mp	—	int	strongly reduced
L148R	mp		—	
L148Q	mp		—	
P150V (I)	sc	int	++	reduced
P150V (II)	sc	—	int	strongly reduced
W200S	mp		+	
Q251K	mp		+	

++ wildtype correction;
+ correction close to wildtype;
— no correction;
int intermediate correction of the 43-3B mutant phenotype.

For most mutants the CFA was determined as described by Westerveld *et al* (1984). (I) and (II) refer to two different populations containing the P150V mutated cDNA.

Table I Summary of ERCC1 missense mutations, their expression and correcting abilities

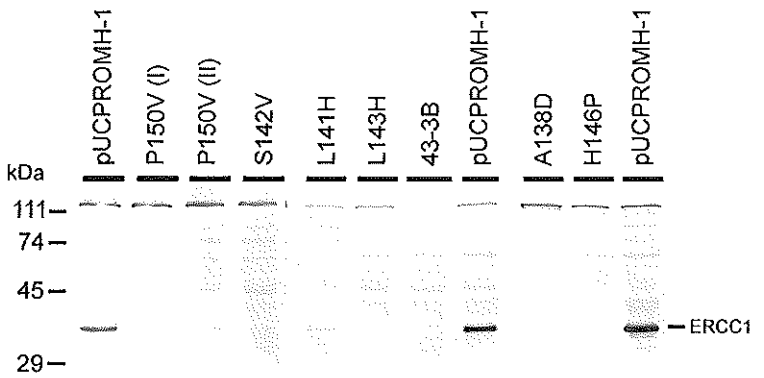


Fig. 6 Expression of wildtype and missense mutated ERCC1 proteins. Equal amounts (90 µg) of whole cell extracts were loaded (pUCPROMH-1, P150V(I), P150V(II), S142V single clones and others mass populations). Blots were incubated with affinity-purified anti-ERCC1 antiserum. ERCC1 protein migrates at 39 kDa. Note that the antiserum does not recognize the endogenous Chinese hamster ERCC1. (I) and (II) refer to two different populations containing the P150V mutated cDNA.

gle clone), the overall expression levels are consistently lower than in cells containing the wildtype cDNA. As with immunoblot analysis, no ERCC1 protein could be detected in transfectant containing the S142V construct that failed to correct.

Amplification and micro-injection

The remarkable absence of ERCC1 protein in most point mutants can be explained in two ways. Either the mutation renders the ERCC1 mRNA or protein unstable or the mutation interferes with correct folding of the protein and prevents it from proper association with the other component(s) of the complex. Uncomplexed (aberrant or wildtype) ERCC1 protein is then rapidly degraded. To investigate these possibilities we assessed the fate of an excess of wildtype ERCC1 protein obtained by *DHFR*-driven amplification of the wildtype gene and by micro-injection of purified ERCC1 protein in primary fibroblasts.

In an attempt to overproduce ERCC1 protein a construct containing the *DHFR* gene and the wildtype human *ERCC1* cosmid (Fig. 8A) together with two dominant selectable marker genes, *gpt* and *agpt*, as positive controls, was transfected into 43-3B cells. Transformants containing the different dominant markers and a functional *ERCC1* gene (as determined by wildtype UV and MMC resistance) were treated with stepwise increasing methotrexate concentrations inducing amplification of the *DHFR* gene together with its flanking sequences. Southern blot analysis and *in situ* hybridization to metaphase chromosomes revealed a massive (100- to more than a 1000-fold) amplification of the *ERCC1* gene in all stable transfectant clones analyzed (*e.g.* clone 41D in Fig. 8B and 8C, respectively). A cor-

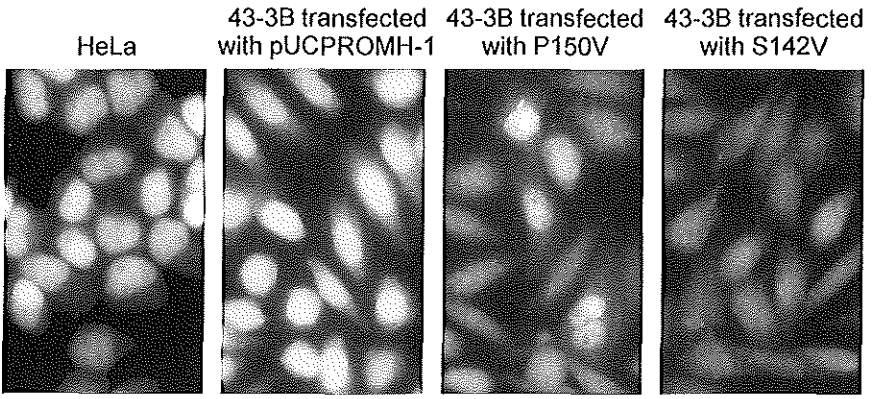


Fig. 7 Expression of (mutated) ERCC1 in single cells. Cells from transfectant P150V are derived from clone (1). Immunofluorescence was carried out using affinity-purified anti-ERCC1 antiserum. HeLa mass population; pUCPROMH-1, P150V and S142V single clones. Like S142V, mock-transfected 43-3B cells displayed no fluorescent staining.

responding dramatic increase in ERCC1 transcripts was found as well (Fig. 8D, compare first lane with last lane). In contrast, ERCC1 induction on the protein level was only approximately 4-fold as estimated by immunoblotting (Fig. 8E). Two-dimensional protein analysis of the transfected cells clearly shows enhanced levels of DHFR and co-amplification of gpt and agpt (two genes not selected for), whereas no protein spot corresponding to ERCC1 could be seen (Fig. 8F). Similar results were obtained with a number of other transformants carrying the amplified functional *ERCC1* gene (data not shown). Apparently, it is not possible to overexpress human ERCC1 protein in mammalian cells.

To analyze the stability of wildtype ERCC1 protein in another manner, purified full-length ERCC1 protein and a ubiquitin-ERCC1 fusion product (Koken *et al* 1993), both overproduced in *E. coli*, were directly injected into the cytoplasm of human primary fibroblasts with the aid of a glass microneedle. Injection of rat serum albumin as a control, resulted in a clear cytoplasmic immunostaining 10 min following injection, which remained fully stable for 1 hr at least. In contrast, a similar number of ERCC1 molecules (representing more than five times the amount normally present in a cell) produced a very weak cytoplasmic staining early (within 10 min) after injection with occasional nuclear staining above background. No exogenous protein could be seen after 1 hr. These micro-injection results strongly suggest that an excess of free ERCC1 is rapidly degraded in the cell. Similar results were obtained for the XPB and XPD repair proteins known to be part of the basal transcription factor TFIIH (our unpublished results).

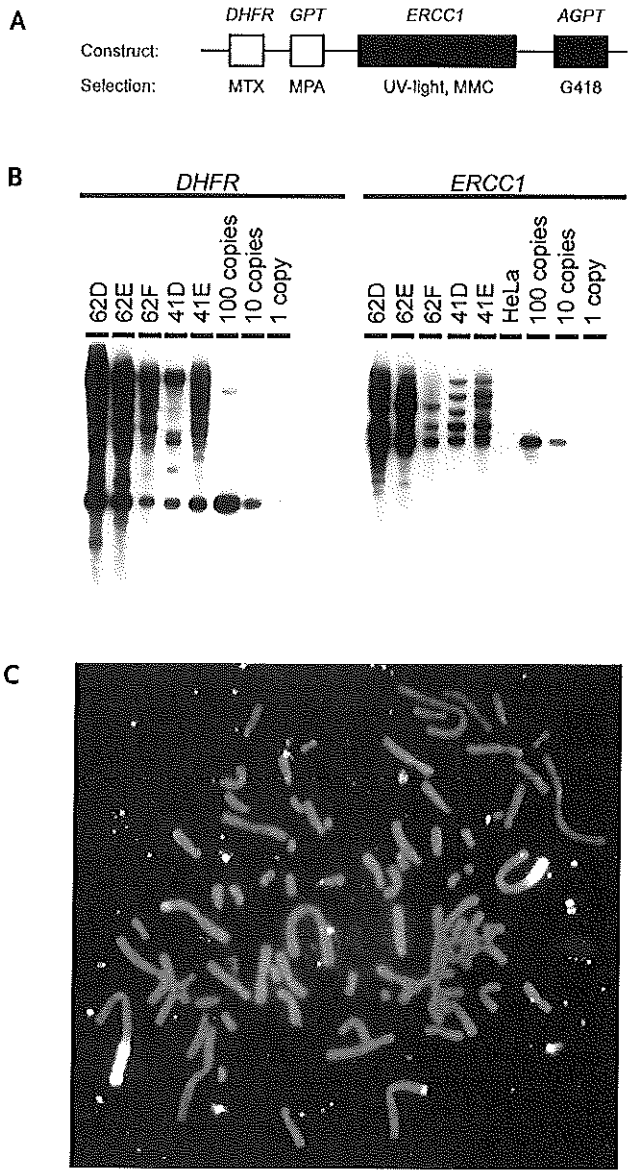
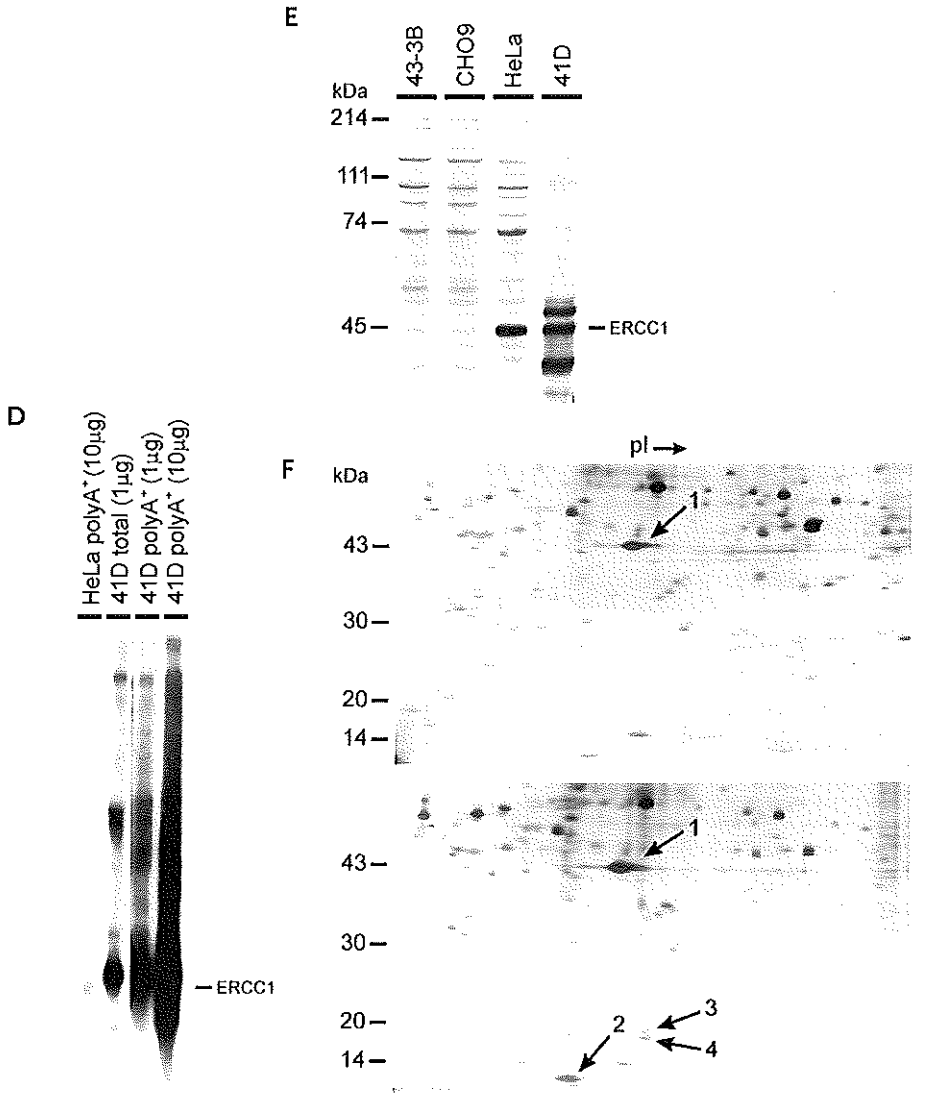


Fig. 8 Characterization of transformant 41D. A. Construct used for the *DHFR*-driven gene-amplification. Black boxes indicate the sequences derived from the *ERCC1* gene on cos34-34 and the pTCF plasmid (containing the *agpt* gene), whereas open boxes represent regions derived from plasmid pHG (carrying the *DHFR* gene) and pSV3gptH (carrying the *gpt* gene). B. Southern blot analysis of *Eco*RI digested DNA (15 µg) from various amplified transformants and 1, 10 and 100 copies of the transfected construct per genome with a ^{32}P-labeled *ERCC1* or *DHFR* probe (pHG) to quantify the amplification. C. *In situ* hybridization on 41D metaphase spreads with a *DHFR* probe.



D. Northern blot analysis of total and polyA$^+$ RNA from HeLa and 41D cells. For hybridization a ^{32}P-labeled *ERCC1* probe was used. Mature ERCC1 transcript is indicated, but note that precursor and incompletely spliced ERCC1 transcripts hybridize as well. Compare the first and last lanes, both containing equal amounts of polyA$^+$ RNA from HeLa and amplificant 41D. E. Immunoblot analysis of 43-3B, CHO9, HeLa, and 41D extracts (20 mg) using crude anti-ERCC1 antiserum. Full-length ERCC1 protein is indicated. F. Two-dimensional protein analysis of 43-3B and 41D whole cell extracts. The 2-D gel was silver stained. (1) actin, (2) DHFR, (3) agpt, (4) gpt. Note that (co-amplified) ERCC1 (39 kDa), supposed to be present in the area between these indicated proteins, is not detectable.

Discussion

ERCC1 mutations were assayed for complementation of the UV sensitivity (NER defect) and MMC sensitivity (recombination defect) of the rodent group 1 mutant 43-3B. In this mutant endogenous ERCC1 protein is hardly detectable (Wood personal communication) and will not compete with the human counterpart for complex formation in the transfectants. By deletion analysis of ERCC1, the minimal essential size of the protein for both of its repair activities could be deduced. From the N-terminus, one third of the ERCC1 protein (91 amino acids) can be removed without loss of correcting ability. This finding indicates that this region (van Duin *et al* 1986) is not required for the NER nor cross-link repair function of ERCC1. Consistently, this region is poorly conserved when compared with *S. cerevisiae* Rad10 (van Duin *et al* 1986) and largely absent in the *S. pombe* homolog Swi10 (Rödel *et al* 1992). However, a cysteine to tryptophan substitution (C76W) within this nonessential part results in a nonfunctional protein (Table 1), pointing to a possible role in protein folding. Removal of 102 N-terminal amino acids fully inactivates ERCC1. This deletion may affect the transient association of ERCC1 with the damage recognition protein XPA, since this interaction involves amino acids within the region of residues 93 to 120 of ERCC1 (Li *et al* 1994). In addition or alternatively, based on the homology between Rad10 and ERCC1, removal of the 102 residues may abolish the formation of a complex of ERCC1 with the human homolog of Rad1. The stretch of residues 90 to 210 in Rad10 has been implicated in the binding of Rad1 (Bardwell *et al* 1992).

Within the central area, missense mutations were introduced affecting the best conserved part between amino acid positions 138 and 150. Most of these mutated *ERCC1* cDNAs produced reduced amounts of protein and could not fully complement the repair defect of the recipient cells. The most plausible interpretation of these findings is that all the different point mutations affect protein stability, probably by interfering with complex formation with ERCC4/ERCC11/XPF. Free ERCC1 molecules are highly unstable inside the cell as was shown for an excess of wildtype ERCC1 introduced transiently by micro-injection or by continuous overexpression in stable amplificants. In line with this observation, the amount of ERCC1 protein in human XP-F and rodent group 4 and 11 cells is strongly reduced (Biggerstaff *et al* 1993, van Vuuren *et al* 1995), whereas the *ERCC1* gene itself does not carry any mutation and is properly expressed at the mRNA level (our unpublished observations). The transfectants expressing detectable (but lowered) levels of ERCC1 protein showed a partial correction. This is consistent with the idea that this central area is needed for the interaction with ERCC4 and stability of the protein. Interestingly, the UV sensitivity of the *S. cerevisiae rad1-20* mutant is caused by a mutation in the Rad10 binding domain of Rad1 and is partially corrected by overexpression of Rad10 protein, presumably increasing the concentration of active Rad1–Rad10 protein complex (Siede *et al* 1993).

In those cases where diminished amounts of mutated protein were detected, the repair of UV damage (NER) was consistently more impaired than the repair of cross-links (recombination). No mutation was found that did affect cross-link re-

pair and not NER. It appears that lower levels of the ERCC1 complex are required for cross-link elimination than for UV lesion removal. Either the number of interstrand cross-links is very low, such that small amounts of ERCC1 complex are sufficient, the ERCC1 complex is more active or not the rate-limiting step in cross-link repair. Some exceptional rodent group 1 and 4 mutants exhibit just a moderate cross-link sensitivity combined with a full UV impairment (Busch *et al* 1994). We have found the same for cells from XP-F patients (our unpublished results), explaining why XP group F presents a NER deficiency rather than a deficiency in cross-link repair.

Several groups reported that increased levels of ERCC1 transcripts correlate with increased cisplatin resistance of human cells (Geleziunas *et al* 1991, Dabholkar *et al* 1992, Dabholkar *et al* 1994). However, we found only an approximately 4-fold increase in ERCC1 protein, despite a massive increase of ERCC1 transcripts (Fig. 8), and no elevated resistance to MMC in overproducing cells (Belt *et al* 1991, our unpublished results). Thus, ERCC1 protein levels should be determined, before conclusions can be drawn with respect to an involvement of this protein in cisplatin resistance. Consistent with this cautious note and with our idea that small amounts of ERCC1 complex are already sufficient for the cross-link repair function, no elevated ERCC1 protein levels were found in nitrogen mustard-resistant cells (Bramson *et al* 1995), indicating that increased ERCC1 levels are not involved in resistance to this cross-linking agent.

At the C-terminal end, no more than four residues appear to be dispensable for both ERCC1 functions. An ERCC1 protein lacking the C-terminal five amino acids, although stable, failed to correct the UV and MMC sensitivity of 43-3B cells. Residue -5 is close to the point where the homology of ERCC1 with the C-terminus of the *E. coli* UvrC repair protein ends (van Duin *et al* 1988). Interestingly, the C-terminus of UvrC itself is also essential for its endonuclease function (Lin and Sancar 1991), though residues that are thought to be directly involved in the incision activity of UvrC may be located elsewhere (Lin and Sancar 1992). It was shown that the *Bacillus subtilis* UvrC protein can substitute the *E. coli* UvrC protein in the UvrABC excinuclease, despite their low homology (38%) (Lin and Sancar 1990). Interestingly, residues conserved between in these two proteins are also present in ERCC1 and are therefore likely to be important for the nuclease activity. A database search revealed the presence of two small subdomains homologous to this essential C-terminal part in a large group of proteins implicated in either DNA break induction or sealing. Representatives of each class of proteins are aligned in Fig. 9. In addition to the known prokaryotic UvrC homologs, inducing the 5' (and possibly also the 3') incision during NER (Lin and Sancar 1991, Lin and Sancar 1992, Moolenaar *et al* 1995), this group includes homologs of RadC, a protein active in recombination-dependent repair of DNA breaks (Felzenszwalb *et al* 1992), and NAD-dependent DNA ligases. Furthermore, residues within subdomain 1 were found to be conserved in a number of other nucleases among which the 5'nuclease domain of *Taq* polymerase (Lawyer *et al* 1989), the human flap-endonuclease FEN-1, equivalent to the 5'→3'endonuclease of *E. coli* DNA polymerase I (Robins *et al* 1994), and many of the bacterial members of the 5'nu-

clease family described by Gutman and Minton (*E. coli* polymerase I amino acids 188-212 Gutman and Minton 1993). The latter region constitutes the last part of the strongly conserved I-region in FEN-1 shared with the XPG and *S. cerevisiae* Rad2 nuclease (Harrington and Lieber 1994a) generating the 3' incision in the eukaryotic NER reaction. The crystal structure of *Taq* polymerase reveals that this area adopts a specific α-helix–turn–α-helix conformation followed by a long loop and two helices (Kim *et al* 1995). Its role in the catalysis of the nuclease reaction has not yet been resolved.

This evolutionary evidence strongly suggests that the UvrC-homologous domain is somehow involved in the activity of the ERCC1 protein, supporting a direct role of ERCC1 in the incision 5' of the DNA lesion. An *ERCC1-UvrC* hybrid gene, however, failed to complement the repair defect of the rodent group 1 mutant, indicating that the C-terminal regions of UvrC and ERCC1 have diverged too much to allow domain swapping. In this regard it should be noted that the C-terminus of UvrC stops at the -6 position in ERCC1 (van Duin *et al* 1986), *i.e.* just beyond the -4 residue critical for both ERCC1 repair functions. The presence of detectable levels of the crucial C-terminally truncated proteins further supports the idea that this area is required for the catalysis rather than for stabilization. An ERCC1 protein with two extra residues at position 208 (see pcDEMP2 in Fig. 1) is also stable in the cell and when strongly overexpressed it exerts a dominant negative effect (Belt *et al* 1991). The poisoning of the ERCC1 complex by this mutant protein suggests that the catalytic domain may extend from the C-terminal end to residue 208 at least. The conservation of the UvrC homology in mammalian ERCC1 and *S. pombe* Swi10 contrasts with its complete absence in *S. cerevisiae* Rad10 (see Fig. 4). Nevertheless, purified Rad1–Rad10 is capable of incision (Sung *et al* 1993b, Tomkinson *et al* 1993, Tomkinson *et al* 1994). A possibility is that the Rad1–Rad10 nuclease can do without this domain. Perhaps more likely, cryptic sequences from the distinct N-terminal part of Rad10 can provide this function, or alternatively, stretches in Rad1 that have no match in its *S. pombe* homolog Rad16.

In conclusion, analysis of mutations introduced throughout the coding area of ERCC1 has revealed dispensability of the poorly conserved N-terminal third of the protein, contrasting with a much more stringent need for the C-terminus. Mutant protein stabilities and local sequence conservation in many DNA break-processing proteins, suggest that the C-terminal domain is primarily required for enzymatic activity of ERCC1, presumed to be a structure-specific endonuclease. The central region of the protein appears to be involved in protein–protein interactions needed for protection against degradation. The repair of cross-links requires lower amounts of ERCC1 than does NER, which could explain the cross-link resistance of XP-F cells and may indicate that the ERCC1-dependent step in this process occurs outside the context of NER. To confirm these findings at the protein level the isolation of the other complex components is underway.

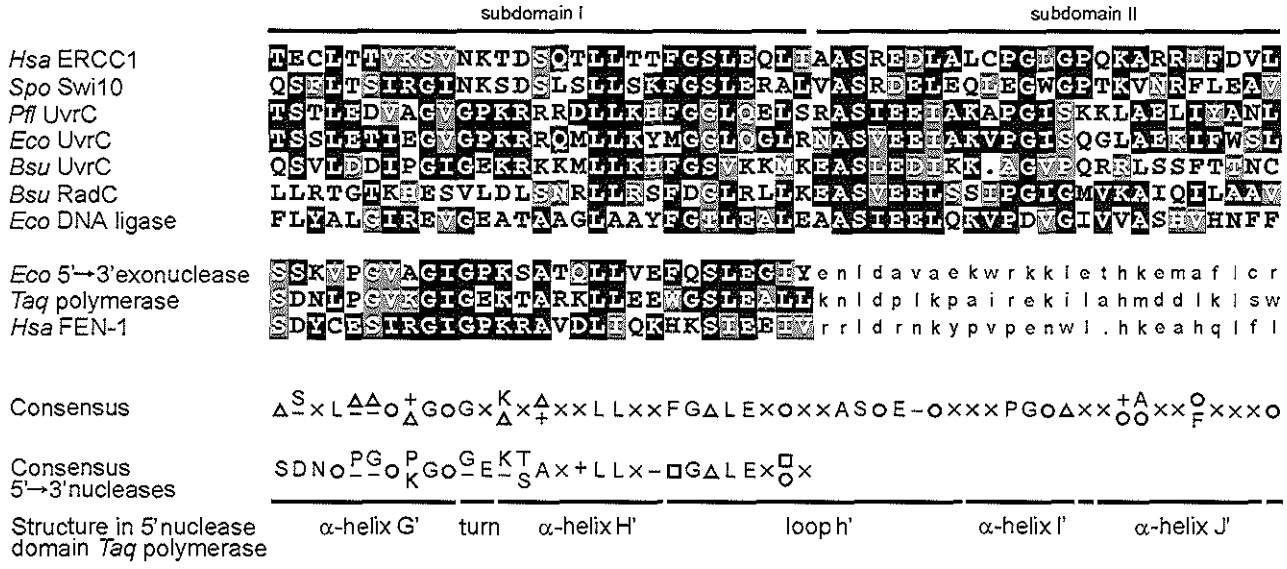


Fig. 9 Homology of the UvrC-like C-terminal domain of ERCC1 with other proteins. Shown is part of the human ERCC1 (297 amino acids) residues 236-290, *S. pombe* Swi10 (252 amino acids) residues 178-232, *Pseudomonas fluorescens* UvrC (607 amino acids) residues 551-605, *E. coli* UvrC (588 amino acids) residues 555-608, *B. subtilis* UvrC (598 amino acids) residues 536-589, *B. subtilis* RadC (231 amino acids) residues 40-94, *E. coli* NAD-dependent DNA ligase (671 amino acids) residues 511-565, an *E. coli* potential 5'→3'exonuclease (251 amino acids) residues 176-230, *Thermus aquaticus* polymerase (832 amino acids) residues 190-244, and the human FEN-1 protein (380 amino acids) residues 232-295 (numbers according to van Duin *et al* 1986, Rödel *et al* 1992, Gaffney *et al* 1994, Sancar *et al* 1984, Chen *et al* 1989, Levin *et al* 1992, Ishino *et al* 1986, Shao and Newman 1993, Lawyer *et al* 1989, and Murray *et al* 1994). Identical amino acids are in black boxes and physicochemically related residues are in gray boxes. The consensus sequence is indicated. Symbol designation: ×, any residue; ○, L V I M; △, S T A G P; −, D E; +, K R H; □, W Y F.

Materials and Methods

Plasmids

The *E. coli* expression construct pETUbi..ERCC1 described earlier (Koken *et al* 1993), encodes an ubiquitin-ERCC1 fusion protein, in which the ubiquitin moiety is thought to protect the N-terminus of ERCC1 against proteolytic degradation. The N-terminal ubiquitin part can be cleaved off by the enzyme ubiquitin lyase (see Koken *et al* 1993). Plasmid pSVL5 is a modification of the pSVL eukaryotic expression vector (Pharmacia) in which the *Eco*RI, *Sal*I, *Kpn*I, and *Hind*III sites are removed. Subsequently, the *ERCC1* cDNA, isolated via PCR, was cloned behind the strong SV40 late promoter giving rise to plasmid pSVL5E (5.8 kb). The PCR-derived *ERCC1* cDNA insert was verified by sequence analysis. Plasmid pUCPROMH-1 (4.2 kb), containing the wildtype *ERCC1* cDNA under the control of its own genomic promoter, has been described previously (van Duin *et al* 1987). Plasmids pRSVneo and pSV3gptH respectively harbor the dominant selectable marker genes *neo* and *gpt* (Gorman *et al* 1983, Westerveld *et al* 1984). Plasmid pHG containing the *DHFR* gene (O'Hare *et al* 1981) was used to drive gene amplification in mammalian cells.

Construction of mutant cDNAs

Missense and C-terminal deletion mutations in *ERCC1* were introduced using site-directed mutagenesis (Kunkel *et al* 1987). The complete *ERCC1* cDNA together with its promoter region was inserted in M13mp18 (Pharmacia) giving rise to Mp18PROM. After mutation induction this insert was used to replace the wildtype *ERCC1* cDNA in plasmid pUCPROMH-1.

The *ERCC1-UvrC* hybrid construct consists of the *ERCC1* cDNA in which the C-terminus, conserved between ERCC1 and the *E. coli* NER protein UvrC, is replaced by the C-terminus of *UvrC* (*ERCC1*(1-708)-*UvrC*(1600-1767)). The *ERCC1* part was amplified using a forward primer containing an optimal translation initiation sequence and a reverse primer containing *ERCC1* (697-708) and *UvrC* (1600-1617) sequences. The complementary oligonucleotide was used (as forward primer) to amplify the C-terminus of *UvrC*. The two amplified fragments were used as template in a subsequent PCR to amplify the *ERCC1-UvrC* hybrid gene.

N-terminal deletion mutations were made via PCR using sense primers containing an optimal translation initiation sequence. The *ERCC1-UvrC* hybrid gene and N-terminal deletion mutants were cloned into plasmid pSVL5E, replacing the wildtype *ERCC1* cDNA.

All mutations were verified by sequence analysis. Furthermore, at least two separate cDNAs were used to assess the biological effects.

DNA transfections

Wildtype and mutated *ERCC1* cDNAs were co-transfected with pRSVneo (in some cases after *in vitro* ligation). 43-3B (ERCC1-deficient CHO) cells (Wood and Burki 1982) were transfected using either the calciumphosphate DNA precipitation procedure (Graham and van der Eb 1973) or lipofectin (BRL) as described (Troelstra *et al* 1992a). Stable transfectants (mass populations or single clones) selected on G418 (800 µg/ml, GIBCO) were checked for the presence of the intact human *ERCC1* cDNA by PCR as described earlier (Troelstra *et al* 1992b).

Survival assays

To determine the colony forming ability (CFA), DNA constructs (5 to 10 μg) were co-transfected with pSV3gptH (2 to 5 μg) into $5 \cdot 10^5$ 43-3B cells in three 90 mm dishes, as described previously (Westerveld *et al* 1984). After 10-14 days of selection on mycophenolic acid (MPA, GIBCO) and mitomycin C (MMC, Kyoma) the cells were fixed, stained and colonies were counted, providing a rough estimate of the survival. To more precisely determine the correcting ability of the mutated *ERCC1* cDNAs, cells of 43-3B, its parental cell line CHO9 and stable transfectants were plated at densities varying from 200 to 1000 cells per 60 mm dish. After attachment, cells were either rinsed with phosphate-buffered saline (PBS) and UV irradiated at various doses (Philips TUV low pressure mercury tube, 15 W, 0.45 J per m^2 per sec, predominantly 254 nm) or incubated with different doses of MMC. The numbers of surviving colonies were counted in triplicate dishes.

In some experiments the presence of non-proliferating giant cells hampered accurate colony counting. Therefore, S phase-dependent ^{3}H-thymidine incorporation, as a measure for the number of proliferating cells, was determined as well. To this end, 500 to 5000 cells were seeded in 30 mm wells and either rinsed and UV irradiated or incubated with MMC or cisplatin (*cis*-diamminedichloroplatinum(II), Lederle) for 1 hr. Seven days later, before reaching confluency, the cells were incubated with ^{3}H-thymidine (2 μCi per ml) and 20 mM HEPES for 1 hr, rinsed twice with PBS and incubated for a further 1 hr in unlabeled medium to deplete radioactive precursor pools. Then, cells were lysed in alkali and radioactivity was quantified by scintillation counting. The two methods used to determine mutagen sensitivity, the classical CFA assay and the rapid and simple ^{3}H-thymidine incorporation assay, gave essentially the same results.

Immunoblotting

Total cell extracts of stable transfectants (90 μg) were checked for the presence of (mutant) ERCC1 protein on immunoblots using affinity-purified anti-ERCC1 antiserum (van Vuuren *et al* 1993).

Two-dimensional electrophoresis was carried out as described by O'Farrell (O'Farrell 1975). The proteins were first separated according to their isoelectric point (pI) and subsequently at right angles en masse by SDS electrophoresis in a polyacrylamide gradient (7.5-20%) gel.

ERCC1 *amplification*

Cosmid 43-34 carrying the *ERCC1* gene, the *gpt* and the *agpt* markers (Westerveld *et al* 1984) was ligated to pHG containing the *DHFR*-gene and transfected to 43-3B cells. Initially, the transfected cells were grown in medium containing MPA (25 μg per ml) and MMC (10^{-8} M) to select for the presence of the *gpt* and the *ERCC1* gene, respectively. In parallel, part of the transfected cells was treated with UV-light and MPA. Both were followed by a selection on 10 μg per ml methotrexate (MTX, Lederle). By stepwise increasing the MTX concentration from 10 to 500 μg per ml amplification of the *DHFR* gene together with its flanking sequences was induced. Stably transfected clones were analyzed for amplification of the *ERCC1* gene, transcript and protein.

Micro-injection

ERCC1 and ubiquitin-ERCC1 proteins (0.1 pg), purified from overproducing *E. coli* (Koken *et al* 1993), were injected into the cytoplasm of human primary

fibroblasts (XP-G cells were used). Rat serum albumin (RSA) was used as a control. Cells were fixed 10 min or 1 hr after injection. Immunofluorescence was carried out using either the anti-RSA or anti-ERCC1 antisera.

In situ *hybridization*

Metaphase spreads of 41D cells were used for *in situ* hybridization with AAF-modified pHG as a probe as described elsewhere (Landegent *et al* 1984). Hybridization was visualized using rabbit anti-AAF and peroxidase-conjugated pig anti-rabbit antisera.

Immunofluorescence

Cells grown on slides were rinsed with PBS and fixed in PBS containing 2% paraformaldehyde for 10 min and in methanol for 20 min. After extensive washing with PBS supplemented with 0.15% glycine and 0.5% BSA the slides were incubated with pre-immune or affinity-purified anti-ERCC1 antiserum (1:100 dilution in PBS) for 1.5 hr at room temperature, rinsed and stained with goat anti-rabbit-FITC antiserum (1:80 dilution) for 1.5 hr. Finally, the slides were rinsed and sealed in Vectashield mounting medium (Vector) containing 4',6'-diamidino-2-phenylindole and propidium iodide (DAPI) as a nuclear marker.

Acknowledgements

We are grateful to M.H.M. Koken, J. de Wit, W. Vermeulen, and G. Weeda for preparing the ubiquitin-ERCC1 construct and for help with transfections and micro-injections. In addition, we acknowledge J. van den Tol, H. Meijers-Heijboer, P. Warmerdam, E. Kootwijk, and A. Overmeer-Graus for construction of several mutated ERCC1 cDNAs and for generation and characterization of the amplificants. M. Kuit is acknowledged for photography. This research was supported in part by the Dutch Foundation for Medical Sciences (GMW, 900-501-113) and by the EEC (project PL00950056).

CHAPTER 3

Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease

Anneke M. Sijbers *
Wouter L. de Laat *
Rafael R. Ariza § *
Maureen Biggerstaff § *
Ying-Fei Wei ‡
Jonathan G. Moggs §
Kenneth C. Carter ‡
Brenda K. Shell ‡
Elizabeth Evans §
Mariska C. de Jong
Suzanne Rademakers
Johan de Rooij
Nicolaas G.J. Jaspers
Jan H.J. Hoeijmakers
Richard D. Wood §

* These authors contributed equally to this work

DEPARTMENT OF CELL BIOLOGY AND GENETICS,
MEDICAL GENETICS CENTER,
ERASMUS UNIVERSITY ROTTERDAM, THE NETHERLANDS

§ IMPERIAL CANCER RESEARCH FUND,
CLARE HALL LABORATORIES SOUTH MIMMS, UNITED KINGDOM

‡ HUMAN GENOME SCIENCES,
INC. ROCKVILLE, UNITED STATES OF AMERICA

Cell 86: 811-822 (1996)

Nucleotide excision repair, which is defective in xeroderma pigmentosum, involves incision of a DNA strand on each side of a lesion. We isolated a human gene homologous to yeast *Rad1* and found that it corrects the repair defects of XP group F as well as rodent groups 4 and 11. Causative mutations and strongly reduced levels of encoded protein were identified in XP-F patients. The XPF protein was purified from mammalian cells in a tight complex with ERCC1. This complex is a structure-specific endonuclease responsible for the 5' incision during repair. These results demonstrate that the *XPF*, *ERCC4*, and *ERCC11* genes are equivalent, complete the isolation of the XP genes that form the core nucleotide excision repair system, and solve the catalytic function of the XPF-containing complex.

Introduction

A coordinated interplay between multiple subunits is required to carry out NER in eukaryotes. The first steps of the process lead to lesion recognition and incision of the damaged strand on each side of a lesion. A 24- to 32-mer oligonucleotide is removed, followed by gap-filling DNA synthesis (Huang *et al* 1992, Friedberg *et al* 1995, Moggs *et al* 1996). In human cells, this repair pathway involves the xeroderma pigmentosum (XP) proteins and associated factors. Individuals with XP show hypersensitivity to sunlight and a greatly increased incidence of skin cancer. Genes encoding the XPA, XPB, XPC, XPD, and XPG proteins have been isolated (Hoeijmakers 1994), and a factor defective in at least some XPE cells has also been identified, although it is not required for the core NER system (reviewed by Wood 1996). XP-F is the only remaining NER-defective XP group for which a complementing cDNA has yet to be identified.

There is evidence that the two incisions made during NER are catalyzed by separate DNA endonucleases. In humans, XPG endonuclease makes the 3' incision relative to the lesion (O'Donovan *et al* 1994a, Matsunaga *et al* 1995). XPG and its yeast homolog Rad2 specifically cleave near junctions of unpaired and duplex DNA, cutting the strand in which the unpaired region moves from 3' to 5' away from the junction (O'Donovan *et al* 1994a, Harrington and Lieber 1994a, Habraken *et al* 1995). In *S. cerevisiae*, the Rad1 and Rad10 proteins form a heterodimeric complex having a structure-specific endonuclease activity with a polarity opposite to XPG and Rad2, leading to the assumption that the Rad1–Rad10 complex makes the 5' incision during NER in yeast (Bardwell *et al* 1994, Davies *et al* 1995). ERCC1 is the mammalian homolog of Rad10 (van Duin *et al* 1986) and has been found to associate with activities that correct human XP-F cell extracts as well as extracts from Chinese hamster cells of repair complementation groups 4 and 11 (Biggerstaff *et al* 1993, van Vuuren *et al* 1993). A polypeptide of relative molecular mass of approximately 115 kDa has been observed to co-purify with ERCC1 by several assays and has been proposed as a candidate for a Rad1 homolog (Aboussekhra *et al* 1995, Park *et al* 1995a, van Vuuren *et al* 1995). We decided to search directly for a human homolog of *S. cerevisiae* Rad1 and test it for possible correcting activity in XP-F and rodent group 4 and 11 cells and extracts.

