Mammalian nucleotide excision repair in vivo

Moné, M.J.

Citation for published version (APA):
Chapter 6

- PROSPECT -
DNA is a chemically reactive molecule. As a result, this carrier of the genetic message has to sustain multiple assaults on its constitution. In the daily life of a cell, these attacks are continuous and are launched from many different sources. For instance, DNA injuries can arise as a side effect of normal cellular events, like oxidative metabolism, mistakes made by machineries that operate on DNA, or even spontaneous modifications through hydrolysis; they can be induced by several types of radiation, or they can result from a broad range of chemical agents from the environment. Hence, DNA damage appears in many guises: single-strand and double-strand breaks, insertions, deletions, mismatches, abasic sites, intra-strand and inter-strand cross-links, and numerous other genotoxic chemical modifications, ranging from small alkyl groups to large "bulky" adducts (Friedberg et al., 1995). If not taken care of, organisms can suffer harmful consequences from these alterations, since DNA damage can interfere with important DNA-associated processes (e.g. transcription, replication). Moreover, DNA damage may lead to heritable mutations, many of which are detrimental to cellular function. Biologically, however, genomic DNA appears to be surprisingly stable, as its genetic integrity and stability are generally well preserved. The universal importance of this genomic safeguarding is underscored by the large number of proteins that have been implicated in cellular responses to DNA damage: some 130 gene products can be distinguished in humans (Wood et al., 2001), and it seems safe to speculate that this is only a cautious lower limit. It has become clear that in the course of evolution cells have become well equipped with an arsenal of sophisticated DNA maintenance systems. These include, among others, base excision repair, nucleotide excision repair, direct reversal of damage, translesion synthesis, mismatch repair, end joining and homologous recombination. These repair processes are generally not "stand-alone", yet they have various interconnections and, in addition, have been shown to interact with many other cellular processes as well, emphasizing the central position of DNA repair in the cell function.

Here we will focus our attention on what is probably the most versatile DNA repair system: nucleotide excision repair (NER). We will briefly highlight some of the general features of this repair mechanism (for extensive overviews, see for instance (Cline and Hanawalt, 2003; Friedberg et al., 1995; Hoeijmakers, 2001) and references therein. For a broad historical review of the DNA repair field, see for instance (Friedberg, 1997)). Subsequently, we will address a few remaining gaps in our current knowledge of this repair mechanism. These include i) the mechanism of DNA lesion recognition, ii) the events that follow the incision of the damaged strand, and iii) the consequences of the chromatin context in which the repair reaction needs to operate in vivo.
6.1 Nucleotide excision repair

NER is a genomic maintenance system that is found in all life forms. Its general mechanism was initially clarified in *Escherichia coli* where it was shown that the damage is removed by the excision of an oligonucleotide of about 12 nucleotides. Subsequently, the residual gap can be resynthesised in an error-free fashion by using the undamaged strand as a template. In mammalian cells this *modus operandi* has remained similar, yet the machinery appears to be more intricate. Mammalian NER involves the action of some 30 gene products that have to operate within a very heterogeneous chromatin environment. The multiple stages that are necessary for a successful repair event encompass recognizing and verifying a damaged site, assembly of a double-incision complex, excision of the damaged DNA strand, repair synthesis of the removed DNA, ligation of the remaining nick, and restoration of the local chromatin structure. Importantly, placental mammals are not equipped with backup repair systems for the removal of for instance UV-induced DNA damage, a naturally abundant classical substrate for NER. This has dramatic clinical effects, as several hereditary diseases are directly linked to deficiencies somewhere in the chain of events during NER. Human patients that suffer from the recessive disorder xeroderma pigmentosum (XP) have extreme photosensitivity and cancer predisposition. Seven complementation groups of XP (designated XP-A through XP–G) reflect distinct NER genes that are associated with this rare syndrome. Another NER-associated heritable disorder is Cockayne syndrome (CS), which is brought about by a mutation in the CSA or CSB gene. The clinical symptoms of CS are very different from those in XP, as CS patients do not exhibit proneness to cancer, yet suffer from a wide range of severe physical and neurological abnormalities. These remarkable differences have been rationalized from the fact that in addition to compromising NER, protein mutations that cause CS affect transcription as well. A paradigm for this multiple functionality is TFIIH, a core factor in both NER and transcription. Mutations in this protein can indeed result in combined clinical symptoms of XP and CS (for detailed reviews of human NER syndromes, see for instance (Berneburg and Lehmann, 2001; de Laat et al., 1999). In addition to sharing certain gene products, NER and transcription are also directly linked at the repair level itself. Besides a global genome NER (GG-NER) mechanism, a distinct transcription-coupled NER (TC-NER) subpathway is employed to resolve DNA injuries in active genes. The CSA and CSB proteins are both required for TC-NER, whereas the XPC-HR23B heterodimer is specific for GG-NER. The main difference between these two systems seems to be the mechanism by which the damage is detected. In TC-NER, DNA damage in the transcribed strand of an active gene is recognized by an elongating RNA polymerase II that stalls on the lesion. This triggers rapid repair which is mediated by the CSA and CSB proteins that somehow act in displacing or removing the blocked polymerase to allow access to the damage by the core NER machinery (Svejstrup, 2002). In GG-NER, recognition of the injury involves XPC-HR23B. Following damage recognition, the later stages of the repair reaction are supposedly identical for TC-NER and GG-NER.

