Incision Coordination in Nucleotide Excision Repair

Coördinatie van incisies tijdens nucleotide excisie herstel

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

vrijdag 12 juni 1998 om 15:45 uur

door

Wouter Leonard de Laat

geboren te Amersfoort

Promotiecommissie

Promotoren:	Prof. dr J.H.J. Hoeijmakers
	Prof. dr D. Bootsma
Overige Leden:	Prof. dr P.C. van der Vliet
	Prof. dr J.A. Grootegoed
	Prof. dr F.G. Grosveld
Co-promotor:	Dr N.G.J. Jaspers

Omslag: Uitzicht van de schrijver bij het vallen van de avond
View from the author's window at dusk

Dit proefschrift kwam tot stand binnen de vakgroep Celbiologie en Genetica van de Faculteit der Geneeskunde en Gezondheidswetenschappen van de Erasmus Universiteit Rotterdam. De vakgroep maakt deel uit van het Medisch Genetisch Centrum Zuid-West Nederland. Het onderzoek is financieel ondersteund door de Stichting Scheikundig Onderzoek in Nederland (SON).

Voor mijn ouders

Voor Evalyne

Contents

Scope of the thesis

This thesis aims to contribute to the understanding of the molecular mechanism that underlies one of the main DNA repair pathways in mammals, nucleotide excision repair. In **chapter 1** the relevance of DNA repair in general is outlined. An overview of mammalian strategies to counteract DNA damage is provided, to show that an intricate network of repair machineries permanently guards the integrity of the genome. In discussing each repair pathway, attention is focussed on how DNA damage is removed and what protein factors are required to accomplish this. Chapter 1 serves as a framework for **chapter 2**, in which one repair pathway, mammalian nucleotide excision repair, is discussed more extensively. In this chapter, a comprehensive overview of the characteristics of each protein factor involved in nucleotide excision repair is given. Our own observations will be included in this overview, which will serve as a fundament to build a model that addresses the molecular steps that lead to DNA incisions around the damage, a key event in nucleotide excision repair. Additionally, a model will be presented for transcription-coupled repair, a phenomenon that accounts for the fast removal of DNA damage from actively transcribed template strands. The contributions of this thesis to the repair field are summarized in **chapters 3**, **4**, **5**, and **6**. For implications of these data, other than those discussed in chapters 1 and 2, the reader is referred to the discussion sections that conclude each of the latter chapters.

Chapter 1

DNA repair mechanisms in mammals

1.1 Introduction

DNA (deoxyribonucleic acid) is the carrier of genetic information. The recepies for all chemical processes in an organism are encoded by the sequence of the four bases (G, A, C and T) that constitute the DNA strands. In the cell, these strands are intertwined to form the typical double helix of DNA. The base sequence is transfered faithfully from generation to generation. However, chemical and physical reactions constantly threaten the integrity of DNA. Intrinsic instability of DNA bonds, errors made by DNA-metabolizing processes, and reactive compounds that are formed during cellular metabolism can all cause 'spontaneous' DNA decay. On top of these intracellular threats, environmental factors such as ionizing radiation, UV radiation and mutagenic molecules present for example in food, medicines, cigarette smoke, car exhaust fumes and industrial pollution, can damage the DNA.

It is of vital importance to cells to counteract DNA damage. Damaged bases that are misinterpreted during DNA replication can result in the incorporation of an incorrect nucleotide leading to a mutation. Mutations in proto-oncogenes or tumor suppressor genes are generally considered to underlie the multistep progression from a normal cell

to a tumor cell. Also, accumulation of DNA damage is assumed to contribute to aging. DNA lesions can block DNA-metabolizing processes such as transcription and replication, which even can lead to cell death. One cross-link between DNA strands is sufficient to kill a yeast cell (45).

Various strategies are used to protect the genetic information against damaging sources. (i) In higher eukaryotes, epithelial tissues (the skin!) form a first natural barrier against environmental factors. (ii) Also in such organisms, organs like the kidney and liver remove toxic compounds from the blood. (iii) A lipid membrane selectively allows molecules to enter the cell. (iv) Scavengers in the cell neutralize reactive compounds. (v) Eukaryotes have directed their genomic DNA to a separate compartment, the nucleus, presumably to minimize exposure to oxidative radicals produced during mitochondrial metabolism. (vi) In the nucleus, DNA is densily packed with proteins, which minimize the accessibility of DNA. On top of these protection strategies, cells are equipped with multiple mechanisms to counteract the deleterious effects of damage that has been induced in the DNA. (i) Upon exposure to damaging agents, checkpoint mechanisms can arrest the cell cycle before replication or division, presumably to allow time to repair. (ii) Various DNA repair pathways exist, each tracing and removing specific types of DNA lesions. (iii) If lesions accumulate above certain thresholds, cells stop dividing and go into apoptosis (programmed cell death). This thesis focusses on DNA repair mechanisms, and particularly on nucleotide excision repair, which will extensively be discussed in chapter 2. In this first chapter, other repair pathways in mammals will briefly be addressed.

1.2 DNA repair mechanisms

An enormous variety in types of DNA lesions exists, ranging from large bulky adducts to DNA cross-links and broken DNA. Nevertheless, a few repair mechanisms can remove the vast majority of these lesions (Fig. 1). *Base excision repair* (BER)

corrects many types of base alterations and is the main pathway for the removal of oxidative damage. Damage removal by BER merely encompasses the replacement of a single nucleotide. *Nucleotide excision repair* (NER) removes lesions that more severely distort the DNA helix, including intrastrand cross-links and bulky adducts. The lesion spectrums of BER and NER partially overlap. NER is the main pathway for repair of UV-induced lesions. A hallmark of NER is that damage is excised as part of a ~30 nt DNA fragment. *Mismatch repair* (MMR) corrects base mispairs and small insertions/

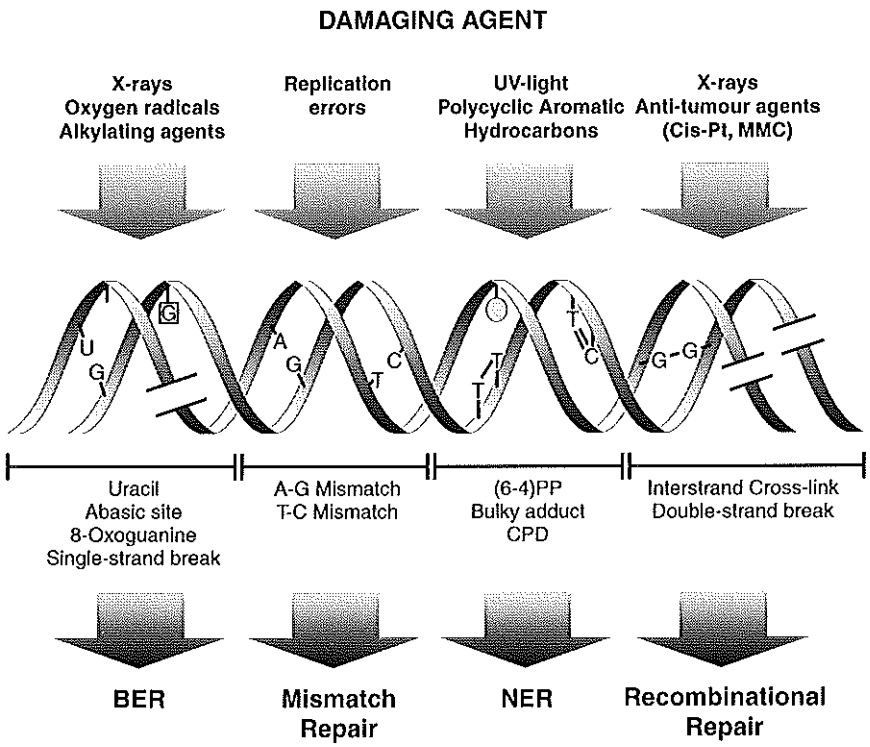


Figure 1. DNA repair mechanisms. Top: common DNA damaging agents. Middle: examples of lesions that can be introduced by these agents into the double helix of DNA. Bottom: the most frequently used repair mechanisms for such lesions. Not depicted but important to realize is that distinct damaging sources can induce similar types of DNA lesions, and that also the lesion spectrums of different repair pathways may overlap.

deletions introduced by the replication machinery. A key event in this process is the discrimination between the template strand and the newly synthesized strand, with the latter by definition containing the wrong sequence. Removal occurs by single-stranded degradation of a (large) region around the mispair. Both BER, NER, and MMR use the 'undamaged' complementary strand of the DNA double helix as a template for DNA synthesis to restore the correct sequence.

Homologous recombination, *DNA-end joining* and *single-strand annealing* are all pathways implicated in the repair of the very noxious double-strand breaks. Homologous recombination involves a search for intact complementary DNA sequences elsewhere in the genome and is in principle error-free. The DNA-end joining pathway merely fuses broken DNA ends and is not necessarily error-free. Single-strand annealing uses regions of homology in the broken DNA fragments to reanneal their ends. Often this coincides with loss of non-homologous sequences (error-prone). Recombinational events are also thought to be required for the repair of interstrand cross-links, which may involve enzymes from various double-strand break repair pathways. The pathways summarized above have in common that multiple enzymes are involved, which rearrange the DNA around the lesion such that replication proteins can restore the integrity of the DNA helix. However, also some monoenzymatic repair processes exist that directly reverse specific DNA adducts.

It should be noted that *in vivo* all DNA repair pathways are faced with the challenge to trace randomly scattered rare lesions in large and highly compacted genomes. 6 x 10^9 base pairs, constituting a linear distance of ~2 meters, are packed into a human cell nucleus with a typical diameter of only 10 microns. On top of this enormous amount of DNA, the nucleus is expected to contain a protein concentration of approximately 250 mg/ml, nearly as high as in many protein crystals (27, 31). Nevertheless, repair in general is very efficient, with most damage being removed within hours. Repair enzymes can be anticipated to function in a highly organized and dynamic fashion in the context of chromatin. As experimental data on this topic are still lacking, this additional level of complexity will largely be ignored in this overview.

Below, the main characteristics of the eukaryotic repair pathways will be outlined, focussing on the molecular mechanisms and discussing briefly the phenotypical consequences of unremoved DNA damage.

1.2.1 **Base excision repair**

Lesion specificity

Base excision repair (BER) corrects base alterations induced by endogenous oxidative events, ionizing radiation and simple alkylating agents. There may well be more than 100 different types of oxidative lesions alone in mammalian DNA, illustrating the tremendous diversity of lesions BER is coping with. Structurally, these lesions vary from deaminated bases and methylated base atoms to abasic sites, which are all alterations causing minor distortion to the DNA helix. Examples of frequently induced and often highly mutagenic BER-lesions are 8-oxoguanine, O^6-methylguanine, deaminated methylated cytosine (uracil!) and thymine glycol.

Molecular mechanism

The complete BER reaction merely involves the replacement of a single nucleotide, and several enzymatic activities are required to accomplish this (Fig. 2). A key event is the release of the damaged base from the deoxyribose-phosphate chain, carried out by a class of enzymes called DNA glycosylases, that each act on a specific subset of damaged bases. The resulting apurinic/apyrimidinic (AP) site is incised at the 5' side by an AP endonuclease, called HAP1 or APE, which opens the deoxyribose-phosphate backbone and creates a (normal) 3' OH group and an (abnormal) 5' deoxyribose residue. The latter needs to be removed in order to complete repair, which proceeds through one of two different pathways. (i) The 5' deoxyribose residue is removed by DNA polymerase β, which harbors an intrinsic AP lyase activity required for this. Incorporation of a single nucleotide by pol β is followed by ligation, involving XRCC1 and DNA ligase III

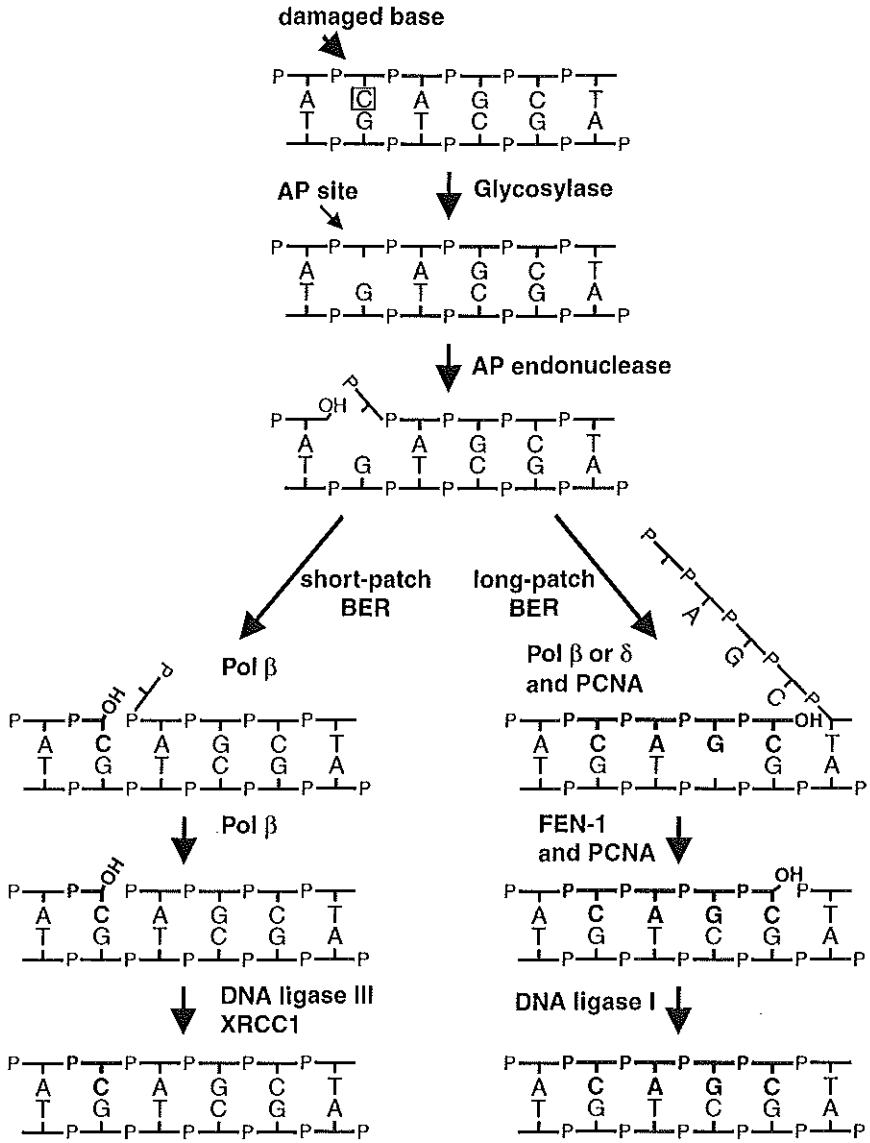


Figure 2. Model for the mechanism of base excision repair. (Adapted from (40)).

or ligase I. This pathway is called short-patch BER, as only one nucleotide is incorporated. (ii) Alternatively, DNA polymerase β or δ incorporates 2-6 nucleotides while replacing the original strand, which is then cleaved off by the structure-specific nuclease FEN1 to allow ligation. PCNA stimulates the latter subpathway, which is referred to as the long-patch BER pathway (40).

Consequences of defective BER

Disruption of genes encoding *XRCC1*, the BER-specific DNA polymerase β, ligase I or *HAP1* in the mouse causes embryonic lethality, reflecting the apparent absence of BER-deficient human syndromes. An alkyladenine DNA glycosylase- (AAG-) deficient mouse, however, developed normally and appeared healthy. Sensitivity of the knockout cells for various alkylating agents revealed a BER-defect (21). Thus, only disruption of one of the many BER-initiating glycosylases, but not a total BER deficiency, is compatible with life. For reviews and original references, see (6, 16, 41, 53, 84).

1.2.2 **Mismatch repair**

Lesion specificity

DNA mismatch repair can recognize and repair all 8 possible base-base mismatches as well as small (1-5 nt) insertion/deletion mispairs that are left behind by DNA polymerases (48). The error-rate of the three mammalian DNA polymerases α, δ and ε in replication of nuclear DNA varies between 10^{-5}-10^{-6}, dependent on proofreading ability (81). Mismatch repair is thought to reduce the amount of errors by 2 or 3 orders of magnitude. Similarly, it is thought to remove (small) non-homologous regions from recombining DNA. Recently, mismatch repair deficiency was found to confer tolerance to certain types of DNA damaging agents such as *N*-Methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) and the chemotherapeutic agent cisplatin. Mismatch repair proteins recognize lesions induced by these agents, and may act as a sensor of damage, required to induce a

G2 checkpoint response or, at high lesion load, cell death (48).

Molecular mechanism

After identification of a mismatch, discrimination between the parental strand and daughter strand is required, with the latter by definition containing the incorrect nucleotide of a mispair. In *E. coli*, this occurs through sensing the methylation status of both strands in the vicinity of the mismatch. An incision, placed in the non-methylated daughter strand, is considered the strand signal that marks the initiation site for exonucleolytic single-stranded DNA degradation (Fig. 3). This strand signal may reside up to 1000 base pairs away on either side of the mispair, necessitating the activity of either a 3' to 5', or a 5' to 3' exonuclease to accomplish DNA degradation across the mispaired site. Excision is followed by DNA repair synthesis. In mammals, mismatch repair is likely to occur in a similar fashion, since mismatch repair genes have been conserved from bacteria to man. However, the nature of the strand signal has not been defined in any eukaryotic organism. The protein complexes MutSα (a heterodimer of Msh2 and Msh6), MutSβ (a heterodimer of Msh2 and Msh3) and MutLα (a heterodimer of Mlh1

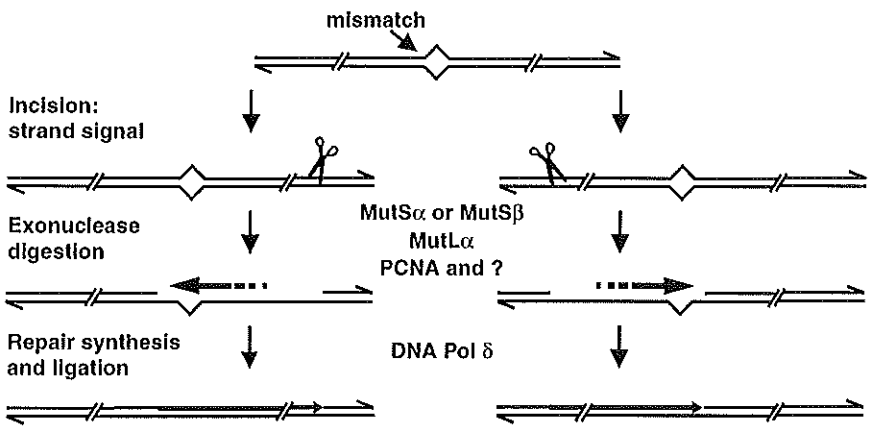


Figure 3. Model for the mechanism of mismatch repair. (Adapted from (48)).

and Pms2) have all been implicated in early events of mammalian mismatch repair, with MutSα binding both base-base and insertion/deletion mispairs and MutSβ binding only the latter category of mispairs, with a preference for larger insertion/deletion mispairs. DNA polymerase δ has been implicated in mismatch repair DNA synthesis, together with the DNA polymerase processivity factor PCNA, which might have an additional engagement in, or prior to, the excision step. Two oppositely directed exonuclease activities have been identified in *E. coli* mismatch repair, whereas at present only one eukaryotic exonuclease has been described. This exonuclease, called Exo1, was recently identified in *S. cerevisiae*, and interacts with Msh2. For reviews on mismatch repair and original references, see (23, 25, 48, 49, 81).

Consequences of defective mmr

Strong support for the idea that elevated mutation rates can cause carcinogenesis comes from the finding that at least 90% of the human hereditary nonpolyposis colon cancers (HNPCC) are mismatch repair-deficient. These tumors are characterized by microsatellite instability (i.e. alterations in short repeated sequences, which presumably are very susceptible to deletion/insertion mispairs) and notably carry mutations in the genes encoding *MLH1*, *MSH2* or *PMS2*. *PMS2*- and *MSH2*-deficient mice also show microsatellite instability and have an increased risk of developing cancer (2, 18, 59).

1.2.3 Alternative pathways for single-strand damage repair

UV damage endonuclease (UVDE)

Both the filamentous fungus *Neurospora crassa* and the fission yeast *S. pombe* appear to use an alternative excision repair pathway for the repair of CPDs and (6-4)PPs, which in *N. crassa* substitutes, and in *S. pombe* functions next to nucleotide excision repair (87, 88). This pathway involves an enzyme called UVDE (UV damage endonuclease), which cleaves immediately 5' to UV-induced dimers. Completion of this

repair pathway can proceed through two separate mechanisms, one of which depends on the structure-specific DNA endonuclease FEN-1. To date, no UVDE-like enzyme has been identified in man.

Monoenzymatic repair pathways

Apart from the classic excision repair pathways, some monoenzymatic pathways exist that directly reverse DNA adducts. These repair enzymes generally act on a small spectrum of frequently occurring lesions. Two examples will be given below.

O^6-methylguanine-DNA methyltransferase (MGMT)

The enzyme MGMT is conserved throughout the animal kingdom, and its main function is to protect DNA against the toxic effects of simple alkylating agents. MGMT acts by directly transfering methyl groups from O^6-methylguanine to an internal cysteine residue, thereby inactivating itself (reviewed by (26, 86)).

Photolyases

Photolyases are found in species ranging from bacteria to plants and aplacental mammals. They act specifically on either cyclobutane pyrimidine dimers (CPDs) or (6-4) photoproducts (PP), the two main UV-induced lesions, and utilize the energy of visible light to directly reduce these base-dimers to the original monomeric form (reviewed by (26, 61)). Two human genes with some sequence homology to (6-4) photolyases were recently identified, but as far as investigated no photoreactivating activity could be detected (76, 77) (André Eker, unpublished observation).

1.2.4 **Double-strand break repair**

Ionizing radiation (IR) and endogenously produced oxidative radicals can both induce double-strand DNA breaks (DSBs), which, together with interstrand DNA cross-

links, are the most disruptive types of DNA damage. At least three pathways can repair broken DNA ends, namely homologous recombination, non-homologous (or illegitimate) DNA-end joining and homology-dependent single-strand annealing. Below, first the mechanistic details of each pathway will briefly be summarized. Then, the relative contribution of these routes to DSB-repair will be discussed.

1.2.5 **Homologous recombination**

Eukaryotic homologous recombination largely depends on the *RAD52* epistasis group of genes, described in *S. cerevisiae*, which includes the genes *RAD50*, *RAD51*, *RAD52*, *RAD54*, *RAD55*, *RAD57*, *ScRAD59*, *MRE11* and *XRS2*. Mutations in these genes cause profound recombination defects and sensitivity to DSB-inducing agents in yeast. DSB-repair via homologous recombination comprises nucleolytic processing of the DNA ends to create single-stranded regions that can be transfered to homologous sequences, present either on homologous chromosomes or on sister chromatids (Fig. 4). Formation of such a joint molecule allows the damaged strands to regain genetic information, as the intact strands can be used as templates for DNA repair synthesis. Branch migration and resolution completes homologous recombination (73).

In *E. coli*, the key enzymes have been characterized, and rapid progress is being made in understanding the action of the eukaryotic *RAD52* epistasis group proteins (see Fig. 4). Both yeast and human Rad51, similar to their RecA homolog in *E. coli*, can form nucleoprotein filaments on single-stranded DNA regions to catalyze strand transfer (56, 70). Rad55 and Rad57, two additional homologs of RecA, form a heterodimer that stimulates Rad51-dependent strand exchange in a reaction that also involves the single-stranded DNA-binding protein RPA (5, 71, 72). Rad52 also interacts with Rad51 to stimulate strand exchange (7, 55, 65, 71). Rad54 has been reported to interact with Rad51 as well (11, 28, 36), but its function is not known yet. As a member of the SNF2/SWI2 DNA-dependent ATPase family, it may be involved in chromatin

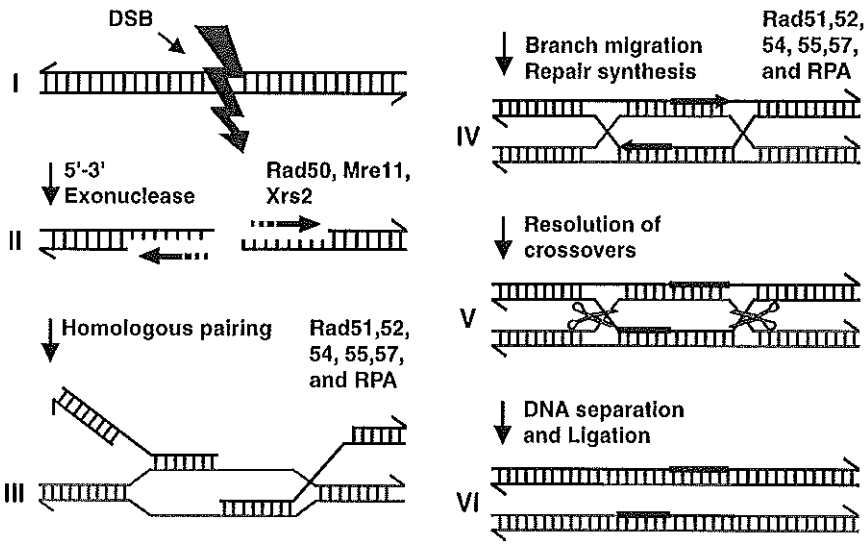


Figure 4. Model for the mechanism of double-strand break (DSB) repair by homologous recombination.

remodelling or act as a strand-exchange processivity factor (38). Although Rad50, Mre11 and Xrs2 belong to the *RAD52* epistasis group and may play a role in the early steps of homologous recombination during meiosis, their main involvement seems to be in another repair pathway, non-homologous DNA-end joining (47, 79, 80).

1.2.6 **Non-homologous DNA-end joining**

Non-homologous DNA-end joining is also known as the process that accounts for V(D)J recombination, which contributes to the genetic variability of immunoglobulins and T-cell receptors required for an adequate immune response in higher eukaryotes. Both to counteract IR-induced breaks and to combine V(ariable), D(iversity) and J(oining) regions, non-homologous DNA-end joining merely fuses juxtaposed DNA ends, using no or very limited sequence homology. For repair of DSBs, this is thought to

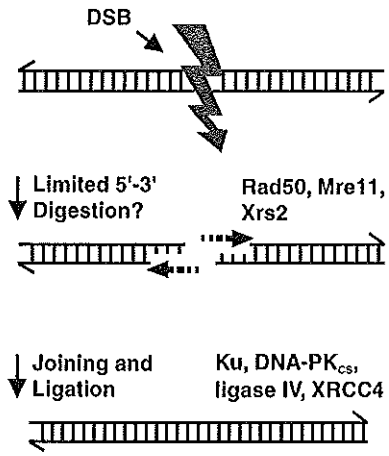


Figure 5. Model for the mechanism of double-strand break (DSB) repair by DNA-end joining.

imply (limited) processing of the DNA-ends, followed by joining and re-ligation (Fig. 5). Importantly, repair by non-homologous DNA-end joining is not an error-free mechanism. The initial step is assumed to require the activities of Rad50, Mre11 and Xrs2. In yeast, ScRad50 and ScMre11 form a complex (20, 37, 46), which, perhaps together with Xrs2 (50, 79, 80), is required for 5' resection of DSBs through nucleolytic DNA-end processing (see Fig. 5) (54). The heterodimer Ku, consisting of the subunits Ku80 and 70, binds to DNA-ends and recruits a DNA-dependent protein kinase called DNA-PK catalytic subunit, or DNA-PK_{CS}. Together, this heterotrimeric complex is referred to as DNA-PK. The function of the activated protein kinase activity of DNA-PK is still unclear. The plethora of proteins that were found to be phosphorylated by DNA-PK *in vitro*, may reflect the initiation of a signal transduction cascade that alerts the cell to the presence of DNA breaks (35). Important clues for the final steps in DNA-end joining were recently obtained by the finding that DNA ligase IV mediates non-homologous DNA-end joining. DNA ligase IV forms a complex with XRCC4, which stimulates its ligase activity several-fold. XRCC4 is also an effective substrate for phosphorylation by DNA-PK *in vitro* (15, 85, 29, 43), but the significance of this notion is unknown. Besides ligase IV, ligase I and III may also be involved (for reviews, see (10, 35, 78)).

1.2.7 **Single-strand annealing**

Single-strand annealing (SSA) is a rather poorly characterized DSB rejoining pathway, with knowledge mainly based on genetic studies in *S. cerevisiae*. It involves enzymes implicated in various other repair pathways, including the NER factor Rad1-Rad10, which is the *S. cerevisiae* homolog of ERCC1-XPF, the major topic of this thesis. Therefore, a relatively detailed description of this pathway will be presented.

SSA utilizes sequence homology on both sides of a DSB to accomplish DNA joining, and is therefore sometimes referred to as a second homologous recombination pathway. However, SSA is error-prone, as intervening non-homologous sequences are deleted (see Fig. 6). In the SSA mechanism, 5'-3' degradation occurs at the ends of a break and continues until homologous sequences become single-stranded and available for base-pairing. The complementary strands anneal, leaving the non-homologous DNA sequences that originally resided between the break and the repeat sequences, as 3'-protruding single-stranded arms. These 3'-arms are cleaved off to allow DNA repair synthesis and ligation of the recombined products (Fig. 6) (24).

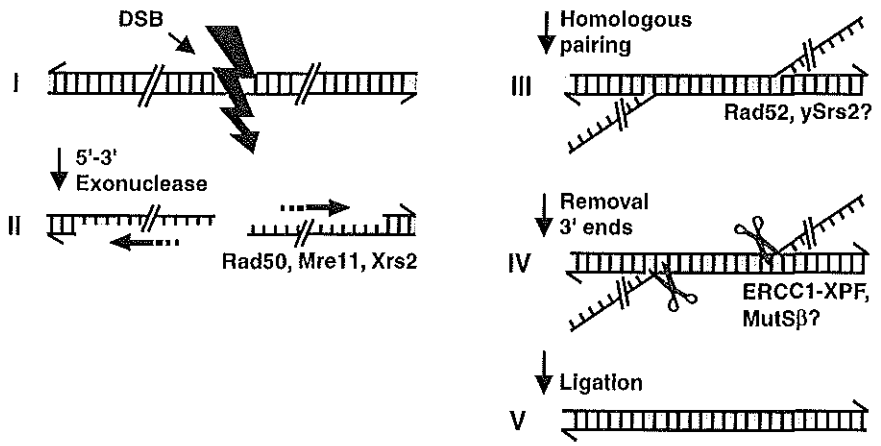


Figure 6. Model for the mechanism of double-strand break (DSB) repair by the single-strand annealing pathway.

To identify factors required for this pathway, mutant *S. cerevisiae* strains were tested for their ability to rearrange an enzymatically induced DSB located between directly repeated sequences of a marker gene. In wild-type cells, these breaks are repaired through two distinct pathways: SSA, which deletes the intervening non-homologous sequences, and homologous recombination, resulting in gene conversion without loss of genetic information (24). Genes from distinct epistasis groups are required for SSA, with *RAD52*, *RAD50* and *XRS2* representing the *RAD52* epistasis group, *RAD1* and *RAD10* representing the *RAD3* epistasis group (nucleotide excision repair, see below) and *MSH2* and *MSH3*, both implicated in mismatch repair. Notably, none of the other tested representatives of these repair pathways play a significant role (24, 33, 34, 44, 62, 69). The activities of the individual gene products involved allow some speculation on the molecular mechanism of SSA. As in non-homologous DNA-end joining, Rad50, Mre11 and Xrs2 may be required for 5' resection of DSBs (54), resulting in 3'-protruding single-stranded arms. Rad52 binds DNA and promotes annealing of complementary DNA strands (52), an activity relevant both for homologous recombination and SSA. Rad1 and Rad10 are present in the cell as a heterodimeric complex (1, 3, 4, 66). Their activity was found to be required for annealing of breaks with non-homologous DNA-ends only; a DSB flanked by completely homologous DNA sequences is fused independent of Rad1-Rad10 (24). This observation led to the idea that Rad1-Rad10 is required for the removal of 3'-protruding non-homologous single-stranded DNA-arms from recombined duplex DNA. Strong support for such a function came from the finding that the Rad1-Rad10 complex is a structure-specific DNA endonuclease that can remove 3'-protruding arms from a variety of artificial DNA substrates, including the proposed recombination intermediate (4). Interestingly, MutSβ (Msh2-Msh3) requirement is also restricted to annealing of breaks with non-homologous ends (62, 69). Since this complex recognizes mispaired bases and small heterologous loops in MMR (30, 39, 57), it may also recognize branched DNA structures with a free 3' tail and, upon binding, facilitate removal of non-homologous DNA ends by Rad1-Rad10 (69). A recent report also implicates a 5'-3' helicase, Srs2, in the SSA reaction, and suggests that this protein

stabilizes and extends the joint between the invading single strand and its complement (see Fig. 6) (58).

1.2.8 Contributions of individual pathways to DSB-repair

Cells are equipped with multiple strategies to join broken DNA-ends. How does each pathway contribute to the repair of exogenously induced DSBs in eukaryotes? Possibly, the pathway preferred varies from species to species. *RAD52*-dependent homologous recombination appears to be the dominant mechanism in *S. cerevisiae*, but in mammals DNA-PK-dependent DNA-end joining is thought to be preferred for DSB repair. Indeed, mammalian mutant cells defective in this pathway show a 2-6 fold increase in sensitivity to ionizing radiation (89), whereas recently generated Rad54-deficient mice do not demonstrate increased radiosensitivity (Jeroen Essers, personal communication). However, both Rad54-deficient ES cells and MEFs do show moderate radiosensitivity, demonstrating that Rad54 plays a significant role in mammalian DSB-repair at least during embryogenesis (22). Furthermore, a recently identified Rad54 homolog in mammals indicates that redundancy may mask the relevance of homologous recombination for DSB-repair. The contribution of each pathway to the removal of DSBs might vary during the cell cycle, since homologous recombination is likely to be most efficient during S phase, when sister chromatids are in close proximity.

It is difficult to assess the relevance of SSA to the repair of DSBs, partly because some of the genes involved participate in multiple DSB-rejoining pathways. Rad1 and Rad10, however, seem to be required for DSB rejoining via SSA only (34). Mutant strains of *S. cerevisiae* defective in either of these genes do not show increased IR-sensitivity, indicating that in this species SSA does not significantly contribute to the repair of IR-induced DSBs. In *Schizosaccharomyces pombe*, the Swi9-Swi10 complex, homologous to Rad1-Rad10, is involved in the recombination that underlies mating-type switching. Besides, *Swi9* and *Swi10* mutant strains are the only NER-defective

strains in *S. pombe* that show cross-sensitivity to ionizing radiation (63, 64). Thus, SSA may contribute significantly to repair of IR-induced damage in *S. pombe*. In human cells, SSA-like events have been shown to occur by sequence analysis of the junctions at break sites in deletions arising after irradiation (51). Some regions of homology used for annealing were found to originate from Alu repeats (51). *ERCC1* and *XPF* are the mammalian homologs of *RAD10* and *RAD1*, respectively (67, 82), and like Rad1-Rad10, their gene products form a stable protein complex that acts as a structure-specific endonuclease (see chapters 3 and 5) (67). In agreement with its presumed function in SSA, ERCC1-XPF was found to remove 3'-protruding single-stranded arms from recombination intermediate-like DNA structures (see chapter 5) (17). Genetic data indicate that also these excision repair genes have an additional role in recombination. ERCC1- and XPF-deficient Chinese hamster mutant cells are extremely sensitive to DNA interstrand cross-links, and also show a slightly increased sensitivity to IR-induced damage (<2 fold) (reviewed by (75)). Although not as profound as observed with DNA-end joining-deficient rodent cells, this sensitivity may indicate that SSA plays a role in mammalian DSB-repair (for reviews, see (32, 35, 38, 60, 75)).

1.2.9 **DNA interstrand cross-link repair**

Interstrand cross-links

The devastating effect of unrepaired DNA interstrand cross-links (hereafter simply referred to as cross-links) is well documented in different species and probably reflects the insuperable obstacles these lesions form for DNA replication. In a cross-link repair-deficient mutant of *S. cerevisiae*, for example, 1 cross-link per cell is already lethal (45). In mammals, some Chinese hamster mutant cells are up to 100 times more sensitive to cross-linking agents than wild-type cells, which far exceeds sensitivities observed to other damaging agents. Mitomycin C (MMC) and cisplatin are two examples of anti-tumour agents that induce cross-links, and their therapeutical effect is based on such

cross-link-induced cell-killing.

Some of the mammalian hypersensitive cell lines were originally isolated on the basis of UV-sensitivity, notably the ERCC1- and XPF-deficient mutants, whereas others were identified as IR-sensitive cells (XRCC2- and XRCC3-deficient mutants), suggesting that both nucleotide excision repair and recombination are required for the removal of cross-links. Cell lines derived from patients with the highly cancer-prone disease Fanconi anaemia (FA), which can be subdivided into eight complementation groups (FA-A to -H), also display extreme sensitivity to cross-linking agents. However, whether the extreme sensitivity of FA cells for MMC reflects a direct or indirect engagement of *FA* gene products in cross-link repair, remains to be elucidated (reviewed by (19)).

Molecular mechanism

Cross-link repair in *E. coli* was found to be dependent on the activities of the excision repair proteins UvrA, -B and -C, the recombination protein RecA and DNA polymerase I (12, 13). Based on these protein requirements, a model was proposed for sequential excision and recombination during cross-link repair (see Fig. 7) (12), and the main principles of this model have now been confirmed by *in vitro* reconstituted reactions with purified factors. Briefly, the endonucleolytic UvrABC complex initiates repair by incising one strand at both sides of the cross-link. RecA-mediated transfer of a homologous DNA strand past the lesion enables retrieval of missing sequence information via a first round of DNA replication. A second pair of incisions made by UvrABC in the opposite strand releases a dsDNA fragment containing the cross-link, and the remaining gap is filled in by repair synthesis (68, 83).

Attempts to study this mechanism in mammalian cell extracts sofar only revealed that a dual incision, very similar to that observed during excision repair of intrastrand cross-links, might be the initial event in higher eukaryotes as well (42). Recently, *in vitro* incisions near a site-specific interstrand cross-link were found to be totally dependent on the NER machinery and could even be reconstituted with purified NER factors (8). Curiously, these incisions occurred both at the 5' side of the cross-link, releasing an

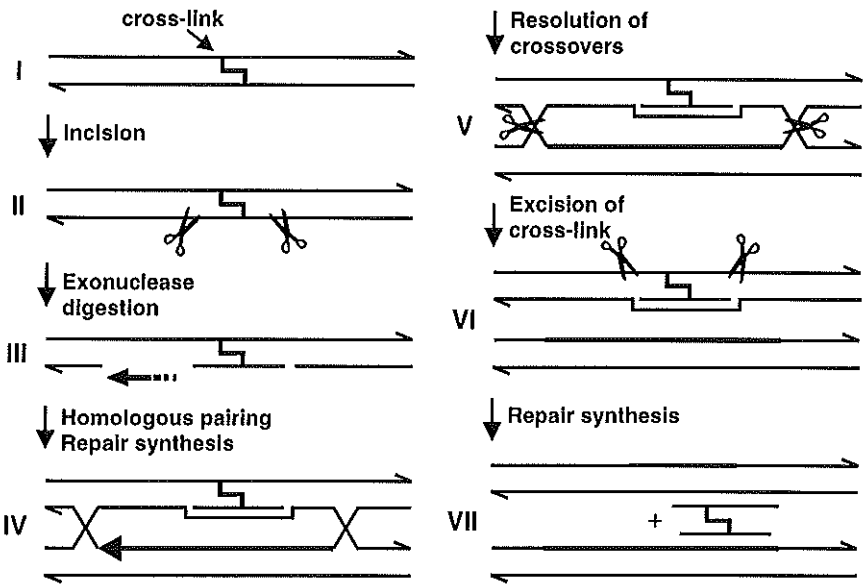


Figure 7. Model for the mechanism of DNA interstrand cross-link repair in *E. Coli*. Adapted from Cole (1973) (12) and Sladek *et al.* (1989) (68)).

undamaged DNA fragment of 22-28 nucleotides and creating a gap at one side of the cross-link (8). The relevance of this observation for the repair of interstrand cross-links is still rather obscure, but a consensus with the previous study on cross-link repair *in vitro* may be that the NER proteins initiate this type of repair by making a dual incision near the DNA lesion (8, 42). The complete NER machinery being involved in mammalian cross-link repair would mimic the situation in *E. coli*, and support for this notion can perhaps be derived from the fact that mutant rodent cell lines representing the NER-genes *XPB*, *XPD* and *XPG* show a moderate sensitivity to MMC and cisplatin (~ 2-5-fold hypersensitive to MMC) (14). However, it needs to be said that MMC, cisplatin and other cross-linking agents induce not only interstrand cross-links but also intrastrand cross-links and various types of DNA adducts (9). Thus, the moderate sensitivity of these typical NER-deficient mutants for such damaging agents may also reflect their inability to remove non-cross-link types of DNA lesions. In any case, these NER factors

are not absolutely required for cross-link repair *in vivo*, unlike ERCC1-XPF, which appears to be essential for this process. Additional insight into the mechanism of cross-link repair awaits *in vitro* assays that measure end-points which reflect the degree of sensitivity of the cell lines tested, as was previously accomplished for the excision repair pathways. Such assays should not only look at cleavage patterns but also take into account later, recombinational, stages to determine the relevance of incisions. The requirement of two Rad51-homologs, XRCC2 and XRCC3 (74, 75), indeed suggests strand transfer during cross-link repair. It will be interesting to see how the crucial function of ERCC1-XPF in cross-link repair relates to the role of this endonuclease complex in other recombinational events.

References

1. **Bailly, V., C.H. Sommers, P. Sung, L. Prakash, and S. Prakash.** 1992. Specific complex formation between proteins encoded by the yeast DNA repair and recombination genes *RAD1* and *RAD10*. *Proc. Natl. Acad. Sci.* **89:** 8273-8277.
2. **Baker, S.M., C.E. Bronner, L. Zhang, A.W. Plug, M. Robatzek, G. Warren, E.A. Elliott, J. Yu, T. Ashley, N. Arnheim, R.A. Flavell, and R.M. Liskay.** 1995. Male mice defective in the DNA mismatch repair gene *PMS2* exhibit abnormal chromosome synapsis in meiosis. *Cell* **82:** 309-319.
3. **Bardwell, A.J., L. Bardwell, D.K. Johnson, and E.C. Friedberg.** 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.
4. **Bardwell, A.J., L. Bardwell, A.E. Tomkinson, and E.C. Friedberg.** 1994. Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. *Science* **265:** 2082-2085.
5. **Baumann, P., and S.C. West.** 1997. The human rad51 protein: polarity of strand transfer and stimulation by hRP-A . *EMBO J* **16:** 5198-5206.
6. **Beckman, K., and B. Ames.** 1997. Oxidative decay of DNA. *J. Biol. Chem.* **272:** 19633-19636.
7. **Benson, F., P. Baumann, and S.C. West.** 1998. Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. *Nature* **391:** 401-404.
8. **Bessho, T., D. Mu, and A. Sancar.** 1997. 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.
9. **Caldecott, K., and P. Jeggo.** 1991. Cross-sensitivity of gamma-ray-sensitive hamster mutants to cross-linking agents. *Mutat. Res.* **255:** 111-121.
10. **Chu, G.** 1997. Double strand break repair. *J. Biol. Chem.* **272:** 24097-24100.
11. **Clever, B., H. Interthal, J. Schmuckli-Maurer, J. King, M. Sigrist, and W. D. Heyer.** 1997. Recombinational repair in yeast: functional interactions between Rad51 and Rad54 proteins. *EMBO J.* **16:** 2535-2544.
12. **Cole, R.** 1973. Repair of DNA containing interstrand crosslinks in Escherichia coli: sequential excision and recombination. *Proc. Natl. Acad. Sci.***70:** 1064-1068.
13. **Cole, R., and R. Sinden.** 1975. Repair of cross-linked DNA in Escherichia coli. *Basic Life Sci.* **5B:** 487-495.
14. **Collins, A.R.** 1993. Mutant rodent cell lines sensitive to ultraviolet light, ionizing radiation and cross-linking agents - A comprehensive survey of genetic and biochemical characteristics. *Mutat. Res.* **293:** 99-118.
15. **Critchlow, S., R. Bowater, and S. Jackson.** 1997. Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. *Curr. Biol.* **7:** 588-598.
16. **Cunningham, R.** 1996. DNA repair: how yeast repairs radical damage. *Curr. Biol.* **6:** 1230-1233.
17. **de Laat, W.L., E. Appeldoorn, N.G.J. Jaspers, and J.H.J. Hoeijmakers.** 1998. DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.* **273:** 7835-7842.
18. **de Wind, N., M. Dekker, A. Berns, M. Radman, and H. te Riele.** 1995. Inactivation of the mouse *MSH2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell* **82:** 321-330.

19. **Digweed, M., and K. Sperling.** 1996. Molecular analysis of Fanconi anaemia. *Bioessays* **18:** 579-585.
20. **Dolganov, G., R. Maser, A. Novikov, L. Tosto, S. Chong, D. Bressan, and J. Petrini.** 1996. Human Rad50 is physically associated with human Mre11: identification of a conserved multiprotein complex implicated in recombinational DNA repair. *Mol. Cell. Biol.* **16:** 4832-4841.
21. **Engelward, B.P., G. Weeda, M.D. Wyatt, J.L. Broekhof, J. de Wit, I. Donker, J.M. Allan, B. Gold, J.H.J. Hoeijmakers, and L.D. Samson.** 1997. Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proc. Natl. Acad. Sci.***25:** 13087-13092.
22. **Essers, J., R. Hendriks, S. Swagemakers, C. Troelstra, J. de Wit, D. Bootsma, J. Hoeijmakers, and R. Kanaar.** 1997. Disruption of mouse RAD54 reduces ionizing radiation resistance and homologous recombination. *Cell* **89:** 195-204.
23. **Fishel, R., and T. Wilson.** 1997. MutS homologs in mammalian cells. *Curr. Opin. Genet. Dev.* **7:** 105-113.
24. **Fishman-Lobell, J., and J.E. Haber.** 1992. Removal of nonhomologous DNA ends in double-strand break recombination: The role of the yeast ultraviolet repair gene *RAD1*. *Science* **258:** 480-484.
25. **Fox, M., J. Radicella, and K. Yamamoto.** 1994. Some features of base pair mismatch repair and its role in the formation of genetic recombinants. *Experientia* **50:** 253-260.
26. **Friedberg, E.C., G.C. Walker, and W. Siede.** 1995. DNA repair and mutagenesis. ASM Press, Washington D.C.
27. **Fulton, A.** 1982. How crowded is the cytoplasm? *Cell* **30:** 345-347.
28. **Golub, E., O. Kovalenko, R. Gupta, D. Ward, and C. Radding.** 1997. Interaction of human recombination proteins Rad51 and Rad54. *Nucleic Acids Res.* **25:** 4106-4110.
29. **Grawunder, U., M. Wilm, X. Wu, P. Kulesza, T. Wilson, M. Mann, and M. Lieber.** 1997. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388:** 492-495.
30. **Habraken, Y., P. Sung, L. Prakash, and S. Prakash.** 1996. Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Curr. Biol.* **6:** 1185-1187.
31. **Hildebrandt, E., and N. Cozzarelli.** 1995. Comparison of recombination in vitro and in E. coli cells: measure of the effective concentration of DNA in vivo. *Cell* **81:** 331-340.
32. **Ivanov, E., and J. Haber.** 1997. DNA repair: RAD alert. *Curr. Biol.* **7:** R492-495.
33. **Ivanov, E., N. Sugawara, C. White, F. Fabre, and J. Haber.** 1994. Mutations in XRS2 and RAD50 delay but do not prevent mating-type switching in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **14:** 3414-3425.
34. **Ivanov, E.L., and J.E. Haber.** 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.
35. **Jeggo, P.** 1997. DNA-PK: at the cross-roads of biochemistry and genetics. *Mutat. Res.* **384:** 1-14.
36. **Jiang, H., Y. Xie, P. Houston, K. Stemke-Hale, U. Mortensen, R. Rothstein, and T. Kodadek.** 1996. Direct association between the yeast Rad51 and Rad54 recombination proteins. *J. Biol. Chem.* **271:** 33181-33186.
37. **Johzuka, K., and H. Ogawa.** 1995. Interaction of Mre11 and Rad50: two proteins required for DNA repair and meiosis-specific double-strand break formation in Saccharomyces cerevisiae.

Genetics **139**:1521-1532.

38. **Kanaar, R., and J.H.J. Hoeijmakers.** 1997. Recombination and joining: different means to the same ends. *Genes and Function* **1**: 165-174.
39. **Kirkpatrick, D., and T. Petes.** 1997. Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature* **387**: 929-931.
40. **Klungland, A., and T. Lindahl.** 1997. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.* **16**: 3341-3348.
41. **Krokan, H., R. Standal, and G. Slupphaug.** 1997. DNA glycosylases in the base excision repair of DNA. *Biochem. J.* **325** (Pt 1): 1-16.
42. **Kumaresan, K., B. Hang, and M. Lambert.** 1995. Human endonucleolytic incision of DNA 3' and 5' to a site-directed psoralen monoadduct and interstrand cross-link. *J Biol Chem* **270**: 30709-30716.
43. **Leber, R., T. Wise, R. Mizuta, and K. Meek.** 1998. The XRCC4 gene product is a target for and interacts with the DNA- dependent protein kinase. *J. Biol. Chem.* **16 (273)**: 1794-1801.
44. **Liefshitz, B., A. Parket, R. Maya, and M. Kupiec.** 1995. The role of DNA repair genes in recombination between repeated sequences in yeast. *Genetics* **140**: 1199-1211.
45. **Magana-Schwencke, N., J. Henriques, R. Chanet, and E. Moustacchi.** 1982. The fate of 8-methoxypsoralen photoinduced crosslinks in nuclear and mitochondrial yeast DNA: comparison of wild-type and repair-deficient strains. *Proc. Natl. Acad. Sci.* **79**: 1722-1726.
46. **Maser, R., K. Monsen, B. Nelms, and J. Petrini.** 1997. hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks. *Mol. Cell. Biol.* **17**: 6087-6096.
47. **Milne, G., S. Jin, K. Shannon, and D. Weaver.** 1996. Mutations in two Ku homologs define a DNA end-joining repair pathway in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **16**: 4189-4198.
48. **Modrich, P.** 1997. Strand-specific mismatch repair in mammalian cells. *J. Biol. Chem.* **272**: 24727-24730.
49. **Modrich, P., and R. Lahue.** 1996. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **65**: 101-133.
50. **Moore, J., and J. Haber.** 1996. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **16**: 2164-2173.
51. **Morris, T., and J. Thacker.** 1993. Formation of large deletions by illegitimate recombination in the HPRT gene of primary human fibroblasts. *Proc. Natl. Acad. Sci.* **90**: 1392-1396.
52. **Mortensen, U., C. Bendixen, I. Sunjevaric, and R. Rothstein.** 1996. DNA strand annealing is promoted by the yeast Rad52 protein. *Proc. Natl. Acad. Sci.* **93**: 10729-10734.
53. **Mullen, G., and S. Wilson.** 1997. DNA polymerase beta in abasic site repair: a structurally conserved helix-hairpin-helix motif in lesion detection by base excision repair enzymes. *Biochemistry* **36**: 4713-4717.
54. **Nairz, K., and F. Klein.** 1997. mre11S-a yeast mutation that blocks double-strand-break processing and permits nonhomologous synapsis in meiosis. *Genes. Dev.* **11**: 2272-2290.
55. **New, J., T. Sugiyama, E. Zaitseva, and S. Kowalczykowski.** 1998. Rad52 protein stimulates DNA strand exchange by Rad51 and replication. *Nature* **391**: 407-410.
56. **Ogawa, T., A. Shinohara, A. Nabetani, T. Ikeya, X. Yu, E. Egelman, and H. Ogawa.** 1993. RecA-like recombination proteins in eukaryotes: functions and structures of RAD51 genes. *Cold*

Spring Harb. Symp. Quant. Biol. **58:** 567-576.

