New factors in nucleotide excision repair : a study in saccharomyces cerevisiae

Dulk, B. den

Chapter

1

General introduction and scope of the thesis

General introduction and scope of the thesis

1.1 Instability of DNA

Scientists have long been puzzled by the remarkable stability of genetic traits. At the beginning of the 20th century it already was established that the hereditary units are harbored in the chromosomes. Which component of the chromosome would carry the actual genetic information was not known at that time, in fact, it could not be imagined that any molecule would be stable enough to preserve the stability exhibited by genetic traits. It would take several more decades before it dawned that DNA is the genetic carrier (Avery *et al.*, 1944), and its stability is the result of maintance by a number of different repair mechanisms.

The DNA molecule itself is certainly not stable. In the oxidative environment within the cell DNA is altered by various chemical reactions, such as hydrolysis, oxidation and base deamination (Lindahl, 1993). In addition to these endogenous threats numerous exogenous agents can potentially alter the structure of our DNA. DNA lesions interfere with essential cellular processes such as transcription and replication and can lead to cellular malfunctioning or cell death. When damages persist in the form of mutations, caused by the erroneous replication of damaged DNA, they can lead to defects such as tissue degeneration, ageing and cancer.

A major source of exogenously induced lesions in DNA is the sun. The emitted ultraviolet radiation can cause various aberrations to DNA. The most frequently occurring types are the *cis-syn* cyclobutane pyrimidine dimer (CPD) and the pyrimidine (6-4) pyrimidone photoproduct ([6-4-]PP) (Mitchell and Nairn, 1989; Pfeifer, 1997; Sage, 1993). Since the sun, and numerous other sources of DNA damage, have been present since (and essential for) the beginning of life on earth, mechanisms evolved, now known as DNA repair pathways, that protect the structural integrity of the DNA. With hindsight, the answer to the mysterious stability of genetic traits is simple; the carrier is not extraordinary stable but the information is preserved by dedicated maintenance of the carrier.

1.2 DNA repair mechanisms

As there are many different types of DNA lesions, several different kinds of DNA repair mechanisms exist. In general, repair of DNA comes in four varieties. (1) Chemical alterations can be directly reversed by photolyases or methylguanine DNA methyltransferases. (2) The ends of double strand breaks can be resealed by non-homologous end joining. (3) Double strand breaks can also be resolved via recombination with a homologous region within the same cell. (4) The damaged base can be excised, after which the DNA structure is restored by DNA synthesis using the undamaged strand as a template. Some of the key repair mechanisms are discussed below.

1.2.1 Direct reversal

Several proteins were identified that possess the ability to bind damaged nucleotides and reverse the modified nucleotide to its original state. A well known example is CPD-photolyase. This flavoprotein contains two chromophore-cofactors. The chromophore at the surface of the protein enables the protein to use energy from near-UV/blue light

as energy source (Mees *et al.*, 2004). This energy is utilized via the second chromophore in order to split the cyclobutane ring, thereby restoring the bases to their undamaged state (Sancar, 2004).

Direct reversal by photolyases is a very efficient way to remove lesions from the DNA. However, the substrate specificity of photolyases is limited to one type of injury. At present, three types of photolyases are known: CPD photolyase, (6-4)PP photolyase and cryptochrome. Whereas the CPD and (6-4)PP photolyases are clearly evolved as DNA repair factors with the sole purpose of removing CPDs or (6-4)PPs from the genome respectively, the cryptochrome proteins are not involved in DNA repair but utilize the same light harvesting mechanism to control the circadian clock and regulate growth and development in animals (Lin and Todo, 2005). Bacteria also possess this enzyme for a yet unknown purpose. Remarkably, CPD and (6-4)PP photolyase are not conserved in placental mammals.

A different type of direct reversal is employed by methylguanine DNA methyltransferase (MGMT), an enzyme that repairs methylguanines that are frequently formed by alkylating agents. MGMT transfers the methyl group from the guanine to an internal cysteine residue. An MGMT enzyme can only be used once, as the methyl group is stably attached to the cysteine, disabling the enzyme for further repair activities.

For the majority of DNA injuries a direct reversal solution is not available and repair of these lesions rely on other, generally more complex, DNA repair mechanisms.

1.2.2 Double strand break repair

Double strand breaks (DSBs) are formed frequently during cellular processes like mitotic recombination, V(D)J recombination and, in yeast, during mating type switching. Double strand breaks can also be induced by exogenous sources, such as ionizing irradiation and cytotoxins like bleomycin. DSBs are obviously hazardous to the genetic integrity and can lead to a wide range of genetic alterations including loss of heterozygosity, translocations, deletions and even chromosome loss (Jackson, 2002). DSBs are dealt with by DSB repair, which is a collective term for two different mechanisms that mend the broken DNA molecule.