Repair deficient cell lines have been employed extensively to clarify the mechanism of NER and, in addition, to scrutinize the molecular basis for the clinical symptoms of hereditary repair diseases. All core NER genes have been cloned and the GG-NER reaction can be successfully performed *in vitro* using reconstituted systems. This has catalysed our understanding of the biochemical characteristics of NER, and we are now aware of many protein interactions, conformational changes and enzymatic properties of this system.
Especially impressive is the large body of knowledge on the anatomy of the double incision complex (e.g., de Laat et al., 1999). The exact mechanism of formation of the incision complex has been a matter of debate. Currently however, the consensus view is that of a sequential assembly mechanism. *In vivo*, this was corroborated by the fact that NER proteins diffuse rapidly in the cell nucleus and become transiently immobilized in the presence of DNA damage. Their effective diffusion coefficients comply with their molecular weights, irrespective of the presence or absence of DNA damage, showing that their majority is not part of large pre-established complexes (Hoogstraten et al., 2002; Houtsmuller et al., 1999; Rademakers et al., 2003). This is in agreement with the observation that NER factors only weakly interact with one another *in vitro* (Araújo et al., 2001).

In this thesis, we show that repair deficient cell lines can still assemble intermediate NER complexes at damaged DNA. For this, we have developed of novel technique that allows one to irradiate only a limited volume of the cell nucleus with UV light (this thesis, Chapter 2) (Moncé et al., 2001). At present, this method has become a standard procedure for studying cellular responses to DNA damage. Our results suggest that the dual-incision complex is formed at sites of damage from the individual NER factors up to the point where the sequential recruitment was interrupted by the lacking protein (this thesis, Chapter 3) (Volker et al., 2001). In addition to providing evidence for a sequential assembly mechanism *per se*, we were able to scrutinise to a large extent the order in which the NER proteins are recruited into the dual-incision complex. XPC appeared to be the earliest core NER factor to bind to sites of DNA damage, confirming its putative role as a damage sensor, and incorporation of the endonuclease ERCC1-XPF seems to be the last step during formation of a functional pre-incision complex (this thesis, Chapter 3). Our *in situ* findings were recently confirmed *in vitro* through the characterization of similar intermediate complexes that could be trapped on an immobilized stretch of damaged DNA (Riedl et al., 2003). Finally, we were able to corroborate that the notion of a sequential assembly mechanism indeed appears to be the case *in vivo*. Employing living cells, we measured assembly rates of GFP-tagged NER factors at sites of DNA damage, reflecting de novo formation of NER complexes. This showed distinct binding kinetics for different NER proteins, providing direct *in vivo* evidence for a stepwise, on-the-spot binding mechanism (this thesis, Chapter 4). In addition, not the mere binding of TFIIH, yet also its helicase activity, appeared to facilitate the recruitment of ERCC1-XPF into the maturing repairosome.

