CHIMIA 49 (1995) Nr. 10 (Oktober)

*Chimia 49 (1995) 365–373*

*ISSN 0009–4293*

# Sunlight-Induced DNA Lesions. Lesion Structure, Mutation Characteristics and Repair

**Thomas Carell***

*Abstract.* The UV part of sunlight is known to induce a variety of genome defects. These lesions are the major cause of skin cancer development. In order to counter such toxic effects cells have developed a number of sophisticated DNA repair systems, like nucleotide excision repair and photoreactivation. The repair machinery is able to specifically recognize sunlight-induced DNA lesions and to subsequently remove this damage. Malfunctioning repair systems are responsible for the three rare genetic diseases *Xeroderma pigmentosum*, *Cockayne's syndrome*, and *trichothiodystrophy*. In this review article, the structure of the major sunlight-induced lesions will be discussed. An overview of the two major repair mechanisms, photoreactivation and excision repair, is given, and the effects of the DNA lesions on the p53 gene and on tumor genesis are discussed.

## 1. Introduction

The construction and function plan for any cell is stored inside the cell nuclei in form of double helical DNA. The ability to recall and translate the information from the DNA molecule whenever necessary is the basic requirement for the existence of life as we know it [1]. Since DNA and RNA are macromolecules possessing only limited stability, decomposition reactions endanger the proper function and survival of every cell, which further threatens the whole organism. To counter these decomposition effects, cells possess efficient genome repair systems to restore partially degraded DNA and thus maintain the genome.

### 1.1. *Instability and Decay of DNA* [1]

A number of decomposition pathways of nucleic acids (DNA and RNA) which continuously damage the genome have been investigated well. UV Irradiation, an ubiquitous damaging agent for species which live exposed to sunlight, causes predominantly the formation of pyrimidine dimers within the DNA double strand [2][3]. DNA hydrolysis of the *N*-glycosidic bond between the sugar-phosphate backbone and the nucleobases causes depurination and to a lesser extent depyrimidation of DNA [4]. This results in the formation of apurinic and apyrimidinic sites within the double helix. It is estimated that *ca.* 2000–10000 depurinations occur every day in the human genome [1]. Furthermore, cytosines and 5-methylcytosines are also vulnerable to hydrolytic deamination which causes the transformation of cytosine (C) into uracil (U) and 5-methylcytosine (5methylC) into thymine (T) [5–7]. Approximately 100–500 C to U transformations are estimated to take place per day in the human genome [1]. C to U and 5-methylC to T transformations are thus widespread genome changes. Other major decomposition reaction are oxidative alteration of the nucleobases in metabolically active and aerobically living cells. Here, the major alterations are the formation of 8-hydroxyguanines, which base pairs preferentially with adenine rather than cytosine [8][9]. Additionally, the formation of thymine and cytosine glycols is observed [10][11]. Especially precarious to spontaneous decomposition is the macromolecule RNA due to the presence of the additional 2'-hydroxyl group. Attack by the 2'-hydroxyl group at the phosphodiester of the backbone causes strand breaks and the liberation of cyclic phosphates. This reaction is further accelerated in the presence of divalent cations like $Mg^{2+}$ or $Ca^{2+}$. Although RNA can serve as a short term carrier for genetic information, its limited stability and, therefore, susceptibility to mutation seems to limit the use of RNA as a molecule for the permanent storage of genetic information. This instability is an argument against the proposal of an 'RNA world' – where RNA is predicted to be the primary carrier of genetic information – in connection with the origin of live [12].

It seems reasonable to assume that with the evolution of life based on DNA as the information carrier, efficient DNA repair mechanisms had to be developed in order to counter toxic spontaneous decomposition effects and damaging environmental influences. Today every living cell possesses efficient repair systems which recognize lesions inside the genome and are subsequently able to repair the damage [13]. During the past four years many exciting research results have shifted DNA repair into the limelight [14]. It now appears well established that inefficient genome repair capabilities of cells are major causes for a variety of cancers [15]. Because a detailed analysis of all repair processes is beyond the scope of this review article, only the lesions induced upon UV irradiation of DNA together with those repair systems developed to remove these lesions (nucleotide excision repair and photoreactivation) will be discussed.

## 2. Skin Tumors and Sunlight-Induced DNA Damage [2][16]

Sunlight-induced DNA lesions are directly involved in the development of a variety of skin cancers like basal cell carcinomas, squamous cell carcinomas and also malignant melanomas [16]. A direct connection between exposure to UV radiation and cancer risk is well established for basal cell and squamous cell carcinomas [17][19]. The available evidence further suggests that the occurrence of the malignant melanoma is also directly linked to the encountered UV dose [2]. Skin tumors are the most prevalent form of cancer and the number of cases have been rising steadily over the past 20 years [17]. It is expected that by the year 2000 nearly 1% of all Americans will be effected by the most serious form of skin cancer: the malignant melanoma. Changing lifestyles are the predominant cause of this dramatic increase. More of the increasingly available free time is spent outside, exposed to UV radiation. No direct link between the

**Correspondence*: Dr. T. Carell
Laboratorium für Organische Chemie
ETH-Zentrum
Universitätstrasse 16
CH–8092 Zürich

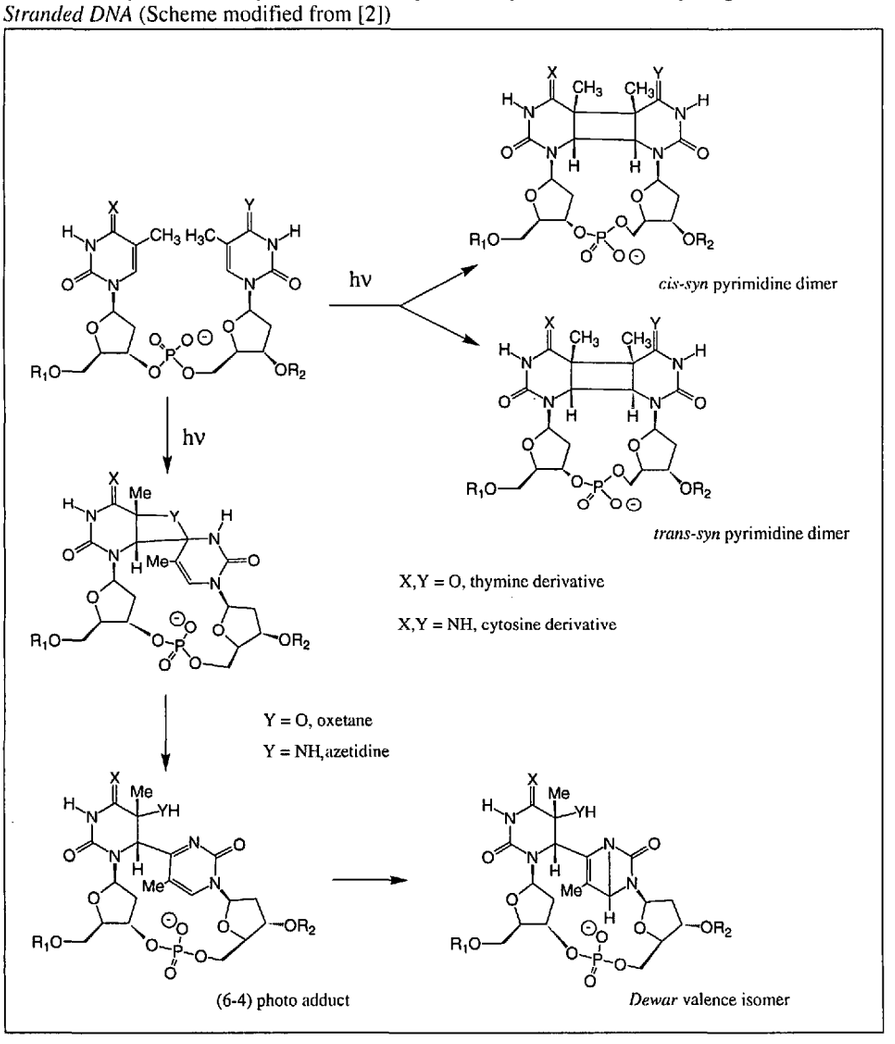
CHIMIA 49 (1995) Nr. 10 (Oktober)

rising incidence of skin cancer and the depletion of the ozone layer has been found yet [19]. Nevertheless, those parts of the world which are directly affected by ozone depletion report a soaring skin cancer risk [18][19]. In addition, the finding that the world wide decline of certain amphibian populations can be linked to insufficient repair capabilities of UV induced lesions, indicate that environmental changes might become more important in future increases in human skin cancer incidence [20].