Results

Isolation of a human homolog of yeast Rad1

A pair of approaches were combined to isolate a human homolog of Rad1. Degenerate primers designed on the basis of homology between *S. cerevisiae* Rad1 (Reynolds *et al* 1987), *S. pombe* Rad16 (Carr *et al* 1994), and *Drosophila melanogaster* Mei-9 (Sekelsky *et al* 1995) were used in a reverse transcriptase-polymerase chain reaction (RT-PCR). The amplified sequence was used as a probe to isolate a human cDNA clone coding for an open reading frame with sequence similarity to the C-terminal half of Rad1 and its homologs. The N-terminal half of the human gene was identified in a database search using *S. cerevisiae* Rad1 and

S. pombe Rad16, which detected two human expressed sequence tag clones encoding an open reading frame that overlapped with the RT-PCR clone. RACE (rapid amplification of cDNA ends)-PCR confirmed the expressed sequence tag sequence. The compiled sequence of the assembled cDNA (shown schematically in Fig. 1A) contains an open reading frame that encodes 905 amino acids (EMBL accession number U64315 for nucleotide sequence). The sequence context of the first ATG in the clone at position 16 matches the consensus translational start site (Kozak 1987), although an initiation site slightly further upstream is not excluded. The predicted protein has a relative molecular mass of 103 kDa and pI of 6.35.
Homology of the protein to Rad1, Rad16, and Mei-9 is most pronounced between residues 699-758 (Fig. 1B), a region located in the Rad10 binding domain of Rad1 (Bardwell *et al* 1993). In the N-terminal half, several leucine-rich repeats are conserved that may be involved in protein–protein interactions (Schneider and Schweiger 1991). The relatively poorly conserved central area harbors putative nuclear targeting sequences (Dingwall and Laskey 1991). Motifs known to be involved in DNA binding or endonuclease function were not found. In a protein sequence property search using the PropSearch algorithm (Hobohm and Sander 1995), the only high scoring hits found with all four Rad1 family members were eukaryotic homologs of DNA mismatch repair proteins MutS and MutL, human Abr proteins, and mouse Rb.
By means of *in situ* hybridization, the *Rad1* homolog was localized to human chromosome 16p13.1-13.2 (Fig. 2A). This corresponds to the locus of a human repair gene complementing rodent NER mutants of group 4, identified using cell hybrids (Liu *et al* 1993) and a genomic clone (Thompson *et al* 1994).

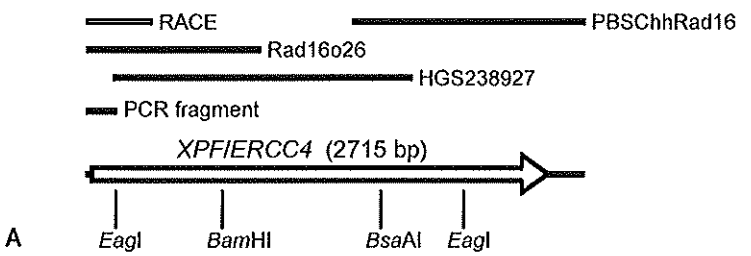
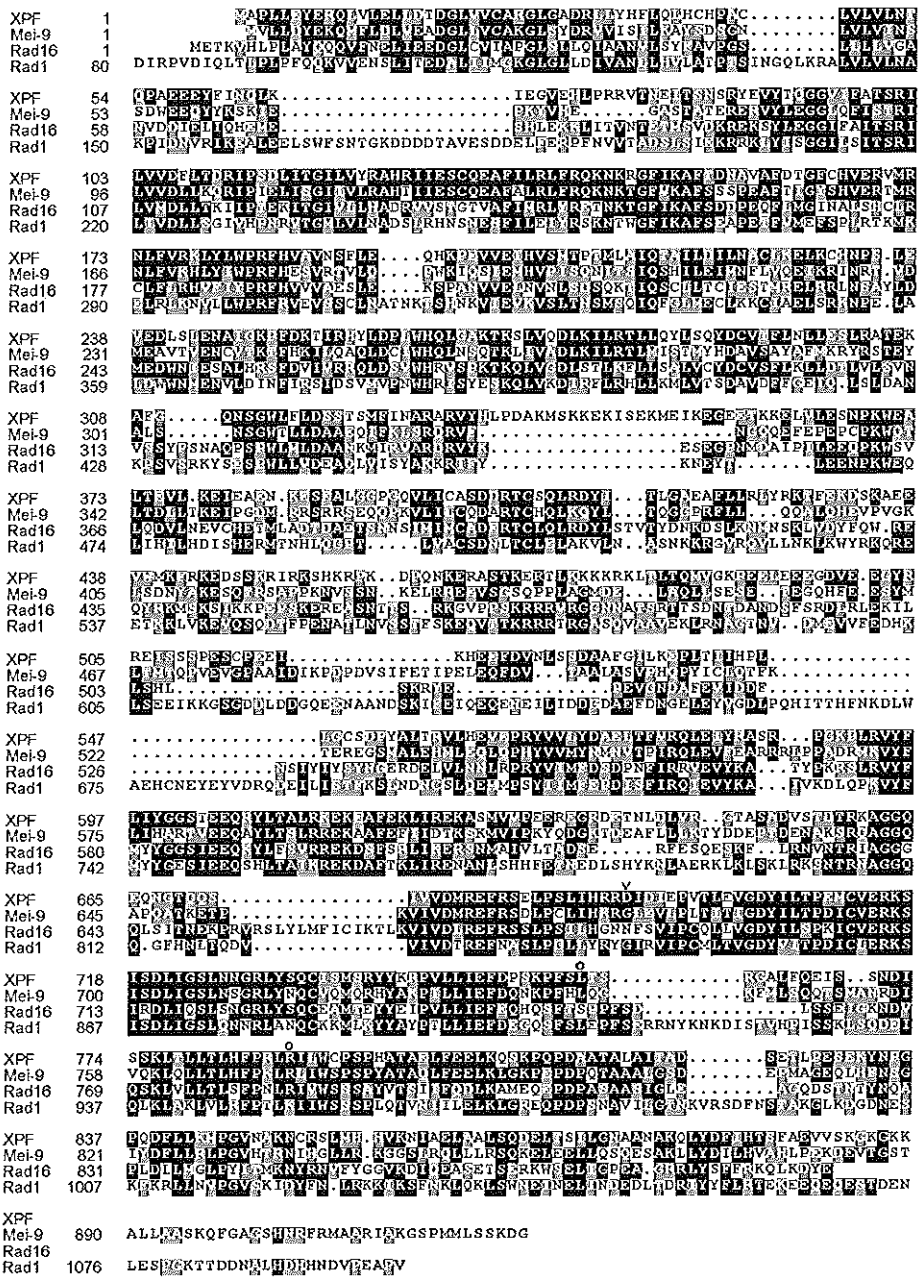


Fig. 1 Cloning of a human *Rad1* homolog. A. Schematic representation of the sources of cDNA used to assemble the functional gene. The open reading frame is indicated by an arrow.



B

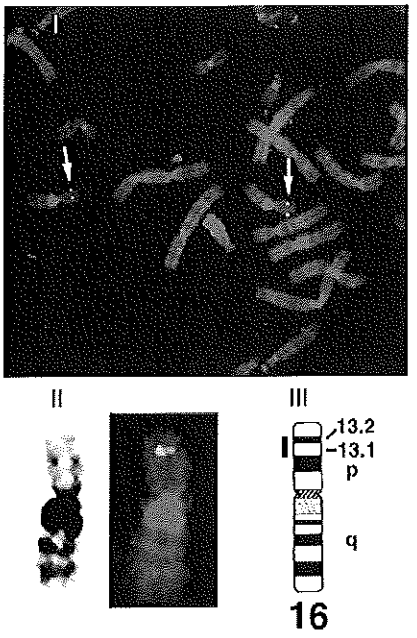
Fig. 1 Cloning of a human *Rad1* homolog. B. Alignment of residues 1 to 905 of the encoded protein with *S. cerevisiae* Rad1, *S. pombe* Rad16, and *D. melanogaster* Mei-9. The starting positions of two mutations in XP-F patient XP126LO are marked with circles above the human sequence, and the position of a polymorphism is marked with a caret.

Correction of XP-F and rodent group 4 and 11 cells by the human gene

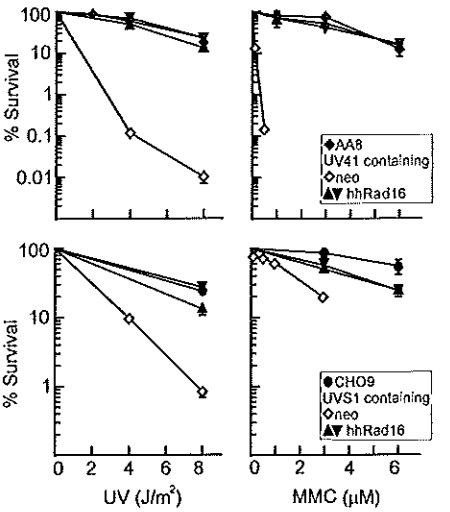
To determine whether the human homolog could correct one of the rodent UV-sensitive mutants, particularly of group 4, the cDNA in a mammalian expression vector (pcDNAhhRad16) was transfected into UV41 (ERCC4$^-$), UVS1 (ERCC11$^-$), and 43-3B (ERCC1$^-$) cells. Stably transfected mass populations of UV41 were obtained that exhibited normal resistance to UV and the cross-linking agent MMC, to which rodent group 4 (and 1) mutants are extremely sensitive (Fig. 2B). This full correction of UV41 was confirmed by the protein expression studies described below and indicates that the cloned cDNA encodes ERCC4. As expected, transfected 43-3B cells did not survive UV selection (data not shown). However, pcDNAhhRad16 also conferred UV- and partial MMC-resistance on UVS1, the only representative of rodent group 11 (Fig. 2B), indicating that it is also the group 11 correcting gene. This was surprising, since complementation between UVS1 cells and two different group 4 mutants, UV47 and UV41, has been previously observed by independent laboratories (Hata *et al* 1991, Busch *et al* 1994). Thus, the two rodent groups represent a unique case of intragenic complementation among mammalian repair genes. The underlying mechanism remains to be elucidated and requires cloning and sequencing of the hamster genes. Interestingly, there is a parallel in *S. pombe*, in which the *rad16.20* allele, which encodes the N-terminal 45% of the protein, was complemented by a plasmid encoding the C-terminal 60% of the *Rad16* gene product (Carr *et al* 1994).

To study further the identity of the human gene product, we raised antibodies in rabbits against recombinant C-terminal fragment. The affinity-purified antiserum recognized a protein band migrating on gels at a relative molecular mass of approximately 115 kDa in extracts from Chinese hamster and normal human cells (Fig. 3A). Chinese hamster mutant UV41 cells transfected with the complementing cDNA regained a band of the same size (data not shown), showing that the cDNA encoded a full-length or near full-length polypeptide. The same protein band was strongly reduced in a HeLa extract depleted for ERCC1 and present in the anti-ERCC1 antiserum-bound fraction (Fig. 3B), confirming its presence in the ERCC1 complex.

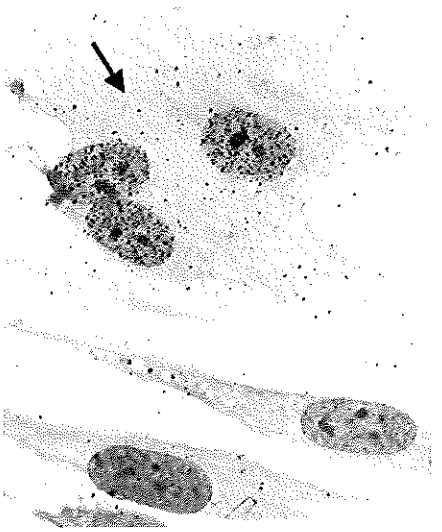
As outlined in the Introduction, ERCC1 and ERCC4 correcting activities are found in a complex that also harbors correcting activity for XP group F. Although we have previously found reduced levels of ERCC1 protein in cells from XP-F patients (Biggerstaff *et al* 1993, van Vuuren *et al* 1995), *ERCC1* was explicitly excluded as the gene responsible for XP-F (van Duin *et al* 1989). To determine whether the Rad1 homolog is involved in XP group F as well, we micro-injected pcDNAhhRad16 into the nucleus of fibroblasts from an XP-F patient. The repair defect was specifically and fully corrected to the level of UV-induced unscheduled DNA synthesis in normal cells (see Fig. 2C, Table 1). Moreover, injection of the ERCC4 antiserum into normal human cells caused specific and complete inactivation of NER (Fig. 2D) but had no effect on transcription (Table 1) or base excision repair (data not shown), two cellular processes distinct from NER. The inhibitory action of the antibodies was competed by pre-incubation with the C-terminal ERCC4 fragment but not by bovine serum albumin (data not shown).


I
II
III
13.2
13.1
p
q
16

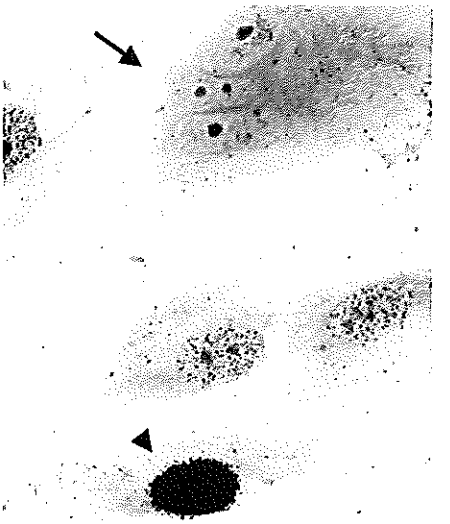
A


100
10
1
0.1
0.01
% Survival
AA8
UV41 containing
neo
hhRad16
CHO9
UVS1 containing
neo
hhRad16
0 2 4 6 8
UV (J/m²)
0 2 4 6
MMC (μM)

B



C



D

Mutations in an XP-F patient

To obtain direct proof that the Rad1 homolog is responsible for the XP-F repair defect, we searched for mutations in the gene. Restriction endonuclease fingerprinting (REF) indicated sequence alterations in the C-terminal part of both alleles from patient XP126LO (Norris *et al* 1988) (Fig. 3E) and in a second unrelated individual with XP-F. RT-PCR clones (2) of each allele were compared by REF analysis with the original PCR mixture, and both were sequenced to rule out PCR-induced mutations. Sequencing both clones of one allele of XP126LO revealed a 4-nucleotide deletion, TCTC, in a repetitive sequence (TTCTCTCTCA) at position 2281, possibly caused by replication slippage, resulting in a frameshift and a truncated protein of 803 amino acids (Fig. 3D). Both clones of the other allele carried a C to T transition at nucleotide 2377, presumably due to deamination of a methylated cytosine at a CpG site, changing arginine residue 788 (conserved in *S. pombe*, Drosophila, and human) into a tryptophan (Fig. 3D, R788W). In addition to these mutational alterations, two sequence polymorphisms were found. One at nucleotide 2090 (A to G) results in Asp or Gly in the human sequence, at a position coding for Gly in *D. melanogaster* and *S. cerevisiae* and Asn in *S. pombe*. A preliminary analysis to estimate the frequency of this polymorphism showed that the Gly residue was present in more than 10 wildtype alleles examined. The second polymorphism at nucleotide 2487 (C to T) does not change the amino acid.
We conclude that the Rad1 homolog is indeed responsible for the repair defect in XP group F. In accordance with nomenclature agreements (Lehmann *et al* 1994), the name *XPF* is recommended for this gene and XPF for its encoded protein.

Fig. 2 The cloned cDNA corrects the rodent complementation group 4 and 11 and XP-F mutant phenotypes. A. (I) Chromosomes (blue) from a single cell showing hybridization to the p arm of each chromosome 16 (red, arrows). (II) Example of a single chromosome 16 with hybridization signal (left: G-band-like diamidophenylindole pattern shown as an inverse black and white image, right: diamidophenylindole-stained chromosome with gene signal). (III) Idiogram of chromosome 16 showing the assigned band position of the *Rad1* homolog. B. Survival of UV41 and UVS1 transfectants after treatment with UV and MMC. Bars represent standard errors of mean. C. Effect of micro-injection of the human *Rad1* homolog on UV-induced unscheduled DNA synthesis in XP-F cells. D. Effect of micro-injection of anti-ERCC4 antiserum on UV-induced unscheduled DNA synthesis in normal cells. Arrows point to injected polynuclear fibroblasts, obtained by cell fusion of XP126LO (XP-F) or repair-proficient normal fibroblasts, both containing three nuclei. Arrowhead indicates a cell in S-phase during incubation with ^{3}H-thymidine to monitor DNA repair synthesis.

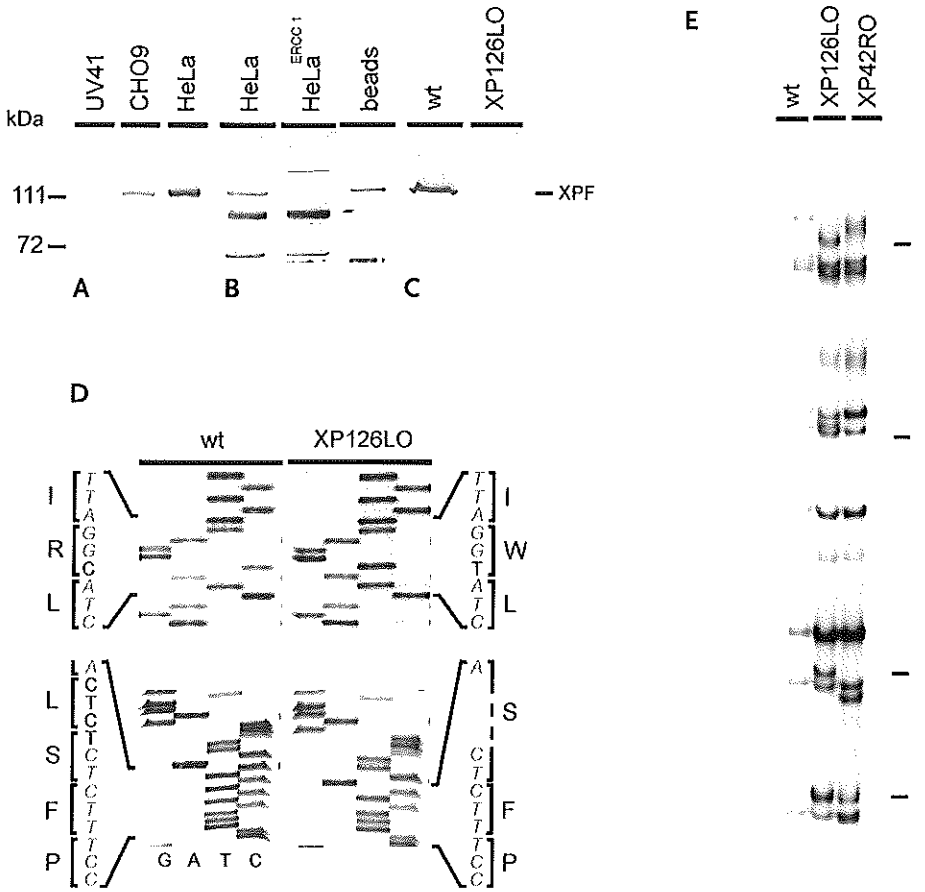


Fig. 3 ERCC4(XPF) protein. A. Expression of ERCC4 protein in UV41 compared with normal Chinese hamster ovary CHO9 cells. B. Immunoblot showing depletion of ERCC4 from a HeLa total cell extract using an anti-ERCC1 antiserum (HeLa^{-ERCC1}) and retention on the antibody beads. The different lanes represent equal amounts of starting material. C. Expression of ERCC4 protein in XP126LO (XP-F) cells compared with SV40-transformed normal cells. The ERCC4(XPF) protein band is indicated. A similar relative difference in the amount of XPF protein was found when total cell extracts of primary fibroblasts from normal and several other XP complementation groups were compared with three XP-F fibroblasts. Equal amounts of whole cell extracts were loaded, and blots were incubated with crude (B) or affinity-purified (A, C) antiserum. Both the hamster and the human XPF protein migrate at 115 kDa. D. Sequence showing mutations in the Rad1 homolog. E. REF analysis in XP-F patient XP126LO. Shown is the C-terminal part after digestion of DNA from XP-F cells (XP126LO and XP42RO) or normal cells with *Eae*I, *Sac*I, and *Stu*I. Aberrant migration is indicated.

Injected substance	Injected cell line	Cellular process assayed	Activity (% of normal)
no injection	XP126LO (XP-F)	NER	15 ± 3
pcDNAhhRad16 (*ERCC4* cDNA)	XP126LO	NER	105 ± 8
pSVL5E (*ERCC1* cDNA)	XP126LO	NER	15 ± 2
no injection	XPCS1BA (XP-B)	NER	7 ± 1
pcDNAhhRad16 (*ERCC4* cDNA)	XPCS1BA	NER	8 ± 1
pre-immune serum	C5RO (normal)	NER	100 ± 8
anti-ERCC4 antibodies	C5RO	NER	4 ± 1
anti-ERCC4 antibodies	C5RO	transcription	104 ± 4

NER activity was determined as UV-induced unscheduled DNA synthesis (see Fig. 2) and transcription was measured as ^{3}H-uridine incorporation. Mean values plus or minus standard errors of the mean are in percentage of normal cells.

Table 1 Micro-injection of cDNA constructs and antibodies into human cells.

Purification of XPF protein in a complex with ERCC1

To study the catalytic function of the ERCC1–XPF protein complex, we purified it from cells producing functional His-tagged ERCC1 protein. ERCC1-defective Chinese hamster 43-3B cells were transfected with His-tagged human *ERCC1* cDNA, selected for repair competence, and extensively characterized to ensure stable expression and function of the His-tagged ERCC1. Immunochemical staining confirmed expression of His-ERCC1 protein in the nuclei of transfected 43-3B cells (Fig. 4A to D), and immunoblotting showed that normal amounts of ERCC1 were produced (Fig. 4E). The transfected His-ERCC1 fully corrected both the UV and MMC sensitivity of 43-3B cells (Fig. 4F and 4G). The use of His-affinity chromatography on chelated nickel columns in combination with five other purification steps resulted in a preparation containing three major polypeptides revealed by silver staining, with relative molecular masses of 42 kDa, 60 kDa, and 115 kDa (Fig. 5A). Immunoblotting identified the 42 kDa band as His-tagged ERCC1 and the 115 kDa band as XPF protein, respectively (Fig. 5B). The ERCC1 and XPF proteins co-eluted at each step, and began to separate from the 60 kDa protein and minor contaminants upon gel filtration chromatography (Fig. 5A, see different peak fractions for ERCC1–XPF and the 60 kDa band), glycerol gradient sedimentation, or on Reactive Yellow 86 agarose (Fig. 5C). These results indicate that the

ERCC1 protein complex is a heterodimer of ERCC1 and XPF, consistent with the suggestion of Park *et al* (1995a). Functional activity was shown by several criteria. The complex could correct the defect in dual incision exhibited by *ERCC1* and *ERCC4* mutant cell extracts, as well as extracts from UVS1 cells (Fig. 6A). Human XP-F cell extracts were also corrected by purified complex, as shown in a repair synthesis assay for NER (Fig. 6B) and in the dual incision assay. The complex did not contain correcting activity for other mutant cell extracts (Fig. 6B). In addition, the complex was active in a fully reconstituted repair system (Aboussekhra *et al* 1995, data not shown).

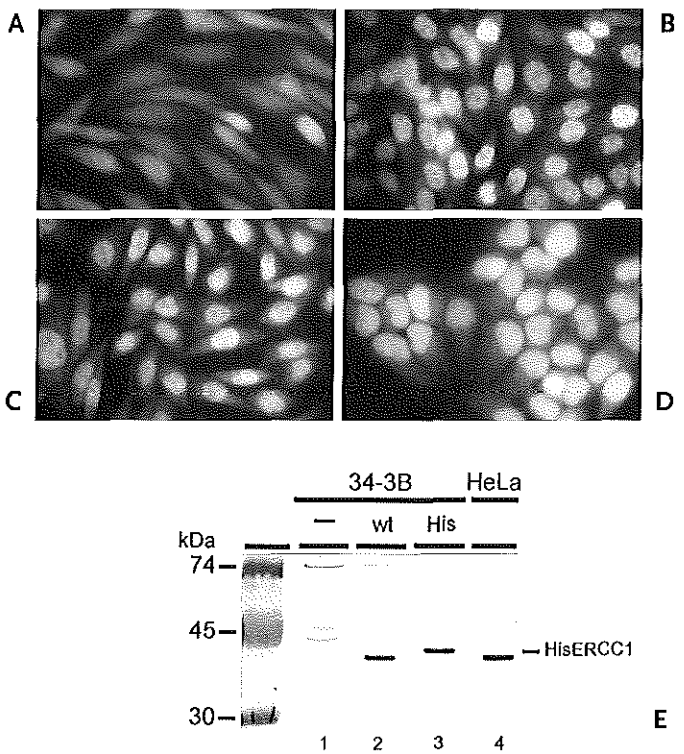


Fig. 4 Functional expression of His-tagged ERCC1 in Chinese hamster 43-3B cells. Nuclear localization of His-ERCC1 protein, detected by immunofluorescence microscopy using anti-ERCC1 antibody in A. 43-3B cells, B. 43-3B cells transfected with His-tagged or C. wildtype ERCC1, and D. HeLa. E. Immunoblot with anti-ERCC1 antibody showing ERCC1 expression levels. Similar protein amounts were loaded. Note that the His-tag causes a mobility-shift on SDS-polyacrylamide gel electrophoresis and that the anti-ERCC1 antibody does not recognize Chinese hamster ERCC1.

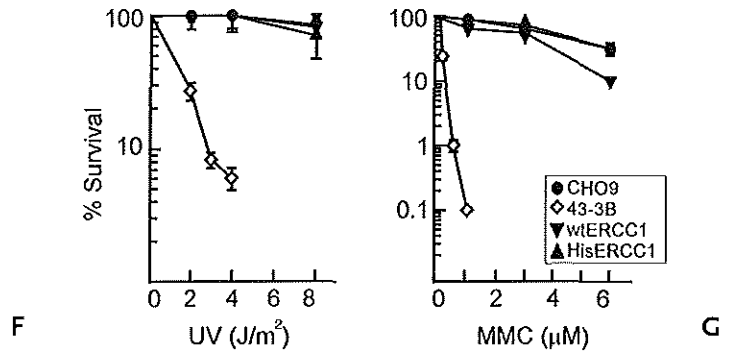


Fig. 4 Functional expression of His-tagged ERCC1 in Chinese hamster 43-3B cells. F and G. Human His-ERCC1 protein fully corrects the UV- and MMC-sensitivity of 43-3B cells.

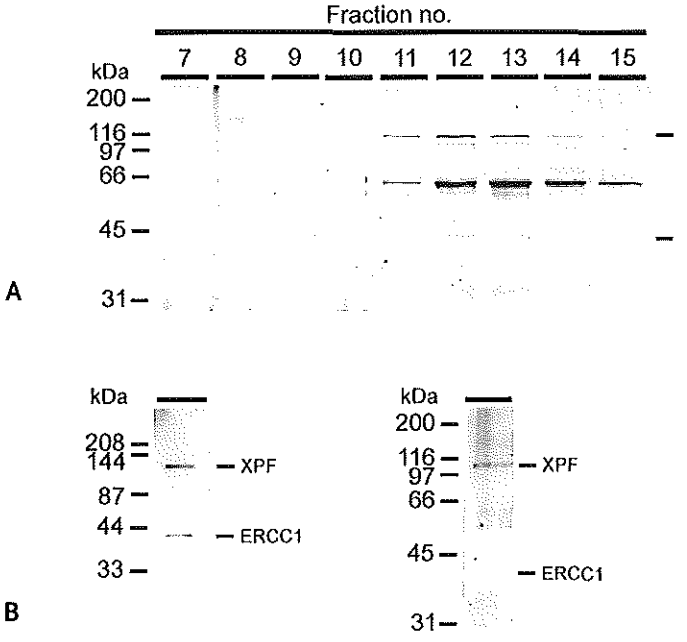
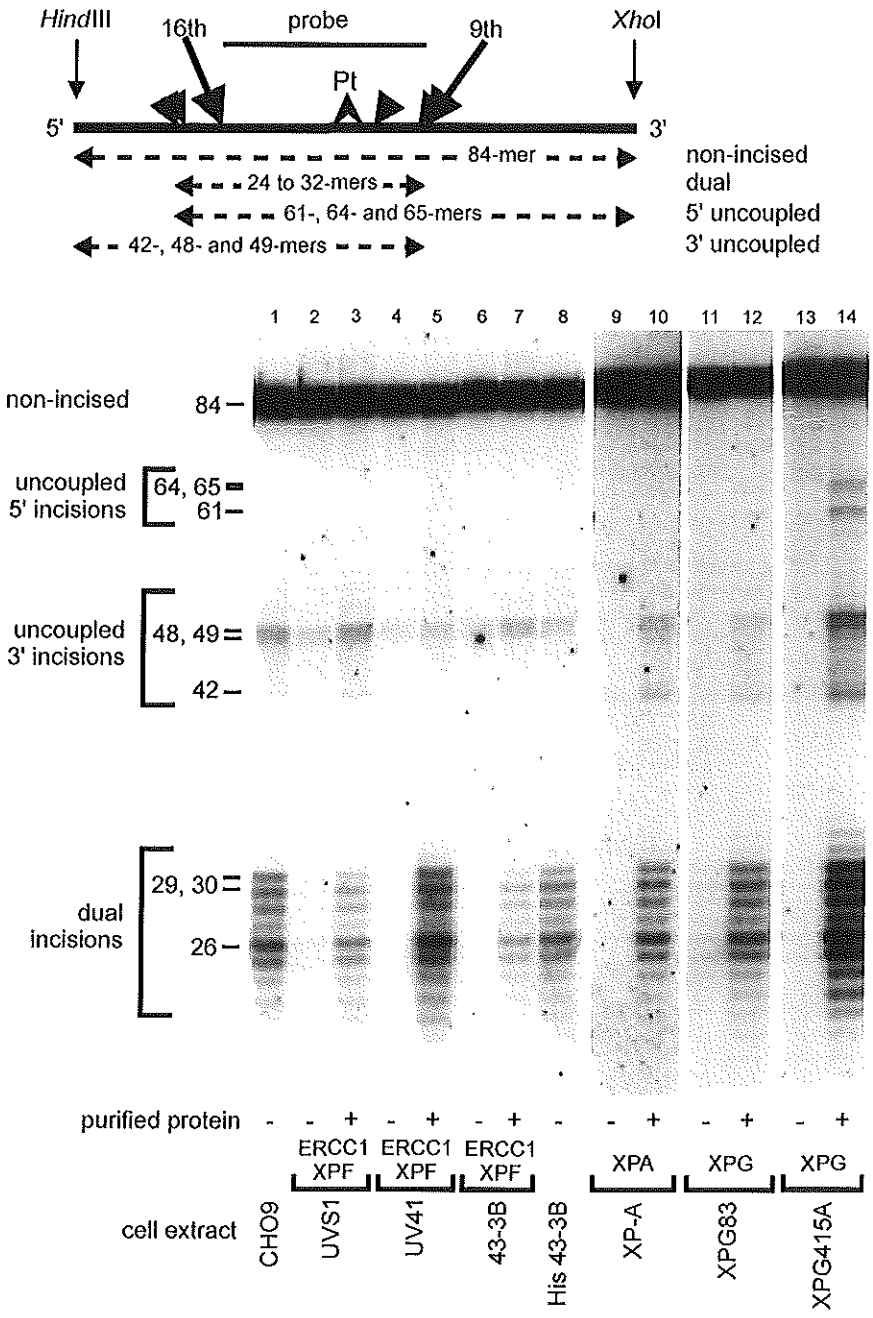
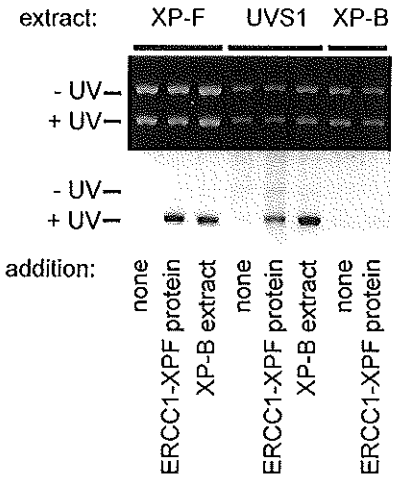


Fig. 5 Purified ERCC1–XPF complex. A. Fraction V was separated on Superose 12 gel filtration column. Fractions were analyzed by SDS-polyacrylamide gel electrophoresis and silver-stained. ERCC1 and XPF are indicated. B. Immunoblot of Fraction V using an antibody raised against full-length ERCC1 and an affinity-purified antibody raised against the C-terminal part of XPF (see Fig. 3). C. Fraction VI after purification on Reactive Yellow 86 agarose.


HindIII
16th
probe
9th
XhoI
Pt
5'
3'
84-mer
non-incised
24 to 32-mers
dual
61-, 64- and 65-mers
5' uncoupled
42-, 48- and 49-mers
3' uncoupled
1 2 3 4 5 6 7 8 9 10 11 12 13 14
non-incised
84
uncoupled 5' incisions
64, 65
61
uncoupled 3' incisions
48, 49
42
dual incisions
29, 30
26
purified protein - - + - + - + - - + - + - +
ERCC1 XPF
ERCC1 XPF
ERCC1 XPF
XPA
XPG
XPG
cell extract
CHO9
UVS1
UV41
43-3B
His 43-3B
XP-A
XPG83
XPG415A

A



B

Fig. 6 Repair-correcting activity of ERCC1–XPF complex. A. Autoradiograph of Southern blot. Platinated DNA fragments corresponding to nonincised DNA (84 nt), uncoupled 5' incisions (61, 64, and 65 nt), uncoupled 3' incisions (42, 48, and 49 nt), and dual incisions (24 to 32 nt) were detected after hybridization with a ^{32}P-labeled 27 mer probe complementary to the DNA sequence surrounding the cisplatin cross-link (Moggs *et al* 1996), as shown in the schematic diagram. Incision reactions were incubated for 30 min using 200 µg of protein from the cell extracts indicated. The complementing factor added to UVS1 (ERCC11), UV41 (ERCC4), and 43-3B (ERCC1) cell extracts (lanes 3, 5, and 7, respectively) was ERCC1–XPF (fraction V, 1 µl, approximately 0.1 pmol complex). Cell extract (200 µg of protein) from 43-3B cells expressing His-tagged ERCC1 was used for lane 8. XPG83 and XPG415A cell extracts were complemented with 50 ng of purified XPG protein (O'Donovan *et al* 1994b, lanes 12 and 14, respectively), and XP-A (GM2345) cell extract was complemented with 90 ng of purified XPA protein (Jones and Wood 1993, lane 10). B. Assay for repair synthesis, with human XP-F cell extract, and extracts from Chinese hamster UVS1 cells and 27-1 cells (a hamster XPB/ERCC3 mutant). Repair synthesis is monitored by incubating cell extracts with a mixture of undamaged (–UV) and UV-damaged (+UV) circular plasmid DNA in a reaction mixture that includes α-^{32}P-dATP (Wood *et al* 1995). Lanes 1, 2, 4, 5, 7, and 8 had 100 µg of the indicated extract (CFII fraction). Lanes 3 and 6 had 50 µg of each CFII fraction. Purified ERCC1–XPF protein (fraction V, 1 µl) was added in lanes 2, 5, and 8.

ERCC1–XPF is a structure-specific DNA endonuclease

To determine whether the purified ERCC1–XPF complex had structure-specific endonuclease activity, we used a partially self-complementary oligonucleotide to form a stem-loop structure consisting of a 22-nucleotide single-stranded loop and a duplex stem of 12 base pairs (Fig. 7A). This substrate was end-labeled on either the 3' or 5' terminus and reaction products were analyzed by comparison to DNA sequencing markers, to map the exact sites of cleavage. The ERCC1–XPF complex specifically cleaved the stem of this substrate 2, 3, and 4 phosphodiester bonds away from the 5' side of the loop (Fig. 7A lanes 2, 3, 6, and 7). These incisions co-localized precisely with those catalyzed by yeast Rad1–Rad10 proteins (Fig. 7A lanes 11, 12, 15, and 16). Conversely, human XPG protein cleaved the substrate on the 3' side of the loop, at the phosphodiester bond on the stem-loop border, and one bond into the stem (Fig. 7A lanes 1 and 8). No cleavage was observed when Mg^{2+} was omitted from reaction mixtures.

Several approaches showed that incision activity was intrinsic to the ERCC1–XPF complex. Structure-specific nuclease activity could be directly followed during the last steps (V and VI) of purification and always co-eluted with the ERCC1 and XPF polypeptides. Furthermore, agarose beads coupled to anti-ERCC1 antibodies could precipitate the nuclease in an active form on the beads, while beads coupled to pre-immune serum could not precipitate active complex (Fig. 7B). These findings clearly demonstrate that incision activity is inherent to the ERCC1–XPF complex. When DNA polymerase I (Klenow fragment) and deoxynucleotides were added after cleavage of the stem-loop, the small DNA incision product was quantitatively converted to approximately 36-mer, indicating the presence of an OH-group at the 3' terminus of the incision product (Fig. 7C).

A "bubble" substrate containing a centrally unpaired region of 30 nucleotides (O'Donovan *et al* 1994a) flanked by duplexes of different sequence was also cleaved by ERCC1–XPF, near the 5' side of the junction between the duplex and unpaired region. However, we did note that ERCC1–XPF cleaved the stem-loop structure more readily than the bubble structure, while XPG preferred the bubble over the stem-loop (data not shown). Additional NER incision step factors such as XPA or RPA were not detectable in the purified preparation and are not required for the structure-specific nuclease activity of ERCC1–XPF with these substrates.

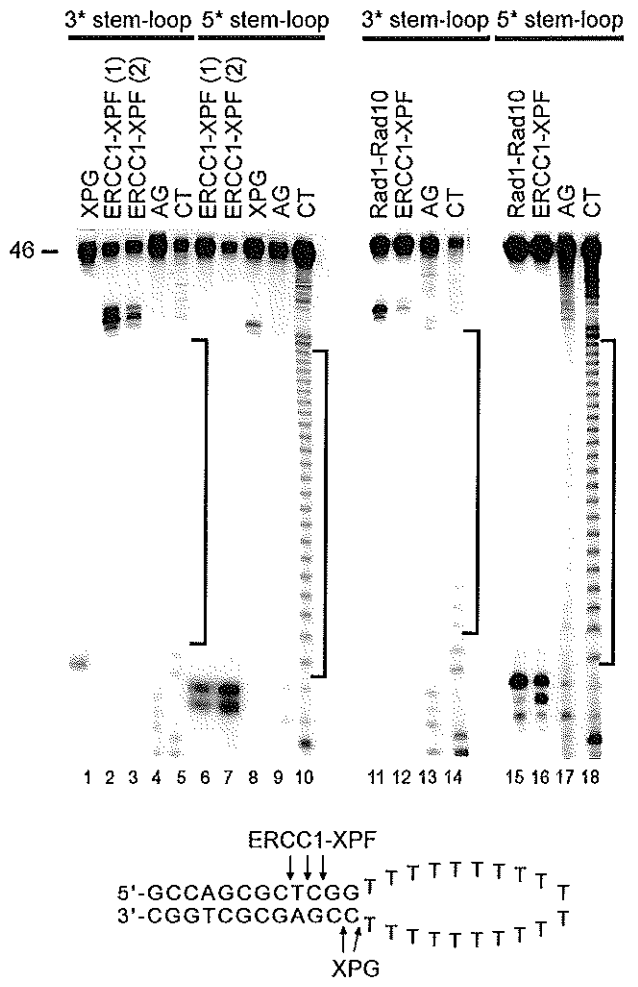


Fig. 7 The ERCC1–XPF complex is a structure-specific endonuclease. A. The structure of the stem-loop substrate is shown below the gel. Arrows indicate the sites of cleavage for the indicated enzymes, as determined from Maxam-Gilbert sequencing ladders. For 5' labeling, Maxam-Gilbert products migrate approximately 1.5 nucleotides faster than the corresponding nuclease products. The 22 T residues are indicated by brackets. Lanes 1 to 5, 3'-labeled stem-loop with lane 1, XPG (50 ng); lanes 2 and 3, ERCC1–XPF (purified fraction V, 1 and 2 µl, respectively); lanes 4 and 5, Maxam-Gilbert G+A and C+T sequencing ladders. Lanes 6 to 10, 5'-labeled stem-loop with lanes 6 and 7, ERCC1–XPF (purified fraction V, 1 and 2 µl, respectively); lane 8, XPG (50 ng), lanes 9 and 10, G+A and C+T sequencing ladders. Lanes 11 to 14, 3'-labeled stem-loop with: lane 11, Rad1 (100 ng) and Rad10 (25 ng); lane 12, ERCC1–XPF (1 µl, fraction V); lanes 13 and 14, Maxam Gilbert G+A and C+T sequencing ladders; Lanes 15 to 18, same as 11 to 14 with 5'-labeled stem-loop.