57. **Palombo, F., I. Iaccarino, E. Nakajima, M. Ikejima, T. Shimada, and J. Jiricny.** 1996. hMutS-beta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr. Biol.* **6:** 1181-1184.
58. **Paques, F., and J. Haber.** 1997. Two pathways for removal of nonhomologous DNA ends during double-strand break repair in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **17:** 6755-6771.
59. **Reitmair, A., M. Redston, J. Cai, T. Chuang, M. Bjerknes, H. Cheng, K. Hay, S. Gallinger, B. Bapat, and T. Mak.** 1996. Spontaneous intestinal carcinomas and skin neoplasms in Msh2-deficient mice. *Cancer Res.* **56:** 3842-3849.
60. **Roeder, G.** 1997. Meiotic chromosomes: it takes two to tango. *Genes Dev.* **11:** 2600-2621.
61. **Sancar, G.B.** 1990. DNA photolyases: physical properties, action mechanism, and roles in dark repair. *Mutat. Res.* **236:** 147-160.
62. **Saparbaev, M., L. Prakash, and S. Prakash.** 1996. Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in Saccharomyces cerevisiae. *Genetics* **142:** 727-736.
63. **Schlake, C., K. Ostermann, H. Schmidt, and H. Gutz.** 1993. Analysis of DNA repair pathways of Schizosaccharomyces pombe by means of swi-rad double mutants. *Mutat. Res.* **294:** 59-67.
64. **Schmidt, H., P. Kapitza-Fecke, E. Stephen, and H. Gutz.** 1989. Some of the swi genes of Schizosaccharomyces pombe also have a function in the repair of radiation damage. *Curr. Genet.* **16:** 89-94.
65. **Shinohara, A., and T. Ogawa .** 1998. Stimulation by Rad52 of yeast Rad51-mediated recombination. *Nature* **391:** 404-407.
66. **Siede, W., A.S. Friedberg, and E.C. Friedberg.** 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.
67. **Sijbers, A.M., W.L. de Laat, R.R. Ariza, M. Biggerstaff, Y.-F. Wei, J.G. Moggs, K.C. Carter, B.K. Shell, E. Evans, M.C. de Jong, S. Rademakers, J. de Rooij, N.G.J. Jaspers, J.H.J. Hoeijmakers, and R.D. Wood.** 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86:** 811-822.
68. **Sladek, F., M. Munn, W. Rupp, and P. Howard-Flanders.** 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.
69. **Sugawara, N., F. Paques, M. Colaiacovo, and J. Haber.** 1997. Role of Saccharomyces cerevisiae Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc. Natl. Acad. Sci.* **94:** 9214-9219.
70. **Sung, P.** 1994. Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* **265:** 1241-1243.
71. **Sung, P.** 1997. Function of yeast Rad52 protein as a mediator between Replication Protein A and the Rad51 recombinase. *J. Biol. Chem.* **272:** 28194-28197.
72. **Sung, P.** 1997. Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase. *Genes Dev.* **11:** 1111-1121.
73. **Szostak, J., T. Orr-Weaver, R. Rothstein, and F. Stahl.** 1983. The double-strand-break repair model for recombination. *Cell* **33:** 25-35.
74. **Tebbs, R., Y. Zhao, J. Tucker, J. Scheerer, M. Siciliano, M. Hwang, N. Liu, R. Legerski, and**

L. Thompson. 1995. Correction of chromosomal instability and sensitivity to diverse mutagens by a cloned cDNA of the XRCC3 DNA repair gene. *Proc. Natl. Acad. Sci.* **92:** 6354-6358.

75. **Thompson, L.** 1996. Evidence that mammalian cells possess homologous recombinational repair pathways. *Mutat. Res.* **363:** 77-88.
76. **Todo, T., H. Ryo, K. Yamamoto, H. Toh, T. Inui, H. Ayaki, T. Nomura, and M. Ikenaga.** 1996. Similarity among the Drosophila (6-4)photolyase, a human photolyase homolog, and the DNA photolyase-blue-light photoreceptor family. *Science* **272:** 109-112.
77. **Todo, T., H. Tsuji, E. Otoshi, K. Hitomi, S.T. Kim, and M. Ikenaga.** 1997. Characterization of a human homolog of (6-4) photolyase. *Mutat. Res.* **384:** 195-204.
78. **Troelstra, C., and N. Jaspers.** 1994. Recombination and repair. Ku starts at the end. *Curr. Biol.* **4:** 1149-1151.
79. **Tsukamoto, Y., J. Kato, and H. Ikeda.** 1997. Budding yeast Rad50, Mre11, Xrs2, and Hdf1, but not Rad52, are involved in the formation of deletions on a dicentric plasmid. *Mol Gen Genet* **255:**543-7.
80. **Tsukamoto, Y., J. Kato, and H. Ikeda.** 1996. Effects of mutations of RAD50, RAD51, RAD52, and related genes on illegitimate recombination in Saccharomyces cerevisiae. *Genetics* **142:** 383-391.
81. **Umar, A., and T. Kunkel.** 1996. DNA-replication fidelity, mismatch repair and genome instability in cancer cells. *Eur. J. Biochem.* **238:** 297-307.
82. **van Duin, M., J. de Wit, H. Odijk, A. Westerveld, A. Yasui, M.H.M. Koken, J.H.J. Hoeijmakers, and D. Bootsma.** 1986. Molecular characterization of the human excision repair gene *ERCC-1:* cDNA cloning and amino acid homology with the yeast DNA repair gene *RAD10. Cell* **44:** 913-923.
83. **van Houten, B., H. Gamper, S.R. Holbrook, J.E. Hearst, and A. Sancar.** 1986. Action mechanism of ABC excision nuclease on a DNA substrate containing a psoralen crosslink at a defined position. *Proc. Natl. Acad. Sci.* **83:** 8077-8081.
84. **Wilson III, D.M., and L.H. Thompson.** 1997. Life without DNA repair. *Proc. Natl. Acad. Sci.* **94:** 12754-12757.
85. **Wilson, T., U. Grawunder, and M. Lieber.** 1997. Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* **388:** 495-498.
86. **Wood, R.D.** 1996. DNA repair in eukaryotes. *Ann. Rev. of Biochem.* **65:** 135-167.
87. **Yajima, H., M. Takao, S. Yasuhira, J. Zhao, C. Ishii, H. Inoue, and A. Yasui.** 1995. A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light. *EMBO J.* **14:** 2393-2399.
88. **Yonemasu, R., S. McCready, J. Murray, F. Osman, M. Takao, K. Yamamoto, A. Lehmann, and A. Yasui.** 1997. Characterization of the alternative excision repair pathway of UV-damaged DNA in Schizosaccharomyces pombe. *Nucleic Acids Res.* **25:** 1553-1558.
89. **Zdzienicka, M.** 1995. Mammalian mutants defective in the response to ionizing radiation-induced DNA damage. *Mutat. Res.* **336:** 203-213.

Chapter 2

Nucleotide excision repair in mammals

2.1 **Introduction**

Nucleotide excision repair, also abbreviated as NER, is a DNA damage removal system, of which the principles have been evolutionary conserved from *Escherichia coli* to man. It counteracts the deleterious effects of a plethora of DNA lesions, including major types of damage induced by environmental sources. The most relevant lesions removed by NER are cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts (6-4PPs), the two major kinds of damage produced by the shortwave UV-component of sunlight (see Fig. 1). Also, bulky DNA adducts, formed by e.g. polycyclic aromatic hydrocarbons present in cigarette smoke, car exhaust fumes and industrial pollution, are removed by NER. Within the divergent spectrum of NER lesions significant distortion of the DNA helix appears the common denominator. A distinguishing mark of NER is the removal of damage as part of a larger DNA fragment, with a typical size of 24-32 nucleotides (73, 119). Overall, the molecular model for NER is thought to comprise (1) damage recognition, (2) lesion demarcation, (3) dual incision around the lesion, (4) release of the damage-containing DNA fragment and (5) gap-filling DNA synthesis and

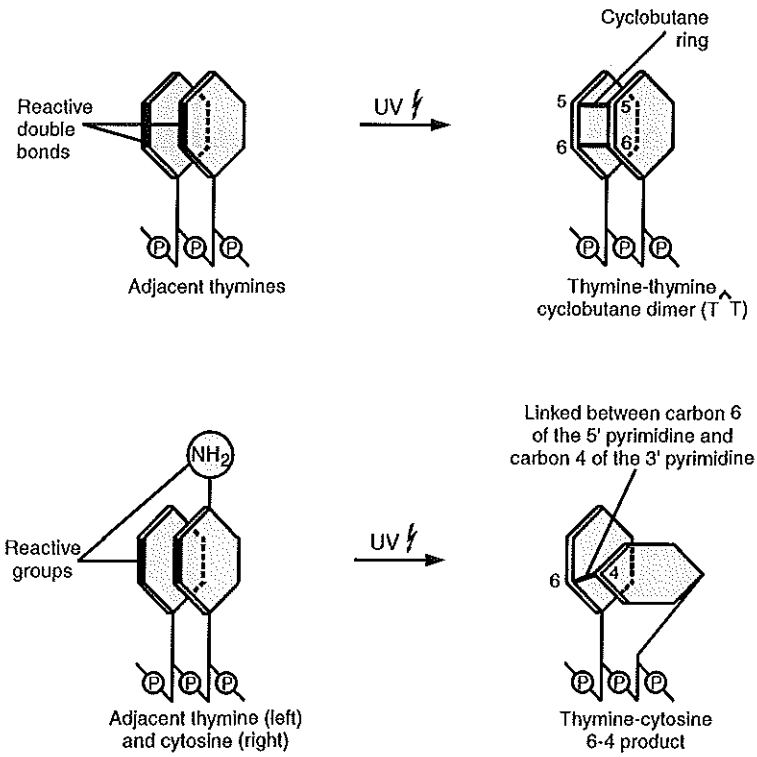


Figure 1. CPDs and (6-4)PPs: UV-induced DNA lesions.

ligation (Fig. 2). Two NER subpathways are distinguished: a transcription-coupled repair pathway (TCR), which accounts for the fast removal of damage from actively transcribed DNA strands, and global genome repair (GGR), which guards the entire genome, but removes some types of lesions more slowly when compared to TCR. In addition to the genomic localization, the type of lesion determines the rate of removal by GGR; (6-4)PPs, for example, are removed approximately 5-fold faster *in vivo* from the overall genome than CPDs (117).

Three rare autosomal recessive human syndromes are connected with a defect in nucleotide excision repair: xeroderma pigmentosum (XP), Cockayne syndrome (CS) and trichothiodystrophy (TTD). The classic NER-defective syndrome XP is characterized by

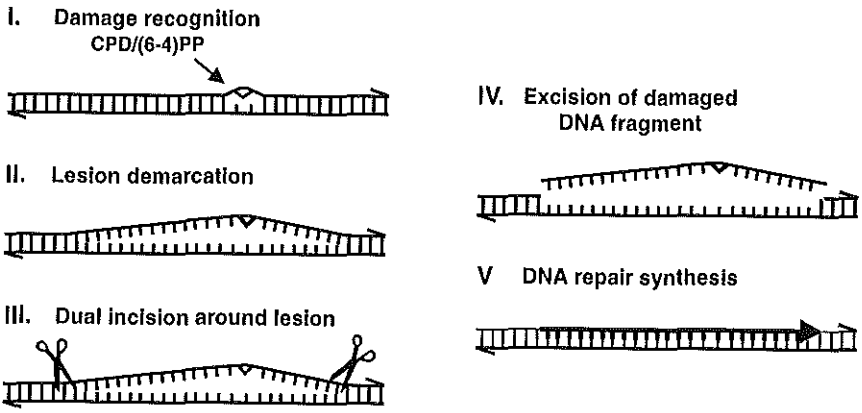


Figure 2. Model for nucleotide excision repair.

extreme photosensitivity, pigmentation abnormalities and pronounced predisposition to skin cancer in sun-exposed parts. Additionally, some XP patients demonstrate neurological degeneration (Table 1). Seven NER-defective XP complementation groups have been identified, each representing defects in distinct repair genes (*XPA-G*). The products of these genes are involved in both subpathways of NER, with the exception of XPC, which is not required for transcription-coupled repair. Patients suffering from Cockayne syndrome are defective in transcription-coupled repair only, and the identification of two CS complementation groups (CS-A and CS-B) shows that at least two additional gene products are needed for the NER machinery to accomplish fast repair of a transcribed strand. Phenotypically, CS is a pleiotropic disorder characterized not only by photosensitivity but also by all sorts of neurological, developmental and aging problems (Table 1). Some XP-B, XP-D and XP-G patients display CS features on top of their XP manifestations. The third syndrome, trichothiodystrophy, is subdivided into a non-photosensitive and a photosensitive form, the latter caused by a defect in NER. TTD is typically characterized by abnormalities in ectodermally derived organs, usually most apparent by the brittle hair of these patients. The remainder of clinical manifestations is

Table 1. Main clinical symptoms of NER syndromes

Clinical Symptoms	XP	XP/CS	CS	TTD
Photosensitivity	++	++	+*	+*
Abnormal pigmentation	++	+	-	-
Skin cancer	++	+	-	-
Progressive mental degeneration	+/-	+	+	+
Neuronal loss	+/-	-	-	-
Neurodysmyelination	-	++	+	+
Wizened facies	-	+	+	+
Growth defect	+/-	+	+	+
Hypogonadism	+/-	+	+	+
Brittle hair and nails	-	-	-	+
Ichthyosis	-	-	-	+

* Also TTD and CS patients exist without photosensitivity and NER defect.
Table adapted from Bootsma *et al.* (1998) (11).

reminiscent of what is observed in CS. Three genes have been associated with TTD: *XPB*, *XPD*, and the as yet uncharacterized *TTDA* gene. Thus, mutations in the former two genes can give rise to XP, TTD, or XP combined with CS features. Similarly, mutations in XPG may result in either an XP or XP combined with CS phenotype. The pleiotropic phenotypical consequences of mutations in these three genes probably are a reflection of additional involvements of the respective gene products in other cellular processes (see below).

The first human nucleotide excision repair gene to be cloned was *ERCC1*, which was reported in 1984 (219). More than a decade later, the cloning of *XPF*, the complexing partner of the ERCC1 protein, completed the isolation of XP genes involved in the core reaction of nucleotide excision repair (13, 171). *In vitro* reconstitution of the GGR pathway of NER with purified proteins revealed that dual incision in NER minimally requires the factors XPC (complexed to hHR23B), XPA, the XPB- and XPD-containing protein complex TFIIH, RPA, XPG and ERCC1-XPF (1, 7, 123, 124). In

addition, RF-C, RPA, PCNA, DNA polymerase δ or ε and DNA ligase I are necessary and sufficient for gap-filling DNA repair synthesis *in vitro* (170).

Below, the main characteristics of the mammalian repair factors will be summarized, including a brief outline of the phenotypical consequences that can arise from defects in these factors. Special attention will be paid to protein-protein interactions, as these can be particularly informative to determine the order and mode of action of these factors during repair. Additional engagements of repair proteins in other processes will be outlined to show that NER, although perfectly suitable to be studied as a distinct reaction mechanism, maintains an intricate communication network with multiple processes in the cell. After introduction of the key players in NER, recent insights into damage recognition, open complex formation and nuclease positioning will be integrated into a detailed model for the molecular events that underlie NER incisions.

2.2 **The repair incision machinery**

2.2.1 **XPA**

Involvement and mutant phenotype

The *XPA* gene product plays a crucial role at an early stage of both transcription-coupled repair and global genome repair (181), but unlike the majority of repair factors, XPA-requirement appears to be restricted to NER only. A relatively large number of XP-A patients have been identified and most of them display a severe clinical form of XP, showing not only high incidence of skin cancer and other cutaneous abnormalities, but also accelerated neurodegeneration, often apparent at an early age. A useful quantitative parameter for the NER capacity of a cell is UV-induced Unscheduled DNA Synthesis (UDS), which measures the incorporation of radiolabeled deoxynucleotides in repair patches during a limited period of time after exposure of the cells to UV-light. Cells from the majority of XP-A patients show UV-induced UDS-levels that are less than 5%

compared to normal cells (reviewed by (11)). The phenotype of *XPA* knockout mice is comparable to that observed in XP-A patients, with the exception that the mice do not show neurological problems, even not at an age of 18 months (32, 131).

Protein and activity

The XPA protein is a 273 amino acid DNA-binding protein with a marked preference for damaged DNA (4, 79, 153). The protein harbors a Zn^{2+}-finger domain (residues 101-141) within the minimal region required for DNA-binding (residues 98-219) (14, 94, 121, 181), which is essential for its function (4, 118). Various lesions are recognized by XPA, including cisplatin-induced intrastrand cross-links and UV-induced lesions, all substrates for the NER machinery (4, 79, 153). XPA demonstrates a much higher affinity for (6-4)PPs than for CPDs *in vitro*, in parallel with the faster removal of the former lesions by NER *in vivo* (79). In general, the affinity of XPA for a lesion seems to correlate with the degree of helical distortion the lesion induces and it has been suggested that the single-strandedness of damaged sites may be the determinant for XPA-recognition (79). On the basis of these findings, XPA is generally assumed to be the damage recognition protein in NER.

Interactions with other proteins

Various protein-protein interactions have been reported between XPA and other NER protein factors. XPA can bind *in vitro* to ERCC1, involving the residues 72-84 in positioning of the ERCC1-XPF endonuclease complex onto the DNA. XPA was also XPA (100, 102, 145), and weakly to XPF (7). Possibly, this interaction facilitates the reported to interact *in vitro* with the p32 and p70 subunit of RPA, and binding involved the XPA-residues 4-29 and 98-187, respectively (66, 101, 158). The relevance of the observed RPA p32-XPA interaction is at present not clear though, as the N-terminal 58 residues of XPA were found to be dispensable for XPA-functioning *in vivo* (118). A complex of XPA and RPA binds cooperatively to DNA lesions, which led to the suggestion that together these factors perform the first step in NER (66). Finally, an interacti-

on between XPA and the helicase complex TFIIH has been reported, accomplished through residues 226-273 of XPA (134, 144).

The network of protein-protein contacts between XPA and other core NER factors suggests that XPA has a central role in coordinating events in NER. However, the established idea that XPA, perhaps in conjunction with RPA, is the NER initiation factor, appears to be wrong. Recent data convincingly demonstrated that XPC acts prior to XPA *in vitro* and serves to recruit the remainder of the NER machinery to the lesion (Sugasawa *et al.*, submitted (176)). Implications of this finding for the role of XPA in the NER reaction will be discussed in the section 'mechanism of nucleotide excision repair'.

Apart from interactions with core NER factors, XPA has also been reported to interact *in vitro* with CSB (163) and with the p34 subunit of TFIIE, the basal transcription factor that recruits TFIIH to the RNA polymerase II initiation complex (144). Whereas these interactions are not required for global genome repair, they may play a role in transcription-coupled repair (144, 163).

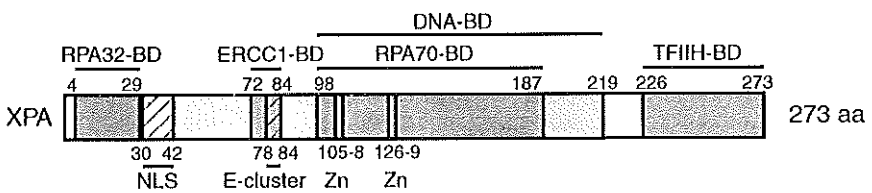


Figure 3. The XPA protein. Schematic presentation of functional domains in the XPA protein. DNA- and protein interaction domains are indicated above (filled boxes), (putative) functional domains based on primary amino acid sequence are indicated below (hatched boxes). *NLS*: nuclear localization signal; *E-cluster*: glutamic acid-repeat; *Zn*: putative Zn-finger.

2.2.2 **XPC-hHR23B**

Involvement and mutant phenotype

XPC is the sole XP-factor that is dispensable for transcription-coupled repair (198, 202, 203). Patients defective in *XPC* are identified relatively frequently and generally show mild clinical symptoms with cutaneous problems (including skin cancer) and eye problems but without neurological abnormalities. The UV-induced UDS levels of most XP-C cell lines vary from 15-30% compared to wild-type cells, and all XP-C cells are less UV-sensitive than typical XP-A strains (reviewed by (11)), consistent with the notion that TCR is still functional. Mice carrying an inactivating mutation in the *XPC* gene display similar clinical manifestations (159).

Protein complex and activity

In vivo, the 125 kDa XPC protein is stably bound to the 58 kDa hHR23B protein, one of the two human homologs of the yeast repair protein Rad23 (109). The second Rad23 homolog, hHR23A, mainly exists in a free form in mammalian cells (109). In *in vitro* repair assays, hHR23B stimulates XPC activity (174), and probably has a structural rather than a catalytic function, as an isolated 54 amino acid XPC-binding domain is already sufficient for XPC stimulation (108). Interestingly, hHR23A can substitute for hHR23B in binding and stimulating XPC *in vitro*, suggesting some functional redundancy (175). All Rad23 derivatives harbor a ubiquitin-like moiety at their N-terminus (109). The relevance of this domain for repair is still not entirely clear though; in mammalian *in vitro* assays it appears dispensable for XPC stimulation, but in *S. cerevisiae* it was reported to be essential for repair *in vivo* (216). hHR23B is much more abundant than XPC in mammalian cells and the majority of hHR23B, like hHR23A, is found in a free form (193). Likely, these proteins have additional engagements outside NER, a notion underscored by the observation that many hHR23B-deficient mice die before birth (Jessica Ng, personal communication).

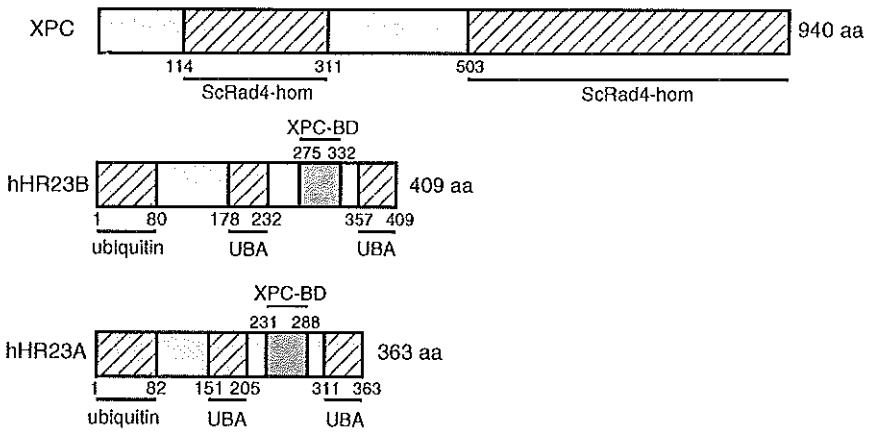


Figure 4. The XPC-hHR23B protein complex. Schematic presentation of functional domains in the XPC, hHR23B and hHR23A protein. Protein interaction domains are indicated above (filled boxes), (putative) functional domains based on primary amino acid sequence are indicated below (hatched boxes). *ScRad4-hom:* domains within XPC with (limited) homology to *S. cerevisiae* Rad4; Ubiquitin: ubiquitin-like region; *UBA*: ubiquitin-associated domain.

Function of the XPC-hHR23B complex

XPC-hHR23B and XPC alone display a similar high affinity for both single-stranded and double-stranded DNA and a preference for UV-damaged DNA (109, 151, 168). The complex is known to be dispensable for transcription-coupled repair, and its function in global genome repair has been obscure for quite some time. XPC-hHR23B is absolutely required for dual incision as well as for open complex formation during NER (1, 41, 124, 126). Assuming that in the XPC-independent TCR pathway the transcription machinery provides the NER proteins with a DNA intermediate pre-opened at the 3' side of the lesion, XPC-hHR23B was speculated to stabilize partially unwound DNA (125, 126). Consistently, it was found that an artificial DNA substrate mimicking such a transcription bubble in front of a lesion could be repaired in the absence of XPC-hHR23B (123). In fact, also bubbles opened across a lesion and bubbles 5' of a lesion alleviated the requirement of XPC-hHR23B (126). Furthermore, whereas most repair factors were dispensable for (limited) opening, both the helicase complex TFIIH and XPC-hHR23B were absolutely required for any helical distortion around a lesion (41).

Also, XPC-hHR23B is not required for *in vitro* repair of certain types of lesions, notably those with a cholesterol moiety, suggesting that certain DNA structures obviate the need for XPC-hHR23B (123). These data are all in line with a role for XPC-hHR23B in open complex formation and/or stabilization.

Evidence for XPC-hHR23B involvement in open complex formation was recently presented by a series of experiments, in which two distinguishable damaged plasmids were pre-incubated each with a different repair factor, then mixed in the presence of a (repair-competent) cell extract after which the repair rate of the two plasmids was measured and compared (Sugasawa *et al.*, submitted (176)). These assays revealed that XPC-hHR23B is in fact the first actor in NER, required even before XPA (and RPA) and capable of recruiting the rest of the repair machinery to the lesion (Sugasawa *et al.*, submitted (176)). By footprinting it was shown that XPC-hHR23B directly binds to DNA damage, changing the DNase I cleavage pattern in both the damaged and the undamaged strand locally around the lesion (Sugasawa *et al.*, submitted (176)). Clearly, this newly identified role of XPC-hHR23B as a damage sensor and repair recruitment factor sheds new light upon many aspects of global genome repair as well as transcription-coupled repair, to be discussed in the model for NER presented below.

Interactions with other proteins

Despite the fact that XPC and hHR23B have been available as purified factors in several laboratories for quite some time now, few interactions with other repair factors have been reported. In fact, only one report claims that XPC co-purifies with TFIIH over 7 chromatographic steps, suggesting an interaction between XPC-hHR23B and this protein complex (36). In *S. cerevisiae*, *in vitro* binding between the XPC homolog Rad4 and TFIIH was observed (5), supporting a direct interaction between these two protein complexes. Also in yeast, a Rad23-affinity matrix was found to retain TFIIH components as well as Rad14, the yeast homolog of XPA (61). However, for comparison of XPC with yeast Rad4 it is important to know that unlike its mammalian counterpart, Rad4 is required for TCR. Similarities and differences between the mammalian XPC-

hHR23B complex and its yeast counterpart Rad4-Rad23 will briefly be addressed in paragraph 2.6.

2.2.3 **TFIIH**

Involvement and mutant phenotype

TFIIH is a multi-subunit protein complex originally purified as a basal transcription factor required for *in vitro* reconstituted RNA polymerase II transcription (24, 44, 49, 54). The subsequent identification of the repair helicases XPB and XPD and the cyclin-dependent kinase- (Cdk-) activating kinase factors Cdk7, cyclin H and MAT1 as components of TFIIH revealed additional engagements of this protein complex in nucleotide excision repair and possibly in cell cycle regulation, respectively (36, 45, 46, 155, 160, 161, 164, 167, 200, 214, 215). TFIIH is a core NER component required for both TCR and GGR. The multiple involvements of TFIIH provide a rationale for the pleiotropic clinical picture observed with XP-B and XP-D patients. XP-B patients are rare, with only three families described worldwide, two of which display the combined XP/CS phenotype and the other showing TTD characteristics. XP-D patients occur more frequently and are also phenotypically diverse, with clinical characteristics varying from XP to XP/CS and TTD (reviewed by (11, 20)). Whereas the classic XP-features such as UV-induced cutaneous abnormalities, photosensitivity, high incidence of skin cancer and neurological degeneration are ascribed to improper functioning of TFIIH in NER, additional CS/TTD-like problems such as growth retardation, neurodysmyelination and brittle hair are thought to arise from a defective role of TFIIH in the transcription of certain genes (209). Support for the latter notion was recently provided by the observation that mice carrying a point mutation in *XPD*, with clinical manifestations resembling UV-sensitive TTD, have reduced levels of transcription of a gene implicated in terminal skin differentiation (de Boer, submitted (29)). An alternative, not mutually exclusive, cause of CS-like features is the accumulation of (endogenously induced) damage in the

genome due to defective TCR; interestingly, not only NER-specific lesions but also other lesions such as oxidative damage appear to be removed in a transcription-coupled fashion (25, 96, 97) (to be discussed extensively in section 2.5).

Protein complex and activity

The activities associated with TFIIH are multiple. XPB and XPD exhibit DNA-dependent ATPase and helicase activities; XPB can unwind DNA in a 3'-5' direction, XPD does this in a 5'-3' direction (156, 160, 161). Cdk7, cyclin H and MAT1 constitute the cdk-activating kinase (CAK) complex that can phosphorylate the cyclin-dependent kinases (cdks) cdk1, cdk2 and cdk4, relevant to cell cycle regulation, as well as the C-terminal domain (CTD) of RNA polymerase II (Pol II), important for transcription initiation (44, 155). Four other subunits have been identified, notably p34, p44, p52 and p62, with the smallest two subunits containing zinc finger motifs and putative DNA-binding capacity (74). No functions have been assigned to p52 and p62 yet (47, 107). At present the consensus number of subunits of TFIIH is nine (220), but alternative forms of TFIIH may exist in the cell. The core protein complex comprises p34, p44, p52, p62 and XPB. CAK is present both as a free ternary complex and as a component of TFIIH, presumably by binding to XPB and XPD (154). Consistent with its affinity for both core TFIIH and CAK, XPD-core TFIIH and XPD-CAK complexes have both been observed in cell extracts (150, 154, 164, 167, 178). CAK is required for transcription initiation but is dispensable for *in vitro* nucleotide excision repair (123, 177, 191, 214). Free CAK preferentially phosphorylates the cdks, whereas CAK-TFIIH prefers the CTD of Pol II as a kinase substrate (154).

TFIIH in transcription

The function of TFIIH in transcription initiation may parallel its role in nucleotide excision repair. In general, four steps can be distinguished in transcription initiation: (i) the formation of an inactive preinitiation complex, composed of the basal transcription factors TFIIA, B, D, F, and Pol II, (ii) transition to an active initiation complex, accom-

panied by promoter opening, (iii) transcription initiation (defined as the synthesis of the first phosphodiester bond of the transcript), and (iv) transition to a productive elongation complex (also known as 'promoter clearance') (reviewed in (139)). TFIIH, the only general transcription factor with known enzymatic activities, is required during step (ii), step (iv) and possibly step (iii). In step (ii), the ATP-dependent helicases of TFIIH mediate melting of the promoter between position -9 and +1 relative to the initiator site, which is necessary to activate the initiation complex (38, 71, 78, 229). In step (iii), expansion of the opening to position +8 is observed, dependent on the formation of the first RNA phosphodiester bond by the activated initiation complex (71, 229). Whether this extra opening requires TFIIH is still a matter of debate. The observation that neither a 6 bp premelted promoter region (melted from position -4 to +2), nor negatively supercoiled templates require TFIIH and ATP to start transcription from, suggests that TFIIH is only required for the initial opening and additional downstream opening may be accomplished by Pol II itself (56, 71). However, others found that the additional opening to +8 does require ATP, suggesting the involvement of an ATPase activity, possibly TFIIH, in this step (229). Step (iv) requires the kinase activity of TFIIH to hyperphosphorylate the C-terminal domain (CTD) of Pol II, enabling its transition into an elongation mode (28, 46, 78, 164, 167). This so-called promoter clearance occurs before the transcript reaches 10 nucleotides (38), and is accompanied by the release of most general transcription factors (230). TFIIH was found to dissociate somewhat later from the elongation complex, between position +30 to +68 (230). Given its involvement in transcription-coupled repair, the fact that TFIIH is released from an elongating Pol II may come as surprise. However, the complex was recently found to be able to reassociate with very early elongation complexes to suppress premature arrest (39). A similar reassociation with an elongating Pol II stalled at a lesion may be anticipated during transcription-coupled repair.

The transcription activities of TFIIH heavily depend on the presence of TFIIE. In transcription initiation, TFIIE recruits TFIIH to the inactive preinitiation complex (49). TFIIE interacts with all general transcription factors as well as with Pol II, whereas

TFIIH only binds to TFIIE and has almost no affinity for Pol II (140, 141). Upon binding to TFIIH, TFIIE stimulates the ATPase and CTD-kinase activity of the complex, but has a negative effect on its DNA-helicase activity (36, 140, 141). TFIIE may also support promoter opening in a direct manner, perhaps by binding and stabilizing melted single-stranded DNA regions (56, 70, 185, 190). After transcription initiation, TFIIE dissociates from Pol II complexes before position +10 (230). Nevertheless, TFIIE, like TFIIH, is required for reinitiation of early arrested Pol II elongation complexes, suggesting that both TFIIH and TFIIE can reassociate with stalled Pol II complexes (39). Thus, although TFIIE is dispensable for global genome repair (144), it still might function in transcription-coupled repair, perhaps to reinitiate elongation or otherwise to recruit TFIIH and other factors to the stalled Pol II complex. In this respect, TFIIE has been found to interact *in vitro* with XPA as well as with CSB (144, 163).

TFIIH in nucleotide excision repair

TFIIH is absolutely required for open complex formation in nucleotide excision repair (1, 40, 41, 124, 126). Permanganate footprinting studies in cell-free extracts revealed that only mutations in XPB and XPD, the two helicase subunits of TFIIH, and XPC can completely abolish unwinding around the lesion, whereas the absence of RPA, XPA or XPG activities still allows limited opening (41). The apparent presence of a partially unwound DNA intermediate may indicate that, as in transcription initiation, open complex formation in NER occurs in two steps, with the first depending on XPC and TFIIH and the second requiring additional factors (discussed in detail below). Demarcation of the lesion through local DNA opening provides the junctions between single-stranded and duplex DNA that are cleaved by the two NER nucleases, XPG and ERCC1-XPF. Studies with premelted DNA substrates containing a site-specific lesion revealed that DNA unwinding is not the only function that TFIIH fulfills in excision repair (125). Unlike what was found for transcription initiation, bubbles of 10 nt at the 3' and 5' side, as well as one of 20 nt spanning the lesion, all still require TFIIH for efficient further processing, indicating an additional structural role for TFIIH in repair.

Whereas hardly any TFIIH interactions were observed with basal transcription factors, a number of interactions with repair factors have been reported, including the already mentioned interactions with XPA and XPC (5, 36, 134, 144). Additionally, direct interactions were observed *in vitro* between multiple subunits of TFIIH and XPG (77), explaining the previously reported extensive copurification of XPG with TFIIH (124). In agreement, in *S. cerevisiae* the homologs of TFIIH and XPG (Rad2) are found in a stable and distinct excision repair complex (5, 62). A mutant in the C-terminal domain of XPB was reported to fully support DNA unwinding and allow 3' incision, but not 5' incision (41). These data suggest that TFIIH has an additional involvement in positioning other repair factors and possibly facilitates the 5' incision by ERCC1-XPF (41). *In vitro* binding was also observed between TFIIH and CSA, an interaction perhaps relevant for transcription-coupled repair (67).

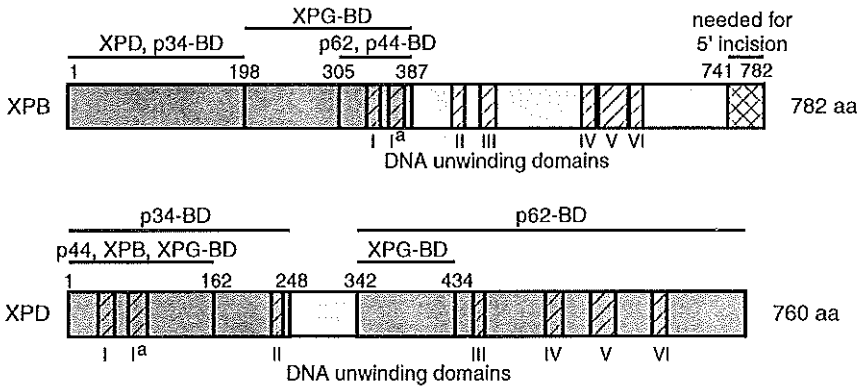


Figure 5. The XPB and XPD subunits of the TFIIH protein complex. Schematic presentation of functional domains in the XPB and XPD proteins. Protein interaction domains are indicated above (filled boxes), (putative) functional domains based on primary amino acid sequence are indicated below (hatched boxes). *p34*, *p44* and *p62* are other TFIIH subunits.

Additional engagements

TFIIH has also been implicated in apoptosis, mainly based on an abnormal p53-induced apoptotic response that was observed in XP-B and XP-D cells (212). Along this line, p53 was found to interact with the TFIIH subunits XPB, XPD and p62, resulting in inhibition of the two helicase activities of TFIIH (99, 212, 213, 227). These findings potentially provide an intriguing link between apoptosis on the one hand and repair and basal transcription on the other hand, but additional experiments are required to better understand this notion.

2.2.4 **Replication Protein A**

Involvement and mutant phenotype

Replication protein A (RPA) was originally identified as a factor required for *in vitro* SV40 DNA replication (42, 221, 224). Now, RPA has an established role not only in DNA replication, but also in nucleotide excision repair and recombination ((26, 27), reviewed by (223)). Indeed, in yeast, RPA mutations were found to interfere with replication, repair and recombination *in vivo* (106). No mammalian mutants defective in RPA are known, presumably because alterations in this protein complex are incompatible with life. The RPA-dependent DNA-metabolizing processes have in common that at a certain stage, complementary DNA strands are separated and action is required along single-stranded (ss) DNA intermediates. RPA is thought to stabilize such intermediates and remove secondary structures by binding to ssDNA.

Protein complex and activity

Human RPA (hRPA) is a single-stranded DNA-binding protein composed of three subunits of 70, 32 and 14 kDa (for a comprehensive review on RPA, see (223)). Its apparent association constant to ssDNA of 10^9-10^{11} M^{-1} is at least three orders of magnitude higher than to double-stranded DNA (89, 90). Binding of an RPA molecule to

ssDNA involves the 70 kDa subunit (55, 87, 92, 222), but single-stranded binding domains are also present in the 32 and 14 kDa subunits (146). Two binding modes have been identified; RPA interacts with a minimal region of 8-10 nucleotides (nt) (9) and with an optimal region of approximately 30 nt (90). Binding to 8-10 nt appears globular shaped under scanning transmission electron microscopy and is thought to precede the almost 100-fold more stable, elongated 30 nt binding mode (10, 89). In this thesis, evidence is provided that RPA binds ssDNA with a defined polarity; initial RPA-binding to 8-10 nt occurs at the 5' side of the presumptive 30 nt binding region, and transition to the full 30 nt binding form likely involves stretching of the RPA molecule along the DNA in the 3' direction (Chapter 6) (de Laat *et al.*, submitted (31)). Perhaps this two-step binding mode, in conjunction with cooperative RPA-binding, underlies the ATP-independent DNA-unwinding capacity of RPA that has been observed with naked DNA templates under certain ionic conditions (53, 110, 187). Cooperativity of RPA-binding to single-stranded DNA is considered low, but human RPA is still 10-20 times more likely to bind adjacent to an already bound RPA molecule than to naked DNA (89, 91).

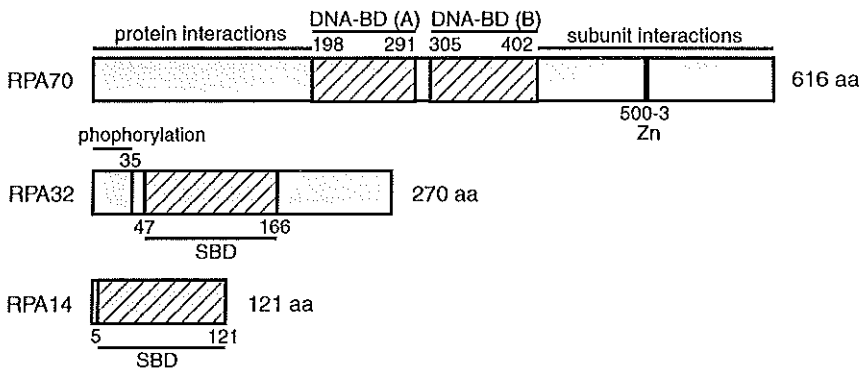


Figure 6. Replication protein A. Schematic presentation of functional domains in the three subunits of human RPA. *SBD*: putative single-stranded DNA binding domain.

Function of the protein complex

RPA is frequently found to facilitate DNA unwinding by ATP-dependent helicases (12, 183, 222, 224). In fact, this activity is one of its major contributions to both replication and repair. In SV40 DNA replication, ATP-dependent opening at the origin of replication by large T antigen requires RPA (222, 224). This requirement does not involve specific protein-protein interactions, as single-stranded DNA binding proteins (SSBs) from other species can substitute for RPA (88). In nucleotide excision repair, RPA is also indispensable for full opening around the lesion (41, 126), and it might help the helicases of TFIIH in doing so. The size of the fully opened repair complex is approximately 30 nt, which corresponds to the size of the optimal DNA-binding region of RPA. Whether other SSBs can substitute for RPA in NER opening is not known, but a subsequent step in NER, dual incision around the demarcated lesion, absolutely requires specific protein-protein interactions with RPA (see below).

Interactions with both XPG and ERCC1-XPF have been reported, indicating that RPA has a function in positioning these NER nucleases (7, 40, 66, 112). Controversy existed, however, whether RPA positively or negatively modulates these nuclease activities (7, 40, 112). The recent finding, described in Chapter 6, that RPA binds ssDNA with a defined polarity clarifies this issue. Bound to single-stranded DNA, the 3' oriented side of RPA specifically interacts with and stimulates the incision activity of ERCC1-XPF but does not interact with XPG, whereas the 5' oriented side of RPA inhibits ERCC1-XPF incisions and specifically interacts with XPG. Thus, each RPA side modulates a given repair endonuclease differently. As apparently the 3' oriented side of RPA positions the 5' nuclease in NER (ERCC1-XPF), and the 5' oriented side of RPA positions the 3' nuclease in NER (XPG), these data pinpoint RPA to the undamaged strand in the repair reaction (de Laat *et al.*, submitted (31)).

RPA and XPA were found to bind cooperatively to damaged DNA (66, 158). RPA itself was also found to have affinity for damaged DNA, notably for cisplatin-induced lesions, AAF damage and 6-4 photoproducts (17, 22, 66), which led to the suggestion

that this factor is involved in DNA damage recognition. In NER, though, RPA is likely to be bound to the undamaged strand (de Laat *et al.*, submitted (31)), indicating that RPA does not recognize the lesion *per se*, but rather has affinity for single-stranded regions exposed by lesion-induced helical distortion. Consistent with this notion, RPA was found to bind to single-stranded bubbles as small as 4 nt (223). Recently, XPC-hHR23B was demonstrated to act prior to XPA, RPA and XPA-RPA together in NER, ruling out that RPA and XPA initiate repair (Sugasawa *et al.*, submitted (176)). Nevertheless, the affinity of RPA and XPA for damage is still expected to be relevant for NER, and may for example assure the correct positioning of other factors with regard to the lesion (see below).

Based on its dual involvement in replication and repair, it can be anticipated that RPA not only acts at pre-incision stages but also during DNA repair synthesis. DNA polymerase δ and ε (Pol δ and ε) have been implicated in repair synthesis, and both were found to be stimulated by RPA (87, 88, 98, 188). Stimulation was not dependent on specific protein-protein interactions, since other SSBs could replace RPA (88, 98, 188). However, genetic evidence in yeast suggests that RPA and Pol δ do have a direct interaction (106). Possibly, RPA remains bound to the undamaged strand after excision, thereby facilitating gap filling DNA repair synthesis.

Finally, the p32 subunit of RPA was found to be phosphorylated in a cell cycle dependent manner and in response to DNA damage (18, 33, 43, 105). Whether the phosphorylation status of RPA has any effect on the efficiency of NER *in vivo*, remains to be established (3, 142).

2.2.5 XPG

Involvement and mutant phenotype

XP-G patients display a heterogeneous clinical picture, varying from mild cutaneous and no neurological abnormalities to severe skin problems and neurological impairment. Generally, cells from these patients are very UV-sensitive and UV-induced UDS levels vary from <5% to 25% (reviewed in (11)). Like XP-B and XP-D patients, a relatively large number of group G patients show typical CS features in addition to XP defects. CS features are thought to arise from a damage-induced transcriptional problem, and therefore a combined XP-CS phenotype indicates that the gene involved has a transcription-related function in addition to its role in the core NER reaction. XPB and XPD, being components of the basal transcription factor TFIIH, may be anticipated to have such a dual function, but for XPG this notion is surprising. A possible specific role for XPG in TCR will be discussed in paragraph 2.5.

XPG in nucleotide excision repair

The *XPG* gene encodes a protein with a predicted molecular mass of 133 kDa (138, 162), which acts as structure-specific endonuclease, cleaving a variety of artificial DNA substrates, including bubbles, splayed arms and stemloops (21, 40, 137, 138, 171). Additionally, the XPG homolog of *S. cerevisiae*, Rad2, was found to cleave flap substrates (63). XPG-mediated incisions always occur in one strand of duplex DNA at the 3' side of a junction with single-stranded DNA. One single-stranded arm, protruding in either the 3' or 5' direction, is necessary and sufficient for the correct positioning of XPG incisions (30). Consistent with this cleavage polarity, XPG makes the 3' incision during NER (137), which is placed at the border of the open DNA intermediate (40). Bubble substrates mimic these open repair complexes, and in such substrates XPG required a minimal opening of 5 nt for incisions (40). The XPG protein is a member of the FEN-1 family of structure-specific endonucleases, which all cut with similar polarity at junctions of duplex and unpaired DNA (reviewed by (104)). The primary amino acid se-

quence of XPG shares three regions of homology with the founder of this family, FEN-1, which is implicated in the processing of Okazaki fragments during replication (157), the completion of a subpathway of BER (93) and possibly DNA-end joining of certain double-stranded DNA breaks (93). Three aspartic acid residues and one glutamic acid residue are absolutely essential for FEN-1 cleavage (165, 166). One of these residues, Asp-181 in FEN-1, is unique in the sense that it is required for cleaving but not for binding DNA (165, 166). In agreement, mutation of the corresponding residue in XPG, Asp-812, gave an active-site mutant, D812A, which was specifically defective in nuclease activity (126, 211).

The XPG-mediated 3' incision precedes the 5' incision made by ERCC1-XPF (123). Interestingly, XPG is not only required for the 3' incision, but also for full open complex formation, indicating a structural role in the core NER reaction (41, 126). Evidence for such a role was provided with the D812A active-site mutant of XPG, which had to be present in order to detect ERCC1-XPF-mediated 5' incisions in an *in vitro* reconstituted repair assay with purified factors (126, 211). Furthermore, this same XPG mutant was found to stabilize a preincision complex containing XPC-hHR23B, TFIIH, XPA and RPA (126). Apparently, independent of its cleavage activity, XPG has a structural function in the assembly of the NER DNA-protein complex. In agreement, several protein interactions have been reported between XPG and other repair factors.

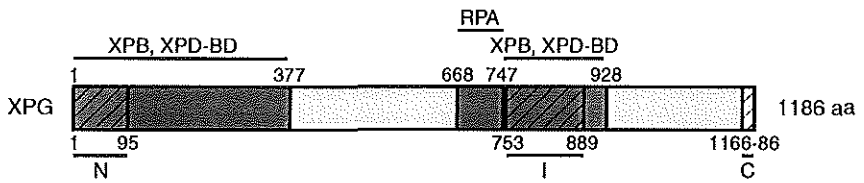


Figure 7. XPG. Schematic presentation of functional domains in the XPG protein. Protein interaction domains are indicated above (filled boxes), (putative) functional domains based on primary amino acid sequence are indicated below (hatched boxes). *N*, *C* and *I* are regions conserved in the FEN-1 family of structure-specific endonucleases.

Protein interactions

XPG interacts with the core NER factors TFIIH (see above) (5, 77), and RPA (see above, and see Chapter 6) (66) (de Laat *et al.*, submitted (31)), as well as with CSB (77). Direct interactions between TFIIH, CSB and XPG may have implications for the mechanism that underlies CS (see below). XPG specifically interacts with the 5' oriented side of RPA, without a pronounced effect on incision activity (see Chapter 6) (de Laat *et al.*, submitted (31)). Given the structural function of XPG in NER, this interaction may assist the assembly of the whole NER complex. Furthermore, similar to FEN-1, XPG was found to interact with Proliferating Cell Nuclear Antigen (PCNA) (52), a factor involved in DNA repair synthesis but dispensable for the incision stage of NER (169). Whereas PCNA stimulates FEN-1 nuclease activity (226), no effect was found when PCNA was added to XPG nuclease assays (40). Perhaps, this interaction allows cross-talk between the *in vitro* separatable incision stage and DNA synthesis stage of NER.