Firstly, the sub-pathway responsible for the repair of DSB in the absence of a homologous donor is termed Non Homologous End Joining (NHEJ), a system that directly joins the disconnected DNA ends by ligation. In yeast, the Ku70/Ku80 and MRX complexes stabilize the ends of the DSB, after which the DNA is sealed by DNA ligase (Lewis and Resnick, 2000). The simplest mode of NHEJ involves DSBs with complementary overhangs including 5' phosphate and 3' hydroxyl groups, which can be re-ligated error free. Yet, the sealing of most breaks requires processing of the loose ends prior to ligation, resulting in deletions or insertions of basepairs. NHEJ is therefore associated with error prone repair of breaks. Despite the error proneness, the NHEJ pathway contributes significantly to the genome stability and suppression of tumorgenesis (Ferguson *et al.*, 2000; Karanjawala *et al.*, 1999).

In the presence of a homologous donor sequence within the same cell, a DSB can be restored via a second sub-pathway, Homologous Recombination (HR). This is a complex procedure, requiring a set of genes in the *RAD52* epistasis group. Repair is established by DNA synthesis using the homologous sequence as template. After the

induction of a DSB, the ends of the DSB are resected 5' to 3'. Once a homologous sequence is detected by means of the Rad51-ssDNA nucleoprotein filament, strand invasion of the 3' single strand tails with a homologous DNA molecule, allowing DNA synthesis using the 3' tail as a priming sequence. The D loop, formed as a consequence of the strand invasion is able to pair with the other side of the DSB resulting in a double Holliday junction. The non-invading strand can now be extended and subsequent filling of the gaps, ligation and resolution of the holliday junction re-establishes the double stranded DNA (Heyer, 2004; Krogh and Symington, 2004). The two DSB-repair systems share the same substrates but the relative activity of the two pathways varies between organisms, cell type and cell stage (Shrivastav *et al.*, 2008).

1.2.3 Nucleotide excision repair

Substrate versatility is a hallmark of the NER system, as it recognizes and removes many different lesions that are mainly generated by exogenous sources. NER substrates include UV induced CPDs and (6-4)PPs, intrastrand crosslinks and various bulky DNA adducts. The *in vitro* reconstituted NER reaction requires at least 16 proteins, each performing a specific step in the reaction leading to the removal of the lesion. The damaged DNA is identified by the NER damage sensors, after which a region of DNA surrounding the lesion is unwound to create a single strand bubble of ~30nt. At the junctions of this bubble, single strand incisions are made and the oligonucleotide containing the lesion is removed. The resulting single stranded gap is then filled by DNA polymerase and sealed by DNA ligase. Given the broad range of substrates it is assumed that NER senses a common feature in the damaged DNA. The NER mechanism is the focus of this thesis and will be discussed further in the following chapters. The question how such a diversity of chemically unrelated lesions is recognized by NER is addressed in chapter 3.

1.2.4 Base excision repair

The base excision repair (BER) pathway deals with the majority of base modifications, inappropriate bases and base losses which are endogenously formed with a high frequency (Holmquist, 1998). Substrates for the BER system are numerous and include the apurinic/apyrimidinic (AP) sites (Boiteux and Guillet, 2004) and the 7,8-dihydro-8-oxoguanine (8-oxoguanine) sites (Fortini *et al.*, 2003), which are both the result from injury to DNA via reactive oxygen species. In contrast to NER, BER does not employ the same proteins for each type of substrate. In fact, the BER pathway refers to a large collection of individually operating glycosylases, each capable of removing only one or a few different types of lesions. The glycosylases remove the damaged base by hydrolysis of the N-glycosylic bond that links the base to the deoxyribose-phosphate backbone. The phosphate backbone of the remaining apurinic/apyrimidinic site is then incised by an AP-endonuclease (Barzilay and Hickson, 1995) and DNA polymerase and DNA ligase subsequently complete the restoration of the DNA.

1.3 DNA damage tolerance

Some DNA lesions will inevitably escape detection by the various damage surveillance proteins and persist into the S phase. Additionally, DNA damage will also be induced during the replication itself. Without assistance, the replication fork can not proceed through damaged DNA and will arrest at the site of the lesion, posing a threat to the viability of the cell. In this case the cell diverts to an alternative means to cope with these lesions. Several mechanisms, collectively referred to as 'DNA damage tolerance', have evolved to resolve the arrested replication machinery on the DNA, some at the cost of inducing mutations. These pathways are also known as 'Post Replication Repair', which is not entirely accurate as the lesion is not removed, but rather bypassed.