Patients suffering from a rare genetic disease, *Xeroderma pigmentosum* lack the ability to repair sunlight induced DNA lesions (see below) [21]. *Xeroderma pigmentosum* cells are, therefore, excellent subjects for studying the effect of UV irradiation on the human genome. In fact, intensive investigation of *Xeroderma pigmentosum* skin tumor cells show that a variety of mutations in one of the major tumor suppresser genes, the p53 gene, are caused by UV radiation [22]. Defects in p53 are thought to be involved in *ca.* 50% of all human cancers. As the result, *Xeroderma pigmentosum* patients show hypersensibility toward light and they generally develop skin tumors before the age of 8. In addition *Xeroderma pigmentosum* patients show neurological abnormalities.

In *Scheme 1* the major photoproducts of irradiation of double-stranded DNA are depicted. In DNA regions where two pyrimidines, such as two thymines, are located directly above each other, a cyclobutane pyrimidine dimer is formed. In double-stranded DNA mainly the *cis-syn* photodimers, depicted in *Scheme 1*, are generated. In single stranded DNA, *e.g.* during transcription or replication, a second pyrimidine photodimer isomer, the *trans-syn* configurated species, is created in up to 10% of the total photodimer content. The *cis-syn* thymine photodimer represents the major lesion induced upon UV irradiation of DNA. The basis for the dimer formation is a photochemically allowed $[2\pi_s+2\pi_s]$ cycloaddition of the C(5)=C(6) bonds of the pyrimidines which are located above each other in double-helical DNA [2]. The photodimers form stable species inside the double helix and are thought to induce cell death through blocking of replication and translation. Whereas the thymine photodimers are removed from the genome by sophisticated DNA repair mechanisms (see below), the cytosine containing photo dimers are considered to be highly pro-mutagenic. Even though they are formed to a lesser extent, cytosine residues within a photodimer are prone to a variety of subsequent mutagenic reactions like deamination which causes ultimately C to T transformations (see below). Additional photoproducts which are formed through irradiation of DNA are the pyrimidine-pyrimidone (6–4) photoadducts (see *Scheme 1*). They are generated through reaction of a pyrimidine C(5)=C(6) bond with the C(4) carbonyl group of a second pyrimidine base. This reaction gives at first oxetanes (Y = O) and azetidines (Y = NH) which isomerise rapidly above –80° to the (6–4) photoadducts which have a covalent bond between position 6 and 4 of two formerly adjacent pyrimidines. Recently, it was found that the (6–4) adducts isomerise upon irradiation with visible light to their *Dewar* valence isomers inside the double strand [23][24].

Scheme 1. *Representation of the Main Photoproducts of UV Irradiation of Single- and Double-Stranded DNA* (Scheme modified from [2])




## 3. Mutation Mechanisms. Two Selected *in vivo* and *in vitro* Experiments

It is of major importance to clarify how gene alterations like the formation of pyrimidine photodimers contribute to tumor development. A better understanding of the mutation mechanisms and the corresponding cell effects would allow the determination of individual risk factors and may lead to the development of preventative treatments. Ultimately, this knowledge can be the basis for the development of compounds with lesion directed artificial repair capabilities. As a consequence, DNA repair is today one of the very hot topics in biochemical research. Unfortunately, the assignment of alterations of cancer cell proteins to certain genome lesions is complicated and very often not possible. A major obstacle derives from the observation that the same DNA damage in different sequences can have very different effects. The bypass of a gene lesion during replication and the insertion of a wrong base opposite the DNA damage is the major mutagenic event (error-prone replication). On the other hand, repair of the lesion prior to the replication removes the damage. Both processes compete with each other. The sequence in

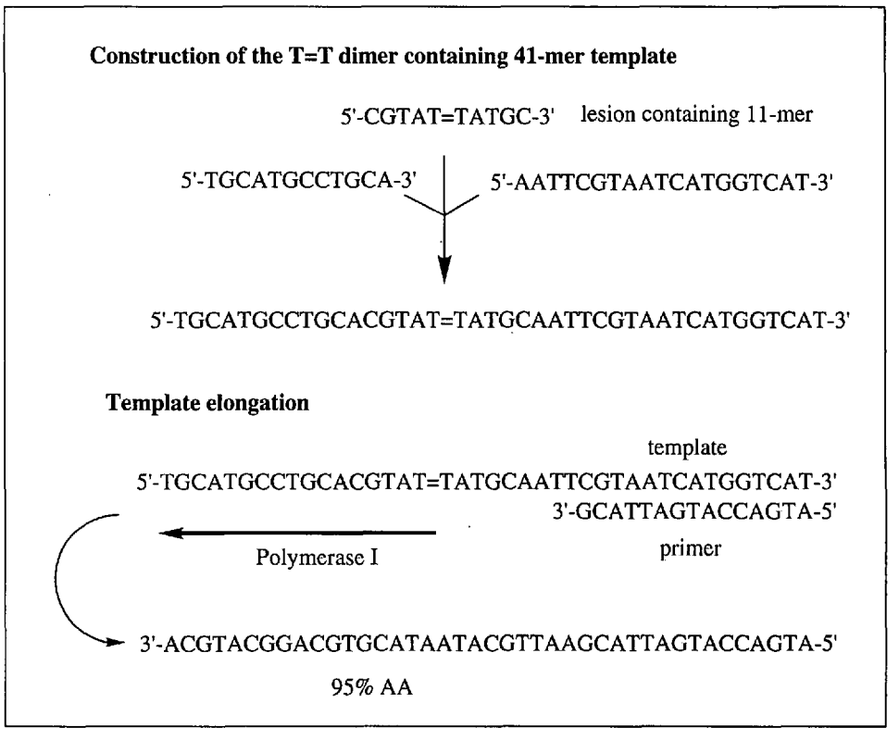
CHIMIA 49 (1995) Nr. 10 (Oktober)

which the lesion is embedded seems to modulate the different process rates. Consequently, the occurrence of a mutation is difficult to assign to a specific photoproduct [25]. Furthermore, due to the efficient repair of most major sunlight induced lesions, minor photoproducts, which might be unknown to date, can account for a significant proportion of the observed mutagenic effects in the cell.