### 6.2 How is DNA damage recognised in GG-NER?

At any time in the daily life of a cell, NER needs to resolve DNA damage that can be whichever from a chemically diverse collection of lesions residing anywhere in the genome within a structurally heterogeneous chromatin context. An important denominator of its efficiency will be how well NER is able to detect the damage. It has been observed that there is a strong positive correlation between the degree of DNA helix distortion and the rate of repair. This is, for instance, illustrated by the different repair kinetics of the two major UV-induced photoproducts: the strong DNA-bending 6-4 photoproduct (6-4PP) is repaired 5 to 10 times faster than the less helix distortive cyclobutane pyrimidine dimer (CPD) (Mitchell and Nairn, 1989). Nonetheless, the answer to the question which factor
recognizes the DNA damage has remained ambiguous to date. Several NER factors have been implicated in the initial detection of lesions. *In vitro*, XPC-HR23B, RPA, XPA, TFI11H and DDB all display a higher binding affinity for damaged over undamaged DNA (Wood, 1999). XPC-HR23B has been shown to bind directly to DNA of which the duplex had been artificially disrupted by the introduction of three to five adjacent mismatches (Sugasawa *et al.*, 2001) indicating that recognition of helix distortion might be the initial event in NER. This notion that actual damage verification takes place after mere DNA deformation detection is strengthened by the observation that GG-NER could even be performed without the need for XPC-HR23B, if such an artificial bubble of at least 5 nucleotides was adjacent to both sides of a UV photoproduct (Mu *et al.*, 1996). Therefore, the role of XPC-HR23B in GG-NER seems to be binding to helix distortions and thereby locally melting the DNA to allow further recruitment of NER factors that assess whether, and in which strand, there is DNA damage. One of the factors that may be involved in the verification of the damage is the core NER factor XPA. Cells that lack XPA are unable to perform any NER (both TC-NER and GG-NER), underlining its central role in this repair pathway. Although this protein has long been implicated in damage recognition because of its high affinity for damaged double-stranded DNA, recent studies have weakened the hypothesis that XPA is involved in bulk recognition of lesions. Using a reconstituted NER system, it was found that repair was much faster if damaged DNA was pre-incubated with XPC-HR23B prior to addition of the other factors, whilst pre-incubation with XPA did not accelerate the removal of lesions compared to the case in which all factors were added at once (Sugasawa *et al.*, 1998). This suggested that among the core factors of NER, XPC-HR23B is the first to associate with DNA damage. Employing a technique to locally irradiate cell nuclei with UV light (this thesis, Chapter 2) (Moné *et al.*, 2001), this perception has gained support in intact cells. The ability of several NER proteins to associate with DNA damage was tested in various NER deficient backgrounds. Whereas in wild type cells all proteins displayed accumulation in UV irradiated nuclear areas, none of the tested NER core proteins — including XPA — were detected at damaged sites in an XPC-deficient cell line (this thesis, Chapter 3) (Volker *et al.*, 2001). Taken together, this confirms that XPC-HR23B is directly involved in the detection of DNA damage.

Another NER factor that possesses very high affinity for various types of DNA lesions is the damaged-DNA binding protein (DDB). So far, its particular role in NER has remained elusive (reviewed by Tang and Chu, 2002). Reminiscent of XPC-HR23B, DDB appears to have a function in GG-NER and is dispensable for TC-NER. It involves two subunits, DDB1 (p127) and DDB2 (p48). Certain mutations in the *DDB2* gene abolish the damage-binding activity of this factor, as can be the case in cells from humans XP-E patients. Although it can readily bind to both CPDs and 6-4 PPs, it was initially found not to be required for NER *in vitro* (Aboussoukha *et al.*, 1995). Recent *in vitro* studies, however, demonstrated that DDB was able to considerably accelerate the removal of CPDs (about 17 times faster) and, although much less efficient, of 6-4PPs (Wakasugi *et al.*, 2001a). This is compatible with the *in vivo* observations that DDB2 significantly stimulated CPD repair (Tang *et al.*, 2000) and also somewhat contributed to resolving 6-4PPs (Hwang *et al.*, 1999). In comparison to the other XP complementation groups, XP-E patients exhibit relatively mild clinical symptoms. This is also reflected by their fairly high levels of residual DNA repair synthesis of about 50%. This has fuelled the assumption that DDB might act as an accessory factor in NER, assisting in the recruitment of XPC-HR23B to DNA damage, especially to CPDs. Indeed, XPC-HR23B has a much lower binding affinity for CPDs in comparison to 6-4PPs. Concurrently, rodent cells do not express DDB2 and are
consequently defective in GG-NER of CPDs (Hwang et al., 1998). Interestingly, UV irradiation of cells can result in the upregulation of the DDB2 gene, a process that is mediated by p53 (Hwang et al., 1999). An increase over time of the available amount of DDB may be necessary to repress mutagenesis by assisting in getting rid of persistent CPDs, which XPC-HR23B cannot properly detect. This, however, underlines that rodent cells, if they are not rescued by DDB2 expression, might not be suitable model systems for human carcinogenesis, since they lack the p53-dependent upregulation of DDB2 (Hanawalt, 2001). Several repair deficient mouse models have indeed shown to exhibit increased cancer predisposition in comparison (de Boer and Hoijmakers, 1999).