Post replication recombination repair involves homologous recombination (HR) using the undamaged sister chromatid as template. This system might be of especial value to solve specific mishaps that can occur following the collision between replication fork and lesion (Li and Heyer, 2008). When a lesion blocks the DNA polymerase but not the helicase unit, the helicase will generate an excess of single stranded DNA, which can result in a DSB after endonucleolytic activity. DSBs are also generated when single stranded DNA breaks induced by reactive oxygen species are encountered by replicating DNA polymerases (Li and Heyer, 2008). In these situations the replication machinery will be displaced to allow repair via HR. Once the generated DSB is resolved DNA synthesis can resume past the site of the lesion. HR is also responsible for filling of the gaps generated in daughter strands opposite base damage (Morimatsu and Kowalczykowski, 2003; Sogo *et al.*, 2002).

Replication fork regression is an alternative mechanism to circumvent the damaged template. Here synthesis switches from the damaged template to the newly synthesized daughter strand. The mechanism is yet poorly understood, however, recent evidence supports a role of Rad5, the caretaker of the error-free branch of damage tolerance, in Replication Fork Regression in yeast (Blastyak *et al.*, 2007).

The best characterized form of replication associated damage tolerance is Translesion Synthesis (TLS). The TLS system consists of various alternative DNA polymerases, all characterized by a compromised fidelity due to a larger active site and/or the absence of exonucleolytic proofreading (For review, see Friedberg (2005)). These polymerases are thereby able to incorporate bases opposite templates containing damaged nucleotides that do not meet the requirement of recognition by replicative DNA polymerases α, δ or ε. TLS is therefore error prone, however, a certain degree of fidelity is realized in this pathway by deploying specific DNA polymerase for specific types of lesions. For example, thymine dimers can not be processed by DNA polymerase a, δ or ε, but DNA polymerase ŋ can effectively pass through these lesions, mainly by correctly inserting two Adenine's opposite the dimer (Washington *et al.*, 2001). Specific usage of DNA polymerase to resolve stalled replication at a CPD linked thymine lesion hence is error free in the great majority of cases. The precise mechanism behind when and how TLS polymerases are recruited is far from fully understood. It is shown however that the proliferating nuclear antigen (PCNA) sliding clamp, which acts as a processivity factor for replicative DNA polymerases, is a key player in the TLS pathway, as it is responsible for the recruitment of certain TLS polymerases (Hoege *et al.*, 2002).

It is not known in detail what feature of the stalled replication fork determines which

of the above described damage tolerance systems is applied. Different post translational modifications of PCNA were shown to act as molecular switches that determine whether the lesion stalling the replication fork will be bypassed via TLS or repaired via post replication repair (Haracska *et al.*, 2004; Hoege *et al.*, 2002; Ulrich *et al.*, 2005; Watts, 2006).

1.4 The scope of this thesis

The research described in this thesis focuses on the role of the Rad4-Rad23 complex, an essential factor in damage recognition of eukaryotic Nucleotide Excision Repair (NER).

After a general introduction in chapter 1, chapter 2 introduces the mechanism of the basic NER reaction and discusses the two sub-pathways of NER, Global Genome Repair (GGR) and Transcription Coupled Repair (TCR).

In chapter 3 an introduction is given on one of the most intriguing aspects of the NER system: its ability to detect many different types of lesions within a huge number of undamaged bases. A possible model of eukaryotic damage recognition is presented, based on the prokaryotic system in which damage recognition is elucidated in considerable detail.

Chapter 4 describes the identification of a homologue of Rad4 in *Saccharomyces cerevisiae,* Rad34. This protein is specifically involved in NER in the relatively small rDNA region. Like their human homologue XPC, Rad4 and Rad34 form a complex with Rad23. Interestingly, in yeast both Rad4 and Rad34 also bind to another small (20kDa) protein that we have identified as a new NER factor, designated Rad33. A study of the NER defect of *rad33* cells is presented in Chapter 5.

Chapter 6 discusses possible analogous roles of Centrin2 in human cells and Rad33 in yeast cells. Although the proteins do not share clear sequence homology, the predicted structures of Rad33 shows resemblance with that of Centrin2. Furthermore, we show that Centrin2 and Rad33 interact with XPC and Rad4, respectively, via the same conserved motif.

Chapter 7 contains a summary of the presented work in this thesis and concluding remarks.