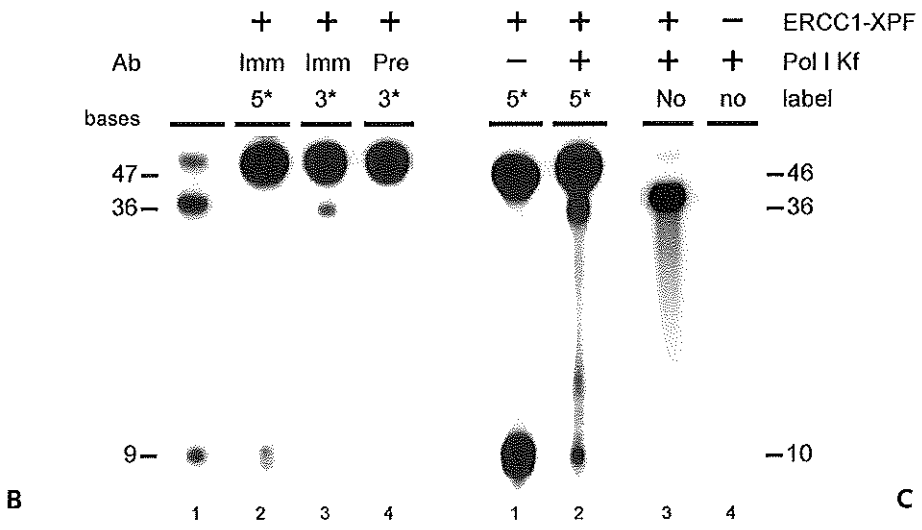


Fig. 7 The ERCC1–XPF complex is a structure-specific endonuclease. B. Nuclease activity isolated on antibody-affinity beads. Lane 1, 5'-^{32}P-labeled marker oligonucleotides. Lane 2, 5'-labeled stem-loop plus affinity-purified anti-ERCC1 antibody-beads (Imm) pre-incubated with purified ERCC1–XPF. Lane 3, as lane 2 but with 3'-labeled stem-loop. Lane 4, as lane 3 but with pre-immune antibody-beads (Pre) pre-incubated with purified ERCC1–XPF. C. Analysis of the 3' end of the cleavage product. Lane 1, 5'-^{32}P-labeled stem-loop plus purified ERCC1–XPF complex. Lane 2, 5'-labeled stem-loop cleaved with ERCC1–XPF, followed by incubation with the Klenow fragment of *E. coli* DNA polymerase I (Pol I Kf) and unlabeled dNTPs. Lane 3, unlabeled stem-loop cleaved with ERCC1–XPF, and then extended with Pol I (Kf) and ^{32}P-labeled dNTPs. Lane 4, control reaction in which unlabeled stem-loop was incubated with Pol I (Kf) and ^{32}P-labeled dNTPs.

ERCC1–XPF makes the 5' incision during repair

These data demonstrate that the ERCC1–XPF complex is a structure-specific endonuclease cleaving one strand of DNA near borders between duplex and single-stranded regions. The polarity of enzymatic cleavage of model substrates strongly suggests that the ERCC1–XPF protein complex makes the 5' incision during NER. This is indeed the case during repair, as shown in Fig. 6A. This assay simultaneously detected excision products resulting from dual incisions as well as uncoupled single incisions in DNA containing a specifically located cisplatin cross-link. Some uncoupled 3' incisions (without the 5' incisions) can be observed in normal cell extracts (see Fig. 6A lane 1, see also Matsunaga *et al* 1995). In very strong reactions, some uncoupled 5' incisions also can be detected (Fig. 6A lane 14, see also Moggs *et al* 1996). XP-G extracts (defective in 3' endonuclease activity) or XP-A extracts

(missing the damage recognition protein XPA) were defective in incision (Fig. 6A lanes 9, 11, and 13). Although ERCC1- or XPF-deficient cell extracts were unable to generate 24-32-nucleotide excision products, fragments corresponding to uncoupled 3' incisions were still detected (Fig. 6A lanes 2, 4, and 6). The ERCC1 and XP-F cell extracts therefore specifically lack the ability to make the 5' incision. Addition of purified ERCC1–XPF complex to the extracts restored 5' incision activity and generated normal excision products (Fig. 6A).

Discussion

Identification of the human XPF *gene*

In this study, a single human cDNA was isolated that can fully correct the faulty repair in XP-F cells as well as in rodent group 4 and 11 cells. Although this has been suggested previously, based on indirect indications (Biggerstaff *et al* 1993, Reardon *et al* 1993b, van Vuuren *et al* 1993, Park *et al* 1995a), this provides a unique demonstration that *XPF*, *ERCC4*, and the group 11 defects are allelic. Intriguingly, the correspondence between rodent groups 4 and 11 is the only known case of intragenic complementation between mammalian NER mutants. The cloning of *XPF* completes the isolation of the set of XP genes that are required for the core process of NER in mammalian cells. The gene is also the final one correcting the six original complementation groups of rodent repair mutants that have been classified by Busch, Thompson, and coworkers (Busch *et al* 1989).

We find that the encoded XPF(ERCC4) protein is intrinsic to an active ERCC1-containing protein complex. The mutations located in the examined XP-F cells occur in the region expected to be involved in ERCC1 binding, based on studies of the homologous yeast complex (Bardwell *et al* 1993). We have noted that ERCC1 mutations interfering with complex formation result in rapid degradation of ERCC1 and that the degree of sensitivity to UV-light and MMC depends on the amount of ERCC1 complex expressed (our unpublished data). The reduced amount of XPF protein detected in extracts of XP-F cells (Fig. 3A), 43-3B, and UVS1 rodent cells (data not shown) closely resembles the strongly reduced levels of ERCC1 protein in these cells that we reported earlier (Biggerstaff *et al* 1993, van Vuuren *et al* 1995) and suggests that complex formation is required for the stability of the XPF component as well. In rodent mutants, hypersensitivity to MMC occurs only at very low levels of ERCC1 (our unpublished data). The residual amount of ERCC1–XPF complex present in XP-F cells may account for their moderate sensitivity to MMC and the slow extended repair characteristic of XP-F (Zelle *et al* 1980).

XPF protein is a subunit of the nuclease that makes the 5' incision

These studies with purified ERCC1–XPF show that the complex has an intrinsic structure-specific endonuclease activity. There are significant parallels with the homologous nuclease complex formed by the Rad1–Rad10 proteins in *S. cerevisiae*. The yeast complex consists of two subunits that are sufficient to perform incision (Sung *et al* 1993b, Tomkinson *et al* 1993, Bardwell *et al* 1994,

Tomkinson *et al* 1994). Our data, including the findings that the 115 kDa component of the human is XPF and is equivalent to ERCC4 and ERCC11, are completely consistent with an heterodimeric composition in mammalian cells as well. Moreover, the cleavage specificities of ERCC1–XPF and Rad1–Rad10 are identical, both on the stem-loop structure (Fig. 7A) and on other substrates. The polarity of cleavage for these enzymes is the opposite of that mediated by XPG nuclease, and the incision sites of ERCC1–XPF are slightly further from the stem-loop junction than those mediated by XPG. Although it was known that XPG and Rad1–Rad10 cleaved near the border of single-stranded and duplex DNA, these data represent the only example in which cleavage sites have been definitively mapped at nucleotide resolution.

In addition, experiments with a specifically placed cisplatin–DNA adduct showed directly that the structure-specific endonuclease activity of ERCC1–XPF is responsible for the 5' incision during the nucleotide excision repair reaction. Finally, it was determined that the product resulting from cleavage by the human enzyme has a 3' OH group, so that gap-filling DNA synthesis can start at the 5' incision site without additional DNA modifications. The fragments produced by dual incision still retain a 5' phosphate (Moggs *et al* 1996), indicating that ERCC1–XPF makes the 5' nick without further processing. Since only the damaged DNA strand is cleaved during nucleotide excision repair, it is likely that lesion-dependent positioning of other incision components such as XPA, RPA, TFIIH, and XPC restricts the action of the two structure-specific repair endonucleases (XPG and ERCC1–XPF) to the damaged strand.

In contrast to the above findings and to those in yeast, a purified preparation containing ERCC1 and a 115 kDa polypeptide from HeLa cells was reported to lack any structure-specific DNA endonuclease activity on a bubble substrate, acting instead as a weak single-stranded endonuclease (Park *et al* 1995a). However, the absence of cleavage in the single-stranded regions of the stem-loop or bubble structures in our assays shows that the ERCC1–XPF complex does not cut single-stranded DNA indiscriminately.

Consequences of XPF–ERCC1 complex inactivation

In the mouse, full inactivation of the *ERCC1* gene results in a severe phenotype causing early death (McWhir *et al* 1993, Weeda and Hoeijmakers unpublished data). This extreme phenotype suggests that the mammalian ERCC1–XPF complex has a function in addition to NER. In yeast, the homologous Rad1–Rad10 proteins are known to engage in a pathway of mitotic recombination by single-stranded DNA annealing (Davies *et al* 1995, Ivanov and Haber 1995). The enzymatic activity demonstrated here for ERCC1–XPF suggests that this complex would also be suitable for such a recombination process in mammalian cells. Impairment of the recombination function of the ERCC1–XPF complex (which is presumably required for repair of MMC-induced DNA cross-links) may underlie the severe phenotype of ERCC1 null mice.

In contrast, the known XP-F patients have relatively mild clinical symptoms, and some are reported to survive into their sixties (Yamamura and Ichihashi 1989). It

remains to be seen whether all group F individuals with the sun-sensitive disorder xeroderma pigmentosum express a low level of ERCC1–XPF complex, and whether a complete absence of ERCC1–XPF is compatible with human viability.

Materials and methods

Isolation of the human XPF cDNA

The 5' part of the gene was isolated by using the *S. pombe* Rad16 and *S. cerevisiae* Rad1 sequences to search a database of human expressed sequence tag sequences (Adams *et al* 1991, Adams *et al* 1995). Of two expressed sequence tag clones detected, the longer (HGS238927) was from a human testis cDNA library. The insert of approximately 2.1 kb encoded a polypeptide with homology to the N-terminal part of Rad16, Rad1, and Mei-9. The HGS238927 insert was used to screen a human testis cDNA library by hybridization, which identified a clone (Rad16o26) with additional 5' terminal sequence. RACE on multiple cDNA libraries confirmed the sequence of Rad16o26 and the presumed AUG initiation codon.

To isolate the 3' part of the gene, numerous degenerate primers were designed for regions in which the *Rad1*, *Rad16* and *D. melanogaster Mei-9* gene products show sequence conservation. In positions where one species differed from two others, the majority rule was followed. At positions where all three differed, the residue in the species most closely related to mammals was utilized. Conditions for PCR amplification were optimized using *S. pombe* cDNA as starting material. Products were evaluated further only if they corresponded to the predicted size. One appropriate primer pair was 5'-TIGTIGAT/CATGA/CGIGAA/GTT-3' and 5'-CIGGIGTIAA/GIATA/GTAA/CTCICCIAC-3'. The protocol for this PCR amplification was 4 min at 94°C (hot start), followed by 35 cycles of 1.5 min at 94°C, 1.5 min at 41°C and 1.5 min at 72°C. The final cycle was concluded with 10 min at 72°C. Reaction mixtures in 60 mM Tris-HCl buffer (pH 8.5) contained 15 mM $(NH_4)_2SO_4$, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 100 pmol primers, and SuperTaq (HT Biotechnology, Ltd.) in a Perkin Elmer Cetus DNA thermocycler. RNA from a non-XP-F primary fibroblast cell line or HeLa cells was used for random hexamer-primed RT-cDNA synthesis using standard conditions. After verification that the amplified product was derived from a human gene with clear homology to Rad1, the PCR product was used as a probe to identify a clone from a human testis cDNA library that encoded a polypeptide with homology to the C-terminal part of the yeast and fly proteins. The insert was cloned into pBluescript, yielding pBSChhRad16. The 5' end of the pBSChhRad16 clone overlapped by 338 bp with the 3' end of the HGS238927 sequence.

To construct a complete cDNA, we first subcloned a 1636 bp *Bsa*AI fragment from pBSChhRad16 into the pBluescript vector carrying HGS238927, to obtain pBSHGSChhRad16. A human fibroblast cDNA library (Keyse and Emslie 1992) was used in a PCR to obtain a fragment containing the most N-terminal sequence flanked by *Sac*I and *Eag*I sites. This PCR fragment was *Sac*I-*Eag*I-digested and separately subcloned into the pBluescript vector carrying HGS238927. An 801 bp *Bam*HI-*Xba*I fragment from the latter construct was inserted into pBSHGSChhRad16 to create a plasmid containing the complete coding sequence, pBShhRad16. The sequence of the assembled cDNA was determined on both strands and matched all partial cDNA clones and RACE sequences. For ex-

pression in mammalian cells, a 2940 bp *NotI-ApaI* fragment from pBShhRad16 was cloned into pcDNA3 (Invitrogen), yielding pcDNAhhRad16.

Mutation analysis

The gene was amplified in two overlapping segments in an RT-PCR using total RNA isolated from XP-F and normal cells. For REF analysis, the products were digested with different sets of restriction enzymes in the presence of shrimp alkaline phosphatase. Pooled digestions were 5' end-labeled with ^{32}P by T4 polynucleotide kinase, denatured, and single strands separated based on their conformation, on a nondenaturing polyacrylamide gel containing 5% glycerol and run at 4°C (Liu and Sommer 1995). Fragments of the two alleles were subcloned and sequenced from XP126LO, XP42RO (both XP-F cell lines), and normal primary fibroblasts (wildtype).

Fluorescence in situ hybridization

The *Rad1* homolog cDNA was nick-translated using Digoxigenin-dUTP (Boehringer Mannheim), and fluorescence *in situ* hybridization was done as detailed (Johnson *et al* 1991). Individual chromosomes were counterstained with diamidophenylindole, and images were recorded using a charged coupled-device camera and analyzed using ISEE software (Inovision Corp.). Spreads (approximately 20) were analyzed by eye, and most had a doublet signal characteristic of genuine hybridization on at least one chromosome 16. Doublet signal was not detected on any other chromosomes. Individual chromosomes (10) were analyzed in detail using a combination of fractional length measurements and fluorescence banding combined with high resolution image analysis.

Selection and characterization of 43-3B His-ERCC1 cells

A tag sequence was introduced at the 3' end of *ERCC1* by PCR methods, to encode a fusion protein consisting of normal full-length human *ERCC1* cDNA followed by Gly-Gly-Ser, a thrombin cleavage site, and six His residues. This construct (pSVL-ERCC1-His) was transfected into 43-3B cells (Wood and Burki 1982) and repair-competent transformants were selected after repeated UV irradiation (4.6 J per m^2). The number of cells surviving treatment with UV or MMC (1 hr) was measured as ^3H-thymidine incorporation 6 days after exposure. Proliferating cells were pulse-labeled with tritiated thymidine (1 hr), followed by a chase (1 hr) to deplete radioactive precursor pools. Cells were lysed in 0.05 M NaOH and transferred to scintillation-counting vials. Survival was calculated as the average ratio of incorporated radiolabel in treated duplicates to that in four untreated control dishes.

Correction of rodent mutants by transfection

pcDNA3 (neo) and pcDNAhhRad16 were transfected into UV41, UVS1 and 43-3B cells using lipofectin, as described previously (Troelstra *et al* 1992a). Stable transfected mass populations were selected on G418 (800 µg/ml) and UV (UV41 three times 8 J per m^2 and UVS1 three times 18 J per m^2). Subsequently, 500 to 5000 cells were seeded in 30 mm wells and either UV irradiated or incubated with the cross-linking agent MMC for 1 hr. Survival was measured as ^3H-thymidine incorporation 7 days after exposure.

Micro-injection

pcDNAhhRad16 or anti-ERCC4 antiserum was injected into one of the nuclei of

XP126LO (XP-F) homopolykaryons or into the cytoplasm of C5RO (normal human primary fibroblasts) homopolykaryons. Repair activity was determined after 24 hr by UV-induced (15 J per m^2) incorporation of ^{3}H-thymidine and autoradiography as described (Vermeulen *et al* 1994a). The number of silver grains above the nuclei is a measure of the level of unscheduled DNA synthesis and reflects the cellular repair capacity.

Immunological methods

An *E. coli* expression construct was made coding for a GST-ERCC4 fusion protein containing the 391 C-terminal residues present in pBSChhRad16. Antibodies raised against fusion protein were isolated on a glutathione column and affinity-purified. Immunofluorescence, SDS-polyacrylamide gel electrophoresis, and immunoblotting were performed according to standard procedures (Sambrook *et al* 1989), using polyclonal antibodies against human ERCC1 and ERCC4. A HeLa whole-cell extract was depleted for ERCC1 using anti-ERCC1 antiserum coupled to protein A beads.

Isolation of ERCC1–XPF protein complex

ERCC1–XPF complex was isolated from 43-3B His-ERCC1 cells, monitoring purification by immunoblotting using an antibody against ERCC1 and by repair-correcting activity. A nuclear extract (Masutani *et al* 1994) was prepared from $1.4 \cdot 10^{11}$ frozen cells and dialyzed against buffer A (20 mM HEPES-KOH pH 7.5, 0.2 mM EDTA, 2 mM $MgCl_2$, 10% glycerol, 0.2 mM 4-(2-aminoethyl)benzenesulfonylfluoride, 5 mM β-mercaptoethanol) containing 0.15 M KCl and 10 µg/ml of aprotinin. The extract (1.5 g) was loaded onto a phosphocellulose column (Whatman P11, 300 ml) equilibrated in the same buffer. The flow-through fraction contains the incision protein RPA and proteins not needed for NER (Shivji *et al* 1992, Biggerstaff *et al* 1993).

Bound proteins (Fraction I, 600 mg protein) were eluted with buffer A containing 1.0 M KCl, supplemented with 1 mM imidazole (Fluka), and applied directly onto a Ni^{2+}-NTA agarose column (Qiagen, 25 ml). The column was washed at 25 ml/hr and eluted sequentially with buffer B (20 mM HEPES-KOH pH 7.5, 2 mM $MgCl_2$, 10% glycerol, 0.2 mM 4-(2-aminoethyl)benzenesulfonylfluoride, 5 mM β-mercaptoethanol, 0.5 M KCl) containing 1 mM, 5 mM (pH 6.6), 20 mM (pH 7.5), and 100 mM imidazole (pH 7.5). ERCC1 was eluted in the last fraction (Fraction II, 35 mg protein), supplemented with 0.02% NP-40 and 1 mM potassium phosphate, and loaded onto a hydroxyapatite column (Bio-Rad, 3 ml). The column was washed with buffer C (25 mM HEPES-KOH pH 7.8, 10% glycerol, 0.2 mM 4-(2-aminoethyl)benzenesulfonylfluoride, 2 mM dithiothreitol, 0.02% NP-40) containing 5 mM potassium phosphate and 0.5 M KCl. ERCC1 was eluted in buffer C containing 30 mM potassium phosphate and 0.5 M KCl (Fraction III, 10 mg). The fraction was dialyzed against buffer C plus 1 mM EDTA and 50 mM KCl, filtered through a 0.45 µm filter (Millipore), and applied onto an FPLC Mono Q HR 5/5 column (Pharmacia). A gradient of 50-300 mM KCl in buffer C was applied, and 0.5 ml fractions were collected. ERCC1 eluted at about 0.2 M KCl. ERCC1-containing peak fractions (Fraction IV, 4 ml) were pooled, diluted to 50 mM KCl with buffer C, loaded onto an FPLC Mono S column and eluted with a gradient of 50-300 mM KCl. Peak fractions (Fraction V, 1.5 ml, approximately 0.1 pmol ERCC1 complex/µl) eluted at approximately 0.22 M KCl. A SMART system (Pharmacia) was used for the Superose 12 gel filtration in Fig. 5A. For Fig. 5C, 20 µl of fraction V was diluted 1:2 in buffer C and added to

10 µl of swollen Reactive Yellow 86 agarose (Sigma) to give 50 µl in 25 mM HEPES-KOH pH 7.8, 10% glycerol, 2 mM dithiothreitol, and 100 mM KCl. After 1 hr at 4°C, the suspension was centrifuged and 30 µl of supernatant loaded onto the gel.

In vitro *DNA repair and nuclease assays*

The assay for dual incision was as described (Moggs *et al* 1996), except that the plasmid was cleaved with *Hind*III and *Xho*I before detection by Southern hybridization. Repair synthesis assays used CFII protein fraction from the indicated cells, purified RPA, and purified proliferating cell nuclear antigen, as previously described (Biggerstaff *et al* 1993). Nuclease reaction mixtures (15 µl) contained 0.2 to 0.5 ng stem-loop DNA and the indicated proteins in buffer as described (O'Donovan *et al* 1994a) for Fig. 7A lanes 1 to 10, or in buffer D (50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 100 µg/ml bovine serum albumin and 0.5 mM β-mercaptoethanol) for lanes 11 to 18. After incubation at 25°C for 2 hr, 15 µl 90% formaldehyde was added, samples were heated at 95°C and loaded onto denaturing 12% polyacrylamide gels. Products were visualized by autoradiography or a phosphorimager. Maxam-Gilbert A+G and C+T reactions were carried out on the 3'- and 5'-labeled stem-loop substrates in order to locate the sites of enzymatic cleavage. For immunoprecipitation, affinity-purified anti-ERCC1 antiserum (30 µl) was coupled to 10 µl protein G beads, and beads were extensively washed and incubated with the indicated amount of ERCC1–XPF complex in 20 µl of buffer D at 4°C. After 2 hr, beads were washed twice with 100 µl of buffer D and added to nuclease reaction mixture. Samples were then loaded on a denaturing 20% polyacrylamide gel. For the end-labeling analysis in Fig. 7C, reaction mixtures containing 5'-labeled (lanes 1 and 2) or unlabeled (lanes 3 and 4) stem-loop substrate were incubated for 16 hr at 16°C, with or without ERCC1–XPF complex (1 µl of fraction V) as indicated. In lane 2, 1 µl of 5 mM dNTPs plus 1 U *E. coli* DNA pol I (Klenow fragment) was added for a further 30 min. For lanes 3 and 4, 5 nmol dGTP, dCTP, and dTTP and 2 nmol α-^{32}P-dATP were added for 30 min, followed by a 90 min chase with 10 nmol unlabeled dATP.

Acknowledgements

We thank the members of our laboratories for advice and reagents, A. A. Davies, R. Sowdhamini, W. Vermeulen, and T. Yagi for discussions and E. C. Friedberg for the Rad1–Rad10 proteins. We thank A. J. van Vuuren, A. de Klein, R. van Os, H. Odijk, and J. de Wit for help with the experiments. This work was supported by the Imperial Cancer Research Fund, the Human Frontiers of Science Program, the Dutch Scientific Organization (Foundation for Chemical Sciences of NWO), and the European Community Human Capital and Mobility Program.

CHAPTER 4

Homozygous R788W point mutation in the XPF gene of a patient with xeroderma pigmentosum and late-onset neurologic disease

Anneke M. Sijbers
Pieter C. van Voorst Vader §
Jos W. Snoek ‡
Anja Raams
Nicolaas G.J. Jaspers
Wim J.Kleijer *

DEPARTMENTS OF CELL BIOLOGY AND GENETICS AND
* CLINICAL GENETICS, ERASMUS UNIVERSITY ROTTERDAM;
§ DERMATOLOGY DEPARTMENT UNIVERSITY HOSPITAL GRONINGEN; AND
‡ NEUROLOGY, MARTINI HOSPITAL GRONINGEN, THE NETHERLANDS

J Invest Dermatol 110: 832-836 (1998)

The second Caucasian xeroderma pigmentosum patient (XP42RO) belonging to complementation group F (XP-F) is described. Mild ocular photophobia was present from childhood, and acute skin reactions occurred upon exposure to sunlight. Basal and squamous cell carcinomas developed after his twenty-seventh year. In his late forties progressive neurologic symptoms emerged, which included intellectual decline, mild chorea and ataxia, and marked cerebral and cerebelar atrophy. Such neurologic abnormalities are very unusual in XP-F. Similar symptoms have been described in only one of 17 other XP-F individuals. His about 5-fold reduced activity of nucleotide excision repair in cultured cells, combined with moderately affected cell survival and DNA replication after UV exposure, are typical of XP-F. The recent cloning of the *XPF* gene allowed a molecular genetic analysis of this unusual patient. XP42RO, representing the second case studied in this respect, turned out to be homozygous for a point mutation in the *XPF* gene, causing an R788W substitution in the encoded protein. Surprisingly, this mutation had also been found in one allele of the other unrelated Caucasian XP-F case. The amount of mutated XPF protein is strongly reduced in cells from XP42RO, presumably due to a conformational change. Biochemical, genetic, and clinical data all indicate the presence of considerable residual repair activity, strongly suggesting that the R788W mutation is leaky.

Introduction

The autosomal recessive disease xeroderma pigmentosum is clinically manifested by photosensitivity of skin and eyes and early onset of freckling and other lesions on sun-exposed skin, culminating in a high propensity to neoplasms, especially carcinomas. Progressive neurologic degeneration occurs in some XP patients, varying from central nervous system abnormalities to, in rare cases, additional peripheral nerve dysmyelination characteristic of Cockayne syndrome (Cleaver and Kraemer 1994). The photosensitivity of XP patients is due to defective NER, a process responsible for the removal of DNA damage induced by UV and diverse chemical mutagens. NER is a complex mechanism, requiring the correct interplay of about 30 different genes (Aboussekhra *et al* 1995, Mu *et al* 1995). Seven of these, named *XPA* to *XPG*, were found to be defective in XP patients. The XPA, XPC, and XPE proteins are probably involved in damage recognition, whereas XPB and XPD mediate subsequent local unwinding of the DNA helix. XPF and XPG are required for strand cleavage at either side of the DNA lesion. In the final stages of NER, the damaged patch is removed and replaced using enzymes from the DNA replication machinery.

Most NER-defective XP patients are affected in *XPA* or *XPC*, followed by *XPD*. *XPF* and *XPG* defects are less frequent and only rare cases with XP-B and XP-E have been described. A patient with a mutated *XPF* gene, first identified in 1979 by genetic complementation analysis (Arase *et al* 1979), showed relatively mild cutaneous symptoms and no neurologic involvement. This picture is characteristic of the XP-F patients reported later on, all from Japan (Nishigori *et al* 1986, Yamamura *et al* 1989) except for one from Europe (Norris *et al* 1988) (see Chapter 6 Table 1). One exceptional Japanese case was identified having neurologic abnormalities (Moriwaki *et al* 1993). Here we describe a European XP-F patient with multiple carcinomas and other mild oculocutaneous symptoms, who developed marked neurodegeneration after his late forties. In view of this unusual clinical picture, we conducted detailed biologic and biochemical studies of his NER capabilities. After the recent cloning of the *XPF* gene (Chapter 3) molecular genetic analysis has also become possible. Here we report the second XP-F patient studied in this respect. To our surprise, we found a homozygous mutation, identical to one of the *XPF* alleles of the first patient.

Results

DNA repair studies

The overall activity of NER was assayed in cultured skin fibroblasts as UDS, *i.e.* UV-induced incorporation of tritiated thymidine was measured by autoradiography. In the patient, NER after exposure to 10 J per m^2 was reduced to 22% of normal levels (see Table 1, top panel) and varied between 15% and 30% on other occasions (not shown). NER in the cells of the father was in the normal range (Table 1). Genetic complementation analysis was performed to identify the gene defect. After fusion of XP42RO fibroblasts with one or two known representatives from all

seven XP groups A through G (for strains see materials and methods), normal or near-normal UDS levels were obtained in all cases except for the XP-F strain XP126LO where UDS remained low in the heterokaryons (data not shown). This result assigns the patient to XP group F. Cell survival studies showed that the relatively low NER activity resulted in only moderate UV cytotoxicity, much less marked than in XP cells from group A (Fig. 1A). This behavior is characteristic of XP-F as shown here for XP2YO (Fig. 1A) and for other XP-F patients elsewhere (Arase *et al* 1979, Nishigori *et al* 1986) (Chapter 6 Table 1). Also the consequences for the overall rates of DNA synthesis, measured 16 hr after UV exposure, were only moderate: the inhibition was much less pronounced than in XP-A cells, and similar to an XP-C strain (Fig. 1B), known to be intermediate in this respect. The cells from the patient's father showed a normal behavior in the latter two assays (only DNA synthesis is shown in Fig. 1B).

Case report

Patient XP42RO, aged 62 years, is one of nine siblings from nonconsanguineous parents of Dutch origin. From about 10 years of age he developed irregularly pigmented macules and telangiectasia on cheeks, temples and dorsa of hands. From his fifteenth to twenty-seventh year he worked outdoors as a farm laborer, experiencing frequent acute sunburn reactions. Since his thirties, nine basal or squamous cell carcinomas have been removed from sun-exposed areas. Mild photophobia and occasional conjunctival vasoinjection were present from childhood. At age 47, his 24 hr and 72 hr MED were 0.08 and 0.04 J per cm^2 of UV-light, respectively, corresponding to the most sensitive type I skin individuals.

After the age of 47, neurologic symptoms started to appear in the form of progressive dysarthria, clumsiness, restlessness, and a mild chorea. Ataxia with unsteady gait, dysmetria of the limbs, and saccadic dysmetria of ocular movement are now present. Nuclear magnetic resonance showed a corresponding marked atrophy of the cerebelum and cerebral cortex. Electroencephalogram, reflexes, and audiogram were normal. Signs of central nerve conduction impairment were noted in somatosensory and motor evoked potentials. Electromyography findings suggested mild axonal polyneuropathy. Neurophysiologic testing revealed deficiencies in attention, concentration, and recall as well as visuospatial and constructional deficits. 1-chloro-2,4-dinitrobenzene testing (Bleumink *et al* 1974) revealed a normal cell-mediated immune response. The patient is married without offspring, due to gonadal dysplasia. The family history revealed that sun sensitivity was also noted in the father, but without the development of severe skin lesions and neoplasms. None of the patient's siblings are known to have similar neurologic or cutaneous problems.

Sample tested	Grains per nucleus[a]	Repair activity[b]
Standard NER assay		
XP42RO (patient)	12.9 ± 0.6 (50)	22
95RD28 (father)	55.4 ± 2.5 (50)	96
C5RO (normal control)	58.7 ± 2.7 (50)	—
Micro-injection of *XPF* cDNA		
XP126LO uninjected	5.5 ± 0.6 (24)	27
XP126LO injected R788W cDNA	12.0 ± 0.7 (27)	58
XP126LO injected wt cDNA	22.9 ± 0.9 (27)	112
C5RO uninjected	20.6 ± 0.5 (29)	—

[a] mean ± SEM (number of nuclei)
[b] % of normal

Table 1 Correction of NER activity in XP42RO

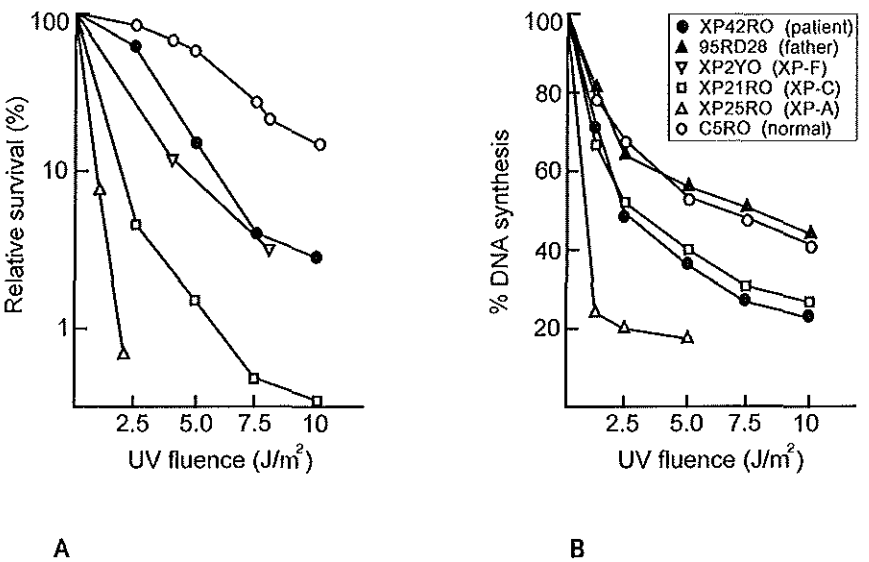


Fig. 1 Cellular responses to UV exposure. A. Survival assay. B. Inhibition of DNA synthesis measured 16 hr after various doses of UV. All points represent averaged duplicate measurements.

Molecular genetic studies

Recently, the *XPF* gene was cloned (Chapter 3), which allows mutational analysis of the affected gene in this patient. Initial REF analysis by single-strand conformation-sensitive gel electrophoresis indicated the presence of specific sequence abnormalities in the downstream end of cDNA amplified from XP42RO mRNA. These were partly similar to the ones observed in XP126LO, another XP-F indi-

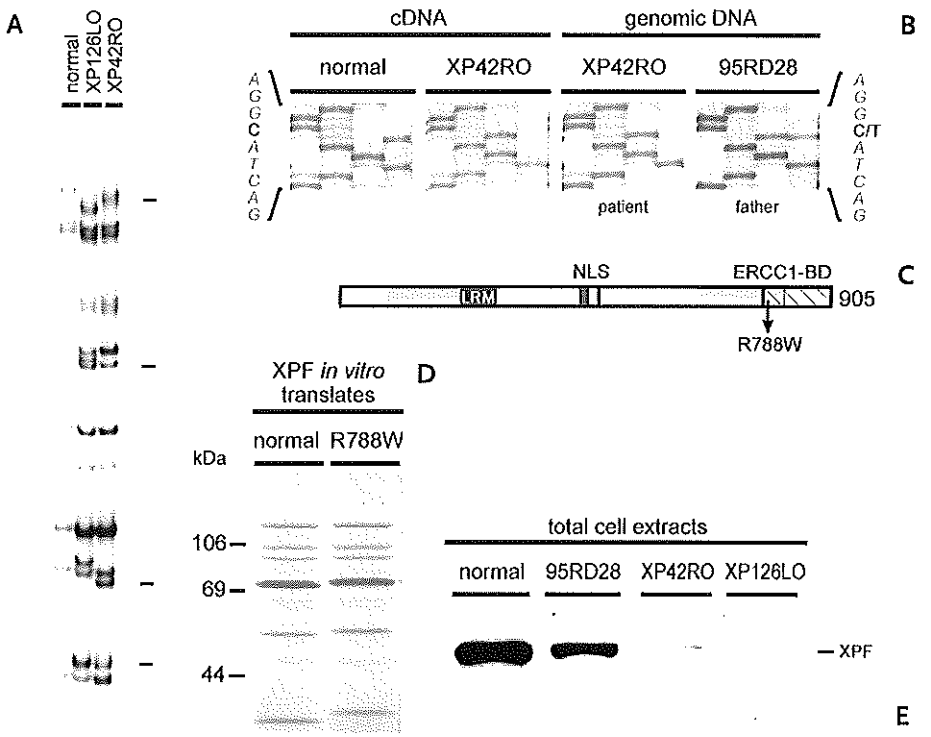


Fig. 2 Mutational analysis of XP42RO. A. REF analysis. The restriction enzyme fingerprint is shown of the 3' end of cDNA from normal, XP126LO, and XP42RO cells after digestion with *Eae*I, *Sac*I and *Stu*I. Aberrant bands are indicated. B. Sequence showing C to T transition in cDNA and genomic DNA. C. Schematic structure of XPF protein with R788W mutation. Domains shown in gray are highly conserved; LRM, leucine-rich motif; NLS, putative nuclear location signal; ERCC1-BD, region required for ERCC1 binding. D. Migration of XPF(R788W) and normal XPF *in vitro* translated proteins in a denaturing gel. Bands smaller than the full-length protein of 115 kDa represent N-terminally truncated species caused by translational starts on internal AUG sites. Aberrant slower mobility of mutated XPF is more readily visible in the shorter C-terminal products consistent with the mutation being situated in the C-terminal part of the protein. E. Immunoblot of equal amounts of cell-free protein extracts with anti-XPF antiserum (Chapter 3).

vidual (Fig. 2A). XP126LO had been shown earlier to carry a 4-nucleotide deletion in one allele coding for a truncated protein, 803 residues in length, and a C to T transition in the other changing amino acid residue 788 from arginine to tryptophan (R788W) (Chapter 3). Subsequent sequence analysis of the region in XP42RO revealed that this patient is homozygous for the R788W mutation (Fig. 2B). Investigation of genomic DNA verified the homozygous (or hemizygous) state of the patient and established the father as carrier (Fig. 2B). Material from the mother was not available for analysis.

Consequences of the R788W mutation

To determine the relevance of the observed change for *in vivo* NER, an *XPF* cDNA carrying the mutation was cloned in a mammalian expression vector with a strong promoter and micro-injected into the nuclei of XP-F strain XP126LO. This overexpressed XPF species could correct the NER defect only very partially, whereas the nonmutated cDNA produces fully normal levels of NER upon microinjection (Table 1, bottom panel) (Chapter 3). However, significant residual activity was evident, which shows that the R788W substitution is not a null-mutation. On the protein level, the mutation causes an aberrant electrophoretic mobility detected in denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis of *in vitro* translated polypeptides (Fig. 2D), indicative of a considerable conformational change. In XP42RO cell extracts, the amount of XPF protein is strongly reduced, as measured by immunoblotting (Fig. 2E). The levels were also low in XP126LO but were intermediate in the father (Fig. 2E), in agreement with his heterozygous state.