In order to study carcinogenic effects of sunlight induced lesions, *in vitro* and *in vivo* studies with modified oligomers containing a defined site specific photo lesion are required. The modified DNA's are then used to investigate how replication and transcription are affected by the presence of the gene lesion. For the *in vitro* experiments, small DNA oligomers which contain a specific photoproduct are prepared. The photoproducts are introducable into those oligonucleotide in two ways: *a*) Through direct irradiation of the oligomer with UV light and subsequent separation of the oligomer which contains the site specific photolesion through HPLC. This method allows the preparation of sub milligram amounts of DNA containing a desired photolesion [26][29]; *b*) A second route involves the synthesis of a photolesion building block, which can then be introduced into the sequence through solid phase phosphoramidite synthesis [30] [32]. This approach has the advantage that multi milligram amounts of a lesion containing oligomer are available in very high purity [33][34]. Unfortunatly, not every photoproduct can be synthesized as its phosphoramidite building block yet. For the *in vitro* study, which is representatively described in this article two other oligonucleotides which contain 12 and 19 base pairs are ligated to the 3'- and the 5'-terminus of the lesion containing oligomer [34]. As depicted schematically in *Scheme 2*, the resulting 41-mer contains a site specific photodimer in the middle of the sequence. This oligomer is subsequently subjected to replication assay experiments designed to clarify if the replication machinery is able to copy the damaged oligomer. Through sequence analysis of the DNA copies one can determine which bases the replication machinery inserted opposite the photolesion. This information is important in order to elucidate which mutations are likely to occur when a damaged strand is copied inside the cell. In this experiment, the 41-mer acts as a template for the elongation of a 15 basepair long oligomer (primer) complementary to the 3'-terminus of the template. The replication enzyme studied was polymerase I. This enzyme is responsible for the filling of single-stranded gap regions in double helical DNA. These single stranded gaps occur during repair when a damaged single-stranded region within the DNA double helix (see below: excision repair) is removed. The experiment revealed that a single thymine dimer represents a major obstacle for replication. Nevertheless, under high polymerase concentration conditions, replication of the damaged template could be enforced. Sequence analysis of the DNA copies showed that in over 95% the T=T was correctly translated into AA. This result indicated that even though simple thymine photodimers are replication obstacles, they are only weakly pro-mutagenic DNA lesions.

*In vivo* studies were conducted to study if a photolesion inside a parent DNA (viral DNA) induces mutations in the genome of descendants. A 11-mer containing a photolesion was produced and inserted into single-stranded virus DNA (plasmid). As depicted in *Scheme 3*, the preparation of the parent virus DNA is a multistep process. In the representative experiment used to clarify the employed techniques, a single-stranded plasmid is linearized in the first step [28]. The second step contains the recirculation of the plasmid using a 51-mer template which possess complementarity to both ends of the single-stranded plasmid. A gap between the two plasmid ends remains. Complementary to this gap region is the photo damaged 11-mer which is inserted into the gap region in a third construction step. Finally, the 11-mer containing the site specific lesion becomes ligated into the virus DNA (pic. D). The *in vivo* experiment involves the infection of DNA repair deficient *Escherichia coli* bacteria. As the result, the *Escherichia coli* become programmed to replicate the virus DNA but they lack the ability to repair the lesion. Through the analysis of the number of progeny plagues, the replication fidelity of the virus DNA inside the *Escherichia coli* can be deduced. Analysis of the progeny virus DNA sequences show which bases the *Escherichia coli* replication apparatus inserted opposite the lesion site. The *in vivo* studies revealed, that a single *cis-syn* thymine dimer inhibited the replication fidelity of the virus DNA nearly completely. As the result of a single thymine dimer inside the virus genome, the virus DNA lost their infectious property. In control experiments where the damaged virus DNA was subjected to an *in vitro* DNA repair reaction prior to the *Escherichia coli* infection, the awaited number of progeny plagues were observed. This result further established that the inability of the *Escherichia coli* replication apparatus to copy the viral DNA was due to the presence of the single thymine photodimer. *Escherichia coli* bacteria possess the ability to replicate even damaged DNA sequences if they contain certain proteins which are expressed during an

Scheme 2. *Schematic Representation of the Construction of a 41-mer Containing a Site-Specific* cis-syn *Cyclobutane Thymine Dimer* (modified from [34]). T=T represents the *cis-syn* cyclobutane thymine photo dimer.



CHIMIA 49 (1995) Nr. 10 (Oktober)

Scheme 3. *Schematic Representation of the Construction of a Single-Stranded Vector Containing a Site-Specific* cis-syn *Thymine Photodimer for* in vivo *Experiments.* ss = Single-Stranded. A: Linearisation with the Enzyme Eco RI, Annealing with the 51-mer Scaffold. B: Treatment of the Plasmid with Excess 11-mer Containing the Photolesion. C: Ligation of the 11-mer into the Plasmid and Denaturing of the Plasmid to Remove the 51-mer Scaffold.

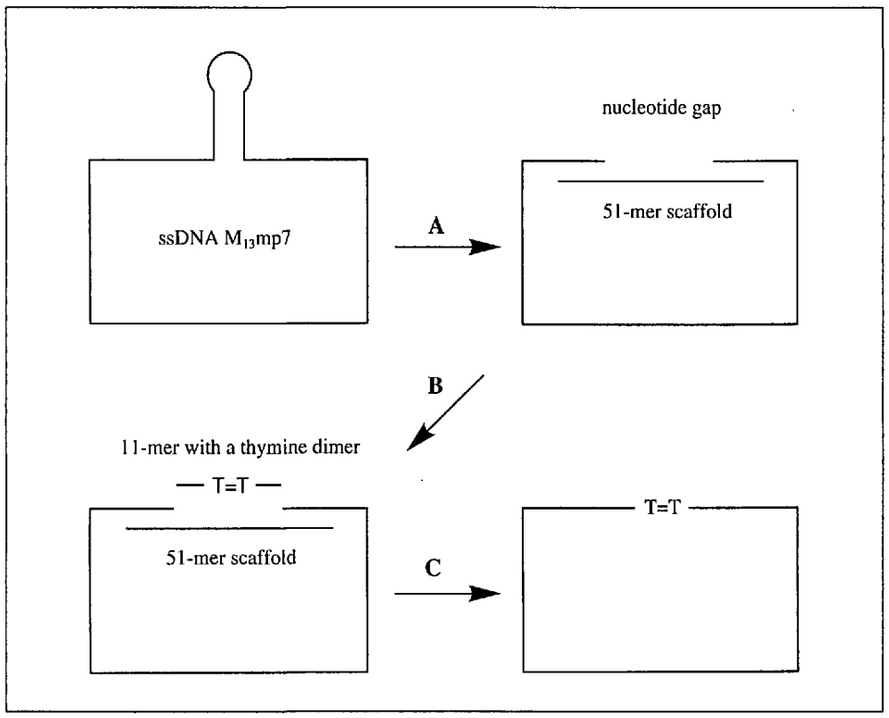



Table. *Percentages and Types of Mutations in Human Skin Cancers* (Data from [22])

| | Frequency of p53 gene mutations [%] | Mutations at dipyrimidine sites [%] | | | Mutations at non-dipyrimidines sites |
|---|---|---|---|---|---|
| | | C to T | CC to TT | Other | [%] |
| Human SCC | 45 | 29 | 11 | 32 | 29 |
| Human VCC | 49 | 38 | 12 | 16 | 34 |
| Humans TX tumors | 42 | 21 | 57 | 21 | 0 |

SOS response [35]. These factors enable the polymerase to bypass lesions. Needless to say, bypassing a lesion introduces replication errors (error-prone replication). Additional experiments, using the thymine photodimer containing plasmid and *Escherichia coli* subjected to strong UV irradiation prior to the infection event – in order to induce the SOS response – were conducted in order to elucidate which bases are inserted during the copy process of the damaged virus DNA under SOS conditions. The experimental results showed, that *Escherichia coli* which contain the SOS response factors are indeed able to replicate the virus DNA. Thymine photodimers were no longer complete blocks of the replication machinery. These results are in agreement with the *in vitro* studies which showed that a thymine photodimer can be bypassed during replication [34]. Analysis of the progeny plagues obtained in the *in vivo* experiments revealed again the astonishing result that in more than 93% of all descendants, the DNA copies of the damaged virus genome contained the right base sequence [28]. Therefore, the translesion replication results in the correct translation of T=T into AA. Thymine photo dimers were again found to be only weakly pro-mutagenic lesions inside the genome. The explanation for this result is based on two arguments. Firstly, in the case of a lesion bypass, polymerases like polymerase III preferentially introduce adenine. This is called A-rule [36]. Secondly, the thymine photodimers are not non-instructional DNA lesions. They still posses to some extent the H-bonding capabilities of two neighboring thymines and therefore induce the incorporation of the correct two adenines in the daughter strand during replication bypass [37]. In fact, the high probability of finding two adenines correctly inserted opposite a thymine dimer cannot be explained with the A-rule alone.