2.2.6 **ERCC1-XPF**

Involvement and mutant phenotype

The gene products of *ERCC1* (excision repair cross complementation group 1) and *XPF* have a predicted molecular mass of 33 kDa and 103 kDa, respectively (13, 171, 194). The proteins form a stable complex *in vivo* and *in vitro* (see Chapters 3 and 4) (8, 143, 171, 199), involving amino acid stretches in the C-terminus of both ERCC1 and XPF (see Chapter 4). Stability of the individual components in the cell is dependent on complex formation (see Chapter 4) (172, 228). Chinese hamster cells defective in either ERCC1 or XPF not only show UV-sensitivity but are also extremely sensitive to interstrand cross-linking agents, a unique feature among the collection of NER-deficient rodent cells (23). Removal of such cross-links (see Chapter 1) is assumed to require recombinational events, and consistent with such a role of ERCC1-XPF in recombination is the finding that the homologous complex in *S. cerevisiae*, Rad10-Rad1, is required

for a specific mitotic recombination pathway called single-strand annealing (SSA) (see Chapter 1) (48). Also, its *Schizosaccharomyces pombe* counterpart Swi10-Rad16 is involved in the recombinational events that underlie mating-type switching (57).

Less than 20 XP-F patients have been described worldwide, who vary only little in their clinical picture. All show mild cutaneous manifestations, only a minority have skin cancers and two display mild late-onset neurological problems (122) (Sijbers, 1998, in press). The UV-induced UDS levels of XP-F cells range between 15% to 30% (11). All these features reflect a mild classic NER-defect. Intriguingly, not a single human patient with a primary defect in ERCC1 has been described. Knockout mice had to be generated to study the consequences of a defective *ERCC1* gene (113, 217). On top of the classic NER-deficient phenotype, ERCC1-deficient mice suffer from liver, spleen and kidney abnormalities, developmental delay, early death and signs of premature senescence (113, 217). Similar severe symptoms were observed with mice carrying a 7 amino acid truncation from the C-terminus of ERCC1 ($ERCC1^{*292}$), although their lifespan was somewhat extended in comparison to the full knock-out (217). The non-NER-specific clinical features presumably are caused by a defect in the recombinational engagement of ERCC1-XPF, and it has been proposed that accumulation of endogenously induced interstrand cross-links underlies these manifestations (217). Clearly, complete disruption of the ERCC1-XPF function severely affects normal development. The apparent absence of ERCC1 patients may indicate that particularly the integrity of this protein is crucial for the function of ERCC1-XPF in normal development, and indeed there are some indications that ERCC1 is the catalytic subunit of the protein complex (see below). All XP-F patients examined still display low levels of ERCC1 and XPF protein (228) (Anja Raams and Anneke Sijbers, personal communication), and their relatively high residual UDS levels suggest that they carry leaky mutations in *XPF*. Residual activity may in fact be the key for the absence of severe developmental problems in these individuals. The observation that less ERCC1 protein is required for the repair of the very toxic cross-links than for nucleotide excision repair, supports this notion (172). Interestingly, an XP-F patient was recently identified who not only showed an unusually severe NER

defect but also developmental problems such as kidney and liver abnormalities and premature aging (unpublished observation), remarkably similar to the symptoms observed in ERCC1 mice. Consistent with the hypothesis, the residual activity in this unusual XP-F patient was much lower than in all other XP-F patients, with UDS levels of less than 5% and no detectable ERCC1 and XPF proteins. Possibly, this patient represents a novel NER-related syndrome with additional involvement of deficient cross-link repair that in the future may also encompass patients with a defect in the ERCC1 protein.

ERCC1-XPF complex and activity

The ERCC1-XPF complex is a structure-specific endonuclease (Chapter 3) (171). Like its yeast homolog Rad1-Rad10, ERCC1-XPF incises a variety of DNA substrates, including bubbles, stem-loops, splayed arms and flaps (7, 30, 171), with the latter possibly representing recombination intermediates (see Chapter 1). Single-stranded loops of 4-8 nt are minimally required to detect ERCC1-XPF incisions (30). Incisions are always made with the same polarity, being in one strand of the duplex at the 5' side of the junction with single-stranded DNA (30, 171). One single-stranded arm protruding in either 3' or 5' direction is necessary and sufficient to correctly position ERCC1-XPF incisions at a DNA junction (see Chapter 5) (30). In NER, ERCC1-XPF makes the 5' incision, consistent with its cleavage polarity (111, 171).

A hint that ERCC1 may be the subunit which catalyzes cleavage, comes from the presence of a double helix-hairpin-helix (HhH) motif in the primary sequence of this protein (172). HhH-motifs are shared by other structure-specific endonucleases and have been implicated in DNA binding (34). Deletion of the double HhH-motif from the *E. coli* NER protein UvrC disrupts 5' incisions without affecting 3' cleavage (120). Possibly, this DNA-binding domain positions the catalytic cleavage site for 5' incision. Curiously, HhH-motifs are apparently lacking in the yeast homolog of ERCC1-XPF, Rad1-Rad10. The relevance of this domain is more extensively discussed in Chapter 4 of this thesis.

The 5' incision by ERCC1-XPF follows the XPG-mediated 3' incision in NER

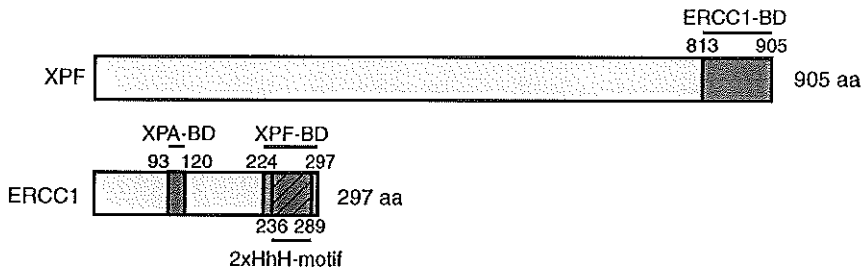


Figure 8. The ERCC1-XPF protein complex. Schematic presentation of functional domains in the ERCC1 and XPF subunits. Protein interaction domains are indicated above (filled boxes), (putative) functional domains based on primary amino acid sequence are indicated below (hatched boxes). *2xHhH-motif*: double helix-hairpin-helix motif.

(123). ERCC1-XPF can be omitted for full open complex formation and 3' incision *in vitro*, and can be added to a preformed incision complex, containing all other factors, to make the 5' incision (41, 123, 126). Thus, unlike XPG, ERCC1-XPF does not appear to have an architectual function in the NER protein-DNA complex. Still, several protein interactions have been reported with ERCC1-XPF, which may account for the positioning of ERCC1-XPF during NER. XPA interacts with the complex (100, 102, 145, 158), mainly via ERCC1, although also a weak affinity for XPF has been reported (7). It has been suggested that RPA and ERCC1 bind sequentially to XPA (158). RPA itself also interacts with ERCC1-XPF, presumably via XPF (7, 112), and this interaction seems particularly important for the positioning of this nuclease complex during the NER reaction. Bound to single-stranded DNA, the 3' oriented side of RPA interacts with ERCC1-XPF and strongly stimulates its nuclease activity, whereas the 5' oriented side of RPA does not interact with the complex and blocks ERCC1-XPF-mediated incisions (see Chapter 6) (de Laat *et al.*, submitted (31)).

2.2.7 **XPE**

XPE is not required for NER reconstituted *in vitro*, but *in vivo* it is, since XP-E patients exhibit mild XP-like skin abnormalities. Cells from XP-E patients are only slightly UV-sensitive and have UDS levels of >50% compared to normal cells (11). The defect in XP-E cells is not unambiguously assigned yet. Some but not all XP-E patients lack a functional damaged DNA-binding (DDB) factor (81, 85, 86, 132). DDB is a heterodimeric protein complex, which, similar to XPA, binds with high affinity to (6-4) photoproducts and modest affinity to CPDs (76, 84, 152), hinting at a function in damage recognition in NER. Possibly, *in vivo* XPE helps remodeling of chromatin at the site of a lesion. A database search revealed the presence of a structural homolog of the large subunit of DDB in humans (van der Spek and Hoeijmakers, unpublished). Functional redundancy may underlie the relatively mild phenotype of XP-E patients. As XPE is apparently not required for the core NER reaction, this factor will be ignored in the models below.

2.3 **The repair synthesis machinery**

The incision and synthesis stages of NER can be separated *in vitro*, and the only factor used in common is RPA, which might remain bound to the undamaged strand to facilitate replication. *In vitro* studies with antibodies and chemical inhibitors against specific DNA polymerases revealed that both DNA polymerases δ and ε (Pol δ and ε) function in NER DNA synthesis (26, 37, 75, 133, 147). A similar observation was made *in vivo* in yeast (16), but the relative contribution of each remains to be determined. The requirement of PCNA (proliferating cell nuclear antigen) is consistent with repair synthesis by these polymerases (169), since it serves as a processivity factor for both, in conjunction with Replication Factor C (RF-C). Indeed, the combination of RPA, PCNA, RF-C (five subunits) and either Pol δ or ε was sufficient for repair synthesis *in vitro*

(170). PCNA can interact with XPG, indicating a possible connection between the incision and replication stage (52). However, PCNA and XPG operate on opposite sides of the repair patch, and more data are required to interpret this link.

DNA synthesis by Pol δ and ε and their cofactors PCNA and RF-C has been studied extensively (for reviews and original references, see (15, 68, 80, 225)). Briefly, RF-C preferentially binds to 3' termini of DNA primers and facilitates the loading of PCNA, which forms a homotrimeric ring-shaped clamp that can track along duplex DNA. This complex serves as a docking platform for both Pol δ and ε, which upon binding form holoenzymes with the cofactors that efficiently can replicate single-stranded DNA.

Noteworthy, PCNA might serve as a mediator between cell cycle control and DNA repair. It interacts with p21, a cyclin-dependent kinase (cdk) inhibitor that is upregulated in a p53-dependent manner upon DNA damage (210). This interaction blocks DNA replication, but does not affect DNA repair, perhaps providing a rationale for the observation that DNA damage induces a replicational arrest to allow repair *in vivo* (103). Also, PCNA interacts with Gadd45, another p53-regulated DNA damage-inducible protein (64, 173). One report claims that Gadd45 inhibits cell cycle entry into S phase and stimulates nucleotide excision repair (173). However, the latter observation could not be confirmed by others (82, 83), so more data are required to determine the functional relevance of this interaction.

The final step in NER is ligation of the 5' end of the newly synthesized patch to the original sequence. This step is thought to be carried out by DNA ligase I. Interestingly, a single case of DNA ligase I-deficiency in humans has been described, and this patient not only suffered from symptoms arising from (mild) defects in semi-conservative replication, but also showed increased sensitivity to several DNA damaging agents, including UV-light (6, 148) (For review on mammalian DNA ligases, see (186)).

2.4 The NER incision stage

In this paragraph, the characteristics of the repair factors described above will serve as a basis to build a model for the individual steps that lead to NER incisions. Historically, dissection of the NER reaction started when Huang *et al.* (1992) elegantly demonstrated that nucleotide excision repair removes damage as part of a 27- to 29-nucleotide oligomer, a finding that established that repair incisions and repair synthesis can be studied separately *in vitro* (73). Reconstitution of the NER reaction with purified protein factors was a subsequent breakthrough which allowed the definition of a minimal set of proteins required for these two NER stages (1, 123, 124). The repair synthesis stage merely involves general replication factors and their action has been discussed above. Here, I will focus on the incision stage of NER. Attempts to further dissect the incision stage by sequentially assembling repair factors onto a defined DNA lesion thusfar failed, and all six factors mentioned above are required for dual incision. The general assumption that repair incision entails (i) damage recognition, (ii) lesion demarcation, and (iii) dual incision (69), was merely based on the individual properties of the proteins involved. Only recently, studies on damage recognition, open complex formation and nuclease positioning have confirmed this assumption and allow a more detailed interpretation of the individual steps that lead to repair incisions.

2.4.1 Damage recognition

Based on its increased affinity for damaged DNA and multiple interactions with other NER proteins, XPA has long been considered the initiating factor of nucleotide excision repair. Recently however, the XPC-hHR23B complex was demonstrated to function even prior to XPA in *in vitro* nucleotide excision repair, and is able to recruit the rest of the repair protein machinery to the damaged site in global genome repair (Sugasawa *et al.*, submitted (176)). In agreement, DNase I footprinting assays revealed

specific binding of the XPC-hHR23B complex to a DNA lesion (Sugasawa *et al.*, submitted (176)), but could not reveal specific binding of XPA, RPA, or the combination of the two, to DNA damage (126).

How does XPC-hHR23B recruit other repair factors in GGR? Evans *et al.* reported that XPC and TFIIH are the only factors absolutely required for any helix distortion around the lesion (41). Helix disruption at a lesion is sufficient for the rest of the repair machinery to act, since locally premelted lesions are efficiently repaired in the absence of XPC-hHR23B (125, 126). Thus, XPC-hHR23B may slightly increase single-strandedness at a damaged site to facilitate entering of other repair factors. Perhaps increased damage-induced helical distortion also underlies the observation that certain cholesterol lesions are repaired in the absence of XPC-hHR23B (123). In addition, XPC-hHR23B may recruit other repair factors through specific protein-protein interactions. In that respect, mammalian XPC-hHR23B has only been reported to interact (weakly) with TFIIH (36). In yeast, Rad4 (XPC)-TFIIH, Rad23-TFIIH and Rad23-Rad14 (XPA) interactions have been observed (5, 61).

Lesions in actively transcribed DNA strands are not expected to be structurally different from those elsewhere in the genome. Nevertheless, XPC-hHR23B is not involved in transcription-coupled repair. This suggests that other factors trace such lesions faster than XPC-hHR23B does, and provide a DNA substrate that can be processed by the rest of the repair machinery. The fact that elongating RNA polymerase II is blocked by lesions in the transcribed strand makes this enzyme an efficient damage sensor (35, 65). The transcription bubble present at the lesion can serve as a substrate for XPC-hHR23B-independent repair (65, 125, 126). We propose that XPC-hHR23B competes with elongating Pol II for detecting lesions in actively transcribed strands; depending on their damage detection rates, lesions will be repaired by either global genome repair or transcription-coupled repair, respectively. In agreement with this model, removal of (6-4) photoproducts by GGR is fast (117), and TCR does not contribute significantly (198). CPDs, on the other hand, are repaired more slowly by GGR, and TCR accounts for faster repair of these lesions from actively transcribed strands

(114, 116).

If XPA does not act as the initial damage sensor, what is the function of its increased affinity for damaged DNA in NER? XPA can interact with many NER factors, including TFIIH, RPA and ERCC1-XPF, indicating a coordinating role in the DNA-protein repair complex (see e.g. table 2). Single-stranded DNA was found to compete 4-fold more effectively than double-stranded DNA for binding of XPA to damaged DNA, and Jones *et al.* proposed that the single-stranded character of a lesion may be the main determinant of XPA-binding (79). In agreement, repair of premelted damaged substrates requires XPA, indicating that XPA is able to bind to lesions outside the helical context (125, 126). XPA may therefore play a role in positioning the repair machinery around the DNA lesion, and it may do so by binding to lesions exposed in a single-stranded context. Additionally, XPA damage-binding can serve to verify NER lesions in TCR (see also paragraph 2.5).

2.4.2 **Lesion demarcation**

Once lesions have been traced, an open DNA complex is formed by the coordinated activities of XPC-hHR23B, TFIIH, XPA and RPA (41, 126). Opening is dependent on the presence of ATP, strongly arguing that the helicases of TFIIH are actively involved (40). The nuclease ERCC1-XPF is not required for opening, but XPG seems to stabilize this opened preincision complex (40, 41, 126, 211). The fully opened repair complex is formed asymetrically around the lesion, with the 5' border further away from the lesion than the 3' border. Evans *et al.* (1997) dissected this process by performing permanganate footprinting studies on damaged DNA substrates in the presence of (repair-deficient) cell extracts. Their data suggest that XPC-hHR23B and TFIIH together are required for an initial opening of <10 nucleotides, and that addition of RPA, XPA and XPG is needed to obtain full opening of approximately 25 nucleotides (40, 41). Using similar techniques but with purified factors rather than cell extracts, Mu *et al.*

Table 2. Physical interactions between NER proteins

	XPA	XPC-hHR23B	TFIIH	RPA	XPG	ERCC1-XPF
XPA			PD (134, 144)	PD, CP (66, 101,158)		PD, 2H (100,102, 145,158)
XPC -hHR23B			CP (36)			
TFIIH	PD (134, 144)	CP (36)			CP, IP (124, 77)	
RPA	PD, CP (66, 101,158)				PD (66,31) **Chapter 6**	PD (7,31) **Chapter 6**
XPG			CP, IP (124, 77)	PD (66,31) **Chapter 6**		
ERCC1-XPF	PD, 2H (100,102, 145,158)			PD (7,31) **Chapter 6**		

PD: Pull-down experiments: (GST-, MBP-) affinity chromatography (techniques using large quantities of 'bait' protein). **CP**: Co-purification or co-precipitation from cell extracts. **2H**: dual hybrid system. **IP**: (Immuno-) co-precipitation of *in vitro* translated gene products (low protein concentrations). **Numbers** indicate relevant references.

(1997) found that all 4 preincision factors are necessary and sufficient to obtain an intermediate opening of 10 to 20 nucleotides positioned rather symmetrically around the lesion; full opening was only observed in conjunction with a dual incision (126). Whether these discrepancies reflect differences in experimental procedures or merely are caused by limited sensitivity due to permanganate acting only on thymine residues, remains to be seen. Both studies, however, indicate a two-step unwinding model with an ATP-dependent TFIIH-mediated initial opening and a subsequent extension of the open complex 5' away from the lesion (see also (41)). Such a mechanism would be analogous to TFIIH-dependent promoter opening in transcription initiation (71, 229).

How do the various factors contribute to opening? (i) TFIIH harbors the two oppositely directed energy-dependent helicase subunits XPB and XPD, and is probably

the motor driving the strand separation. In transcription initiation, TFIIH-dependent opening spans initially ~10 nucleotides (71), which is similar to the size of initial opening in repair (41, 126). The fact that opening is restricted to 10 (or 10 to 20) nucleotides may reflect an intrinsic limitation of TFIIH-mediated strand separation. Alternatively, in repair the encountering of a lesion might prevent TFIIH from further unwinding, as was observed for the helicase activity of purified Rad3, the *S. cerevisiae* homolog of XPD (129, 130). In addition to unwinding, TFIIH has an a structural role in the preincision complex, as premelted lesions still require TFIIH for repair (125, 126). (ii) XPA might account for correct positioning of the opened preincision complex, since it can bind the DNA adduct and interacts with both TFIIH and RPA (see above). (iii) RPA may stabilize the unwound DNA intermediate. Data presented in this thesis pinpoint RPA to the undamaged strand in repair (see Chapter 6) (de Laat *et al.*, submitted (31)), and it is tempting to implicate the single-stranded DNA-binding characteristics of RPA in the creation of a full open repair complex. The 5' oriented side of RPA contains a strong DNA-binding domain which accounts for initial association to 8-10 nt DNA regions (de Laat *et al.*, submitted (31)) (10). Stable binding of RPA to DNA requires ~30 nt single-stranded regions (90). Interestingly, initial opening in NER occurs rather symmetrically around the lesion and exposes ~10 nt of the undamaged strand (41, 126), thus creating a perfect docking site for the 5' oriented side of RPA. We propose that subsequent RPA-stretching in the 3' direction contributes to the formation of a fully opened complex, which indeed is expected to extend in that direction and encompass ~30 nucleotides. Possibly, RPA's interaction with XPG serves to define the 5' border of its 30 nt binding region (see Chapter 6) (de Laat *et al.*, submitted (31)).

It is not known whether repair *in vivo* involves sequential assembly of individual factors or loading of a complete 'repairosome' onto a DNA lesion. In either case, the repair factors are likely to act in a defined order. It is interesting to note that in mammals only TFIIH has been reported to interact with the repair recruitment factor XPC-hHR23B (36). Since XPC and TFIIH are the only factors indispensable for any distortion around a lesion, Evans *et al.* (1997) proposed that XPC-hHR23B and TFIIH may

accomplish initial repair opening (41). Thus, TFIIH may well be the second factor acting at the site of damage. This would imply that TFIIH facilitates the recruitment of XPA to the lesion, rather than the other way around (134, 144). Although merely speculation, TFIIH also seems an attractive candidate to be the first 'repair' factor acting in the XPC-hHR23B-independent TCR pathway; TFIIH has been shown to re-enter early stalled Pol II complexes (39) and is known to interact with both transcription (TFIIE, Pol II?) (49, 140, 141) and repair factors (XPA, XPG) (77, 124, 134, 144).

2.4.3 **Dual repair incision**

Following lesion demarcation, the actual incisions are made by the structure-specific endonucleases XPG and ERCC1-XPF. XPG makes the 3' incision and ERCC1-XPF makes the 5' incision (111, 137, 171). Incisions are made asymetrically around the lesion, with the 3' incision 2-8 nt, and the 5' incision 15-24 nt away from the lesion, corresponding indeed to the borders of the open complex (40, 73, 119). The exact incision positions seem to depend on the type of lesion (111, 119, 179). Although incisions occur near-synchronously, consensus exists that the 3' incision precedes the 5' incision (123). In agreement with this order, XPG-mediated cleavage can be detected in the absence of ERCC1-XPF, but ERCC1-XPF incision activity requires the structural presence, but not the catalytic activity, of XPG (123, 126, 211). Also, limited opening of 10-20 nt is sufficient for XPG cleavage, whereas ERCC1-XPF cutting in NER requires full opening of 25-30 nt (41, 123, 126).

In principle, XPG and ERCC1-XPF are able to cut both strands of an opened DNA intermediate (30, 137, 171), but during repair the nucleases are directed to the damaged strand only. What factors account for the positioning of XPG and ERCC1-XPF? As outlined, RPA plays a crucial role in nuclease positioning. Each side of this molecule, oriented on single-stranded DNA, interacts with a distinct nuclease. In fact, bound to the undamaged strand, RPA alone is sufficient to confer strand specificity to

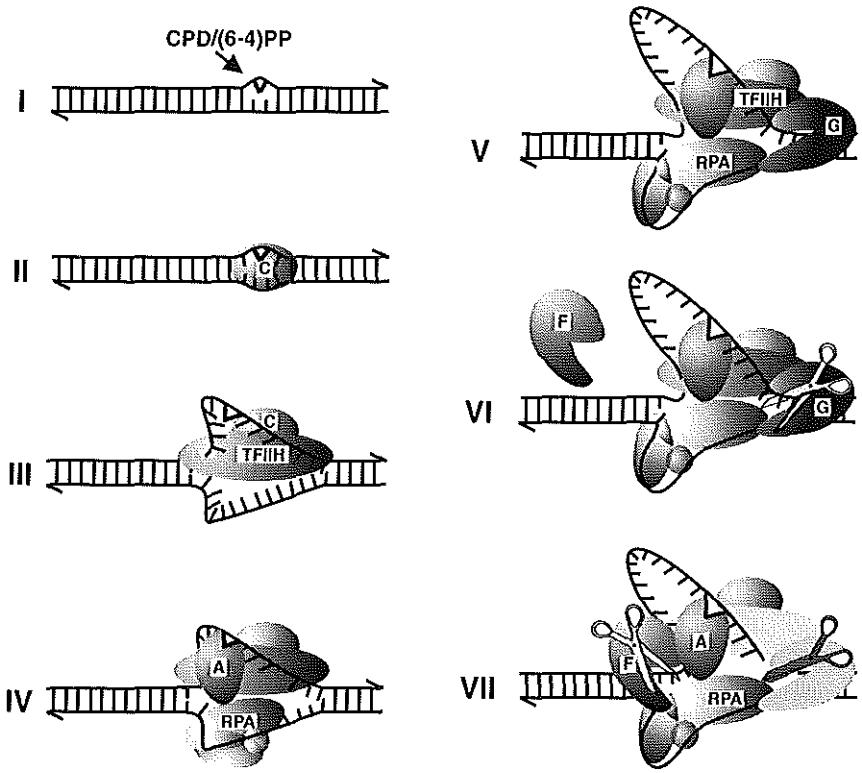


Figure 9. Molecular model for the incision stage of nucleotide excision repair. [I] Typical NER lesions cause local distortion of the DNA helix. [II] XPC-hHR23B (C) senses the lesion, leading to additional disruption around the lesion. [III] XPC-hHR23B at lesion attracts TFIIH, which creates a 10-20 nt opened DNA complex around the lesion by virtue of its helicases XPB and XPD; this step requires ATP. [IV] XPA (A) and RPA (its 3 subunits are depicted) stabilize the 10-20 nt opening; XPA binds to the damaged nucleotides and positions other factors. RPA binds in its 8-10 nt binding mode to the undamaged strand. Step III and IV may occur concomitantly. [V] Creation of a fully opened DNA complex (25-30 nt); we propose that RPA's transition to a 30 nt binding mode (RPA 'stretching') plays an important role in full open complex formation. XPG (G) stabilizes open complex. TFIIH (ATP?) may also actively be involved in this step. (VI) XPG, positioned by TFIIH and RPA, makes the 3' incision. Step V and VI may occur concomitantly. (VII) ERCC1-XPF (F), positioned by RPA and XPA, makes the 5' incision. Dual incision is followed by gap-filling DNA synthesis and ligation (not shown). Drawn contacts between molecules reflect reported protein-protein interactions.

ERCC1-XPF-mediated incisions during NER. XPA's interaction with both RPA and ERCC1-XPF may facilitate or stabilize the positioning of ERCC1-XPF onto the damaged strand (100, 102, 145, 158).

RPA presumably contributes, but is not sufficient to confer strand specificity to XPG. For one, RPA is not found to inhibit XPG incisions in the undamaged strand. Second, XPG incisions in the damaged strand are not stimulated by RPA, despite specific protein-protein contacts. We proposed that RPA, instead of XPG, might have more profit from this interaction, by using it as the 5' border of its 30 nt binding region (de Laat *et al.*, submitted (31)) (Chapter 6). Other factors can be anticipated to contribute to the positioning of XPG during NER. Given the intimate link between XPG and TFIIH, demonstrated not only by the observed physical interaction between the two (5, 62, 77), but also reflected by the CS features that are commonly shared by mutated forms of both factors, TFIIH is an attractive candidate to be involved in XPG positioning as well (see also paragraph 2.5).

The 5' incision by ERCC1-XPF, which completes the incision stage of nucleotide excision repair, leaves an OH-group at the 3' terminus of the primer strand and principally no additional DNA modifications are required to start DNA synthesis at this side of the gap (see Chapter 3) (171). *In vitro*, the oligonucleotide containing the damage is found to be released by the NER incision factors in the absence of DNA repair synthesis (123, 126). Presumably, RPA stays attached to the template strand to protect it against undesirable nucleases and to facilitate repair synthesis. Replication of the undamaged strand does not result in strand displacement beyond the patch, but rather stops at the 3' cleavage site. RPA may coordinate this, but perhaps also the XPG-PCNA interaction participates in restricting patch synthesis.

2.5 XP factors in transcription-coupled repair

Some types of lesions are found to be removed faster from actively transcribed template strands than from other parts of the genome, a phenomenon referred to as transcription-coupled repair. Importantly, this not only holds for typical NER lesions such as CPDs, but also for other types of lesions, including oxidative damage such as thymine glycols (72, 96, 97). In general, all lesions that interfere with transcription elongation may well be a substrate for TCR (see also (184, 189)). Cells from classical XP-A and XP-F patients, and from XP-G patients without CS, are defective in TCR of typical NER lesions, but are normal in TCR of oxidative damage (25). On the other hand, cells from CS-A and CS-B patients are defective in transcription-coupled removal of both CPDs and oxidative damage, indicating a general TCR defect (96, 97, 197, 201). Thus, rather than TCR being a subpathway of nucleotide excision repair, NER seems to be one of more repair pathways that can utilize the fast damage-sensing capacity of elongating Pol II.

A number of XP-B, XP-D and XP-G patients display CS features on top of their XP manifestations, and consistent with CS reflecting a total TCR defect, cells from these patients are deficient in the transcription-coupled removal of both UV-induced lesions and oxidative damage (25). This distinguishes TFIIH and XPG from the other core NER factors and indicates an engagement of these enzymes in the coupling of transcription to other repair pathways as well. Hanawalt and Mellon argued that in order to couple NER to transcription, the stalled Pol II complex has to retract or dissociate to allow access of repair proteins to the lesion (65). This model can be expanded to other repair pathways as well. We propose that defects in this process underlie the total TCR defect observed in CS cells, implicating TFIIH and XPG, like CSA and CSB, in the release of Pol II from the damaged site. Possible roles for CSA and CSB, with the latter found to be associated to elongating Pol II complexes (163, 182, 195), have recently been discussed by van Gool *et al.* (1997), and will not be reiterated here (see (196)). However, in parallel with the reflections on their function in global genome repair, it is

interesting to speculate on the role of XPG and TFIIH in TCR.

Irrespective of the type of damage, TCR has to make a dual incision around the lesion. The general requirement of XPG for TCR indicates that all lesions that block transcription are incised at their 3' side by this nuclease. Elongating Pol II complexes track along the template strand in a 3' to 5' direction, and are therefore expected to position a transcription bubble at the 3' side of such lesions. As relatively small bubbles (>5 nt) were found to be sufficient for XPG cleavage (40), these transcription bubbles potentially serve as good substrates for XPG incisions. XPG interacts with TFIIH, which is anticipated to be important for the positioning of this nuclease. In global genome repair, recruitment of TFIIH to the damaged site requires at least XPC-hHR23B, and presumably depends on XPC-hHR23B-mediated DNA distortion and protein interactions. XPG cleavage depends on the formation of a stable open complex, which in global genome repair comprises the activities of XPC-hHR23B, TFIIH, XPA and RPA. A major difference between TCR and GGR is that a stalled Pol II complex already provides a stable DNA opening 3' of the lesion, which may be accessible to TFIIH and XPG in the absence of other NER factors. We propose that a stalled Pol II complex attracts TFIIH and XPG to catalyze a 3' incision independent of the type of lesion that causes the block. The incision enables retraction or dissociation of Pol II, which concomitantly provides access of other repair proteins to the lesion. Protein-protein interactions have been reported *in vitro* between TFIIH and TFIIE (140, 141), isolated TFIIH subunits and CSA (67), and XPG and CSB (77), all possibly relevant for recruitment or coordination of TFIIH and XPG in TCR. Defects in the proposed functions of XPG and TFIIH in TCR will block the release of trapped transcription and cause accumulation of (oxidative) damage, two problems that have been suggested to underlie CS (65, 192, 196). The exposed, 3' incised, lesion is anticipated to be further processed by the relevant repair pathway. If TCR is to be completed by the NER machinery, TFIIH and XPG can facilitate the positioning of XPA, RPA and ERCC1-XPF, which will allow 5' incision. The implication of mismatch repair proteins in TCR-completion of UV-induced lesions (95, 115) may, if correct, be explained by the mispaired nature of these lesions at the trans-

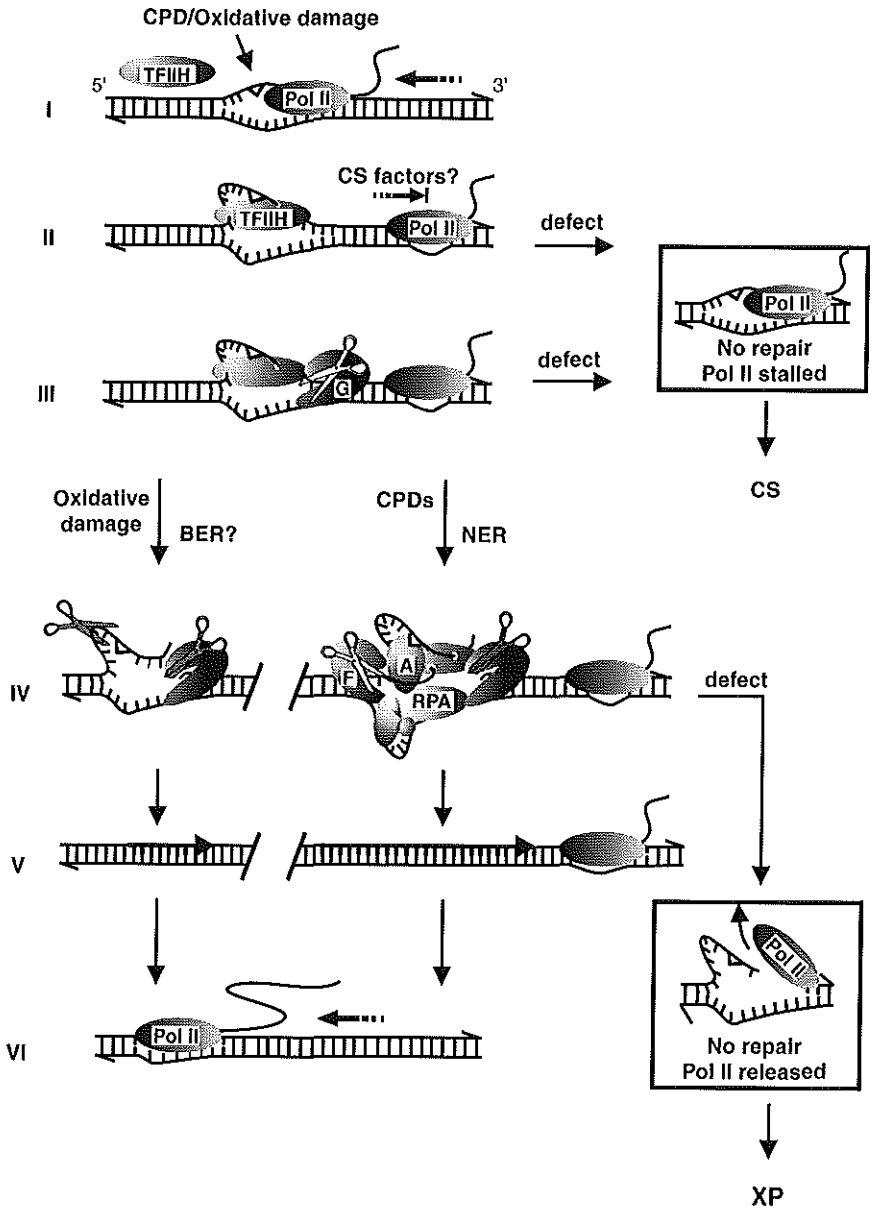

CPD/Oxidative damage
TFIIH
5'
Pol II
3'
I
CS factors?
TFIIH
Pol II
II
defect
Pol II
No repair
Pol II stalled
III
G
defect
CS
Oxidative damage
BER?
CPDs
NER
IV
F
A
RPA
defect
V
Pol II
VI
No repair
Pol II released
XP

Figure 10. Model for XP factors in transcription-coupled repair. [I] Translocating RNA polymerase II (**Pol II**) is blocked by a lesion and attracts TFIIH. Stalled Pol II causes distortion of the helix at site of the lesion. [II] TFIIH (**TFIIH**) enters the partially unwound damaged site, possibly through interactions with other transcription factors or CS factors, and creates an open complex. Pol II retracts from the damaged site; CS-factors may facilitate retraction. [III] TFIIH positions XPG, which makes an incision 3' of the lesion. Defects in step II and III can cause CS features. [IV] Multiple repair pathways can complete TCR, dependent on the type of lesion that caused the transcription block. In the case of oxidative damage, BER may make the 5' incision (left). In the case of CPDs and other typical NER-lesions, XPA (A), RPA (3 subunits depicted) and ERCC1-XPF (F) enter the damaged site, and ERCC1-XPF makes the 5' incision; TFIIH and XPG position these factors (middle panel). Defects in the latter events can lead to XP features (right). [V] DNA repair synthesis. [VI] Reinitiation of Pol II transcription.

cription bubble and the absence of XPC-hHR23B, which normally dedicates these lesions to the NER machinery. In that case, the XPG-mediated incision may serve as the signal required for strand discrimination in mismatch repair.

The classical XP-G patient XP125LO is inactive in GGR, shows no TCR of UV-induced lesions but still displays TCR of oxidative damage (25). The NER-defect of this patient is caused by an Ala-792 to Val missense mutation (135). As this mutation resides next to a presumed catalytic residue, Glu-791, in nuclease domain I of XPG (see paragraph 2.2.5) (135, 165, 166), this XPG protein has been assumed to be totally inactive in cleavage (25, 135, 136, 149). In fact, based on this assumption, Cooper *et al.* suggested that the requirement of XPG for TCR of oxidative damage is independent of its incision activity (25). However, XP125LO has 15% UDS, indicative for residual incision activity (208). Also, Glu-791 has not formally been proven to be essential for cleavage. A corresponding mutation (E160A) in FEN-1, which shares common nuclease domains with XPG, indeed yields a nuclease inactive FEN-1 protein (165, 166). However, this cannot be extrapolated to XPG gratuitously, since for example the D86A substitution in FEN-1 inactivates this nuclease (165, 166), whereas the corresponding D77A mutation in XPG only mildly affects cutting activity (211). Furthermore, a fairly conservative Ala-Val substitution next to a catalytic residue does not necessarily yield a nuclease inactive protein. We anticipate that XP125LO carries XPG protein with residual nuclease activity, which is sufficient to release RNA Pol II in TCR, but fails to fully support removal of UV-induced damage.

2.6 Parallels and discrepancies between human and yeast NER

The general principles of NER have well been conserved from yeast to man, and in fact parallels between these species greatly contributed to rapid progress in the repair field. In this paragraph, the most apparent differences between the yeast and mammalian NER mechanism will briefly be highlighted.

In *S. cerevisiae*, the same minimal set of (homologous) proteins is found to be required for *in vitro* reconstitution of NER (58). Different protein subassemblies are defined though, designated as nucleotide excision repair factors, or NEFs, with NEF1 containing Rad14 (XPA), Rad1 (XPF) and Rad10 (ERCC1) (59), NEF2 containing Rad4 (XPC) and Rad23 (58), and NEF3 containing TFIIH and Rad2 (XPG) (62). The heterotrimer RPA functions as a separate entity (58). NEF1 is *in vitro* preassembled by mixing purified Rad14 with Rad1-Rad10, as only a fraction of Rad14 is found in complex with Rad1-Rad10 in cell extracts (59). NEF2 is purified as such from yeast cells overexpressing Rad4 (58) and NEF3 purification requires the overproduction of Rad2 (62). Thus, the isolation of these entities does not necessarily mean that different NER factor subassemblies exist *in vivo* in yeast and mammals.

A comparison between the *in vitro* reconstituted NER reactions of the two species reveals some mechanistic differences. Whereas in the mammalian system ERCC1-XPF was found to be dispensable for 3' incisions, in yeast the presence of the homologous complex Rad1-Rad10 was absolutely required to detect any cleavage, indicating that Rad1-Rad10 has a structural role in the repair complex not shared by ERCC1-XPF (58, 123). More differences are observed between these two homologous complexes. Both interact with the damage recognition protein XPA (Rad14), however, ERCC1-XPF binds XPA mainly by its ERCC1 subunit (7, 100, 102, 145), whereas Rad1-Rad10 mediates the Rad14-interaction mainly via Rad1 (XPF) (59). The primary sequences of Rad10 and ERCC1 show a striking difference, as ERCC1 contains an additional domain of approximately 80 residues at its C-terminus (194). The relevance of this region in ERCC1 is extensively discussed elsewhere in Chapter 4 of this thesis (see also (172)).

Whereas the 3 NEFs from yeast, together with RPA, are sufficient to detect damage-dependent incisions, a fourth NEF was identified that markedly stimulated cleavage (60). NEF4 comprises the Rad7 and Rad16 proteins, which together bind UV-damaged DNA in an ATP-dependent manner (60). Rad7 and Rad16 both function in global genome repair but are dispensable for transcription-coupled repair (128, 206, 207). Based on sequence homology, Rad16 belongs to the Swi2/Snf2 subfamily of DNA-dependent ATPases, and may therefore be anticipated to function in the remodeling of chromatin. No mammalian homologs of yeast Rad7 and Rad16 have been identified thusfar.

A striking difference between mammalian and yeast NER seems the differential requirement of XPC and its homolog Rad4 for TCR. XPC-hHR23B is dispensable for transcription-coupled repair of RNA Pol II transcribed strands in mammals (202), whereas Rad4-Rad23 is required for such repair in yeast (127, 206). A reciprocal requirement is observed for RNA Pol I transcribed strands; mammalian repair of actively transcribed rDNA strands depends on a functional XPC-hHR23B complex (19), whereas in yeast RNA Pol I-mediated repair occurs in the absence of Rad4-Rad23 (205). Perhaps this discrepancy is fortuitous. As discussed, the damage sensor XPC-hHR23B, and by analogy Rad4-Rad23, is expected to be dispensable when an RNA polymerase detects lesions more rapidly and is able to provide a DNA structure that can be further processed by the rest of the repair machinery. Possibly in mammals only RNA Pol II, and in yeast only RNA Pol I meets these requirements (see also (204)).

As stated, the involvement of mammalian mismatch repair proteins in Pol II-mediated TCR of UV-induced lesions may be explained by the mispaired nature of these lesions and the abolished requirement of XPC-hHR23B. Interestingly, TCR of RNA Pol II transcribed strands in yeast requires Rad4-Rad23 but does not involve mismatch repair proteins (180). We predict that the involvement of XPC-hHR23B and its yeast homolog dedicate lesions to the NER machinery, and that only in their absence mismatch repair proteins are able to participate in TCR of UV-induced lesions.

2.7 Final considerations

Over the past few years, insight into the core NER reaction has increased enormously. On top of the availability of defined mutants in both yeast and man, purified protein factors, site-specific DNA lesions, artificial DNA substrates and sensitive methods to trap reaction intermediates all contributed to the detailed knowledge we have todate on the mechanism of the NER reaction. However, many fundamental issues are still unresolved. Some of the ones directly related to the presented models will briefly be addressed in this paragraph.

Multiple proteins have been implicated in damage recognition, including XPC-hHR23B, XPA, XPE, possible human homologs of Rad7-Rad16, and the RNA polymerases in transcription-coupled repair. An important challenge will be to determine how these individual factors function and cooperate in NER. Does XPA, for example, preferentially bind to lesions exposed in a single-stranded DNA surrounding, as suggested by our model? Probably studies on naked DNA are insufficient and chromatin needs to be included to fully appreciate the contributions of all these factors to NER.

Whereas attempts to sequentially assemble proteins on a DNA lesion could not reveal an order of enzymatic activities, the recent development of a sophisticated competition assay has lead to the surprising finding that *in vitro* XPC-hHR23B is the first factor to act during NER (Sugasawa *et al.*, submitted (176)). Possibly, this same type of assay will also allow the identification of the next factor recruited to the lesion. As outlined, TFIIH is considered an attractive candidate. However, XPA, RPA and TFIIH may also concomitantly be needed to further demarcate the lesion.

Lesion demarcation is still a poorly understood step in the NER reaction. The generation of ATPase-inactive XPB and XPD mutants in the context of TFIIH may provide direct evidence that TFIIH is the motor driving open complex formation. Also, such mutants seem ideal to test whether repair opening, analogous to promoter opening, occurs in two steps, and, if so, whether TFIIH helicase activity is only required for 10-20 nt opening or also for full open complex formation (see also (41)).

Mutant proteins specifically defective in protein-protein interactions should allow a better appreciation of the functional significance of a given interaction and may provide insight into the structural role of each factor in the repair complex. It will be interesting to see, for example, whether and how XPA and TFIIH contribute to the positioning of ERCC1-XPF and XPG, respectively.

TFIIH and XPG were postulated here to play a crucial role in an early stage of transcription-coupled repair, and a challenge will be to determine this function. Introduction of an established active site mutant of XPG (e.g. XPG D812A (211)) into XP-G cells with combined XP/CS should address whether general TCR depends on the structural presence (25) or the catalytic activity (this thesis) of XPG. In the former case, this mutant should correct the TCR defect of oxidative damage, whereas our model predicts the opposite.

ERCC1-XPF cleavage 5' of the lesion completes the incision stage of NER. This cut also marks the initiation site of DNA repair synthesis. It will be interesting to know how the stages are coupled. Also, *in vitro* excision of the damaged oligonucleotide was found to be accomplished by the NER incision machinery and coincides with the release of a subset of repair factors (123, 126). Assuming that RPA remains bound to the undamaged template strand to coordinate subsequent DNA synthesis, protein release likely involves the interruption of RPA-contacts with XPA, XPG and ERCC1-XPF. An active-site mutant of ERCC1-XPF may determine whether ERCC1-XPF cleavage or -binding is required for triggering the release of repair factors from the site of damage.

Finally, a major challenge will be to understand the NER process in the context of chromatin. Compaction of DNA into nucleosomes and higher order structures will surely affect the accessibility of lesions. In agreement, repair in the non-transcribed strand of a gene was found to be fast in linker DNA and slow in internal protected regions of the nucleosome, whereas repair of the transcribed strand was independent of chromatin organization *in vivo* (218). Given the high degree of organization and the crowdedness in the nucleus, random scanning of DNA damage sensors such as XPC-hHR23B along DNA templates seems very unlikely. Repair rates probably depend on

both the concentration of repair factors and their affinity for lesions. For some of the basal repair factors we estimated the presence of 10^4-10^5 molecules per nucleus (unpublished observations). This would indicate that one repair molecule, or one repair complex, still has to guard 10^4-10^5 base pairs of DNA. As repair *in vivo* is highly efficient, repair proteins can be anticipated to act in a highly coordinated fashion in the context of chromatin. Also, lesions, and the repair of lesions, are expected to disrupt the genomic integrity (2). Restoration of nucleosome formation was found to initiate at target sites for NER and appeared to propagate bidirectionally, indicating a still poorly defined link between repair and DNA compaction (50, 51). Thus, although *in vitro* the principle mechanisms of NER seem fairly well characterized, we are still far from understanding the dynamics and requirements of this lifeguarding process in the context of a living cell.

References

1. **Aboussekhra, A., M. Biggerstaff, M.K.K. Shivji, J.A. Vilpo, V. Moncollin, V.N. Podust, M. Protic, U. Hubscher, J.-M. Egly, and R.D. Wood.** 1995. Mammalian DNA nucleotide excision repair reconstituted with purified components. *Cell* **80:** 859-868.
2. **Althaus, F.** 1992. Poly ADP-ribosylation: a histone shuttle mechanism in DNA excision repair. *J. Cell. Sci.* **102 (Pt 4):** 663-670.
3. **Ariza, R., S. Keyse, J. Moggs, and R.D. Wood.** 1996. Reversible protein phosphorylation modulates nucleotide excision repair of damaged DNA by human cell extracts. *Nucleic Acids Res.* **24:** 433-440.
4. **Asahina, H., I. Kuraoka, M. Shirakawa, E.H. Morita, N. Miura, I. Miyamoto, E. Ohtsuka, Y. Okada, and K. Tanaka.** 1994. The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage. *Mutat. Res.* **315:** 229-237.
5. **Bardwell, A.J., L. Bardwell, N. Iyer, J.Q. Svejstrup, W.J. Feaver, R.D. Kornberg, and E.C. Friedberg.** 1994. Yeast nucleotide excision repair proteins rad2 and rad4 interact with RNA polymerase II basal transcription factor b (TFIIH). *Mol. Cell. Biol.* **14:** 3569-3576.
6. **Barnes, D.E., A.E. Tomkinson, A.R. Lehmann, A.D.B. Webster, and T. Lindahl.** 1992. Mutations in the *DNA ligase I* gene of an individual with immunodeficienies and cellular hypersensitivity to DNA-damaging agents. *Cell* **69:** 495-503.
7. **Bessho, T., A. Sancar, L.H. Thompson, and M.P. Thelen.** 1997. Reconstitution of human excision nuclease with recombinant XPF-ERCC1complex. *J. Biol. Chem.* **272:** 3833-3837.
8. **Biggerstaff, M., D.E. Szymkowski, and R.D. Wood.** 1993. Co-correction of ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects *in vitro*. *EMBO J.* **12:** 3685-3692.
9. **Blackwell, L.J., and J.A. Borowiec.** 1994. Human replication protein A binds single-stranded DNA in two distinct complexes. *Mol. Cell. Biol.* **14:** 3993-4001.
10. **Blackwell, L.J., J.A. Borowiec, and I.A. Masrangelo.** 1996. Single-stranded-DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase. *Mol. Cell. Biol.* **16:** 4798-4807.
11. **Bootsma, D., K. Kraemer, J. Cleaver, and J. Hoeijmakers.** 1998. Nucleotide Excision Repair Syndromes: Xeroderma Pigmentosum, Cockayne Syndrome, and Trichothiodystrophy, p. 245-274. *In* B. Vogelstein, and K. Kinzler (ed.), The Genetic Basin of Human Cancer, vol. Chapter 3. McGraw-Hill Book Co., New York.
12. **Brill, S.J., and B. Stillman.** 1989. Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication. *Nature* **342:** 92-95.
13. **Brookman, K., J. Lamerdin, M. Thelen, M. Hwang, J. Reardon, A. Sancar, Z. Zhou, C. Walter, C. Parris, and L. Thompson.** 1996. ERCC4 (XPF) encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Mol Cell Biol* **16:** 6553-6562.
14. **Buchko, G., and M. Kennedy.** 1997. Human nucleotide excision repair protein XPA: 1H NMR and CD solution studies of a synthetic peptide fragment corresponding to the zinc-binding domain (101-141). *J. Biomol. Struct. Dyn.* **14:**677-690.
15. **Budd, M., and J. Campbell.** 1997. The roles of eukaryotic DNA polymerases in DNA repair synthesis. *Mut Res* **384:**157-167.
16. **Budd, M.E., and J.L. Campbell.** 1995. DNA polymerases required for repair of UV-induced

damage in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **15:** 2173-2179.

