Photoinduced processes in DNA: basic theoretic and experimental features

Roberto Improta, Thierry Douki

Chapter 1: Photoinduced processes in DNA: basic theoretic and experimental features.

Roberto Improta [a*] and Thierry Douki [b*].

[a]CNR--Consiglio Nazionale delle Ricerche, Istituto di Biostrutture e Bioimmagini (IBB-CNR), via Mezzocannone 16,

[b] Univ. Grenoble Alpes, CEA, CNRS, IRIG, SyMMES, F-38000 Grenoble

*corresponding email address: robimp@unina.it, thierry.douki@cea.fr

ABSTRACT

This chapter is intended to provide a concise overview on some basic and general concepts of photophysics/photochemistry useful for the understanding of the following chapters. After a general and easy introduction on the theoretical basis of the interaction between light and molecules, we shall briefly describe the principle of the most used steady-state and Time-resolved experimental spectroscopy techniques. Then we shall introduce the most used analytical techniques to quantify and characterize the photodamage and describe the main biological phenomena associated with DNA damage and repair.

1.1. Introduction

This book describes the main processes triggered by the interaction of UV light with nucleic acids, either direct or indirect through the mediation of other compounds, and discuss some of the resulting biological consequences. It is not possible, obviously, to treat, even rapidly, all the scientific and technical aspects involved in the study of such

a large variety of phenomena. On the other hand, it can be helpful for the readers to concisely review some basic features of the techniques, experimental or computational, most commonly used to investigate these processes, and the general concepts/models employed for their interpretation, focussing essentially on those most thoroughly used in the following chapters. This is the goal of the present introductory chapter. Some parts will be obvious to chemists and some to biologists. Hopefully, the all readers will share a common language and will be able to gather useful information from the whole volume.

1.2. Some theoretical basis

According to Quantum Mechanics (QM), the behaviour of atom and molecules is governed by the Hamiltonian operator, which contains all the energy terms associated to electrons and nuclei [1, 2]. The solution of the associated Schrödinger equation, which allows predicting the time evolution of the system, is much simpler by resorting to the Born-Oppenheimer (BO) approximation, which provides the separation of the motion of electrons and nuclei, the latter being much heavier and, therefore, much slower. In this framework, for a molecule we can identify different electronic states (i.e. the solution of the electronic Hamiltonian) differing on the distribution of the electrons in the Molecular Orbitals (MO). It is then possible to compute for each molecule a ground electronic state with its equilibrium geometry, corresponding to the structure where the total energy (electron + nuclei) of the molecule reaches its minimum. Please remind that the total energy is negative so a decrease of the energy corresponds to a stabilization of the system. When the ground electronic state (sometimes labelled as

GS) is a singlet (i.e. all the electrons are spin paired in their MOs) is very often labelled as S_0.

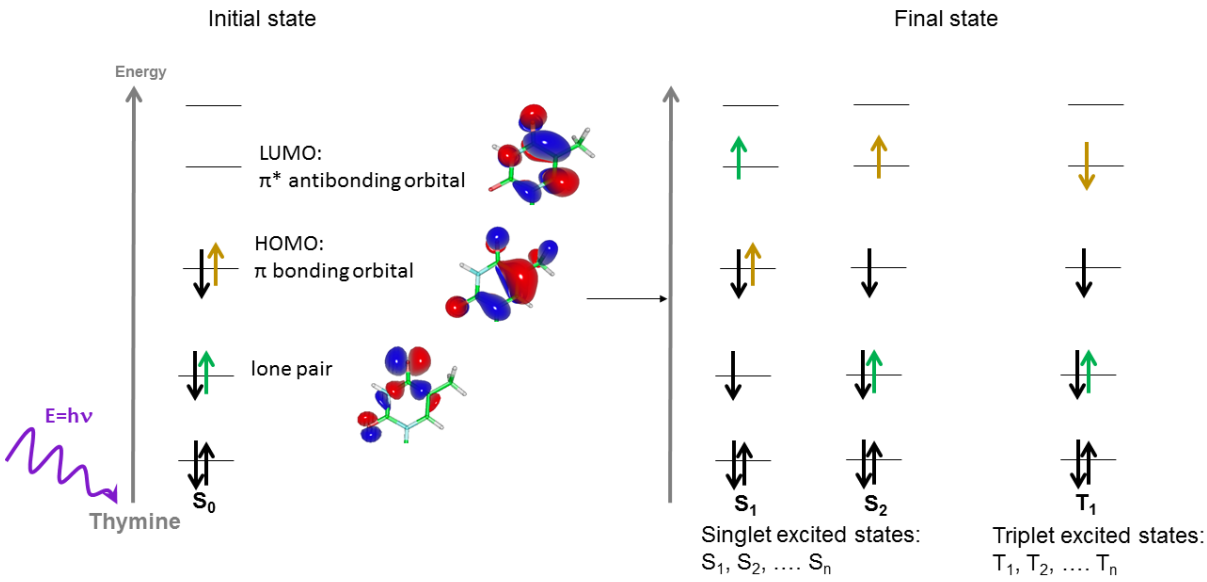


Figure 1.1. *Schematic description of some of the changes in the electronic states population induced by light irradiation of a molecule (using thymine as example).*