The characteristics of damage removal thus argue that there might be more proteins that can assist in detecting DNA damage. Recently, various alternative models for damage recognition have been proposed. The single-strand DNA binding protein RPA, which is indispensable for NER, has been suggested to bind to DNA damage prior to XPC-HR23B recruitment (Reardon and Sancar, 2002). However, we found that XPC-deficient cells that had been locally UV-irradiated exhibited no accumulation of RPA in the damaged nuclear areas (Rademakers et al., 2003). Another study put forward the idea that DDB might be directly responsible for the recruitment of XPA and RPA to sites of DNA damage, hence circumventing the requirement for XPC-HR23B (Wakasugi et al., 2001b). Nonetheless, in this thesis we show that XPC-deficient (but DDB-proficient) cells failed to accumulate XPA in damaged nuclear areas (this thesis, Chapter 3) (Volker et al., 2001). It seems therefore likely that, in vivo, the bulk of DNA damage is detected by XPC-HR23B, sometimes aided by DDB. At this point it cannot be ruled out that there are certain lesions for which NER does not require the involvement of XPC-HR23B, but this would probably constitute only a small fraction of lesions. In addition, a random binding model for lesion recognition has been proposed, in which RPA, XPA, and XPC cooperatively establish damage detection (Reardon and Sancar, 2003). This would have certain ramifications for repair efficiency, as we could predict from a mathematical model for NER (this thesis, Chapter 5). For instance, if one assumes that cells attempt to get rid of DNA damage as quickly as possible in order to prevent cytotoxicity and mutagenesis, the choice for a damage recognition system that involves multiple damage detectors can harbour some risks. This is especially the case for the NER core proteins, since their simultaneous presence in the maturing repairosome is required at some point during complex formation. If they could all bind DNA damage individually, these factors could become sequestered on separate lesions, which, in turn, will lead to a decrease in repair efficiency. The mathematical model for NER could show that repair efficiency does not benefit from random assembly of factors (this thesis, Chapter 5). Nonetheless, specific types of lesions might be recognized differently. As described above, DDB is likely to aid in damage recognition. However, it does probably not need to stay at the damage after XPC-HR23B has arrived, since there are no indications that DDB has a function at a later stage of the incision reaction. Mathematical modelling predicts that such displacement mechanism could be kinetically beneficial for rapid removal of DNA injuries (this thesis, Chapter 5).

6.3 What happens after dual incision?

The dual incision complex is assembled by sequential recruitment of XPC-HR23B, TFIIH, RPA, XPA, and the endonucleases XPG and ERCC1-XPF (this thesis). For the 3' and 5'
incisions to be accomplished, the bi-directional helicase activity of TFIIH is required to unwind the duplex around the lesion. In contrast to our relatively detailed knowledge up to this point (de Laat et al., 1999), the events that follow have remained far less explored. Yet, as ERCC1-XPF is likely to be the last factor that is recruited to the incision complex (this thesis), it is noticeable that this 5' endonuclease was found to remain at the damage for several minutes (Houtsmuller et al., 1999). Hence, it may be that (part of) the incision complex stays behind after having performed the double incision, feasibly to mediate the events that follow. After incisions, the oligonucleotide containing the damage is removed and RFC, PCNA and DNA polymerase δ/ε are recruited to the site to resynthesise the DNA, after which Ligase I can seal it. It is feasible that their recruitment is facilitated by the core NER factors. Since the replication machinery involves RPA, one could envisage that at least the already bound RPA protein might need to remain at the gap. This, in turn, could prevent other proteins that constituted the double incision complex from leaving, whether functionally relevant or not. Recent in vitro experiments indeed showed that RPA remains associated to become part of the resynthesis apparatus. Recruitment of the repair machinery caused the remaining incision complex to dissociate (Riedl et al., 2003). It will be important to assess the order of recruitment and dissociation, and the dynamical properties, of these late events in NER. These data would be a great step forwards to come to an all-embracing model of the NER pathway.