## 4. Sunlight-Induced Alterations of the p53 Tumor Supressor Gene [22]

Another excellent tool for the analysis of how the UV irradiation of cells is linked to tumor development is the analysis of mutated p53 tumor suppresser genes from different cancer cells. p53 is of major importance for cells because it halts the cell cycle after a damage has been detected [38][39]. This gives the cell time to repair the lesion before the cell cycle continues. In addition the p53 protein seems to play an important role in apoptosis (programmed cell death) if the damage is too severe for an efficient repair [40–42]. New results indicate that p53 is, in addition, indirectly involved in the stimulation of the repair machinery [43]. The multitude of important functions of the p53 protein make defects in it a serious threat for the cell and explains why a mutated p53 gene is found in many cancer cells. Through comparison of mutated p53 genes from cell of various internal cancers and p53 genes from skin tumors developed by *Xeroderma pigmentosum* patients, it is possible to deduce indirectly, which modifications are caused through UV irradiation [22][44]. Such a comparison showed that in more then 90% of all mutated p53 genes in skin cancer cells from *Xeroderma pigmentosum* patients – which lack the ability to remove sunlight induced DNA lesions – the mutations occurred at dipyrimidine sites, which is a strong indication that pyrimidine dimers like T=T, C=T, or C=C are responsible for these mutations. In internal cancers, the mutation distribution is proportional to the frequency of bipyrimidine sequences in the p53 gene. More than 50% of the gene alteration found in *Xeroderma pigmentosum* skin tumor cells were determined to be CC to TT double base substitutions which further supports the idea that cytosine photodimers are the most pro-mutagenic UV induced lesions and therefore the major causes of skin cancer. In fact tandem CC to TT transitions are fingerprints for UV induced genome lesion. Tandem CC to TT transitions are also common (11%) in basal cell carcinomas (BCC) and squamous cell car-

CHIMIA 49 (1995) Nr. 10 (Oktober)

cinomas (SCC) developed by normal individuals not affected by *Xeroderma pigmentosum* (XP). In contrast, the CC to TT gene alteration is rarely observed (0.8%) in internal malignancies [22].

The *Table* summarizes the percentages and the types of detected p53 gene alterations in various skin tumors. The study involved the investigation of BCC, SCC and skin tumors developed by *Xeroderma pigmentosum* patients. In *ca.* 40–50% of the analyzed tumors, mutations within the p53 gene had occurred. In SCC and BCC developed by normal individuals 60–70% of these mutations are located at dipyrimidine sites. Even more dramatic is the finding that in skin tumors from *Xeroderma pigmentosum* patients nearly all p53 gene alterations involve a dipyrimidine sequence. In *Xeroderma pigmentosum* skin tumors approximatly 55% of the dipyrimidine alterations are CC to TT double base substitutions. These investigations demonstrate that sunlight induced DNA lesions which are responsible for tumorgenesis occur predominantly at dipyrimidine sites. In 39% (SCC), 50% (BCC), and 79% (XP) of all detected p53 gene mutations, the lesions involved a C to T transition. This makes cytosine within a dipyrimidine sequence a 'hotspot' for sunlight induced mutagenesis.

### 4.1. *C to T Mutation Mechanisms*

An explanation for the observed C to T mutations is depicted in *Fig. 1* and *Scheme 4*. Two major mutation mechanisms have been proposed [27]. Both mechanistic proposals use the fact that a cytosine within a photodimer contains a saturated C(5),C(6) bond. The first mechanism, proposed by *Bockrath* and coworkers [45][46] (tautomer bypass mechanism) is based on the observation that cytosines without aromatic stabilization exist to a higher degree as the (*E*)-imino tautomers (see *Fig. 1*) [47–49]. Due to the different H-bonding properties of the *E*-imino tautomer compared to the amino form of cytosine, the *E*-imino tautomer cause the incorporation of adenine and not guanine during replication. Spectroscopic investigation show that in solutions of cytosines with a saturated C(5),C(6) bond (5,6-dihydrocytosine) in water nearly 1% of all cytosines belong to the *E*-imino form. In chloroform all C(5),C(6) saturated cytosine derivatives (100%) exit as this alternative tautomer [48]. Even though it is unknown how many cytosines of a photodimer within the double helix exist in the mutagenic *E*-imino form, even a slight increase is thought to facilitate the incorporation of an adenine during replication opposite to a cytosine (*E*)-imino tautomer within a C=C photodimer (see *Scheme 4*). Subsequent DNA repair would remove the C=C photodimer and incorporate a thymine, complementary to the wrongly introduced adenine.

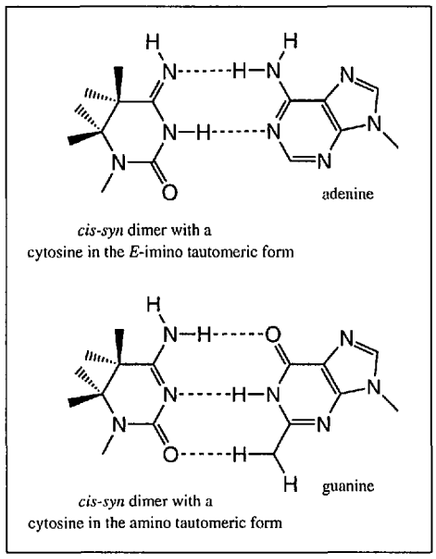
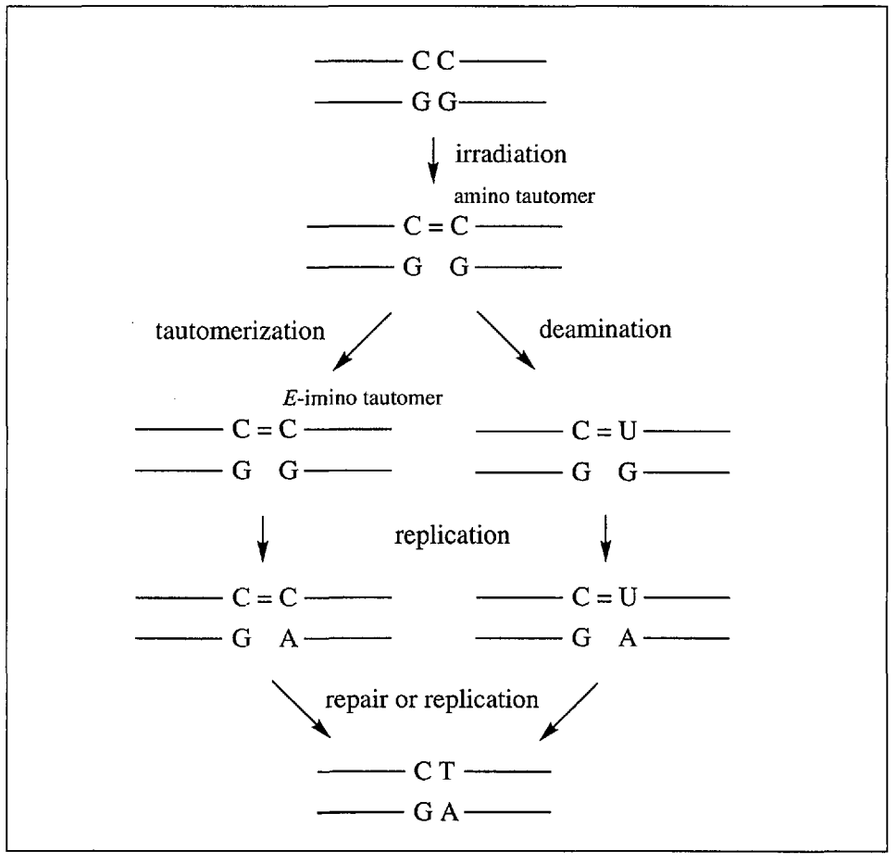



Fig. 1. *Schematic representation of the amino tautomer and (E)-imino tautomer of cytosine together with the prefered base pairing*

The second mechanism (deamination bypass mechanism, depicted in *Scheme 4*) is thought to play a major role in slow growing cells [2][27]. Based on the observation that cytosines with a saturated C(5),C(6) bond are vulnerable to deamination it is assumed that cytosines within a photodimer are able to deaminate before repair occurs [50–52]. It is estimated that a cytosine within a photodimer deaminates with a halflife of 12 h *in vivo* [51]. In comparison, the half live for unmodified cytosine within the DNA double helix is 30000 years [7]. Such a deamination converts the cytosine into a uracil which initiates the incorporation of adenine in the daughter strand during replication (for the removal of uracil from the genome see [53][54]). Again, subsequent repair removes the photodimer and inserts a thymine opposite the wrongly introduced adenine.