Discussion

Patients with a mutated *XPF* gene are mostly found in Japan, where it is the most common form of XP after XP-A. Outside Japan, XP-F is rare. This patient XP42RO represents only the second Caucasian case ever reported. The first case was an English patient XP126LO (Norris *et al* 1988). Characteristic of XP-F are mild cutaneous manifestations, relatively late onset of cutaneous carcinomas, and the absence of neurologic abnormalities. Clinical, photobiological, and repair hallmarks of published XP-F cases are summarized in Chapter 6 (Table 1).

Acute sun sensitivity reactions were reported in 14 of 18 patients. The subnormal MED peak at 24 hr in our patient corresponded with that in seven of 10 other XP-F patients, whereas it was low in only three. Delayed or morphologically abnormal photosensitivity reactions have been reported in the majority of XP patients, consistent with the delayed reaction at 72 hr seen in our patient.

Malignant skin tumors first appeared between 27 and 47 years of age, which is relatively late for XP patients in general, but early among the nine of 18 XP-F cases where tumors have arisen. The prevalence of only basal and squamous carcinomas, characteristic for XP-F, is also lower than in XP-A and XP-D patients. Obviously, tumor risk also depends on total UV exposure and medical care. Our patient, who had been extensively exposed to sunlight for at least twelve years, contrasts with

nine tumor-free patients of whom four were still very young (Chapter 6 Table 1). Residual repair activity is similar to the other XP-F patients. This four to six times reduction of overall NER rate seems to be in contradiction with an only moderate UV cytotoxicity and nucleic acids synthesis inhibition, a unique feature typical for the XP-F group also found in our patient. This behavior is thought to be caused by a partially impaired NER activity which proceeds longer and is less reduced at later times after exposure (Zelle *et al* 1980, Arase *et al* 1990). The rather mild cutaneous and ocular manifestations of XP-F are in line with these repair characteristics.
Whereas the UV-sensitive phenotype of our patient is rather typical of XP-F, the neurologic abnormalities that started to develop in the patient's late forties are unusual and have been reported in only one exceptional Japanese case so far (Moriwaki *et al* 1993), with a type and onset of neurologic symptoms grossly similar to our case. Five of the other patients were much younger at the time of report, but in three aged 50 to 73 there was no evidence of neurologic problems. In general, central nervous system dysfunction is not uncommon in XP. However, usually only cases with the most severe photosensitivity are affected, such as those in groups A and D (Andrews *et al* 1978), where the problems develop at an early age. Robbins *et al* (1993) reported neurologic degeneration in a mildly sensitive XP-C patient who was in her forties. However, in our case with comparable mild NER impairment the origin may be distinct and specifically related to the *XPF* defect as the course of the neurologic disease was different from that of the XP-C patient. Peripheral nervous system problems such as dysmyelination, which are not present in our case, are found in another repair disorder Cockayne syndrome and in rare cases that combine the features of XP and CS (Nance and Berry 1992). Such patients occur in XP groups B, D, and G only (Hamel *et al* 1996, Vermeulen *et al* 1997). Therefore, these abnormalities were proposed to be associated with subtle disturbances in basal transcription, a process requiring the proper functioning of the XPD and XPB proteins (Vermeulen *et al* 1994b).
The nonstandard clinical picture of our patient prompted us to perform a more detailed molecular genetic analysis. The assignment to group F by complementation analysis was confirmed on the molecular level by the identification of a homozygous mutation in the *XPF* gene. A C to T transition (probably due to deamination of a methylated cytosine at a CpG site) at nucleotide 2377 changes the basic arginine residue 788 of the encoded protein into a hydrophobic tryptophan. This mutation is physiologically relevant and most likely responsible for the patient's NER defect, because overexpression of micro-injected mutated cDNA cannot fully correct UDS in XP-F cells. Surprisingly, the same R788W mutation had been found earlier in one of the alleles of the unrelated English patient XP126LO (Chapter 3). Both this point mutation and the C-terminal truncation encoded by the second allele of XP126LO affect the region of the XPF protein that is required for stable physical interaction with another repair enzyme called ERCC1 (Chapter 5, Fig. 2C). Uncomplexed free XPF and ERCC1 proteins are unstable *in vivo* (Chapter 2). Therefore, the strongly reduced levels of XPF protein that we observed in both patients' cells (Chapter 3) and the deficiency of XP-F cells in ERCC1 protein (Biggerstaff *et al* 1993, van Vuuren *et al* 1995, Yagi *et al* 1997)

would be readily explained by this interaction failure. However, *in vitro* the R788W protein could still bind ERCC1 (our unpublished data). It follows that the apparent XPF protein instability *in vivo* is related to the conformational changes that we detected in a denaturing gel.
In the process of NER, the ERCC1–XPF protein complex catalyses strand incision at the 5' side of the DNA lesion (Chapter 3, Bessho *et al* 1997b). Because this structure-specific endonuclease activity is essential for reconstitution of NER *in vitro* (Aboussekhra *et al* 1995, Mu *et al* 1995), the presence of significant residual repair in the XP-F cells suggests that in most patients the mutations are leaky. Intrinsic residual activity in the R778W mutated protein was apparent from its capability to partially restore repair in micro-injected XP-F cells that overexpress it.
In conclusion, a homozygous R788W mutation in the *XPF* gene underlies the NER defect of patient XP42RO. The mutation causes strongly reduced but still detectable levels of XPF protein. The biochemical defect results in clinical and cellular UV-responses that are mild and typical of other XP-F cases reported so far. Whether the R788W substitution is also responsible for the unusual late-onset neurologic symptoms in the patient remains to be seen, because the other European XP-F patient with the same mutation and showing no neurologic symptoms (Norris *et al* 1988) was still young (22 years old) at the time of report. Moreover, extended case finding and mutational analysis, currently in progress, will also shed light on the possibility that R788W represents a Caucasian founder mutation.

Materials and methods

Cell strains and culture

Primary skin fibroblasts were cultured routinely in Ham's F10 medium supplemented with 15% fetal bovine serum and antibiotics. Cell strains used were XP42RO (patient), 95RD28 (patient's father), XP25RO (XP-A), XP11BE (XP-B), XP20RO, XP21RO, and XP3MA (XP-C), XP1BR, and XP2CS (XP-D), XP2RO (XP-E), XP2YO, and XP126LO (XP-F), XP3BR (XP-G), and C5RO (normal control).

DNA repair studies

The UV source was a standard germicidal mercury tube emitting UV-C rays of predominantly 254 nm. NER activity was assayed in cultured fibroblasts as UDS, by measuring UV-induced incorporation of tritiated thymidine in autoradiograms, as described elsewhere (Vermeulen *et al* 1994a). The number of grains above the nuclei represents the cellular repair activity. The short-term effect of UV on nucleic acids synthesis was similarly assayed by ^{3}H-thymidine incorporation 16 hr after irradiation (Hamel *et al* 1996). A simplified assay was used for cellular survival (Hamel *et al* 1996). In short, sparsely seeded, 2-day-old cultures were exposed to graded UV doses. Three to five days after irradiation, the number of proliferating cells was estimated by liquid scintillation counting of tritiated thymidine incorporated during a 3 hr pulse. For complementation analysis, XP cells were fused using Sendai virus and assayed for UDS after 24 to 48 hr, as described (Keijzer *et al* 1979).

General biochemical procedures

Nucleic acids purification, restriction enzyme digestions, gel electrophoresis, and immunoblot analysis were performed according to standard procedures (Sambrook *et al* 1989). Coupled *in vitro* transcription and translation reactions using the TNT reticulocyte lysate system were performed as described by the manufacturer (Promega). Labeled translation products were analyzed by standard sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Mutational analysis

The *XPF* gene was amplified in two overlapping segments by PCR from cDNA prepared from total cell RNA isolated from cultured fibroblasts using random hexamers. Products were digested with different sets of restriction enzymes and 5' end-labeled with ^{32}P-phosphate (Chapter 3). Labeled single strands were separated based on conformation in nondenaturing polyacrylamide gels containing 5% glycerol and run at 4°C, as described (Liu and Sommer 1995). Amplified genomic or cDNA fragments were directly sequenced.

Micro-injection

XPF cDNA subcloned in a mammalian expression vector pcDNA3 was microneedle-injected into one nucleus of XPF-deficient XP126LO homopolykaryons. After a 24 hr expression period, UV-UDS was measured in the injected cells (Vermeulen *et al* 1994a).

Acknowledgements

We thank S. Rademakers for the micro-injection experiments. The DNA repair studies were supported by the Louis Jeantet Award and by the CEC Concerted Action Program on DNA repair.

CHAPTER 5

Mapping of interaction domains between human repair proteins ERCC1 and XPF

Wouter L. de Laat
Anneke M. Sijbers
Hanny Odijk
Nicolaas G.J. Jaspers
Jan H.J. Hoeijmakers

DEPARTMENT OF CELL BIOLOGY AND GENETICS,
MEDICAL GENETICS CENTER,
ERASMUS UNIVERSITY ROTTERDAM, THE NETHERLANDS

Nucleic Acids Res 26: 4146-4152 (1998)

ERCC1–XPF is a heterodimeric protein complex involved in nucleotide excision repair and recombinational processes. Like its homologous complex in *S. cerevisiae*, Rad10–Rad1, it acts as a structure-specific DNA endonuclease, cleaving at duplex–single-stranded DNA junctions. In repair, ERCC1–XPF and Rad10–Rad1 make an incision on the 5' side of the lesion. No humans with a defect in the ERCC1 subunit of this protein complex have been identified and ERCC1-deficient mice suffer from severe developmental problems and signs of premature aging on top of a repair-deficient phenotype. Xeroderma pigmentosum group F patients carry mutations in the XPF subunit and generally show the clinical symptoms of mild DNA repair deficiency. All XP-F patients examined demonstrate reduced levels of XPF and ERCC1 protein, suggesting that proper complex formation is required for stability of the two proteins. To better understand the molecular and clinical consequences of mutations in the ERCC1–XPF complex, we decided to map the interaction domains between the two subunits. The XPF-binding domain comprises C-terminal residues 224 to 297 of ERCC1. Intriguingly, this domain resides outside the region of homology with its yeast Rad10 counterpart. The ERCC1-binding domain in XPF maps to C-terminal residues 814 to 905. ERCC1–XPF complex formation is established by a direct interaction between these two binding domains. A mutation from an XP-F patient that alters the ERCC1-binding domain in XPF indeed affects complex formation with ERCC1.

Introduction

Nucleotide excision repair is a cellular process that guards the integrity of the genome. It removes a wide variety of lesions from the DNA, including bulky DNA adducts and the most prominent UV-induced damages. During NER, a dual incision is made asymmetrically around a lesion to allow its release as part of a larger (24 to 32 nucleotides) DNA fragment (Huang *et al* 1992, Mu *et al* 1995, Moggs *et al* 1996). The remaining gap is filled by DNA synthesis and ligation (Aboussekhra *et al* 1995, comprehensively reviewed in Friedberg *et al* 1995). The mammalian heterodimeric protein complex ERCC1–XPF is a structure-specific endonuclease that catalyzes incision on the 5' side of the lesion during NER (Chapter 3, Matsunaga *et al* 1995, de Laat *et al* 1998a). Like XPG, which makes the 3' incision (O'Donovan *et al* 1994a), ERCC1–XPF is thought to be positioned through protein–protein interactions with other repair factors around a partially unwound DNA intermediate (Evans *et al* 1997a, 1997b, Mu *et al* 1997b). Defects in one of the XP factors (XPA-G) involved in the incision stage of NER can cause the typical UV-sensitive, cancer-prone phenotype observed with xeroderma pigmentosum patients.

Chinese hamster cell lines defective in either ERCC1 or XPF show not only sensitivity to UV-light, but are also extremely sensitive to DNA interstrand cross-linking agents, a phenomenon not observed with any other NER-deficient cell line (Busch *et al* 1989). Mutational analysis of the *ERCC1* gene *in vivo* showed that indeed most mutations affecting its NER function also disrupt its presumed function in the repair of interstrand cross-links (Chapter 2). Moreover, severe symptoms like liver and kidney abnormalities, developmental delay, reduced life span, and signs of premature senescence are typically observed with ERCC1 knock-out mice and are absent in NER-deficient XPA and XPC knock-out mice (McWhir *et al* 1993, de Vries *et al* 1995, Nakane *et al* 1995, Sands *et al* 1995, Weeda *et al* 1997a). In *S. cerevisiae*, strains defective in the homologs of ERCC1 and XPF, Rad10 and Rad1, fail to complete recombination between direct repeated DNA sequences (Fishman-Lobell and Haber 1992) and mutations in the homologous proteins of *S. pombe*, Swi10 and Rad16, can affect mating-type switching (Gutz and Schmidt 1985). An additional engagement of ERCC1–XPF and their homologs in recombinational pathways might commonly underlie these non-NER-related phenotypes.

There is limited knowledge of structural and functional domains within the ERCC1–XPF complex. Protein–protein interactions have been reported with the putative damage recognition protein XPA (Robins *et al* 1991, Li *et al* 1994, Park *et al* 1995a) and with the single-stranded DNA-binding protein RPA (Matsunaga *et al* 1996, Bessho *et al* 1997b), which may stabilize the opened DNA complex (Evans *et al* 1997b, Mu *et al* 1997b). The interaction with XPA occurs through residues 93 to 120 in the ERCC1 protein (Li *et al* 1994). Recently, we demonstrated that RPA can modulate ERCC1–XPF incision activity such that cleavage is restricted to the damaged strand (de Laat *et al* 1998b). It is as yet unknown which region in ERCC1–XPF is responsible for the interaction with RPA. At the C-terminus of the

ERCC1 protein, a region of 53 residues shows extensive homology to the C-terminus of the *E. coli* NER protein UvrC (Doolittle *et al* 1986, van Duin *et al* 1988). In UvrC, this region was found to be essential for its endonuclease activity (Lin and Sancar 1991) and deletion of this region specifically disrupted the 5' incision during NER in *E. coli* (Moolenaar *et al* 1998). Possibly, these residues are required to position the active cleavage site correctly onto the DNA (Moolenaar *et al* 1998). In agreement, the region comprises a so-called helix-hairpin-helix (HhH) motif, which has been implicated in non-sequence specific DNA binding and was found to be present in many DNA break processing enzymes (Chapter 2, Doherty *et al* 1996), including the structure-specific DNA nucleases FEN-1 and XPG (Harrington and Lieber 1994b, Lieber 1997). The 20 amino acids long HhH motif, present twice in this C-terminal part of ERCC1, is the only domain that ERCC1–XPF shares with other structure-specific nucleases. Intriguingly, the *S. cerevisiae* Rad1–Rad10 complex lacks this UvrC-like domain, including the HhH motifs, but incises DNA at exactly the same positions as ERCC1–XPF (Chapter 3). A more detailed map of functional domains within the ERCC1 and XPF protein might provide insight into the relevance of sequence motifs within this complex. Also, it would allow a more accurate interpretation of the phenotypical consequences of mutations in the encoding genes. Here we report the mapping of the interaction domains between ERCC1 and XPF and we demonstrate directly that a naturally occurring XP-F mutation affects complex formation.

Results

ERCC1 and XPF efficiently reconstitute a protein complex in vitro

In mammalian cells, endogenous ERCC1 and XPF are associated in a stable heterodimeric protein complex (Chapter 3, van Vuuren *et al* 1993, Biggerstaff *et al* 1993, Park *et al* 1995a) and complex formation *in vivo* was also observed when recombinant ERCC1 and XPF proteins were overproduced together in insect cells (Bessho *et al* 1997b, de Laat *et al* 1998a) or *E. coli* (unpublished observation). To determine whether ERCC1 and XPF associate *in vitro*, immunoprecipitations were performed on *in vitro* translated gene products. For this purpose, affinity-purified polyclonal antibodies against ERCC1 and XPF were used that had previously been shown to be able to deplete the complex from whole cell extracts (Chapter 3, van Vuuren *et al* 1993). *In vitro* translation of XPF in a reticulocyte lysate-based transcription-translation system resulted not only in the 115 kDa full-length gene product but also in a series of truncated polypeptides which could be precipitated with an anti-XPF antibody (Fig. 1, lane 4). Consistent with their molecular weight, these fragments appeared to originate from in-frame alternative start codons. When XPF protein was incubated with anti-ERCC1 antibodies alone, no significant XPF-precipitation was observed, showing minimal cross-reactivity with the anti-ERCC1 antibody (Fig. 1, lane 2). However, addition of *in vitro* translated ERCC1 (Fig. 1, lane 1) resulted in efficient XPF precipitation with anti-ERCC1 antibodies (Fig. 1, lane 3), demonstrating complex formation between ERCC1 and XPF. Similarly, precipitation of wildtype ERCC1 with anti-XPF antibodies was de-

tected only in the presence of XPF protein (Fig. 1, compare lanes 5 and 6). In all these tests, the corresponding pre-immune antiserum and an unrelated antiserum raised against XPA protein were not able to precipitate any complex. Patterns were not different from the beads-only controls (Fig. 1, lanes 2 and 5 and data not shown). We conclude that ERCC1 and XPF efficiently reconstitute a protein complex *in vitro*. As this assay detects binding of small amounts of proteins amidst a vast excess of reticulocyte lysate-derived proteins, it has to be considered a stringent binding assay, detecting high affinity interactions only. As such, it will provide a conservative estimate of the domains responsible for complex formation.

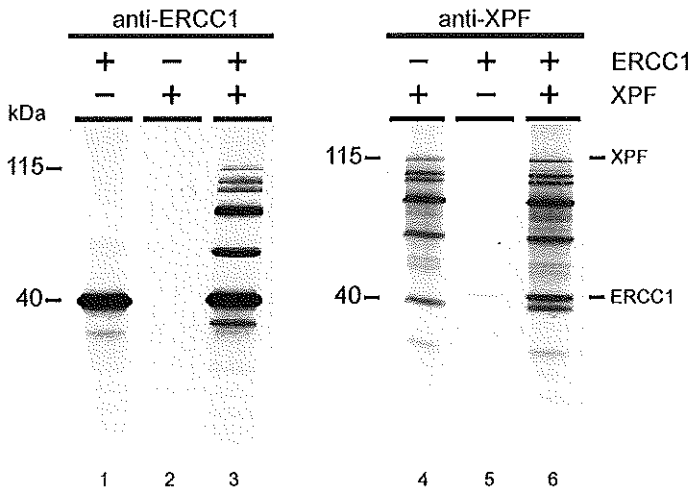


Fig. 1 *In vitro* translated ERCC1 and XPF reconstitute a complex. Autoradiogram of (antibody bead-bound) ^{35}S-labeled proteins separated on an 11% SDS-polyacrylamide gel. Input is indicated above each lane. ERCC1 and XPF *in vitro* translates directly analyzed on gel show identical products as observed in lanes 1 and 4, respectively. The molecular weights of the full-length proteins are indicated.

The XPF-binding domain is localized to the C-terminal region of ERCC1

Localized protein–protein interaction domains were previously assigned in Rad1–Rad10, the homologous counterpart of the ERCC1–XPF complex in *S. cerevisiae* (Bardwell *et al* 1992, 1993). Almost two thirds of the Rad10 protein, stretching from residue 90 to 210 at the very C-terminus, was shown to be required for Rad1 binding. This region corresponds to amino acids 98 to 214 in ERCC1, *i.e.* the middle part of this protein. In order to identify the XPF-binding domain in ERCC1, initial truncations from both sides of the ERCC1 protein were based on this putative interaction domain.

The first 92 amino acids at the N-terminus of ERCC1 were found to be dispensable for XPF binding and, in contrast to Rad10, deleting the N-terminal 103 residues

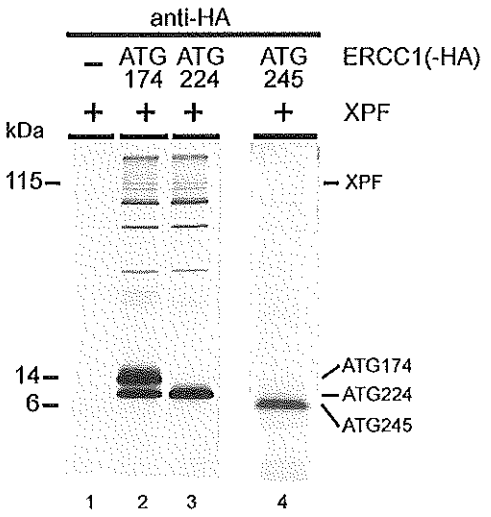


Fig. 2 N-terminal truncations of ERCC1. Autoradiogram of (antibody bead-bound) ^{35}S-labeled proteins separated on a 16.5% SDS-polyacrylamide gel. Note that only the ERCC1 protein fragments carry an HA epitope tag. The lower band synthesized along with ERCC1-ATG174 (lane 2) originates from the in-frame alternative translational start at position 224 in ERCC1 and also carries an HA epitope tag at the C-terminus. Production and visualization of the very small (6 kDa) ERCC1-ATG245 fragment was difficult. Co-precipitation of XPF with this fragment was performed in a separate experiment and is shown as a distinct panel (lane 4). The altered migration pattern of *XPF* gene products (compared with Fig. 1) is due to different percentage of gel. The molecular weights of the proteins are indicated.

(ERCC1-ATG103) and 118 residues (ERCC1-ATG118) did not seem to affect complex formation either. Even subsequent truncations from the N-terminus did not abolish the XPF binding capacity of ERCC1 and, much to our surprise, we found that an ERCC1 peptide lacking the complete region of homology to Rad10 (ERCC1-ATG224) still bound to XPF (Fig. 2, lanes 2 and 3). This interaction was observed with both an anti-XPF antibody (data not shown) and with an antibody directed against an HA epitope tag introduced at the C-terminus of the ERCC1 fragments (Fig. 2). The latter had to be used because the anti-ERCC1 antibody failed to precipitate such small C-terminal ERCC1 peptides. Deleting the N-terminal 245 residues of ERCC1, yielding a peptide with a molecular weight of only 6 kDa, abolished the affinity for XPF (Fig. 2, lane 4). We conclude that the N-terminal border of the domain responsible for initial and stable binding to XPF resides between residues 224 and 245 in the ERCC1 protein (Fig. 4).

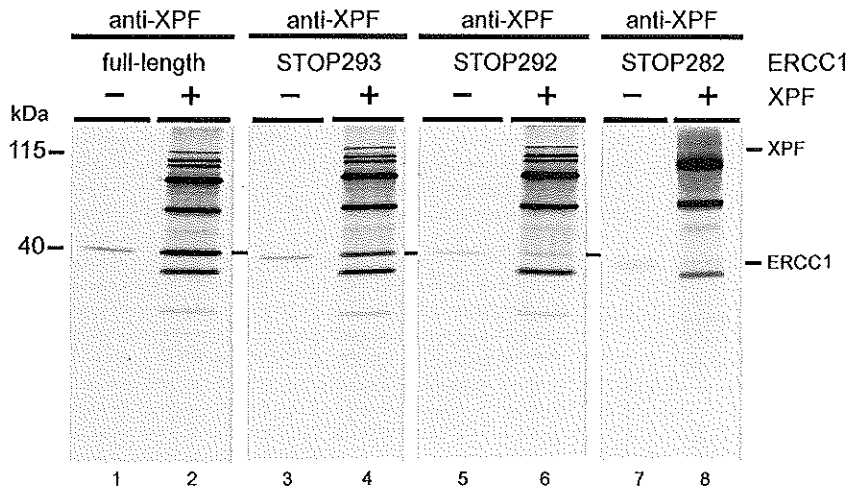


Fig. 3 C-terminal truncations of ERCC1. Autoradiogram of (antibody bead-bound) ^{35}S-labeled proteins separated on an 11% SDS-polyacrylamide gel. The odd lanes show non-specific binding of ERCC1 fragments to anti-XPF antibody beads. Interaction with XPF is scored positive if the amount of precipitated ERCC1 is significantly more in the presence (even lanes) than in the absence (odd lanes) of XPF. In each precipitation (lanes 1 to 8), similar amounts of ERCC1 products were used. Note that ERCC1-STOP293 specifically binds to XPF, but with reduced affinity compared with full-length ERCC1 (compare lanes 4 and 2). The molecular weights of the proteins are indicated.

In agreement with this conclusion, a C-terminally truncated ERCC1 protein containing only the first 215 residues (ERCC1-STOP215) was deficient in XPF binding. The same was found for the proteins ERCC1-STOP235, ERCC1-STOP257 (data not shown) and ERCC1-STOP282 (Fig. 3, lanes 7 and 8). Even ERCC1-STOP292, lacking only five amino acids from the C-terminus of full-length ERCC1, did not show XPF binding (Fig. 3, lanes 5 and 6). Interestingly, the addition of residue Phe293 to ERCC1-STOP292, yielding ERCC1-STOP293, reproducibly restored partial affinity for XPF, possibly demonstrating a direct involvement of this phenylalanine in XPF binding (Fig. 3, lanes 3 and 4). However, ERCC1-STOP293 never co-precipitated with XPF as efficiently as did full-length ERCC1, showing that even the last four residues contained sequence information required for optimal XPF binding. We conclude therefore that the C-terminal border of the XPF-binding domain in ERCC1 is located between residues 293 and 297, which is the last amino acid of full-length ERCC1 (Fig. 4).

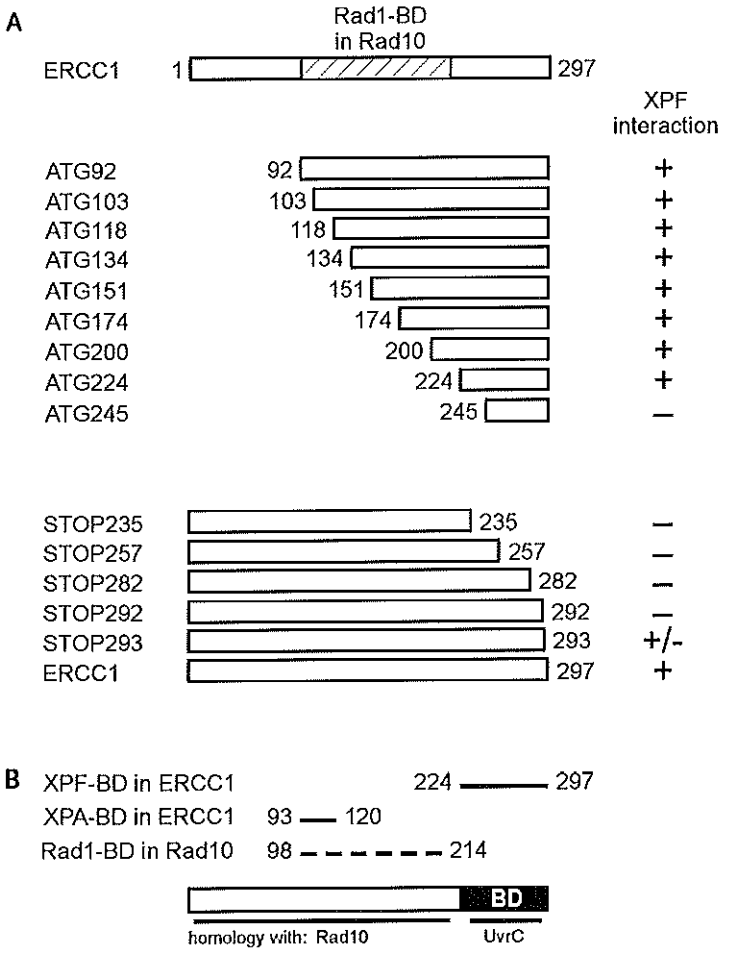


Fig. 4 Schematic presentation of the XPF-binding domain in ERCC1. A. Overview of XPF interactions obtained with truncated ERCC1 fragments. At the top is shown full-length ERCC1 with the domain that corresponds to the Rad1-binding domain (BD) in Rad10. +, interaction with XPF; +/–, intermediate interaction with XPF; –, no specific interaction with XPF. B. Mapping of XPF-binding domain in ERCC1. A summary of reported interaction domains in ERCC1. The dotted line represents Rad1-BD in Rad10. Note that Rad1 and ERCC1 do not physically interact (see discussion).

The ERCC1-binding domain is localized to the C-terminal region of XPF

Figure 1 shows that a complete series of truncated XPF fragments originating from alternative translational start sites co-precipitate with wildtype ERCC1 when an anti-ERCC1 antibody is used (Fig. 1, lane 3). This indicates that the C-terminal part of XPF is responsible for ERCC1 binding, which would be in agreement with the Rad10-binding domain in Rad1, which was mapped to a region corresponding to residues 662 to 827 in XPF (Bardwell *et al* 1993). To identify the ERCC1-interacting region in the XPF protein, initially a set of N-terminally truncated XPF cDNAs was made by systematically removing coding DNA on the 5' side of in-frame ATG codons. In this way we obtained the constructs XPF-ATG629, XPF-ATG677, XPF-ATG711 and XPF-ATG737. On polyacrylamide gels, the *in vitro* translated products of these constructs co-migrated exactly with the truncated fragments that were synthesized along with full-length XPF. Precipitation studies with full-length ERCC1 showed that not only XPF-ATG629, lacking the N-terminal 629 residues, but also the other N-terminally truncated XPF fragments were fully capable of binding ERCC1 (data not shown, but see below for further trunca-

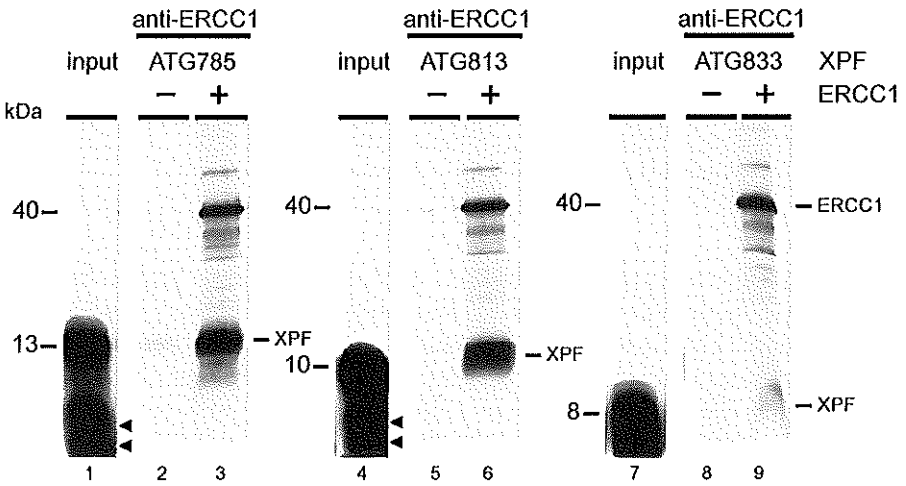


Fig. 5 N-terminal truncations of XPF. Autoradiogram of (antibody bead-bound) ^{35}S-labeled proteins separated on a 15% SDS-polyacrylamide gel. Arrowheads (lanes 1 and 4) indicate XPF products originating from the alternative starts at position 844 and 856. Note that these two truncated XPF fragments do not co-precipitate with ERCC1 (lanes 3 and 6), confirming the mapping of the N-terminal border of the ERCC1-binding domain in XPF. Lanes 2, 5, and 8 show non-specific binding of XPF fragments to anti-ERCC1 antibody beads. Input is shown in this figure to demonstrate that although some XPF-ATG833 specifically co-precipitates with ERCC1 (compare lanes 9 and 8) the ERCC1-affinity of this XPF fragment is strongly reduced (compare lanes 9 and 7 with lanes 6 and 4, and 3 and 1). The molecular weights of the proteins are indicated.

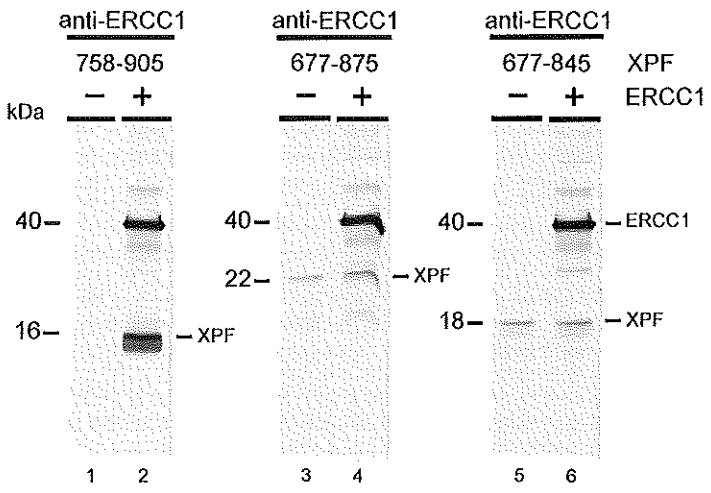


Fig. 6 C-terminal truncations of XPF. Autoradiogram of (antibody bead-bound) ^{35}S-labeled proteins separated on a 15% SDS-polyacrylamide gel. Odd lanes show non-specific binding of XPF fragments to anti-ERCC1 antibody beads. Equal amounts of XPF products were used in each precipitation (lanes 1 to 6). Note that although the XPF-677–875 (lanes 3 and 4) and XPF-677–845 (lanes 5 and 6) are larger than XPF-758–905 (lanes 1 and 2), they do not precipitate with ERCC1, due to C-terminal deletions. The molecular weights of the proteins are indicated.

tions). This demonstrated directly that, like Rad1, XPF contains a large N-terminal region that is dispensable for complex formation, but it also showed that the interaction domain in XPF is located more towards the C-terminus. By means of PCR, we further truncated the XPF protein and found that ERCC1 affinity was unaffected even when almost 90% of residues were missing from the N-terminus of the XPF protein. The peptides XPF-ATG758 (Fig. 6, lanes 1 and 2), XPF-ATG785 (Fig. 5, lanes 1 to 3), and XPF-ATG813 (Fig. 5, lanes 4 to 6) all bound strongly to the ERCC1 protein. However, ERCC1 binding capacity was severely reduced when 20 more residues were deleted from the N-terminus of XPF (XPF-ATG833) (Fig. 5, lanes 7 to 9). Thus, the N-terminal border of the ERCC1 interaction domain is located between residues 814 and 834 in XPF.

To map the C-terminal border of the ERCC1-binding domain in XPF, premature translational stops at positions 845 and 875 were introduced into the binding-proficient XPF-ATG677 construct, reducing its length by 60 and 30 amino acids, respectively. This yielded the peptides XPF-677–845 and XPF-677–875. Unlike a (smaller) peptide that contained the very C-terminal residues of XPF (XPF-758–905), neither XPF-677–845 nor XPF-677–875 showed specific binding to ERCC1 (Fig. 6), demonstrating that the C-terminal border of the ERCC1-binding domain in XPF resides between the amino acids 875 and 905 (Fig. 7).

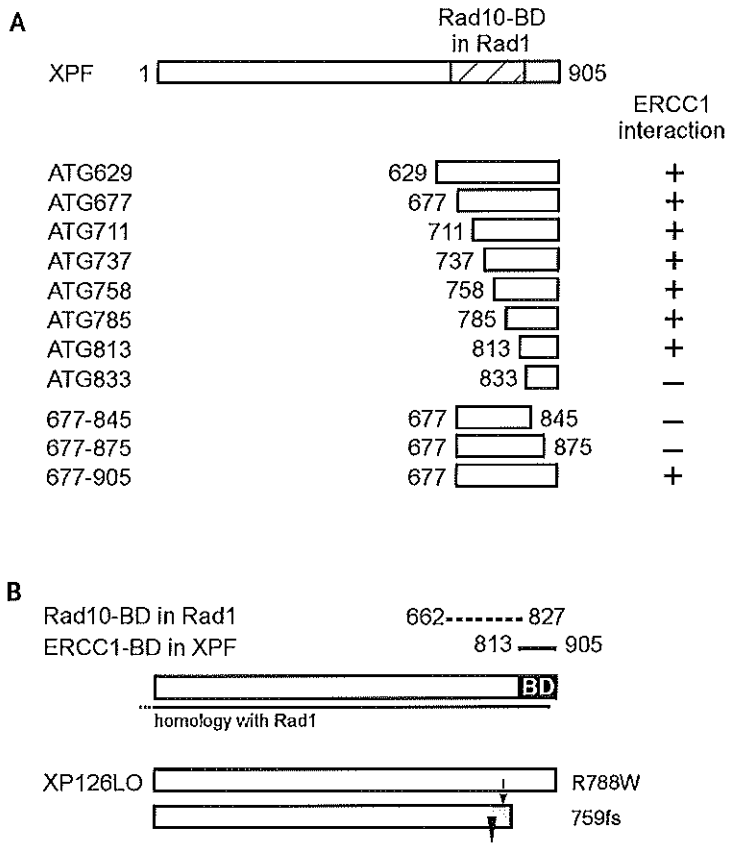


Fig. 7 Schematic presentation of the ERCC1-binding domain in XPF. A. Overview of ERCC1 interactions obtained with truncated XPF fragments. At the top is shown full-length XPF with the domain that corresponds to the Rad10-binding domain in Rad1. +, interaction with ERCC1; –, no specific interaction with ERCC1. B. Mapping of ERCC1-BD in XPF, and schematic presentation of the XP126LO alleles. The arrow indicates the position of the R788W substitution. The triangle indicates position of the frameshift resulting in a truncated protein.

Direct interaction between the binding domains of ERCC1 and XPF

The minimal domain required for initial and stable XPF binding spans residues 224 to 297 in ERCC1, whereas residues 814 to 905 in XPF were found to be necessary and sufficient for stable binding to ERCC1. To find out whether these domains physically interact, we mixed the HA-tagged peptide ERCC1-224–297 with XPF-813–905 and precipitated them with an HA-antibody. Since these two peptides migrate similarly on SDS-polyacrylamide gels, radiolabeled methionines were only incorporated into XPF-813–905. XPF-813–905 was found to co-precipitate with ERCC1-224–297 (Fig. 8, lane 2), but not with the smaller ERCC1-245–297 fragment (Fig. 8, lane 3, compare with lane 1), demonstrating that ERCC1–XPF complex formation is established by a direct interaction between residues 224 to 297 of ERCC1 and 814 to 905 of XPF.

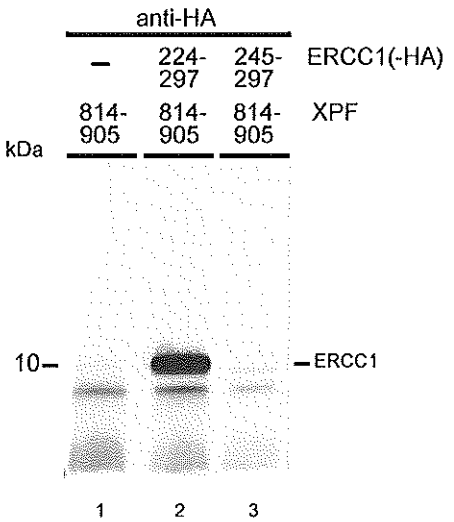


Fig. 8 Direct interaction between the binding domains of ERCC1 and XPF. Only ERCC1 fragments contained an HA epitope tag at the C-terminus. Precipitations were performed with an anti-HA antibody. No radiolabel was incorporated into ERCC1 fragments. The molecular weight of XPF-813–905 is indicated.