17. **Burns, J., S. Guzder, P. Sung, S. Prakash, and L. Prakash.** 1996. An affinity of human replication protein A for ultraviolet-damaged DNA. *J. Biol. Chem.* **271:** 11607-11610.
18. **Carty, M.P., M. Zernik-Kobak, S. McGrath, and K. Dixon.** 1994. UV light-induced DNA synthesis arrest in HeLa cells is associated with changes in phosphorylation of human single-stranded DNA-binding protein. *EMBO J.* **13:** 2114-2123.
19. **Christians, F., and P. Hanawalt.** 1994. Repair in ribosomal RNA genes is deficient in xeroderma pigmentosum group C and in Cockayne's syndrome cells. *Mutat. Res.* **323:** 179-187.
20. **Chu, G., and L. Mayne.** 1996. Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy: do the genes explain the diseases? *Trends in Gen.* **12:** 187-192.
21. **Cloud, K., B. Shen, G. Strniste, and M. Park.** 1995. XPG protein has a structure-specific endonuclease activity. *Mutat. Res.* **347:** 55-60.
22. **Clugston, C., K. McLaughlin, M. Kenny, and R. Brown.** 1992. Binding of human single-stranded DNA binding protein to DNA damaged by the anticancer drug cis-diamminedichloroplatinum (II). *Cancer Res.* **52:** 6375-6379.
23. **Collins, A.R.** 1993. Mutant rodent cell lines sensitive to ultraviolet light, ionizing radiation and cross-linking agents - A comprehensive survey of genetic and biochemical characteristics. *Mutat. Res.* **293:** 99-118.
24. **Conaway, R.C., and J.W. Conaway.** 1989. An RNA polymerase II transcription factor has an associated DNA-dependent ATPase (dATPase) activity strongly stimulated by the TATA region of promoters. *Proc. Natl. Acad. Sci.* **86:** 7356-7360.
25. **Cooper, P., T. Nouspikel, S. Clarkson, and S. Leadon.** 1997. Defective transcription-coupled repair of oxidative base damage in cockayne syndrome patients from XP group G. *Science* **275:** 990-3.
26. **Coverley, D., M.K. Kenny, D.P. Lane, and R.D. Wood.** 1992. A role for the human single-stranded DNA binding protein HSSB/RPA in an early stage of nucleotide excision repair. *Nucl. Acids Res.* **20:** 3873-3880.
27. **Coverley, D., M.K. Kenny, M. Munn, W.D. Rupp, D.P. Lane, and R.D. Wood.** 1991. Requirement for the replication protein SSB in human DNA excision repair. *Nature* **349:** 538-541.
28. **Dahmus, M.E.** 1995. Phosphorylation of C-terminal domain of RNA polymerase II. *Bioch. Biophys. Acta* **1261:** 71-82.
29. **de Boer, J., J. de Wit, H. van Steeg, R.J.W. Berg, H. Morreau, P. Visser, A.R. Lehmann, M. Duran, J.H.J. Hoeijmakers, and G. Weeda.** 1998. mouse model for the basal transcription/ DNA repair syndrome trichothiodystrophy. *Mol. Cell. In press.*
30. **de Laat, W.L., E. Appeldoorn, N.G.J. Jaspers, and J.H.J. Hoeijmakers.** 1998. DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.* **273:** 7835-7842.
31. **de Laat, W.L., E. Appeldoorn, K. Sugasawa, N.G.J. Jaspers, and J.H.J. Hoeijmakers.** 1998. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Submitted*.
32. **De Vries, A., C.T.M. van Oostrom, F.M.A. Hofhuis, P.M. Dortant, R.J.W. Berg, F.R. de Gruijl, P.W. Wester, C.F. van Kreijl, P.J.A. Capel, H. van Steeg, and S.J. Verbeek.** 1995. Increased susceptibility to ultraviolet-B and carcinogens of mice lacking the DNA excision repair gene *XPA*. *Nature* **377:** 169-173.
33. **Din, S., S.J. Brill, M.P. Fairman, and B. Stillman.** 1990. Cell-cycle-regulated phosphorylation

of DNA replication factor A from human and yeast cells. *Genes Dev.* **4**: 968-977.

34. **Doherty, A.S., 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.
35. **Donahue, B.A., S. Yin, J.-S. Taylor, D. Reines, and P.C. Hanawalt.** 1994. Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc. Natl. Acad. Sci.* **91**:8502-8506.
36. **Drapkin, R., J.T. Reardon, A. Ansari, J.C. Huang, L. Zawel, K. Ahn, A. Sancar, and D. Reinberg.** 1994. Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* **368**: 769-772.
37. **Dresler, S., and M. Frattini.** 1986. DNA replication and UV-induced DNA repair synthesis in human fibroblasts are much less sensitive than DNA polymerase alpha to inhibition by butylphenyl-deoxyguanosine triphosphate. *Nucleic Acids Res.* **14**: 7093-7102.
38. **Dvir, A., R. Conaway, and J. Conaway.** 1996. Promoter escape by RNA polymerase II. A role for an ATP cofactor in suppression of arrest by polymerase at promoter-proximal sites. *J. Biol. Chem.* **271**: 23352-23356.
39. **Dvir, A., R. Conaway, and J. Conaway.** 1997. A role for TFIIH in controlling the activity of early RNA polymerase II elongation complexes. *Proc. Natl. Acad. Sci.* **94**: 9006-9010.
40. **Evans, E., J. Fellows, A. Coffer, and R.D. Wood.** 1997. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J.* **16**: 625-38.
41. **Evans, E., J. Moggs, J. Hwang, J. Egly, and R. Wood.** 1997. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J.* **16**: 6559-6573.
42. **Fairman, M.P., and B. Stillman.** 1988. Cellular factors required for multiple stages of SV40 DNA replication in vitro. *EMBO J.* **7**: 1211-1218.
43. **Fang, F., and J. Newport.** 1993. Distinct roles of cdk2 and cdc2 in RP-A phosphorylation during the cell cycle. *J. Cell. Sci.* **106 (Pt 3)**: 983-994.
44. **Feaver, W.J., O. Gileadi, and D. Kornberg.** 1991. Purification and characterization of yeast RNA polymerase II transcription factor b. *J. Biol. Chem.* **266**: 19000-19005.
45. **Feaver, W.J., J.Q. Svejstrup, L. Bardwell, A.J. Bardwell, S. Buratowski, K.D. Gulyas, T.F. Donahue, E.C. Friedberg, and R.D. Kornberg.** 1993. Dual roles of a multiprotein complex from S. cerevisiae in transcription and DNA repair. *Cell* **75**: 1379-1387.
46. **Feaver, W.J., J.Q. Svejstrup, N.L. Henry, and R.D. Kornberg.** 1994. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell* **79**: 1103-1109.
47. **Fischer, L., M. Gerard, C. Chalut, Y. Lutz, S. Humbert, M. Kanno, P. Chambon, and J.-M. Egly.** 1992. Cloning of the 62-kilodalton component of basic transcription factor BTF2. *Science* **257**: 1392-1395.
48. **Fishman-Lobell, J., and J.E. Haber.** 1992. Removal of nonhomologous DNA ends in double-strand break recombination: The role of the yeast ultraviolet repair gene *RAD1*. *Science* **258**: 480-484.
49. **Flores, O., H. Lu, and D. Reinberg.** 1992. Factors involved in specific transcription by mammalian RNA polymerase II. *J.Biol.Chem.* **267**: 2786-2790.
50. **Gaillard, P., E. Martini, P. Kaufman, B. Stillman, E. Moustacchi, and G. Almouzni.** 1996. Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor I. *Cell* **86**: 887-896.

51. **Gaillard, P., J. Moggs, D. Roche, J. Quivy, P. Becker, R. Wood, and G. Almouzni.** 1997. Initiation and bidirectional propagation of chromatin assembly from a target site for nucleotide excision repair. *EMBO J.* **16:** 6281-6289.

52. **Gary, R., D. Ludwig, H. Cornelius, M. MacInnes, and M. Park.** 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.

53. **Georgaki, A., B. Strack, V. Podust, and U. Hubscher.** 1992. DNA unwinding activity of replication protein A. *FEBS Lett* **308:** 240-244.

54. **Gerard, M., L. Fischer, V. Moncollin, J.-M. Chipoulet, P. Chambon, and J.-M. Egly.** 1991. Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *J. Biol. Chem.* **266:** 20940-20945.

55. **Gomes, X., and M. Wold.** 1996. Functional domains of the 70-kilodalton subunit of human replication protein A. *Bioch.* **35:**10558-10568.

56. **Goodrich, J.A., and R. Tjian.** 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* **77:** 145-156.

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

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

59. **Guzder, S., P. Sung, L. Prakash, and S. Prakash.** 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.

60. **Guzder, S., P. Sung, L. Prakash, and S. Prakash.** 1997. Yeast Rad7-Rad16 complex, specific for the nucleotide excision repair of the nontranscribed DNA strand, is an ATP-dependent DNA damage sensor. *J. Biol. Chem.* **272:** 21665-21668.

61. **Guzder, S.A., V. Bailly, P. Sung, L. Prakash, and S. Prakash.** 1995. Yeast DNA repair protein Rad23 promotes complex formation between transcription factor TFIIH and DNA damage recognition factor RAD14. *J. Biol. Chem.* **270:** 8385-8388.

62. **Habraken, Y., P. Sung, S. Prakash, and L. Prakash.** 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.* **93:** 10718-10722.

63. **Habraken, Y.S., P; Prakash, L; Prakash, S.** 1995. Structure-specific nuclease activity in yeast nucleotide excision. *J. Biol.Chem.* **270:** 30194-30198.

64. **Hall, P., J. Kearsey, P. Coates, D. Norman, E. Warbrick, and L. Cox.** 1995. Characterisation of the interaction between PCNA and Gadd45. *Oncogene* **10:** 2427-2433.

65. **Hanawalt, P., and I. Mellon.** 1993. Stranded in an active gene. *Curr. Biol.* **3:** 67-69.

66. **He, Z., L.A. Henricksen, M.S. Wold, and C.J. Ingles.** 1995. RPA involvement in the damage-recognition and incision step of nucleotide excision repair. *Nature* **374:** 566-569.

67. **Henning, K.a., L. Li, N. Iyer, L. McDaniel, M.S. Reagan, R. Legerski, R.A. Schultz, M. Stefanini, A.R. Lehmann, L.V. Mayne, and E.C. Friedberg.** 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.

68. **Hindges, R., and U. Hubscher.** 1997. DNA polymerase delta, an essential enzyme for DNA transactions. *Biol. Chem.* **378:** 345-362.

69. **Hoeijmakers, J.H.J.** 1994. Human nucleotide excision repair syndromes: molecular clues to unexpected intricacies. *Eur. J. of Cancer* **30A:** 1912-1921.

70. **Holstege, F.C.P., D. Tantin, M. Carey, P.C. van der Vliet, and H.T.M. Timmers.** 1995. The requirement for the basal transcription factor IIE is determined by the helical stability of promotor DNA. *EMBO J.* **14:** 810-819.

71. **Holstege, F.C.P., P.C. Van der Vliet, and H.T.M. Timmers.** 1996. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors TFIIE and TFIIH. *EMBO J.* **15:** 1666-1677.

72. **Htun, H., and B. Johnston.** 1992. Mapping adducts of DNA structural probes using transcription and primer extension approaches. *Methods Enzymol.* **212:** 272-294.

73. **Huang, J.C., D.L. Svoboda, J.T. Reardon, and A. Sancar.** 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.* **89:** 3664-3668.

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

75. **Hunting, D., B. Gowans, and S. Dresler.** 1991. DNA polymerase delta mediates excision repair in growing cells damaged with ultraviolet radiation. *Biochem. Cell. Biol.* **69:** 303-308.

76. **Hwang, B.J., and G. Chu.** 1993. Purification and characterization of a human protein that binds to damaged DNA. *Biochem.* **32:** 1657-1666.

77. **Iyer, N., M.S. Reagan, K.-J. Wu, B. Canagarajah, and E.C. Friedberg.** 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. *Biochem.* **35:** 2157-2167.

78. **Jiang, Y., M. Yan, and J. Gralla.** 1996. A three-step pathway of transcription initiation leading to promoter clearance at an activation RNA polymerase II promoter. *Mol Cell Biol* **16:** 1614-1621.

79. **Jones, C.J., and R.D. Wood.** 1993. Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochem.* **32:** 12096-12104.

80. **Jonsson, Z., and U. Hubscher.** 1997. Proliferating cell nuclear antigen: more than a clamp for DNA polymerases. *Bioessays* **19:** 967-975.

81. **Kataoka, H., and Y. Fujiwara.** 1991. UV damage-specific DNA-binding protein in xeroderma pigmentosum complementation group E. *Biochem. Biophys. Res. Comm.* **175:** 1139-1143.

82. **Kazantsev, A., and A. Sancar.** 1995. Does the p53 up-regulated Gadd45 protein have a role in excision repair? *Science* **270:** 1003-1004.

83. **Kearsey, J., P. Coates, A. Prescott, E. Warbrick, and P. Hall.** 1995. Gadd45 is a nuclear cell cycle regulated protein which interacts with p21Cip1. *Oncogene* **11:** 1675-1683.

84. **Keeney, S., G.J. Chang, and S. Linn.** 1993. Characterization of human DNA damage binding protein implicated in xeroderma pigmentosum E. *J. Biol. Chem.* **268:** 21293-21300.

85. **Keeney, S., A.P.M. Eker, T. Brody, W. Vermeulen, D. Bootsma, J.H.J. Hoeijmakers, and S. Linn.** 1994. correction of the DNA repair defect in xeroderma pigmentosum group E by injection of a DNA damage-binding protein. *Proc. Natl. Acad. Sci.* **91:** 4053-4056.

86. **Keeney, S., H. Wein, and S. Linn.** 1992. Biochemical heterogeneity in xeroderma pigmentosum

complementation group E. *Mutat. Res.* **273:** 49-56.

87. **Kenny, M., U. Schlegel, H. Furneaux, and J. Hurwitz.** 1990. The role of human single-stranded DNA binding protein and its individual subunits in simian virus 40 DNA replication. *J. Biol. Chem.* **265:** 7693-7700.

88. **Kenny, M.K., S.H. Lee, and J. Hurwitz.** 1989. Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta. *Proc. Natl. Acad. Sci.* **86:** 9757-9761.

89. **Kim, C., B.F. Paulus, and M.S. Wold.** 1994. Interactions of human replication protein A with oligonucleotides. *Biochem.* **33:** 14197-14206.

90. **Kim, C., R.O. Snyder, and M.S. Wold.** 1992. Binding properties of replication protein-A from human and yeast cells. *Mol. Cell. Biol.* **12:** 3050-3059.

91. **Kim, C., and M. Wold.** 1995. Recombinant human replication protein A binds to polynucleotides with low cooperativity. *Biochem.* **34:** 2058-2064.

92. **Kim, D., E. Stigger, and S. Lee.** 1996. Role of the 70-kDa subunit of human replication protein A (I). Single-stranded dna binding activity, but not polymerase stimulatory activity, is required for DNA replication. *J. Biol. Chem.* **271:** 15124-15129.

93. **Klungland, A., and T. Lindahl.** 1997. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.* **16:** 3341-3348.

94. **Kuraoka, I., E. Morita, M. Saijo, T. Matsuda, K. Morikawa, M. Shirakawa, and K. Tanaka.** 1996. Identification of a damaged-DNA binding domain of the XPA protein. *Mutat. Res.* **362:** 87-95.

95. **Leadon, S., and A. Avrutskaya.** 1997. Differential involvement of the human mismatch repair proteins, hMLH1 and hMSH2, in transcription-coupled repair. *Cancer Res.* **57:** 3784-3791.

96. **Leadon, S.A., and P.K. Cooper.** 1993. Preferential repair of ionizing-radiation induced damage in the transcribed strand of an active human gene is defective in Cockayne syndrome. *Proc.Natl.Acad.Sci.* **90:** 10499-10503.

97. **Leadon, S.A., and D.A. Lawrence.** 1992. Strand-specific repair of DNA damage in the yeast *GAL7* gene requires RNA polymerase II. *J. Biol. Chem.* **267:** 23175-23182.

98. **Lee, S., Z. Pan, A. Kwong, P. Burgers, and J. Hurwitz.** 1991. Synthesis of DNA by DNA polymerase epsilon in vitro. *J. Biol. Chem.* **266:** 22707-22717.

99. **Leveillard, T., L. Andera, N. Bissonnette, L. Schaeffer, L. Bracco, J. Egly, and B. Wasylyk.** 1996. Functional interactions between p53 and the TFIIH complex are affected by tumour-associated mutations. *EMBO J.* **15:**1615-1624.

100. **Li, L., S.J. Elledge, C.A. Peterson, E.S. Bales, and R.J. Legerski.** 1994. Specific association between the human DNA repair proteins XPA and ERCC1. *Proc. Natl. Acad. Sci.* **91:** 5012-5016.

101. **Li, L., X. Lu, C.A. Peterson, and L. R.J.** 1995. An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Mol. Cell. Biol.* **15:** 5396-5402.

102. **Li, L., C.A. Peterson, X. Lu, and R.J. Legerski.** 1995. Mutations in XPA that prevent association with ERRC1 are defective in nuclear excision repair. *Mol. Cell. Biol.* **15:** 1993-1998.

103. **Li, R., S. Waga, G.J. Hannon, D. Beach, and B. Stillman.** 1994. Differential effects by the p21 CDK inhibitor on PCNA-dependent DNA replication and repair. *Nature* **371:** 534-537.

104. **Lieber, M.** 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication,

recombination and repair. *Bioessays* **19:** 233-240.

105. **Liu, V.F., and D.T. Weaver.** 1993. The ionizing radiation-induced replication protein A phosphorylation response differs between ataxia telangiectasia and normal human cells. *Mol. Cell. Biol.* **13:** 7222-7231.

106. **Longhese, M.P., P. Plevani, and G. Lucchini.** 1994. Replication factor A is required *in vivo* for DNA replication, repair, and recombination. *Mol. Cell. Biol.* **14:** 7884-7890.

107. **Marinoni, J., R. Roy, W. Vermeulen, P. Miniou, Y. Lutz, G. Weeda, T. Seroz, D. Gomez, J. Hoeijmakers, and J. Egly.** 1997. Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH. *EMBO J.* **16:** 1093-1102.

108. **Masutani, C., M. Araki, K. Sugasawa, P. van der Spek, A. Yamada, A. Uchida, T. Maekawa, D. Bootsma, J. Hoeijmakers, and F. Hanaoka.** 1997. Identification and characterization of XPC-binding domain of hHR23B. *Mol. Cell. Biol.* **17:** 6915-6923.

109. **Masutani, C., K. Sugasawa, J. Yanagisawa, T. Sonoyama, M. Ui, T. Enomoto, K. Takio, K. Tanaka, P.J. van der Spek, D. Bootsma, J.H.J. Hoeijmakers, and F. Hanaoka.** 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.

110. **Matsumoto, K., and Y. Ishimi.** 1994. Single-stranded-DNA-binding protein-dependent DNA unwinding of the yeast ARS1 region. *Mol. Cell. Biol.* **14:** 4624-4632.

111. **Matsunaga, T., D. Mu, C.-H. Park, J.T. Reardon, and A. Sancar.** 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.

112. **Matsunaga, T., C.H. Park, T. Bessho, D. Mu, and A. Sancar.** 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.

113. **McWhir, J., J. Seldridge, D.J. Harrison, S. Squires, and D.W. Melton.** 1993. Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat. Gen.* **5:** 217-223.

114. **Mellon, I., and P.C. Hanawalt.** 1989. Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* **342:** 95-98.

115. **Mellon, I., D.K. Rajpal, M. Koi, C.R. Boland, and G.N. Champe.** 1996. transcription-coupled repair deficiency and mutations in human mismatch repair genes. *Science* **272:** 557-560.

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

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

118. **Miyamoto, I., N. Miura, H. Niwa, J. Miyazaki, and K. Tanaka.** 1992. Mutational analysis of the structure and function of the xeroderma pigmentosum group A complementing protein. *J. Biol. Chem.* **267:**12182-12187.

119. **Moggs, J.G., K.J. Yarema, J.M. Essigmann, and R.D. Wood.** 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.

120. **Moolenaar, G., R. Uiterkamp, D. Zwijnenburg, and N. Goosen.** 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.

121. **Morita, E., T. Ohkubo, I. Kuraoka, M. Shirakawa, K. Tanaka, and K. Morikawa.** 1996. Implications of the zinc-finger motif found in the DNA-binding domain of the human XPA protein. *Genes Cells* **1:** 437-442.

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

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

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

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

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

127. **Mueller, J., and M. Smerdon.** 1996. Rad23 is required for transcription-coupled repair and efficient overrall repair in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **16:** 2361-2368.

128. **Mueller, J., and M. Smerdon.** 1995. Repair of plasmid and genomic DNA in a rad7 delta mutant of yeast. *Nucleic Acids Res.* **23:** 3457-3464.

129. **Naegeli, H., L. Bardwell, and E. Friedberg.** 1993. Inhibition of Rad3 helicase activity by DNA adducts and abasic sites: implications for the role of a DNA helicase in damage-specific incision of DNA. *Biochem.* **32:** 613-621.

130. **Naegeli, H., L. Bardwell, and E.C. Friedberg.** 1992. The DNA helicase and adenosine triphosphatase activities of yeast Rad3 protein are inhibited by DNA damage. A potential mechanism for damage-specific recognition. *J. Biol. Chem.* **267:** 392-398.

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

132. **Nichols, A., P. Ong, and S. Linn.** 1996. Mutations specific to the xeroderma pigmentosum group E Ddb- phenotype. *J. Biol. Chem.* **271:** 24317-24320.

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

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

135. **Nouspikel, T., and S.G. Clarkson.** 1994. Mutations that disable the DNA repair gene *XPG* in a xeroderma pigmentosum group G patient. *Hum. Mol. Gen.* **3:** 963-967.

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

137. **O'Donovan, A., A.A. Davies, J.G. Moggs, S.C. West, and R.D. Wood.** 1994. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371:** 432-435.

138. **O'Donovan, A., and R.D. Wood.** 1993. Identical defects in DNA repair in xeroderma pigmentosum group G and rodent ERCC group 5. *Nature* **363:** 185-188.
139. **Okhuma, Y.** 1997. Multiple functions of general transcription factors TFIIE and TFIIH: possible points of regulation by *trans*-acting factors. *J. Biochem.* **122:**4 81-489.
140. **Okhuma, Y., S. Hashimoto, C. Wang, M. Horikoshi, and R. Roeder.** 1995. Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha. *Mol. Cell. Biol.* **15:** 4856-4866.
141. **Okhuma, Y., and R. Roeder.** 1994. Regulatory effect of TFIIE on both ATPase and kinase activities of TFIIH during active initiation complex formation. *Nature* **368:** 160-163.
142. **Pan, Z.-Q., C.-H. Park, A.A. Amin, J. Hurwitz, and A. Sancar.** 1995. Phosphorylated and unphosphorylated forms of human single-stranded DNA-binding protein are equally active in simian virus 40 DNA replication and in nucleotide excision repair. *Proc.Natl.Acad.Sci.* **92:**4636-4640.
143. **Park, C.-H., T. Bessho, T. Matsunaga, and A. Sancar.** 1995. Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *J. Biol. Chem.* **270:** 22657-22660.
144. **Park, C.-H., D. Mu, J.T. Reardon, and A. Sancar.** 1995. 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.
145. **Park, C.-H., and A. Sancar.** 1994. Formation of a ternary complex by human XPA, ERCC1 and ERCC4(XPF) excision repair proteins. *Proc. Natl. Acad. Sci.* **91:** 5017-5021.
146. **Philipova, D., J. Mullen, H. Maniar, J. Lu, C. Gu, and S. Brill.** 1996. A hierarchy of SSB protomers in replication protein A. *Genes Dev.* **10:** 2222-2233.
147. **Popanda, O., and H. Thielmann.** 1992. The function of DNA polymerases in DNA repair synthesis of ultraviolet-irradiated human fibroblasts. *Biochim. Biophys. Acta* **1129:** 155-160.
148. **Prigent, C., M. Satoh, G. Daly, D. Barnes, and T. Lindahl.** 1994. Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I. *Mol. Cell. Biol.* **14:** 310-317.
149. **Reardon, J., T. Bessho, H. Kung, P. Bolton, and A. Sancar.** 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.* **94:** 9463-9468.
150. **Reardon, J., H. Ge, E. Gibbs, A. Sancar, J. Hurwitz, and Z. Pan.** 1996. Isolation and characterization of two human transcription factor IIH (TFIIH)-related complexes: ERCC2/CAK and TFIIH. *Proc. Natl. Acad. Sci.***93:**6482-6487.
151. **Reardon, J., D. Mu, and A. Sancar.** 1996. Overproduction, purification, and characterization of the XPC subunit of the human DNA repair excision nuclease. *J. Biol. Chem.* **271:** 19451-19456.
152. **Reardon, J.T., A.F. Nichols, S. Keeney, C.A. Smith, J.S. Taylor, S. Linn, and A. Sancar.** 1993. 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.
153. **Robins, P., C.J. Jones, M. Biggerstaff, T. Lindahl, and R.D. Wood.** 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.
154. **Rossignol, M., I. Kolb-Cheynel, and J. Egly.** 1997. Substrate specificity of the cdk-activating

kinase (CAK) is altered upon association with TFIIH. *EMBO J.* **16:** 1628-1637.

155. **Roy, R., J.P. Adamczewski, T. Seroz, W. Vermeulen, J.-P. Tassan, L. Schaeffer, E.A. Nigg, J.H.J. Hoeijmakers, and J.-M. Egly.** 1994. The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. *Cell* **79:** 1093-1101.
156. **Roy, R., L. Schaeffer, S. Humbert, W. Vermeulen, G. Weeda, and J.-M. Egly.** 1994. The DNA-dependent ATPase activity associated with the class II basic transcription factor BTF2/TFIIH. *J. Biol. Chem.* **269:** 9826-9832.
157. **Rumbaugh, J., R. Murante, S. Shi, and R. Bambara.** 1997. Creation and removal of embedded ribonucleotides in chromosomal DNA during mammalian Okazaki fragment processing. *J. Biol. Chem.* **272:** 22591-22599.
158. **Saijo, M., I. Kuraoka, C. Masutani, F. Hanaoka, and K. Tanaka.** 1996. Sequential binding of DNA repair proteins RPA and ERCC1 to XPA in vitro. *Nucleic Acids Res.* **24:** 4719-4724.
159. **Sands, A.T., A. Abuin, A. Sanchez, C.J. Conti, and A. Bradley.** 1995. High susceptibility to ultraviolet-induced carcinogenesis in mice lacking *XPC*. *Nature* **377:** 162-165.
160. **Schaeffer, L., V. Moncollin, R. Roy, A. Staub, M. Mezzina, A. Sarasin, G. Weeda, J.H.J. Hoeijmakers, and J.M. Egly.** 1994. The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.* **13:** 2388-2392.
161. **Schaeffer, L., R. Roy, S. Humbert, V. Moncollin, W. Vermeulen, J.H.J. Hoeijmakers, P. Chambon, and J. Egly.** 1993. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* **260:** 58-63.
162. **Scherly, D., T. Nouspikel, J. Corlet, C. Ucla, A. Bairoch, and S.G. Clarkson.** 1993. Complementation of the DNA repair defect in xeroderma pigmentosum group G cells by a human cDNA related to yeast RAD2. *Nature* **363:** 182-185.
163. **Selby, C., and A. Sancar.** 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.
164. **Serizawa, H., T.M. Mäkelä, J.W. Conaway, R.C. Cnaway, R.A. Weinberg, and R.A. Young.** 1995. Association of cdk-activating kinase subunits with transcription factor TFIIH. *Nature* **374:** 280-282.
165. **Shen, B., J. Nolan, L. Sklar, and M. Park.** 1996. Essential amino acids for substrate binding and catalysis of human flap endonuclease 1. *J. Biol. Chem.* **271:** 9173-9176.
166. **Shen, B., J. Nolan, L. Sklar, and M. Park.** 1997. Functional analysis of point mutations in human flap endonuclease-1 active site. *Nucleic Acids Res.* **25:** 3332-3338.
167. **Shiekhattar, R., F. Mermelstein, R.P. Fisher, R. Drapkin, B. Dynlacht, H.C. Wesselin, D.O. Morgan, and D. Reinberg.** 1995. Cdk-activating kinase complex is a component of human transcription factor TFIIH. *Nature* **374:** 283-287.
168. **Shivji, M.K.K., A.P.M. Eker, and R.D. Wood.** 1994. DNA repair defect in xeroderma pigmentosum group C and complementing factor from HeLa cells. *J. Biol. Chem.* **269:** 22749-22757.
169. **Shivji, M.K.K., M.K. Kenny, and R.D. Wood.** 1992. Proliferating cell nuclear antigen is required for DNA excision repair. *Cell* **69:** 367-374.
170. **Shivji, M.K.K., V.N. Podust, U. Hubscher, and R.D. Wood.** 1995. Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochem.* **34:** 5011-5017.
171. **Sijbers, A.M., W.L. de Laat, R.R. Ariza, M. Biggerstaff, Y.-F. Wei, J.G. Moggs, K.C. Carter,**

B.K. Shell, E. Evans, M.C. de Jong, S. Rademakers, J. de Rooij, N.G.J. Jaspers, J.H.J. Hoeijmakers, and R.D. Wood. 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86:** 811-822.

172. **Sijbers, A.M., P.J. van der Spek, H. Odijk, J. van den Berg, M. van Duin, A. Westerveld, N.G.J. Jaspers, D. Bootsma, and J.H.J. Hoeijmakers.** 1996. Mutational analysis of the human nucleotide excision repair gene *ERCC1*. *Nucleic Acids Res.* **24:** 3370-3380.

173. **Smith, M.L., I.-T. Chen, Q. Zhan, I. Bae, C.-Y. Chen, T.M. Gilmer, M.B. Kastan, P.M. O'Connor, and A.J. Fornace Jr.** 1994. Interaction of the p53-regulated protein Gadd45 with proliferating cell nuclear antigen. *Science* **266:** 1376-1380.

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

175. **Sugasawa, K., J. Ng, C. Masutani, T. Maekawa, A. Uchida, P. van der Spek, A. Eker, S. Rademakers, C. Visser, A. Aboussekhra, R. Wood, F. Hanaoka, D. Bootsma, and J. Hoeijmakers.** 1997. Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity. *Mol. Cell. Biol.* **17:** 6924-6931.

176. **Sugasawa, K., J.M.Y. Ng, C. Masutani, P.J. van der Spek, A.P.M. Eker, F. Hanaoka, D. Bootsma, and J.H.J. Hoeijmakers.** 1998. Xeroderma pigmentosum group C is the initiator of global genome nucleotide excision repair. *Submitted*.

177. **Sung, P., S. Guzder, L. Prakash, and S. Prakash.** 1996. Reconstitution of TFIIH and requirement of its DNA helicase subunits, Rad3 and Rad25, in the incision step of nucleotide excision repair. *J. Biol. Chem.* **271:** 10821-10826.

178. **Svejstrup, J.Q., Z. Wang, W.J. Feaver, X. Wu, D.A. Bushnell, T.F. Donahue, E.C. Friedberg, and R.D. Kornberg.** 1995. Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* **80:** 21-28.

179. **Svoboda, D.L., J.S. Taylor, J.E. Hearst, and A. Sancar.** 1993. DNA repair by eukaryotic nucleotide excision nuclease. *J. Biol. Chem.* **268:** 1931-1936.

180. **Sweder, K.S., R.A. Verhage, D.J. Crowley, G.F. Crouse, J. Brouwer, and P.C. Hanawalt.** 1996. Mismatch repair mutants in yeast are not defective in transcription-coupled DNA repair of UV-induced DNA damage. *Genetics* **143:** 1127-1135.

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

182. **Tantin, D., A. Kansal, and M. Carey.** 1997. Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Mol Cell Biol* **17:** 6803-6814.

183. **Thommes, P., E. Ferrari, R. Jessberger, and U. Hubscher.** 1992. Four different DNA helicases from calf thymus. *J. Biol. Chem.* **267:** 6063-6073.

184. **Tijsterman, M., R. Verhage, P. van de Putte, J. Tasseron-de Jong, and J. Brouwer.** 1997. Transitions in the coupling of transcription and nucleotide excision repair within RNA polymerase II-transcribed genes of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.* **94:** 8027-8032.

185. **Timmers, H.T.M.** 1994. Transcription initiation by RNA polymerase II does not require hydrolysis of the beta-gamma phosphoanhydride bond of ATP. *EMBO J.* **13:** 391-399.

186. **Tomkinson, A., and D. Levin.** 1997. Mammalian DNA ligases. *Bioassays* **19:** 893-901.

187. **Treuner, K., U. Ramsperger, and R. Knippers.** 1996. Replication protein A induces the unwinding of long double-stranded DNA regions. *J. Mol. Biol.* **259:** 104-112.

188. **Tsurimoto, T., M. Fairman, and B. Stillman.** 1989. Simian virus 40 DNA replication in vitro: identification of multiple stages of initiation. *Mol. Cell. Biol.* 9:3839-49.

189. **Tu, Y., S. Bates, and G. Pfeifer.** 1997. Sequence-specific and domain-specific DNA repair in xeroderma pigmentosum and Cockayne syndrome cells. *J Biol Chem* **272:** 20747-55.

190. **Tyree, C.M., C.P. George, L.M. Lira-De Vito, S.L. Wampler, M.E. Dahmus, L. Zawel, and J.T. Kadonaga.** 1993. Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. *Genes Dev.* **7:** 1254-1265.

191. **Valay, J.G., M. Simon, M.F. Dubois, O. Bensuade, C. Facca, and G. Faye.** 1995. The *KIN28* gene is required both for RNA polymerase II mediated transcription and phosphorylation of the Rpb1 CTD. *J. of Mol.Biol.* **249:** 535-544.

192. **van der Horst, G., H. van Steeg, R. Berg, A. van Gool, J. de Wit, G. Weeda, H. Morreau, R. Beems, C. van Kreijl, F. de Gruijl, D. Bootsma, and J. Hoeijmakers.** 1997. Defective transcription-coupled repair in Cockayne syndrome B mice is associated with skin cancer predisposition. *Cell* **89:** 425-435.

193. **van der Spek, P.J., A. Eker, S. Rademakers, C. Visser, K. Sugasawa, C. Masutani, F. Hanaoka, D. Bootsma, and J.H. Hoeijmakers.** 1996. XPC and human homologs of RAD23: intracellular localization and relationship to other nucleotide excision repair complexes. *Nucleic Acids Res.* **24:** 2551-2559.

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

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

196. **van Gool, A., G. van der Horst, E. Citterio, and J. Hoeijmakers.** 1997. Cockayne syndrome: defective repair of transcription? *EMBO J.* **16:** 4155-4162.

197. **van Hoffen, A., A.T. Natarajan, L.V. Mayne, A.A. van Zeeland, L.H.F. Mullenders, and J. Venema.** 1993. Deficient repair of the transcribed strand of active genes in Cockayne syndrome. *Nucleic Acids Res.* **21:** 5890-5895.

198. **van Hoffen, A., J. Venema, R. Meschini, A.A. van Zeeland, and L.H.F. Mullenders.** 1995. Transcription-coupled repair removes both cyclobutane pyrimidine dimers and 6-4 photoproducts with equal efficiency and in a sequential way from transcribed DNA in xeroderma pigmentosum group C fibroblasts. *The EMBO Journal* **14:** 360-367.

199. **van Vuuren, A.J., E. Appeldoorn, H. Odijk, A. Yasui, N.G.J. Jaspers, D. Bootsma, and J.H.J. Hoeijmakers.** 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.

200. **van Vuuren, A.J., W. Vermeulen, L. Ma, G. Weeda, E. Appeldoorn, N.G.J. Jaspers, A.J. van der Eb, D. Bootsma, J.H.J. Hoeijmakers, S. Humbert, L. Schaeffer, and J.-M. Egly.** 1994. Correction of xeroderma pigmentosum repair defect by basal transcription factor BTF2(TFIIH).

EMBO J. **13**: 1645-1653.

201. **Venema, J., L.H.F. Mullenders, A.T. Natarajan, A.A. Van Zeeland, and L.V. Mayne.** 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.* **87**: 4707-4711.
202. **Venema, J., A. van Hoffen, V. Karcagi, A.T. Natarajan, A.A. van Zeeland, and L.H.F. Mullenders.** 1991. Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. *Mol. Cell. Biol.* **11**: 4128-4134.
203. **Venema, J., A. Van Hoffen, A.T. Natarajan, A.A. Van Zeeland, and L.H.F. Mullenders.** 1990. The residual repair capacity of xeroderma pigmentosum complementation group C fibroblasts is highly specific for transcriptionally active DNA. *Nucleic Acids Res.* **18**: 443-448.
204. **Verhage, R.A.** 1996. Transcription-coupled and globalg enome repair in *Saccharomyces cerevisiae*.Thesis. Leiden University, Leiden.
205. **Verhage, R.A., P. van der Putte, and J. Brouwer.** 1996. Repair of rDNA in *Saccharomyces cerevisiae*: *RAD4*-independent strand-specific nucleotide excision repair of RNA polymerase I transcribed strands. *Nucleic Acid Res.* **24**: 1020-1025.
206. **Verhage, R.A., A. Zeeman, N. de Groot, F. Gleig, D. Bang, P. van der Putte, and J. Brouwer.** 1994. The *RAD7* and *RAD16* genes are essential for repair of non-transcribed DNA in *Saccharomyces ceriviae*. *Mol. Cell. Biol.* **14**: 6135-6142.
207. **Verhage, R.A., A.J. Van Gool, N. De Groot, J.H.J. Hoeijmakers, P. Van de Putte, and J. Brouwer.** 1996. Double mutants of *Saccharomyces cerevisiae* with alterations in global genome and transcription-coupled repair. *Mol. Cell. Biol.* **16**: 496-502.
208. **Vermeulen, W., J. Jaeken, N.G.J. Jaspers, D. Bootsma, and J.H.J. Hoeijmakers.** 1993. Xeroderma pigmentosum complementation group G associated with Cockayne's syndrome. *Am. J. Human Genet.* **53**: 185-192.
209. **Vermeulen, W., A.J. van Vuuren, M. Chipoulet, L. Schaeffer, E. Appeldoorn, G. Weeda, N.G.J. Jaspers, A. Priestley, C.F. Arlett, A.R. Lehmann, M. Stefanini, M. Mezzina, A. Sarasin, D. Bootsma, J.-M. Egly, and J.H.J. Hoeijmakers.** 1994. 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.
210. **Waga, S., G.J. Hannon, D. Beach, and B. Stillman.** 1994. The p21 inhibitor of cylcin-dependent kinases controls DNA replication by interaction with PCNA. *Nature* **369**: 574-578.
211. **Wakasugi, M., J. Reardon, and A. Sancar.** 1997. The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair. *J. Biol. Chem.* **272**: 16030-16034.
212. **Wang, X., W. Vermeulen, J. Coursen, M. Gibson, S. Lupold, K. Forrester, G. Xu, L. Elmore, H. Yeh, J. Hoeijmakers, and C. Harris.** 1996. The XPB and XPD DNA helicases are components of the p53-mediated apoptosis pathway. *Genes Dev.* **10**: 1219-1232.
213. **Wang, X.W., H. Yeh, L. Schaeffer, R. Roy, V. Moncollin, J.-M. Egly, Z. Wang, E.C. Friedberg, M.K. Evans, B.G. Taffe, V.A. Bohr, G. Weeda, J.H.J. Hoeijmakers, K. Forrester, and C.C. Harris.** 1995. p53 modulation of TFIIH-associated nucleotide excision repair activity. *Nat. Gen.* **10**: 188-195.
214. **Wang, Z., S. Buratowski, J.Q. Svejstrup, W.J. Feaver, X. Wu, R.D. Kornberg, T.D. Donahue, and E.C. Friedberg.** 1995. The yeast *TFB1* and *SSL1* genes, which encode subunits of transcription factor IIH, are required for nucleotide excision repair and RNA polymerase II transcription. *Mol. Cell. Biol.* **15**: 2288-2293.

215. **Wang, Z., J.Q. Svejstrup, W.J. Feaver, X. Wu, R.D. Kornberg, and E.C. Friedberg.** 1994. Transcription factor b (TFIIH) is required during nucleotide excision repair in yeast. *Nature* **368:** 74-76.

216. **Watkins, J.F., P. Sung, L. Prakash, and S. Prakash.** 1993. The *Saccharomyces cerevisiae* DNA repair gene *RAD23* encodes a nuclear protein containing a ubiquitin-like domain required for biological function. *Mol. Cell. Biol.* **13:** 7757-7765.

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

218. **Wellinger, R., and F. Thoma.** 1997. Nucleosome structure and positioning modulate nucleotide excision repair in the non-transcribed strand of an active gene. *EMBO J.* **16:** 5046-5056.

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

220. **Winkler, G., W. Vermeulen, F. Coin, J. Egly, J. Hoeijmakers, and G. Weeda.** 1998. Affinity purification of human DNA repair/transcription factor TFIIH using epitope-tagged xeroderma pigmentosum B protein. *J. Biol. Chem.* **273:**1092-1098.

221. **Wobbe, C., L. Weissbach, J. Borowiec, F. Dean, Y. Murakami, P. Bullock, and J. Hurwitz.** 1987. Replication of simian virus 40 origin-containing DNA in vitro with purified proteins. *Proc. Natl. Acad. Sci.* **84:** 1834-1838.

222. **Wold, M., D. Weinberg, D. Virshup, J. Li, and T. Kelly.** 1989. Identification of cellular proteins required for simian virus 40 DNA replication. *J. Biol. Chem.* **264:** 2801-2809.

223. **Wold, M.S.** 1997. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* **66:** 61-92.

224. **Wold, M.S., and T. Kelly.** 1988. Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. *Proc. Natl Acad. Sci.* **85:** 2523-2527.

225. **Wood, R.D., and M.K.K. Shivji.** 1997. Which DNA polymerases are used for DNA-repair in eukaryotes? *Carcinog.* **18:** 605-610.

226. **Wu, X., J. Li, X. Li, C. Hsieh, P. Burgers, and M. Lieber.** 1996. Processing of branched DNA intermediates by a complex of human FEN-1 and PCNA. *Nucleic Acids Res.* **24:** 2036-43.

227. **Xiao, H., A. Pearson, B. Coulombe, R. Truant, S. Zhang, J. Regier, S. Triezenberg, D. Reinberg, O. Flores, and C. Ingles.** 1994. Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. *Mol. Cell. Biol.* **14:** 7013-7024.

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

229. **Yan, M., and J. Gralla.** 1997. Multiple ATP-dependent steps in RNA polymerase II promoter melting and initiation. *EMBO J.* **16:**7457-7467.

230. **Zawel, L., K. Kumar, and D. Reinberg.** 1995. Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* **9:** 1479-1490.

Outline of the work

The work presented in this thesis aims to contribute to the understanding of the molecular mechanisms that underlie mammalian nucleotide excision repair. More specifically, it addresses the questions: what enzymes are responsible for incisions around DNA damage and how are these enzymes coordinated? Initially, attention was focussed on a protein complex containing the repair factor ERCC1.

In **Chapter 3**, the purification of the ERCC1-containing complex from mammalian cells is reported. A key step was the introduction of an affinity-tag at one of the termini of the ERCC1 protein, which allowed the purification of this repair complex to homogeneity. ERCC1 appeared to be tightly associated to another protein, XPF. We show that purified ERCC1-XPF is a structure-specific endonuclease that cuts one strand of duplex DNA at junctions with single-stranded DNA. Consistent with its polarity of cleavage, we demonstrate that ERCC1-XPF is responsible for the 5' incision relative to the damage during NER.

In **Chapter 4** we define the regions in the ERCC1 and XPF proteins that are required for complex formation. Immunoprecipitation studies with truncated protein fragments revealed that the main domains responsible for interaction between ERCC1 and XPF reside in the C-terminal parts of both proteins. Mutated proteins from an XP-F patient are shown to be partially disrupted in ERCC1-complex formation.

In **Chapter 5** we further characterize the enzymatic activity of this repair complex by determining the DNA structural elements required for ERCC1-XPF cleavage. This is particularly relevant, since ERCC1-XPF has an additional, but poorly defined role in recombinational processes. We report that one single-stranded arm, protruding in either the 3' or 5' direction, is necessary and sufficient for correct positioning of incisions by ERCC1-

XPF. We also show that ERCC1-XPF cleaves presumed recombination intermediates.

Finally, we try to integrate the ERCC1-XPF activity in the context of other repair factors. In **Chapter 6**, we show that RPA has an important role in positioning ERCC1-XPF and XPG, the nuclease responsible for the 3' incision in nucleotide excision repair. We demonstrate that RPA binds single-stranded DNA with a defined polarity, which allows the definition of a 3' and a 5' oriented side of RPA. ERCC1-XPF specifically interacts with, and is stimulated by, the 3' oriented side of RPA. Conversely, XPG specifically interacts with the 5' oriented side of RPA. We argue that RPA is bound to the undamaged strand during nucleotide excision repair and confers strand-specificity to ERCC1-XPF and possibly XPG.

The implications of this work for nucleotide excision repair have extensively been discussed in **Chapter 2**. For more detailed considerations of specific findings I refer to the discussion sections present at the end of each of the following chapters.

Chapter 3

The ERCC1-XPF complex is a structure-specific DNA endonuclease that makes the 5' incision in nucleotide excision repair

Abstract

During nucleotide excision repair a dual incision is made at both sides of a DNA lesion to allow removal of the damaged oligonucleotide followed by gap filling DNA synthesis. Humans suffering from xeroderma pigmentosum (XP) are defective in this repair pathway. Here we report the purification of a His-tagged ERCC1 protein complex from mammalian cells. We show that this complex is a heterodimer of ERCC1 and the recently identified XPF protein. Using synthetic DNA substrates with a duplex stem and a single-stranded loop, we demonstrate that the ERCC1-XPF complex is a structure-specific endonuclease, cleaving one strand of duplex DNA specifically at the 5' side of a single stranded border. ERCC1-deficient cell extracts are shown to be defective in 5' incision activity; this activity is restored upon addition of purified ERCC1-XPF complex. These data show that the ERCC1-XPF complex is a structure-specific endonuclease making the 5' incision during nucleotide excision repair.

Data have been published in Cell, vol. 86, 811-822 (1996)

Introduction

A coordinated interplay between multiple subunits is required to carry out nucleotide excision repair (NER) in eucaryotes. The first steps of the process lead to to lesion recognition and incision of the damaged strand on each side of the lesion. A 24-32 mer oligonucleotide is removed, followed by gap-filling DNA synthesis (6, 10, 14). In human cells, this process 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 (9), 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 (reviewed by (26)). Simultaneous with the work described here, the *XPF* gene was cloned, which appeared to encode a 115 kD protein (19).

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 (13, 15). XPG and its yeast homolog Rad2 specifically cleave near junctions of unpaired and duplex DNA, cutting the strand in which the unpaired region moves 3' to 5' away from the junction (7, 8, 15). In *Saccharomyces 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 (2, 5). ERCC1 is the mammalian homolog of Rad10 (23) 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 (3, 25). A polypeptide of relative molecular mass of approximately 115 kD has been observed to copurify with ERCC1 by several assays and has been proposed as a candidate for a Rad1 homolog (1, 17, 24). We decided to study directly the composition and activity of the ERCC1 protein complex by purifying it from Chinese hamster cells.

Results

ERCC1 and XPF form a heterodimeric protein complex in mammalian cells

To facilitate purification of the ERCC1 protein complex, a 6xHis affinity tag was introduced at the C-terminus of human ERCC1, and cDNA encoding this His-tagged ERCC1 was transfected into ERCC1-deficient Chinese Hamster (43-3B) cells. The transformed cells were selected for repair competence and extensively characterized to ensure proper expression and functioning of the His-tagged ERCC1 protein. Immunochemical staining confirmed expression of His-ERCC1 protein in the nuclei of transfected 43-3B cells (Fig. 1a-d), and immunoblotting showed that normal amounts of

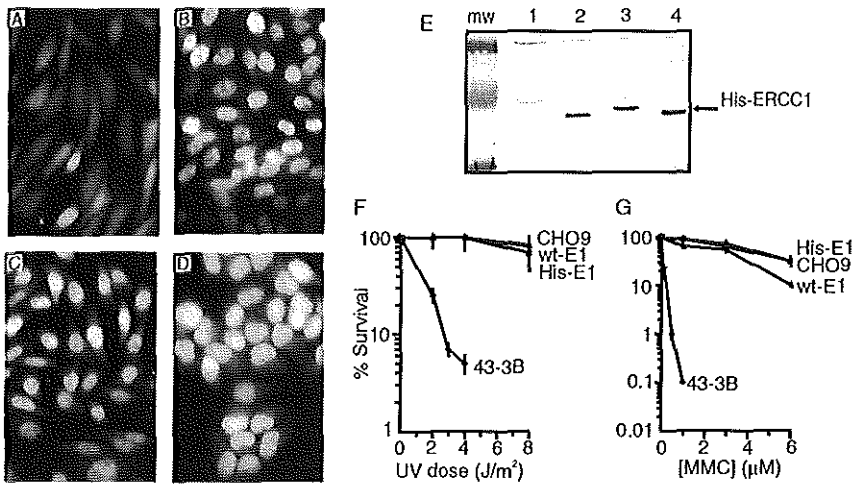


Figure 1. Functional expression of His-tagged ERCC1 in Chinese hamster 43-3B cells. (A-D) Nuclear localization of His-ERCC1 protein, detected by immunofluorescence microscopy using affinity purified anti-ERCC1 antibody. (A) 43-3B cells, (B) 43-3B + His-ERCC1, (C) 43-3B + wt-ERCC1, (D) HeLa. (E) Immunoblot with anti-ERCC1 antibody, showing ERCC1 expression levels in lane 1, 43-3B; lane 2, 43-3B plus wild-type ERCC1; lane 3, 43-3B plus His-ERCC1; lane 4, HeLa. Similar protein amounts were loaded in the different lanes. Note that the His-tag causes a mobility-shift on SDS-polyacrylamide gel electrophoresis and that this anti-ERCC1 antibody does not recognize Chinese Hamster ERCC1. (F-G) Human His-ERCC1 protein fully corrects the UV- and mitomycin C (mmc) sensitivity of 43-3B cells.

ERCC1 were produced (Fig. 1e). The transfected His-ERCC1 fully corrected both the UV and MMC sensitivity of 43-3B cells (Fig. 1f-g). 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, visualized by silver staining, with apparent molecular weights of 42 kD, 60 kD and 115 kD (Fig. 2a). Immunoblotting identified the 42 kD band as His-tagged ERCC1 and the 115 kD band as XPF protein, respectively (Fig. 2b). The ERCC1 and XPF proteins coeluted at each step and began to separate from the 60 kD protein and minor contaminants upon gel filtration chromatography (Fig. 2a; see different peak fractions for ERCC1-XPF and the 60 kD band), glycerol gradient sedimentation, or on Reactive Yellow 86 agarose (Fig. 2c). These results indicate that the ERCC1 protein complex is a heterodimer of ERCC1 and XPF, consistent with the suggestion of Park et al. (17).