Sometimes repair reactions are very slow and the time lag between the occurrence of a photo lesion and its removal can be more than 10 h. This would be sufficient time for the deamination of a considerable amount of the cytosines within a cyclobutane pyrimidine dimer [1][2].

Scheme 4. *Schematic Presentation of the Deamination Bypass Mechanism and the Tautomer Bypass Mechanism which Are Both Responsible for C to T Transformations*



### 4.2. *A Short Summary of the Mutagenic Potential of Various Photo Dimers*

It is widely accepted that the UV-induced thymine dimers are only weakly pro-mutagenic. Thymine dimers block cell replication and transcription and therefore cause cell death, but during bypass they are correctly (in over 95%) translated into two adenines. More problematic regarding tumor genesis are photodimers which contain a cytosine. Due to the rapid tautomerisation and deamination of cytosines with a saturated C(5),C(6) bond, a cytosine within a dimer is translated incorrectly during a replication bypass into adenine which after repair of the photolesion results in the replacement of one or two of the cytosines within the photodimer by one or two thymines complementary to these wrongly introduced adenines. This results in C to T transitions and CC to TT double base substitutions which are the predominant gene alterations in mutated skin cancer p53 genes.

## 5. DNA Structure Alteration in the Vicinity of a Cyclobutane Pyrimidine Dimer

*cis-syn* Thymine photodimers are genetic lesions which are efficiently removed from the genome by either excision repair or photoreactivation. In both cases proteins which scan the DNA for thymine dimers recognize the lesions with high specificity. How the proteins recognize the damaged DNA strand is not currently understood. In addition, considerable disagreement exists about how strongly a single thymine dimer within the DNA double helix bends and unwinds the double strand. These structural distortions are thought to play a major role during damage recognition. Several different experiments have been performed in order to address the structure alteration question. The migration behavior during gel electrophoresis of DNA containing three site specific thymine dimers in phase with the helix screw axis where performed in order to obtain data about the overall change in the DNA structure [55]. Bending of DNA reduces the mobility of the DNA oligomer in the gel electropheresis experiments. Melting point measurements of oligomeric double-helical DNA were performed with DNA built up by $dA_{250}$ and $dT_n$ with $n$ varying between 4 and 10 [56]. Into the $dT_n$ oligomer a site specific pyrimidine photo dimer was introduced through photolysis of a thymine dimer (TpT) and subsequent elongation of the dimer segment. The melting temperatures were analyzed in order to estimate how many H-bonds are broken in the vicinity of a cyclobutane thymine dimer [56]. Molecular mechanics calculations [57][58] and NMR investigations [37][59] of double helical DNA containing a site specific photodimer were conducted to gain deeper insight into the structure alteration caused by a cyclobutane thymine photodimer. The NMR studies revealed that the double helix is only slightly distorted and the thymine dimer seems to be able to build weak H-bonds to the opposite adenines. A molecular mechanics calculation [58] and a newer gel electrophoresis study [60] of $dA_n \cdot dT_n$ sequences containing a thymine photodimer supports the notion that thymine photodimers cause only minor structural alterations. They are still instructional lesions with the correct H-bonding capabilities. This result would explain the low mutagenic potential of the thymine photodimer. In contrast, other molecular mechanics calculation [57] together with the melting point measurements indicate a much stronger distortion of the DNA [56]. A kink in the DNA of *ca.* 30° and an unwinding of 7° was determined with the DNA being bent into the major groove. The structural distortion of the DNA was estimated to be more pronounced at the 3'-end of the dimer. The melting temperature measurements revealed that the H-bonding of the cyclobutane thymine dimer and the two flanking nucleobases are completly disrupted. Unfortunately, up to now a detailed picture of the DNA structure around a photodimer has not been obtained. One reason could be that the structural alteration caused by the photodimer is modulated by the exact sequence of the DNA close to the dimer. Because the different investigations were performed with different DNA sequences a direct comparison of the research results might be difficult. As the result, the exact structure of the DNA double helix in the vicinity of a photodimer is unclear. However, all experiments established that the H-bonds between the two DNA single strands in the vicinity of the photodimer are significantly weakened.

Newest X-ray crystallographic investigations of repair enzymes indicate that not the structural distortion of the double helix, but the weakened H-bonds, might be responsible for the damage recognition by the repair machinery. Since the discovery that the HhaI methyltransferase flips its target base out of the DNA helix into an active site cavity, it is speculated that certain repair enzymes scan the DNA for areas with weakened H-bonds [61]. These areas are thought to become turned out of the double helix prior to the repair process. For two repair proteins, the uracil-DNA glycosylase and the DNA-photolyase such a binding mode was postulated based on their structures determined by X-ray crystallography [62][63]. In both cases, the structure of the binding cavity was found to possess the right geometric properties to accomodate the corresponding flipped out DNA-lesion.

## 6. Damage Repair

Two major repair pathways for the removal of photolesions inside the genome can be distinguished. The nucleotide excision repair process, which is an ubiquitous and a very complicated multistep repair reaction [64]. Special damage recognition proteins recognize a variety of different lesions, among them the cyclobutane containing pyrimidine dimers. A variety of other repair proteins are subsequently attracted to the damaged site to perform the repair. In man the excision repair process seems to be the only process capable of removing pyrimidine dimers from the genome. A deficiency in excision repair capabilities is responsible for three rare genetic diseases. *Xeroderma pigmentosum*, *Cockayne's syndrome* and *trichothiodystrophy*. Investigation of these three diseases recently revealed a more detailed understanding of how the excision repair process works. The picture obtained is still very vague and proposals about the exact mechanism are often based on preliminary experimental results. Nevertheless, an overall picture seems to emerge [64].

In contrast, photoreactivation is a relatively 'simple' process able to repair only cyclobutane pyrimidine dimers [65]. For the whole repair reaction including damage recognition, only a single monomeric protein (photolyase) is required. Photolyases are found in a variety of different species but are absent in others in a seemingly unpredictable manner. In plants however, photoreactivation seems to be the major repair reaction for the removal of UV induced lesions from the genome.