Naturally occurring XP-F mutations affect complex formation

Mutational analysis of an XP-F patient, XP126LO (Norris *et al* 1988), demonstrated sequence alterations in the C-terminal part of both *XPF* alleles, one being a point mutation resulting in an amino acid change at position 788 (R788W) and the other being a 4-nucleotide-deletion causing a frameshift at residue 759 (759fs) and a premature truncation at position 803. Strongly reduced protein levels of both ERCC1 and XPF were observed in cells of this patient, apparently as a consequence of these XP-F mutations. Previous observations indicated that ERCC1 and XPF molecules residing outside the complex are rapidly degraded in the cell. We therefore hypothesized that these mutations would interfere with stable complex formation (Chapter 2). To test this directly, cDNAs encoding the two XP126LO

alleles were cloned into expression vectors and *in vitro* translated. Co-precipitations with wildtype ERCC1 translates revealed that the amino acid substitution R788W did not alter the ERCC1 binding capacity of the XPF protein (Fig. 9, lane 3), whereas the product of the other allele had completely lost it (Fig. 9, lane 4). Thus, one of the alleles of this XP-F patient indeed carries a mutation that affects ERCC1 binding *in vitro*. However, under the conditions used here, no altered ERCC1 affinity was observed for the gene product carrying the point mutation at position 788.

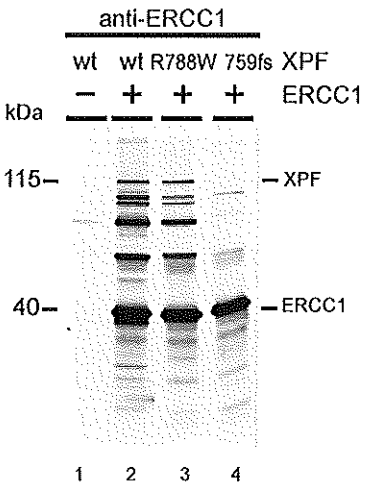


Fig. 9 A naturally occurring XP-F mutation affects complex formation with ERCC1. Autoradiogram of (antibody bead-bound) ^{35}S-labeled proteins separated on an 11% SDS-polyacrylamide gel. Similar amounts of XPF were used in each precipitation (lanes 1 to 4). The molecular weights of the proteins are indicated; wt, wildtype XPF protein.

Discussion

In this study we mapped the interaction domains between two polypeptides by performing immunoprecipitations on *in vitro* translated proteins. In such assays, small amounts of both proteins are mixed in the context of a huge excess of unrelated, reticulocyte lysate-derived, proteins. Co-precipitation, therefore, requires that proteins bind each other with relatively high affinity. A well-established interaction between ERCC1 and XPA for example, with a dissociation constant of $2.5 \cdot 10^{-7}$ M (Saijo *et al* 1996), cannot be visualized with this procedure and co-precipitation of *in vitro* translated XPA is only detected with an excess of (recombinantly overproduced) ERCC1 on antibody beads (our unpublished observation, Bessho *et al* 1997a). Thus, our binding assay is a conservative one that will only reveal domains required for initial and stable complex formation, but fails to show residue stretches of minor importance for protein–protein interac-

tions. We define residues 224 to 297 in ERCC1 and residues 814 to 905 in XPF as the regions responsible for initial and stable complex formation between ERCC1 and XPF.

Comparison between ERCC1–XPF and Rad10–Rad1

Rad1–Rad10 from *S. cerevisiae* and mammalian ERCC1–XPF not only display extensive amino acid sequence homology, but are also functionally very similar. Both complexes interact with the DNA damage-binding NER protein XPA, which is known as Rad14 in *S. cerevisiae* (Li *et al* 1994, Park and Sancar 1994, Guzder *et al* 1996). Also, both display identical incision patterns on stem–loop substrates (Chapter 3) and require similar DNA structural elements for nuclease activity (Bardwell *et al* 1994, de Laat *et al* 1998a). Interaction domains responsible for complex formation have previously been assigned to the Rad1–Rad10 complex, using a two-hybrid system and immunoprecipitation assays similar to those described in this report (Bardwell *et al* 1992, 1993). The binding domains in both Rad1 and Rad10 appeared to map to protein regions that are well conserved between the yeast and mammalian homologs. Surprisingly, however, we find that the interaction domains in ERCC1 and XPF locate elsewhere.

The location of the Rad1-binding domain in Rad10 corresponds to residues 98 to 214 in ERCC1 (Bardwell *et al* 1993), but the XPF-binding domain in ERCC1 comprises a stretch of amino acids (224 to 297) that resides outside the region of homology with Rad10 (Fig. 4B). This C-terminal extension in ERCC1 is predominantly composed of a double HhH motif (residues 236 to 289). HhH motifs have been found in many DNA break processing enzymes, including the *E. coli* NER protein UvrC, and are thought to be involved in DNA binding. Previous mutagenesis studies showed that a 'Rad10-like' ERCC1 protein, with a stop at residue 214, was functionally inactive (van Duin *et al* 1988). This can now be explained by the inability of this protein to form a complex with XPF. However, addition of a double HhH motif alone is not sufficient to restore XPF-binding capacity. A hybrid protein containing the first 236 amino acids of ERCC1 fused to the double HhH motif of bacterial UvrC, composing a peptide of 291 amino acids, failed to bind XPF (data not shown). Similarly, the ERCC1-STOP292 peptide containing the complete double HhH motif of ERCC1 did not bind XPF (Fig. 3). Apparently, the double HhH motif in ERCC1 is not directly involved in XPF binding. Our data demonstrate that residues 293 to 297 in ERCC1 are important for XPF binding. Interestingly, an as yet unobserved homology exists between the C-terminal residues 293 to 296 in ERCC1, comprising the amino acids Phe-Leu-Lys-Val, and the final four residues of Rad10 (207 to 210), which are Tyr-Leu-Asn-Leu. Although the Rad1-binding motif in Rad10 was reported to extend to residue 210 (Bardwell *et al* 1993), to our knowledge no subtle truncations from the C-terminus of Rad10 have been made. Hence, it is not clear whether these particular four residues of Rad10 are indispensable for Rad1 binding, but it is possible that the two motifs in ERCC1 and Rad10 fulfil similar roles in binding the complexing protein partner. Along this line of argument, the double HhH motif present in ERCC1 may support correct structural positioning of the very C-terminal residues of ERCC1.

The functional relevance of an additional, putative DNA-binding domain in this part of ERCC1 still has to be resolved though. It is worth mentioning that also *S. pombe* Swi10 (Rödel *et al* 1992) and plant homologs of ERCC1 (Xu *et al* 1998) also contain a UvrC-like C-terminal domain, which suggests that *S. cerevisiae* Rad10 is an exception in lacking this region.
XPF and Rad1 use partially different domains for binding their respective complexing partners as well. The Rad10-binding domain in Rad1 has homology to residues 662 to 827 in XPF, but we find that residues 814 to 905 in XPF are required for actual ERCC1 binding. Motifs with a putative function have not been found in this region (Chapter 3). The poorly conserved localization of interaction domains between *S. cerevisiae* Rad1–Rad10 and human ERCC1–XPF explains why neither *in vivo* nor *in vitro* an interspecies protein–protein interaction was observed between Rad1 and ERCC1 (Bardwell *et al* 1993).

Interpretation of ERCC1 and XP-F mutations

Recently, we reported a mutational analysis of the *ERCC1* gene in which a series of truncated ERCC1 proteins were tested for repair capacity *in vivo* (Chapter 2). The N-terminal 92 amino acids were found to be dispensable for ERCC1 to function in repair *in vivo*. Disruption of the N-terminal 103 residues of ERCC1, however, completely destroyed the repair capacity of this peptide. As in our precipitation assay this protein fragment is capable of binding XPF, affects on ERCC1–XPF complex formation seem not to be involved. Rather, a reduced affinity for XPA, whose binding domain maps to residues 93 to 120 in ERCC1, may underlie this phenotype. C-terminal ERCC1 truncations demonstrated a direct link between complex formation ability and repair capacity. A deletion of five amino acids from the C-terminus of ERCC1 (ERCC1-STOP292) completely destroyed both the protein's ability to bind to XPF (Fig. 3) and to support repair *in vivo* (Chapter 2). However, addition of only one residue to this peptide, Phe293 in ERCC1-STOP293, not only partially restored XPF binding (Fig. 3), but was also sufficient for partial restoration of the *in vivo* repair capacity of ERCC1 (Chapter 2). Apparently, Phe293 in ERCC1 is crucial for XPF affinity and complex formation is a prerequisite for ERCC1 to function in repair.
Reduced protein levels, not only of XPF but also of ERCC1, are frequently observed in XP-F patients (Yagi *et al* 1997, personal observation). The instability of these proteins is thought to be caused by the lack of a stably bound partner. A similar phenomenon has been observed with other tight protein complexes, for example with Ku70–Ku80 (Chen *et al* 1996). XP126LO is an XP-F patient with such reduced XPF and ERCC1 protein levels whose mutations have been mapped. One allele carries a frameshift at residue 759, which leads to a premature stop in front of the ERCC1-binding domain, and in agreement, we find that the encoded protein does not interact with ERCC1. The R788W point mutation in the other allele resides outside the ERCC1-binding region in XPF and we find that this mutated XPF protein binds normally to ERCC1. However, as residual XPF and ERCC1 protein levels in this patient are less than 50%, this mutation is also anticipated to affect complex formation. Apparently, additional low affinity interaction domains

not picked up in our conservative assay can still be of significant relevance for *in vivo* complex stability. Of course, the possibility that R788W makes XPF proteolytically unstable by itself cannot be dismissed here. Either way, the residual complexes observed in this patient must contain XPF molecules carrying the R788W substitution. This mutated form of XPF is responsible for the phenotypic features of the XP126LO patient, which is supported by our recent finding of a homozygous R788W defect in another XP-F patient with similar responses to UV (Chapter 4). Residual ERCC1–XPF protein amounts are likely to be required for normal development and viability, since mice carrying a homozygous null mutation in *ERCC1* show developmental delay, severe liver and kidney abnormalities, reduced life span and signs of premature senescence (McWhir *et al* 1993, Weeda *et al* 1997a). As most XP-F patients tested have reduced ERCC1 and XPF protein levels (Yagi *et al* 1997, unpublished observations), we expect that partially disrupted ERCC1 affinity frequently underlies the XP-F phenotype.

Materials and methods

Construction of mutant cDNAs

N-terminal truncations were made via PCR using sense primers containing (5' to 3') a T7 polymerase recognition sequence, an optimal translational initiation sequence and 18 to 24 nucleotides complementary to the sequence of insertion. Exceptions were the constructs XPF-ATG629, XPF-ATG677, XPF-ATG711, and XPF-ATG737. Here, the restriction sites *Hind*III, *Rsa*I, *Fok*I and *Eag*I were used, respectively, to remove upstream coding cDNA in front of these in-frame ATG codons. HA epitope tags were introduced as described previously (de Laat *et al* 1998a). C-terminal truncations were made via PCR using anti-sense primers containing (3' to 5') 18 to 24 complementary nucleotides, a translational stop sequence and 6 to 9 random nucleotides. cDNA from the two alleles of XP126LO was amplified, subcloned and sequenced as described (Chapter 3).

In vitro *translations and immunoprecipitations*

ERCC1 and XPF constructs were transcribed and translated separately *in vitro*, following the instruction manual of the TnT coupled transcription-translation system of Promega. PCR products were phenol extracted and ethanol precipitated. 2 to 10 µg DNA were added per 50 µl *in vitro* transcription-translation reaction mix. The polyclonal antibodies against ERCC1 and XPF (affinity-purified) used for immunoprecipitations have been described before (Chapter 3, van Vuuren *et al* 1993). For immunoprecipitations, (truncated) ERCC1 protein and (truncated) XPF protein were mixed with 5 µl NETT buffer (100 nM NaCl, 5 mM EDTA, 50 mM Tris pH 7.5, and 0.5% Triton-X100) (total volume 15 to 20 µl) and incubated for 30 min at 30°C. Antibody was added and after 2 hr at 4°C, 100 µl NETT buffer containing 10% protein A-sepharose beads were added. Incubation proceeded for another 2 hr at 4°C (with tumbling), then the beads were washed four times with 0.5 ml NETT and suspended in sample buffer. Samples were boiled and protein fragments were separated on SDS-polyacrylamide gels, which varied between 5 to 16.5% acrylamide, depending on the size of the peptides studied. In the case of 15 and 16.5% gels, 0.1 M NaAc was added to the lower running buffer for optimal separation of small protein fragments and the

dye *m*-cresol purple was added to the sample buffer instead of coomassie (Christy *et al* 1989). Dried gels were analyzed by autoradiography or with a phosphorimager.

Acknowledgements

This work was financially supported by the Dutch Scientific Organization (Foundation for Chemical Sciences of NWO).

CHAPTER 6

General discussion

Considerable progress has been made in recent years in understanding the roles of the XP and ERCC factors in NER and the molecular basis of xeroderma pigmentosum. Insights into the NER pathway began through complementation studies. Yeast as well as rodent repair mutants and cells from XP patients have been of great importance for the analysis of NER. To date, most of the genes complementary to these mutant groups have been cloned and their role is being elucidated. The cloning of the *XPF* gene, presented in this thesis, completed the isolation of the core NER factors. The *XPF* gene product was found to constitute an endonuclease together with ERCC1. The XPF and ERCC1 proteins and their role in DNA repair and xeroderma pigmentosum are further discussed in this chapter.

ERCC1–XPF complex

Composition

The *ERCC1* gene is the first human repair gene cloned. The gene was isolated after transfection and correction of an UV-sensitive rodent group 1 mutant (Westerveld *et al* 1984, van Duin *et al* 1986). It encodes a protein with homology to the yeast repair protein Rad10. *In vitro* complementation studies with mutant cell extracts had demonstrated that ERCC1 resides in a complex together with the ERCC4, ERCC11, and XPF-correcting activities (Biggerstaff *et al* 1993, Reardon *et al* 1993b, van Vuuren *et al* 1993). Information on the homologous complex in *S. cerevisiae* has been of great help for the characterization of the ERCC1-containing complex. The yeast complex consists of two proteins: the ERCC1 homolog Rad10, and Rad1 (Bailly *et al* 1992, Bardwell *et al* 1992). Homologs of Rad1 had already been identified in *S. pombe* and the fruitfly, Rad16 and Mei-9, respectively (Carr *et al* 1994, Sekelsky *et al* 1995, Araj and Smith 1996). On the basis of the homology in their protein sequences, the human homolog was isolated. Interestingly, the gene could restore the repair defect not only in XP-F cells but also in rodent ERCC4- and ERCC11-deficient cells, suggesting that only one and the same gene is involved in these different mutant cells. About the same time, the isolation of the ERCC1-containing protein complex using a histidine-tagged ERCC1 protein was accomplished. Analogous to the complex in yeast, the mammalian complex appeared to consist of only two proteins, ERCC1 and XPF, which supports the idea, that the encoding genes *XPF*, *ERCC4*, and *ERCC11* are equivalent. Complementation studies, however, had previously assigned the mutants to distinct groups, 4 and 11 respectively (Hata *et al* 1991, Riboni *et al* 1992). Apparently, the mutated *ERCC4* and *ERCC11* genes, or better: the mutated *XPF* genes of Chinese hamster groups 4 and 11, can compensate each others defect, a phenomenon known as intragenic complementation. Interestingly, there is a parallel in *S. pombe*, in which the *rad16.20* allele, encoding the N-terminal half of the protein, was complemented by a construct encoding the C-terminal half of the *Rad16* gene product (Carr *et al* 1994).

However, no mutations have been found in the open reading frame of the *XPF* gene of UVS1, the Chinese hamster group 11 mutant (Hayashi and Yasui personal communication). Furthermore, fusion of group 4 mutants with the group 11 mutant seems to restore the expression of the XPF transcript. These results would suggest that UVS1 represents a distinct gene that may encode a transcriptional regulator (Hayashi and Yasui personal communication). This might also explain the fact that overexpression of the *XPF* cDNA could correct the UV-sensitivity of UVS1, whereas transfection of the genomic sequence could not (Brookman *et al* 1996). Overexpression of the complex partner ERCC1 could not restore the ERCC11 mutant phenotype, excluding the possibility that addition of (ERCC1–)XPF stimulates an alternative repair pathway.

On the other hand, it is possible that ERCC11 is needed in an auxiliary way, perhaps to stabilize the complex. In that case, an excess of XPF might overcome a lack of (functional) ERCC11 supporting complex formation. A similar phenomenon

has been observed with the double-strand break repair-deficient *rad55* and *rad57* mutants. Their X-ray sensitivity could be suppressed by the overexpression of Rad51 and/or Rad52 (Hays *et al* 1995). The exact mechanism remains to be elucidated.

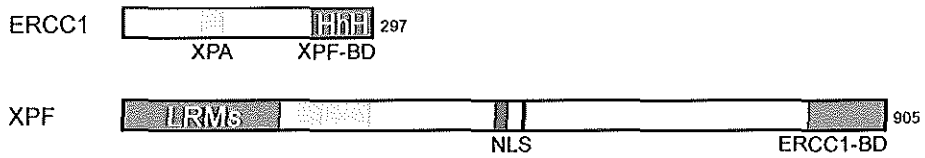


Fig. 1 The ERCC1–XPF complex. Schematic representation of functional domains within the ERCC1 and XPF proteins. HhH, helix-hairpin-helix motif; LRM, leucine-rich repeat; LZ, leucine zipper; NLS, nuclear localization signal; BD, binding domain.

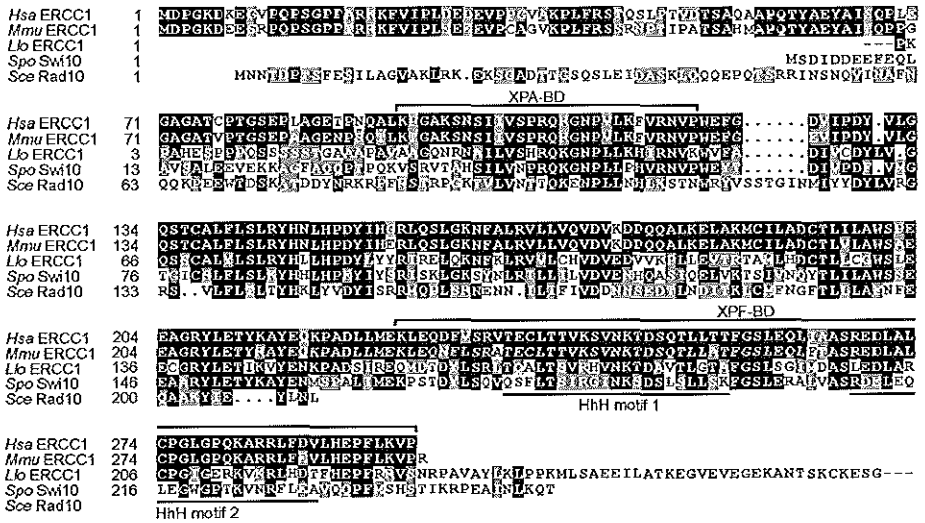


Fig. 2 Alignment of human ERCC1 with its homologs. Displayed are human ERCC1 (297 amino acids), its mouse homolog (298 amino acids), the *Lilium longiflorum* homolog (fragment of 278 amino acids) residues 1-272, *S. pombe* Swi10 (252 amino acids) and *S. cerevisiae* Rad10 (210 amino acids). Identical amino acids are in black boxes and physicochemically related residues are in gray boxes. Indicated are the XPA- and XPF-binding domain (BD) including the two HhH motifs (numbers according to van Duin *et al* 1986, 1988, Xu *et al* 1998, Rödel *et al* 1992, and Reynolds *et al* 1985).

Functional domains

ERCC1 and XPF form a tight complex with structure-specific endonuclease activity. The ERCC1–XPF endonuclease performs the 5' incision during NER. It specifically recognizes branched DNA structures with 3' single-stranded DNA ends. ERCC1–XPF cuts near the transition of double-stranded to single-stranded DNA (Chapter 3). The nuclease is also required for removal of nonhomologous DNA from 3' single-stranded ends of recombining DNA, which structure resembles the damage-containing DNA substrate cut by ERCC1–XPF during NER (Sargent *et al* 1997). A close examination of the protein sequences and the study of mutated proteins revealed several domains important for their function (Fig. 1).

In ERCC1, the N-terminus is the least conserved part of the protein (Fig. 2). Accordingly, up to one third of the protein can be deleted without loss of function. The region contains a putative nuclear location signal sequence, which is apparently not essential for nuclear transport of the protein. Instead, the putative bipartite NLS in XPF (Dingwall and Laskey 1991), located in the middle of the protein, may be responsible for transport of the complex to the nucleus. The central region in ERCC1 is the best conserved (van Duin *et al* 1986). It contains the XPA-binding site (Li *et al* 1994, Park and Sancar 1994, Li *et al* 1995b, Saijo *et al* 1996). In *S. cerevisiae*, the Rad1 and Rad10 proteins together interact with the XPA homolog, Rad14, much stronger than each protein alone (Guzder *et al* 1996). Accordingly, the presence of XPF may enhance the affinity of ERCC1 for XPA. The complex also interacts with RPA and with DNA, interactions not found with the ERCC1 subunit alone (Saijo *et al* 1996, de Laat *et al* 1998b). So far, no other interacting proteins have been reported, neither for ERCC1 nor for XPF.

The N-terminal half of XPF contains a large number of leucine residues that may form two structural motifs involved in protein binding (Fig. 3). At the very beginning, the sequence starts with a repeat of leucine-rich motifs (or leucine-rich repeats). These LRMs are short sequences of 22 to 30 residues with hydrophobic amino acids at conserved positions. Leucine-rich motifs have a degenerate consensus sequence as they contain gaps, and vary in length and in amino acid composition. LRMs are found in a large variety of proteins and are implicated in tight and highly specific protein–protein interactions, such as receptor–ligand binding (Schneider *et al* 1988, Windisch *et al* 1995a). The three-dimensional structure of the ribonuclease inhibitor protein (RI) revealed the structural properties of LRMs responsible for the binding function. The buried leucines in LRMs play a structural role and are not directly involved in protein–protein interactions. The sequence of RI consists of 15 LRMs. All repeats adopt very similar structures, consisting of a short β-strand and an α-helix approximately parallel to each other, resulting in a flexible, arch-shaped binding surface suitable for strong and distinctive interactions with ligand proteins (Fig. 4) (Kobe and Deisenhofer 1993). One of the three motifs in the extracellular domain of the neurotrophin receptor TrkA was found to specifically bind to nerve growth factor, indicating that LRMs can represent independent functional entities (Windisch *et al* 1995b). LRM structures are also present in the yeast homologs Rad1 and Rad16 and resemble the motifs found in yet another yeast repair protein Rad7 (Schneider and Schweiger 1991). The protein(s) that can bind to the LRMs in XPF

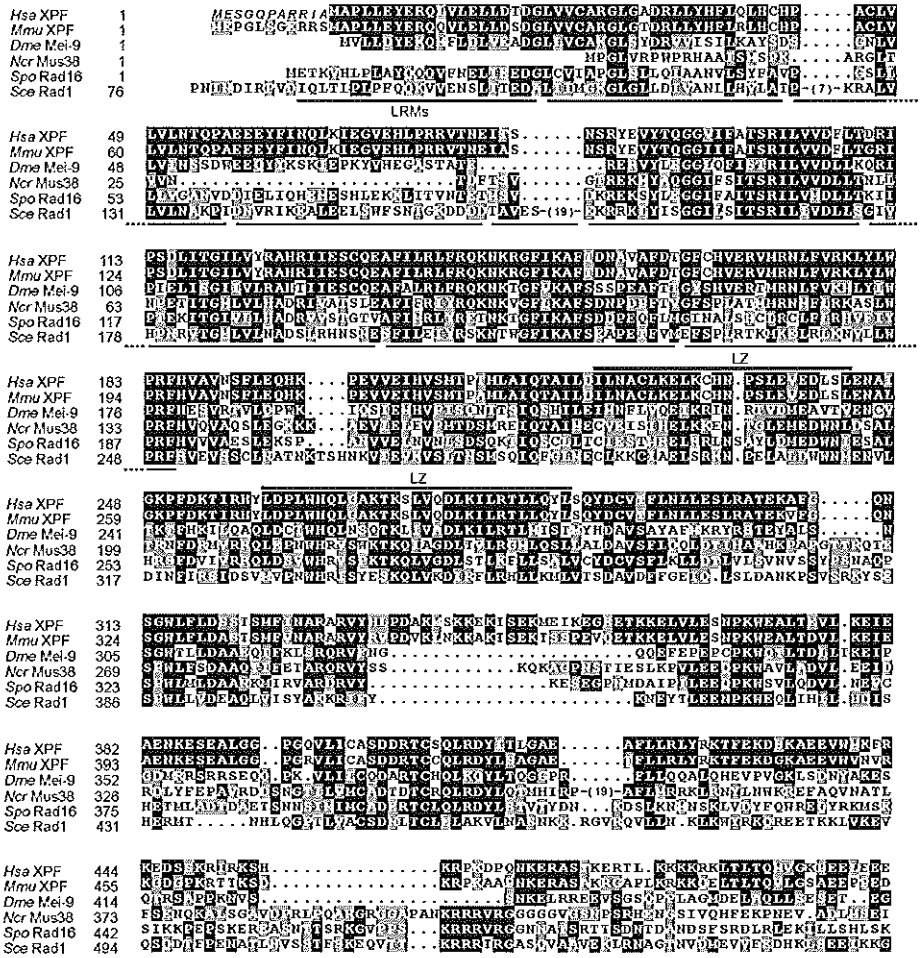
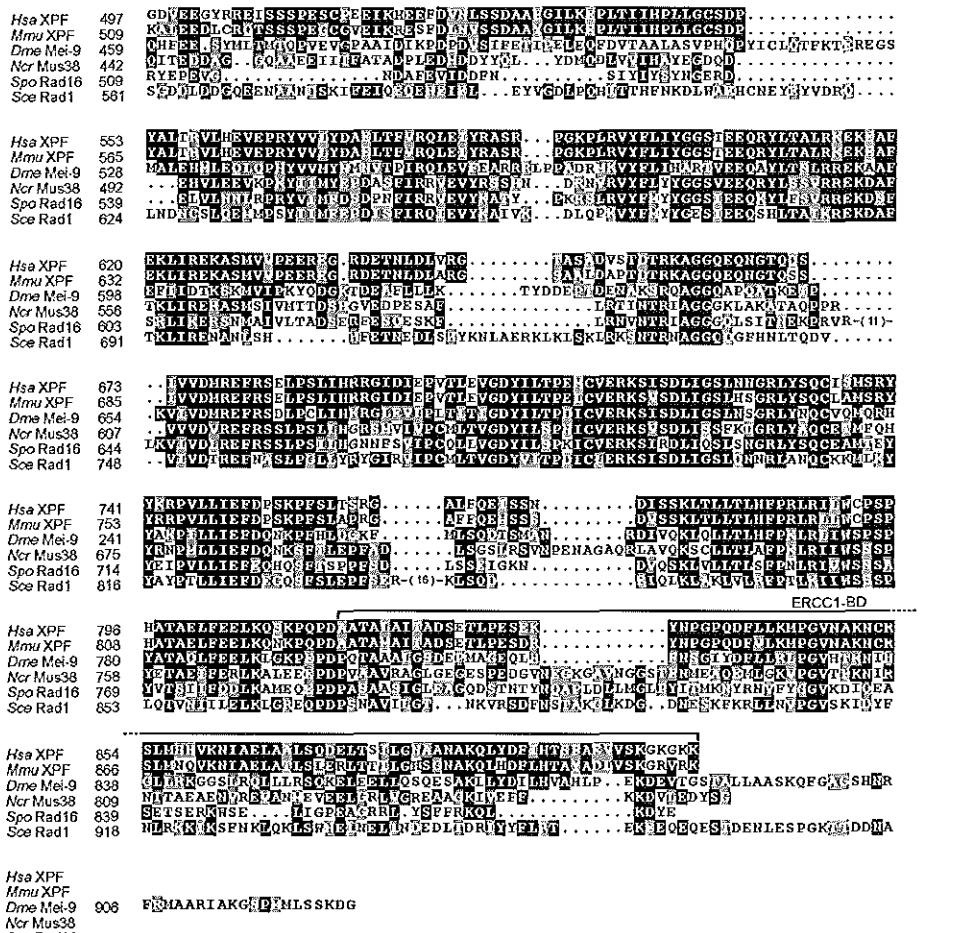


Fig. 3 Alignment of XPF with its homologs. Displayed are human XPF (905 amino acids), its *Mus musculus* homolog (917 amino acids), *D. melanogaster* Mei-9 (946 amino acids), *N. crassa* Mus38 (892 amino acids), *S. pombe* Rad16 (892 amino acids) and *S. cerevisiae* Rad1 (1100 amino acids) residues 76-1100. Occasionally, stretches of residues present in only one of the peptides were deleted for optimal alignment. Identical amino acids are in black boxes and physicochemically related residues are in gray boxes. Indicated are the ERCC1-binding domain (BD) and leucine zippers (LZ) in XPF as well as the leucine-rich motifs (LRMs) in Rad1. Additional residues (italic) upstream of human XPF were reported by Brookman *et al* (1996) (numbers according to Chapter 3, Lamerdin unpublished results, Sekelsky *et al* 1995, Hatakeyama *et al* 1998, Carr *et al* 1994, and Reynolds *et al* 1987).



remain to be identified. Further downstream, the N-terminus contains several leucine-zipper motifs (Brookman *et al* 1996). The only similarity between LRMs and leucine zippers is the conservation of leucines at particular spacings. Leucine zippers are built of heptad repeats with leucines present at every seventh position. These leucines form a hydrophobic ridge on one side of a helix that interacts with a complementary helix. The leucine residues in leucine zippers participate in dimerization (Landschulz *et al* 1988). Dimerization may occur within the XPF polypeptide. Alternatively, an ERCC1–XPF dimer–dimer could be formed by association of leucine zipper motifs of two separate XPF peptides. It is not known, however, whether ERCC1–XPF is active as a monomeric complex or acts in multimeric form. Separation of a cellular protein extract on a size fractionation column revealed that most ERCC1–XPF eluted in the fractions with a molecular weight corresponding with the

size of a monomeric ERCC1–XPF complex, whereas a small amount of ERCC1–XPF eluted in fractions corresponding to a much higher molecular weight (van Gool *et al* 1997), indicating that ERCC1–XPF can be part of a larger complex.
Complex formation involves the C-terminal domains in both proteins (Chapter 5). Also in yeast, the domains in Rad1 and Rad10 required for their mutual interaction, are located in their C-terminal regions (Bardwell *et al* 1992, Bardwell *et al* 1993). The ERCC1-binding domain in XPF, however, resides, different from the Rad10-interaction domain in Rad1, in the very C-terminal end (see Chapter 5 for further details). Similarly, the XPF-binding domain in ERCC1 is located in the very C-terminal end, a region that is even lacking in the yeast Rad10 protein. The ERCC1 C-terminus shows some homology with the C-terminal part of the *E. coli* NER endonuclease UvrC (van Duin *et al* 1988). The homologous region contains two putative helix-hairpin-helix motifs. HhH motifs are found in a large variety of DNA break-generating and processing enzymes (Chapter 2, Doherty *et al* 1996). These motifs are implicated in non-sequence-specific DNA binding and it is likely that the HhH-containing proteins recognize a specific structure of the DNA helix rather than a specific nucleotide sequence. Six crystal structures of predicted HhH motifs are currently known: a series of two in rat and in human DNA polymerase β (Sawaya *et al* 1994, Pelletier *et al* 1996) and one each in Endonuclease III (Thayer *et al* 1995), AlkA (Labahn *et al* 1996), and in the 5'nuclease domain of *Taq* polymerase I (Kim *et al* 1995). Superposition of polymerase β, Endo III and AlkA HhH three dimensional structures reveals remarkable similarity in the type II β-turn that bridges the helices and their crossing angle. The HhH motifs bind DNA via the formation of hydrogen bonds between protein backbone nitrogens and DNA phosphate groups. The 20 residue-long motif contains several positions with high propensities for hydrophobic residues at positions 3, 6, 9, 14, 17, and 18 (Fig. 5). Residues at positions 3, 17, and

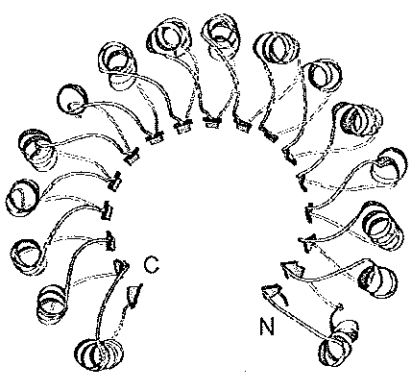


Fig. 4 Structure of ribonuclease inhibitor. Ribbon diagram of the RI structure (adapted from Kobe and Deisenhofer 1995).

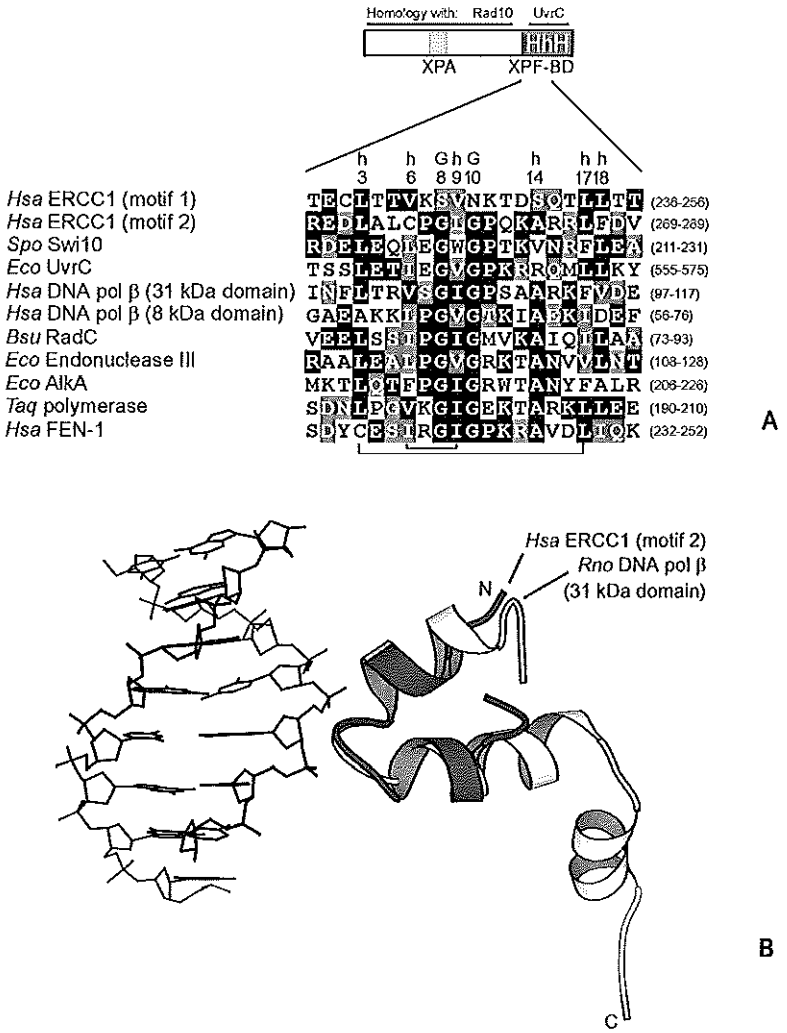


Fig. 5 The helix-hairpin-helix (HhH) motif. A. Alignment of the two HhH motifs from ERCC1 with the motifs from other HhH-containing proteins. Identical amino acids are in black boxes and physicochemically related residues are in gray boxes. B. Shown are the second HhH motif in human ERCC1 (dark gray) on top of the second HhH motif in rat DNA polymerase β (31 kDa domain) (light gray) together with the DNA molecule. The three dimensional structure of the ERCC1 HhH was modeled with help of the HhH structure in rat DNA polymerase β using the SwissModel program. *Hsa, H. sapiens; Spo, S. pombe; Eco, E. coli; Bsu, B. subtilis; Taq, T. aquaticus; Rno, Rattus norvegicus;* h, hydrophobic residue (numbers according to van Duin *et al* 1986, Rödel *et al* 1992, Sancar *et al* 1984, Chyan *et al* 1994, Levin *et al* 1992, Asahara *et al* 1989, Itoh *et al* 1996, Lawyer *et al* 1989, Murray *et al* 1994, and Kumar *et al* 1990b).

18 stabilize helical packing, whereas residues at positions 6 and 9 stabilize the β-turn and the hairpin. Glycines at positions 8 and 10 form important elements of the hairpin loop. Glycine HhH position 8 is important for the formation of the type II β-turn, whereas glycine HhH position 10 contributes to a pronounced extended surface that mediates DNA-binding. A high propensity for lysine at HhH position 12 and threonine or serine at HhH13 suggests that these interact with DNA phosphate groups. An alanine or small hydrophobic residue is preferred at HhH position 14.
The HhH motif in the 8 kDa domain of polymerase β is implicated in the short-gap filling activity of the enzyme. This domain has been shown to bind single-stranded DNA (Kumar *et al* 1990b). The second HhH motif occurs in the 31 kDa domain that binds only to double-stranded DNA (Kumar *et al* 1990a, Pelletier *et al* 1996). The putative HhH motif in endonuclease III is associated with thymine glycol binding (Kuo *et al* 1992), whereas uracil-DNA glycosylase has a DNA-binding groove and an adjacent pocket with HhH-configuration that tightly fits a deoxyuridine residue flipped-out from the damaged DNA (Roberts 1995). The HhH motif in *Taq* polymerase I protrudes into a major cleft in the 5'nuclease domain adjacent to metal binding sites. In this way, DNA bound to the motif is presented to the nuclease active site.
Like ERCC1, most HhH proteins contain (multiples of) two copies of the motif. The presence of two motifs probably enables the protein to bind DNA twice; to structures that contain transitions from double-stranded to single-stranded DNA, such as loops, hairpins, flipped out bases, and cruciforms, generated during processes like DNA repair and recombination, as illustrated by the *B. subtilis* RecR protein. RecR contains two HhH motifs. The protein is involved in postreplication repair and homologous recombination. It binds ssDNA and dsDNA. In fact, it is a sequence-independent DNA-binding protein that shows some preference for damaged DNA (Alonso *et al* 1993). Recently, it was found that RecR can bind two DNA molecules at the same time with preference for the cross over of two helices (Ayora *et al* 1997). Because of these HhH motifs, ERCC1 is likely to be responsible for DNA binding.
Like other structure-specific nucleases ERCC1 may also harbor the nuclease domain. In comparison with FEN-1 and XPG, however, the positions of the possible active site residues (conserved acidic amino acids D and E) relative to the HhH motif seem not to be conserved in ERCC1 (Mueser *et al* 1996, Shen *et al* 1997). Intriguingly, the yeast Rad1–Rad10 complex appears to lack the HhH-containing UvrC-like domain, but incises DNA at the same positions as ERCC1–XPF (Chapter 3). Insight into the function of the HhH motifs in ERCC1 will have to come from structural and detailed mutation analysis.