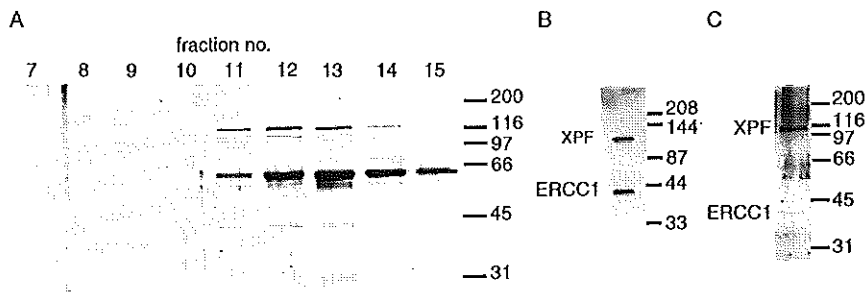


Figure 2. Purified ERCC1-XPF complex. (A) Fraction V was separated on a Superose 12 gel filtration column. Fractions were analyzed by SDS-polyacrylamide gel electrophoresis and silver-stained. ERCC1 and XPF are indicated by arrows. (B) Immunoblot of fraction V, using anti-ERCC1 antibody and an antibody raised against the C-terminal part of XPF (19). (C) Fraction VI after purification on Reactive Yellow 86 agarose.

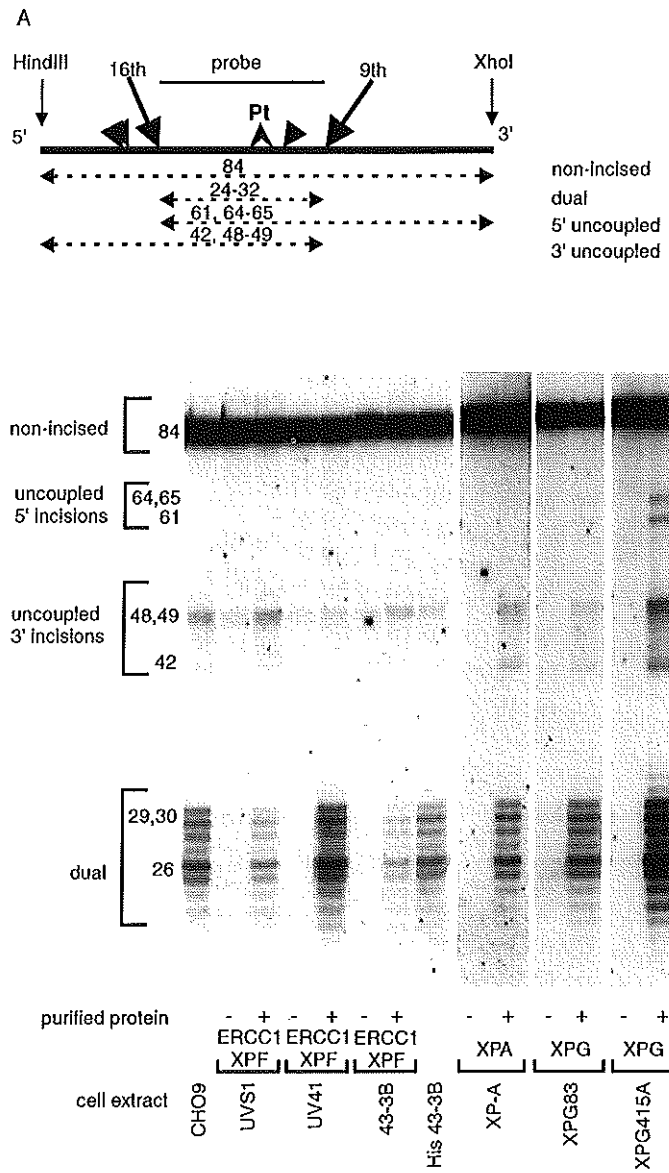
Functional activity was shown by several criteria. The complex could correct the repair defect of human XP-F cell extracts, as shown in a repair synthesis assay for NER (Fig. 3b) and in a dual incision assay (Fig. 3a). Also, complementation of the repair

defects of Chinese Hamster group 1 and 4 mutant cell extracts as well as a UVS1 cell extract was found with purified ERCC1-XPF (Fig. 3a-b). The latter cell line was formerly known as the sole representative of complementation group 11 but was recently shown to be defective in XPF functioning (19). The complex did not contain correcting activity for other mutant cell lines (Fig. 3b). In addition, the complex was active in a fully reconstituted repair system (data not shown) (1).

ERCC1-XPF is a structure-specific DNA endonuclease

To determine whether the purified ERCC1-XPF complex possessed structure-specific endonuclease activity, we used a partially self-complementary oligonucleotide to form a stem-loop structure consisting of a 22 nt single-stranded loop and a duplex stem of 12 base pairs (Fig. 4a). 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 this substrate in one strand of the stem 2, 3, and 4 phosphodiester bonds away from the 5' side of the loop (Fig. 4a, lanes 2, 3, 6, and 7). These incisions colocalized precisely with those catalyzed by yeast Rad1-Rad10 (Fig. 4a, lanes 11, 12, 15, and 16). Conversely, human XPG 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. 4a, 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 coeluted 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 preimmune serum could not precipitate active complex (Fig. 4b). These findings clearly demonstrate that structure-specific 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 approxi-


A
HindIII
16th
probe
9th
XhoI
Pt
5'
3'
84
24-32
61, 64-65
42, 48-49
non-incised
dual
5' uncoupled
3' uncoupled
non-incised
84
uncoupled
5' incisions
64,65
61
uncoupled
3' incisions
48,49
42
dual
29,30
26
purified protein
- + - + - + - + - + - +
ERCC1 ERCC1 ERCC1
XPF XPF XPF
XPA XPG XPG
cell extract
CHO9
UVS1
UV41
43-3B
His 43-3B
XP-A
XPG83
XPG415A

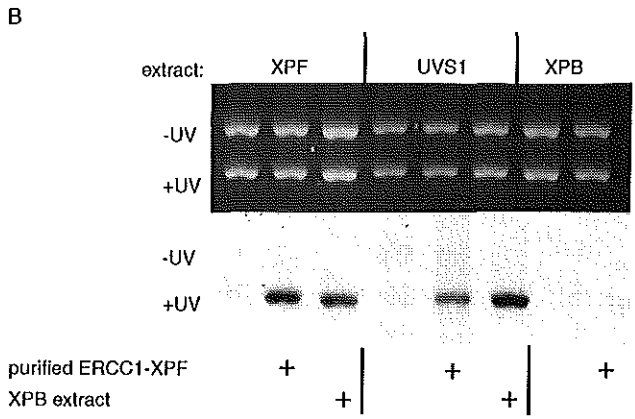


Figure 3. Repair-correcting activity of ERCC1-XPF complex. (A) Autoradiograph of Southern blot. Platinated DNA fragments corresponding to nonincised DNA (84 nt), uncoupled 3' incisions (42, 48 and 49 nt), and dual incisions (24-32 nt) were detected after hybridization with a ^{32}P-labeled 27 mer probe complementary to the DNA sequence surrounding the cisplatin cross-link (14), 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 (CHO group 11), 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 of complex). Cell extract (200 μg of protein) from 43-3B cells expressing His-tagged ERCC1 was used in lane 8. XPG83 and XPG415A cell extracts were complemented with 50 ng of purified XPG protein (16) (lanes 12 and 14, respectively), and XP-A (GM2345) cell extract was complemented with 90 ng of purified XPA protein (11) (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 (27). Lanes 1, 2, 4, 5, 7, and 8 had 100 μg of each CFII fraction. Purified ERCC1-XPF protein (fraction V, 1 μl) was added in lanes 2, 5, and 8.

mately 36 mer, indicating the presence of an OH group at the 3' terminus of the incision product (Fig. 4c).

A "bubble" substrate containing a centrally unpaired region of 30 nt (15) 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,

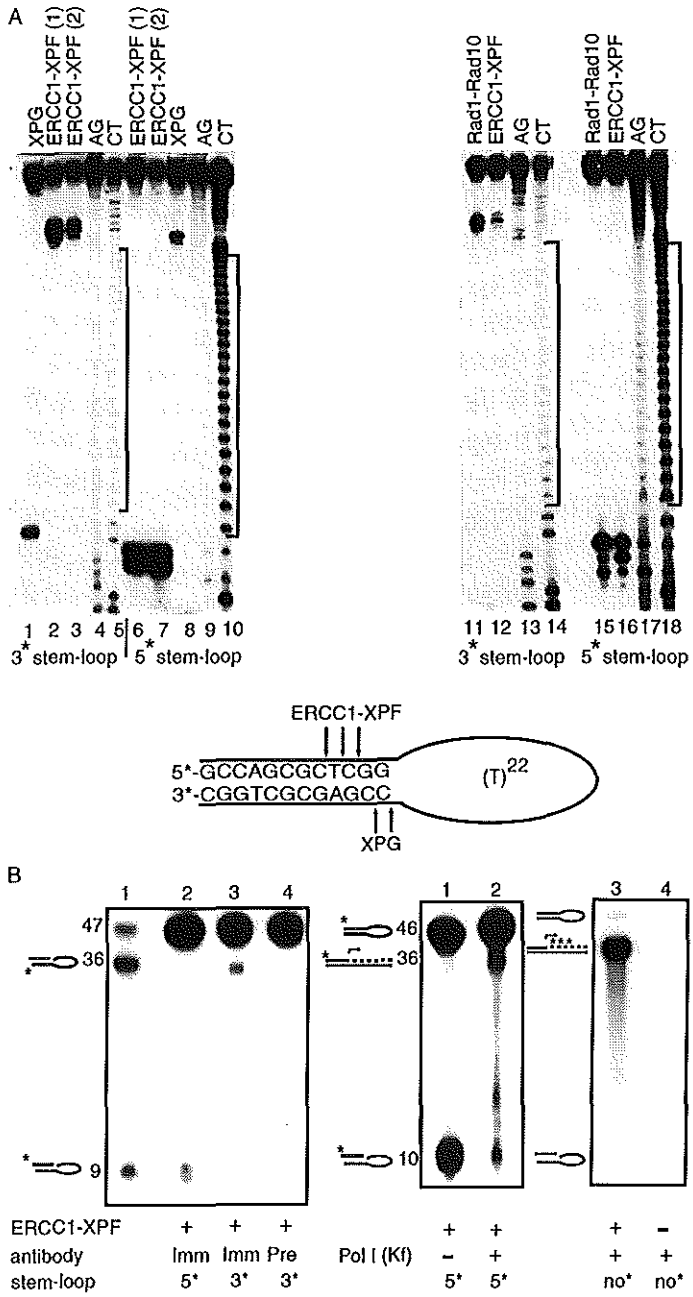

A
XPG
ERCC1-XPF (1)
ERCC1-XPF (2)
AG
CT
ERCC1-XPF (1)
ERCC1-XPF (2)
XPG
AG
CT
Rad1-Rad10
ERCC1-XPF
AG
CT
Rad1-Rad10
ERCC1-XPF
AG
CT
1 2 3 4 5 6 7 8 9 10
3* stem-loop
5* stem-loop
11 12 13 14 15 16 17 18
3* stem-loop
5* stem-loop
ERCC1-XPF
5*-GCCAGCGCTCGG
3*-CGGTCGCGAGCC
(T)22
XPG
B
1 2 3 4
47
36
9
1 2
46
36
10
3 4
ERCC1-XPF + + + + + -
antibody Imm Imm Pre Pol I (Kf) - + + +
stem-loop 5* 3* 3* 5* 5* no* no*

Figure 4. The ERCC1-XPF complex is a structure-specific DNA endonuclease. (A) The structure of the stem-loop substrate is shown below thegels. 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 nt faster that the corresponding nuclease products. The 22 T residues are indicated by brackets. Lanes 1-5, 3'-labeled stem-loop with lane 1, XPG (50 ng); lanes 2 and 3, ERCC1-XPF (purified fraction V, 1 and 2 µl, respectively); lane 4 and 5, Maxam-Gilbert G+A and C+T sequencing ladders. Lanes 6-10, 5'-labeled stem-loop, with lanes 6 and 7, ERCC1-XPF (1 and 2 µl of fraction V, respectively); lane 8, XPG (50 ng); lanes 9 and 10, G+A and C+T sequencing ladders. Lanes 11-14, 3'-labeled stem-loop, with lane 11, Rad1 (100 ng) and Rad10 (25 ng); lane 12, ERCC1-XPF (1 µl, fraction V); lanes 13 an 14, Maxam-Gilbert G+A and C+T sequencing ladders. Lanes 15-18, same as 11-14 with 5'-labeled stem-loop.
(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) preincubated with purified ERCC1-XPF; lane 3, as lane 2 but with 3'-labeled lasso; lane 4, as lane 3 but with preimmune antibody-beads (Pre) preincubated 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'-(^{32}P)-labeled stem-loop cleaved with ERCC1-XPF, followed by incubation with Klenow fragment of *E. coli* DNA polymerase I (Pol I Kf) and unlabelled 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.

while XPG preferred the bubble over the stem-loop (data not shown). Additional NER incision stage 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.

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 duplex DNA specifically at the 5' side of single-stranded borders. The polarity of enzymatic cleavage of model substrates strongly suggests that ERCC1-XPF makes the 5' incision during NER. This is indeed the case during repair, as shown in Figure 3a. 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. 3a, lane 1; see

also (13)). In very strong reactions, some uncoupled 5' incisions also can be detected (see Fig. 3a, lane 14; see also (14)). XP-G extracts (defective in 3' endonuclease activity) or XP-A extracts (missing the damage recognition protein XPA) were defective in incision (Fig. 3a, lanes 9, 11, 13). Although ERCC1- or XPF-deficient cell extracts were unable to generate 24-32 nt excision products, fragments corresponding to uncoupled 3' incisions were still detected (Fig. 3a, lanes 2, 4, 6). The ERCC1 and XP-F cell extracts therefore specifically lack the ability to make the 5' incision during repair. Addition of purified ERCC1-XPF complex to the extracts restored 5' incision activity and generated normal excision products (Fig. 3a, lanes 3, 5, 7).

Discussion

ERCC1 and XPF form an endonuclease complex that makes the 5' incision in NER

Simultaneous with the study presented here, the human homolog of Rad1 was cloned, which turned out to be the *XPF* gene (4, 19). An antibody raised against ERCC1 not only precipitated ERCC1 protein from normal human cell extracts, but also XPF protein, strongly suggesting that ERCC1 and XPF, like their yeast counterparts Rad10 and Rad1, form a stable protein complex in the cell (19). Here we demonstrate directly that ERCC1 and XPF are tightly associated *in vivo*. After extensive purification, only one protein copurifies exactly with ERCC1, and with an antibody raised against XPF we were able to show that this 115 kD protein is XPF. Futhermore, the purified ERCC1-XPF complex corrected the repair defects of both ERCC1- and XPF-deficient cell extracts. 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 (2, 20-22). Our data, together with the findings by Sijbers et al. that XPF is the human equivalent of ERCC4 in Chinese hamster (19), 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. 4a) 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 (14), 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 kD polypeptide from HeLa cells was reported to lack any structure-specific endonuclease activity on a bubble substrate, acting instead as a weak single-stranded endonuclease (17). 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.

Materials and methods

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 6 His residues. This construct (pSVL-ERCC1-His) was transfected into 43-3B cells (28), and repair-competent transformants were selected after repeated UV-treatment (4.6J/m^2). To test survival after induction of DNA damage, cells were treated with UV or MMC (1hr), allowed to proliferate for 6 days, followed by pulse-labeling with tritiated thymidine (1 hr), and a chase (1 hr) to deplete radioactive precursor pools. Cells were lysed in 0.05M 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.

Purification of the ERCC1-XPF complex

ERCC1 protein 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 (12) was prepared from 1.4 x 10^{11} frozen cells, dialyzed against buffer A (20 mM Hepes-KOH pH7.5, 0.2 mM EDTA, 2 mM $MgCl_2$, 10% glycerol, 0.2 mM 4-(2-aminoethyl)benzenosulfonyl fluoride (AEBSF), 5 mM ß-mercaptoethanol) containing 0.15 M KCl and 10 µg/ml 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 contained the incision protein RPA and proteins not needed for NER (3, 18).

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) (25 ml/hr). The column was washed and eluted sequentially with buffer B (20 mM Hepes-KOH pH7.5, 2 mM $MgCl_2$, 10% glycerol, 0.2 mM AEBSF, 5 mM ß-mercaptoethanol, 0.5 M KCl) containing 1 mM, 5

mM (pH6.6), 20 mM (pH7.5) and 100 mM imidazole (pH7.5). ERCC1 was eluted in the last fraction (Fraction II; 35 mg protein), supplemented with 0.02% NP40 and 1 mM potassium phosphate, and loaded onto a hydroxyapatite column (Biorad; 3 ml). The column was washed with buffer C (25 mM Hepes-KOH pH7.8, 10% glycerol, 0.2 mM AEBSF, 2 mM dithiothreitol (DTT), 0.02% NP40) 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 Figure 2a. For Figure 2c, 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 DTT, and 100 mM KCl. After 1 hr at 4^0C, the suspension was centrifuged and 30 μl of supernatant was loaded onto the gel.

In vitro DNA repair and nuclease assays

The assay for dual incision was performed with closed circular (double-stranded) M13 DNA molecules containing a site-specific 1,3-intrastrand d(GpTpG)-cisplatin cross-link. In brief, these DNA substrates were constructed by treating a 24-mer oligonucleotide with cisplatin, annealing the platinated 24-mer to plus strand M13mp18GTGx and allowing DNA synthesis (T4 DNA polymerase gp43 subunit) and ligation (T4 DNA ligase); for details, see (14). Dual incision assays were performed as described (14), except that the plasmid was cleaved with HindIII and XhoI before detection by Southern hybridization. Repair synthesis assays used CFII fractions from

the indicated cells, purified RPA, and purified proliferating cell nuclear antigen (PCNA), as previously described (3).

Nuclease reaction mixtures (15 µl) contained 0.2-0.5 ng of stem-loop DNA and the indicated proteins in buffer as described (15) for figure 4a, lanes 1-10, or in buffer D (50 mM Tris pH8, 10 mM $MgCl_2$, 0.1 mg/ml bovine serum albumin (BSA) and 0.5 mM β-mercaptoethanol) for lanes 11-18. After incubation at 25^0C for 2 hrs, 15 µl 90% formaldehyde was added, and samples were heated at 95^0C 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 immunoprecipitations, affinity-purified anti-ERCC1 antiserum (30µl) was coupled to 10 µl of protein G beads, which were extensively washed and incubated with the indicated amount of ERCC1-XPF complex in 20 µl of buffer D at 4^0C. After 2 hrs, beads were washed twice with 100 µl of buffer D and added to the nuclease reaction mixture. Samples were then loaded on a denaturing 20% polyacrylamide gel. For the endlabeling analysis in figure 4c, reaction mixtures containing 5'-labeled (lanes 1 and 2) or unlabeled (lanes 3 and 4) stem-loop substrate were incubated for 16 hr at 16^0C, 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 of dGTP, dCTP, and dTTP and 2 nmol of α-^{32}P-dATP were added for 30 min, followed by a 90 min chase with 10 nmol of unlabeled dATP.

References

1. **Aboussekhra, A., M. Biggerstaff, M.K.K. Shivji, J.A. Vilpo, V. Moncollin, V.N. Podust, M. Protic, U. Hubscher, J.-M. Egly, and R.D. Wood.** 1995. Mammalian DNA nucleotide excision repair reconstituted with purified components. *Cell* **80:** 859-868.
2. **Bardwell, A.J., L. Bardwell, A.E. Tomkinson, and E.C. Friedberg.** 1994. Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. *Science* **265:** 2082-2085.
3. **Biggerstaff, M., D.E. Szymkowski, and R.D. Wood.** 1993. Co-correction of ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects *in vitro*. *EMBO J.* **12:** 3685-3692.
4. **Brookman, K., J. Lamerdin, M. Thelen, M. Hwang, J. Reardon, A. Sancar, Z. Zhou, C. Walter, C. Parris, and L. Thompson.** 1996. ERCC4 (XPF) encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Mol. Cell. Biol.* **16:** 6553-6562.
5. **Davies, A.F., 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.
6. **Friedberg, E.C., G.C. Walker, and W. Siede.** 1995. DNA repair and mutagenesis. ASM Press, Washington D.C.
7. **Habraken, Y.S., P; Prakash, L; Prakash, S.** 1995. Structure-specific nuclease activity in yeast nucleotide excision. *J. Biol.Chem.* **270:** 30194-30198.
8. **Harrington, J.J., and M.R. Lieber.** 1994. Functional domains within FEN-1 and RAD2 define a family of structure-specific endonucleases: implications for nucleotide excision repair. *Genes Dev.* **8:** 1344-1355.
9. **Hoeijmakers, J.H.J.** 1994. Human nucleotide excision repair syndromes: molecular clues to unexpected intricacies. *Eur. J. of Cancer* **30A:** 1912-1921.
10. **Huang, J.C., D.L. Svoboda, J.T. Reardon, and A. Sancar.** 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.* **89:** 3664-3668.
11. **Jones, C.J., and R.D. Wood.** 1993. Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochem.* **32:** 12096-12104.
12. **Masutani, C., K. Sugasawa, J. Yanagisawa, T. Sonoyama, M. Ui, T. Enomoto, K. Takio, K. Tanaka, P.J. van der Spek, D. Bootsma, J.H.J. Hoeijmakers, and F. Hanaoka.** 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.
13. **Matsunaga, T., D. Mu, C.-H. Park, J.T. Reardon, and A. Sancar.** 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.
14. **Moggs, J.G., K.J. Yarema, J.M. Essigmann, and R.D. Wood.** 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.
15. **O'Donovan, A., A.A. Davies, J.G. Moggs, S.C. West, and R.D. Wood.** 1994. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371:** 432-435.
16. **O'Donovan, A., D. Scherly, S.G. Clarkson, and R.D. Wood.** 1994. Isolation of active recombinant XPG protein, a human DNA repair endonuclease. *J. Biol. Chem.* **269:** 15965-15968.

17. **Park, C.-H., T. Bessho, T. Matsunaga, and A. Sancar.** 1995. Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *J. Biol. Chem.* **270:** 22657-22660.
18. **Shivji, M.K.K., M.K. Kenny, and R.D. Wood.** 1992. Proliferating cell nuclear antigen is required for DNA excision repair. *Cell* **69:** 367-374.
19. **Sijbers, A.M., W.L. de Laat, R.R. Ariza, M. Biggerstaff, Y.-F. Wei, J.G. Moggs, K.C. Carter, B.K. Shell, E. Evans, M.C. de Jong, S. Rademakers, J. de Rooij, N.G.J. Jaspers, J.H.J. Hoeijmakers, and R.D. Wood.** 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86:** 811-822.
20. **Sung, P., P. Reynolds, L. Prakash, and S. Prakash.** 1993. Purification and characterization of the *Saccharomyces cerevisiae* RAD1/RAD10 endonuclease. *J. Biol. Chem.* **268:** 26391-26399.
21. **Tomkinson, A.E., A.J. Bardwell, L. Bardwell, N.J. Tappe, and E.C. Friedberg.** 1993. Yeast DNA repair and recombination proteins RAD1 and RAD10 constitute a single-stranded-DNA endonuclease. *Nature* **362:** 860-862.
22. **Tomkinson, A.E., A.J. Bardwell, N. Tappe, W. Ramos, and E.C. Friedberg.** 1994. Purification of Rad1 protein from *Saccharomyces cerevisiae* and further characterization of the Rad1/Rad10 endonuclease complex. *Biochem.* **33:**5305-5311.
23. **van Duin, M., J. de Wit, H. Odijk, A. Westerveld, A. Yasui, M.H.M. Koken, J.H.J. Hoeijmakers, and D. Bootsma.** 1986. Molecular characterization of the human excision repair gene *ERCC-1:* cDNA cloning and amino acid homology with the yeast DNA repair gene *RAD10. Cell* **44:** 913-923.
24. **van Vuuren, A.J., E. Appeldoorn, H. Odijk, S. Humbert, V. Moncollin, A.P.M. Eker, N.G.J. Jaspers, J.-M. Egly, and J.H.J. Hoeijmakers.** 1995. Partial characterization of the DNA repair protein complex, containing the ERCC1, ERCC4, ERCC11 and XPF correcting activities. *Mutat. Res.* **337:** 25-39.
25. **van Vuuren, A.J., E. Appeldoorn, H. Odijk, A. Yasui, N.G.J. Jaspers, D. Bootsma, and J.H.J. Hoeijmakers.** 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.
26. **Wood, R.D.** 1996. DNA repair in eukaryotes. *Ann. Rev. of Biochem.* **65:** 135-167.
27. **Wood, R.D., M. Biggerstaff, and M.K.K. Shivji.** 1995. Detection and measurement of nucleotide excision repair synthesis by mammalian cell extracts *in vitro*, p. 163-175, *Methods enzymol.*, vol. 7.
28. **Wood, R.D., and H.J. Burki.** 1982. Repair capability and the cellular age response for killing and mutation induction after UV. *Mutat. Res.* **95:** 505-514.

Chapter 4

Mapping of interaction domains between the human nucleotide excision repair proteins ERCC1 and XPF

Abstract

ERCC1-XPF is a heterodimeric protein complex involved in nucleotide excision repair and recombinational processes. Like its homologous complex in *Saccharomyces 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 at the 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 the 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 the C-terminal residues 813 to 905. A mutation from an XP-F patient that alters these residues indeed affects complex formation with ERCC1.

Introduction

Nucleotide excision repair (NER) 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 (21, 28, 30). The remaining gap is filled by DNA synthesis and ligation (1, 17). The mammalian heterodimeric protein complex ERCC1-XPF is a structure-specific endonuclease that catalyzes the incision at the 5' side of the lesion during NER (11, 25, 41). Like XPG, which makes the 3' incision (34), ERCC1-XPF is thought to be positioned through protein-protein interactions with other repair factors around a partially unwound DNA intermediate (14, 15, 31). 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 (8). 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 (42). 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 XPA- and XPC- NER-deficient knock-out mice (12, 27,

32, 40, 45). In *Saccharomyces cerevisiae*, strains defective in the homologs of ERCC1 and XPF, Rad10 and Rad1, fail to complete recombination between direct repeated DNA sequences (16), and mutations in the homologous proteins of *Schizosaccharomyces pombe*, Swi10 and Rad16, can affect mating-type switching (18). 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 (22, 35, 37), and with the single-stranded DNA-binding protein RPA (Replication Protein A) (6, 26), which may stabilize the opened DNA complex (15, 31). The interaction with XPA occurs through the residues 93 to 120 in the ERCC1 protein (22). Recently, we demonstrated that RPA can modulate ERCC1-XPF incision activity such that cleavage is restricted to the damaged strand (de Laat *et al.*, submitted). 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 *Escherichia coli* NER protein UvrC (43). In UvrC, this region was found to be essential for its endonuclease activity (24), and deletion of this region specifically disrupted the 5' incision during NER in *E. coli* (29). Possibly, these residues are required to position the active cleavage site correctly onto the DNA (29). 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 (13, 42), including the structure-specific DNA nucleases FEN-1 and XPG (20, 23). 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 (41).

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 XPF 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 (7, 35, 41, 44), and complex formation *in vivo* was also observed when recombinant ERCC1 and XPF proteins were overproduced together in insect cells (6, 11) or *Escherichia 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

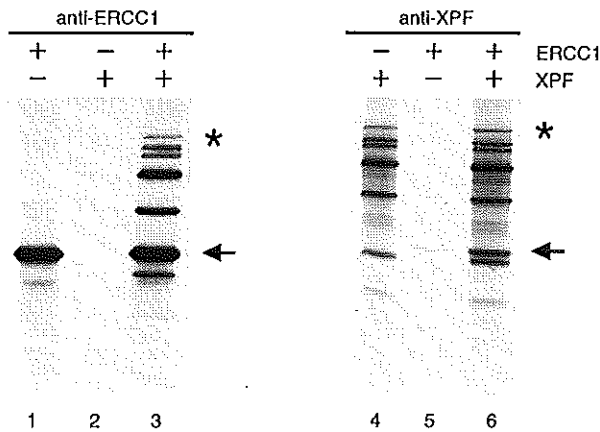


Figure 1. ***In vitro*** **translated ERCC1 and XPF reconstitute a complex.** Autoradiogram of (antibody beads-bound) ^{35}S-labeled proteins separated on 11% SDS-polyacrylamide gel. Input is indicated above each lane. ERCC1 and XPF *in vitro* translates directly analyzed on gel show identical products as observed in lane 1 and 4, respectively. Asterisk indicates full-length XPF protein, arrow indicates full-length ERCC1 protein.

ERCC1 and XPF were used that had previously been shown to be able to deplete the complex from whole cell extracts (41, 44). *In vitro* translation of XPF in a reticulocyte lysate-based transcription/translation system resulted not only in the 115K 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 wild-type ERCC1 with anti-XPF antibodies was detected only in the presence of XPF protein (Fig. 1, compare lanes 5 and 6). 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 that will give a conservative estimate of the domains responsible for complex formation.

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* (2, 4). Almost two-third of the Rad10 protein, stretching from residue 90 to residue 210 at the very C-terminus, was shown to be required for Rad1 binding. This region corresponds to amino acids 98-214 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, different from Rad10, deleting the N-terminal 103 residues (ERCC1-ATG103) and 118 residues (ERCC1-ATG118) did not seem to affect complex formation

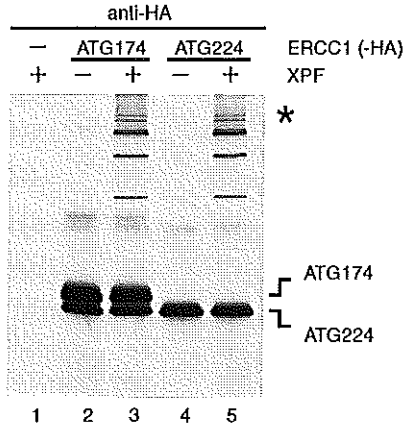


Figure 2. N-terminal truncations of ERCC1. Autoradiogram of (antibody beads-bound) ^{35}S-labeled proteins separated on 16.5% SDS-polyacrylamide gel. Note that only the ERCC1 protein fragments carry an HA-epitope tag. The lower band synthesized along with ERCC1-ATG174 (lane 2 and 3) originates from the in-frame alternative translational start at position 224 in ERCC1 and also carries an HA-epitope tag at the C-terminus. Altered migration pattern of XPF products (compared to figure 1) is due to different percentage of gel. Asterisk indicates full-length XPF protein.

either. Even subsequent truncations from the N-terminus did not abolish the XPF-binding capacity of ERCC1, and much to our suprise we found that an ERCC1 peptide lacking the complete region of homology to Rad10 (ERCC1-ATG224) still bound to XPF (Fig. 2). This interaction was both observed with an anti-XPF antibody (data not shown) and with an antibody directed against the 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. Unfortunately, we were unable to produce sufficient amounts of even smaller ERCC1 fragments with this synthesis procedure (ERCC1-ATG245 and -265, with predicted molecular weights of approximately 6 kD and 4 kD, respectively), and therefore did not find an N-terminally truncated ERCC1 fragment without affinity for XPF. Nevertheless, we can conclude that the N-terminal border of domain responsible for initial and stable binding to XPF resides beyond residue 224 in the ERCC1 protein (see Fig. 4).

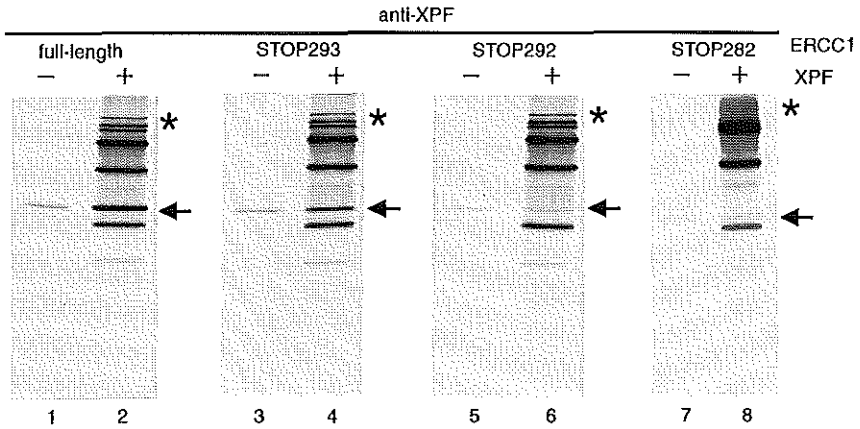


Figure 3. C-terminal truncations of ERCC1. Autoradiogram of (antibody beads-bound) ^{35}S-labeled proteins separated on 11% SDS-polyacrylamide gel. Arrows indicate (truncated) ERCC1 products, asterisks indicate full-length XPF. The odd lanes show aspecific binding of ERCC1 fragments to anti-XPF antibody 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 of XPF (odd lanes). In each precipitation (lane 1-8), similar amounts of ERCC1 products were used. Note that ERCC1-STOP293 specifically binds to XPF, but with reduced affinity compared to full-length ERCC1 (compare lanes 4 and 2).

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 5 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 Phe-293 to ERCC1-STOP292, yielding ERCC1-STOP293, reproducably restored partial affinity for XPF, demonstrating possibly a direct involvement of this phenylalanine in XPF binding (Fig. 3, lanes 3 and 4). However, ERCC1-STOP293 never coprecipitated with XPF as efficiently as full-length ERCC1 did, 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 residue 293 and residue 297, which is the last amino acid of full-length ERCC1 (see Fig. 4).

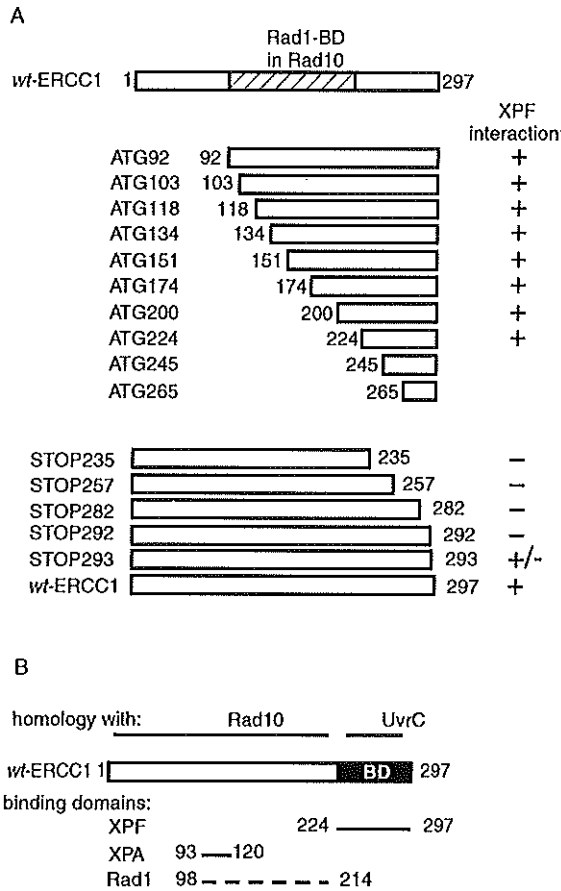


Figure 4. Schematic presentation of the XPF-binding domain in ERCC1. (A) Overview of XPF-interactions obtained with truncated ERCC1 fragments. On top is shown full-length ERCC1 with the domain that corresponds to the Rad1 binding domain (BD) in Rad10. +, interaction with XPF; +/-, intermediate interaction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

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

Figure 1 shows that a complete series of truncated XPF fragments originating from alternative translational start sites coprecipitate with wild-type ERCC1 when an anti-ERCC1 antibody is used (see Fig. 1, lane 3). This already indicates that the C-terminal part

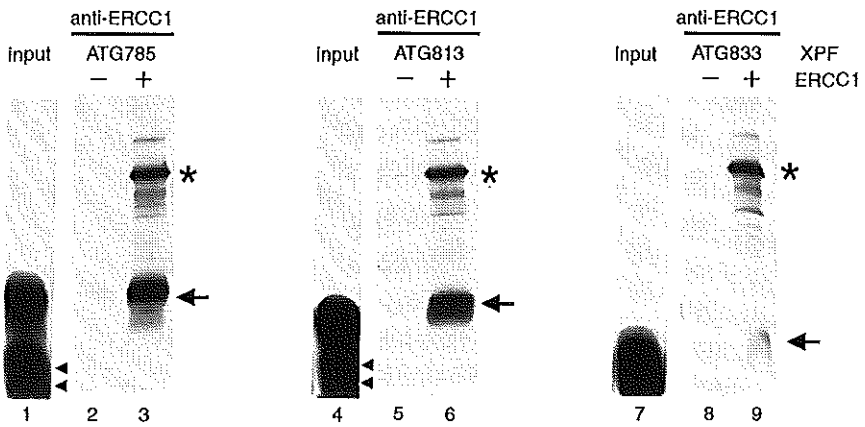


Figure 5. N-terminal truncations of XPF. Autoradiogram of (antibody beads-bound) ^{35}S-labeled proteins separated on 15% SDS-polyacrylamide gel. Arrows indicate truncated XPF product, asterisks indicate full-lenght ERCC1 protein, and arrowheads (Lanes 1 and 4) indicate XPF products originating from the alternative starts at position 844 and 856. Note that these two truncated XPF fragments do not coprecipitate 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 aspecific binding of XPF fragments to anti-ERCC1 antibody beads. Input is shown in this figure to demonstrate that although some XPF-ATG833 specifically coprecipitates with ERCC1 (compare lanes 9 and 8), the ERCC1-affinity of this XPF fragment is strongly reduced (compare lanes 9 and 7, versus lanes 6 and 4, and 3 and 1).

of XPF is responsible for ERCC1-binding, which would be in agreement with the Rad10-binding domain in Rad1, that was mapped to a region corresponding to residues 662-827 in XPF (2). 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 at 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 indeed comigrated 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 truncations). This demonstrated directly

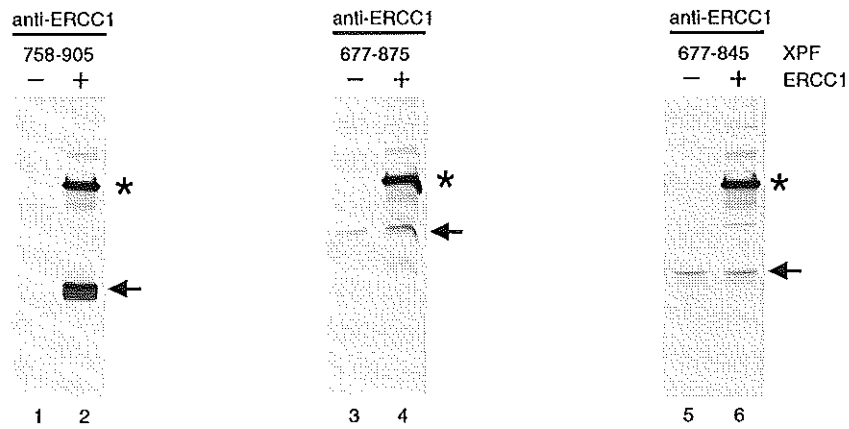


Figure 6. C-terminal truncations of XPF. Autoradiogram of (antibody beads-bound) ^{35}S-labeled proteins separated on 15% SDS-polyacrylamide gel. Arrows indicate XPF fragments, asterisks indicate full-length ERCC1 protein. Odd lanes show aspecific binding of XPF fragments to anti-ERCC1 antibody beads. Equal amounts of XPF products were used in each precipitation (lanes 1-6). Note that although the XPF fragments 677-875 (lanes 3 and 4) and 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.

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 the residues was missing from the N-terminus of the XPF protein; the peptides XPF-ATG758 (see Fig. 6, lanes 1 and 2), XPF-ATG785 (Fig. 5, lanes 1-3), and XPF-ATG813 (Fig. 5, lanes 4-6) 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 7-9). 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 with 60 and 30 amino acids, respectively. This

yielded the peptides XPF677-845 and XPF677-875. Unlike a (smaller) peptide that contained the very C-terminal residues of XPF (XPF758-905), neither XPF677-845 nor XPF677-875 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 (see Fig. 8).

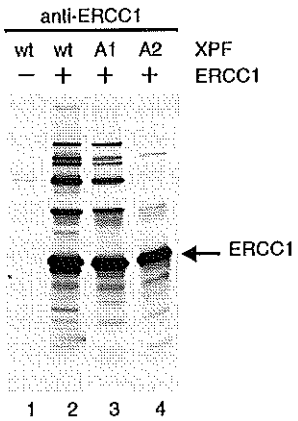


Figure 7. Naturally occurring XPF-mutation affects complex formation with ERCC1. Autoradiogram of (antibody beads-bound) ^{35}S-labeled proteins separated on 11% SDS-polyacrylamide gel. Arrow indicates ERCC1 protein. wt, wild-type XPF protein; A1, allele 1 of XP126LO patient, carrying the R788W mutation,; A2, allele 2, carrying the frameshift at position 757. Similar amounts of XPF were used in each precipitation (lanes 1-4).

Naturally occurring XPF mutations affect complex formation

Mutational analysis of an XP-F patient, XP126LO (33), 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 nt deletion causing a frameshift at residue 757 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 XPF mutations. Previous observations indicated that

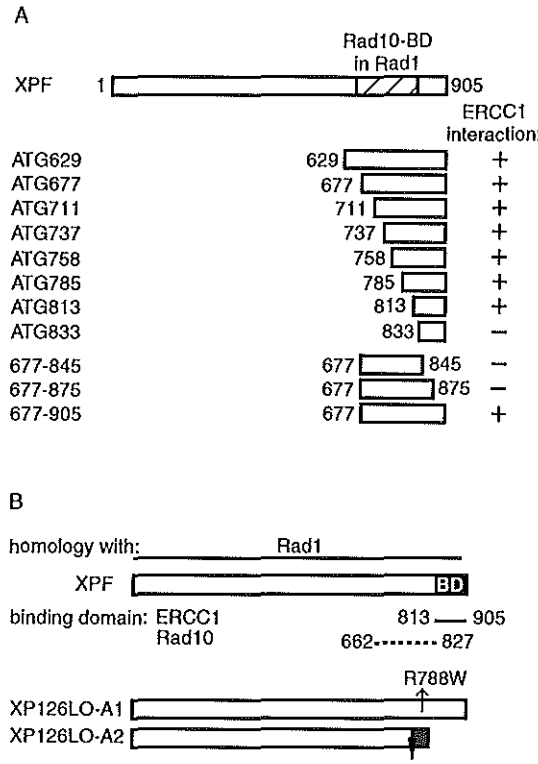


Figure 8. Schematic presentation of the ERCC1-binding domain in XPF. (A) Overview of ERCC1-interactions obtained with truncated XPF fragments. On top is shown full-length XPF with the domain that corresponds to the Rad10-binding domain in Rad1. +, interaction with ERCC1; -, no specific interaction with ERCC1. (B) Mapping of ERCC1-binding domain in XPF (BD), and schematic presentation XP126LO alleles. Arrow indicates position of R788W substitution in allele 1 of XP126LO. Triangle indicates position of frameshift, resulting in truncated protein in allele 2 of XP126LO.

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 (42). To test this directly, cDNAs encoding the two XP126LO alleles, were cloned into expression vectors and *in vitro* translated. Coprecipitations with wild-type ERCC1 translates revealed that the amino acid substitution R788W did not alter the

ERCC1-binding capacity of the XPF protein (Fig. 7, lane 3), whereas the product of the other allele had completely lost it (Fig. 7, 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.

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. Coprecipitation, therefore, requires that proteins bind eachother with relatively high affinity. A well-established interaction between ERCC1 and XPA for example, with a dissociation constant of 2.5×10^{-7} M (39), cannot be visualized with this procedure and coprecipitation of *in vitro* translated XPA is only detected with an excess of (recombinantly overproduced) ERCC1 on antibody beads (5) (unpublished observation). Thus, our binding-assay is a stringent 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 interactions. We define residues 224-297 in ERCC1 and residues 813-905 in XPF as the regions responsible for initial and stable complex formation between ERCC1 amd XPF.

Comparison between ERCC1-XPF and Rad1-Rad10

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* (19, 22, 36). Also, both display identical incision patterns on stem-loop substrates (41), and require similar DNA structural elements for nuclease activity (3,

11). 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 (2, 4). The binding domains in both Rad1 and Rad10 appeared to map to protein regions that are well conserved between the yeast and mammalian homologs. Suprisingly 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 98-214 in ERCC1 (2), but the XPF-binding domain in ERCC1 comprises a stretch of amino acids (224-297) that resides even outside the region of homology with Rad10 (see Fig. 4b). This C-terminal extension in ERCC1 is predominantly composed of a double helix-hairpin-helix (HhH-) motif (residues 236-289). 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 had shown that a 'Rad10-like' ERCC1 protein, with a stop at residue 214, was functionally inactive (43). 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 helix-hairpin-helix motif in ERCC1 is not directly involved in XPF-binding. Our data demonstrate that residues 293-297 in ERCC1 are important for XPF-binding. Interestingly, yet unobserved homology exists between the C-terminal residues 293-296 in ERCC1, comprising the amino acids Phe-Leu-Lys-Val, and the final four residues of Rad10 (207-210), which are Tyr-Leu-Asn-Leu. Although the Rad1-binding motif in Rad10 was reported to extend to residue 210 (2), 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 helix-hairpin-helix motif present in ERCC1 may support the correct structural positioning of the very C-terminal residues of ERCC1. The functional relevance of an additional, putative DNA-binding domain in ERCC1 still has to be resolved, though. It is worth mentioning that also *S. Pombe* Swi10 (38) and plant homologs of ERCC1 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 662-827 in XPF, but we find that residues 813-905 in XPF are required for actual ERCC1 binding. Motifs with a putative function have not been found in this region (41). 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 (2).

Interpretation of ERCC1 and XPF 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* (42). 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, affected ERCC1-XPF complex formation seems not involved. Rather, a reduced affinity for XPA, which binding domain maps to residues 93-120 in ERCC1, may underly this phenotype. C-terminal ERCC1 truncations demonstrated a direct link between complex formation ability and repair capacity. A deletion of 5 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* (42). However, addition of only one residue to this peptide, Phe-293 in ERCC1-STOP293, not only partially restored the XPF-binding (Fig. 3), but was also sufficient for partial restoration of the *in vivo* repair capacity of ERCC1 (42). Apparently, the phenylalanine at position 293 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 (46) (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 (9). 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 757, 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. Possibly, our assay is not sensitive enough to detect the impact of this mutation on ERCC1-affinity, or alternatively, the R788W mutation alters the XPF turn-over rate *in vivo*. In either way, it can be expected that the residual complexes observed in this patient contain XPF molecules carrying the R788W substitution and that this mutated form of XPF is responsible for the phenotypical features of the XP126LO patient. 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, reduced life span and signs of premature senescence (27, 45). As most XP-F patients tested have reduced ERCC1 and XPF protein levels (46) (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-

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 HindIII, RsaI, FokI and EagI were used, respectively, to remove upstream coding cDNA in front of these in-frame ATG codons. HA-epitope tags were introduced as described previously (11). C-terminal truncations were made via PCR using anti-sense primers containing (3' to 5') 18-24 complementary nucleotides, a translational stop sequence and 6-9 random nucleotides. cDNA from the two alleles of XP126LO was amplified, subcloned and sequenced as described (41).

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. For maximal protein yield, the amount of each PCR product, phenol extracted and ethanol precipitated, varied between 2-10 µg per 50 µl reaction mix. Polyclonal antibodies against ERCC1 and XPF (affinity-purified) used for immunoprecipitations were described before (41, 44). For immunoprecipitations, (truncated) ERCC1 protein and (truncated) XPF protein were mixed with 5 µl NETT-buffer (100 mM NaCl, 5 mM EDTA, 50 mM Tris pH 7.5 and 0.5% Triton-X100) (total volume 15-20 µl) and incubated for 30 min at 30^{0}C. Antibody was added and after 2 hours at 4^{0}C, 100 µl NETT-buffer containing 10% protein A-sepharose beads was added. Incubation was proceeded for another 2 hours at 4^{0}C (tumbling), then beads were washed 4 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 sample buffer, instead of coomassie (10). Dried gels were analyzed by autoradiography or a phosphorimager.

References

1. **Aboussekhra, A., M. Biggerstaff, M.K.K. Shivji, J.A. Vilpo, V. Moncollin, V.N. Podust, M. Protic, U. Hubscher, J.-M. Egly, and R.D. Wood.** 1995. Mammalian DNA nucleotide excision repair reconstituted with purified components. *Cell* **80:** 859-868.
2. **Bardwell, A.J., L. Bardwell, D.K. Johnson, and E.C. Friedberg.** 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.
3. **Bardwell, A.J., L. Bardwell, A.E. Tomkinson, and E.C. Friedberg.** 1994. Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. *Science* **265:** 2082-2085.
4. **Bardwell, L., A.J. Cooper, and E.C. Friedberg.** 1992. Stable and specific association between the yeast recombination and DNA repair proteins RAD1 and RAD10 *in vitro*. *Mol. Cell. Biol.* **12:** 3041-3049.
5. **Bessho, T., D. Mu, and A. Sancar.** 1997. 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.
6. **Bessho, T., A. Sancar, L.H. Thompson, and M.P. Thelen.** 1997. Reconstitution of human excision nuclease with recombinant XPF-ERCC1complex. *J. Biol. Chem.* **272:** 3833-7.
7. **Biggerstaff, M., D.E. Szymkowski, and R.D. Wood.** 1993. Co-correction of ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects *in vitro*. *EMBO J.* **12:** 3685-3692.
8. **Busch, D., C. Greiner, K. Lewis, R. Ford, G. Adair, and L. Thompson.** 1989. Summary of complementation groups of UV-sensitive CHO mutants isolated by large-scale screening. *Mutagen.* **4:** 349-354.
9. **Chen, F., S. Peterson, M. Story, and D. Chen.** 1996. Disruption of DNA-PK in Ku80 mutant xrs-6 and the implications in DNA double-strand break repair. *Mutat Res* **362:**9-19.
10. **Christy, K.J., D. LaTart, and W. Osterhoudt.** 1989. Modifications for SDS-PAGE of proteins. *BioTechn.* **7:** 692-693.
11. **de Laat, W.L., E. Appeldoorn, N.G.J. Jaspers, and J.H.J. Hoeijmakers.** 1998. DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.* **273:** 7835-7842.
12. **De Vries, A., C.T.M. van Oostrom, F.M.A. Hofhuis, P.M. Dortant, R.J.W. Berg, F.R. de Gruijl, P.W. Wester, C.F. van Kreijl, P.J.A. Capel, H. van Steeg, and S.J. Verbeek.** 1995. Increased susceptibility to ultraviolet-B and carcinogens of mice lacking the DNA excision repair gene *XPA*. *Nature* **377:** 169-173.
13. **Doherty, A.S., 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.
14. **Evans, E., J. Fellows, A. Coffer, and R.D. Wood.** 1997. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J.* **16:** 625-638.
15. **Evans, E., J. Moggs, J. Hwang, J. Egly, and R. Wood.** 1997. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J.* **16:** 6559-6573.
16. **Fishman-Lobell, J., and J.E. Haber.** 1992. Removal of nonhomologous DNA ends in double-strand break recombination: The role of the yeast ultraviolet repair gene *RAD1*. *Science* **258:** 480-484.