### 6.1. Escherichia coli *Excision Repair* [64]

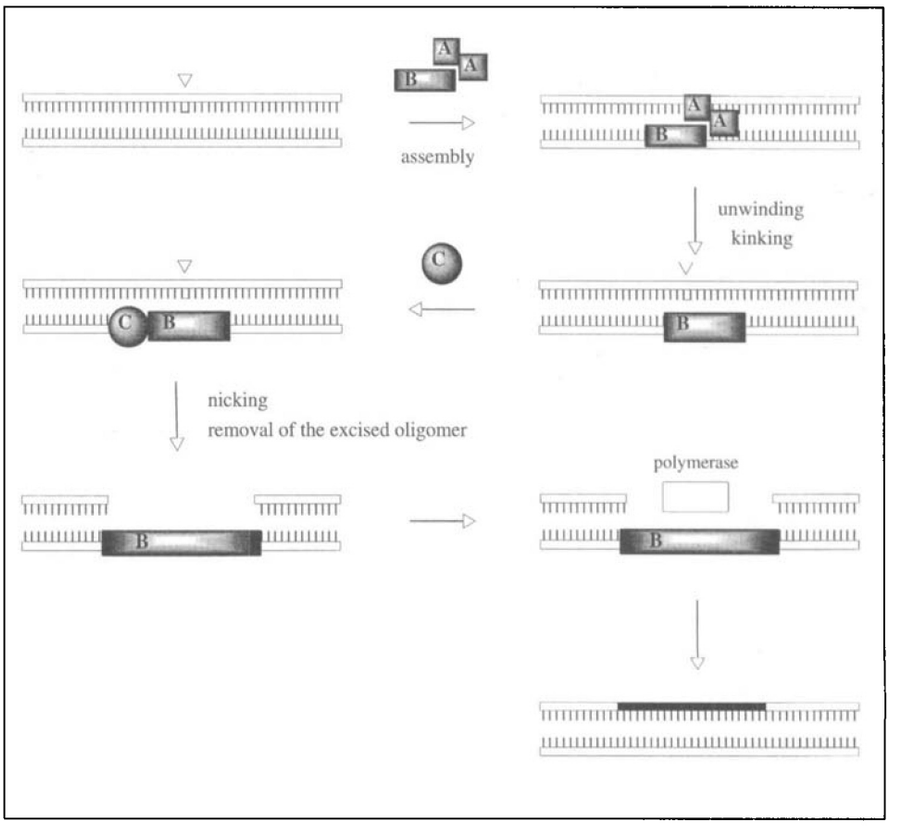
In *Escherichia coli* the excision repair process is well documented by a wealth of experimental results. In *Escherichia coli* three proteins coded by the UvrA-UvrC genes are necessary and sufficient for the repair process. A schematic representation of the mechanism is depicted in *Scheme 5*. The UvrA acts as a damage

recognition protein. It forms a $A_2B$ complex with UvrB. This complex binds to the damaged part of the DNA, unwinding and kinking this DNA region. UvrB forms a tight complex with the conformational altered DNA strand which is subsequently recognized by UvrC. The UvrB and UvrC proteins of the UvrB-UvrC-DNA complex nick the fifth phosphodiester bond on the 3' end and the eighth phosphodiester bond at the 5' end. Helicase II (UvrD) removes the 12 to 13 basepair long excised oligomer and polymerase I fills in the excision gap. Finally, the patch is ligated and in consequence, the damaged site is removed.

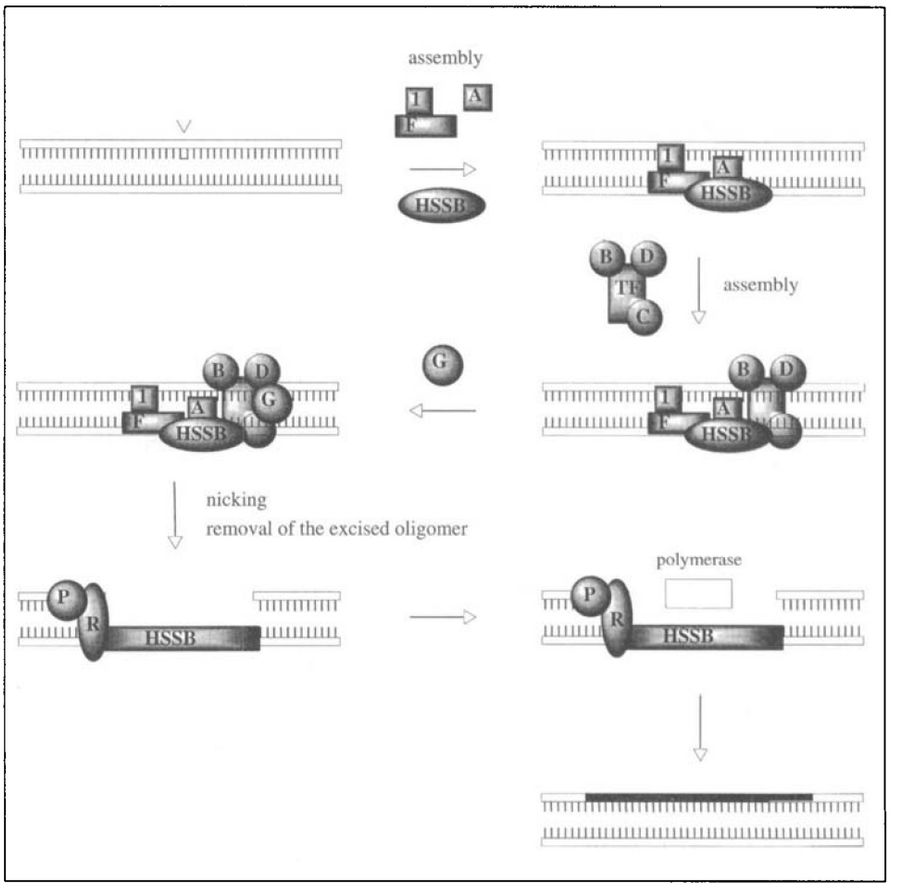
### 6.2. *Human Excision Repair* [64][66]

The human excision repair reaction requires more than 17 polypeptides. The proteins involved in prokaryotic and eukaryotic excision repair share no homologies. Nevertheless, the overall repair mechanism, depicted schematically in *Scheme 6*, seems to be very similar. Damage recognition in humans is initiated by the XPA protein which forms a tight complex with the damaged DNA strand and a number of additional factors necessary for the repair. The entire complex seems to consists of *a*) the damaged DNA strand and XPA; *b*) a hetero dimer, comprised of XPF and ERCC1; *c*) the replication protein HSSB. XPA further recruits the general transcription factor TFIIH to the damaged site. TFIIH contains six polypeptides and in addition the XPB and XPD proteins which both possess helicase activity. The last two proteins which merge with the repair complex are XPC and XPG which may be loosly associated to TFIIH prior to assembly. In this preincision complex the DNA seems to be kinked and unwound like in *Escherichia coli*. Incision occurs through action of XPG and XPF which both possess nuclease activity. After the incision only a subset of repair proteins remains associated with the DNA. The release of the excised oligomer is thought to be induced through the action of two other proteins, the proliferating cell nuclear antigen (PCNA), and the RFC replication protein. Finally, the gap is filled in by Pol$\delta$ or Pol$\varepsilon$ and the newly synthesized oligomer is ligated. The mechanism described is, as stated earlier, currently supported partially by preliminary results and further investigation is necessary to deepen our understanding of the underlying repair reactions. These experimental results will show if the repair process follows the proposed reaction sequence. Additional support for the repair mechanism might be available through analysis of the excision

Scheme 5. *Schematic Representation of the* Escherichia coli *Excision Repair Mechanism* (Modified from [65])




Scheme 6. *Schematic Representation of the Human Excision Repair Mechanism* (Modified from [64])