As schematically depicted in Figure 1.1. When a molecule absorbs a photon of appropriate energy, i.e. corresponding to the energy difference between the 'ground' and an 'excited' electronic state, an electron is promoted to some vibrational level in the excited singlet manifold (S_1, S_2, S_3 in order of increasing energy) [3]. However, not all the transitions are 'allowed', i.e. for symmetry reasons light absorption can populate some excited states (labelled as bright excited states) and not others (dark excited states). For example, in the molecule depicted above (thymine), the lowest energy bright excited state corresponds to S_2 and can be described as due to a $\pi\pi^*$ (indicating that the transfer is from a π to a π^* MO symmetry of the involved MOs) transition. In thymine it involves the transfer of an electron from the Highest Occupied Molecular Orbital (HOMO, a π MO) to the lowest unoccupied MO (LUMO, a π^* MO). A similar transition is usually described as HOMO→ LUMO. S_1, i.e. the lowest energy excited

state in the ground state minimum, is a dark $n\pi^*$, involving a Lone Pair of the carbonyl oxygen, which is the second highest Occupied Molecular Orbital (HOMO-1) and the same π^* LUMO. This transition is usually labelled as HOMO-1$\rightarrow$ LUMO. In addition to excited states having the same multiplicity as the ground electronic state (a singlet in our example), we can have excited electronic states with different multiplicity, as the Triplets (T_1, T_2, T_3 in order of increasing energy), where the two electrons in the 'half filled' MO's have the same spin. These excited states, which can be described by following a similar framework used below for the singlets, i.e. based on the occupied MO's, cannot be reached directly by light irradiation but, as discussed below, can also play a role in the photoactivated dynamics. We highlight that we have just described the different adiabatic states (S_1, S_2, T_1, etc) according to their 'diabatic' nature, related to the occupancy of the different electronic states. The 'adiabatic' description simply denote the relative energy of the different 'adiabatic' electronic state for a give nuclear structure. As a consequence, the mapping between adiabatic and diabatic picture can thus depend on the considered geometry. In other words, as we'll see below, we can find nuclear arrangements where, in the example we have used, S_1, i.e. the lowest energy adiabatic state, corresponds to a $\pi\pi^*$ transition, i.e. to a different diabatic state than in the ground state minimum.

After light absorption, our molecule is now on S_2 electronic state. The process of light absorption is extremely rapid (≤1 fs) and the nuclei can be considered frozen on that time-scale, i.e. they still have the 'optimal' arrangement for S_0. Within the BO approximation, as we discussed above for S_0, we can associate to each electronic state a Potential Energy Surface, which tells how the energy of this state changes due to the motion of the nuclei. The molecule thus starts evolving on the S_2 PES, i.e. the nuclei start readjusting to find the new minimum energy structure and the total energy

of S_2 starts decreasing, and can reach its minimum. In its minimum can be the $\pi\pi^*$ diabatic state is the lowest energy excited state (S_1), whereas the $n\pi^*$ state now corresponds to S_2.