6.4 Repair on a chromatin template

Over the recent years it has become apparent that for comprehending DNA-associated processes one needs to take into account the chromatin environment in which they occur. Maybe in this context one should not speak of DNA repair, but rather of chromatin repair. The many different types of lesions that a cell has to sustain can reside in many different types of chromatin structures: transcriptionally active or silent, heterochromatic or euchromatic, in linker DNA or at nucleosomal DNA, etc. Chromatin has multiple levels of organization and can be modified in numerous ways (van Driel et al., 2003; Vermaak et al., 2003). Important for NER in this respect is to appreciate that not only does it need to detect DNA damage and restore the genetic information, but that the epigenetic information has to be maintained as well. Clearly, this will have repercussions for the action of DNA repair in vivo, which we are only beginning to understand. In general, the repair efficiency of NER is considerably decreased on nucleosomal DNA. Seemingly, the wrapping of the DNA around the histone octamer sterically hampers damage recognition and/or processing events. The nucleosome itself is a heterogeneous entity that can constitute a range of distinct histone variants. Moreover, nucleosomes can be altered in various ways, including numerous post-translational histone modifications, ATP-dependent chromatin remodelling, and interaction with histone chaperones. These changes can affect the accessibility of lesions that are part of nucleosomal DNA. Many relationships between chromatin structure and DNA accessibility have been positively established for a variety of DNA-interacting events, including different types of DNA repair (for a detailed evaluation, see for instance (Green and Almouzni, 2002; Meijer and Smerdon, 1999; Thoma, 1999). Here we will highlight a few noteworthy links between NER and chromatin structure.

NER appears to harbour its own chromatin remodelling activity. The TC-NER specific CSB protein has been identified as a DNA-dependent ATPase of the SWI2/SNF2 family.
and appeared to be able perform nucleosome remodelling in vitro (Citterio et al., 2000). Since CSB may play a non-essential role in transcription elongation as well (Selby and Sancar, 1997), it remains to be solved whether its remodelling activity is directly supporting its repair function. Furthermore, the GG-NER specific DDB2 protein has been found to associate with two different families of histone acetyltransferases, the STAGA (Martinez et al., 2001) and CPB complex (Datta et al., 2001). Both histone acetyltransferases and ATP-dependent chromatin remodelling factors can enhance the accessibility of nucleosomal DNA (Peterson and Logie, 2000). This makes it tempting to consider that NER proteins could be directly involved in rendering accessibility to DNA damage by the modification of chromatin structure, either through themselves (CSB) or by recruiting the appropriate machinery (DDB). In agreement with this suggestion is the observation that the DDB complex was found to associate with UV damaged chromatin (Groisman et al., 2003).

As mentioned, it will be important to conserve epigenetic information that is stored on nucleosomes. If the repair process disrupts chromatin in any way, it needs to be restored at some point. NER might be directly connected to such chromatin reinstatement, as chromatin assembly factor 1 (CAF-1) was shown to promote chromatin assembly upon repair (Gaillard et al., 1996). This was recently supported in vivo, where CAF-1 associated with sites of UV-induced DNA damage, and this binding was fully depending on functional NER (Green and Almouzni, 2003). Thus, following DNA repair synthesis, chromatin restoration is likely to be the last event of a full NER cycle.

A logical aim for NER research would be to come to a comprehensive kinetic model of the NER process. In the end, one would like to understand mammalian NER where it actually takes place: on a chromatin template in living cells. This requires detailed quantitative information of the dynamic behaviour of all players. Although this may seem fanciful, using a combination of enhanced genetic and biochemical tools, the availability of fluorescence-tagged proteins, and powerful modern imaging techniques and computing, this should be achievable in the near future. The work described in this thesis offers a proof of principle for such an integrated approach, providing novel mechanistic insights (Chapters 2, 3 and 4), detailed dynamic behaviour (Chapter 4) and in-depth kinetic characteristics (Chapter 5) for the intricate NER pathway.

6.5 References