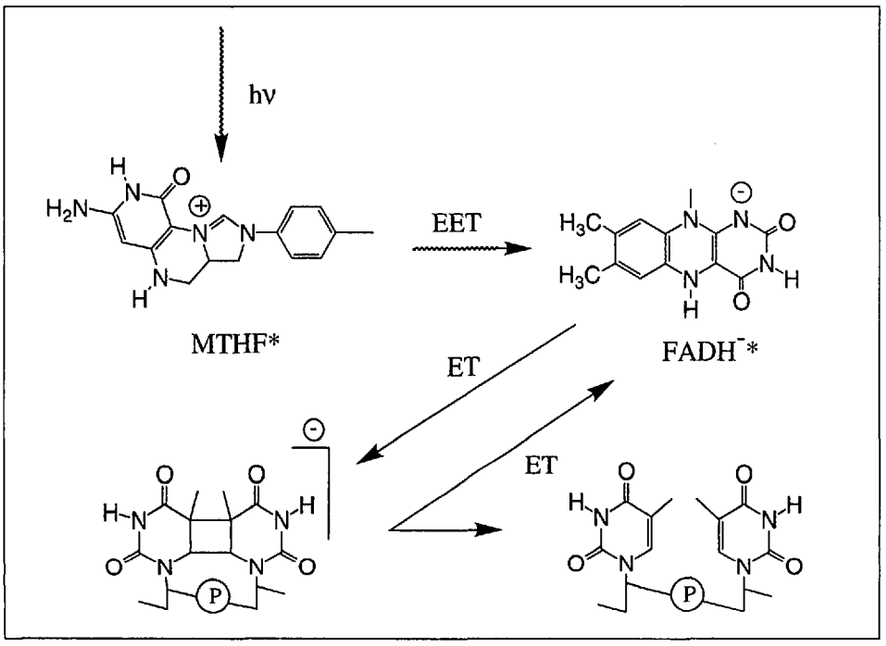
CHIMIA 49 (1995) Nr. 10 (Oktober)

repair system in yeast. Recently, the proteins involved, RAD1–4, 10, 14, and 25, have been purified and characterized. These proteins show sequence homologies to the factors involved in the human repair reaction [67].

## 7. Photoreactivation [65][68]

Photoreactivation is a repair mechanism differing completely from excision repair. During photoreactivation only a single protein, the DNA photolyase, is required for the efficient repair of pyrimidine photodimers within the DNA double helix. Photolyases are proteins with a molecular weight between 55000 and 60000 Daltons. All photolyases contain two cofactors. One is a flavin which has to be in the $FADH^-$ form for the repair reaction. As the second chromophor, photolyases contain either an 8-hydroxydeazaflavin (8-HDF) or a 5,10-methenyltetrahydrofolate (5,10-MTHF). These cofactors without their side chains are depicted in *Fig. 2*. In the first step of the repair reaction the DNA photolyase recognizes the damaged DNA region and forms a tight complex with this region. The fact that photolyases bind nearly equally strongly to single and double stranded DNA currently fuels speculations that the photolyases belong to the class of damage recognition proteins which recognize not a structural DNA alteration but the weakend H-bonded area. In this respect, photolyases are thought to flip the damaged nucleotides out of the double helix into a special active site cavity, where the repair of the lesion occurs. The X-ray structure of the DNA photolyase from *Escherichia coli*, solved recently, reveals a number of interesting features [63]. One is the active center, a cavity which possesses the right dimensions and polarity distribution to accommodate such a flipped out thymine photodimer.

Scheme 7. *Schematic Representation of the DNA-Photolyase Repair Mechanism* (EET = Excitation Energy Transfer; ET = Electron Transfer)



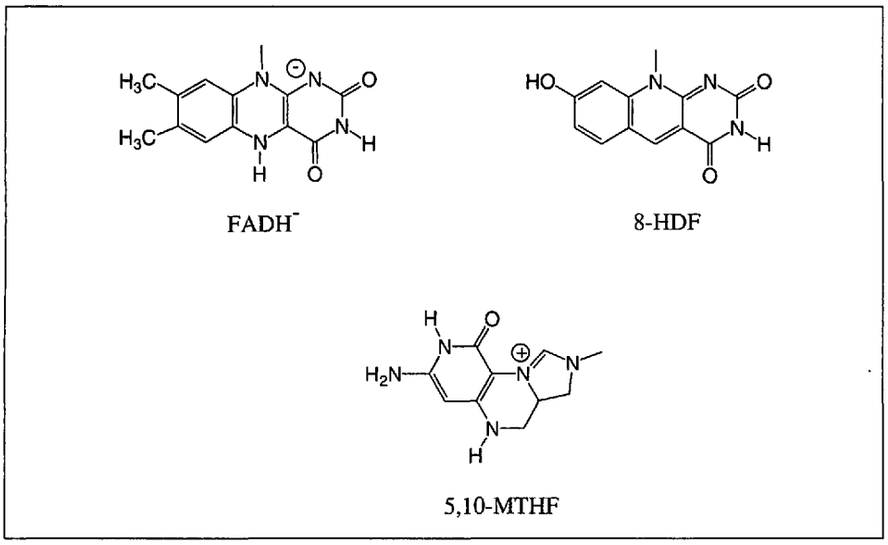




Fig. 2. *Representation of the cofactors present in DNA-photolyases without their side chains*

The repair reaction, schematically described in *Scheme 7*, involves electron and energy transfer steps. For the initiation of the repair reaction, the photolyase DNA complex needs sunlight (300–500 nm). Upon irradiation of the photolyase-DNA complex the flavin is transformed into the excited state either through direct absorption of a light quantum or through energy transfer from the excited second chromophor (in *Fig. 2* the 5,10-MTHF). The second cofactor therefore acts predominantely as a light harvesting chromophor and is not directly involved in the repair reaction. The excited and reduced flavin is located inside the photolyase cavity and would come into *van der Waals* contact with the thymine in the flipped out conformation [63]. It is widely assumed that the repair reaction requires an electron transfer from the excited flavin to the thymine dimer. This electron donation converts the thymine dimer into a radical anion [69]. As the radical anion, thymine dimers decompose spontaneously into the two thymines. The electron is subsequently transferred back to the flavin and the photolyase DNA complex dissociates. Since this repair reaction depends on the instability of the thymine dimer radical anion the photolyase driven process can only repair UV induced lesions. Other damage in photolyase containing organisms have to be repaired with different repair systems.

## 8. Conclusion

The investigation of how certain gene lesions, like sunlight-induced damage, induce mutations have deepen our understanding of the mechanisms that lead to tumor development. T=T, T=C, and C=C photo dimers show widely different promutagenic properties. C to T transitions and CC to TT double base substitutions are often encountered in mutated p53 genes in various skin tumors. Cytosines within a

photo dimer are thought to be responsible for these mutations. Cytosine within a photo dimer is vulnerable to a variety of subsequent decomposition reactions like deamination. Deaminations and the tautomerisations of cytosines within a photodimer are therefore found to be highly deleterious genome processes.

Whether the organism uses photoreactivation or excision repair to remove the induced lesions in order to protect the genome, both systems are sophisticated repair machineries. They are able to recognize the damaged DNA regions and to remove the lesions. Without those repair systems the information stored in the form of genetic code would readily fade away. Cells would not be able to exist in sunlight. Malfunctioning repair is a major cause for the accumulation of DNA defects and hence for the occurrence of mutations. The mutations are responsible for the development of cancer and are also thought to be involved in the aging process. Our understanding of these important reactions is in many case preliminary and further research to elucidate the repair mechanisms are interesting and challenging research goals. Recent research justify optimism that in the near future even details of the single reactions underlying damage recognition and lesion removal might be fully understandable. With this knowledge in hand researchers can start to synthesize molecules able to complement the human repair system and to restore a proper repair in cells with pathological repair defects.