17. **Friedberg, E.C., G.C. Walker, and W. Siede.** 1995. DNA repair and mutagenesis. ASM Press, Washington D.C.
18. **Gutz, H., and H. Schmidt.** 1985. Switching genes in *Schizosaccharomyces pombe*. *Curr. Genet.* **9:** 325-331.
19. **Guzder, S., P. Sung, L. Prakash, and S. Prakash.** 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.
20. **Harrington, J.J., and M.R. Lieber.** 1994. The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J.* **13:** 1235-1246.
21. **Huang, J.C., D.L. Svoboda, J.T. Reardon, and A. Sancar.** 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.* **89:** 3664-3668.
22. **Li, L., S.J. Elledge, C.A. Peterson, E.S. Bales, and R.J. Legerski.** 1994. Specific association between the human DNA repair proteins XPA and ERCC1. *Proc. Natl. Acad. Sci.* **91:** 5012-5016.
23. **Lieber, M.** 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* **19:** 233-240.
24. **Lin, J.J., and A. Sancar.** 1991. The C-terminal half of UvrC protein is sufficient to reconstitute (A)BC excinuclease. *Proc. Natl. Acad. Sci.* **88:** 6824-6828.
25. **Matsunaga, T., D. Mu, C.-H. Park, J.T. Reardon, and A. Sancar.** 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.
26. **Matsunaga, T., C.H. Park, T. Bessho, D. Mu, and A. Sancar.** 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.
27. **McWhir, J., J. Seldridge, D.J. Harrison, S. Squires, and D.W. Melton.** 1993. Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat. Gen.* **5:** 217-223.
28. **Moggs, J.G., K.J. Yarema, J.M. Essigmann, and R.D. Wood.** 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.
29. **Moolenaar, G., R. Uiterkamp, D. Zwijnenburg, and N. Goosen.** 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.
30. **Mu, D., C.-H. Park, T. Matsunaga, D.S. Hsu, J.T. Reardon, and A. Sancar.** 1995. Reconstitution of human DNA repair excision nuclease in a highly defined system. *J. Biol. Chem.* **270:** 2415-2418.
31. **Mu, D., M. Wakasugi, D. Hsu, and A. Sancar.** 1997. Characterization of reaction intermediates of human excision repair nuclease. *J. Biol. Chem.* **272:** 28971-28979.
32. **Nakane, H., S. Takeuchi, S. Yuba, M. Saijo, Y. Nakatsu, T. Ishikawa, S. Hirota, Y. Kitamura, Y. Kato, Y. Tsunoda, H. Miyauchi, T. Horio, T. Tokunaga, T. Matsunaga, O. Nikaido, Y. Nishimune, Y. Okada, and K. Tanaka.** 1995. High incedence of ultraviolet-B- or chemical-carcinogen-induced skin tomours in mice lacking the xeroderma pigmentosum group A gene. *Nature* **377:** 165-168.
33. **Norris, P.G., J.L.M. Hawk, and F. Giannelli.** 1988. Xeroderma pigmentosum complementation group F in a non-Japanese patient. *J. Am. Acad. Dermatol.* **18:** 1185-1188.

34. **O'Donovan, A., A.A. Davies, J.G. Moggs, S.C. West, and R.D. Wood.** 1994. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371:** 432-435.
35. **Park, C.-H., T. Bessho, T. Matsunaga, and A. Sancar.** 1995. Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *J. Biol. Chem.* **270:** 22657-22660.
36. **Park, C.-H., and A. Sancar.** 1994. Formation of a ternary complex by human XPA, ERCC1 and ERCC4(XPF) excision repair proteins. *Proc. Natl. Acad. Sci.* **91:** 5017-5021.
37. **Robins, P., C.J. Jones, M. Biggerstaff, T. Lindahl, and R.D. Wood.** 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.
38. **Rodel, C., S. Kirchhoff, and H. Schmidt.** 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.
39. **Saijo, M., I. Kuraoka, C. Masutani, F. Hanaoka, and K. Tanaka.** 1996. Sequential binding of DNA repair proteins RPA and ERCC1 to XPA in vitro. *Nucleic Acids Res* **24:**4719-24.
40. **Sands, A.T., A. Abuin, A. Sanchez, C.J. Conti, and A. Bradley.** 1995. High susceptibility to ultraviolet-induced carcinogenesis in mice lacking *XPC. Nature* **377:** 162-165.
41. **Sijbers, A.M., W.L. de Laat, R.R. Ariza, M. Biggerstaff, Y.-F. Wei, J.G. Moggs, K.C. Carter, B.K. Shell, E. Evans, M.C. de Jong, S. Rademakers, J. de Rooij, N.G.J. Jaspers, J.H.J. Hoeijmakers, and R.D. Wood.** 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86:**811-822.
42. **Sijbers, A.M., P.J. van der Spek, H. Odijk, J. van den Berg, M. van Duin, A. Westerveld, N.G.J. Jaspers, D. Bootsma, and J.H.J. Hoeijmakers.** 1996. Mutational analysis of the human nucleotide excision repair gene *ERCC1. Nucleic Acids Res.* **24:** 3370-3380.
43. **van Duin, M., J. van den Tol, P. Warmerdam, H. Odijk, D. Meijer, A. Westerveld, D. Bootsma, and J.H.J. Hoeijmakers.** 1988. Evolution and mutagenesis of the mammalian excision repair gene ERCC-1. *Nucleic Acids Res.* **16:** 5305-5322.
44. **van Vuuren, A.J., E. Appeldoorn, H. Odijk, A. Yasui, N.G.J. Jaspers, D. Bootsma, and J.H.J. Hoeijmakers.** 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.
45. **Weeda, G.D., I; de Wit, J; Morreau, H; Janssens, R; Vissers, CJ; Nigg, A; van Steeg, H; Bootsma, D; Hoeijmakers, JHJ.** 1997. Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr. Biol.* **7:** 427-439.
46. **Yagi, T., R.D. Wood, and H. Takebe.** 1997. A low content of ERCC1 and a 120 kDa protein is a frequent feature of group F xeroderma pigmentosum fibroblast cells. *Mutagen.* **12:** 41-44.

Chapter 5

DNA structural elements required for ERCC1-XPF endonuclease activity

Abstract

The heterodimeric complex ERCC1-XPF is a structure-specific endonuclease responsible for the 5' incision during mammalian nucleotide excision repair (NER). Additionally, ERCC1-XPF is thought to function in the repair of interstrand DNA crosslinks and, by analogy to the homologous Rad1-Rad10 complex in *Saccharomyces cerevisiae*, in recombination between direct repeated DNA sequences. To gain insight into the role of ERCC1-XPF in such recombinational processes and in the NER reaction, we studied in detail the DNA structural elements required for ERCC1-XPF endonucleolytic activity. Recombinant ERCC1-XPF, purified from insect cells, was found to cleave stem-loop substrates at the DNA junction in the absence of other proteins like RPA, showing that the structure-specific endonuclease activity is intrinsic to the complex. Cleavage depended on the presence of divalent cations and was optimal in low Mn^{2+}-concentrations (0.2 mM). A minimum of 4 to 8 unpaired nucleotides was required for incisions by ERCC1-XPF.

Published in J. Biol. Chem., vol. 273 (14), 7835-7842 (1998)

Splayed arm and flap substrates were also cut by ERCC1-XPF, resulting in the removal of 3' protruding single-stranded arms. All incisions occurred in one strand of duplex DNA at the 5' side of a junction with single-stranded DNA. The exact cleavage position varied from 2 to 8 nucleotides away from the junction. One single-stranded arm, protruding either in the 3' or 5' direction, was necessary and sufficient for correct positioning of incisions by ERCC1-XPF. Our data specify the engagement of ERCC1-XPF in NER and allow a more direct search for its specific role in recombination.

Introduction

Nucleotide excision repair (NER[1]) guards the integrity of the genome by removing bulky adducts and the most prominent UV-induced lesions from the DNA. During NER, the damaged DNA strand is incised asymmetrically around the lesion to allow release of the damage as part of a 24 to 32 nucleotides DNA fragment, followed by resynthesis and ligation (9, 14, 24). Whereas the latter steps are accomplished by general replication factors such as PCNA, RFC, DNA polymerase δ and/or ϵ and ligase, the incision stage of NER requires the action of XP-factors (XPA-G) (1). XPA, and perhaps XPC complexed with hHR23B, are thought to be involved in damage recognition; XPB and XPD, two helicases present in the basal transcription factor complex TFIIH, and replication protein A (RPA) presumably demarcate the lesion by local opening of the DNA; XPG and the heterodimeric complex ERCC1-XPF perform the actual DNA cleavage. The role of XPE is, as yet, unclear. Mutations in either of the XP-factors can cause the extreme UV-sensitive and cancer-prone phenotype observed with *Xeroderma Pigmentosum* (XP) patients. Some human XP-genes cross-complement the NER-defect of laboratory-induced UV-sensitive Chinese hamster mutant cell lines, subdivided into 10 complementation groups. Human XPF, for example, corrects the repair-defect of rodent group 4 cell lines. The human Excision Repair Cross-Complementing gene ERCC1 corrects rodent group 1 cell lines, but no ERCC1-deficient patients are known. ERCC1 and XPF form an endonuclease that

incises the damaged strand 15 to 24 nucleotides away at the 5' side of the lesion (26, 31). Cleavage by ERCC1-XPF and XPG, which makes the 3' incision during NER, presumably occurs in a structure-specific manner, as both NER nucleases cut artificial DNA substrates preferentially at duplex-single stranded junctions (28, 31). Incisions made by ERCC1-XPF and XPG during NER colocalize to the borders of the opened DNA intermediate (7) and protein-protein interactions with other NER factors are likely to determine the exact positioning of both nucleases. For ERCC1-XPF, interactions with XPA and RPA have been reported (3, 19, 22, 30).

The counterpart of ERCC1-XPF in *Saccharomyces cerevisiae*, Rad10-Rad1, is required not only for NER but also for recombination between direct repeated DNA sequences, a process which is also called single strand annealing (8, 16). By analogy, ERCC1-XPF is expected to fulfil a similar role in mammalian cells, and the two homologous protein complexes have been proposed to remove 3' protruding single-stranded arms from recombined duplex regions (5, 8). Chinese hamster mutant cell lines defective in either ERCC1 (group 1) or XPF (group 4) are unique among the other NER-deficient mutants in that they show an extreme sensitivity to DNA crosslinking agents such as cisplatin and mitomycin C (4). Repair of interstrand crosslinks probably also requires a recombinational event.

In order to gain insight in the role of ERCC1-XPF in such recombinational processes and in the NER reaction, it is of importance to define its DNA substrate specificity. Similar studies with the FEN-1 family of structure-specific nucleases revealed interesting characteristics of individual members. Flap endonuclease-1 (FEN-1), implicated in both *in vitro* replication (15) and *in vitro* repair (18), was found to cleave branched DNA structures with 5' protruding single-stranded arms (flaps, splayed arms), but not with closed single-stranded regions (loops, bubbles) (12). XPG, also a member of this family and as yet only known to be involved in NER, cleaves both stem-loop and bubble substrates as well as flaps and splayed arms (7, 28). Similar results were obtained with the *S. cerevisiae* homolog of XPG, Rad2, and all three members of this nuclease family cleaved DNA with similar polarity, i.e. in one strand of duplex DNA at the 3' side of the junction with single-stranded DNA (11,

13). The differences in substrate specificity between XPG and Rad2 on the one hand and FEN-1 on the other, may point at their different roles in the cell. The well-defined structural DNA requirements of FEN-1 have been the key to the recent disclosure of its role *in vitro* in a base excision repair subpathway (18). ERCC1-XPF, and its yeast homolog Rad1-Rad10, are not part of a larger family of structure-specific nucleases. Rad1-Rad10 was previously shown to cleave flaps, splayed arms, stem-loops and bubbles at the 5' side of duplex-single-stranded junctions (2). Intriguingly, the sole domain shared between ERCC1 and FEN-1 nucleases, a presumed helix-hairpin-helix motif present twice at the C-terminus of ERCC1, is lacking in Rad10 (32, 35). This double motif, in total approximately 50 residues long, is assumed to be involved in non-sequence specific DNA binding (6), and might endow ERCC1-XPF with additional properties or altered substrate specificity compared to Rad1-Rad10.

Here we report a detailed analysis of the DNA structural elements required for ERCC1-XPF cleavage. We found that ERCC1-XPF needs at least 4 to 8 unpaired nucleotides to cleave stem-loop substrates. Single-stranded arms protruding in the 3' direction were efficiently removed by ERCC1-XPF from splayed arm and flap substrates. Incisions occurred in duplex DNA at the 5' side of a junction with single-stranded DNA. Both 3' and 5' protruding arms could be used by ERCC1-XPF for correctly positioning its nuclease activity. These data provide a framework to envisage ERCC1-XPF activity in repair as well as in recombinational pathways.

Results

Characterization of purified recombinant ERCC1-XPF

To reconstitute recombinant ERCC1-XPF protein complex, Sf21 insect cells were co-infected with ERCC1- and XPF-containing baculovirus expression constructs. Both ERCC1 and XPF cDNAs encoded C-terminal affinity tags, a 6xHis-tag and a 6xHis-HA-epitope double-tag respectively, to facilitate purification of the proteins. Neither of these tags

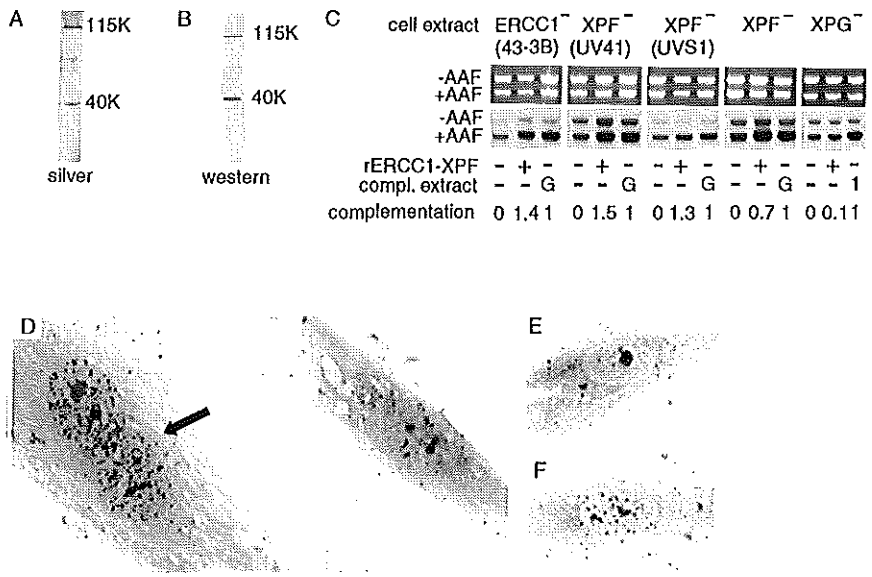


Figure 1. Characterization of purified recombinant ERCC1-XPF. (A) Silver-stained protein profile of top-fraction from glycerol gradient. (B) Immunoblot analysis of same fraction, using anti-ERCC1 and anti-XPF antibodies. (C) *In vitro* repair synthesis assay. Upper panels show ethidium bromide stained DNA, lower panels show autoradiogram of same agarose gel. 43-3B, Chinese hamster group 1 cell line; UV41, group 4 cell line; UVS1, group 4 cell line (former group 11); other cell lines originate from XP-patients. 'G' and '1' indicate that an XPG- and an ERCC1-deficient cell extract was used for complemention, respectively. Below each lane is indicated the amount of complementation obtained with purified ERCC1-XPF protein relative to that obtained with a complementing cell extract. Numbers were obtained by quantitation of specific incorporation of radiolabel into the AAF-damaged plasmid (i.e. incorporation in damaged minus undamaged plasmid), taking into account the amount of DNA present in each lane. Note that purified ERCC1-XPF complex is able to complement 43-3B and UVS1 cell extracts, as reflected by the increased levels of incorporation into damaged DNA.
(D-F) *In vivo* repair synthesis assay. (D) XPF fibroblasts (arrow indicates injected multikaryon). (E) Uninjected and (F) injected ERCC1 mouse embryonal fibroblasts (MEFs). Nuclei were micro-injected with purified ERCC1-XPF. Number of grains above the nuclei, used as a measure for repair capacity, was increased 5-10 times both in XPF⁻ primary fibroblasts and in ERCC1⁻ MEFs upon injection of purified ERCC1-XPF.

interfered with the function of the proteins *in vivo*, as micro-injection of XPF-(6xHis-HA) cDNA into the nucleus of XPF-deficient human fibroblasts and transfection of ERCC1-(6xHis) cDNA into ERCC1-deficient Chinese hamster (CHO-) cells fully corrected the

repair defects of these mutant cell types (data not shown). A three-step purification protocol, which included phosphocellulose, nickel affinity chromatography and a glycerol gradient, resulted in highly purified fractions containing 2 major proteins on silverstain with apparent molecular weight of approximately 40K and 115K (Fig. 1A). Immunoblot analysis with polyclonal antibodies confirmed these proteins to be ERCC1 and XPF (Fig. 1B). On the glycerol gradient the majority of ERCC1 and XPF was found to co-migrate, which demonstrated that the proteins reconstitute a stable ERCC1-XPF protein complex in insect cells. As only little further purification was obtained with this gradient and no significant differences were detected between the enzymatic activities of Ni-purified- versus glycerol gradient-purified-protein fractions, most experiments were performed with Ni-purified proteins. During the progress of this work also recombinant ERCC1-XPF was purified from E. coli, which was used to confirm that the activities described below were intrinsic to the ERCC1-XPF complex (data not shown).

The activity of recombinant ERCC1-XPF was tested in *in vitro* and *in vivo* repair assays. *In vitro*, ERCC1-XPF was found to correct the repair-deficient phenotype of cell extracts from human XP-F and Chinese hamster complementation groups 1 and 4, but not from XP-G (Fig. 1C). Correction was also found with an extract from the Chinese hamster mutant cell line UVS1, formerly known as the sole representative of complementation group 11, but recently shown to be functionally deficient in XPF as well (31). To test the activity of recombinant ERCC1-XPF *in vivo*, purified complex was micro-injected into primary fibroblasts derived from an XP-F patient. The injected cells showed a significant increase of unscheduled DNA synthesis after UV treatment, indicative for restored NER capacity (Fig. 1D). Chinese hamster cells are not very suitable for micro-injection for technical reasons, but the establishment of Mouse Embryonal Fibroblasts (MEFs) from a repair-deficient ERCC1^{-} mouse (37) enabled us to analyse protein fractions for ERCC1-correcting activity *in vivo*. Micro-injection of purified ERCC1-XPF in these cells resulted in a significant increase in repair activity as well (Fig. 1D). These data demonstrate that the purified recombinant ERCC1-XPF complex effectively supports NER both *in vitro* and *in vivo*.

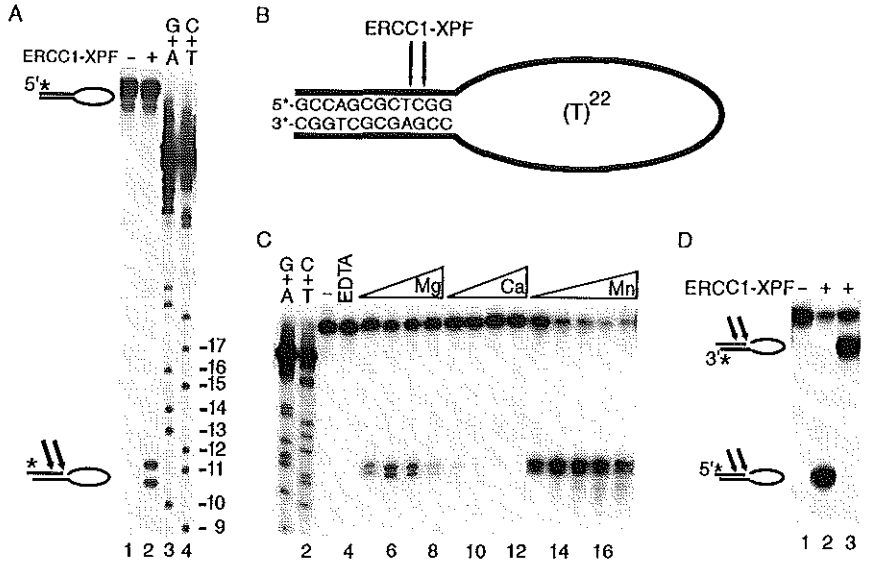


Figure 2. Recombinant ERCC1-XPF cleaves stem-loop substrate at DNA junction. (A) Incision assay with stem12-loop22. *Lane 1*: no protein added. *Lane 2*: 20 ng ERCC1-XPF. *Lane 3* and *4*: Maxam-Gilbert (MG) sequence ladders, obtained from stem21-loop4. Note that 5' labeled MG sequence products run approximately 1.5 nt faster than ERCC1-XPF cleavage products. Asterisks indicate position of radioactive label in the substrate. (B) Schematic presentation of ERCC1-XPF incisions in stem12-loop22 substrate. (C) Divalent cation requirements for ERCC1-XPF cleavage. *Lane 1-2*: MG-ladders. *Lane 3*: no protein. *Lane 4-17*: 20 ng ERCC1-XPF. *Lane 4*: 5 mM EDTA. *Lane 5-8*: 1, 2, 4, 8 mM $MgCl_2$, respectively. *Lane 9-12*: 1, 2, 4, 8 mM $CaCl_2$. *Lane 13-17*: 0.1, 0.2, 0.5, 1, 2 mM $MnCl_2$. *D*, absence of aspecific (single-stranded) DNA degradation in 0.2 mM Mn^{2+}-nuclease buffer. *Lane 1*: 5' labeled stem-loop substrate, no ERCC1-XPF. *Lane 2*: 5' labeled stem-loop substrate with 20 ng ERCC1-XPF. *Lane 3*: 3' labeled stem-loop substrate with 20 ng ERCC1-XPF.

Recombinant ERCC1-XPF incises DNA at duplex-single strand junctions

Previously, we have shown that a His-tagged ERCC1-XPF complex purified from CHO cells efficiently incises artificial DNA substrates containing a duplex-single stranded DNA junction (31). The observed ERCC1-XPF endonuclease activity was structure-specific and independent of other proteins like RPA. To test nuclease activity of our recombinant ERCC1-XPF, similar assays with such stem-loop substrates were performed. Rather than

random DNA sequences, our artificial substrates exclusively contained thymine residues in the single-stranded regions to minimize formation of secondary DNA structures. Two major incisions were detected when recombinant ERCC1-XPF was applied to these substrates, being in one strand of the duplex region, two and three nucleotides (nt) away at the 5' side of the junction (Fig. 2A and 2B). This cleavage pattern was identical to that observed with endogenous ERCC1-XPF purified from CHO cells (31).

To define optimal conditions for the nuclease activity we analysed scissions at various concentrations of divalent metal ions. As also observed with the CHO ERCC1-XPF complex, incision activity was absolutely dependent on the presence of divalent cations, as addition of EDTA completely blocked incisions (Fig. 2C, *lane* 4). Both Mg^{2+} and Mn^{2+} stimulated the endonuclease activity and, surprisingly, Mn^{2+} appeared to stimulate much better than Mg^{2+}. Optimal activity was observed in low concentrations of Mn^{2+} (0.2 mM) (Fig. 2C, *lane* 14) and was approximately 3 times higher than in the most optimal Mg^{2+}-buffer (Fig. 2C, *lane* 6). In contrast, incision activity was inhibited by increasing amounts of Ca^{2+} (*lanes* 9-12), showing that Ca^{2+} cannot substitute for Mg^{2+} and Mn^{2+} in these reactions. The reaction products appearing at 0.2 mM Mn^{2+} were similar to those observed in all Mg^{2+}-buffers tested (Fig. 2C), demonstrating that low Mn^{2+}-concentrations were suitable for studying ERCC1-XPF activity. A control reaction in 0.2 mM $MnCl_2$ without ERCC1-XPF (but with a corresponding amount of ERCC1-XPF dilution buffer B) showed that the DNA substrate is not subject to aspecific degradation under these conditions (Fig. 2D, *lane* 1); furthermore, when the same stem-loop substrate was labeled at the 3' end, ERCC1-XPF cleavage at 0.2 mM Mn^{2+} resulted in a relatively large incision product consisting of the complete single-stranded loop and one strand of the stem, which demonstrated the absence of single-stranded endonucleolytic degradation (Fig. 2D, *lane* 3). In view of these results, all assays described below were performed in a standard buffer containing 0.2 mM $MnCl_2$.

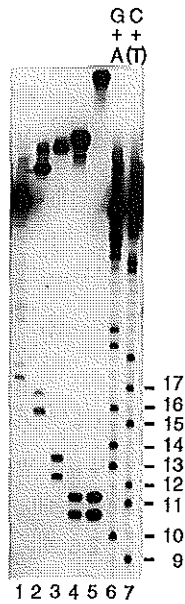


Figure 3. Determination of the minimal single-stranded loop size. *Lane 1-5*: 20 ng ERCC1-XPF added. *Lane 1*: Stem19-loop8. *Lane 2*: stem17-loop12. *Lane 3*: stem14-loop18. *Lane 4*: stem-12-loop22. *Lane 5*: stem12-loop40.

Minimal loop-size incised by ERCC1-XPF

Determination of the minimal size of the single-stranded loop in a stem-loop substrate can provide indications about the minimal degree of helix opening required for ERCC1-XPF incision, a relevant parameter for NER. In order to investigate this, stem-loops with varying loop- and stem-sizes but with a similar DNA context around the junction were constructed (see table 1). If possible, the total amount of nucleotides was kept the same among these substrates. Incisions were observed on stems with a loop of 12, 18, 22 and 40 nucleotides, whereas a loop of 8 nucleotides only gave minor cleavage products (Fig. 3). All incisions were made at identical positions relative to the junction, 2 and 3 nucleotides away in the stem. Occasional incisions were seen in substrates with a loop-size of 4

nucleotides (data not shown), but none were found in a 2 nucleotides loop and a fully paired hairpin structure (see also Fig. 4). This demonstrated that DNA substrates with less than 4 to 8 unpaired bases become poor DNA substrates for ERCC1-XPF mediated incision. Comparison of equally large stems (12 bp) with different loop sizes (22nt *vs* 40nt) showed that incision efficiency increased with the size of the single-stranded loop (Fig. 3, *lanes* 4 and 5). We could not determine the minimal duplex-size requirements for ERCC1-XPF activity as base-pairing will be lost with very small stems , but apparently a 12 bp stem is still sufficient for ERCC1-XPF mediated cleavage.

Minimal DNA structural elements required for ERCC1-XPF mediated incision

A role for ERCC1-XPF has been anticipated not only in NER, but also in repair of interstrand crosslinks and in single-strand annealing. The DNA intermediates that are recognized and cut by ERCC1-XPF during these events will probably differ from eachother but they most likely share the simplest DNA structural motif that can be cleaved by the complex. Since the exact DNA structural elements required for endonucleolytic activity of ERCC1-XPF are still largely unknown, we decided to dissect the basic stem-loop substrate as listed in Figure 4A.

A partially self-complementary oligonucleotide with two unpaired arms (splayed arm substrate) was incised by ERCC1-XPF only on one side of the duplex near the junction, - causing removal of the 3' protruding single-stranded arm. The polarity of this incision corresponded to those observed in stem-loop substrates. Comparison to Maxam-Gilbert ladders revealed that the major incision was made 4 nucleotides away from the DNA junction in the stem region (Fig. 4B). To investigate whether ERCC1-XPF required two or just one single-stranded arm for recognition and incision, self-annealing substrates were used having a constant duplex DNA region with either a 3' or 5' protruding single-stranded overhang (Fig. 4A). Consistent with the activity of Rad1-Rad10, ERCC1-XPF was found to remove the 3' protruding single-stranded arm by incising this particular DNA strand in the duplex near the junction (Fig. 4C). Interestingly, incisions near the DNA junction were also detected in the substrate containing the 5' protruding arm (Fig. 4D). In this case,

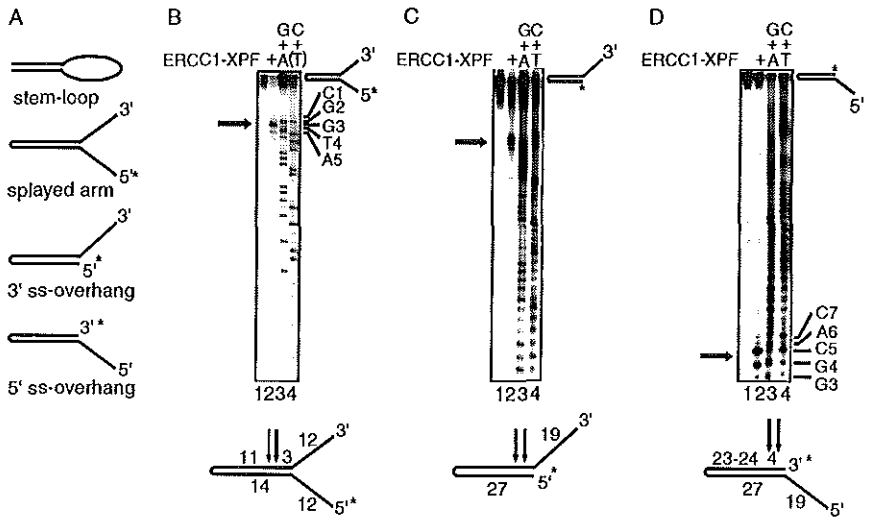


Figure 4. Minimal DNA structural elements required for ERCC1-XPF activity. (A) Schematic presentation of 'dissected' stem-loop substrate. (B) Splayed arm substrate, 5' radiolabeled. (C) 3'-overhang substrate, 5' radiolabeled. (D) 5'-overhang substrate, 3' radiolabeled. For each panel (B-D): *Lane 1*: no protein added. *Lane 2*: 20 ng ERCC1-XPF. *Lane 3-4*: MG-ladders of corresponding substrate; counting starts at junction (note that in (C), the 3'-overhang substrate, separation of MG-ladders is insufficient to exactly determine cleavage positions).

cleavage did not result in removal, but rather extension of the single-stranded portion, as incisions were solely made in the opposite, non-protruding DNA strand. The polarity of cleavage was again consistent with the ERCC1-XPF mediated incisions observed in stem-loops and splayed arms, being in one strand of duplex DNA at the 5' side of a duplex-single stranded junction. Since we never observed incisions at blunt-ended duplex regions (for example in the stem-loop substrates), we concluded that one single-stranded arm, moving away from the junction in either 3' or 5' direction, is necessary and sufficient for ERCC1-XPF to correctly position its nuclease activity on the DNA.

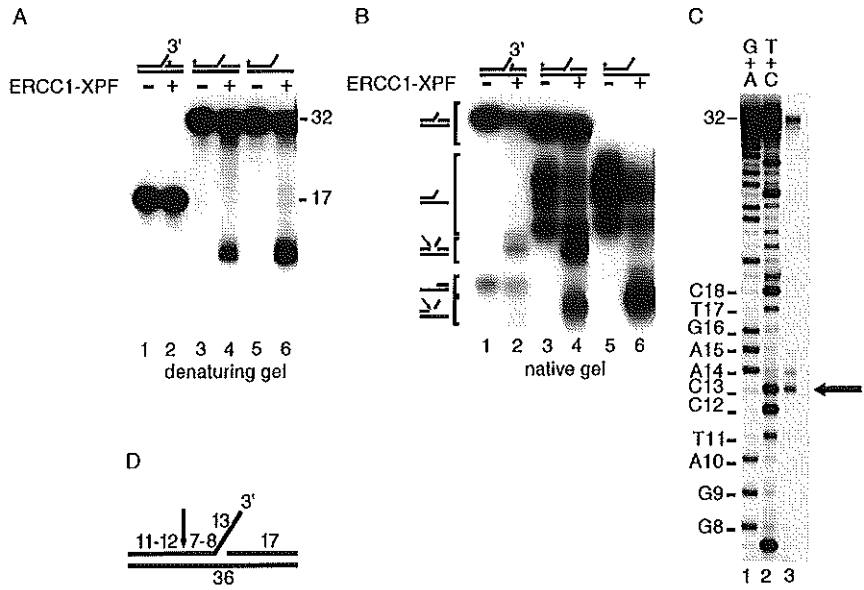


Figure 5. ERCC1-XPF cleaves 3'flap substrates. (A) Denaturing PAGE analysis of 3'flap incisions. (B) Native PAGE analysis of 3'flap incisions. For (A) and (B): *Lane 1*: 3'flap, F17 radiolabeled, no protein. *Lane 2*: 3'flap, F17 labeled, 20 ng ERCC1-XPF. *Lane 3*: 3' flap, F32 labeled, no protein. *Lane 4*: 3'flap, F32 labeled, 20 ng ERCC1-XPF. *Lane 5*: splayed arm, F32 labeled, no protein. *Lane 6*: splayed arm, F32 labeled, 20 ng ERCC1-XPF. Note that especially splayed arm substrate shows a heterogeneous migration pattern on native PAGE (*b. Lane 3* and *5*). (C) Determination of exact cleavage positions. *Lane 1-2*: MG-ladders of F32. *Lane 3*: ERCC1-XPF cleavage pattern of 3'flap, F32 radiolabeled. (D) Schematic presentation of ERCC1-XPF incisions in 3'flap substrates.

ERCC1-XPF cleaves 3' flap structures

To envisage its role in the single strand annealing pathway, ERCC1-XPF has been proposed to remove 3' protruding single-stranded arms from recombined duplex regions (8). To investigate whether ERCC1-XPF is able to cut these so-called flap structures, three (partially) complementary oligos were mixed and annealed (Fig. 5D). Analysis on native gels showed that this resulted in a mixture of all possible annealing products. Therefore both native and denaturing gels were used to analyse incisions (Fig. 5B and C). In this way we found that ERCC1-XPF not only incised the splayed arm substrates present in the

mixture, but also the proposed intermediate in single strand annealing, the flap structure, thereby removing the 3' protruding single-stranded arm from the duplex region. Major incisions in both this splayed arm and flap substrate were made seven and eight nucleotides away from the junction (Fig. 5D). The flap substrate was cleaved less efficiently than the splayed arm substrate (Fig. 5B), which shows that ERCC1-XPF prefers two single-stranded rather than two duplex arms at a DNA junction. No release of a single-stranded arm protruding 5' from a flap substrate was observed, showing that also these substrates are recognized and cleaved with a defined polarity.

Discussion

ERCC1-XPF is a structure-specific endonuclease

Structure-specific DNA nucleases have been implicated in a variety of DNA metabolizing processes, including replication, recombination and repair (20). Recently, we showed that ERCC1-XPF purified from Chinese hamster cells specifically incised DNA at duplex-single stranded borders, hence defining it as a structure-specific endonuclease (31). Others, however, reported that ERCC1-XPF cleaved single-stranded DNA (29) and that replication protein A (RPA) confered structure-specific endonuclease activity to ERCC1-XPF (22). Here we demonstrate that a recombinant ERCC1-XPF complex, highly purified from insect cells, makes incisions specifically near DNA junctions in stem-loops, splayed arms, flaps and single-stranded overhanging DNA substrates. Incisions were found in the absence of other proteins, showing that this activity is intrinsic to ERCC1-XPF. No scissions were observed in the single-stranded DNA regions of our substrates. We would like to propose two possible explanations for the fact that we find intrinsic structure-specific endonucleolytic activity with purified ERCC1-XPF, whereas others do not. Firstly, by using stretches of thymines rather than random sequences we think we have minimized the formation of secondary structure in the single-stranded DNA portions of the substrates. Well-defined single-stranded character was shown to be an important parameter for optimal

activity of structure-specific nucleases (7). Secondly, in order to detect incisions, optimized buffer conditions appeared to be crucial. Hardly any incisions were detectable when reactions were performed in the presence of 10 mM Mg^{2+}, whereas clear activity was observed with low Mg^{2+} (1 - 2 mM) and especially with low ^{2}Mn (0.2 - 1 mM) concentrations. Thus, our data firmly establish ERCC1-XPF, like its yeast homolog Rad1-Rad10 (2), as a protein complex with intrinsic structure-specific endonuclease activity. Concerning replication protein A, we found that the effect of RPA on ERCC1-XPF activity was strongly dependent on the DNA substrate used. Although RPA was never found to be required for confering structure-specific endonuclease activity to ERCC1-XPF, we did observe a stimulatory effect on the nuclease activity, specifically in those cases where ERCC1-XPF cleavage occurred in the strand opposite the one RPA was bound to. Conversely, RPA blocked incision activities that were directed to the same strand it was bound to. Details on this intricate cross-talk between RPA and ERCC1-XPF will be published elsewhere.

Comparison with other structure-specific nucleases

An extensively studied family of structure-specific nucleases is the FEN-1 family. Its members, which include the repair proteins Rad2 in *S. cerevisiae* and mammalian XPG, share three large stretches of homologous amino acids with flap endonuclease-1 (FEN-1) (13). Like ERCC1-XPF, these enzymes cleave branched DNA structures, but they do this at the opposite side of the junction, where the ssDNA is moving away in a 3' to 5' direction (12, 28). Consistent with this cleavage polarity, XPG has been shown to make the 3' incision during NER (28). Although FEN-1 and XPG act similarly on certain DNA structures (flaps, splayed arms), distinct activities on other DNA substrates hint at their different activities in the cell. For example, FEN-1 prefers flap substrates and does not cleave bubble substrates nor other substrates with a closed loop, which probably reflects the requirement of free single-stranded DNA ends for FEN-1 DNA-binding (12, 27). Also, no FEN-1 incisions were observed in substrates containing one single-stranded overhang, showing that FEN-1 only cleaves certain types of single-stranded-duplex DNA junctions.

A functional relevance for FEN-1-mediated incisions was recently provided with the finding that completion of a subpathway of DNA base excision repair (BER) *in vitro* requires FEN-1-dependent removal of splayed arm intermediates; XPG could not substitute for FEN-1 in this reaction (18).

As the engagement of ERCC1-XPF in recombinational events is still poorly understood, it is of particular importance to define the DNA substrate specificity of this protein complex. On the amino acid sequence level, ERCC1 shares a relatively small region of homology with FEN-1 and XPG, encompassing a helix-hairpin-helix motif. This motif, present twice in ERCC1, has been implicated in non-sequence-specific DNA binding (6). Functionally, the ERCC1-XPF complex is most readily compared with XPG. Like XPG, and perhaps expected from its involvement in NER, ERCC1-XPF efficiently cuts bubble substrates (31) and stem-loops, demonstrating that these NER nucleases do not require free single-stranded DNA ends for their activities. We found that ERCC1-XPF incises splayed arms and flaps, thereby cleaving off the 3' protruding single-stranded arm in each case. ERCC1-XPF incisions were also observed in substrates with a 3' or a 5' overhanging arm only, resulting in removal or extension of the single-stranded portion, respectively. In contrast to FEN-1, we found that also XPG incised these substrates, and in agreement with its reported cleavage polarity, this caused extension of the 3' arm and removal of the 5' arm, respectively (data not shown). Apparently, contribution of DNA breathing in these cases is limited and the observed incision activities of ERCC1-XPF and XPG are specific for these overhanging substrates, as we could not detect comparable incisions at the blunt-ended duplex DNA of, for example, stem-loop substrates. Thus, a junction with one single-stranded arm, branching away from duplex DNA either in the 3' or the 5' direction, is necessary and sufficient for correct positioning of incisions by ERCC1-XPF and XPG. We postulate that in *S. cerevisiae*, the same will hold for Rad2 and Rad1-Rad10. The latter complex was previously shown to cut branched DNA substrates (2), but an activity on 5' overhanging substrates has not been reported yet. Intriguingly, Rad10 lacks the helix-hairpin-helix motif that ERCC1 shares with other structure-specific nucleases. As the substrate specificities of ERCC1-XPF and Rad1-Rad10 appear to be similar, either this

motif is not absolutely required for structure-specific nuclease activity, or alternatively, it is cryptically present elsewhere in the Rad1-Rad10 complex.

Incision position partially depends on DNA sequence

Running Maxam-Gilbert sequence ladders along our assays allowed a detailed analysis of the cleavage patterns. ERCC1-XPF-mediated incisions always ocurred in the duplex some nucleotides away from the junction, varying from 2-3 nucleotides in stem-loop substrates, to 7-8 nucleotides in the case of the flap substrate. Given the fact that the DNA sequence around the junction was identical for the different stem-loop substrates, the observation that changing the structural properties by altering loop- and stem-sizes did not affect the cleavage pattern, may indicate that the site of cleavage is to some extent sequence-dependent. In support of this notion, the splayed arm and the flap substrates, composed of the same oligonucleotides, were also cleaved at identical positions (see Fig. 5). A splayed arm substrate with a different sequence composition around the junction was cleaved at 3 nt, rather than 7-8 nt, away from the junction (see Fig. 4). In all these cases, the major incisions were made at the 3' side of a pyrimidine residue, suggesting that ERCC1-XPF prefers to cleave at cytosines and thymines. Two major incisions were observed with the 5' overhanging substrate, one at a pyrimidine and the other at a purine residue. Whether the latter incision shows that the suggested pyrimidine preference is less strict in substrates with one single-stranded arm only or that this preference is based on coincidence, remains to be investigated.

ERCC1-XPF in nucleotide excision repair

During NER, the two incisions around the lesion are made in a near-synchronous manner. Depending on the type of lesion, XPG incisions have been found 2 to 9 nucleotides away from the lesion, whereas ERCC1-XPF cuts appear 15 to 24 nucleotides away on the opposite side of the lesion (21, 24, 25, 34). The positions of these incisions correspond to the borders of the open DNA complex, which is formed during NER and spans approximately 25 nucleotides across the lesion (7). The question why such, relatively, large

region is unwound and excised is still unanswered, but one possibility could be that the size of the open complex reflects the minimal amount of unwinding required for XPG and ERCC1-XPF cutting activity. We found however, that ERCC1-XPF incised substrates with loops as small as 4-8 nucleotides. Also XPG was previously shown to require only 5 to 10 unpaired nucleotides for efficient cleavage (7). Therefore, most likely other repair factors determine the actual size of the excised patch and they may do so by positioning the nucleases onto the borders of the unwound DNA intermediate. For ERCC1-XPF, two obvious candidates are the lesion-recognizing protein XPA and the single-stranded binding protein RPA, which both interact with this nuclease complex (3, 19, 22, 30). The fact that the size of the patch (24-32 nt) strongly correlates to the length of the optimal single-stranded binding region of RPA (30 nt) (17) also suggests an important role of RPA in positioning the NER incisions.

ERCC1-XPF in recombinational pathways

ERCC1-XPF and its homologs have been implicated in such diverse processes as mating type switching in *Schizosaccharomyces pombe* (10) and DNA interstrand crosslink repair in mammals. It has been suggested that common recombinational mechanisms underly these processes. Direct evidence for a role of ERCC1-XPF in recombination comes from studies in *S. cerevisiae*, where it was shown that strains defective in Rad1 or Rad10 failed to complete recombination between direct DNA repeats, presumably due to their inability to remove nonhomologous DNA from the 3' ends of recombined DNA intermediates (8). Here we tested directly whether ERCC1-XPF was able to remove protruding DNA sequences from duplex regions and we found that the complex, like Rad1-Rad10, can indeed cut flap substrates with single-stranded arms protruding in the 3' direction. The relevance of investigating the role of ERCC1-XPF in processes other than NER becomes most apparent from studies on ERCC1-deficient mice. On top of the classic NER-deficient phenotype observed with XPA- as well as XPC-knock-out mice, characterized most profoundly by photosensitivity and predisposition to UV-induced skin cancer, ERCC1-deficient mice suffer from liver, spleen and kidney abnormalities,

developmental delay, reduced life span and signs of premature senescence (23, 37). Accumulation of endogenously generated interstrand crosslinks has been proposed to underly these non-NER-specific clinical manifestations (37). Whether flap structures are intermediates in for example interstrand crosslink repair remains to be shown, but the DNA substrate specificity described in this paper will allow a more direct search for the role of ERCC1-XPF in recombinational events.

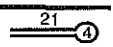
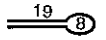
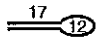
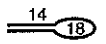
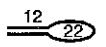
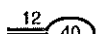
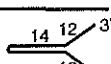
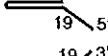
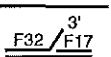
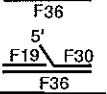
Substrate	Sequence	Structure
stem21-loop4	5'-TCCTGGGTCGCCAGCGCTCGG(T)$_4$ CCGAGCGCTGGCGACCCAGGa-3'	
stem19-loop8	5'-TCGGGTCGCCAGCGCTCGG(T)$_8$ CCGAGCGCTGGCGACCCGa-3'	
stem17-loop12	5'-TGGTCGCCAGCGCTCGG(T)$_{12}$ CCGAGCGCTGGCGACCa-3'	
stem14-loop18	5'-TCGCCAGCGCTCGG(T)$_{18}$ CCGAGCGCTGGCGa-3'	
stem12-loop22	5'-GCCAGCGCTCGG(T)$_{22}$ CCGAGCGCtggc-3'	
stem12-loop40	5'-GCCAGCGCTCGG(T)$_{40}$ CCGAGCGCtggc-3'	
splayed-arm	5'-(T)$_{12}$GCCATCGCGAGTCC-GGACTCGCGATGGC(T)$_{12}$-3'	
5'-overhang	5'-(T)$_{19}$-GGCCGTGCTCTCCTGGATGTTCGAAAG- CTTTCGAACATCCAGGAGAGCACggcc-3'	
3'-overhang	5'-GGCCGTGCTCTCCTGGATGTTCGAAAG- CTTTCGAACATCCAGGAGAGCACGGCC(T)$_{19}$-3'	
F19	5'-GCGATGCGGATCCAAGTCT-3'	
F32	5'-GCGATGCGGATCCAAGTCTATTAGCGACAATG-3'	
F36	5'-CCTAGACTTAAGAGGCCAGACTTGGATCCGCATCGC-3'	
F17	5'-GGCCTCTTAAGTCTAGG-3'	
F30	5'-GTAACAGCGATTAGGCCTCTTAAGTCTAGG-3'	

Table 1. Summary of artificial DNA substrates used. Nucleotides denoted in small letter type were filled-in by DNA polymerase I Klenow fragment, optionally with radioactive labeled dNTPs. Sizes of duplex regions are in base-pairs, sizes of loops and overhangs in nucleotides. For the 3'flap and the 5'flap, the oligonucleotide composition is given. Underlined sequences in F32 and F30 represent unpaired protruding arms.

Materials and methods

Construction of recombinant baculoviruses

The construction of a cDNA encoding ERCC1 with a C-terminal 6xHis-tag was described previously (31). The 3' end (NcoI-BamHI) of this cDNA was used to replace the 3' end (NcoI-BamHI) of wild-type ERCC1 cDNA in the vector pET3C-ERCC1 which contains a unique NdeI site at the translation initiation site. This NdeI site and a ClaI site were used to isolate ERCC1His cDNA, recessive ends were filled in by Klenow fragment, and the cDNA was cloned into the blunt-ended BamHI site of the donor plasmid pFastBacI (BAC-TO-BAC Baculovirus Expression System, GIBCO BRL), resulting in pFastBacI-ERCC1His. To introduce a 6xHis-HA epitope encoding DNA sequence at the 3' end of XPF cDNA, first a (blunt) in-frame NruI site was introduced at the 3' end of the XPF coding region by means of PCR using the antisense primer 5'-CGATCGATTCGGAAGCG-CTGCCTCCCTTTTTCCCTTTTCCTTTT-GAT-3'. Two primers ([1] HIS-HA-sense: 5'-CACCACCATCACCATCACGGAGGCAGCGCTTACCCATAGATGTTCCAGATTA CGCTAGCTGAATCGATG-3', and [2] HIS-HA-antisense: 5'-GATCCATCGATTCAGC-TAGCGTAATCTGGAACATCGTATGGGTAAGCGCTGCCTCCGTGATGGTGATG GTGGTG-3') were annealed and cloned into the NruI site, resulting in the construct pBL-XPFHisHA. This construct was sequenced to exclude the presence of cloning and PCR mistakes. An NcoI (present on translation initiation site) - ClaI digest (both blunt-ended by Klenow fragment) released the XPF-HisHA cDNA, which was cloned into the BamHI site (also made blunt by Klenow fragment) of the donor plasmid pFastBacI (GIBCO BRL), resulting in the construct pFastBacI-XPFHisHA. pFastBACI-ERCC1His and pFastBACI-XPFHisHA plasmids were transfected separately into competent DH10Bac cells to allow site-specific transposition with the Bacmid bMON14272. The recombinant Bacmids (rBac-ERCC1His and rBac-XPFHisHA) were isolated and transfected into Sf21 cells, which resulted in the production of recombinant viruses, designated virBac-ERCC1His and virBacXPFHisHA respectively.

Baculovirus infection and extract preparation from the infected cells

Virus amplification was performed as described (instruction manual BAC-TO-BAC system, GIBCO BRL). For production of the proteins, seven 175 cm^2 dishes with monolayers of Sf21 cells were co-infected with virBac-ERCC1His and virBacXPFHisHA, at a multiplicity of infection of 5 to 10. Two days post-infection, cells were collected by low-speed centrifugation and washed twice with ice-cold phosphate-buffered saline. Cell pellets were fractionated as described before (33). Briefly, pellets were lysed in 8 times packed cell volume (8x PCV) of NP-lysisbuffer (25 mM Tris pH8, 1 mM EDTA, 10% glycerol, 1% NP-40, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 µg/ml aprotinin, 1 µg/mg leupeptin and 1 µg/ml antipain), incubated on ice for 30 min and centrifuged at 800x g for 10 min. The resulting supernatant (S1), containing most of the ERCC1-XPF complex, was used for further purification.