Triple function

Nucleotide excision repair

Mutations in the ERCC1 and XPF proteins result in UV-sensitivity, which reflects their role in NER. However, ERCC1 and XPF are involved in the repair of a wider

range of DNA damage. The rodent complementation groups 1 and 4 are also extremely sensitive to DNA cross-linking agents while the others are not (Hoy *et al* 1985). Intrastrand cross-links are easily removed via NER. The damaged nucleotides are removed and replaced with normal ones, using the complementary undamaged strand as a template. However, in the case of interstrand cross-links, both strands are damaged. The repair of these cross-links probably requires an additional step that involves homologous recombination.

Cross-link repair and homologous recombination

In *E. coli*, the removal of interstrand cross-links is thought to involve first dual incision in one strand around the cross-linked base. In this way, the NER excinuclease UvrABC produces an oligonucleotide cross-linked to an intact DNA strand (van Houten *et al* 1986). Following excision, a 5'exonuclease activity (DNA polymerase I) probably generates a single-stranded DNA region to allow strand invasion of a homologous sequence (Sladek *et al* 1989) (Fig. 6). Subsequent new DNA synthesis using the homologous sequence as a template and ligation result in a triple-stranded DNA structure. Finally, the cross-link is eliminated in a second excision reaction (Cheng *et al* 1988, Sladek *et al* 1989, Cheng *et al* 1991). In *E. coli*, interstrand cross-links are completely repaired by the NER proteins with help of a 5'exonuclease and RecA. A different model has been proposed for the repair of interstrand cross-links in mammalian cells (Bessho *et al* 1997a). With a substrate containing a defined interstrand cross-link, the initial incisions *in vitro* were found in one of the two strands, both on the same side of the cross-linked base. Remarkably, the oligonucleotide excised, does not contain one of the cross-linked bases. In the model, at least two more rounds of excision will be required to complete the repair reaction. Moreover, the excision reaction seems to depend on NER proteins other than ERCC1 and XPF, which is not expected from the NER mutants' phenotypes. The role of ERCC1–XPF in cross-link repair, which seems to occur outside the context of NER, remains to be determined. Their function may resemble the role of ERCC1–XPF during homologous recombination, which is better understood.

In yeast, two distinct homology-dependent pathways are initiated by a DSB: homologous recombination and single-strand annealing. In both pathways, the ends of the DSB are processed by a 5'→3'exonuclease to expose long 3' single-stranded tails. During homologous recombination, the Rad51, assisted by RPA and Rad52 (all single-stranded DNA-binding proteins), mediates the search for homologous sequences and annealing of these single-stranded DNA ends to a homologous duplex (Ogawa *et al* 1993). During SSA, annealing occurs within repeats and these intermediates may have protruding nonhomologous 3' ends. Such DNA structures are substrates for the NER endonuclease Rad1–Rad10 (Fishman-Lobell and Haber 1992, Bardwell *et al* 1994, Davies *et al* 1995, Ivanov and Haber 1995). Accordingly, the Rad1–Rad10 proteins, in contrast to other NER proteins, are involved in mitotic recombination between direct repeats (Klein 1988, Schiestl and Prakash 1988, Schiestl and Prakash 1990, Saffran *et al* 1994, Ivanov and Haber 1995). Likewise, ERCC1–XPF seems to be involved in SSA between direct repeats in human cells

only if nonhomologous ends are present (Yao *et al* 1997).
The analogous complex in *S. pombe*, Swi10–Rad16, is required for the resolution of DNA intermediates during mating-type switching. Mating-type switching is a process related to homologous recombination. Switching is initiated by a site-specifically induced double strand break. Subsequent strand exchange generates a heteroduplex DNA molecule, but does not produce a Holliday junction (Egel *et al* 1984, Gutz and Schmidt 1985, Formosa and Alberts 1986, Nassif *et al* 1994). *S. pombe Swi10* and *Rad16* (allelic to *Swi9*) have been assigned to the class II mating-type switching genes (Egel 1989, Fleck *et al* 1992). In class II switching mutants the reduced frequency of mating-type switching is caused by aberrant resolution of recombination intermediates. Interestingly, the human *ERCC1* gene was found to correct the switching defect of a *swi10* mutant (Rödel *et al* 1997), which points to a role of ERCC1–XPF in homologous recombination. Remarkably, also *swi4*, a mismatch repair protein MutS homolog, belongs to the class II mutants.
In *S. cerevisiae*, the removal of nonhomologous ends in recombination intermediates not only depends on Rad1–Rad10, but also on the mismatch repair proteins Msh2 and Msh3 (Saparbaev *et al* 1996, Sugawara *et al* 1997). Mutations in Msh2 reduce spontaneous mitotic recombination with certain direct repeat substrates to the same extent as mutations in Rad1. A similar reduction is seen with the double mutations *msh2rad1* and *msh3rad1*, suggesting that Rad1, Msh2 and Msh3 function in a single pathway (Saparbaev *et al* 1996). Only Msh2 and Msh3, and no other mismatch repair proteins appear to be involved. The requirement for Msh2 and Msh3 differs with the different DNA structures formed during homologous recombination and SSA. In SSA, Msh2 and Msh3 are especially required for structures with short homologous regions. When the annealing sequences are long, Msh2 and Msh3 have only little effect, while Rad1 and Rad10 are still required. In contrast, Msh2 and Msh3 are as necessary as Rad1 during homologous recombination, independent of the length of flanking homologous sequences. Msh2 and Msh3 may recognize and stabilize the branched recombination structure. In addition or alternatively, the heterodimer may load Rad1–Rad10 onto the recombination site. Either way, Msh2 and Msh3 may act to facilitate Rad1–Rad10-dependent removal of nonhomologous ends from double-strand breaks. The proteins have no subsequent role in mitotic DSB-induced recombination (Sugawara *et al* 1997), although Msh2 has been shown to bind Holliday junctions (Alani *et al* 1997).
Interestingly, *S. cerevisiae exo1* mutants are defective in recombination between direct repeats (Fiorentini *et al* 1997). Exo1, which interacts with Msh2, is a double-stranded DNA-specific 5'→3'exonuclease active in the Msh2 pathway of mismatch repair (Tishkoff *et al* 1997). It probably degrades the newly synthesized strand from the initial nick towards the mismatch during mismatch repair. Exo1 may have a similar function in mitotic recombination in concerted action with the mismatch repair proteins and the Rad1–Rad10 complex.
It is likely that ERCC1–XPF resides in a large complex, different from the 'repairosome' involved in NER, with other proteins such as Msh2 and Msh3 and the human homolog of Exo1. An intriguing speculation would be that the mismatch repair factors Msh2 and Msh3 might be suitable partners for dimerization

with ERCC1–XPF. They were high scoring proteins in the protein sequence property search PropSearch (Hobohm and Sander 1995) with XPF, indicating that they resemble the XPF protein. Moreover, Msh2 and Msh3 play a role in mismatch recognition as heterodimers. By way of dimer formation, the availability of ERCC1–XPF in the NER, SSA, and homologous recombination pathways could be regulated. As discussed below, there are indications that the NER and mismatch repair pathways interact with one another as well.

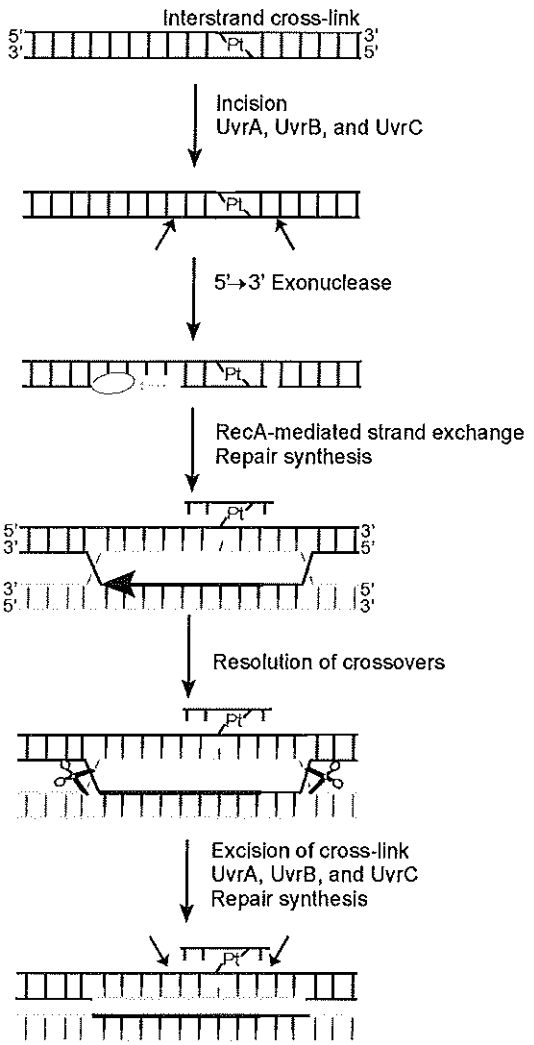


Fig. 6 Model for interstrand cross-link repair in *E. coli*.

Mismatch repair

The Msh2 factor is involved in the repair of mispaired bases and small loops that arise by replication slippage and during recombination between related but diverged sequences. The various Msh-dimers have distinct but partially overlapping substrate specificity. While, the yeast Msh2–Msh3 heterodimer primarily recognizes small loops, the Msh2–Msh6 dimer prefers single basepair mismatches (Alani 1996, Johnson *et al* 1996, Marsischky *et al* 1996). Mismatches also occur in heteroduplexes formed during meiotic recombination (Modrich and Lahue 1996). Repair of these mismatches either restores the sequence originally on that chromatid or replaces it with the sequence of the homologous chromatid. When a mismatch within heteroduplex DNA is not repaired through meiosis, both sequences become fixed in the first round of DNA synthesis, an event also known as post-meiotic segregation (PMS). Along with Msh2, Rad1 specifically affects the repair of large DNA loops formed during meiotic recombination (Kirkpatrick and Petes 1997). The Drosophila XPF homolog Mei-9 is involved in meiotic recombination (Baker and Carpenter 1972). *Mei-9* mutants are sensitive to UV and ionizing radiation due to a defect in the NER and DSB repair pathways (Dusenbery *et al* 1983). In addition, these mutants are defective in the repair of mismatches within meiotic heteroduplex DNA, apparent from high levels of PMS, suggesting that Mei-9 has a function in a mismatch repair homologous recombination pathway (Romans 1980, Carpenter 1982). Moreover, Mei-9 has been shown to be required for certain types of DNA mismatch repair throughout the Drosophila life span (Bhui-Kaur *et al* 1998). Thus, Mei-9 is involved in NER as well as mismatch repair. Mismatch repair proteins can also bind to intrastrand cisplatin cross-links, which are normally repaired via the NER pathway (Duckett *et al* 1996, Mello *et al* 1996). This could explain the cellular resistance to cisplatin in mismatch repair mutants (Aebi *et al* 1996, Drummond *et al* 1996). Binding of human MSH2 may shield the damaged bases from repair. Alternatively, MSH2 may provoke incomplete, and therefore lethal, mismatch repair. Mismatched cisplatin lesions, produced by post replication repair DNA synthesis past the damaged base, are also bound by mismatch repair proteins (Yamada *et al* 1997). These compound lesions are substrates for NER too (Hess *et al* 1997, Moggs *et al* 1997). Probably, the increased structural alteration enhances damage recognition. Repair synthesis after excision of the cross-link, however, will copy the former mismatch, fixing a mutation in both strands. The non-complementary residues opposite the intrastrand cross-link enhance *in vitro* nucleotide excision repair of the lesion, which is, however, independent of the MSH2–MSH6 proteins (Moggs *et al* 1997). Thus, no functional overlap between the mismatch and nucleotide excision repair pathways for the repair of these lesions has been observed *in vitro* (Moggs *et al* 1997, Mu *et al* 1997a).

Repair interplay

In short, the ERCC1–XPF proteins seem to be involved in at least three different repair pathways: NER, recombination and mismatch repair. The means of repair, for which the ERCC1–XPF endonuclease activity is required, will be regulated by different protein–protein interactions. So far, only one association has been deter-

mined, the interaction with the damage recognition protein XPA. Other candidate factors are MSH2 and MSH3. Especially the XPF protein contains a large region with several motifs implicated in protein binding that will be interesting for further analysis of the role of ERCC1–XPF in the different repair processes.

ERCC1 and xeroderma pigmentosum group F

XP group F

Xeroderma pigmentosum is associated with a defect in one of the XP factors involved in NER. As mentioned before, XP is characterized by severe photosensitivity, pigmentation abnormalities and predisposition to skin cancer. The median age of 830 reported XP patients, reviewed in 1987, was only twelve years (Kraemer *et al* 1987). Between one and two years of age on average, XP patients begin to show sun-sensitivity and freckling. At the moment of recording, half of the patients had developed skin cancer, predominantly basal and squamous cell carcinomas, almost exclusively in sun-exposed areas of the skin. The average age of onset was eight years. Ocular abnormalities were reported in 40% of the cases, whereas neurological signs appeared in nearly 20% of the patients. Progressive mental retardation, hypo- or areflexia and/or deafness appeared in nearly all these cases before the age of five. Neurological degeneration may be caused by an accumulation of unrepaired endogenous oxidative damage leading to premature death of nerve cells (Satoh *et al* 1993, Reardon *et al* 1997). DNA damage is likely to be produced in extra large amounts in neurons due to their high metabolic activity, and may therefore contribute to progressive neurodegeneration in XP. XP group F patients show a rather mild phenotype. Their clinical features are summarized in Table 1. XP-F patients are only moderately sensitive to sunlight and the average age of onset of skin cancer was 40 years. Only three out of 25 individuals had ocular problems. Mild neurological signs, also detected in three patients, appeared in two of the patients much later in life (after their forties). Compared with other groups, the clinical features of group F XP patients are relatively mild.

XP-F mutations

The number of patients with group F xeroderma pigmentosum is relatively low (Cleaver and Kraemer 1989). To date, 18 cases with XP-F have been reported (Table 1). As the clinical features in most XP-F individuals are relatively mild, these may often remain unnoticed. Mutational analysis revealed that only two mutations reside in the N-terminal region of XPF. Similarly, only few mutations affect the C-terminal part close to the ERCC1-binding domain. In an English patient XP126LO, both mutations reside in the C-terminal part of the *XPF* gene. The frameshift mutation, 759fs, codes for a truncated XPF protein lacking the ERCC1-binding domain. The R788W point mutation resides outside the binding domain, however, alters the protein structure and is therefore thought to affect ERCC1 binding (Chapter 4). Nevertheless, the mutated XPF protein could bind ERCC1 *in vitro* and has been shown to have residual activity *in vivo*, which can explain the mild clinical features. The R788W mutation has also been found one Dutch and in

Patient description			Skin characteristics			Tumors
Code	Gender	Age	Acute erythema	MED / peak	Freckles (Age of onset)	no and type (Age of onset)
Japanese						
XP23OS	F	45			+ (6)	–
XP101OS	F	49	+[d]	low	+	1 MF (30), 2 BCC, 1 KA (47)
XP2YO[a]	F	64			+	1 SSC (62), BDC († 65)
XP3YO[a]	M	29			+	1 KA (26)
XP25KO	F	8	+	low / 24h	+ (1.5)	–
XP27KO[b]	F	11	+	nor / 24h	+ (3)	–
XP28KO[b]	F	8	+		+ (3)	–
XP41KO[b]	M	5	+		+ (3)	–
XP38KO	F	44	+	nor / 24h	+	–
XP46KO	F	61	+	nor / 24h	+ (14)	–
XP7KA	F	42	+	low / 24+72h	+ (4)	2 BCC (40)
XP90TO[c]	M	42	+	nor / 48h	+ (10)	6 BCC (42)
XP92TO[c]	F	41	+	nor / 48h	+ (5)	1 BCC (41)
XP107TO[1]	F	73	+	low / 48+72h	+ (5)	2 BCC, 1 TF (72)
Kps6	F	18	±		+ (11)	1 BCC (18)
XP24KY[2]	M	48			+ (6)	1 BDC († 50)
Nps1	M	34				–
Nps8	M	63				–
European						
XP126LO	F	22	+	nor / 24h	+ (10)	–
XP42RO	M	62	+	nor / 24+72h	+ (10)	8 BCC, 1 SCC, 1 KA (27)
XP7NE	M	28		low / 48 h	+	–
XP24BR	F	29	+	low / 24	+	–
XP26BR	F	20	+	low / 24	+	–
XP32BR	M	7	+		+	
96RD239	M	15	+		+ (15)	–

[1] also known as XP1TS
[2] also known as MNNH
[a, b, c] Family relationship
[d] not after UV
[e] 1 (Arase *et al* 1979)
2 (Nishigori *et al* 1986)
3 (Nishigori *et al* 1991)
4 (Takebe *et al* 1980)
5 (Fujiwara *et al* 1985b)
6 (Yamamura and Ichihashi 1989)
7 (Fujiwara *et al* 1985a)
8 (Arase *et al* 1988)
9 (Kondo *et al* 1989)
10 (Itoh *et al* 1995)
11 (Moriwaki *et al* 1993)
12 (Itoh *et al* 1994)
13 (Norris *et al* 1988)
14 Chapter 4
[f] Jaspers personal communication

Table 1 Overview of XP-F cases

Other features				Mutation(s)		
Ocular	Neurological	UV-sensitivity (fold more sensitive than normal)	%NER	1	2	reference [a]
	–	4.3	10	444fs	444fs	1
–	–	2	15	I214M	G502R	2
	–	3.4	17	T556A	646fs	3,4
	–	2.8	15	R479Q	L599P	3,4
–	–	3	10-15			5,6
–	–	3	10-15			5,6
–	–	3	10-15			5,6
–	–					5,6
–	–	2.3	20-25			6,7
–	–	2.8	12			6
–	–	4.3	19	E491K	I518T	6,8
±	–	3	12			9
–	–	3	12			9
	–	2.3	19	594del	594del	3
–	–	1.3	20.5			10
–	+	3.1	7	R443W	525fs	11
						12
			17			12
–	–		13	R788W	759fs	13
±	+	2	15-30	R788W	R788W	14
	possibly		20-30	P368S	nd	-[f]
		3	20-30	R788W	R577W	-[f]
–	–	2	15	R788W	R788W	-[f]
	–	2	10			-[f]
+	+	5	<5	R142P	R142P	-[f]

BCC basal cell carcinoma
SCC sqamous cell carcinoma
KA keratoacanthoma
TF trichofolliculoma
MF mammary fibroadenoma
BDC bile duct cancer
nd not determined

other English XP-F patients. Most patients carry the point mutation in at least one allele, which suggests that the mutation originates from a common ancestor (Fig. 7). In contrast, various unique mutations were found among the Japanese XP-F patients (Matsumura *et al* 1998). Different from the European mutations, these are concentrated in the central part of the XPF protein (Fig. 7).

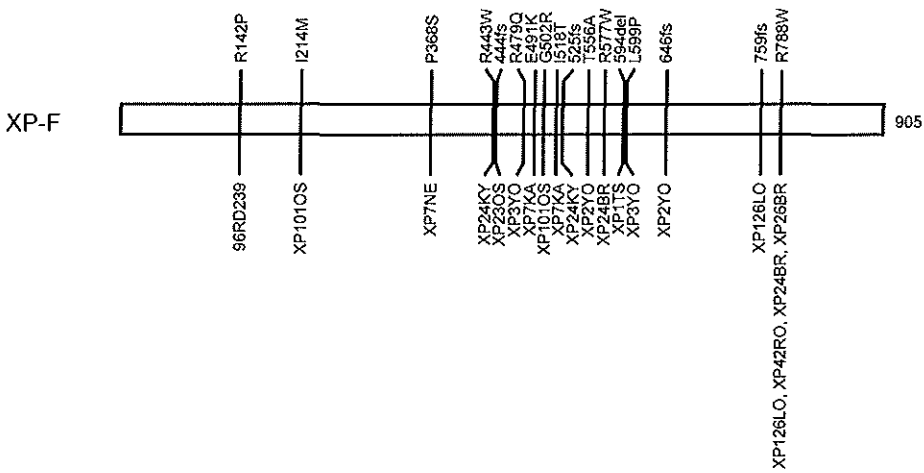


Fig. 7 XP-F mutations. Sites of mutations in the *XPF* gene product. Gray boxes indicate best conserved regions.

Except for the XP23OS cells, all examined XP-F cells contain at least one allele coding for a mutated, though, full-length XPF protein. Although, these mutations did not affect transcript stability, the mutant proteins were hardly detectable (Matsumura *et al* 1998). Like R788W, small amounts of these mutated proteins may still be able to interact with ERCC1 and as such responsible for some residual repair activity. In XP23OS cells, only one mutated transcript has been found, which produces a truncated protein lacking the entire C-terminal half including the ERCC1-binding domain. Despite this severe mutation, the patient's phenotype is as mild as described for all other group F patients. The mutation, however, has not been verified at the level of the gene, and it is therefore still possible that also in this patient a small amount of full-length (mutated) XPF is expressed from the other allele.

ERCC1–XPF residual activity

The residual amount of complex present in XP-F cells probably accounts for slow but long-lasting repair (Zelle *et al* 1980). Previous mutational analysis in rodent mutants revealed that the degree of sensitivity to UV-light and MMC depends on the amount of ERCC1(–XPF complex) expressed (Chapter 2). The residual repair activity may, therefore, account for the moderate UV-sensitivity associated with XP-F, contrasting with the extremely low cellular repair activity, immediately after

UV. Moreover, small amounts of ERCC1–XPF are probably required for normal development and viability, since mice carrying a homozygous null-mutation in *ERCC1* are severely affected.

ERCC1-syndrome

A human syndrome associated with a defect in ERCC1 has not yet been identified. ERCC1 is not involved in any of the known XP, CS or TTD complementation groups (van Duin *et al* 1989 and unpublished results). It is possible, though, that an ERCC1 defect leads to a different phenotype, not represented among existing XP groups. Candidate diseases could be Fanconi's anaemia or Roberts syndrome, both associated with severe cross-link sensitivity (Burns and Tomkins 1989). Mutations in ERCC1 may, on the other hand, interfere with viability because of its additional function in recombination.

Mice with an XPA defect show many of the characteristics of XP. They are photosensitive and predisposed to UV-induced skin cancers. These mice are viable, develop normally and have a normal life span. In contrast, mice carrying non-functional *ERCC1* alleles show a more severe and complex phenotype. They have a markedly reduced life span, severely runted phenotype, abnormalities in liver nuclei, and reduced liver functions (McWhir *et al* 1993, Weeda *et al* 1997a). No pathological abnormalities are noticed in the heart, lung, brain and eyes, nor obvious neurological abnormalities (Weeda *et al* 1997a). Sexual development seems normal, although most mice die before the reproductive age and no mice have been shown to be fertile.

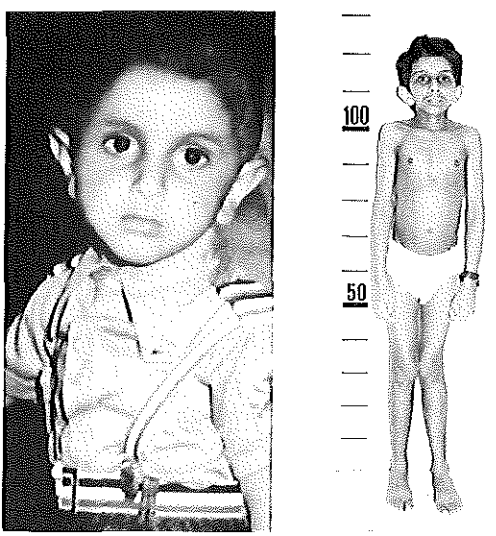


Fig. 8 Individual with the unusual severe form of XP-F at the age of 8 (left) and 16 (right) years old. Courtesy of Meinecke.

The mice have a thin skin and lack subcutaneous fat. Their kidneys contain dilated tubules associated with renal dysfunction. Nuclear abnormalities were found not only in liver, but also in kidney (increased octaploidy, coarse chromatin and prominent nucleoli with abnormal organization, nuclear invagination and intranuclear inclusions, abnormal nuclear size). Primary fibroblasts showed no increase in chromosomal instability nor programmed cell death.

Case report

The proposita is a 16-year-old Afghan boy (Fig. 8). His parents are cousins. His two younger brothers are normal. Maternal age was 17 years. As reported by his parents, his birth measurements were quite normal and early development was only slightly delayed. At age 8 months, he was treated for a hepatitis A infection. Several years later, he again suffered from recurrent fevers (possibly tuberculosis) and was treated with rifampicin for 2 to 3 years. Already in his first years of life, he showed marked sun-sensitivity.

From the age of 10, his parents noticed his oldish appearance, narrow face and loss of subcutaneous fat, and not before the age of 12, his failure to grow. He was examined in India for growth hormones. In the following years, he lost weight and suffered from dizziness. He needed assistance with dressing, but could eat himself. After consumption of milk products he had intermittent diarrhea. According to his father, he made water frequently during day and night. He had learning difficulties and hearing loss.

At the age of 15, he was evaluated for short stature and developmental delay. He presented at Altonaer Kinder-Krankenhaus (Germany) with dystrophy, dwarfism (height of 120.2 cm, weight of 17.6 kg), microcephaly (head circumference of 46.9 cm), birdlike face, normal hair, and an overall prematurely aged appearance. He had no carries, but enamel dysplasia. His skin was dry and atrophic with irregular pigmentation, after severe sun burn, mainly in face and neck, and no skin cancer. Secondary sexual characteristics were present with descended testis. According to his parents he was mentally delayed. Furthermore, mild neurological impairment with disturbed coordination and mild ataxia were observed. The electroencephalogram (EEG) demonstrated mild generalized slowing. He had a high pitched voice. Ophthalmologic examination revealed optic atrophy and narrow retinal vessels (from age 6, he wear glasses.). Results of examination of the lungs were unremarkable, whereas the heart was enlarged. Spleen and liver were not clearly separated. Liver malfunctioning (not due to an infection) was apparent from elevated serum liver enzyme levels (AP 1288 U/l, γ-GT 76 U/l, GPT 29 U/l), an elevated blood ferritin level, pointing to chronic liver cell injury, as well as low levels of serum phosphate and calcium. He also suffered from severe kidney insufficiency with secondary osteopenia (bone reduction), anemia, acidosis. Both kidneys were relatively small and condensed with limited glomerulus filtrate due to reduced levels of serum albumin. Blood pressure was 160/110 mm Hg (hypertension). Sonographic studies revealed not only the presence of ascites but also pleural fluid. His fibroblasts showed a complete loss of cellular repair activity (UV-induced UDS was less than 5% of normal), normal karyotype and no increased spontaneous chromosomal breakage. Complementation analysis assigned the patient to xeroderma pigmentosum group F. Due to progressive kidney problems (increasing levels of creatinine) and a pneumonia, he died from a multi-organ failure at the age of 16.

Features	New XP-F patient	ERCC1-deficient mice
Clinical		
Growth	retarded	retarded
Pigmentation abnormalities	+	
Skin cancer	–	
Lack of subcutaneous fat	+	+
Prematurely aged appearance	+	+
Renal dysfunction	+	+ (dilated tubules)
Liver malfunction		
– alanine aminotransferase		15-fold increase[1]
– γ-glutamyl transferase	elevated (76 U/l)	2-fold increase
– alkaline phosphatase	elevated (1288 U/l)	2-fold increase[2]
– bilirubin		7-fold increase[2]
– serum albumin	reduced (22.0 g/l)	reduced (20.5 g/l)
– serum iron	elevated[1]	
Hemolytic anemia	+	+
Hypertension	+	
Optic atrophy	+	
Neurological abnormalities	+ (mild impairment)	
Reduced life span	+ (16 yr)	+ (3 wk - 6 mnd, depending on genetic background)
Cellular		
Nuclear abnormalities	nd	liver and kidney
Chromosomal instability	–	–
UV-sensitivity	5x	+
MMC-sensitivity	3x	+
Repair activity (UV-induced UDS)	<5%	4%

[1] indicative of liver cell injury
[2] indicative of reduced liver excretory functions
[nd] not determined

Table 2 Comparison XP-F patient with ERCC1-deficient mice

The mutant mice have an increase in Howel Jolly bodies (granular structures within erythrocytes) in peripheral blood, pointing to hemolytic anemia, as well as an elevated ferritin deposition in spleen, which is indicative of an increased turnover of erythrocytes. Some of the features point to premature senescence.
The UV-induced repair activity in their fibroblasts was extremely low (4% of normal). These cells were hypersensitive to UV irradiation and MMC treatment as well and showed early seizure in proliferation (replicative senescence) (Weeda *et al* 1997a). Furthermore, ERCC1-deficient mouse embryonic stem cells carry a defect in gene targeting by homologous recombination when targeting constructs con-

tain heterologous ends (Weeda personal communication).

Liver, kidney and brain are metabolically very active. The involvement of liver and kidney suggests that ERCC1–XPF is important in the repair of some endogenous damage associated with high metabolism. Candidates for endogenous metabolites are carbonyl compounds, which are found in the liver and can produce DNA interstrand cross-links, such as malondialdehyde, a side product of the fatty acid metabolism (Smith and Pereira-Smith 1996). The absence of ERCC1 or XPF probably results in the accumulation of interstrand cross-links giving rise to early onset of cell cycle arrest and polyploidy, explaining the runted phenotype, the poor and delayed development and early death associated with ERCC1–XPF deficiency.

Novel repair syndrome

Recently, we have come across an exceptional XP-F patient, whose clinical and cellular characteristics strikingly resemble those of the *ERCC1* null mutant mice (see case report and Table 2). In addition to the severe sun-sensitivity and neurological degeneration normally associated with XP, and developmental retardation reminiscent of CS, this XP-F patient has severe liver and kidney abnormalities. The patient is homozygous for a point mutation in the N-terminus of the XPF protein. A conserved arginine residue present in one of the leucine-rich motifs is changed into a proline (R142P in Fig. 7). The mutation results in an undetectable amount of ERCC1–XPF. The additional symptoms not associated with XP, such as the kidney and liver problems might be due to the absence of the possible recombination function of ERCC1–XPF.

These new features associated with a defect in ERCC1–XPF may be of help identifying ERCC1 and XP-F patients in the future.

References

Aboussekhra A, Biggerstaff M, Shivji MKK, Vilpo JA, Moncollin V, Podust VN, Protic M, Hübscher U, Egly J-M, Wood RD (1995) Mammalian DNA nucleotide excision repair reconstituted with purified components. *Cell* 80: 859-868

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252: 1651-1656

Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Buit CJ, Lee NH, Kirkness EF, Weinstock KG, Gocayne JD, White O, Sutton G, Blake JA, Brandon RC, Chiu MW, Clayton RA, Cline RT, Cotton MD, Earle-Hughes J, Fine LD, Fitzgerald LM, FitzHugh WM, Fritchman JL, Geoghagen NSM, Glodek A, Gnehm CL, Hanna MC and others (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* 377S: 173-174

Aebi S, Kurdi-Haidar B, Gordon R, Cenni B, Zheng H, Fink D, Christen RD, Boland CR, Koi M, Fishel R, Howell SB (1996) Loss of DNA mismatch repair in acquired resistance to cisplatin. *Cancer Res* 56: 3087-3090

Alani E (1996) The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Mol Cell Biol* 16: 5604-5615

Alani E, Lee S, Kane MF, Griffith J, Kolodner RD (1997) *Saccharomyces cerevisiae* Msh2, a mispaired base recognition protein, also recognizes Holliday junctions in DNA. *J Mol Biol* 265: 289-301

Alonso JC, Stiege AC, Dobrinski B, Lurz R (1993) Purification and properties of the RecR protein from *Bacillus subtilis* 168. *J Biol Chem* 268: 1424-1429

Andrews AD, Barrett SF, Robbins JH (1978) Xeroderma pigmentosum neurological abnormalities correlate with colony-forming ability after ultraviolet radiation. *Proc Natl Acad Sci USA* 75: 1984-1988

Araj H, Smith PD (1996) Positional cloning of the *Drosophila melanogaster* mei-9 gene, the putative homolog of the *Saccharomyces cerevisiae Rad1* gene. *Mutation Res* 364: 209-215

Arase S, Kozuka T, Tanaka K, Ikenaga M, Takebe H (1979) A sixth complementation group in xeroderma pigmentosum. *Mutation Res* 59: 143-146

Arase S, Takada M, Watanabe Y, Nakanishi H (1988) A case of xeroderma pigmentosum complementation group F. *Hifubyoh-Shinryoh* 10: 57-60

Arase S, Nakanishi H, Kodama S, Ishizaki K (1990) Liquid-holding recovery of sister chromatid exchanges in UV-irradiated normal and xeroderma pigmentosum complementation groups A and F fibroblasts. *Arch Dermatol Res* 282: 68-70

Asahara H, Wistort PM, Bank JF, Bakerian RH, Cunningham RP (1989) Purification and characterization of *Escherichia coli* endonuclease III from the cloned nth gene. *Biochemistry* 28: 4444-4449

Asahina H, Kuraoka I, Shirakawa M, Morita EH, Miura N, Miyamoto I, Ohtsuka E, Okada Y, Tanaka K (1994) The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage. *Mutation Res* 315: 229-237

Au KG, Welsh K, Modrich P (1992) Initiation of methyl-directed mismatch repair. *J Biol Chem* 267: 12142-12148

Ayora S, Stiege AC, Lurz R, Alonso JC (1997) *Bacillus subtilis* 168 RecR protein–DNA complexes visualized as looped structures. *Mol Gen Genet* 254: 54-62

Bailly V, Sommers CH, Sung P, Prakash L, Prakash S (1992) Specific complex formation between proteins encoded by the yeast DNA repair and recombination genes *Rad1* and *Rad10*. *Proc Natl Acad Sci USA* 89: 8273-8277

Baker BS, Carpenter AT (1972) Genetic analysis of sex chromosomal meiotic mutants in *Drosophila melanogaster*. *Genetics* 71: 255-286

Bardwell L, Cooper AJ, Friedberg EC (1992) Stable and specific association between the yeast recombination and DNA repair proteins Rad1 and Rad10 *in vitro*. *Mol Cell Biol* 12: 3041-3049

Bardwell AJ, Bardwell L, Johnson DK, Friedberg EC (1993) Yeast DNA recombination and repair proteins Rad1 and Rad10 constitute a complex *in vivo* mediated by localized hydrophobic domains. *Mol Microbiol* 8: 1177-1188

Bardwell AJ, Bardwell L, Tomkinson AE, Friedberg EC (1994) Specific cleavage of model recombination and repair intermediates by the yeast Rad1–Rad10 DNA endonuclease. *Science* 265: 2082-2085

Belt PB, van Oosterwijk MF, Odijk H, Hoeijmakers JHJ, Backendorf C (1991) Induction of a mutant phenotype in human repair proficient cells after overexpression of a mutated human DNA repair gene. *Nucleic Acids Res* 19: 5633-5637

Bessho T, Mu D, Sancar A (1997a) Initiation of DNA interstrand cross-link repair in humans: the nucleotide excision repair system makes dual incisions 5′ to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand. *Mol Cell Biol* 17: 6822-6830

Bessho T, Sancar A, Thompson LH, Thelen MP (1997b) Reconstitution of human excision nuclease with recombinant XPF–ERCC1 complex. *J Biol Chem* 272: 3833-3837

Bhui-Kaur A, Goodman MF, Tower J (1998) DNA mismatch repair catalyzed by extracts of mitotic, postmitotic, and senescent Drosophila tissues and involvement of mei-9 gene function for full activity. *Mol Cell Biol* 18: 1436-1443

Biggerstaff M, Szymkowski DE, Wood RD (1993) Co-correction of ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects *in vitro*. *EMBO J* 12: 3685-3692

Bleumink E, Nater JP, Schraffordt-Koops H, The TH (1974) A standard method for DNCB sensitization testing in patients with neoplasms. *Cancer* 33: 911-915

Bootsma D, Hoeijmakers JHJ (1993) Engagement with transcription. *Nature* 363: 114-115

Bootsma D, Kraemer KH, Cleaver JE, Hoeijmakers JHJ (1998) Nucleotide excision repair syndromes: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. In *The genetic basis of human cancer*. Vogelstein B, Kinzler K (eds) McGraw-Hill Book Co. New York: 245-274

Bowman KK, Sidik K, Smith CA, Taylor JS, Doetsch PW, Freyer GA (1994) A new ATP-independent DNA endonuclease from *Schizosaccharomyces pombe* that recognizes cyclobutane pyrimidine dimers and 6-4 photoproducts. *Nucleic Acids Res* 22: 3026-3032

Bramson J, McQuillan A, Panasci LC (1995) DNA repair enzyme expression in chronic lymphocytic leukemia *vis-a-vis* nitrogen mustard drug resistance. *Cancer Lett* 90: 139-148

Brookman KW, Lamerdin JE, Thelen MP, Hwang M, Reardon JT, Sancar A, Zhou ZQ, Walter CA, Parris CN, Thompson LH (1996) *ERCC4(XPF)* encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Mol Cell Biol* 16: 6553-6562

Broughton BC, Thompson AF, Harcourt SA, Vermeulen W, Hoeijmakers JHJ, Botta E, Stefanini M, King MD, Weber CA, Cole J, Arlett CF, Lehmann AR (1995) Molecular and cellular analysis of the DNA repair defect in a patient in xeroderma pigmentosum complementation group D who has the clinical features of xeroderma pigmentosum and Cockayne syndrome. *Am J Human Genet* 56: 167-174

Burns MA, Tomkins DJ (1989) Hypersensitivity to mitomycin C cell-killing in Roberts syndrome fibroblasts with, but not without, the heterochromatin abnormality. *Mutation Res* 216: 243-249

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

Busch D, Greiner C, Rosenfeld KL, Ford R, de Wit J, Hoeijmakers JHJ, Thompson LH (1994) Complementation group assignments of moderately UV-sensitive CHO mutants isolated by large-scale screening. *Mutagenesis* 9: 301-306

Carpenter AT (1982) Mismatch repair, gene conversion, and crossing-over in two recombination-defective mutants of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 79: 5961-5965

Carr AM, Schmidt H, Kirchhoff S, Muriel WJ, Sheldrick KS, Griffiths DJ, Basmacioglu CN, Subramani S, Clegg M, Nasim A, Lehmann AR (1994) The Rad16 gene of *Schizosaccharomyces pombe*: a homolog of the *Rad1* gene of *Saccharomyces cerevisiae*. *Mol Cell Biol* 14: 2029-2040

Chen F, Peterson S, Story M, Chen D (1996) Disruption of DNA-PK in Ku80 mutant xrs-6 and the implications in DNA double-strand break repair. *Mutation Res* 362: 9-19

Chen N-Y, Zhang J-J, Paulus H (1989) Chromosomal location of the *Bacillus subtilis* aspartokinase II gene and nucleotide sequence of the adjacent genes homologous to uvrC and trx of *Escherichia coli*. *J Gen Microbiol* 135: 2931-2940

Cheng S, van Houten B, Gamper HB, Sancar A, Hearst JE (1988) Use of psoralen-modified oligonucleotides to trap three-stranded RecA–DNA complexes and repair of these cross-linked complexes by ABC excinuclease. *J Biol Chem* 263: 15110-15117