Received: August 28, 1995

[1] T. Lindahl, *Nature (London)* **1993**, *362*, 709.
[2] J.-S. Taylor, *Acc. Chem. Res.* **1994**, *27*, 76; J.-S. Taylor, *J. Chem. Ed.* **1990**, *67*, 835.
[3] F. Hutchinson, *Photochem. Photobiol.* **1987**, *45*, 897.
[4] T. Lindahl, A. Andersson, *Biochemistry* **1972**, *11*, 3618; L.A. Loeb, B.D. Preston, *A. Rev. Genet.* **1986**, *20*, 201.
[5] T. Lindahl, B. Nyberg, *Biochemistry* **1974**, *13*, 3405.
[6] M. Ehrlich, X.-Y. Zhang, N.M. Inamdar, *Mutat. Res.* **1990**, *238*, 277.
[7] L.A. Frederico, T.A. Kunkel, B.R. Shaw, *Biochemistry* **1990**, *29*, 2532.
[8] H. Kasai, S. Nishimura, *Nucleic Acids Res.* **1984**, *12*, 2137.
[9] S. Shibutani, M. Takeshita, A. Grollman, *Nature (London)* **1991**, *349*, 431.
[10] S.S. Wallace, *Envir. Molec. Mutag.* **1988**, *12*, 431.
[11] L.H. Breimer, *Molec. Carcinog.* **1990**, *3*, 188.
[12] N.R. Pace, *Cell* **1991**, *65*, 531.
[13] E.C. Friedman 'DNA-Repair', W. H. Freeman & Co, New York, 1985.
[14] DNA repair works its way to the top: *Science* **1994**, *266*, 1926.
[15] P. Modrich, *Science* **1994**, *266*, 1959.
[16] E. Hecker, E.G. Jung, F. Marks, W. Tilgen, Ed., Series: 'Recent Results in Cancer Research 128', Skin carcinogenesis in man and in experimental models, Springer Verlag, Heidelberg, 1993.
[17] G.G. Glass, R.N. Hoover, *J. Am. Med. Assoc.* **1989**, *262*, 2097.
[18] R. Stolarski, R. Bojkov, L. Bishop, C. Zerefos, J. Staehelin, J. Zawodny, *Science* **1992**, *256*, 342.
[19] J.C. van der Leun, *J. Photochem. Photobiol., B: Biol.* **1988**, *1*, 493.
[20] A.R. Blaustein, P.D. Hoffman, D.G. Hokit, J.M. Kiesecker, S.C. Wallis, J.B. Hays, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1791.
[21] H.W. Thielmann in [16], pp. 275–297.
[22] L. Daya-Grosjean, N. Dumaz, A. Sarasin, *J. Photochem. Photobiol. B: Biol.* **1995**, *28*, 115; A.J. Natarag, J.C. Trent III, H.N. Ananthaswamy, *Photochem. Photobiol.* **1995**, *62*, 218.
[23] J.-S. Taylor, M.P. Cohrs, *J. Am. Chem. Soc.* **1987**, *109*, 2834.
[24] J.-S. Taylor, H.-F. Lu, J.J. Kotyk, *Photochem. Photobiol.* **1990**, *51*, 161.
[25] D.E. Brash, S. Seetharam, K.H. Kraemer, M.M. Seidman, A. Bredberg, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3782.
[26] J.E. LeClerk, A. Borden, C.W. Lawrence, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9685.
[27] N. Jiang, S.-T. Taylor, *Biochemistry* **1993**, *32*, 472.
[28] S.K. Banerjee, R.B. Christensen, C.W. Lawrence, J.E. LeClerk, *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8141.
[29] C.A. Smith, J.-S. Taylor, *J. Biol. Chem.* **1993**, *268*, 11143.
[30] J.-S. Taylor, I.R. Brockie, C.L. O'Day, *J. Am. Chem. Soc.* **1987**, *109*, 6735.
[31] J.-S. Taylor, I.R. Brockie, *Nucleic Acids. Res.* **1988**, *16*, 5123.
[32] J.-S. Taylor, S. Nadji, *Tetrahedron* **1991**, *47*, 2579.
[33] C.-I. Wang, J.-S. Taylor, *Biochemistry* **1992**, *31*, 3671.
[34] J.-S. Taylor, C.L. O'Day, *Biochemistry* **1990**, *29*, 1624.
[35] R.H. Tamarin, 'Principles of Genetics', WCB, Oxford, 1991, p. 470.
[36] I. Tessman, *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6614; K.L. Larson, B.S. Strauss, *Biochemistry* **1987**, *26*, 2471.
[37] J.-S. Taylor, D.S. Garrett, I.R. Brockie, D.L. Svoboda, J. Telser, *Biochemistry* **1990**, *29*, 8858.
[38] S.J. Kuerbitz, B.S. Plunkett, W.V. Walsh, M.B. Kastan, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7491.
[39] M.B. Kastan, Q. Zhan, W.S. El-Deiry, F. Carrier, T. Jacks, W.V. Walsh, B.S. Plunkett, B. Vogelstein, A.J. Fornace, Jr., *Cell* **1992**, *71*, 587.
[40] Jean Marx, *Science*, **1994**, *266*, 1321.
[41] A.R. Clarke, C.A. Purdie, D.J. Harrison, R.G. Morris, *Nature (London)* **1993**, *362*, 849.
[42] S.W. Lowe, H.E. Ruley, T. Jacks, D.E. Houseman, *Cell* **1993**, *74*, 957.
[43] M.L. Smith, I.T. Chen, Q. Zhan, I. Bae, C.-Y. Chen, T.M. Gilmer, M.B. Kastan, P.M. O'Connor, A.J. Fornace, Jr., *Science* **1994**, *266*, 1376.
[44] M.S. Greenblatt, W.P. Bennett, L. Hollstein, C.C. Harris, *Cancer Res.* **1994**, *54*, 4855.
[45] R.C. Bockrath, M.K. Cheung, *Mutat. Res.* **1973**, *19*, 23.
[46] S. Person, J.A. McClosky, W. Snipes, R.C. Bockrath, *Genetics* **1974**, *78*, 1035.
[47] N. Dupuy-Mamelle, B. Pullman, *J. Chim. Phys. Phys.-Chim. Biol.* **1967**, *64*, 708.
[48] D.M. Brown, M.J.E. Hewlins, *J. Chem. Soc. C* **1968**, 2050.
[49] L. Grossman, K. Kato, L. Orce, *Fed. Proc.* **1966**, *25*, 276.
[50] R.B. Setlow, W.L. Carrier, *J. Mol. Biol.* **1966**, *17*, 237.
[51] D. Fix, R.C. Bockrath, *Mol. Gen. Genet.* **1981**, *182*, 7.
[52] F.-T. Liu, N.C. Yang, *Biochemistry* **1978**, *17*, 4865.
[53] T. Lindahl, *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3649.
[54] C. Upton, D.T. Stuart, G. McFadden, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4518.
[55] I. Husain, J. Griffith, A. Sancar, *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2558.
[56] F.N. Hayes, D.L. Williams, R.L. Ratliff, A.J. Varghese, C.S. Rupert, *J. Am. Chem. Soc.* **1971**, *93*, 4940.
[57] D.A. Pearlman, S.R. Holbrook, D.H. Pirkle, S.-H. Kim, *Science* **1985**, *227*, 1304.
[58] S.N. Rao, J.W. Keepers, P. Kollman, *Nucleic Acids Res.* **1984**, *12*, 4789.
[59] J. Kemmink, R. Boelens, T. Koning, G.A. van der Marel, J.H. vanBoom, R. Kaptein, *Nucleic Acids Res.* **1987**, *15*, 4045.
[60] C.-I. Wang, J.-S. Taylor, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9072.
[61] S. Klimasauskas, S. Kumas, R.J. Roberts, X. Cheng, *Cell* **1994**, *76*, 357.
[62] R. Savva, K. McAuley-Hecht, T. Brown, L. Pearl, *Nature* **1995**, *373*, 487.
[63] H.-W. Park, S.-T. Kim, A. Sancar, J. Deisenhofer, *Science* **1995**, *268*, 1866.
[64] A. Sancar, *Science* **1994**, *266*, 1954.
[65] A. Sancar, *Biochemistry* **1994**, *33*, 1.
[66] A. Sancar, *J. Biol. Chem* **1995**, *270*, 15915.
[67] S. Prakash, P. Sung, L. Prakash, *Ann. Rev. Genet.* **1993**, *27*, 33.
[68] S.-T. Kim, A. Sancar, *Photochem. Photobiol.* **1993**, *57*, 895.
[69] P.F. Heelis, S.-T. Kim, T. Okamura, A. Sancar, *J. Photochem. Photobiol. B: Biol.* **1993**, *17*, 219.