Purification of recombinant ERCC1-XPF complex

Prior to loading, NaCl concentration of S1 protein extracts was adjusted to 0.15 M. Protein extracts were applied to a phosphocellulose column (1 ml beads per 7 mg protein, Whatman P11), equilibrated in buffer A (20 mM HEPES-KOH pH7.8, 2 mM $MgCl_2$, 10% glycerol, 1 mM PMSF, 5 mM ß-mercaptoethanol, 0.02% NP-40) containing 0.15 M KCl. Bound proteins were eluted in buffer A containing 0.6 M KCl. After co-incubation of this fraction overnight at 4^0C with Ni^{2+}-NTA agarose (Qiagen; 1ml/5mg protein) in the presence of 1 mM imidazole, the protein-beads slurry was packed in a column, which was then washed with buffer A containing 20 mM imidazole (0.6 M KCl) and buffer A containing 40 mM imidazole (0.1 M KCl). ERCC1 and XPF were eluted in buffer A containing 200 mM imidazole and 0.1 mM KCl. Insulin (0.1 mg/ml) was added and this fraction was dialysed for 2 x 1 hour against buffer B (20 mM HEPES-KOH pH7.8, 10% glycerol, 2 mM dithiothreitol, 0.02% NP-40, 50 mM KCl, 1 mM EDTA, 2 mM $MgCl_2$). Typically, 50-100 µg (20 ng/µl) purified complex was obtained from 15 mg S1 protein extract.

To analyse complex formation between ERCC1 and XPF, 200 µl of the purified fraction was applied onto a 4.2 ml 15-35% glycerol gradient in buffer B. The gradients were

centrifuged for 20 hours at 300.000 x g and 160 µl fractions were collected.

In vitro and in vivo repair correction assays

To test repair-correcting activity *in vitro*, 60 ng of purified ERCC1-XPF complex was added to Manley-type cell extracts derived from the indicated mutant cell lines; the protein mixture was incubated with an undamaged and a AAF- (*N*-acetoxy-2-acetylaminofluorene-) damaged plasmid in the presence of [^{32}P]-labeled deoxynucleotides. Specific[^{3}H]-incorporation into the damaged plasmid was used as a measure for DNA repair synthesis (for details, see Wood 1995) (38).

In vivo repair activity was determined by micro-injection of purified ERCC1-XPF into the cytoplasm of XPF homopolykaryons (XP126LO fibroblasts) or into the cytoplasm of mouse embryonal fibroblasts (MEFs) derived from ERCC1-deficient mice (37). Repair activity was analysed after 4 hours by UV-induced (15 J/m^2) incorporation of[^{3}H]-thymidine and autoradiography as described (36). The number of grains above the nuclei is a measure of unscheduled DNA synthesis and reflects the cellular repair capacity.

Nuclease assays

DNA oligonucleotides were purified by denaturing polyacrylamide gelelectrophoresis (DPAGE) and 100-150 ng DNA substrate was [^{32}P]-labeled using polynucleotide kinase or Klenow fragment, followed by phenol extraction in 100 µl and G50-sepharose column centrifugation. To allow self-annealing, oligonucleotides were heated for 3 min at 95^{0}C and put on ice. Labeled substrates were kept at 4^{0}C. For the construction of flap substrates, one of the oligonucleotides was labeled, a 2-fold molar excess of the unlabeled complementary oligonucleotides was added and the mixture was heated at 95^{0}C for 5 min, followed by stepwise cooling to allow annealing (30 min at 60^{0}C, 30 min at 37^{0}C, 30 min at 25^{0}C and 30 min on ice).

Nuclease reactions (15 µl) were carried out in optimized nuclease buffer (50 mM Tris [pH8.0], 0.2 mM $MnCl_2$, 0.1 mg/ml bovine serum albumin and 0.5 mM ß-mercaptoethanol), containing 100 fmol of substrate DNA and 100 fmol recombinant

ERCC1-XPF complex. Reactions were incubated for 90 min at 28°C and either stopped by adding 15 µl 90% formamide, heated at 95°C and loaded onto 10-20% denaturing polyacrylamide gels, or 5% glycerol was added and samples were immediately loaded onto 15-20% non-denaturing polacrylamide gels. Reaction products were visualized by autoradiography or phosphorimager.

References

1. **Aboussekhra, A., M. Biggerstaff, M.K.K. Shivji, J.A. Vilpo, V. Moncollin, V.N. Podust, M. Protic, U. Hubscher, J.-M. Egly, and R.D. Wood.** 1995. Mammalian DNA nucleotide excision repair reconstituted with purified components. *Cell* **80:** 859-868.
2. **Bardwell, A.J., L. Bardwell, A.E. Tomkinson, and E.C. Friedberg.** 1994. Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. *Science* **265:** 2082-2085.
3. **Bessho, T., A. Sancar, L.H. Thompson, and M.P. Thelen.** 1997. Reconstitution of human excision nuclease with recombinant XPF-ERCC1complex. *J. Biol. Chem.* **272:** 3833-3837.
4. **Busch, D., C. Greiner, K. Lewis, R. Ford, G. Adair, and L. Thompson.** 1989. Summary of complementation groups of UV-sensitive CHO mutants isolated by large-scale screening. *Mutagen.* **4:** 349-354.
5. **Davies, A.F., 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.
6. **Doherty, A.S., 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.
7. **Evans, E., J. Fellows, A. Coffer, and R.D. Wood.** 1997. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J.* **16:** 625-638.
8. **Fishman-Lobell, J., and J.E. Haber.** 1992. Removal of nonhomologous DNA ends in double-strand break recombination: The role of the yeast ultraviolet repair gene *RAD1*. *Science* **258:** 480-484.
9. **Friedberg, E.C., G.C. Walker, and W. Siede.** 1995. DNA repair and mutagenesis. ASM Press, Washington D.C.
10. **Gutz, H., and H. Schmidt.** 1985. Switching genes in *Schizosaccharomyces pombe*. *Curr. Genet.* **9:** 325-331.
11. **Habraken, Y.S., P; Prakash, L; Prakash, S.** 1995. Structure-specific nuclease activity in yeast nucleotide excision. *J. Biol.Chem.* **270:** 30194-30198.
12. **Harrington, J.J., and M.R. Lieber.** 1994. The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J.* **13:** 1235-1246.
13. **Harrington, J.J., and M.R. Lieber.** 1994. Functional domains within FEN-1 and RAD2 define a family of structure-specific endonucleases: implications for nucleotide excision repair. *Genes Dev.* **8:** 1344-1355.

14. **Huang, J.C., D.L. Svoboda, J.T. Reardon, and A. Sancar.** 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.* **89:** 3664-3668.
15. **Ishimi, Y.C., A; Bullock, P; Hurwitz, J.** 1988. Complete enzymatic synthesis of DNA containing the SV40 origin of replication. *J. Biol. Chem.* **263:** 19723-19733.
16. **Ivanov, E.L., and J.E. Haber.** 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.
17. **Kim, C., R.O. Snyder, and M.S. Wold.** 1992. Binding properties of replication protein-A from human and yeast cells. *Mol. Cell. Biol.* **12:** 3050-3059.
18. **Klungland, A., and T. Lindahl.** 1997. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.* **16:** 3341-3348.
19. **Li, L., S.J. Elledge, C.A. Peterson, E.S. Bales, and R.J. Legerski.** 1994. Specific association between the human DNA repair proteins XPA and ERCC1. *Proc. Natl. Acad. Sci.* **91:** 5012-5016.
20. **Lieber, M.** 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* **19:** 233-240.
21. **Matsunaga, T., D. Mu, C.-H. Park, J.T. Reardon, and A. Sancar.** 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.
22. **Matsunaga, T., C.H. Park, T. Bessho, D. Mu, and A. Sancar.** 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.
23. **McWhir, J., J. Seldridge, D.J. Harrison, S. Squires, and D.W. Melton.** 1993. Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat. Gen.* **5:** 217-223.
24. **Moggs, J.G., K.J. Yarema, J.M. Essigmann, and R.D. Wood.** 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.
25. **Mu, D., D.S. Hsu, and A. Sancar.** 1996. Reaction mechanism of human DNA repair excision nuclease. *J. Biol. Chem.* **271:** 8285-8294.
26. **Mu, D., C.-H. Park, T. Matsunaga, D.S. Hsu, J.T. Reardon, and A. Sancar.** 1995. Reconstitution of human DNA repair excision nuclease in a highly defined system. *J. Biol. Chem.* **270:** 2415-2418.
27. **Murante, R.S., L. Rust, and R.A. Bambara.** 1995. Calf 5' to 3' exo/endonuclease must slide from a 5' end of the substrate to perform structure-specific cleavage. *J. Biol. Chem.* **270:** 30377-30383.
28. **O'Donovan, A., A.A. Davies, J.G. Moggs, S.C. West, and R.D. Wood.** 1994. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371:** 432-435.
29. **Park, C.-H., T. Bessho, T. Matsunaga, and A. Sancar.** 1995. Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *J. Biol. Chem.* **270:** 22657-22660.
30. **Park, C.-H., and A. Sancar.** 1994. Formation of a ternary complex by human XPA, ERCC1 and ERCC4(XPF) excision repair proteins. *Proc. Natl. Acad. Sci.* **91:** 5017-5021.
31. **Sijbers, A.M., W.L. de Laat, R.R. Ariza, M. Biggerstaff, Y.-F. Wei, J.G. Moggs, K.C. Carter, B.K. Shell, E. Evans, M.C. de Jong, S. Rademakers, J. de Rooij, N.G.J. Jaspers, J.H.J. Hoeijmakers, and R.D. Wood.** 1996. Xeroderma pigmentosum group F caused by a defect in a

structure-specific DNA repair endonuclease. *Cell* **86:** 811-822.

32. **Sijbers, A.M., P.J. van der Spek, H. Odijk, J. van den Berg, M. van Duin, A. Westerveld, N.G.J. Jaspers, D. Bootsma, and J.H.J. Hoeijmakers.** 1996. Mutational analysis of the human nucleotide excision repair gene *ERCC1*. *Nucleic Acids Res.* **24:** 3370-3380.
33. **Sugasawa, K., C. Masutani, A. Uchida, T. Maekawa, P.J. van der Spek, D. Bootsma, J.H.J. Hoeijmakers, and F. Hanaoka.** 1996. HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro. *Mol. Cell. Biol.* **16:** 4852-4861.
34. **Svoboda, D.L., J.S. Taylor, J.E. Hearst, and A. Sancar.** 1993. DNA repair by eukaryotic nucleotide excision nuclease. *J. Biol. Chem.* **268:** 1931-1936.
35. **van Duin, M., J. van den Tol, P. Warmerdam, H. Odijk, D. Meijer, A. Westerveld, D. Bootsma, and J.H.J. Hoeijmakers.** 1988. Evolution and mutagenesis of the mammalian excision repair gene ERCC-1. *Nucleic Acids Res.* **16:** 5305-5322.
36. **Vermeulen, W., R.J. Scott, S. Potger, H.J. Muller, J. Cole, C.F. Arlett, W.J. Kleijer, D. Bootsma, J.H.J. Hoeijmakers, and G. Weeda.** 1994. Clinical heterogeneity within xeroderma pigmentosum associated with mutations in the DNA repair and transcription gene ERCC3. *Am. J. Human Gen.* **54:** 191-200.
37. **Weeda, G.D., I; de Wit, J; Morreau, H; Janssens, R; Vissers, CJ; Nigg, A; van Steeg, H; Bootsma, D; Hoeijmakers, JHJ.** 1997. Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr. Biol.* **7:** 427-39.
38. **Wood, R.D., M. Biggerstaff, and M.K.K. Shivji.** 1995. Detection and measurement of nucleotide excision repair synthesis by mammalian cell extracts *in vitro*. *Meth. Enzymol.* **7:** 163-175.

Chapter 6

DNA-binding polarity of human replication protein A (RPA) positions nucleases in nucleotide excision repair

Abstract

The human single-stranded DNA-binding protein RPA is involved in various DNA-processing events. By comparing the affinity of hRPA for artificial DNA hairpin structures with 3' or 5' protruding ss-arms, we found that hRPA binds ssDNA with a defined polarity; a strong ssDNA-interaction domain of hRPA is positioned at the 5' side of its binding region, a weak ssDNA-binding domain resides at the 3' side. Polarity appears crucial for positioning of the excision repair nucleases XPG and ERCC1-XPF on the DNA; with the 3'-oriented side of hRPA facing a duplex single-stranded DNA junction, hRPA interacts with and stimulates ERCC1-XPF, whereas the 5'-oriented side of hRPA at a DNA junction allows stable binding of XPG to hRPA. Our data pinpoint hRPA to the undamaged strand during nucleotide excision repair. Polarity of hRPA on ssDNA is likely to contribute to the directionality of other hRPA-dependent processes as well.

Introduction

Human replication protein A (hRPA) is a single-stranded DNA-binding protein composed of three subunits of approximately 70, 32 and 14 kDa (15, 47). It is involved in multiple DNA-metabolizing processes, including (SV40-) DNA replication, recombination and repair (reviewed by (45)). At a certain stage during all these processes, complementary DNA strands are separated and action is required along single-stranded (ss) DNA intermediates. This action always occurs with a defined directionality; DNA polymerases synthesize new DNA strands in a 5' to 3' direction, Rad51-catalysed recombinational strand exchange proceeds in a 3' to 5' direction (1), and repair incisions are made by nucleases that specifically cleave either at the 3' or the 5' side of a lesion (28, 33, 40). Two properties of hRPA have generally been acknowledged to play a role in these processes: firstly, hRPA's high affinity for ssDNA, and secondly, its ability to specifically interact with a variety of proteins.

hRPA binds to ssDNA with an apparent association constant of 10^9-10^{11} M^{-1}, which is at least three orders of magnitude higher than its affinity for double-stranded (ds) DNA (23, 24). Binding of a monomer of hRPA to ssDNA occurs through the 70 kDa subunit, and two different binding modes have been identified; hRPA interacts with a minimal occluded binding site of 8-10 nucleotides (nt) (4) and with a 100-fold more stable occluded binding site of approximately 30 nt (5, 23, 24). In general, the high affinity of hRPA for ssDNA is thought to contribute to the stabilization of ssDNA intermediates and the removal of secondary structure from ss-regions. More specifically, binding of hRPA to ssDNA is required for unwinding of (SV40) replication origins and elongation of replication forks (22, 43, 44, 46). In recombination, hRPA stimulates DNA strand exchange between Rad51-coated single-stranded DNA and duplex DNA (1, 32, 38, 41). In nucleotide excision repair (NER), binding of hRPA to ssDNA might help the formation and stabilization of a locally unwound DNA intermediate. Interestingly, the extent to which DNA is opened around a lesion corresponds to the 30 nt optimal binding region of hRPA (14, 21). Most of the activities mentioned above are non-specific, in the sense that other single-stranded binding

proteins (SSBs) can carry out these functions as well.

The specificity of hRPA's contribution to the different DNA-metabolizing events comes from its ability to interact with other proteins. hRPA specifically interacts with DNA polymerase α, probably to recruit this DNA polymerase to the unwound replication origin to initiate DNA replication (7, 13, 22). Also, hRPA binds to hRad52, an interaction that seems to be essential for homologous recombination (35). Furthermore, hRPA-interactions have been described with the NER-proteins XPA, XPG and ERCC1-XPF (19, 29). A complex between hRPA and the damage-recognition protein XPA was reported to show cooperative binding to DNA lesions (19).

XPG and ERCC1-XPF are structure-specific endonucleases that cut the damaged strand respectively at the 3'- and 5'-side of the locally unwound NER intermediate (14, 28, 33, 40). It is unknown what directs the nucleases to the damaged strand and prevents incisions in the non-damaged strand during NER. hRPA is indispensable for NER incisions (11), but its physical localization with regard to the DNA and the remainder of the repair machinery is not clear. Interactions of hRPA with XPG and ERCC1-XPF were found to modulate the activity of both NER nucleases. Controversy exists, however, as to whether hRPA has a positive or negative effect on cleavage (3, 14, 29). Here, we tested the hypothesis that the orientation of hRPA with regard to the cleavage site determines this outcome.

Results

hRPA binds to single-stranded DNA with a defined polarity

Several indications exist that the 8-10 nt binding mode of hRPA is a precursor of the more stable 30 nt binding mode (4, 5, 17). To investigate whether hRPA binds ssDNA with a defined polarity, we considered the possibility that initial DNA recognition occurs at a fixed position at one of the borders of the 30 nt binding region. This would imply that, in the case of short (8-30 nt) single-stranded arms protruding from duplex regions, hRPA

shows a binding-preference for arms protruding in either the 3' or the 5' direction.

To study this, we used partially self-complementary oligonucleotides containing identical duplex regions with either 3'- or 5'-protruding single-stranded arms. The rationale to use hairpin DNA substrates, rather than combinations of two annealed oligonucleotides, is that their strong tendency to self-anneal counteracts the ability of hRPA to unwind duplex DNA regions (16, 42), which otherwise would complicate the analysis. As hRPA preferentially binds to ss-polypyrimidine tracts (24), the single-stranded overhangs of these substrates solely consisted of thymine residues, which also minimized the formation of secondary structures. Using gel retardation assays, a dramatic difference in hRPA-affinity for 3'- and 5'-protruding 19 nt ss-arms (19-d(T) substrates) was observed. hRPA (Fig. 1H) bound very efficiently to the 3'-protruding 19-d(T) substrate, with almost 80% of the substrate complexed to hRPA monomers at approximately equimolar hRPA-DNA concentrations (Fig. 1A). In contrast, virtually no monomeric protein-DNA complexes were observed with the 5'-protruding substrate, not even when a >3-fold molar excess of hRPA was present (Fig. 1B). With these hRPA-concentrations, only some hRPA trimers complexed to (possibly partially unwound) DNA hairpins were detected (Fig. 1B).

The effect of the size of the ss-stretch on the hRPA-binding preference was examined. Using substrates with 10 nt overhangs, essentially similar binding characteristics were observed as with the 19 nt ss-overhangs. Very few protein-DNA complexes were formed with 5'-protruding 10-d(T) substrates, whereas hRPA-DNA complexes were clearly detectable with the 3'-protruding 10-d(T) substrate (Fig. 1C and 1D). However, the affinity of hRPA for the 3'-protruding 10-d(T) substrate was reduced compared to its 3'-protruding 19 nt counterpart, as at approximately equimolar concentrations 10%, instead of 80% of the substrate was complexed to an hRPA-monomer (Fig. 1G).

Similar results were obtained with ss-overhangs composed of a random sequence instead of a poly-dT tract, demonstrating that our results were not a peculiarity of specific oligonucleotides. Virtually no hRPA-binding was observed to a 5'-protruding 13-d(N) substrate, whereas at approximately equimolar hRPA-DNA concentrations 30-40% of a 3'-protruding 14-d(N) substrate was bound to hRPA-monomers (data not shown). Changing

the divalent ion-concentration from 0.75 mM Mn^{2+} to 5 mM Mg^{2+} or adding 5 mM EDTA did not alter the affinity of hRPA for any of the substrates (data not shown).

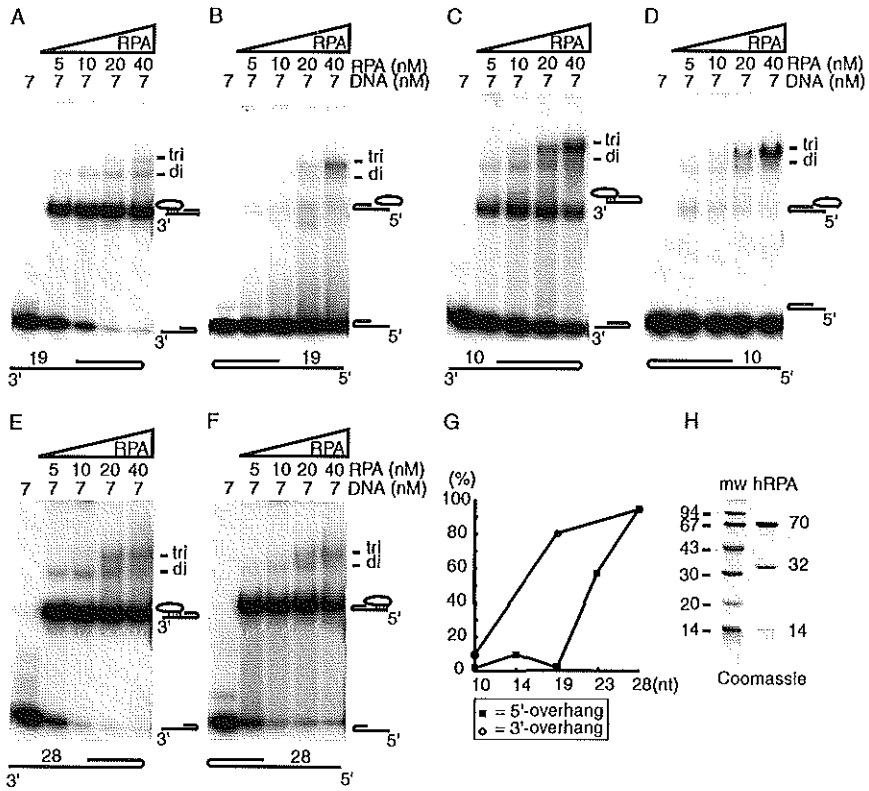


Figure 1. RPA preferentially binds to 3'-protruding single-stranded arms. (A-F) Gel retardation assays showing the binding characteristics of hRPA to (A) 3'-protruding 19-d(T) substrate, (B) 5'-protruding 19-d(T) substrate, (C) 3'-protruding 10-d(T) substrate, (D) 5'-protruding 10-d(T) substrate, (E) 3'-protruding 28-d(T) substrate, (F) 5'-protruding 28-d(T) substrate. 'Di' and 'tri' indicate dimeric and trimeric hRPA-complexes bound to DNA, respectively. No glutaraldehyde was added to any of the reactions.
(G) Graphic presentation showing DNA-binding efficiency of hRPA (in percentage of bound substrate at 10 nM hRPA, which is the hRPA-concentration that is referred to in the text as the approximately equimolar concentration) versus the size of single-stranded overhang. The DNA-binding efficiencies presented here are not meant to give a quantative measure of hRPA's affinity for the different substrates.
(H) Coomassie-stained protein gel of purified hRPA.

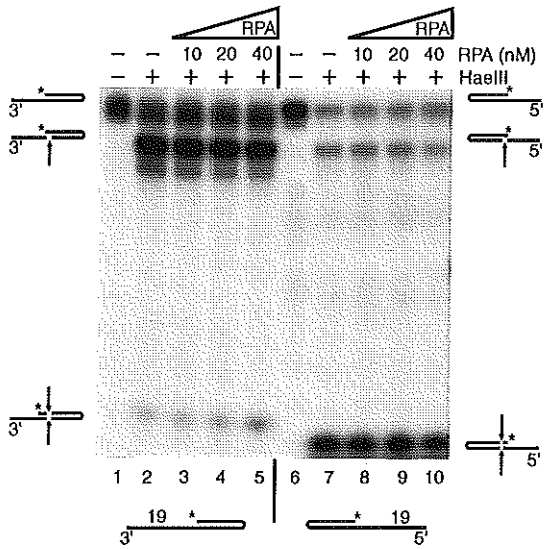


Figure 2. hRPA does not significantly unwind hairpin structures. Cleavage pattern of HaeIII in 3'- and 5'-protruding 19 nt substrates is not altered when substrates are preincubated with hRPA (compare lanes 3-5 with lane 2, and lanes 8-10 with lane 7, respectively). Note that, independent of the presence of hRPA, HaeIII is hardly able to cleave the non-protruding strand of the 3'-protruding substrate (lanes 2-5). The dissimilar migration pattern of the released dinucleotides in lanes 2-5 versus lanes 7-10 we attribute to different nucleotide compositions (CC versus GG).

To find out whether unwinding of duplex DNA underlies the binding-preference of hRPA for 3'-protruding ss-regions, we performed a footprinting-type of analysis with restriction enzymes. Cleavage of 3'- and 5'-protruding substrates by HaeIII, which recognizes the four basepairs immediately bordering the ssDNA portion of both substrates, was not affected by the presence of hRPA in the concentration-range used above (Fig. 2). Similar results were obtained with BstB1, a restriction enzyme that recognizes six basepairs located near the turn of the hairpins (data not shown). These data demonstrated that the duplex region of these hairpins is not significantly unwound by hRPA monomers. Thus, the observation that hRPA monomers only bind stably to 3'-protruding ss-arms and not to 5'-protruding ss-arms, at least when these arms range in size from 10 to 19 nucleotides, can

only be explained by defined ssDNA-binding polarity of hRPA. Apparently, with arms protruding in the 5' direction, the neighbouring duplex DNA physically hinders the oriented interaction of hRPA with ssDNA.

Next, we increased the size of the ss-arms to that of the optimal hRPA-binding site. Equally efficient interactions were observed with both 3'- and 5'-protruding 28-d(T) substrates. More than 90% of both substrates were complexed to hRPA-monomers at approximately equimolar hRPA-DNA concentrations (Fig. 1E-G). A partial loss of binding-preference was already observed with a 23 nt long 5'-protruding substrate, which was bound with moderate affinity by hRPA (Fig. 1G). We conclude that hRPA preferentially binds 3'-protruding single-stranded arms, a phenomenon that is abolished when ss-overhangs reach the size of the optimal hRPA-binding site. The preference in affinity for ss-arms protruding in a certain direction implies that hRPA binds ssDNA with a defined polarity. We designate the strong ssDNA-binding side of hRPA as the '*5'-oriented side*' of hRPA, since it is located at the 5' side of the bound ssDNA-region (see also Fig. 8). Conversely, the weak ss-interaction subdomain is defined as the '*3'-oriented side*' of hRPA. Obviously, the orientation of hRPA with respect to the ssDNA has important implications for the proteins with which it interacts.

Defined ssDNA-binding polarity of hRPA modulates the activity of the excision repair nuclease ERCC1-XPF

To investigate possible functional consequences of the defined ssDNA-binding polarity of hRPA, we focussed on its role in nucleotide excision repair. Human RPA has been claimed to modulate the activity of XPG and the heterodimeric complex ERCC1-XPF (3, 14, 29). Both are structure-specific NER endonucleases that direct their incisions to one strand of duplex DNA at ds-ss DNA junctions, with ERCC1-XPF cutting the strand that continues as a 3'-protruding ss-region and XPG cleaving the opposite strand (3, 14, 29) (see Fig. 3B, 4A and 5B). Previously, we have shown that these enzymes require only one single-stranded arm protruding from duplex DNA to correctly position their incisions (12). As such, the 3'- and 5'-protruding constructs used above are suitable substrates for cleavage

by ERCC1-XPF and XPG. To determine whether the orientation of hRPA and the strand to which it binds influence the activity of the NER nucleases we performed incision assays in the presence of hRPA. The data presented below were obtained with a homogeneous preparation of recombinant ERCC1-XPF purified from baculovirus infected insect cells

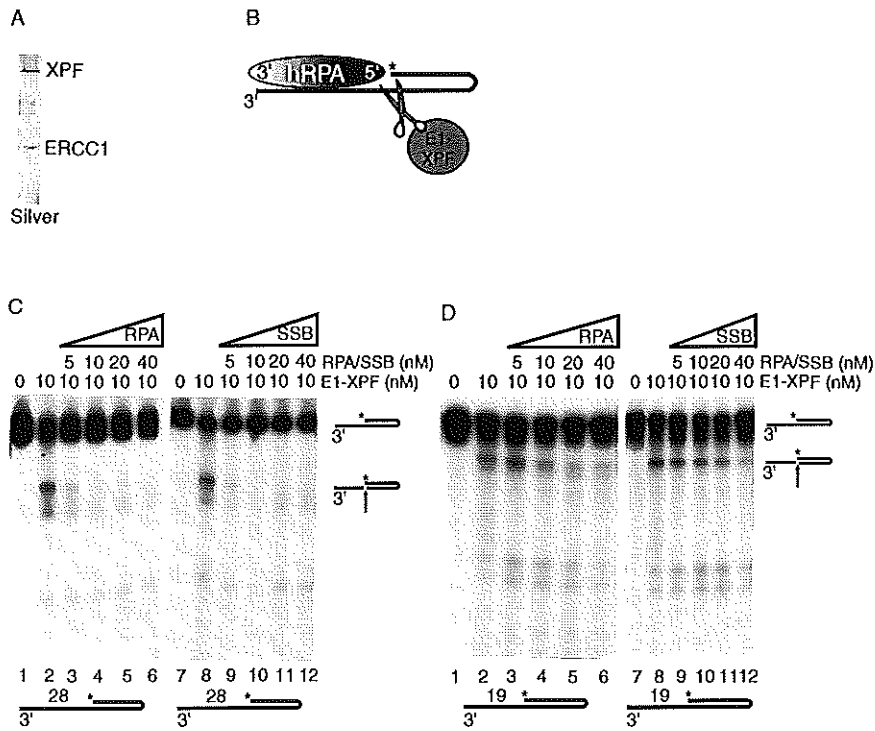


Figure 3. hRPA inhibits ERCC1-XPF endonuclease activity on 3' -protruding ss-arms. (A) Silver-stained protein gel of purified ERCC1-XPF. (B) Schematic presentation of ERCC1-XPF and hRPA acting on a 3'-protruding substrate. Asterisk indicates position of radioactive label. The 3'- and 5'-oriented side of hRPA, representing the weak and the strong ssDNA-binding side of hRPA, respectively, are indicated.
(C-D) Denaturing polyacrylamide gels analyzing DNA incision products. (C) ERCC1-XPF nuclease assays on 3'-protruding 28-d(T) substrates, with increasing amounts of hRPA (lanes 3-6) and *E. coli* SSB (lanes 9-12). (D) ERCC1-XPF nuclease assays on 3'-protruding 19-d(T) substrates, with increasing amounts of hRPA (lanes 3-6) and *E. coli* SSB (lanes 9-12). Note that *E. coli* SSB-concentration is given in nmoles of tetramer per liter. The minor incision products visible in figure c and d do not correspond to known duplex-single stranded DNA junctions and probably arise from weak cutting activity near uncharacterized secondary structures in the DNA hairpin substrates.

(Fig. 3A). The results of key experiments were verified with recombinant ERCC1-XPF purified from *E. coli* to confirm that the observed nuclease activities were due to ERCC1-XPF and not a contaminating nuclease, and to exclude the possibility that other repair factors copurifying with ERCC1-XPF from insect cells obscured the outcome (see also (12)).

ERCC1-XPF cleaves 3'-protruding substrates in the strand that continues as a single-stranded arm, thereby removing the ss-portion (Fig. 3B). Pre-incubation of the 3'-protruding 28-d(T) substrate with increasing amounts of hRPA caused almost complete inhibition of ERCC1-XPF cleavage, a phenomenon also observed with the shorter 3'-protruding 19-d(T) substrates (Fig. 3C and 3D, lanes 1-6). To test whether this inhibitory effect was specific for hRPA, similar assays were performed with *E. coli* single-stranded binding protein (SSB). *E. coli* SSB blocked ERCC1-XPF activity approximately to the same extent as hRPA did (Fig. 3C and 3D, lanes 7-12), suggesting that hRPA-mediated inhibition of ERCC1-XPF cleavage on 3'-protruding DNA substrates does not require specific protein-protein interactions and is most likely due to steric hindrance.

Similarly, we tested 5'-protruding 28-d(T) DNA substrates. In the absence of hRPA, ERCC1-XPF cleaves these substrates near the junction in the non-protruding strand, which results in extension of the single-stranded portion (Fig. 4A). In contrast to 3' protrusions, preincubation of hRPA with the 5'-protruding 28-d(T) substrate had a strong stimulatory effect on ERCC1-XPF activity (Fig. 4B, lanes 1-6). No stimulation of ERCC1-XPF activity was observed with *E. coli* SSB (Fig. 4B, lanes 7-12), demonstrating that this effect was specific for hRPA and suggesting a direct protein-protein interaction between hRPA and ERCC1-XPF on 5'-protruding substrates. Suprisingly, even on 5'-protruding 19-d(T) substrates, for which bandshift experiments failed to reveal hRPA-binding (see Fig. 1B), ERCC1-XPF cleavage was strongly stimulated by preincubation with hRPA (Fig. 4C). Apparently, hRPA does transiently interact with short 5'-protruding ss-arms, but this interaction is not stable enough to withstand bandshift conditions. ERCC1-XPF-mediated cleavage of 5'-protruding substrates yields a DNA product that can serve again as a substrate for cleavage (Fig. 4A), and indeed with 5'-protruding hairpin substrates labeled

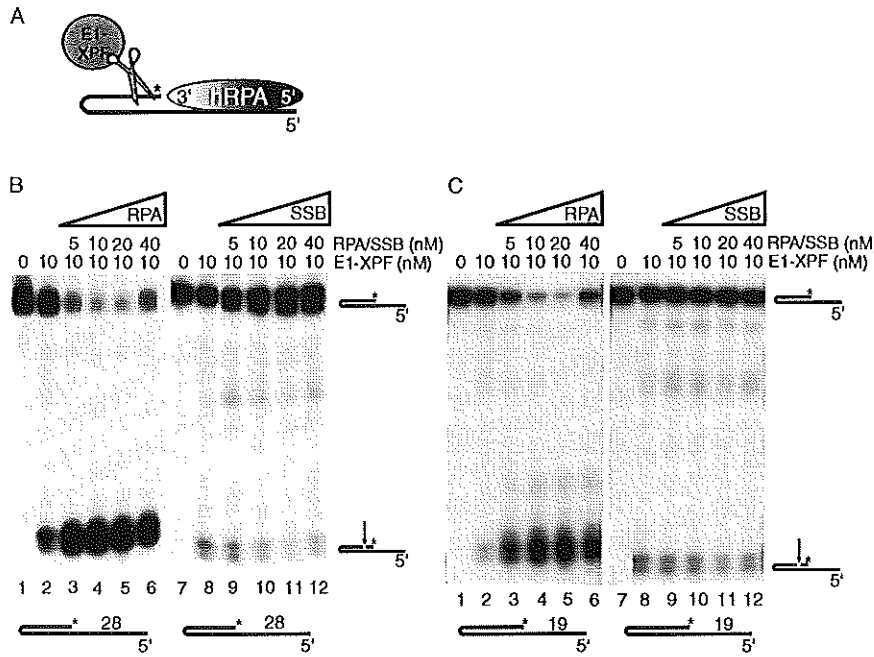


Figure 4. hRPA stimulates ERCC1-XPF endonuclease activity on 5' -protruding ss-arms. (A) Schematic presentation of ERCC1-XPF and hRPA acting on a 5'-protruding substrate. Asterisk indicates position of radioactive label. The 3'- and 5'-oriented side of hRPA, representing the weak and the strong ssDNA-binding side of hRPA, respectively, are indicated.

(B) ERCC1-XPF nuclease assays on 5'-protruding 28-d(T) substrates, with increasing amounts of hRPA (lanes 3-6) and *E. coli* SSB (lanes 9-12).

(C) ERCC1-XPF nuclease assays on 5'-protruding 19-d(T) substrates, with increasing amounts of hRPA (lanes 3-6) and *E. coli* SSB (lanes 9-12). *E. coli* SSB-concentration is given in nmoles of tetramer per liter. The minor incision products visible in figure B and C do not correspond to known duplex-single stranded DNA junctions and probably arise from weak cutting activity near uncharacterized secondary structures in the DNA hairpin substrates.

at the 5' end we observed that in the presence of hRPA most of the incision products are incised again by ERCC1-XPF some nucleotides further away in the duplex (data not shown). The original positions of ERCC1-XPF incisions in both 3'- and 5'-protruding substrates were not altered by hRPA though (Fig. 3C and D, and 4B and C), indicating

again that hRPA-binding does not significantly modify the hairpin structure. The specific stimulation of ERCC1-XPF activity on 5'-protruding ss-arms, not observed with 3'-protruding hairpin substrates, suggests orientation-specific interactions between hRPA and ERCC1-XPF.

hRPA binding orientation has minor implications on XPG endonuclease activity

Similar incision assays were performed to determine the effect of differently oriented hRPA molecules on the activity of XPG, the complementary NER endonuclease. For this purpose, a homogeneous preparation of XPG protein was obtained from baculovirus infected insect cells (Fig. 5A), using a standard purification protocol (with one

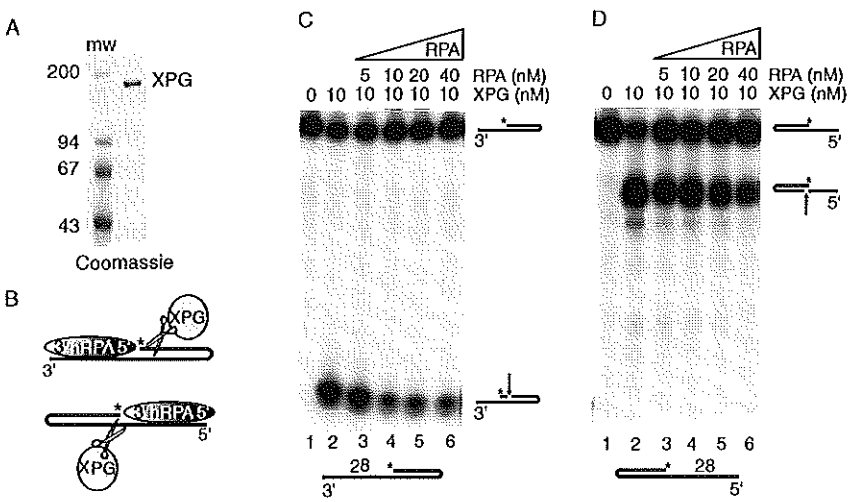


Figure 5. Polarity of hRPA on ssDNA has minor implications for XPG endonuclease activity. (A) Coomassie-stained protein gel of purified XPG. (B) Schematic presentation of XPG and hRPA acting on a 3'-protruding substrate (upper drawing) and a 5'-protruding substrate (lower drawing). Asterisks indicate positions of radioactive label. The 3'-oriented and 5'-oriented side of hRPA, representing the weak and the strong ssDNA-binding side of hRPA, respectively, are indicated.
(C-D) XPG nuclease assays on (c) 3'-protruding 28-d(T) substrates, -/+ hRPA; (d) 5'-protruding 28-d(T) substrates, -/+ hRPA.

additional purification step) that removes other repair factors and contaminating nucleases (34). XPG-mediated cleavage in 3'-overhanging hairpin substrates occurs in the non-protruding strand, whereas hairpins with a 5' ss-arm are cleaved in the protruding strand (Fig. 5B). Both on 3'- and 5'-protruding 28-d(T) substrates we found an inhibitory effect of hRPA on XPG-incision activity (Fig. 5C and D). Inhibition, however, was not as pronounced as observed with ERCC1-XPF on 3'-protrusions, and was incomplete, even with an excess of hRPA. The partial inhibitory effect was also found with *E. coli* SSB replacing hRPA (data not shown). Again, hRPA did not change the position of XPG-mediated incisions in either substrate, once more confirming the footprinting analysis that demonstrated the structural integrity of these hairpins in the presence of hRPA. We conclude that hRPA, irrespective of its orientation with regard to the ds-ss junction, shows a weak and non-specific modulation of XPG activity on the substrates tested.

hRPA-ERCC1-XPF complex formation occurs specifically on 5'-protruding substrates

Obviously, the stimulatory effect of hRPA on ERCC1-XPF activity could imply that hRPA facilitates the positioning of this nuclease onto DNA, in line with the reported protein-protein interactions between them (3, 29). Selective stimulation of ERCC1-XPF activity on 5'-protruding DNA substrates suggests that such a physical interaction strongly depends on the orientation of hRPA with regard to the ds-ss junction. To further investigate this, ERCC1-XPF binding to pre-formed DNA-hRPA complexes was monitored in gel retardation assays. In the absence of hRPA, we could not detect DNA-ERCC1-XPF complexes with any of the substrates, neither in standard nuclease buffers nor in the presence of divalent ions such as Ca^{2+} (Fig. 6A, lane 1). After co-incubation of ERCC1-XPF with hRPA and the 5'-protruding 28-d(T) substrate in standard nuclease buffers, a supershifted band was vaguely detectable, possibly representing a complex of DNA, hRPA and ERCC1-XPF (data not shown). However, the strong stimulation by hRPA of ERCC1-XPF-mediated cleavage is expected to hamper the trapping of transient complexes. To circumvent this problem, a similar co-incubation was performed in a standard buffer containing 5 mM Ca^{2+} instead of Mn^{2+} or Mg^{2+}. In the presence of calcium, which was

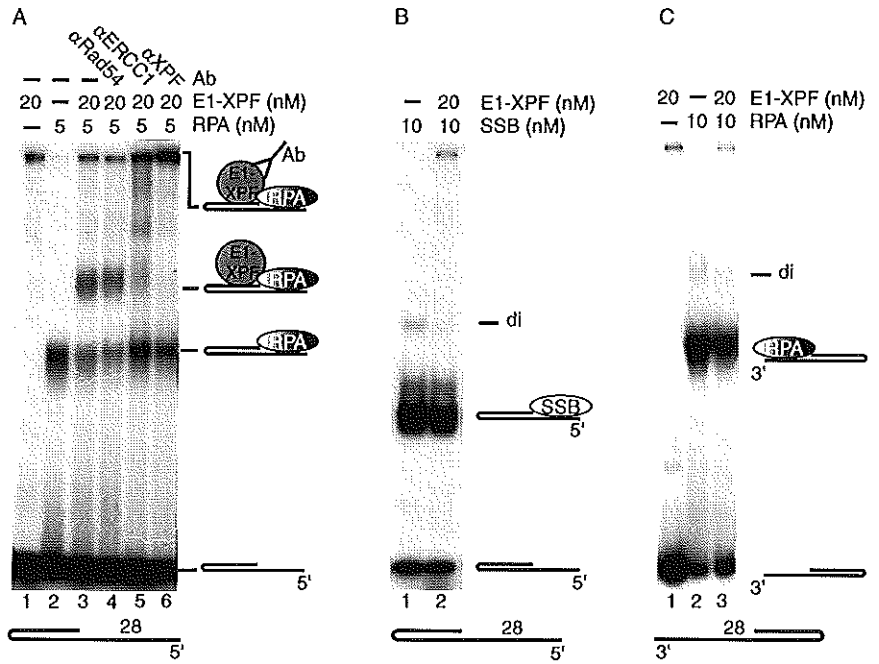


Figure 6. hRPA-ERCC1-XPF complex formation occurs specifically on 5'-protruding substrates. (A-C) Gel retardation assays; all incubations were performed in 5 mM $CaCl_2$, in the presence of 0.05% glutaraldehyde. (A) hRPA-ERCC1-XPF complex formation on 5'-protruding 28-d(T) substrates. For clarity, only 5 nM hRPA was used; higher concentrations of hRPA induced some hRPA-dimer formation on these substrates, which migrated similar to the ternary hRPA-ERCC1-XPF complexes. However, ternary complex formation was also observed at 10 nM hRPA (data not shown). (B) No complex formation between *E. coli* SSB and ERCC1-XPF on 5'-protruding 28-d(T) substrates. (C) No hRPA-ERCC1-XPF complex formation on 3'-protruding 28-d(T) substrates.

previously shown to block ERCC1-XPF incision activity (12), a clear supershifted complex was detectable in addition to the DNA-hRPA complex, which was dependent on the presence of DNA, hRPA and ERCC1-XPF and originated from the DNA-hRPA complex (Fig. 6A, lane 3, compare with lane 1 and 2). Antibodies against ERCC1 and XPF further shifted this particular complex to the slot of the gel, whereas an unrelated antibody did not (Fig. 6A, lane 4-6). Thus, the supershifted band indeed represented a ternary complex,

consisting of ERCC1-XPF and hRPA bound to the 5'-protruding DNA substrate. The use of EDTA instead of Ca^{2+} completely disrupted ternary complex formation but did not alter the binding of hRPA to DNA (data not shown), demonstrating that not only a direct protein-interaction with hRPA, but also divalent cation-dependent DNA-binding is a prerequisite for ERCC1-XPF to interact with hRPA-DNA complexes. Calcium, apparently, allows ERCC1-XPF to bind DNA but prevents further endonucleolytic processing. Ternary complex formation required specific interactions between ERCC1-XPF and hRPA, since *E. coli* SSB was not able to produce such a supershift (Fig. 6B). Thus, it appears that hRPA stimulates ERCC1-XPF activity by facilitating, via direct protein-protein interactions, the positioning of this nuclease onto the DNA.

Under the optimized conditions described above, ERCC1-XPF could not form a ternary complex on a 3'-protruding 28-d(T) substrate pre-incubated with hRPA (Fig. 6C), not even when twice the amount of enzyme was added (data not shown). This demonstrated that physical interaction between ERCC1-XPF, hRPA and DNA is dependent on the orientation of hRPA with regard to the ds-ss DNA junction. As on 5'-protruding substrates the 3'-oriented side of hRPA is positioned towards the junction, we conclude that hRPA, bound to ssDNA, can only position and stimulate ERCC1-XPF-mediated incisions in duplex DNA that borders directly at its 3'-oriented side (see also Fig. 8).

hRPA-XPG complex formation occurs specifically on 3'-protruding substrates

Although hRPA did not dramatically affect XPG-activity under the conditions tested here, the possibility of a specific engagement when acting in close vicinity on DNA had to be considered. In the presence of Ca^{2+}, no ternary complex was detected when XPG was co-incubated with hRPA bound to either DNA substrate (data not shown). When binding was allowed in standard nuclease buffer containing 0.75 mM Mn^{2+} however, clearly a new, slowly migrating complex was observed. Its formation required the presence of hRPA, XPG and the 3'-protruding 28-d(T) substrate (Fig. 7A, lane 3, compare with lanes 1 and 2). Addition of an antibody against XPG specifically shifted this complex to the slot of the gel, whereas an unrelated antibody did not, demonstrating the presence of XPG in this complex

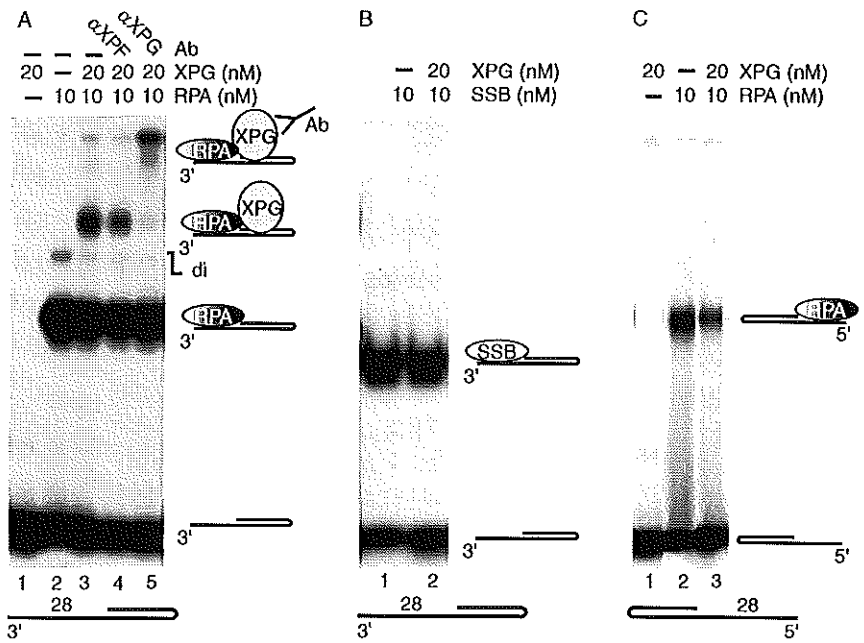


Figure 7. hRPA-XPG complex formation occurs specifically on 3'-protruding substrates. (A-C) Gel retardation assays; all incubations were performed in 0.75 mM $MnCl_2$, in the presence of 0.05% glutaraldehyde. (A) hRPA-XPG complex formation on 3'-protruding 28-d(T) substrates. Note that, at 10 mM hRPA some hRPA dimer formation, 'di', occurs (lane 2). (B) No complex formation between *E. coli* SSB and XPG on 3'-protruding 28-d(T) substrates. (C) No hRPA-XPG complex formation on 5'-protruding 28-d(T) substrates.

(Fig. 7A , lanes 4 and 5). Complex formation could not be realized by *E. coli* SSB, showing the specificity of the XPG-hRPA interaction (Fig. 7B). Furthermore, no interaction between XPG and hRPA-DNA complexes was detected on 5'-protruding 28-d(T) substrates (Fig. 7C), revealing that also the interaction between XPG and hRPA depends on the orientation of hRPA with regard to the ds-ss junction. Thus, hRPA can only specifically interact with XPG on DNA when the duplex region starts immediately at its 5'-oriented side, which is exactly opposite to what was found for ERCC1-XPF.

Discussion

Human replication protein A (hRPA) is a DNA-binding protein involved in multiple DNA transactions, including replication, recombination, repair and possibly transcription (45). A common denominator in all these processes is a single-stranded DNA intermediate along which action takes place with a defined directionality. Two properties of hRPA have generally been acknowledged to play an important role in these processes. Firstly, its high affinity for ssDNA is thought to be required for the stabilization of single-stranded DNA intermediates and the removal of secondary structures from these regions. Secondly, its ability to specifically interact with many proteins, such as DNA polymerases and repair enzymes, is assumed to coordinate the action of these factors. Here, we introduced a new parameter important for hRPA functioning. Using a set of defined DNA structures, containing a duplex region with one single-stranded arm protruding either 3' or 5', we demonstrated that hRPA binds ssDNA with a defined polarity. A strong interaction domain of hRPA, required for initial ssDNA-binding, is positioned at the 5' side of the ss-binding region and a weaker binding domain resides at the 3' side. Functional implications were demonstrated for the NER system. The oriented binding of hRPA to ssDNA determines the positioning of the two NER nucleases on neighbouring duplex DNA and modulates the activity of ERCC1-XPF. Below, we deduce a model for NER in which hRPA is situated on the non-damaged strand and directs incisions to the damaged strand. We anticipate that the defined ssDNA-binding polarity of hRPA contributes to the directionality of other hRPA-dependent processes as well.

hRPA binding to ssDNA initiates at the 5' side of the prospective 30 nt binding region

Of the three hRPA-subunits, only RPA70 is able to stably interact with ssDNA on its own (17, 18). Within RPA70, two subdomains are present that together are required for binding to ssDNA (36). Co-crystallisation of these domains with ssDNA revealed that they form a relatively straight channel that can stably bind to an 8-mer ss-oligonucleotide (6). As binding of this hRPA fragment had a defined polarity, the interaction of the complete

hRPA trimer with ssDNA was considered likely to have polarity as well (6). Our finding that the affinity of hRPA for small single-stranded protrusions depends on their polarity strongly supports this hypothesis. Formally, the bandshift experiments in Figure 1 do not exclude the possibility that the observed binding preference is due to the hydrogen-bonded DNA-terminus rather than the polarity of the single strand. In this scenerio, a hydrogen-bonded 5', but not 3', end stabilizes the binding of hRPA molecules to short (3') protrusions. However, since we also find that modulation of the NER nucleases by hRPA strongly depends on the polarity of the single-stranded region and others have shown that the DNA-binding domain of hRPA binds single-stranded DNA with a defined polarity (6), we believe that polarity of the single strand is the main determinant of the observed binding preference. The fact that hRPA prefers short 3' protrusions suggests that initial binding occurs at the *5' side* of the prospective 30 nt ssDNA binding region (see Fig. 8, drawing [iii] and [iv]). This predicts that the two RPA70 subdomains, responsible for initial DNA-binding, are also located at the 5' side of the 30 nt binding region, and suggests a 5' to 3' directed progression from initial binding to the stable 30 nt occluded binding mode (Fig. 8). Stable binding may require a conformational change of hRPA and/or additional contacts between hRPA and the DNA.