Cheng S, Sancar A, Hearst JE (1991) RecA-dependent incision of psoralen-crosslinked DNA by (A)BC excinuclease. *Nucleic Acids Res* 19: 657-663

Christy KJ, LaTart D, Osterhoudt W (1989) Modifications for SDS-PAGE of proteins. *Biotechniques* 7: 692-693

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

Chu G, Chang E (1990) Cisplatin-resistant cells express increased levels of a factor that recognizes damaged DNA. *Proc Natl Acad Sci USA* 87: 3324-3328

Chu G (1997) Double strand break repair. *J Biol Chem* 272: 24097-24100

Chyan YJ, Ackerman S, Shepherd NS, McBride OW, Widen SG, Wilson SH, Wood TG (1994) The human DNA polymerase β gene structure. Evidence of alternative splicing in gene expression. *Nucleic Acids Res* 22: 2719-2725

Cleaver JE, Kraemer KH (1989) Xeroderma pigmentosum and Cockayne syndrome. In *The metabolic basis for inherited disease.* Scriver CR, Beaudet AL, Sly WS, Valle D (eds) McGraw-Hill Book Co. New York: 2949-2971

Coverley D, Kenny MK, Lane DP, Wood RD (1992) A role for the human single-stranded DNA binding protein HSSB/RPA in an early stage of nucleotide excision repair. *Nucleic Acids Res* 20: 3873-3880

Dabholkar M, Bostick-Bruton F, Weber C, Bohr VA, Egwuagu C, Reed E (1992) ERCC1 and ERCC2 expression in malignant tissues from ovarian cancer patients. *J Natl Cancer Inst* 84: 1513-1517

Dabholkar M, Vionnet J, Bostick-Bruton F, Yu JJ, Reed E (1994) Messenger RNA levels of XPAC and ERCC1 in ovarian cancer tissue correlate with response to platinum-based chemotherapy. *J Clin Invest* 94: 703-708

Davies AA, Friedberg EC, Tomkinson AE, Wood RD, West SC (1995) Role of the Rad1 and Rad10 proteins in nucleotide excision repair and recombination. *J Biol Chem* 270: 24638-24641

de Laat WL, Appeldoorn E, Jaspers NGJ, Hoeijmakers JHJ (1998a) DNA structural elements required for ERCC1–XPF endonuclease activity. *J Biol Chem* 273: 7835-7842

de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NGJ, Hoeijmakers JHJ (1998b) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* 12: 2598-2609

de Vries A, van Oostrom CTM, Hofhuis FMA, Dortant PM, Berg RJW, de Gruijl FR, Wester PW, van Kreijl CF, Capel PJA, van Steeg H, Verbeek SJ (1995) Increased susceptibility to ultraviolet-B and carcinogens of mice lacking the DNA excision repair gene *XPA*. *Nature* 377: 169-173

Dingwall C, Laskey RA (1991) Nuclear targeting sequences – a consensus? *Trends Biochem Sci* 16: 478-481

Dodson ML, Michaels ML, Lloyd RS (1994) Unified catalytic mechanism for DNA glycosylases. *J Biol Chem* 269: 32709-32712

Doetsch PW (1995) What's old is new: an alternative DNA excision repair pathway. *Trends Biochem Sci* 20: 384-386

Doherty AJ, Serpell LC, Ponting CP (1996) The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res* 24: 2488-2497

Doolittle RF, Johnson MS, Husain I, van Houten B, Thomas DC, Sancar A (1986) Domainal evolution of a prokaryotic DNA repair protein and its relationship to active transport proteins. *Nature* 323: 451-453

Drapkin R, Reardon JT, Ansari A, Huang JC, Zawel L, Ahn K, Sancar A, Reinberg D (1994) Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* 368: 769-772

Drummond JT, Li GM, Longley MJ, Modrich P (1995) Isolation of an hMSH2–p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science* 268: 1909-1912

Drummond JT, Anthoney A, Brown R, Modrich P (1996) Cisplatin and adriamycin resistance are associated with MutLα and mismatch repair deficiency in an ovarian tumor cell line. *J Biol Chem* 271: 19645-19648

Duckett DR, Drummond JT, Murchie AI, Reardon JT, Sancar A, Lilley DM, Modrich P (1996) Human MutSα recognizes damaged DNA base pairs containing O^6-methylguanine, O^4-methylthymine, or the cisplatin-d(GpG) adduct. *Proc Natl Acad Sci USA* 93: 6443-6447

Dusenbery RL, McCormick SC, Smith PD (1983) Drosophila mutations at the mei-9 and mus(2)201 loci which block excision of thymine dimers also block induction of unscheduled DNA synthesis by methyl methanesulfonate, ethyl methanesulfonate, N-methyl-N-nitrosourea, UV light and X-rays. *Mutation Res* 112: 215-230

Egel R, Beach DH, Klar AJ (1984) Genes required for initiation and resolution steps of mating-type switching in fission yeast. *Proc Natl Acad Sci USA* 81: 3481-3485

Egel R (1989) Mating-type genes, meiosis, and sporulation. In *Molecular biology of the fission yeast.* Nasim A, Young P, Johnson BF (eds) Academic Press San Diego CA: 31-73

Evans E, Fellows J, Coffer A, Wood RD (1997a) Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J* 16: 625-638

Evans E, Moggs J, Hwang J, Egly J-M, Wood RD (1997b) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J* 16: 6559-6573

Felzenszwalb I, Boiteux S, Laval J (1992) Molecular cloning and DNA sequencing of the radC gene of *Escherichia coli* K-12. *Mutation Res* 273: 263-269

Fiorentini P, Huang KN, Tishkoff DX, Kolodner RD, Symington LS (1997) Exonuclease I of *Saccharomyces cerevisiae* functions in mitotic recombination *in vivo* and *in vitro*. *Mol Cell Biol* 17: 2764-2773

Fishel R, Kolodner RD (1995) Identification of mismatch repair genes and their role in the development of cancer. *Curr Opin Genet Dev* 5: 382-395

Fishman-Lobell J, Haber JE (1992) Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene *Rad1*. *Science* 258: 480-484

Fleck O, Michael H, Heim L (1992) The Swi4 gene of *Schizosaccharomyces pombe* encodes a homologue of mismatch repair enzymes. *Nucleic Acids Res* 20: 2271-2278

Formosa T, Alberts BM (1986) DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. *Cell* 47: 793-806

Freyer GA, Davey S, Ferrer JV, Martin AM, Beach D, Doetsch PW (1995) An alternative eukaryotic DNA excision repair pathway. *Mol Cell Biol* 15: 4572-4577

Friedberg EC, Walker GC, Siede W (1995) *DNA repair and mutagenesis*. ASM Press Washington DC

Frosina G, Fortini P, Rossi O, Carrozzino F, Raspaglio G, Cox LS, Lane DP, Abbondandolo A, Dogliotti E (1996) Two pathways for base excision repair in mammalian cells. *J Biol Chem* 271: 9573-9578

Fujiwara Y, Ichihashi M, Uehara Y, Matsumoto A, Yamamoto Y, Kano Y, Tanakura Y (1985a) Xeroderma pigmentosum groups C and F: additional assignments and a review of the subjects in Japan. *J Radiat Res* 26: 443-449

Fujiwara Y, Uehara Y, Ichihashi M, Nishioka K (1985b) Xeroderma pigmentosum complementation group F: more assignments and repair characteristics. *Photochem Photobiol* 41: 629-634

Gaffney TD, Lam ST, Ligon J, Gates K, Frazelle A, Maio J, Hill S, Goodwin S, Torkewitz N, Allshouse AM, Kempf HJ, Becker JO (1994) Global regulation of expression of antifungal factors by a *Pseudomonas fluorescens* biological control strain. *Mol Plant Microbe Interact* 7: 455-463

Gary R, Ludwig DL, Cornelius HL, MacInnes MA, Park MS (1997) The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21. *J Biol Chem* 272: 24522-24529

Geleziunas R, McQuillan A, Malapetsa A, Hutchinson M, Kopriva D, Wainberg MA, Hiscott J, Bramson J, Panasci L (1991) Increased DNA synthesis and repair-enzyme expression in lymphocytes from patients with chronic lymphocytic leukemia resistant to nitrogen mustards. *J Natl Cancer Inst* 83: 557-564

Gorman C, Padmanabhan R, Howard BH (1983) High efficiency DNA-mediated transformation of primate cells. *Science* 221: 551-553

Graham FL, van der Eb A (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52: 456-467

Grilley M, Welsh KM, Su SS, Modrich P (1989) Isolation and characterization of the *Escherichia coli* mutL gene product. *J Biol Chem* 264: 1000-1004

Grilley M, Griffith J, Modrich P (1993) Bidirectional excision in methyl-directed mismatch repair. *J Biol Chem* 268: 11830-11837

Gutman PD, Minton KW (1993) Conserved sites in the 5'→3'exonuclease domain of *Escherichia coli* DNA polymerase. *Nucleic Acids Res* 21: 4406-4407

Gutz H, Schmidt H (1985) Switching genes in *Schizosaccharomyces pombe*. *Curr Genet* 9: 325-331

Guzder S, Habraken Y, Sung P, Prakash L, Prakash S (1995) Reconstitution of yeast nucleotide excision repair with purified Rad proteins, replication protein A, and transcription factor TFIIH. *J Biol Chem* 270: 12973-12976

Guzder SN, Sung P, Prakash L, Prakash S (1996) Nucleotide excision repair in yeast is mediated by sequential assembly of repair factors and not by a pre-assembled repairosome. *J Biol Chem* 271: 8903-8910

Habraken Y, Sung P, Prakash L, Prakash S (1995) Structure-specific nuclease activity in yeast nucleotide excision repair protein Rad2. *J Biol Chem* 270: 30194-30198

Habraken Y, Sung P, Prakash S, Prakash L (1996) Transcription factor TFIIH and DNA endonuclease Rad2 constitute yeast nucleotide excision repair factor 3: implications for nucleotide excision repair and Cockayne syndrome. *Proc Natl Acad Sci USA* 93: 10718-10722

Hamel BC, Raams A, Schuitema-Dijkstra AR, Simons P, van der Burgt I, Jaspers NGJ, Kleijer WJ (1996) Xeroderma pigmentosum–Cockayne syndrome complex: a further case. *J Med Genet* 33: 607-610

Hanawalt PC (1994) Transcription-coupled repair and human disease. *Science* 266: 1957-1958

Hang H, Hager DN, Goriparthi L, Hopkins KM, Shih H, Lieberman HB (1996) *Schizosaccharomyces pombe Rad23* is allelic with *Swi10*, a mating-type switching/radioresistance gene that shares sequence homology with human and mouse *ERCC1*. *Gene* 170: 113-117

Harrington JJ, Lieber MR (1994a) Functional domains within Fen-1 and Rad2 define a family of structure-specific endonucleases: implications for nucleotide excision repair. *Genes Dev* 8: 1344-1355

Harrington JJ, Lieber MR (1994b) The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J* 13:1235-1246

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

Hatakeyama S, Ito Y, Shimane A, Ishii C, Inoue H (1998) Cloning and characterization of the yeast Rad1 homolog gene (*Mus-38*) from *Neurospora crassa*: evidence for involvement in nucleotide excision repair. *Curr Genet* 33: 276-283

Hays SL, Firmenich AA, Berg P (1995) Complex formation in yeast double-strand break repair: participation of Rad51, Rad52, Rad55, and Rad57 proteins. *Proc Natl Acad Sci USA* 92: 6925-6929

He Z, Henricksen LA, Wold MS, Ingles CJ (1995) RPA involvement in the damage-recognition and incision step of nucleotide excision repair. *Nature* 374: 566-569

He Z, Ingles CJ (1997) Isolation of human complexes proficient in nucleotide excision repair. *Nucleic Acids Res* 25: 1136-1141

Henning KA, Li L, Iyer N, McDaniel LD, Reagan MS, Legerski R, Schultz RA, Stefanini M, Lehmann AR, Mayne LV, Friedberg EC (1995) The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. *Cell* 82: 555-564

Hess MT, Schwitter U, Petretta M, Giese B, Naegeli H (1997) Bipartite substrate discrimination by human nucleotide excision repair. *Proc Natl Acad Sci USA* 94: 6664-6669

Hobohm U, Sander C (1995) A sequence property approach to searching protein databases. *J Mol Biol* 251: 390-399

Hoeijmakers JHJ (1994) Human nucleotide excision repair syndromes: molecular clues to unexpected intricacies. *Eur J Cancer* 30A: 1912-1921

Hoy CA, Thompson LH, Mooney CL, Salazar EP (1985) Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. *Cancer Res* 45: 1737-1743

Huang JC, Svoboda DL, Reardon JT, Sancar A (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc Natl Acad Sci USA* 89: 3664-3668

Humbert S, van Vuuren AJ, Lutz Y, Hoeijmakers JHJ, Egly J-M, Moncollin V (1994) Characterization of p44/SSL1 and p34 subunits of the BTF2/TFIIH transcription/repair factor. *EMBO J* 13: 2393-2398

Ishino Y, Shinagawa H, Makino K, Tsunasawa S, Sakiyama F, Nakata A (1986) Nucleotide sequence of the lig gene and primary structure of DNA ligase of *Escherichia coli*. *Mol Gen Genet* 204: 1-7

Itoh T, Ono T, Yamizumi M (1994) A new UV-sensitive syndrome not belonging to any complementation groups of xeroderma pigmentosum or Cockayne syndrome: siblings showing biochemical characteristics of

Cockayne syndrome without typical clinical manifestations. *Mutation Res* 314: 233-248

Itoh T, Watanabe H, Yamaizumi M, Ono T (1995) A young woman with xeroderma pigmentosum complementation group F and a morphoeic basal cell carcinoma. *Br J Dermatol* 132: 122-127

Itoh T, Aiba H, Baba T, Hayashi K, Inada T, Isono K, Kasai H, Kimura S, Kitakawa M, Kitagawa M, Makino K, Miki T, Mizobuchi K, Mori H, Mori T, Motomura K, Nakade S, Nakamura Y, Nashimoto H, Nishio Y, Oshima T, Saito N, Sampei G, Seki Y, Horiuchi T and others (1996) A 460-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 40.1-50.0 min region on the linkage map. *DNA Res* 3: 379-392

Ivanov EL, Haber JE (1995) *Rad1* and *Rad10*, but not other excision repair genes, are required for double-strand break-induced recombination in *Saccharomyces cerevisiae*. *Mol Cell Biol* 15: 2245-2251

Iyer N, Reagan MS, Wu K-J, Canagarajah B, Friedberg EC (1996) Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochemistry* 35: 2157-2167

Johnson CV, Singer RH, Lawrence JB (1991) Fluorescent detection of nuclear RNA and DNA: implications for genome organization. *Methods Cell Biol* 35: 73-99

Johnson RE, Kovvali GK, Prakash L, Prakash S (1996) Requirement of the yeast *Msh3* and *Msh6* genes for Msh2-dependent genomic stability. *J Biol Chem* 271: 7285-7288

Jones CJ, Wood RD (1993) Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochemistry* 32: 12096-12104

Kanaar R, Hoeijmakers JHJ (1997) Recombination and joining: different means to the same ends. *Genes Funct* 1: 165-174

Keijzer W, Jaspers NGJ, Abrahams PJ, Taylor AMR, Arlett CF, Zelle B, Takebe H, Kinmont PDS, Bootsma D (1979) A seventh complementation group in excision deficient xeroderma pigmentosum. *Mutation Res* 62: 183-190

Keyse SM, Emslie EA (1992) Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase. *Nature* 359: 644-646

Kim C, Snyder RO, Wold MS (1992) Binding properties of replication protein-A from human and yeast cells. *Mol Cell Biol* 12: 3050-3059

Kim Y, Eom SH, Wang J, Lee D-S, Suh SW, Steitz TA (1995) Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature* 376: 612-616

Kirkpatrick DT, Petes TD (1997) Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature* 387: 929-931

Klein HL (1988) Different types of recombination events are controlled by the *Rad1* and *Rad52* genes of *Saccharomyces cerevisiae*. *Genetics* 120: 367-377

Klungland A, Lindahl T (1997) Second pathway for completion of human DNA base excision repair: reconstitution with purified proteins and requirement for DNase IV (FEN-1). *EMBO J* 16: 3341-3348

Kobe B, Deisenhofer J (1993) Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* 366: 751-756

Kobe B, Deisenhofer J (1995) Proteins with leucine-rich repeats. *Curr Opin Struct Biol* 5: 409-416

Koken MHM, Odijk H, van Duin M, Fornerod M, Bootsma D (1993) Augmentation of protein production by a combination of the T7 RNA polymerase system and ubiquitin fusion. Overproduction of the human DNA repair protein, ERCC1, as a ubiquitin fusion protein in *Escherichia coli*. *Biochem Biophys Res Comm* 195: 643-653

Kondo S, Mamada A, Miyamoto C, Keong CH, Satoh Y, Fujiwara Y (1989) Late onset of skin cancers in two xeroderma pigmentosum group F siblings and a review of 30 Japanese xeroderma pigmentosum patients in groups D, E and F. *Photodermatol* 6: 89-95

Kozak M (1987) An analysis of 5' non-coding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* 15: 8125-8148

Kraemer KH, Lee MM, Scotto J (1987) Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases. *Arch Dermatol* 123: 241-250

Kumar A, Abbotts J, Karawya EM, Wilson SH (1990a) Identification and properties of the catalytic domain of mammalian DNA polymerase β. *Biochemistry* 29: 7156-7159

Kumar A, Widen SG, Williams KR, Kedar P, Karpel RL, Wilson SH (1990b) Studies of the domain structure of mammalian DNA polymerase β. Identification of a discrete template binding domain. *J Biol Chem* 265: 2124-2131

Kunkel TA, Robetts JD, Zakour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol* 154: 367-382

Kuo CF, McRee DE, Fisher CL, O'Handley SF, Cunningham RP, Tainer JA (1992) Atomic structure of the DNA repair [4Fe-4S] enzyme endonuclease III. *Science* 258: 434-440

Labahn J, Scharer OD, Long A, Ezaz-Nikpay K, Verdine GL, Ellenberger TE (1996) Structural basis for the excision repair of alkylation-damaged DNA. *Cell* 86: 321-329

Landegent JE, Jansen in de Wal N, Baan RA, Hoeijmakers JHJ, van der Ploeg M (1984) 2-Acetylaminofluorene-modified probes for the indirect hybridocytochemical detection of specific nucleic acid sequences. *Exp Cell Res* 153: 61-72

Landschulz WH, Johnson PF, McKnight SL (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240: 1759-1764

Lawyer FC, Stoffel S, Saiki RK, Myambo K, Drummond R, Gelfand DH (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J Biol Chem* 264: 6427-6437

Lehmann AR, Bootsma D, Clarkson SG, Cleaver JE, McAlpine PJ, Tanaka K, Thompson LH, Wood RD (1994) Nomenclature of human DNA repair genes. *Mutation Res* 315: 41-42

Lehmann AR (1995) Nucleotide excision repair and the link with transcription. *Trends Biochem Sci* 20: 402-405

Levin PA, Margolis PS, Setlow P, Losick R, Sun D (1992) Identification of *Bacillus subtilis* genes for septum placement and shape determination. *J Bacteriol* 174: 6717-6728

Li L, Elledge SJ, Peterson CA, Bales ES, Legerski RJ (1994) Specific association between the human DNA repair proteins XPA and ERCC1. *Proc Natl Acad Sci USA* 91: 5012-5016

Li L, Lu X, Peterson CA, Legerski RJ (1995a) An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Mol Cell Biol* 15: 5396-5402

Li L, Peterson CA, Lu X, Legerski RJ (1995b) Mutations in XPA that prevent association with ERRC1 are defective in nuclear excision repair. *Mol Cell Biol* 15: 1993-1998

Lieber MR (1997) The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* 19:233-240

Lin JJ, Sancar A (1990) Reconstitution of nucleotide excision nuclease with UvrA and UvrB proteins from *Escherichia coli* and UvrC protein from *Bacillus subtilis*. *J Biol Chem* 265: 21337-21341

Lin JJ, Sancar A (1991) The C-terminal half of UvrC protein is sufficient to reconstitute (A)BC excinuclease. *Proc Natl Acad Sci USA* 88: 6824-6828

Lin JJ, Sancar A (1992) Active site of (A)BC excinuclease. I. Evidence for 5' incision by UvrC through a catalytic site involving Asp399, Asp438, Asp466, and His538 residues. *J Biol Chem* 267: 17688-17692

Liu P, Siciliano J, White B, Legerski R, Callen D, Reeders S, Siciliano MJ, Thompson LH (1993) Regional mapping of human DNA excision repair gene *ERCC4* to chromosome 16p13.13-p13.2. *Mutagenesis* 8: 199-205

Liu QQ, Sommer SS (1995) Restriction endonuclease fingerprinting (REF): a sensitive method for screening mutations in long, contiguous segments of DNA. *Biotechniques* 18: 470-477

Marsischky GT, Filosi N, Kane MF, Kolodner R (1996) Redundancy of *Saccharomyces cerevisiae* Msh3 and Msh6 in Msh2-dependent mismatch repair. *Genes Dev* 10: 407-420

Masutani C, Sugasawa K, Yanagisawa J, Sonoyama T, Ui M, Enomoto T, Takio K, Tanaka K, van der Spek PJ, Bootsma D, Hoeijmakers JHJ, Hanaoka F (1994) Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homolog of yeast Rad23. *EMBO J* 13: 1831-1843

Matsuda T, Saijo M, Kuraoka I, Kobayashi T, Nahatssu Y, Nagai A, Enjoji T, Masutani C, Sugasawa K, Hanaoka F, Yasui A, Tanaka K (1995) DNA repair protein XPA binds to replication protein A (RPA). *J Biol Chem* 270: 4152-4157

Matsumoto Y, Kim K, Bogenhagen DF (1994) Proliferating cell nuclear antigen-dependent abasic site repair in *Xenopus laevis* oocytes: an alternative pathway of base excision DNA repair. *Mol Cell Biol* 14: 6187-6197

Matsumura Y, Nishigori C, Yagi T, Imamura S, Takebe H (1998) Characterization of molecular defects in xeroderma pigmentosum group F in relation to its clinically mild symptoms. *Hum Mol Genet* 7: 969-974

Matsunaga T, Mu D, Park C-H, Reardon JT, Sancar A (1995) Human DNA repair excision nuclease. Analysis of the roles of the subunits involved in dual incisions by using anti-XPG and anti-ERCC1 antibodies. *J Biol Chem* 270: 20862-20869

Matsunaga T, Park C-H, Bessho T, Mu D, Sancar A (1996) Replication protein A confers structure-specific endonuclease activities to the XPF–ERCC1 and XPG subunits of human DNA repair excision nuclease. *J Biol Chem* 271: 11047-11050

Mazur SJ, Grossman L (1991) Dimerization of *Escherichia coli* UvrA and its binding to undamaged and ultraviolet light damaged DNA. *Biochemistry* 30: 4432-4443

McCready S, Carr AM, Lehmann AR (1993) Repair of cyclobutane pyrimidine dimers and 6-4 photoproducts in the fission yeast *Schizosaccharomyces pombe*. *Mol Microbiol* 10: 885-890

McWhir J, Seldridge J, Harrison DJ, Squires S, Melton DW (1993) Mice with DNA repair gene (*ERCC1*) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nature Genet* 5: 217-224

Mello JA, Acharya S, Fishel R, Essigmann JM (1996) The mismatch-repair protein hMSH2 binds selectively to DNA adducts of the anticancer drug cisplatin. *Chem Biol* 3: 579-589

Mimaki T, Nitta M, Saijo M, Tachi N, Minami R, Tanaka K (1996) Truncated XPA protein detected in atypical group A xeroderma pigmentosum. *Acta Paediatr* 85: 511-513

Miura M, Domon M, Sasaki T, Takasaki Y (1992) Induction of proliferating cell nuclear antigen (PCNA) complex formation in quiescent fibroblasts from a xeroderma pigmentosum patient. *J Cell Physiol* 150: 370-376

Modrich P (1994) Mismatch repair, genetic stability, and cancer. *Science* 266: 1959-1960

Modrich P, Lahue R (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem* 65: 101-133

Moggs JG, Yarema KJ, Essigmann JM, Wood RD (1996) Analysis of incision sites produced by human cells extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *J Biol Chem* 271: 7177-7186

Moggs JG, Szymkowski DE, Yamada M, Karran P, Wood RD (1997) Differential human nucleotide excision repair of paired and mispaired cisplatin–DNA adducts. *Nucleic Acids Res* 25: 480-491

Moolenaar GF, Franken KLMC, Dijkstra DM, Thomas-Oates JE, Visse R, van de Putte P, Goosen N (1995) The C-terminal region of the UvrB protein of *Escherichia coli* contains an important determinant for UvrC binding to the preincision complex but not the catalytic site for 3' incision. *J Biol Chem* 270: 30508-30515

Moolenaar GF, Uiterkamp R, Zwijnenburg D, Goosen N (1998) The C-terminal region of the *Escherichia coli* UvrC protein, which is homologous to the C-terminal region of the human ERCC1 protein is involved in DNA binding and 5' incision. *Nucleic Acids Res* 26: 462-468

Moriwaki S, Nishigori C, Imamura S, Yagi T, Takahashi C, Fujimoto N, Takebe H (1993) A case of xeroderma pigmentosum complementation group F with neurological abnormalities. *Br J Dermatol* 128: 91-94

Mu D, Park C-H, Matsunaga T, Hsu DS, Reardon JT, Sancar A (1995) Reconstitution of human DNA repair excision nuclease in a highly defined system. *J Biol Chem* 270: 2415-2418

Mu D, Hsu DS, Sancar A (1996) Reaction mechanism of human DNA repair excision nuclease. *J Biol Chem* 271: 8285-8294

Mu D, Sancar A (1997) Model for XPC-independent transcription-coupled repair of pyrimidine dimers in humans. *J Biol Chem* 272: 7570-7573

Mu D, Tursun M, Duckett DR, Drummond JT, Modrich P, Sancar A (1997a) Recognition and repair of compound DNA lesions (base damage and mismatch) by human mismatch repair and excision repair systems. *Mol Cell Biol* 17: 760-769

Mu D, Wakasugi M, Hsu DS, Sancar A (1997b) Characterization of reaction intermediates of human excision repair nuclease. *J Biol Chem* 272: 28971-28979

Mueser TC, Nossal NG, Hyde CC (1996) Structure of bacteriophage T4 RNase H, a 5'→3' RNA–DNA and DNA–DNA exonuclease with sequence similarity to the Rad2 family of eukaryotic proteins. *Cell* 85: 1101-1112

Murray JM, Tavassoli M, Al-Harithy R, Sheldrick KS, Lehmann AR, Carr AM, Watts FZ (1994) Structural and functional conservation of the human homolog of the *S. pombe Rad2* gene, which is required for chromosome segregation and recovery from DNA damage. *Mol Cell Biol* 14: 4878-4888

Myles GM, Sancar A (1991) Isolation and characterization of functional domains of UvrA. *Biochemistry* 30: 3834-3840

Nagai A, Saijo M, Kuraoka I, Matsuda T, Kodo N, Nakatsu Y, Mimaki T, Mino M, Biggerstaff M, Wood RD, Sijbers A, Hoeijmakers JHJ, Tanaka K (1995) Enhancement of damage-specific DNA binding of XPA by interaction with the ERCC1 DNA repair protein. *Biochem Biophys Res Comm* 211: 960-966

Nakane H, Takeuchi S, Yuba S, Saijo M, Nakatsu Y, Ishikawa T, Hirota S, Kitamura Y, Kato Y, Tsunoda Y, Miyauchi H, Horio T, Tokunaga T, Matsunaga T, Nikaido O, Nishimune Y, Okada Y, Tanaka K (1995) High incidence of ultraviolet-B- or chemical-carcinogen-induced skin tumours in mice lacking the xeroderma pigmentosum group A gene. *Nature* 377: 165-168

Nance MA, Berry SA (1992) Cockayne syndrome: review of 140 cases. *Am J Med Genet* 42: 68-84

Nassif N, Penney J, Pal S, Engels WR, Gloor GB (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol* 14: 1613-1625

Nichols AF, Sancar A (1992) Purification of PCNA as a nucleotide excision repair protein. *Nucleic Acids Res* 20: 2441-2446

Nishigori C, Ishizaki K, Takebe H, Imamura S, Hayakawa M (1986) A case of xeroderma pigmentosum group F with late onset of clinical symptoms. *Arch Dermatol* 122: 510-511

Nishigori C, Fujisawa H, Uyeno K, Kawaguchi T, Takebe H (1991) Xeroderma pigmentosum patients belonging to complementation group F and efficient liquid-holding recovery of ultraviolet damage. *Photodermatol Photoimmunol Photomed* 8: 146-150

Nocentini S, Coin F, Saijo M, Tanaka K, Egly J-M (1997) DNA damage recognition by XPA protein promotes efficient recruitment of transcription factor II H. *J Biol Chem* 272: 22991-22994

Norris PG, Hawk JLM, Giannelli F (1988) Xeroderma pigmentosum complementation group F in a non-Japanese patient. *J Am Acad Dermatol* 18: 1185-1188

Nouspikel T, Lalle P, Leadon SA, Cooper PK, Clarkson SG (1997) A common mutational pattern in Cockayne syndrome patients from xeroderma pigmentosum group G: implications for a second XPG function. *Proc Natl Acad Sci USA* 94: 3116-3121

O'Donovan A, Davies AA, Moggs JG, West SC, Wood RD (1994a) XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* 371: 432-435

O'Donovan A, Scherly D, Clarkson SG, Wood RD (1994b) Isolation of active recombinant XPG protein, a human DNA repair endonuclease. *J Biol Chem* 269: 15965-15968

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250: 4007-4021.

Ogawa T, Shinohara A, Nabetani A, Ikeya T, Yu X, Egelman EH, Ogawa H (1993) RecA-like recombination proteins in eukaryotes: functions and structures of *Rad51* genes. *Cold Spring Harb Symp Quant Biol* 58: 567-576

O'Hare K, Benoist C, Breathnach R (1981) Transformation of mouse fibroblasts to methotrexate resistance by a recombinant plasmid expressing a prokaryotic dihydrofolate reductase. *Proc Natl Acad Sci USA* 78: 1527-1531

Otrin VR, McLenigan M, Takao M, Levine AS, Protic M (1997) Translocation of a UV-damaged DNA binding protein into a tight association with chromatin after treatment of mammalian cells with UV-light. *J Cell Sci* 110: 1159-1168

Palombo F, Gallinari P, Iaccarino I, Lettieri T, Hughes M, DqArrigo A, Truong O, Hsuan JJ, Jiricny J (1995) GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science* 268: 1912-1914

Palombo F, Iaccarino I, Nakajima E, Ikejima M, Shimada T, Jiricny J (1996) hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr Biol* 6: 1181-1184

Park C-H, Sancar A (1994) Formation of a ternary complex by human XPA, ERCC1 and ERCC4(XPF) excision repair proteins. *Proc Natl Acad Sci USA* 91: 5017-5021

Park C-H, Bessho T, Matsunaga T, Sancar A (1995a) Purification and characterization of the XPF–ERCC1 complex of human DNA repair excision nuclease. *J Biol Chem* 270: 22657-22660

Park C-H, Mu D, Reardon JT, Sancar A (1995b) The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIE transcription factor. *J Biol Chem* 270: 4896-4902

Paz-Elizur T, Takeshita M, Goodman M, O'Donnell M, Livneh Z (1996) Mechanism of translesion DNA synthesis by DNA polymerase II. Comparison to DNA polymerases I and III core. *J Biol Chem* 271: 24662-24669

Pelletier H, Sawaya MR, Wolfle W, Wilson SH, Kraut J (1996) Crystal structures of human DNA polymerase β complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* 35: 12742-12761

Podust LM, Podust VA, Floth C, Hübscher U (1994) Assembly of DNA polymerase δ and ε holoenzymes depends on the geometry of the DNA template. *Nucleic Acids Res* 22: 2970-2975

Protic M, Levine AS (1993) Detection of DNA damage-recognition proteins using the bandshift assay and southwestern hybridization. *Electrophoresis* 14: 682-692

Reardon JT, Nichols AF, Keeney S, Smith CA, Taylor JS, Linn S, Sancar A (1993a) Comparative analysis of binding of human damaged DNA-binding protein (XPE) and *Escherichia coli* damage recognition protein (UvrA) to the major ultraviolet photoproducts - T[CS]TT[Ts]TT[6-4]T and T[Dewar]T. *J Biol Chem* 268: 21301-21308

Reardon JT, Thompson LH, Sancar A (1993b) Excision repair in man and the molecular basis of xeroderma pigmentosum. *Cold Spring Harb Symp Quant Biol* 58: 605-617

Reardon JT, Mu D, Sancar A (1996) Overproduction, purification, and characterization of the XPC subunit of the human DNA repair excision nuclease. *J Biol Chem* 271: 19451-19456

Reardon JT, Bessho T, Kung HC, Bolton PH, Sancar A (1997) *In vitro* repair of oxidative DNA damage by human nucleotide excision repair system: possible explanation for neurodegeneration in xeroderma pigmentosum patients. *Proc Natl Acad Sci USA* 94: 9463-9468

Reynolds P, Prakash L, Dumais D, Perozzi G, Prakash S (1985) Nucleotide sequence of the *Rad10* gene of *Saccharomyces cerevisiae*. *EMBO J* 4: 3549-3552

Reynolds P, Prakash L, Prakash S (1987) Nucleotide sequence and functional analysis of the *Rad1* gene of *S. cerevisiae*. *Mol Cell Biol* 7: 1012-1120

Riboni R, Botta E, Stefanini M, Numata M, Yasui A (1992) Identification of the eleventh complementation group of UV-sensitive excision repair-defective rodent mutants. *Cancer Res* 52: 6690-6691

Robbins JH, Brumback RA, Moshell AN (1993) Clinically asymptomatic xeroderma pigmentosum neurological disease in an adult: evidence for a neurodegeneration in later life caused by defective DNA repair. *Eur Neurol* 33: 188-190

Roberts RJ (1995) On base flipping. *Cell* 82: 9-12

Robins P, Jones CJ, Biggerstaff M, Lindahl T, Wood RD (1991) Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA. *EMBO J* 10: 3913-3921

Robins P, Pappin DJC, Wood RD, Lindahl T (1994) Structural and functional homology between mammalian DNase IV and the 5'nuclease domain of *Escherichia coli* DNA polymerase I. *J Biol Chem* 269: 28535-28538

Rödel C, Kirchhoff S, Schmidt H (1992) The protein sequence and some intron positions are conserved between the switching gene *Swi10* of *Schizosaccharomyces pombe* and the human excision repair gene *ERCC1*. *Nucleic Acids Res* 20: 6347-6353

Rödel C, Jupitz T, Schmidt H (1997) Complementation of the DNA repair-deficient *swi10* mutant of fission yeast by the human *ERCC1* gene. *Nucleic Acids Res* 25: 2823-2827

Romans P (1980) Gene conversion in *mei-9*^{1}, a crossover defective mutant in *Drosophila melanogaster*. *Dros Inf Serv* 55: 130-132

Roy R, Schaeffer L, Humbert S, Vermeulen W, Weeda G, Egly J-M (1994) The DNA-dependent ATPase activity associated with the class II basic transcription factor BTF2/TFIIH. *J Biol Chem* 269: 9826-9832

Saffran WA, Greenberg RB, Thaler MS, Jones MM (1994) Single strand and double strand DNA damage-induced reciprocal recombination in yeast. Dependence on nucleotide excision repair and *Rad1* recombination. *Nucleic Acids Res* 22: 2823-2829.