In this study, we used *E. coli* SSB to verify the specificity of protein-protein interactions. *E. coli* SSB and eucaryotic RPA carry out many of the same essential roles in different DNA-metabolizing processes but show little amino acid sequence similarity. The 19kD *E. coli* SSB protein preferentially binds ssDNA as a homotetramer (26); RPA, on the other hand, is composed of three distinct subunits, that together form a stable ssDNA-binding monomer. A recent study revealed the presence of four potential ssDNA-binding domains (SBDs) within the three subunits of RPA, all showing weak sequence homology to the *E. coli* SSB protomer (37). Two of the SBDs reside in the RPA70 subunit and correspond to the DNA-binding domains mentioned above, the others reside in RPA14 and RPA32. Thus, binding of RPA to ssDNA may structurally resemble the *E. coli* SSB-tetramer bound to ssDNA. This notion, combined with the fact that *E. coli* SSB cannot replace hRPA in human *in vitro* repair reactions (11), makes *E. coli* SSB a proper control

to determine the specificity of interactions between hRPA and human NER proteins.

hRPA is bound to the undamaged strand during NER

To investigate possible functional implications of the defined ssDNA-binding polarity of hRPA, we focussed on its role in NER. During NER, the structure-specific endonucleases ERCC1-XPF and XPG cleave the damaged strand at the borders of a partially unwound DNA intermediate, respectively at the 5' and the 3' side of the lesion (14, 28, 33, 40). Protein-protein interactions are thought to position the two nucleases specifically onto the damaged strand and prevent incisions in the undamaged strand. hRPA was reported to interact with both ERCC1-XPF (3, 29)and XPG (19). Here, we demonstrate that the orientation of hRPA with regard to the ss-ds DNA junction plays a crucial role in the positioning of the two nucleases. On 5'-protruding substrates, the 3'-oriented side of hRPA faces the DNA junction. This allows a specific interaction with ERCC1-XPF, resulting in strong stimulation of ERCC1-XPF-mediated incisions in the DNA strand not bound by hRPA (Fig. 8, drawing [iii]). No physical interaction is observed in this orientation between hRPA and XPG, which cleaves these substrates in the same strand hRPA is bound to (Fig. 8, [iii]). Conversely, on 3'-protruding substrates the 5'-oriented side of hRPA is positioned towards the DNA junction. Now hRPA does not physically interact with ERCC1-XPF, but sterically blocks incisions by this nuclease in the strand hRPA is bound to (Fig. 8, [iv]). The 5'-oriented side of hRPA facing a DNA junction allows specific interaction with XPG, which cleaves the strand opposite the one that is bound by hRPA (Fig. 8, [iv]). Thus, hRPA positions both NER nucleases, and even stimulates ERCC1-XPF activity, when cleavage is to occur in the strand opposite the one hRPA is bound to. This implies that during NER hRPA is bound to the undamaged strand, and directs incisions to the damaged strand (Fig. 8, [v]). Obviously, it is essential that incisions are restricted to the damaged strand, as the non-damaged strand later in the NER reaction has to serve as a template for gap-filling DNA synthesis. hRPA bound to the non-damaged strand even protects this strand from inadvertent incision by ERCC1-XPF, as we demonstrated that ERCC1-XPF is not able to make incisions in a DNA strand that is bound nearby by hRPA.

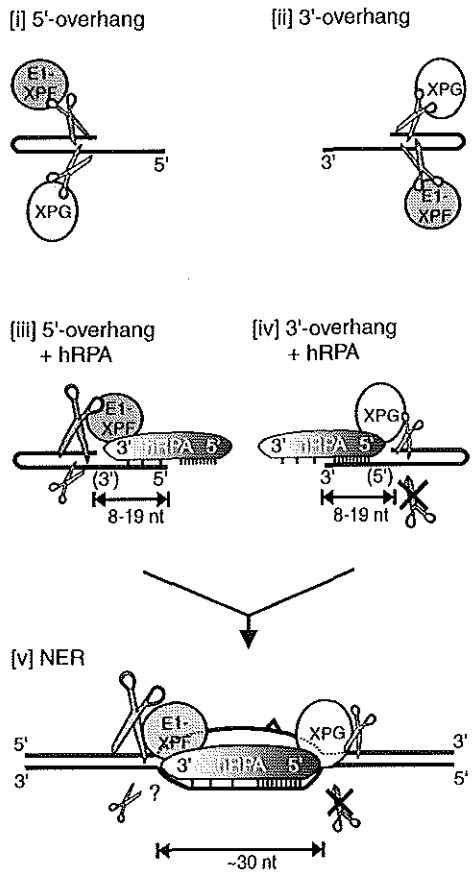


Figure 8. Model of hRPA-binding to the undamaged strand during NER. [i] and [ii] ERCC1-XPF and XPG cleavage of 5'- and 3'-protruding substrates, respectively. [iii] On short 5'-protruding arms, hRPA weakly interacts with the ss-stretch, as bordering duplex DNA physically hinders stable binding of the strong DNA-interaction domain at the 5'-oriented side of hRPA. The 3'-oriented side of hRPA positions and strongly stimulates ERCC1-XPF-mediated incisions in the non-hRPA-bound strand, whereas it does not interact with XPG and slightly inhibits its activity. [iv] On short 3'-protruding arms, the strong binding domain of hRPA stably interacts with the ss-portion. This 5'-oriented side of hRPA positions XPG-mediated incisions (and slightly inhibits XPG-activity, but see text), and completely blocks ERCC1-XPF-mediated incisions. [v] Extrapolation to nucleotide excision repair. Upon the formation of an opened DNA intermediate, which requires the activities of XPA, XPC-HHR23B, TFIIH and hRPA, an hRPA monomer is bound to the undamaged DNA strand and positions both nucleases onto the damaged strand; bound as such, it is able to strongly stimulate ERCC1-XPF incisions in the damaged strand and completely block ERCC1-XPF incisions in the undamaged strand. For clarity, XPA, XPC-HHR23B and TFIIH are not depicted.

On the other hand, inhibition of XPG-mediated incisions in hRPA-bound DNA strands was much less pronounced, indicating that possibly other factors are required to prevent XPG from cutting the template strand during NER. Curiously, hRPA-mediated positioning of XPG on the opposite strand does not result in stimulation of XPG-incision activity, but rather in slight inhibition, which seems to obscure the functional relevance of this interaction. Possibly, this relevance becomes more apparent on less-defined DNA substrates, where hRPA's ability to remove secondary structures from ssDNA may provide XPG with a better cleavable substrate. This notion is supported by the fact that stimulation of XPG-activity by hRPA was observed on substrates which are very poorly incised by XPG (29). Alternatively, hRPA may be the more active profiteer of this interaction during NER by utilizing XPG as the 5' border of its 30 nt binding region. In this respect it is interesting to note that the distance from the lesion to the XPG-mediated 3' incision, which is thought to preceed the ERCC1-XPF-mediated 5' incision, varies in a lesion type-dependent manner, whereas the 30 nt distance between the two incisions remains virtually constant (21, 30, 31).

hRPA has been reported to bind to *cis*-platin-, AAF-, and UV-damaged DNA (8, 9, 19). Together with the damage-recognition protein XPA, it showed a striking cooperativity in binding to DNA lesions (19). Our data pinpoint hRPA to the non-damaged strand during NER, strongly arguing that it is not the binding of hRPA to the damage *per se*, but rather local helical distortion and partial single-stranded character induced by the lesion that allows hRPA to bind to damaged DNA. The interaction of hRPA with XPA is likely to facilitate or stabilize the binding of the latter protein to the opposite strand containing the damage.

Implications of the defined ssDNA-binding polarity of hRPA for other DNA-metabolizing processes

Association of hRPA to ssDNA is likely to be most efficient when binding is initiated at the strong ssDNA-affinity side of the molecule, which we defined as the 5'-oriented side of hRPA. This ssDNA-binding parameter of hRPA may well have an impact

on other hRPA-dependent processes. In an *in vitro* recombination assay, binding of hRPA to single-stranded DNA was found to inhibit the formation of Rad51 nucleoprotein filaments on that same DNA template, and inhibition could be alleviated by Rad52, involving specific Rad52-hRPA and Rad52-Rad51 protein-protein interactions (2, 32, 38, 41). However, when hRPA was added after ssDNA-Rad51 nucleoprotein filaments had been allowed to form, it was found to stimulate Rad51-catalyzed strand exchange between a linear duplex and a circular single-stranded DNA fragment (1, 2, 32, 38, 41). Strand exchange involved the transfer of the 5'-terminus of a strand of linear duplex DNA to circular ssDNA (1), and our data suggest that such hRPA-mediated stimulation can be explained by preferential binding of this protein factor to the released 3'-protruding ss-arm from the linear duplex to prevent re-annealing (1, 25).

After the incision stage of NER hRPA has an additional involvement in gap-filling DNA repair synthesis, probably similar to its role during regular DNA replication (39). Both DNA polymerase δ, which is stimulated by hRPA (22), and DNA polymerase ε have been implicated in repair synthesis (10). To allow synthesis of new DNA strands, hRPA has to leave the template strand. Interestingly, the DNA polymerase, proceeding in a 3' to 5' direction on the template strand, faces the weak ssDNA-binding side, the 3'-oriented side, of hRPA. It is probably more efficient to peel off a bound hRPA molecule from ssDNA starting from the side of the protein that is not so strongly attached to the template. DNA polymerase α, involved in the initiation of DNA replication, is also stimulated by hRPA. In contrast to DNA polymerase δ however, this requires a specific interaction with hRPA (7, 22, 27). For this protein and other hRPA-interacting proteins that act with a defined directionality on DNA, like ERCC1-XPF and XPG, it will be interesting to see where the interacting subdomains reside in an hRPA molecule bound to ssDNA. Most likely, however, this awaits co-crystallisation of the complete hRPA trimer on ssDNA.

Materials and methods

Purified proteins

Recombinant human RPA, expressed from the construct p11d-tRPA (a generous gift of Dr. M. Wold), was purified from *E. coli* as described by Henricksen (20). For biochemical assays, fresh dilutions were made from a 125 ng/µl purified hRPA stock in buffer A (25 mM Tris pH7.5, 1 mM EDTA, 0.3 M NaCl, 50% glycerol, 1 mM dithiothreitol (DTT) and 0.2 mg/ml bovine serum albumin). Purification of recombinant human ERCC1-XPF complex from baculovirus-infected insect cells has been described previously (12). Purified fractions, directly used for assays, contained approximately 20 ng/µl ERCC1-XPF complex in buffer B (20 mM Hepes pH7.8, 50 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 2 mM DTT and 0.02% NP-40). Recombinant human XPG protein was purified from baculovirus-infected insect cells, using a protocol described by O'Donovan (34), with one additional purification step. After the hydroxyapatite column, the peak fractions were pooled, dialysed against buffer C (25 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.01% NP-40 and 10% glycerol) containing 0.1 M KCl, and loaded onto an HPLC Mono Q HR5/5 column, equilibrated in the same buffer. After washing, a gradient from 0.1 to 0.5 M KCl (20 ml) was applied at 0.5ml/min and 0.5 ml fractions were collected. XPG eluted between 0.4 M and 0.5M KCl, and peak fractions contained approximately 250 ng/µl of purified XPG protein. For biochemical assays, fresh dilutions of these fractions were made in buffer B.

DNA substrates

The sequences of the DNA substrates used, were as follows: 3'-protruding 10-d(T) substrate: 5'-**GGCC**GTGCTCTGAATTCCTGGATG**TTCGAAAG**/C**TTTCGAA**CATCC-AGGAATTCAGAGCAC**GGCC**$(T)_{10}$-3'; 5'-protruding 10-d(T) subtrate: 5'-$(T)_{10}$ **GGCC**-GTGCTCTGAATTCCTGGATG**TTCGAAAG**/C**TTTCGAA**CATCCAGGAATTCAG AGCAC**GGCC**-3'; 3'-protruding 19-d(T) substrate: 5'-GGCCGTGCTCTCCTGGATG**TT-CGAAAG**/C**TTTCGAA**CATCCAGGA-GAGCAC**GGCC**$(T)_{19}$-3'; 5'-protruding 19-d(T)

substrate:5'-$(T)_{19}$**GGCC**GTGCTCTCCTGGATG**TTCGAAAG**/C**TTTCGAA**CATCCAG-GAGAGCAC**GGCC**-3';3'-protruding28-d(T)substrate:5'-**GGCC**GTGCTCTGATG**TTC-GAAAG**/C**TTTCGAA**CATCAGAGCAC**GGCC**$(T)_{28}$-3'; 5'-protruding 28-d(T) substrate: 5'-$(T)_{28}$**GGCC**GTGCTCTGATG**TTCGAAAG**/C**TTTCGAA**CATCAGAGCAC**GGCC**-3'; 3'-protruding 14-d(N) substrate: 5'-GTGCTCTCCTGGATGTTCGAAAGCTGGGCGT/A-CGCCCAGCTTTCGAACATCCAGGAGAGCAC-(AGACTTGGACCCGC)-3'; 5'-protruding13-d(N)substrate:5'-(CCTAGACTTAAGA)-GGCCGTGCTCTCCTGGATGT-TCGAAAGCTGGG/CCCAGCTTTCGAACATCCAGGAGAGCACGGCC-3'.Underlined sequences were filled in by Klenow fragment, using dGTP and (radiolabeled) dCTP. HaeIII restriction sites (GG/CC) and BstB1 restriction sites (TT/CGAA) are indicated in bold.

Nuclease assays

DNA oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis (DPAGE) and 150 ng DNA substrate was ^{32}P-labeled using polynucleotide kinase or Klenow fragment, followed by phenol extraction in 100 µl and G50-sepharose column centrifugation. To allow self-annealing, oligonucleotides were heated for 3 min at 95^{0}C and put on ice. Labeled substrates were kept at 4^{0}C.

Nuclease assays (15µl) were carried out in nuclease buffer D (50 mM Tris (pH 8.0), 0.75 mM $MnCl_2$, 0.1 mg/ml bovine serum albumin and 0.5 mM β-mercaptoethanol), containing approximately 100 fmol DNA substrate. First, 1 µl with the indicated amount of hRPA (or 1 µl buffer A) was pre-incubated with the DNA substrate in buffer D for 10 min at 37^{0}C. Then 150 fmol ERCC1-XPF or 150 fmol XPG was added on ice and reactions were continued for 50 min at 37^{0}C. Reactions were stopped by adding 90% formamide, heated for 3 min at 95^{0}C and applied onto 10-15% DPAGE. Reactions products were visualized by autoradiography or phosphorimager. In the case of the footprinting-type of analysis, 10U HaeIII or 10U BstBI were added after 10 min hRPA pre-incubation, and incubation was proceded for another 4 hrs at 37^{0}C. Similar results as shown in Figure 2 were obtained when HaeIII cleavage was allowed for 50 min at 37^{0}C with 28 nt hairpin substrates.

Gel retardation assays

To analyse hRPA-DNA complex formation, the indicated amounts of hRPA were co-incubated with 100 fmol DNA substrate and 1µl buffer B in buffer D (total reaction-volume 15 µl; final KCl concentration was approximately 25 mM). No glutaraldehyde was added. After incubation for 1 hr at 37°C, reactions were put on ice, 3 µl of ice-cold 30% glycerol was added and samples were loaded immediately onto a pre-cooled 5% (1 : 40 acrylamide: bisacrylamide) native gel. Gels were run in 0.5x TBE for 2-3 hrs at 150 V (20 mA) at 4°C.

To analyse ternary complex formation, 100 fmol of 3'- or 5'-protruding 28-d(T) substrate was pre-incubated with the indicated amounts of hRPA (or 1µl buffer A) in buffer D containing either 0.75 mM $MnCl_2$, 5 mM $CaCl_2$, 5 mM $MgCl_2$ or 5 mM EDTA. After 10 min at 30°C, the indicated amounts of ERCC1-XPF or XPG (or corresponding volumes of buffer B) were added on ice, and incubation was continued for 10 min at 30°C. Then, 1.5 µl of 0.5% glutaraldehyde (freshly diluted in 50 mM Tris pH7.5) was added at RT (final concentration: 0.05% in 15 µl reaction volume) and crosslinking was allowed for 40 min (30°C). In reactions containing antibody, after 25 min of glutaraldehyde-crosslinking, 2 µl of antibody was added and incubation was continued for another 15 min at 30°C. Reactions were stopped on ice, 3 µl of ice-cold 30% glycerol was added and samples were loaded immediately onto a pre-cooled 5% (1 : 40) native PAGE (0.5x TBE) (see above). Products were visualized by autoradiography or phosphorimager.

References

1. **Baumann, P., and S.C. West.** 1997. The human rad51 protein: polarity of strand transfer and stimulation by hRP-A . *EMBO J.* **16:** 5198-5206.
2. **Benson, F., P. Baumann, and S. West.** 1998. Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. *Nature* **391:** 401-404.
3. **Bessho, T., A. Sancar, L.H. Thompson, and M.P. Thelen.** 1997. Reconstitution of human excision nuclease with recombinant XPF-ERCC1complex. *J. Biol. Chem.* **272:** 3833-3837.
4. **Blackwell, L.J., and J.A. Borowiec.** 1994. Human replication protein A binds single-stranded DNA in two distinct complexes. *Mol. Cell. Biol.* **14:** 3993-4001.
5. **Blackwell, L.J., J.A. Borowiec, and I.A. Masrangelo.** 1996. Single-stranded-DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase. *Mol. Cell. Biol.* **16:** 4798-4807.
6. **Bochkarev, A., R. Pfuetzner, A. Edwards, and L. Frappier.** 1997. Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385:** 176-181.
7. **Braun, K., Y. Lao, Z. He, C. Ingles, and M. Wold.** 1997. Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochem.* **36:** 8443-8454.
8. **Burns, J., S. Guzder, P. Sung, S. Prakash, and L. Prakash.** 1996. An affinity of human replication protein A for ultraviolet-damaged DNA. *J. Biol. Chem.* **271:** 11607-11610.
9. **Clugston, C., K. McLaughlin, M. Kenny, and R. Brown.** 1992. Binding of human single-stranded DNA binding protein to DNA damaged by the anticancer drug cis-diamminedichloroplatinum (II). *Cancer Res.* **52:** 6375-6379.
10. **Coverley, D., M.K. Kenny, D.P. Lane, and R.D. Wood.** 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.
11. **Coverley, D., M.K. Kenny, M. Munn, W.D. Rupp, D.P. Lane, and R.D. Wood.** 1991. Requirement for the replication protein SSB in human DNA excision repair. *Nature* **349:** 538-541.
12. **de Laat, W.L., E. Appeldoorn, N.G.J. Jaspers, and J.H.J. Hoeijmakers.** 1998. DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.* **273:** 7835-7842.
13. **Dornreiter, I., L. Erdile, I. Gilbert, D. von Winkler, T. Kelly, and E. Fanning.** 1992. Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *EMBO J.* **11:** 769-776.
14. **Evans, E., J. Fellows, A. Coffer, and R.D. Wood.** 1997. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J.* **16:** 625-638.
15. **Fairman, M.P., and B. Stillman.** 1988. Cellular factors required for multiple stages of SV40 DNA replication in vitro. *EMBO J.* **7:** 1211-1218.
16. **Georgaki, A., B. Strack, V. Podust, and U. Hubscher.** 1992. DNA unwinding activity of replication protein A. *FEBS Lett* **308:** 240-244.
17. **Gomes, X., L. Henricksen, and M. Wold.** 1996. Proteolytic mapping of human replication protein A: evidence for multiple structural domains and a conformational change upon interaction with single-stranded DNA. *Biochem.* **35:** 5586-5595.

18. **Gomes, X., and M. Wold.** 1996. Functional domains of the 70-kilodalton subunit of human replication protein A. *Biochem.* **35:** 10558-10568.
19. **He, Z., L.A. Henricksen, M.S. Wold, and C.J. Ingles.** 1995. RPA involvement in the damage-recognition and incision stepa of nucleotide excision repair. *Nature* **374:** 566-569.
20. **Henricksen, L., C. Umbricht, and M. Wold.** 1994. Recombinant replication protein A: expression, complex formation, and functional characterization. *J. Biol. Chem.* **269:** 11121-11132.
21. **Huang, J.C., D.L. Svoboda, J.T. Reardon, and A. Sancar.** 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.* **89:** 3664-3668.
22. **Kenny, M.K., S.H. Lee, and J. Hurwitz.** 1989. Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta. *Proc. Natl. Acad. Sci.* **86:** 9757-61.
23. **Kim, C., B.F. Paulus, and M.S. Wold.** 1994. Interactions of human replication protein A with oligonucleotides. *Biochem.* **33:** 14197-14206.
24. **Kim, C., R.O. Snyder, and M.S. Wold.** 1992. Binding properties of replication protein-A from human and yeast cells. *Mol. Cell. Biol.* **12:** 3050-3059.
25. **Lavery, P., and S. Kowalczykowski.** 1992. A postsynaptic role for single-stranded DNA-binding protein in recA protein-promoted DNA strand exchange. *J. Biol. Chem.* **267:** 9315-9320.
26. **Lohman, T., and L. Overman.** 1985. Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J. Biol. Chem.* **260:** 3594-3603.
27. **Matsumoto, T., T. Eki, and J. Hurwitz.** 1990. Studies on the initiation and elongation reactions in the simian virus 40 DNA replication system. *Proc. Natl. Acad. Sci.* **87:** 9712-9716.
28. **Matsunaga, T., D. Mu, C.-H. Park, J.T. Reardon, and A. Sancar.** 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.
29. **Matsunaga, T., C.H. Park, T. Bessho, D. Mu, and A. Sancar.** 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.
30. **Moggs, J.G., K.J. Yarema, J.M. Essigmann, and R.D. Wood.** 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.
31. **Mu, D., C.-H. Park, T. Matsunaga, D.S. Hsu, J.T. Reardon, and A. Sancar.** 1995. Reconstitution of human DNA repair excision nuclease in a highly defined system. *J. Biol. Chem.* **270:**2415-2418.
32. **New, J., T. Sugiyama, E. Zaitseva, and S. Kowalczykowski.** 1998. Rad52 protein stimulates DNA strand exchange by Rad51 and replication. *nature* **391:**407-10.
33. **O'Donovan, A., A.A. Davies, J.G. Moggs, S.C. West, and R.D. Wood.** 1994. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371:**432-435.
34. **O'Donovan, A., D. Scherly, S.G. Clarkson, and R.D. Wood.** 1994. Isolation of active recombinant XPG protein, a human DNA repair endonuclease. *J. Biol. Chem.* **269:** 15965-15968.
35. **Park, M.S., D.L. Ludwig, E. Stigger, and S.H. Lee.** 1996. Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J. Biol. Chem.* **271:** 18996-19000.
36. **Pfuetzner, R., A. Bochkarev, L. Frappier, and A. Edwards.** 1997. Replication protein A.

Characterization and crystallization of the DNA binding domain. *J. Biol. Chem.* **272**:430-434.

37. **Philipova, D., J. Mullen, H. Maniar, J. Lu, C. Gu, and S. Brill.** 1996. A hierarchy of SSB protomers in replication protein A. *Genes Dev.* **10**: 2222-2233.
38. **Shinohara, A., and T. Ogawa .** 1998. Stimulation by Rad52 of yeast Rad51-mediated recombination. *Nature* **391**:404-407.
39. **Shivji, M.K.K., V.N. Podust, U. Hubscher, and R.D. Wood.** 1995. Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochem.* **34**: 5011-5017.
40. **Sijbers, A.M., W.L. de Laat, R.R. Ariza, M. Biggerstaff, Y.-F. Wei, J.G. Moggs, K.C. Carter, B.K. Shell, E. Evans, M.C. de Jong, S. Rademakers, J. de Rooij, N.G.J. Jaspers, J.H.J. Hoeijmakers, and R.D. Wood.** 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**: 811-822.
41. **Sung, P.** 1997. Function of yeast Rad52 protein as a mediator between Replication Protein A and the Rad51 recombinase. *J. Biol. Chem.* **272**: 28194-28197.
42. **Treuner, K., U. Ramsperger, and R. Knippers.** 1996. Replication protein A induces the unwinding of long double-stranded DNA regions. *J. Mol. Biol.* **259**: 104-112.
43. **Tsurimoto, T., and B. Stillman.** 1991. Replication factors required for SV40 DNA replication in vitro. II. Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand synthesis. *J. Biol. Chem.* **266**: 1961-1968.
44. **Weinberg, D.H., K.L. Collins, P. Simancek, A. Russo, M.S. Wold, D.M. Virshup, and T.J. Kelly.** 1990. Reconstitution of simian virus 40 DNA replication with purified proteins. *Proc. Natl. Acad. Sci.* **87**: 8692-8696.
45. **Wold, M.S.** 1997. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* **66**: 61-92.
46. **Wold, M.S., J.J. Li, and T.J. Kelly.** 1987. Initiation of simian virus 40 DNA replication in vitro: large-tumor-antigen- and origin-dependent unwinding of the template. *Proc. Natl. Acad. Sci.* **84**: 3643-3647.
47. **Wood, R.D., P. Robins, and T. Lindahl.** 1988. Complementation of the xeroderma pigmentosum DNA repair defect in cell-free extracts. *Cell* **53**: 97-106.

Summary

DNA, the carrier of genetic information that encodes the recepies for all chemical processes in an organism, is constantly subject to undesirable changes. Environmental factors such as UV and ionizing radiation and compounds present in e.g. cigarette smoke and industrial pollution, as well as endogenous factors like oxygen radicals that are produced during cellular metabolism, can all cause harm to the DNA. Persisting damage can result in DNA mutations, and when mutations arise in genes controlling cell growth this may ultimately lead to cancer (uncontrolled cell growth) or cell death. In order to maintain the integrity of genetic material, cells have an intricate network of DNA repair mechanisms at their disposal.

Chapter 1 provides an overview of the different repair mechanisms that are currently known to function in higher organisms such as man. Each repair pathway recognizes and removes structurally different types of DNA lesions. *Base excision repair*, *nucleotide excision repair*, and *mismatch repair* are all multi-enzymatic processes that remove single-stranded damage as part of a larger DNA fragment and use the complementary DNA strand to restore the genetic information. Base excision repair restores base alterations, nucleotide excision repair deals with lesions that more dramatically distort the DNA helix, and mismatch repair removes replication mistakes. *O^6-methylguanine-DNA methyltransferase* and *photolyases* are examples of enzymes that directly reverse DNA adducts. Such enzymes usually act on a small spectrum of frequently occurring lesions. Double strand DNA breaks and interstrand DNA cross-links require recombinational events to be repaired. *Homologous recombination*, non-homologous *DNA-end joining*, homology-dependent *single-strand annealing* and *interstrand cross-link repair* all have been implicated in the repair of such noxious lesions. In this chapter, the enzymes involved in the various repair pathways are summarized, the phenotypical consequences of repair defects are discussed and models are presented for the mechanisms that underlie DNA repair.

In **chapter 2**, the mechanism of nucleotide excision repair (NER) is discussed in more detail. NER is the main pathway for the repair of UV-induced DNA lesions, but also removes, for example, bulky adducts introduced into DNA by polycyclic aromatic hydrocarbons, which are present in e.g. industrial pollution. The repair process encompasses five steps: (i) damage recognition, (ii) lesion demarcation, (iii) dual incision around the lesion, (iv) removal of a ~30 nucleotides DNA fragment containing the damage, and (v) gap-filling DNA synthesis and ligation. The latter step is carried out by general replication enzymes, whereas incision around the lesion requires the concerted action of repair factors. Defects in the incision stage of NER can cause the NER-deficient phenotype observed with xeroderma pigmentosum (XP) patients, characterized mainly by photosensitivity and a highly increased risk of getting skin cancer. Six factors, XPA, XPC-hHR23B, TFIIH, RPA, XPG and ERCC1-XPF, are necessary and sufficient to make a dual incision around a lesion *in vitro*. XPC-hHR23B is the factor that recognizes the damage and recruits the remainder of the repair machinery to the lesion. Also XPA has affinity for damaged DNA and might position the other factors around the lesion. The basal transcription factor complex TFIIH is able to unwind duplex DNA during NER; this activity is important to form the observed opened DNA complex that demarcates the lesion. RPA, a heterotrimeric protein complex that binds single-stranded DNA, may stabilize the unwound DNA intermediate. It also has an important function in positioning the nucleases XPG and ERCC1-XPF, which cleave the damaged strand at the 3' and 5' side of the lesion, respectively. Chapter 2 extensively discusses the activities of these factors and presents models for the molecular events that underlie the different steps of nucleotide excision repair.

Chapters 3 to 6 summarize the contributions of this thesis to the repair field. In **chapter 3**, ERCC1 is demonstrated to be tightly associated to a 115 kD protein in mammalian cells, which turns out to be XPF. The ERCC1-XPF protein complex acts as a structure-specific DNA endonuclease that cleaves at duplex single-stranded DNA junctions. In nucleotide excision repair, this complex is shown to make the incision at the 5' side of the lesion. Furthermore, we demonstrate that no additional DNA modifications are required

to start gap-filling DNA synthesis at the site of ERCC1-XPF cleavage.

In **chapter 4**, the interaction between the two proteins ERCC1 and XPF is studied in more detail. It is demonstrated that the XPF-binding domain maps to the C-terminal residues 224-297 of ERCC1. The ERCC1-binding domain is found to comprise the amino acids 813-905 of XPF. A mutation observed in an XP-F patient that disrupts this binding domain indeed affects complex formation with ERCC1.

In **chapter 5**, the DNA structural elements required for ERCC1-XPF cleavage *in vitro* are determined to better define the additional engagement of ERCC1-XPF in recombinational processes. We demonstrate that ERCC1-XPF cleaves putative recombination intermediates with 3' protruding single-stranded arms. Furthermore, it is demonstrated that one single-stranded arm, protruding either 3' or 5', is necessary and sufficient for ERCC1-XPF to correctly position its incisions in DNA junctions.

In **chapter 6**, the ERCC1-XPF activity is put in the context of other repair factors. We demonstrate that RPA has an important function in directing ERCC1-XPF-mediated incisions to the damaged strand during repair. RPA is shown to bind to single-stranded DNA in a defined orientation. This allows the definition of a 5'- and 3'-oriented side of RPA on DNA. Polarity appears crucial for positioning of ERCC1-XPF and XPG, the complementary repair nuclease, on the DNA; with the 3'-oriented side of RPA facing a duplex single-stranded DNA junction, RPA interacts with and stimulates ERCC1-XPF, whereas the 5'-oriented side of RPA at a DNA junction allows stable binding of XPG to RPA, but inhibits ERCC1-XPF cleavage. As is discussed in this chapter, these data pinpoint RPA to the undamaged strand during nucleotide excision repair.

Samenvatting (voor de leek)

Alle levensvormen zijn opgebouwd uit cellen. Bacteriën en gisten zijn eencellige organismen, terwijl mensen, dieren en planten uit een groot aantal cellen bestaan. Elke levende cel bevat een aantal enorm lange molekuulketens die samen het DNA van zo'n cel vormen. In het DNA liggen de 'recepten' opgeslagen voor de chemische processen die zich binnen een cel, of een organisme, kunnen afspelen. Telkens wanneer een cel zich deelt wordt het DNA nauwgezet verdubbeld en precies verdeeld over de twee dochtercellen. Hetzelfde gebeurt bij de voortplanting, waarbij individuen hun genetische informatie (DNA) doorgeven aan hun nakomelingen. DNA kan dus vanuit biologisch oogpunt worden beschouwd als het fundament waarop het leven is gebouwd. Veranderingen in de DNA samenstelling (mutaties) zijn maar in beperkte mate verenigbaar met 'normaal' leven. In uiterste gevallen kunnen DNA mutaties leiden tot ongecontroleerde celgroei (kanker) of celdood.

DNA mutaties kunnen ontstaan doordat DNA molekulen beschadigd worden. Zowel in het milieu als in de cel zelf zijn allerlei faktoren aanwezig die een constante bedreiging vormen voor de integriteit van het DNA. Uitlaatgassen van auto's, uitstoot van fabrieken en rook van sigaretten, bijvoorbeeld, bevatten chemische verbindingen die het DNA kunnen beschadigen. UV-licht, dat door de zon wordt uitgezonden, kan de structuur van het DNA aantasten. In de cel zelf worden stoffen (bijvoorbeeld zuurstof radicalen) aangemaakt die het DNA kunnen beschadigen. Samen zorgen deze faktoren ervoor dat er voortdurend een enorm aantal beschadigingen in het DNA worden aangebracht. Geschat wordt bijvoorbeeld dat in normale menselijke cellen telkens 150 duizend (1.5×10^5) van de in totaal 6 miljard (6×10^9) DNA bouwstenen (de zogenaamde basen) beschadigd zijn door zuurstof radicalen. Dat maar een zeer kleine fraktie van deze schades tot een definitieve verandering in de DNA samenstelling leidt komt, doordat cellen beschikken over uitermate efficiënte DNA reparatie enzymen. Dit proefschrift tracht inzicht te verschaffen in de werking van deze DNA reparatie enzymen.

In de eerste twee, inleidende, hoofdstukken wordt de huidige kennis omtrent DNA herstel samengevat. De enzymen die voor DNA herstel zorgen worden behandeld, en er wordt bediscussieerd hoe deze enzymen samenwerken om schades in het DNA op te sporen en te herstellen.

In **hoofdstuk 1** wordt duidelijk gemaakt dat er meerdere reparatie mechanismen (enzym systemen) naast elkaar bestaan die ieder een ander type schade verwijderen. Meestal bepaalt de structuur van het beschadigde DNA welk reparatie systeem gewenst is. Er zijn maar een paar voorbeelden bekend waarbij een schade door een enkel enzym direct wordt hersteld. De meeste schades vereisen samenwerking van groepen enzymen. Men spreekt dan van multi-enzymatisch reparatie processen, waarbij elk enzym een specifiek aandeel levert in het stapsgewijs verloop van het herstel.

DNA molekulen bestaan uit twee strengen van bouwstenen die onderling contact maken. Voor dit contact is het belangrijk dat de juiste bouwstenen tegenover elkaar liggen, en daarom bepaalt de volgorde van bouwstenen van de ene streng die van de ander, en vice versa. Multi-enzymatische DNA herstel mechanismen maken gebruik van dit gegeven voor de reparatie van schades die zich beperken tot één streng van het DNA. Zij beschikken over enzymen die de schade kunnen herkennen en die ervoor zorgen dat alleen de aangetaste DNA streng wordt doorgeknipt aan weerszijden van de schade. Het stukje DNA van deze streng met de schade erop wordt verwijderd, waarna het ontstane gat wordt opgevuld met de bouwstenen die passen op die van de tegenoverliggende, onbeschadigde, streng. Voor het herstel van zogenaamde dubbelstrengs breuken en kruisverbindingen tussen twee DNA strengen zijn andere groepen enzymen nodig, die ervoor moeten zorgen dat verschillende stukken DNA aan elkaar geplakt worden. In dit hoofdstuk wordt uiteengezet welke enzymen er bij de verschillende reparatie mechanismen betrokken zijn, hoe deze enzymen individueel en met elkaar funktioneren, en wat de eventuele consequenties zijn van defecten in deze herstel systemen.

In **hoofdstuk 2** wordt dieper ingegaan op één van deze herstel systemen, namelijk nucleotide excisie reparatie (afgekort tot 'NER'). NER is een voorbeeld van zo'n multi-enzymatisch reparatie systeem, dat schades in één streng van het DNA herstelt met behulp van de tegenoverliggende, onbeschadigde, streng. Het is het belangrijkste herstel mechanisme voor DNA schades die door UV licht worden gemaakt, maar het verwijdert ook schades die door stoffen in uitlaatgassen en sigarettenrook in het DNA kunnen worden aangebracht. Patienten met de ernstige erfelijke ziekte xeroderma pigmentosum (XP) hebben een defect in één van de enzymen die zorgen voor NER, en zijn dus niet in staat om dergelijke schades uit het DNA te verwijderen. Zij zijn erg gevoelig voor zonlicht en hebben bijvoorbeeld een meer dan 1000-maal verhoogde kans op huidkanker. Het reparatieproces in de cel omvat vijf stappen: (i) schade herkenning, (ii) markering van de schade, (iii) twee knippen in de beschadigde DNA streng, aan iedere kant van de schade één, (iv) verwijderen van het DNA strengetje met de schade, en (v) opvullen van het onstane gat met de bouwstenen die passen op de tegenoverliggende, onbeschadigde, streng. *In vitro*, in een reageerbuisje, zijn zes enzymen voldoende om de eerste drie stappen van NER te bewerkstelligen. Zij heten XPC-hHR23B, XPA, TFIIH, RPA, XPG en ERCC1-XPF. XPC-hHR23B bestaat uit twee eiwit componenten en is in staat om schades in het DNA op te sporen en vervolgens de rest van de reparatie enzymen aan te trekken. Het XPA eiwit hecht zich ook aan DNA schades, en gedacht wordt dat dit eiwit de overige enzymen op de juiste wijze ten opzichte van de schade plaatst. TFIIH bestaat uit 9 eiwit componenten en kan de twee DNA strengen ontwinden, waardoor rondom de schade enkelstrengs DNA ontstaat. RPA zoekt dit soort enkelstrengs stukken DNA op en voorkomt dat ze weer dubbelstrengs worden. Zoals wordt aangetoond in hoofdstuk 6, is RPA ook belangrijk voor het op de juiste plaats zetten van de twee enzymen die de knippen in het DNA zetten, de zogenaamde nucleases (DNA knippers) XPG en ERCC1-XPF.

De **hoofdstukken 3 tot en met 6** vatten het praktisch werk van dit promotieonderzoek samen. Twee vragen stonden centraal bij dit onderzoek: Welke enzymen maken de knippen rondom de schade en hoe worden deze enzymen gecoördineerd?

In **hoofdstuk 3** laten we zien dat het ERCC1 eiwit in de cel vastzit aan het XPF eiwit. Dit ERCC1-XPF enzym complex knipt DNA daar waar dubbelstrengs DNA overgaat in enkelstrengs DNA. Dergelijke DNA overgangen worden ook gevormd tijdens de eerste twee stappen van het DNA herstel, en inderdaad vinden we dat ERCC1-XPF één van de twee knippen maakt tijdens NER.

In **hoofdstuk 4** wordt dit eiwit complex meer in detail bestudeerd. Onderzocht is met welke bouwstenen (aminozuren) de twee eiwitten aan elkaar gebonden zijn. Zowel in ERCC1 als in XPF blijken aminozuren in de C-terminus (de 'achterkant' van de eiwitten) de eiwit-eiwit binding te bewerkstelligen. Het in kaart brengen van welke aminozuren in een eiwit waarvoor verantwoordelijk zijn is met name interessant omdat dan beter mogelijk is te bepalen wat er mis gaat met de kapotte eiwitten van patienten. We vinden bijvoorbeeld dat een mutatie in een xeroderma pigmentosum groep F (XP-F) patient de aminozuren aantast die nodig zijn voor de eiwit-eiwit contacten met ERCC1, en inderdaad zijn deze kapotte XPF eiwitten niet in staat om ERCC1 te binden. Heel waarschijnlijk dus ligt verminderde affiniteit van XPF voor het ERCC1 eiwit ten grondslag aan de klinische verschijnselen die deze patient karakteriseren.

ERCC1-XPF is niet alleen bij NER betrokken, maar funktioneert ook bij het DNA herstel van de zeer schadelijke kruisverbindingen tussen twee DNA strengen. Over dit reparatie proces is nog erg weinig bekend, en de rol van ERCC1-XPF erin is slecht beschreven. Het ligt echter voor de hand dat dit enzym complex ergens tijdens het proces nodig is om het DNA te knippen. In **hoofdstuk 5** definieren we precies wat voor soorten overgangen tussen dubbelstrengs en enkelstrengs DNA ERCC1-XPF allemaal wel en niet knipt, waardoor in de toekomst een gerichtere zoektocht naar de rol van dit

enzym complex in het herstel van kruisverbindingen mogelijk is.

In **hoofdstuk 6**, tenslotte, onderzoeken we welke eiwitten tijdens NER de knip-enzymen ERCC1-XPF en XPG rondom de schade zetten, zodat de knip op de juiste plaats wordt gezet. Net als ERCC1-XPF is XPG een enzym dat knippen maakt bij overgangen van dubbelstrengs naar enkelstrengs DNA. Dergelijke overgangen worden tijdens het reparatie proces gevormd, maar het is ook uiterst belangrijk dat knippen alleen worden gezet in de beschadigde DNA streng, en dat de onbeschadigde streng intact blijft. RPA, het eiwit dat aan enkelstrengs DNA bindt, blijkt een cruciale rol in de coördinatie van de knip-enzymen te spelen. We laten zien dat RPA altijd op dezelfde manier, in één oriëntatie, aan DNA bindt en dat tijdens DNA herstel RPA aan de onbeschadigde DNA streng gebonden is. Doordat de ene kant van RPA kan binden aan ERCC1-XPF, en de andere kant van RPA kan binden aan XPG, is RPA in staat om de twee enzymen op de juiste posities, d.w.z. in de beschadigde DNA streng, ieder aan één kant van de schade, te laten knippen.

List of publications

Medema R.H., de Laat, W.L., Martin, G.A., McCormick, F., and Bos, J.L. (1992). GTPase-activating protein SH2-SH3 domains induce gene expression in a Ras-dependent fashion. *Mol. Cell. Biol.* **12**: 3425-3430.

Cox, M., van Tilborg, P.J., de Laat W.L., Boelens, R., van Leeuwen, H.C., van der Vliet, P.C., and Kaptein, R. (1995). Solution structure of the Oct-1 POU homeodomain determined by NMR and restrained molecular dynamics. *J. Biomol. NMR.* 6: 23-32.

Sijbers, A.M.#, de Laat, W.L.#, Ariza, R.R.#, Biggerstaff, M.#, Wei, Y.-F., Moggs, J.G., Carter, K.C., Shell, B.K., Evans, E., de Jong, M.C., Rademakers, S., de Rooij J., Jaspers, N.G.J., Hoeijmakers, J.H.J., and Wood, R.D. (1996). Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**: 811-822. (*# These authors contributed equally to this work*).

van Leeuwen, H.C., Strating, M., Rensen, M., de Laat, W.L., and van der Vliet, P.C. (1997). Linker length and composition influence the flexibility of Oct-1 DNA binding. *EMBO J.* **16(8)**: 2043-2053.

Hoeijmakers, J.H.J., van der Horst, G.T.J., Weeda, G., Vermeulen, W., Winkler, G.S., de Boer, J., de Laat, W.L., Sijbers, A.M., Citterio, E., Jaspers, N.G.J., Egly, J-M., and Bootsma, D. (1997). Recombining DNA damage repair, basal transcription and human syndromes. *In: 'Genomics instability and immortality in cancer'* (eds. E. Mihich and L. Hartwell) Plenum Publishers, New York, pp. 83-96.

Vermeulen, W., de Boer, J., Citterio, E., van Gool, A.J., van der Horst, G.T.J., Jaspers, N.G.J., de Laat, W.L., Sijbers, A.M., van der Spek, P.J., Sugasawa, K., Weeda, G., Winkler, G.S., Bootsma, D., Egly, J-M., and Hoeijmakers, J.H.J. (1997). Mammalian nucleotide excision repair and syndromes. *Bioch. Soc. Transact.* **25**: 309-315.

de Laat, W.L., Appeldoorn, E., Jaspers, and N.G.J., Hoeijmakers, J.H.J. (1998). DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.* **273**: 7835-7842.

de Laat, W.L., Appeldoorn, E., Sugasawa, K., Jaspers, N.G.J., and Hoeijmakers, J.H.J. (1998). DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Submitted.*

de Laat, W.L., Sijbers, A.M., Odijk, H., Japers, N.G.J., and Hoeijmakers, J.H.J. (1998). Mapping of interaction domains between human repair proteins ERCC1 and XPF. *Submitted.*

Curriculum vitae

Name: Wouter Leonard de Laat

Born: 14 February 1970, Amersfoort, The Netherlands.

Sept. 1982 - June 1988:	Stedelijk Gymnasium 'Johan van Oldenbarnevelt', Amersfoort.
Sept. 1988-Nov. 1993:	Undergraduate Student, Faculty of Biology, Rijksuniversiteit Utrecht, The Netherlands.
May 1990 - Aug. 1990:	Department of Pharmacology, New York University, New York, USA (Dr A.M. Honegger, Prof. dr J. Schlessinger).
Sept. 1991 - Aug. 1992:	Laboratory of Physiological Chemistry, Rijksuniversiteit Utrecht (Drs R.H. Medema, Prof. dr J.L. Bos).
Nov. 1992 - March 1993:	Bijvoet Center for Biomolecular Research, Rijksuniversiteit Utrecht (Drs M. Cox, Prof.dr R. Kaptein).
April 1993 - Aug. 1993:	Laboratory of Physiological Chemistry, Rijksuniversiteit Utrecht (Drs H.C. van Leeuwen, Prof. dr P.C. van der Vliet).
Dec. 1993-June 1998:	PhD. Student, Faculty of Medicine, Department of Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands. Promotors: Prof. dr J.H.J. Hoeijmakers, Prof. dr D. Bootsma Co-promotor: Dr N.G.J. Jaspers

Nawoord

Op deze plek wil ik nog even een aantal mensen bedanken die een bijdrage hebben geleverd aan de totstandkoming van dit proefschrift. Ik realiseer me dat iedereen die op enigerlei wijze betrokken is geweest bij het verwekken van dit geesteskind (lees: proefschrift), bij de bevalling (lees: promotie) eigenaardigerwijs vooral gaat roeren in de nageboorte (lees: nawoord) op zoek naar tekens van herkenning. Echter, de afdeling is me wat te groot om iedereen persoonlijk te bedanken, en ik wil me tot een aantal personen beperken. Desalniettemin: iedereen bij genetica...bedankt voor de prettige en leerzame tijd! Hetzelfde geldt natuurlijk voor de Celbio's next door (ik moet aan m'n toekomst denken).

Allereerst wil ik mijn ouders bedanken voor de steun die zij mij altijd geven. Sieg en Ria, ik vind het geweldig zoals jullie met mij, Cas en Joost en onze niet altijd even voordehand liggende keuzes omgaan, die variëren van 'ik wil bioloog worden' tot 'ik wil in het buitenland wonen'. Ik beloof jullie: deze jongen blijft (voorlopig nog even) in Nederland! Sieg, ondanks dat deze bioloog niet verschrikkelijk veel kaas van de ontwikkelingsbiologie heeft gegeten, vind ik het toch leuk dat ik kan zeggen dat wij vakgenoten zijn.

Mijn beide promotoren Jan Hoeijmakers en Dick Bootsma en mijn co-promotor Koos Jaspers wil ik bedanken voor het vertrouwen dat zij in mij stelden en de mogelijkheden die zij mij hebben geboden tijdens de afgelopen vier jaar. Jan, ik heb het werken onder jouw bezielende (bege-)leiding als stimulerend en leerzaam ervaren, en ik waardeer het zeer dat jij altijd spontaan weer tijd vrij wist te maken om over de laatste proeven te praten. Datzelfde geldt voor Koos, met wie ik het meest regelmatig knip-en-shift evaluaties hield. Je nuttige presentatietips mogen niet onvermeld blijven. Koos, bedankt voor de prettige samenwerking.

Esther en Hanny wil ik bedanken voor hun onmisbare praktische bijdrage aan dit proefschrift. Esther, het recombinante eiwit complex dat jij uit insekten cellen wist te

peuteren heeft het mogelijk gemaakt om hoofdstuk 5 en 6 bij elkaar te pipetteren. Hanny heeft hoofdstuk 4 in z'n geheel bij elkaar gepipetteerd. Het was goed toeven met jullie in het 'kippenhok' 734, en met Anneke natuurlijk, aan wiens zijde ik in vier jaar het ERCC1 complex groot heb zien worden. I would like to acknowledge Kaoru Sugasawa, who purified XPG and RPA in his spare time, which allowed me to obtain the results presented in chapter 6.

I would also like to thank Dr Rick Wood for having me over in his lab (ICRF, South Mimms, UK) in the summer of '95, and Maureen Biggerstaff, together with whom I purified the ERCC1-XPF complex from an amazing amount of Chinese hamster cells.

Mijn studenten Eric (taxol), Maarten (contact-advertentie) en Jos (ski-les), bedankt. Eric, van jouw resultaten wordt nog regelmatig gebruik gemaakt.

Jan en Jeroen, twee koningen van de afdeling en tevens paranimfen van ondergetekende, het is niet meer te achterhalen wat ons drieën vanuit Utrecht in Rotterdam heeft doen belanden, maar wat mij betreft was het een feest om deze periode met jullie te delen.

In de categorie nog niet vermeld, (vaak) wetenschappelijk niets wijzer van geworden, maar toch meer dan collega's: Anja (laat ik nou met jou beginnen), Wim (de ouwe), Tol (koning), Fokke (koning), Wim (de jonge, tevens koning), Marella (haring) en Ronald (in den beginne); mede dankzij jullie aanwezigheid ging ik ook om niet-wetenschappelijke redenen met veel plezier naar het lab.

Rest mij nog om vrienden te bedanken voor het begrip voor mijn langdurig onderduiken in de laatste fase van mijn promotie onderzoek. Of deze Salman Rushdie-achtige aktie ook werkelijk geaccepteerd wordt merk ik wel wanneer ik weer eens opbel. Cas en Joost, al gaan jullie op de Noordpool zitten, ik kom lekker toch langs. En natuurlijk Evalyne, aan wie ik dit boekje heb opgedragen, niet in de eerste plaats omdat ik weet dat je zoveel van lezen houdt.