Saijo M, Kuraoka I, Masutani C, Hanaoka F, Tanaka K (1996) Sequential binding of DNA repair proteins RPA and ERCC1 to XPA *in vitro*. *Nucleic Acids Res* 24: 4719-4724

Sambrook J, Fitsch E, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press Cold Spring Harbor NY

Sancar GB, Sancar A, Rupp WD (1984) Sequences of the *E. coli* uvrC gene and protein. *Nucleic Acids Res* 12: 4593-4608

Sancar GB (1990) DNA photolyases: physical properties, action mechanism, and roles in dark repair. *Mutation Res* 236: 147-160

Sands AT, Abuin A, Sanchez A, Conti CJ, Bradley A (1995) High susceptibility to ultraviolet-induced carcinogenesis in mice lacking *XPC*. *Nature* 377: 162-165

Saparbaev M, Prakash L, Prakash S (1996) Requirement of mismatch repair genes *Msh2* and *Msh3* in the Rad1–Rad10 pathway of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics* 142: 727-736

Sargent RG, Rolig RL, Kilburn AE, Adair GM, Wilson JH, Nairn RS (1997) Recombination-dependent deletion formation in mammalian cells deficient in the nucleotide excision repair gene *ERCC1*. *Proc Natl Acad Sci USA* 94: 13122-13127

Satoh MS, Jones CJ, Wood RD, Lindahl T (1993) DNA excision-repair defect of xeroderma pigmentosum prevents removal of a class of oxygen free radical-induced base lesions. *Proc Natl Acad Sci USA* 90: 6335-6339

Sawaya MR, Pelletier H, Kumar A, Wilson SH, Kraut J (1994) Crystal structure of rat DNA polymerase β: evidence for a common polymerase mechanism. *Science* 264: 1930-1935

Schaeffer L, Roy R, Humbert S, Moncollin V, Vermeulen W, Hoeijmakers JHJ, Chambon P, Egly J-M (1993) DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* 260: 58-63

Schaeffer L, Moncollin V, Roy R, Staub A, Mezzina M, Sarasin A, Weeda G, Hoeijmakers JHJ, Egly J-M (1994) The ERCC2 DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J* 13: 2388-2392

Schiestl RH, Prakash S (1988) *Rad1*, an excision repair gene of *Saccharomyces cerevisiae*, is also involved in recombination. *Mol Cell Biol* 8: 3619-3626

Schiestl RH, Prakash S (1990) *Rad10*, an excision repair gene of *Saccharomyces cerevisiae*, is involved in the *Rad1* pathway of mitotic recombination. *Mol Cell Biol* 10: 2485-2491

Schmidt H, Kapitza-Fecke P, Stephen ER, Gutz H (1989) Some of the *Swi* genes of *Schizosaccharomyces pombe* also have a function in the repair of radiation damage. *Curr Genet* 16: 89-94

Schneider R, Schneider-Scherzer E, Thurnher M, Auer B, Schweiger M (1988) The primary structure of human ribonuclease/angiogenin inhibitor (RAI) discloses a novel highly diversified protein superfamily with a common repetitive module. *EMBO J* 7: 4151-4156

Schneider R, Schweiger M (1991) The yeast DNA repair proteins Rad1 and Rad7 share similar putative functional domains. *FEBS Lett* 283: 203-206

Sekelsky JJ, McKim KS, Chin GM, Hawley RS (1995) The Drosophila meiotic recombination gene mei-9 encodes a homologue of the yeast excision repair protein Rad1. *Genetics* 141: 619-627

Selby CP, Sancar A (1997) Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *J Biol Chem* 272: 1885-1890

Seroussi E, Lavi S (1993) Replication protein A is the major single-stranded DNA binding protein detected in mammalian cell extracts by gel retardation assays and UV cross-linking of long and short single-stranded DNA molecules. *J Biol Chem* 268: 7147-7154

Shao Z, Newman EB (1993) Sequencing and characterization of the sdaB gene from *Escherichia coli* K-12. *Eur J Biochem* 212: 777-784

Shen B, Nolan JP, Sklar LA, Park MS (1997) Functional analysis of point mutations in human flap endonuclease-1 active site. *Nucleic Acids Res* 25: 3332-3338

Shivji MKK, Kenny MK, Wood RD (1992) Proliferating cell nuclear antigen is required for DNA excision repair. *Cell* 69: 367-374

Shivji MKK, Podust VN, Hubscher U, Wood RD (1995) Nucleotide excision repair DNA synthesis by DNA polymerase ε in the presence of PCNA, RF-C, and RPA. *Biochemistry* 34: 5011-5017

Siede W, Friedberg AS, Friedberg EC (1993) Evidence that the Rad1 and Rad10 proteins of *Saccharomyces cerevisiae* participate as a complex in nucleotide excision repair of UV radiation damage. *J Bacteriol* 175: 6345-6347

Sladek FM, Munn MM, Rupp WD, Howard-Flanders P (1989) *In vitro* repair of psoralen–DNA cross-links by RecA, UvrABC, and the 5'exonuclease of DNA polymerase I. *J Biol Chem* 264: 6755-6765

Smith JR, Pereira-Smith OM (1996) Replicative senescence: implications for *in vivo* aging and tumor suppression. *Science* 273: 63-67

States JC, Myrand SP (1996) Splice site mutations in a xeroderma pigmentosum group A patient with delayed onset of neurological disease. *Mutation Res* 363: 171-177

Stefanini M, Lagomarisini P, Gilliani S, Nardo T, Botta E, Peserico A, Kleyer WJ, Lehmann AR, Sarasin A (1993a) Genetic heterogeneity of the excision repair defect associated with trichothiodystrophy. *Carcinogenesis* 14: 1101-1105

Stefanini M, Vermeulen W, Weeda G, Giliani S, Nardo T, Mezzina M, Sarasin A, Harper JL, Arlett CF, Hoeijmakers JHJ, Lehmann AR (1993b) A new nucleotide excision repair gene associated with the disorder trichothiodystrophy. *Am J Hum Genet* 53: 817-821

Sugasawa K, Masutani C, Uchida A, Maekawa T, van der Spek PJ, Bootsma D, Hoeijmakers JHJ, Hanaoka F (1996) HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair *in vitro*. *Mol Cell Biol* 16: 4852-4861

Sugasawa K, Ng JMY, Masutani C, van der Spek PJ, Eker APM, Hanaoka F, Bootsma D, Hoeijmakers JHJ (1998) Xeroderma pigmentosum group C is the initiator of global genome nucleotide excision repair. *Mol Cell* 2: 223-232

Sugawara N, Haber JE (1992) Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. *Mol Cell Biol* 12: 563-575

Sugawara N, Paques F, Colaiacovo M, Haber JE (1997) Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc Natl Acad Sci USA* 94: 9214-9219

Sung P, Bailly V, Weber C, Thompson LH, Prakash L, Prakash S (1993a) Human xeroderma pigmentosum group D gene encodes a DNA helicase. *Nature* 365: 852-855

Sung P, Reynolds P, Prakash L, Prakash S (1993b) Purification and characterization of the *Saccharomyces cerevisiae* Rad1–Rad10 endonuclease. *J Biol Chem* 268: 26391-26399

Svejstrup JQ, Wang Z, Feaver WJ, Wu X, Bushnell DA, Donahue TF, Friedberg EC, Kornberg RD (1995) Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* 80: 21-28

Svoboda DL, Taylor JS, Hearst JE, Sancar A (1993) DNA repair by eukaryotic nucleotide excision nuclease. *J Biol Chem* 268: 1931-1936

Takebe H, Nikaido O, Ishizaki K, Yagi T, Sasaki M, Ikenaga M, Kozuka T, Fujiwara Y, Satoh Y (1980) Genetic aspects of xeroderma pigmentosum and other cancer-prone diseases. In *Genetic and environmental factors in experimental and human cancers*. Gelboin H, Magmahorn B, Matsushima T, Sugimura T, Takayama S, Takebe H (eds) Japan Science Societies Press Tokyo: 259-270

Teo I, Sedgwick B, Demple B, Li B, Lindahl T (1984) Induction of resistance to alkylating agents in *E. coli*: the ada+ gene product serves both as a regulatory protein and as an enzyme for repair of mutagenic damage. *EMBO J* 3: 2151-2157

Thayer MM, Ahern H, Xing D, Cunningham RP, Tainer JA (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J* 14: 4108-4120

Thompson LH, Brookman KW, Weber CA, Salazar EP, Reardon JT, Sancar A, Deng ZM, Siciliano MJ (1994) Molecular cloning of the human nucleotide excision repair gene *ERCC4*. *Proc Natl Acad Sci USA* 91: 6855-6859

Tishkoff DX, Boerger AL, Bertrand P, Filosi N, Gaida GM, Kane MF, Kolodner RD (1997) Identification and characterization of *Saccharomyces cerevisiae Exo1*, a gene encoding an exonuclease that interacts with Msh2. *Proc Natl Acad Sci USA* 94: 7487-7492

Todo T, Takemori H, Ryo H, Ihara M, Matsunaga T, Nikaido O, Sato K, Nomura T (1993) A new photoreactivating enzyme that specifically repairs ultraviolet light-induced (6-4)photoproducts. *Nature* 361: 371-374

Todo T, Ryo H, Yamamoto K, Toh H, Inui T, Ayaki H, Nomura T, Ikenaga M (1996) Similarity among the Drosophila (6-4)photolyase, a human photolyase homolog, and the DNA photolyase-blue-light photoreceptor family. *Science* 272: 109-112

Tomkinson AE, Bardwell AJ, Bardwell L, Tappe NJ, Friedberg EC (1993) Yeast DNA repair and recombination proteins Rad1 and Rad10 constitute a single-stranded-DNA endonuclease. *Nature* 362: 860-862

Tomkinson AE, Bardwell AJ, Tappe N, Ramos W, Friedberg EC (1994) Purification of Rad1 protein from *Saccharomyces cerevisiae* and further characterization of the Rad1–Rad10 endonuclease complex. *Biochemistry* 33: 5305-5311

Treiber DK, Chen Z, Essigmann JM (1992) An ultraviolet light-damaged DNA recognition protein absent in xeroderma pigmentosum group E cells binds selectively to pyrimidine (6-4) pyrimidone photoproducts. *Nucleic Acids Res* 20: 5805-5810

Troelstra C, Landsvater RM, Wiegant J, van der Ploeg M, Viel G, Buys CHCM, Hoeijmakers JHJ (1992a) Localization of the nucleotide excision repair gene *ERCC6* to human chromosome 10q11-q21. *Genomics* 12: 745-749

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

van der Spek PJ, Kobayashi K, Bootsma D, Takao M, Eker AP, Yasui A (1996) Cloning, tissue expression, and mapping of a human photolyase homolog with similarity to plant blue-light receptors. *Genomics* 37: 177-182

van Duin M, de Wit J, Odijk H, Westerveld A, Yasui A, Koken MHM, Hoeijmakers JHJ, Bootsma D (1986) Molecular characterization of the human excision repair gene *ERCC1:* cDNA cloning and amino acid homology with the yeast DNA repair gene *Rad10. Cell* 44: 913-923

van Duin M, Koken MHM, van den Tol J, ten Dijke P, Odijk H, Westerveld A, Bootsma D, Hoeijmakers JHJ (1987) Genomic characterization of the human DNA excision repair gene *ERCC1. Nucleic Acids Res* 15: 9195-9213

van Duin M, van den Tol J, Warmerdam P, Odijk H, Meijer D, Westerveld A, Bootsma D, Hoeijmakers JHJ (1988) Evolution and mutagenesis of the mammalian excision repair gene *ERCC1. Nucleic Acids Res* 16: 5305-5322

van Duin M, Vredeveldt G, Mayne LV, Odijk H, Vermeulen W, Klein B, Weeda G, Hoeijmakers JHJ, Bootsma D, Westerveld A (1989) The cloned human DNA excision repair gene *ERCC1* fails to correct xeroderma pigmentosum complementation groups A through I. *Mutation Res* 217: 83-92

van Gool AJ, Citterio E, Rademakers S, van Os R, Vermeulen W, Constantinou A, Egly J-M, Bootsma D, Hoeijmakers JHJ (1997a) The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerase II-containing complex. *EMBO J* 16: 5955-5965

van Hoffen A, Natarajan AT, Mayne LV, van Zeeland AA, Mullenders LHF, Venema J (1993) Deficient repair of the transcribed strand of active genes in Cockayne syndrome. *Nucleic Acids Res* 21: 5890-5895

van Houten B, Gamper H, Holbrook SR, Hearst JE, Sancar A (1986) Action mechanism of ABC excision nuclease on a DNA substrate containing a psoralen cross-link at a defined position. *Proc Natl Acad Sci USA* 83: 8077-8081

van Vuuren AJ, Appeldoorn E, Odijk H, Yasui A, Jaspers NGJ, Bootsma D, Hoeijmakers JHJ (1993) Evidence for a repair enzyme complex involving ERCC1 and complementing activities of ERCC4, ERCC11 and xeroderma pigmentosum group F. *EMBO J* 12: 3693-3701

van Vuuren AJ, Vermeulen W, Ma L, Weeda G, Appeldoorn E, Jaspers NGJ, van der Eb AJ, Bootsma D, Hoeijmakers JHJ, Humbert S, Schaeffer L, Egly J-M (1994) Correction of xeroderma pigmentosum repair defect by basal transcription factor BTF2(TFIIH). *EMBO J* 13: 1645-1653

van Vuuren AJ, Appeldoorn E, Odijk H, Humbert S, Moncollin V, Eker APM, Jaspers NGJ, Egly J-M, Hoeijmakers JHJ (1995) Partial characterization of the DNA repair protein complex, containing the ERCC1, ERCC4, ERCC11 and XPF correcting activities. *Mutation Res* 337: 25-39

Venema J, Mullenders LHF, Natarajan AT, van Zeeland AA, Mayne LV (1990) The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. *Proc Natl Acad Sci USA* 87: 4707-4711

Vermeulen W, Stefanini M, Giliani S, Hoeijmakers JHJ, Bootsma D (1991) Xeroderma pigmentosum complementation group H falls into complementation group D. *Mutation Res* 255: 201-208

Vermeulen W, Jaeken J, Jaspers NGJ, Bootsma D, Hoeijmakers JHJ (1993) Xeroderma pigmentosum complementation group G associated with Cockayne's syndrome. *Am J Human Genet* 53: 185-192

Vermeulen W, Scott RJ, Potger S, Muller HJ, Cole J, Arlett CF, Kleijer WJ, Bootsma D, Hoeijmakers JHJ, Weeda G (1994a) Clinical heterogeneity within xeroderma pigmentosum associated with mutations in the DNA repair and transcription gene *ERCC3*. *Am J Human Gen* 54: 191-200

Vermeulen W, van Vuuren AJ, Chipoulet M, Schaeffer L, Appeldoorn E, Weeda G, Jaspers NGJ, Priestley A, Arlett CF, Lehmann AR, Stefanini M, Mezzina M, Sarasin A, Bootsma D, Egly J-M, Hoeijmakers JHJ (1994b) Three unusual repair deficiencies associated with transcription factor BTF2(TFIIH). Evidence for the existence of a transcription syndrome. *Cold Spring Harb Symp Quant Biol* 59: 317-329

Vermeulen W, de Boer J, Citterio E, van Gool AJ, van der Horst GT, Jaspers NGJ, de Laat WL, Sijbers AM, van der Spek PJ, Sugasawa K, Weeda G, Winkler GS, Bootsma D, Egly JM, Hoeijmakers JHJ (1997) Mammalian nucleotide excision repair and syndromes. *Biochem Soc Trans* 25: 309-315

Weeda G, Donker I, de Wit J, Morreau H, Janssens R, Vissers CJ, Nigg A, van Steeg H, Bootsma D, Hoeijmakers JHJ (1997a) Disruption of mouse *ERCC1* results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr Biol* 7: 427-439

Weeda G, Eveno E, Donker I, Vermeulen W, Chevallier-Lagente O, Taieb A, Stary A, Hoeijmakers JHJ, Mezzina M, Sarasin A (1997b) A mutation in the *XPB/ERCC3* DNA repair transcription gene, associated with trichothiodystrophy. *Am J Hum Genet* 60: 320-329

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

White CI, Haber JE (1990) Intermediates of recombination during mating type switching in *Saccharomyces cerevisiae*. *EMBO J* 9: 663-673

Windisch JM, Auer B, Marksteiner R, Lang ME, Schneider R (1995a) Specific neurotrophin binding to leucine-rich motif peptides of TrkA and TrkB. *FEBS Lett* 374: 125-129

Windisch JM, Marksteiner R, Schneider R (1995b) Nerve growth factor binding site on TrkA mapped to a single 24-amino acid leucine-rich motif. *J Biol Chem* 270: 28133-28138

Wood RD, Burki HJ (1982) Repair capability and the cellular age response for killing and mutation induction after UV. *Mutation Res* 95: 505-514

Wood RD, Biggerstaff M, Shivji MKK (1995) Detection and measurement of nucleotide excision repair synthesis by mammalian cell extracts *in vitro*. *Methods* 7: 163-175

Wood RD (1996) DNA repair in eukaryotes. *Ann Rev Biochem* 65: 135-167

Wood RD (1997) Nucleotide excision repair in mammalian cells. *J Biol Chem* 272: 23465-23468

Wood RD, Shivji MKK (1997) Which DNA polymerases are used for DNA repair in eukaryotes? *Carcinogenesis* 18: 605-610

Xu H, Swoboda I, Bhalla PL, Sijbers AM, Zhao C, Ong E-K, Hoeijmakers JHJ, Singh MB (1998) Plant homologue of human excision repair gene *ERCC1* points to conservation of DNA repair mechanisms. *Plant J* 13: 823-829

Yagi T, Wood RD, Takebe H (1997) A low content of ERCC1 and a 120 kDa protein is a frequent feature of group F xeroderma pigmentosum fibroblast cells. *Mutagenesis* 12: 41-44

Yajima H, Takao M, Yasuhira S, Zhao JH, Ishii C, Inoue H, Yasui A (1995) A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light. *EMBO J* 14: 2393-2399

Yamada M, O'Regan E, Brown R, Karran P (1997) Selective recognition of a cisplatin–DNA adduct by human mismatch repair proteins. *Nucleic Acids Res* 25: 491-496

Yamamura K, Ichihashi M (1989) Clinical and photobiological characteristics of xeroderma pigmentosum complementation group F – a review of cases from Japan. *Br J Dermatol* 121: 471-480

Yao XD, Matecic M, Elias P (1997) Direct repeats of the herpes simplex virus a sequence promote nonconservative homologous recombination that is not dependent on XPF/ERCC4. *J Virol* 71: 6842-6849

Yasui A, Eker AP, Yasuhira S, Yajima H, Kobayashi T, Takao M, Oikawa A (1994) A new class of DNA photolyases present in various organisms including aplacental mammals. *EMBO J* 13: 6143-6151

Yasui A, Eeker APM (1998) DNA photolyases. In *DNA damage and repair*. Vol II *DNA repair in higher eukaryotes*. Nickoloff JA, Hoekstra MF (eds) Humana Press Inc. Totowa NJ: 9-32

Zelle B, Berends F, Lohman PH (1980) Repair of damage by ultraviolet radiation in xeroderma pigmentosum cell strains of complementation groups E and F. *Mutation Res* 73: 157-169

Abbreviations

BER	base excision repair
bp	base pair(s)
cDNA	complementary DNA
CFA	colony-forming ability
CHO	Chinese hamster ovary
CPD	cyclobutane pyrimidine dimer
CS	Cockayne syndrome
DDB	damaged DNA binding
DNA	deoxyribonucleic acid
DSB	double-strand break
ERCC	excision repair cross complementing
EST	expressed sequence tag
HhH	helix-hairpin-helix
kb	kilobases
kDa	kilodaltons
LRM	leucine-rich motif
Mei	meiosis
MMC	mitomycin C
Msh	MutS homolog
neo	neomycin
NER	nucleotide excision repair
NLS	nuclear location signal
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PMS	post-meiotic segregation
RACE	rapid amplification of cDNA ends
Rad	radiation
REF	restriction enzyme fingerprint
RNA	ribonucleic acid
RPA	replication protein A
RT-PCR	reverse transcriptase polymerase chain reaction
SEM	standard error of the mean
SSA	single-strand annealing
SSB	single-strand break
Swi	(mating-type) switching
TC-NER	transcription-coupled nucleotic excision repair
TFIIH	transcription initiation factor II H
TTD	trichothiodystrophy
UDS	unscheduled DNA synthesis
UV	ultraviolet
Uvr	ultraviolet resistance
wt	wildtype
XP	xeroderma pigmentosum

Summary

DNA is continuously under threat of numerous naturally occurring and man-made agents producing a wide variety of DNA lesions. These can cause cell death by blocking essential cellular processes, whereas persistent DNA lesions may bring about mutations in the DNA. Accumulating mutations play an important role in the development of tumors and processes like aging. It is clear that removal of DNA damage is of vital importance to the cell. Different repair pathways have evolved to protect the cell against DNA damage. These are summarized in Chapter one.

One of the main repair systems is the nucleotide excision repair (NER) pathway. A coordinated interplay between multiple proteins is required to carry out NER in eukaryotes. The first steps of the process lead to lesion recognition, local opening of the DNA double helix and incision of the damaged strand on each side of a lesion. A 23 to 32-mer oligonucleotide containing the lesion is removed, followed by gap-filling DNA synthesis. In human cells, this repair pathway involves the xeroderma pigmentosum (XP) proteins and associated factors. Individuals with the repair disorder xeroderma pigmentosum show hypersensitivity to sunlight, a greatly increased incidence of skin cancer, and in some cases neurological disease. Genes encoding the XPA, XPB, XPC, XPD, and XPG proteins have been isolated, and a factor defective in at least some XP-E cells has also been identified, although it is not required for the core NER system.

The two incisions made during NER are catalyzed by separate DNA endonucleases. In humans, XPG endonuclease makes the 3' incision relative to the lesion. In *Saccharomyces cerevisiae*, the Rad1 and Rad10 proteins form a heterodimeric complex having a structure-specific endonuclease activity with a polarity opposite to XPG, leading to the assumption that the Rad1–Rad10 complex makes the 5' incision during NER in yeast. The Rad1–Rad10 complex has an additional engagement in a mitotic recombination pathway, probably employed for the repair of DNA cross-links. Like Rad10, its mammalian homolog ERCC1 is involved in both NER and cross-link repair. Mutational analysis revealed that the poorly conserved N-terminal 91 residues of ERCC1 are dispensable for both repair functions in contrast to a deletion of only four residues from the C-terminus. These findings are reported in Chapter two. A database search revealed a strongly conserved motif in this C-terminus sharing sequence homology with many DNA break-producing and processing proteins, indicating that this part may be required for the presumed structure-specific endonuclease activity of ERCC1. Most missense mutations in the central region give rise to an unstable protein (complex). Accordingly, we found that free ERCC1 is very rapidly degraded, suggesting that protein–protein interactions provide stability. It seems that the removal of cross-links requires less ERCC1 than UV-damage repair.

ERCC1 has been found to associate with activities that correct human XP-F cell extracts as well as extracts from Chinese hamster cells of repair complementation groups 4 and 11. Furthermore, a polypeptide observed to co-purify with ERCC1 by several assays has been proposed as a candidate for a Rad1 homolog. We isolated a human gene homologous to yeast Rad1 and found that it corrects the repair defects of XP group F as well as rodent groups 4 and 11. These results demonstrate that the *XPF, ERCC4,* and *ERCC11* genes are equivalent and complete the isolation of the XP genes that form the core NER system as reported in Chapter three. The XPF protein was purified from mammalian cells in a tight complex with ERCC1. The complex indeed is a structure-specific endonuclease responsible for the 5' incision during repair. Causative mutations and strongly reduced levels of encoded protein were identified in XP-F patients.

Xeroderma pigmentosum group F patients generally show the clinical symptoms of a mild DNA repair deficiency. All XP-F patients examined demonstrate reduced levels of XPF and ERCC1 protein, suggesting that proper complex formation is required for stability of the two proteins. No humans with a defect in the ERCC1 subunit have been identified, and ERCC1-deficient mice suffer from severe developmental problems and signs of premature aging on top of a repair-deficient phenotype. To better understand the molecular and clinical consequences of mutations in the ERCC1 and XPF proteins, their mutual interaction domains were mapped (Chapter five). The XPF-binding domain comprises the final 73 C-terminal residues of ERCC1, whereas the ERCC1-binding domain in XPF maps to the last 91 C-terminal residues. Formation of the ERCC1–XPF complex is established by a direct interaction between these two domains. The frame shift mutation found in one allele of the first Caucasian XP-F patient described (XP126LO) results in a protein lacking the ERCC1-binding domain, which is shown to affect complex formation with ERCC1.

The identification of a second Caucasian XP patient belonging to complementation group F (XP42RO) is described in Chapter four. Mild ocular photophobia was present from childhood and acute skin reactions occurred upon exposure to sunlight. Basal and squamous cell carcinomas developed after his twenty-seventh year. In his late forties, progressive neurological symptoms emerged, which included intellectual decline, mild chorea and ataxia and marked cerebral, and cerebelar atrophy. Such neurologic abnormalities are very unusual in XP-F. Similar symptoms have been described in only one of 17 other reported XP-F individuals. His approximately 5-fold reduced activity of NER in cultured cells combined with moderately affected cell survival and DNA replication after UV exposure are typical of XP-F. The patient turned out to be homozygous for a point mutation in the *XPF* gene, causing an R788W substitution in the encoded protein. This mutation had also been found in one allele of the other unrelated Caucasian XP-F case. The amount of mutated XPF protein is strongly reduced in cells from XP42RO, presumably due to a conformational change. However, biochemical and clinical data indicate the presence of residual repair activity.

We have only recently identified an exceptional XP-F patient who strikingly resembles the ERCC1-deficient mice (Chapter six). His UV-sensitivity is very severe and his cells appear to have lost all excision repair activity. His homozygous point mutation results in an undetectable amount of ERCC1–XPF, which probably interferes with the possible recombination function of the endonuclease. The recombination defect may explain the additional symptoms not associated with XP, such as the kidney and liver problems found in the patient. The new repair syndrome described here is more reminiscent of a cell cycle arrest or premature aging disorder than of XP.

Samenvatting

VOOR DE LEEK

Alle levensvormen zijn opgebouwd uit cellen. Bacterien en gisten zijn eencellige organismen, terwijl planten, dieren en mensen uit een groot aantal cellen bestaan. In elke levende cel bevindt zich DNA dat er uit ziet als twee om elkaar gewikkelde strengen. Het DNA is de drager van de erfelijke informatie van een organisme. Het DNA van de mens bevat ongeveer 100.000 genen. Een gen is een stukje DNA dat informatie bevat voor een erfelijke eigenschap. Het vormt als het ware een recept voor het maken van een eiwit dat helpt die eigenschap tot stand te brengen. Deze recepten worden gespeld met vier letters: A, C, T en G, die op alle mogelijke manieren achter elkaar kunnen staan. In de twee strengen van de dubbele DNA keten vormen de letters tegenoverliggende paren en dat kan maar op één manier: A paart met T en G paart met C.
Helaas kan dit kostbare erfelijk materiaal schade oplopen door gebeurtenissen binnen in de cel, maar ook door invloeden van buitenaf, zoals chemische en radioactieve stoffen of de UV-straling van de zon. UV-straling kan bijvoorbeeld T's die naast elkaar zitten in de DNA streng met elkaar verkleven. Röntgenstraling kan de DNA strengen breken en sommige medicijnen tegen kanker knopen ze juist verkeerd aan elkaar. Doordat deze schades belangrijke processen in de cel blokkeren, kunnen de cellen afsterven en als ze dan toch overleven blijven er permanente veranderingen in de recepten achter, die men mutaties noemt. Een opeenhoping van mutaties speelt een belangrijke rol in het ontstaan van kanker en veroudering. Deze problemen kunnen enigszins worden ondervangen als schade tijdig kan worden gerepareerd.
Dat DNA-reparatie erg belangrijk is kan men zien bij patiënten die lijden aan de erfelijke ziekte xeroderma pigmentosum (XP). Ze kunnen de beschadigingen veroorzaakt door UV-licht niet repareren. Hun huid is extreem gevoelig voor de zon en daardoor erg droog (*xeroderma*) en vol onregelmatige pigmentvlekken (*pigmentosum*). De gevoeligheid voor UV-licht gaat bij deze patiënten vaak gepaard met een sterk verhoogde kans op huidkanker. In sommige gevallen is er ook sprake van ernstige problemen van het zenuwstelsel. In Nederland heeft minder dan één op de 200.000 mensen deze erfelijke ziekte. Verspreid over de wereld kunnen we op dit moment 7 verschillende typen xeroderma pigmentosum (A tot en met G) onderscheiden die berusten op afwijkingen in verschillende DNA-reparatie-genen. Dat zijn dus de genen die zelf het recept voor de reparatie-eiwitten dragen.
De reparatie-eiwitten (XPA, XPB, XPC etc genoemd) zorgen er samen voor dat de schades uit het DNA verwijderd worden. Dit proces, nucleotide excisie reparatie genoemd, omvat vijf stappen:

Herkenning van de DNA schade. De reparatie-eiwitten XPA, XPC, en mogelijk ook XPE, zoeken naar schades in het DNA, waar ze vervolgens op gaan zitten en andere reparatie-eiwitten naar de beschadigde plek lokken.
Opening van de DNA dubbel helix. Rondom de beschadiging worden vervolgens de twee strengen van het DNA uit elkaar gehaald door een combinatie van eiwitten, waar XPB en XPD onderdeel van uitmaken.
Knippen van het DNA. Daardoor krijgen XPG en een eiwitpaar (met daarin het reparatie-eiwit ERCC1) de ruimte om elk aan een kant van de schade een knip te zetten in de beschadigde streng.
Schade verwijdering. Vervolgens wordt het deel met de schade uit het DNA gewipt. Het is nog niet precies bekend hoe dat laatste in z'n werk gaat.
Opvulling. Uiteindelijk wordt het achtergebleven gat gedicht met nieuwe letters door DNA-producerende enzymen.

De inleiding van dit proefschrift beschrijft een aantal verschillende soorten van DNA-reparatie, maar gaat dieper in op nucleotide excisie reparatie (NER). NER van de mens lijkt sterk op NER van insekten en zelfs van eenvoudige gisten. Het mechanisme van herstel is dus gedurende de evolutie weinig veranderd. Aan de hand van een bekend gen voor een reparatie eiwit in gist hebben we het vergelijkbare gen in de mens gevonden. Het bleek overeen te komen met het gen dat verantwoordelijk is voor de ziekte xeroderma pigmentosum type F (beschreven in hoofdstuk drie). Het eiwit (XPF genoemd) vormt een vast koppel met een ander reparatie-eiwit ERCC1. In hoofdstuk vijf wordt beschreven op welke manier ERCC1 en XPF aan elkaar vastzitten. Het ERCC1–XPF paar zet één van de knippen in de beschadigde DNA-streng tijdens het reparatieproces. Behalve hun rol bij het herstel van UV-schade, zijn ERCC1 en XPF ook betrokken bij het losmaken van kruisverbindingen tussen DNA-strengen. Het feit dat ERCC1 en XPF betrokken zijn bij het repareren van DNA-kruisverbindingen voorspelt dat ze ook een rol spelen in andere reparatieprocessen die schades op een andere manier uit het DNA verwijderen dan NER, bijvoorbeeld via re-combinatie van stukken DNA die erg op elkaar lijken (hoofdstuk zes). We hebben de delen van ERCC1 die belangrijk zijn voor het goed functioneren van het eiwit in kaart gebracht. Dit is beschreven in hoofdstuk twee. Het blijkt dat ERCC1 en XPF elkaar beschermen. In ongebonden toestand worden ze snel door de cel afgebroken. In cellen met een te kleine hoeveelheid (minder dan normaal) van het ERCC1-XPF paar wordt UV-schade heel langzaam verwijderd (NER), terwijl kruisverbindingen tussen DNA-strengen normaal gerepareerd worden (re-combinatie).

Om de relatie tussen de erfelijke veranderingen in het XPF gen (het 'genotype' van de patiënt) en het daaruit volgende ziektebeeld (het 'fenotype') te begrijpen, hebben we van een aantal patiënten de afwijkingen in de XPF genen bepaald. Ieder mens beschikt over twee kopieen van het XPF gen, één van de vader en één van de moeder. Bij een van de onderzochte patiënten mist de ene kopie vier letters. Hierdoor wordt een korter dan normaal XPF eiwit geproduceerd. De andere bevat een één-letter verandering die waarschijnlijk een andere opbouw van het eiwit tot ge-

volg heeft. Die mutatie werd ook gevonden in een andere Europese XP-F patiënt die wordt beschreven in hoofdstuk vier. Hoewel beide gemuteerde eiwitten van deze patiënt snel worden afgebroken, kan er toch een kleine hoeveelheid van het XPF eiwit in zijn cellen worden aangetoond. Ze hebben dan ook een klein deel van de oorspronkelijke reparatieactiviteit behouden. Dit verklaart het relatief milde ziektebeeld dat karakteristiek is voor alle XP-F patiënten die we tot nu toe kennen. In tegenstelling tot wat we verwachtten, hebben nog geen patiënt gevonden met een genetische verandering in het ERCC1 eiwit. Anders dan de milde verschijnselen veroorzaakt door een mutatie in XPF bij de mens, blijkt dat muizen met een ERCC1 mutatie een zeer ernstig fenotype hebben. Behalve een reparatiedefect en UV-gevoeligheid, hebben deze muizen lever- en nierproblemen en vertraagde groei. De muizen gaan binnen enkele weken na de geboorte dood. We hebben recent een uitzonderlijke XP-F patiënt gevonden, waarvan het ziektebeeld veel lijkt op dat van de ERCC1-mutante muis (hoofdstuk zes). Door de mutatie is er geen aantoonbare hoeveelheid ERCC1–XPF en nauwelijks nog reparatieactiviteit aanwezig in zijn cellen. We verwachten daarom dat nu ook de eventuele functie in het recombinatie herstel van kruisverbindingen is verstoord. Dit laatste zou de symptomen die niet karakteristiek zijn voor XP, zoals de nier- en leverproblemen, kunnen verklaren. Deze nieuwe 'reparatieziekte' heeft meer weg van voortijdige veroudering dan van XP en mogelijk helpt dit nieuwe ziektebeeld in de toekomst bij het vinden van ERCC1 patiënten.

List of publications

Nagai A, Saijo M, Kuraoka I, Matsuda T, Kodo N, Nakatsu Y, Mimaki T, Mino M, Biggerstaff M, Wood RD, Sijbers AM, Hoeijmakers JHJ, Tanaka K (1995) Enhancement of damage-specific DNA binding of XPA by interaction with the ERCC1 DNA repair protein. *Biochem Biophys Res Commun* 211: 960-966

Sijbers AM, van den Berg J, van der Spek PJ, Odijk H, van Duin M, Jaspers NGJ, Hoeijmakers JHJ, Bootsma D (1995) Mutational analysis of the human *ERCC1* gene. *J Cell Biochem* supplement 21A: 284

Sijbers AM, van der Spek PJ, Odijk H, van den Berg J, van Duin M, Westerveld A, Jaspers NGJ, Bootsma D, Hoeijmakers JHJ (1996) Mutational analysis of the human nucleotide excision repair gene *ERCC1*. *Nucleic Acids Res* 24: 3370-3380

Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC, Rademakers S, de Rooij J, Jaspers NGJ, Hoeijmakers JHJ, and Wood RD (1996) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* 86: 811-822

Vermeulen W, de Boer J, Citterio E, van Gool AJ, van der Horst GTJ, Jaspers NGJ, de Laat WL, Sijbers AM, van der Spek PJ, Sugasawa K, Weeda G, Winkler GS, Bootsma D, Egly JM, Hoeijmakers JHJ (1997) Mammalian nucleotide excision repair and syndromes. *Biochem Soc Trans* 25: 309-315

Hoeijmakers JHJ, van der Horst GTJ, Weeda G, Vermeulen W, Winkler GS, de Boer J, de Laat WL, Sijbers AM, Citterio E, Jaspers NGJ, Egly J-M, Bootsma D (1997) Recombining DNA damage repair, basal transcription and human syndromes. In *Genomic instability and immortality in cancer*. Mihich E, Hartwell L (eds) Plenum Publishers New York: 83-96

Xu H, Swoboda I, Bhalla PL, Sijbers AM, Zhao C, Ong E-K, Hoeijmakers JHJ, Singh MB (1998) Plant homologue of human excision repair gene *ERCC1* points to conservation of DNA repair mechanisms. *Plant J* 13: 823-829

Sijbers AM, van Voorst Vader PC, Snoek JW, Raams A, Jaspers NGJ, Kleijer WJ (1998) Xeroderma pigmentosum complementation group F: a European patient with adult-onset of neurologic disease. *J Invest Derm* 110: 832-836

Yagi T, Matsumura Y, Sato M, Nishigori C, Mori T, Sijbers AM, Takebe H (1998) Complete restoration of normal DNA repair characteristics in group F xeroderma pigmentosum cells by over-expression of transfected *XPF* cDNA. *Carcinogenesis* 19: 55-60

de Laat WL, Sijbers AM, Odijk H, Jaspers NGJ, Hoeijmakers JHJ (1998) Mapping of interaction domains between human repair proteins ERCC1 and XPF. *Nucleic Acids Res* 26: 4146-4152

Curriculum Vitae

Anneke Sijbers werd op 3 oktober 1968 geboren in Venray. In 1987 behaalde zij het VWO diploma aan het Rijksatheneum te Helmond. Aansluitend studeerde zij Moleculaire Wetenschappen aan de Landbouw universiteit Wageningen. Voor de afstudeerrichting chemisch-biologisch liep ze onderzoekstages bij de afdeling Biochemie (Dr I.M.C.M Rietjens) en afdeling Celbiologie en Genetica van de Erasmus Universiteit Rotterdam (Prof. Dr J.H.J. Hoeijmakers). In maart 1993 behaalde ze het doctoraal diploma. In april 1993 begon ze met een promotieonderzoek bij diezelfde afdeling Celbiologie en Genetica van de Erasmus Universiteit Rotterdam. Hier voerde zij onder begeleiding van Prof. Dr J.H.J. Hoeijmakers en Dr N.G.J. Jaspers het in dit proefschrift beschreven onderzoek uit. Van februari tot augustus 1997 bracht zij een werkbezoek aan de groep van Prof. Dr F. Hanaoka in het Institute of Molecular and Cellular Biology van Osaka University in Japan. Zij is sinds maart 1998 werkzaam als postdoc op de Target Discovery afdeling van N.V. Organon Oss.

Dankwoord

Het schrijven van de laatste woorden van mijn boekje geeft mij de gelegenheid om iedereen te bedanken die heeft bijgedragen aan de totstandkoming van dit proefschrift.
Allereerst wil ik mijn promotoren Prof. Dr Hoeijmakers en Prof. Dr Bootsma bedanken. Tijdens mijn onderzoeksstage in Rotterdam werd ik pas echt enthousiast voor het onderzoek. Ik vond het dan ook erg leuk dat ik kon blijven om te beginnen aan het promotieonderzoek dat uiteindelijk heeft geleid tot dit proefschrift. Jan bedankt voor je onophoudelijke enthousiasme en dat je altijd voor me klaar stond.
Ook wil ik mijn co-promotor Koos Jaspers bedanken. Na al die jaren heb je me volgens mij nog niet alles kunnen vertellen van wat jij allemaal weet. Jouw brede belangstelling heb ik altijd zeer gewaardeerd. Wouter wil ik bedanken voor de prikkelende vragen en leerzame discussies. Onze samenwerking was geknipt voor het vinden en karakteriseren van ons complex. Hanny, bedankt voor al die mooie plaatjes die jij bij elkaar gepipetteerd hebt. Ik vond het erg leuk om met je samen te werken. Esther en Anja, hartelijk bedankt voor de lol en relativerende woorden. Het doet mij een groot plezier dat jullie mij op de dag van mijn promotie als paranimfen terzijde willen staan.
Johan en Mariska wil ik bedanken voor hun enthousiasme en waardevolle bijdrage aan dit werk tijdens hun stage. Succes met jullie eigen promotieonderzoek. Suzanne, bedankt voor de micro-injecties. Verder wil ik alle andere collega's van het lab –Alain, André, Annelies, Bas, Berna, Bert, Betty, Cécile, Christine, Coen, Dick, Geert, Hanneke, Henk, Ingrid, Jan (dB), Jan (dW), Jan (vK), Jan-Huib, Jeroen, Joke, Kaoru, Manja, Marianne, Michael, Mies, Orlando, Ria, Roland, Rosalinde, Sigrid en Wim– en daarbuiten –Rita, Mirko, Melle, Rein, Mieke, Jan Jos, Ton, Piet en de dames en heer van de keuken– bedanken voor de helpende hand en een gezellige tijd. Peer jij wist hoe graag ik naar Japan wilde. Leuk dat je nu weer mijn collega bent.
Ik wil ook graag diegenen bedanken die belangrijk zijn geweest voor de klinische kant van het verhaal. Wim Kleijer (*Instituut Klinische Genetica, Erasmus Universiteit Rotterdam*), Dr Snoek (*Neurologie, Martini Ziekenhuis, Groningen*), Dr van Voorst Vader (*Dermatologie, Academisch Ziekenhuis, Groningen*) en Dr Meinecke (*Altonaer KinderKrankenhaus, Hamburg, Germany*).
I would like to thank Dr Wood and Rafael Ariza (*Imperial Cancer Research Fund, South Mimms, United Kingdom*), Dr Singh and Huiling Xu (*Department of Agriculture and Resource Management, University of Melbourne, Australia*), Dr Yagi (*Department of Radiation Genetics, Kyoto University, Japan*), and Dr Yasui and Ms Hayashi (*Department of Molecular Genetics, Tohoku University, Japan*) for our pleasant collaboration.

I am grateful to Prof. Dr Hanaoka for giving me the opportunity to work in his laboratory (*Institute of Molecular and Cellular Biology, Osaka University, Japan*) and to visit the very beautiful places of Japan. I would like to thank all members in his lab, especially Dr Yokoi and Dr Masutani. I am also indepted to Prof. Dr Tanaka who gave me the opportunity to meet with xeroderma pigmentosum patients.
Niet te vergeten mijn nieuwe collega's van de Target Discovery afdeling (N.V. Organon Oss), hen wil ik bedanken voor hun belangstelling en begrip. Dat geldt ook voor mijn vrienden die het vaak zonder mij moesten doen.
Lieve pap en mam, bedankt voor jullie onvoorwaardelijke steun en stimulans de afgelopen jaren en last but not least Paul, bedankt voor alles en dat is heel veel.