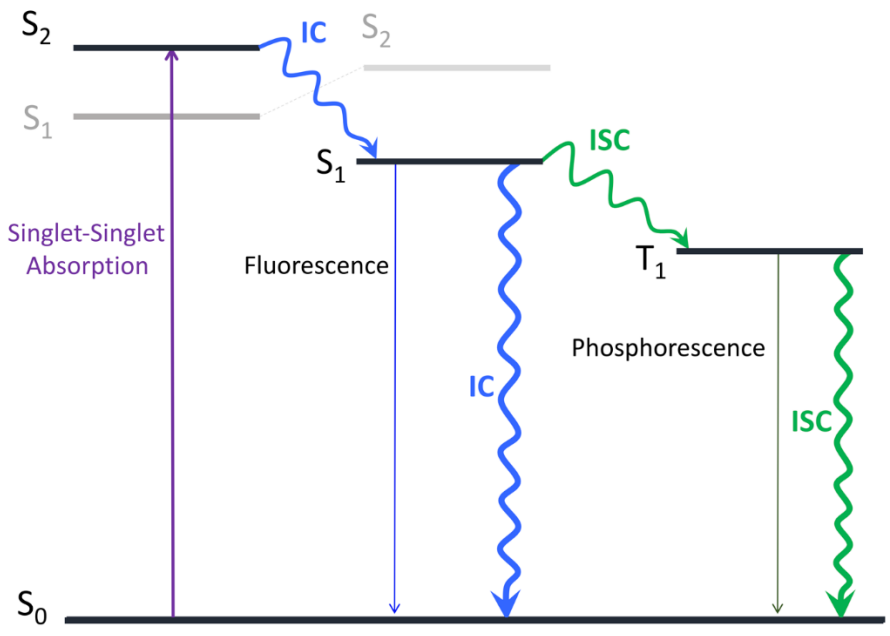


Figure 1.2. *Jablonski diagram of the main photophysical processes in a molecule photoexcited on S_2.*

One possibility is then that the molecule emits light and so we have a radiative decay to S_0. This process is known as fluorescence, when emission takes place from a single state, but eventually it can also happen from a triplet state called phosphorescence (Figure 1.2 straight arrows). The ratio between the photons absorbed and emitted is defined as Fluorescence Quantum Yield (often abbreviated as QY and sometimes as ϕ). The concept of QY is very important in the study of radiation-induced processes and it is not limited to fluorescence, but it can be extended to any possible photoactivated event, as the ratio between the number of events occurring in the

system and that of photon absorbed. For example, in Chapter 2, 3 etc, we shall often refer to the QY of the different photochemical processes.

As depicted in Figure 1.2 (curved arrows) a bright excited state can also decay non-radiatively by internal conversion to dark excited states (both singlet and triplets) or, directly, to S_0. This process is usually described as 'internal conversion' (IC) or as intersystem crossing (ISC), depending if involves states with the same or with different spin multiplicity.

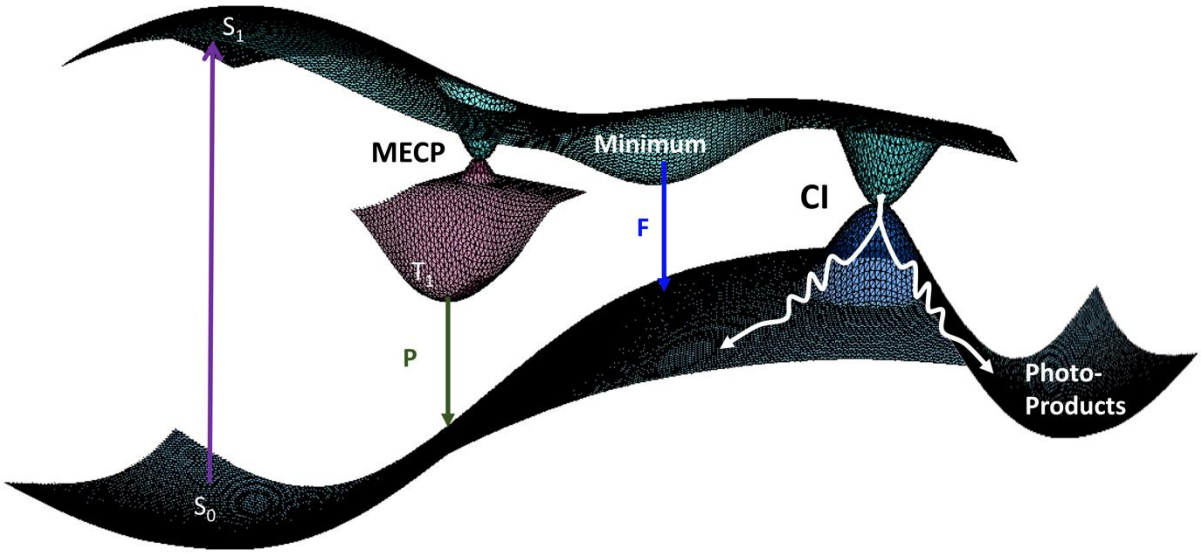


Figure 1.3. *Schematic picture of the Potential Energy Surfaces of the ground (S_0) and an excited electronic state (S_1) and of the main photophysical and photochemical processes involved (see text for details).*

The non-radiative decay is particularly effective in the proximity of the surface crossings, i.e. when two different PES are degenerate (Figure 1.3). In these regions, the BO approximation is not valid, and the so-called non-adiabatic coupling effects and, for singlet/triplets crossings, Spin-Orbit couplings need to be taken in account to correctly describe the evolution of the molecular system. For only two coordinates, the

intersection between the PES of two states of different symmetry is a single point and the PES can adopt a typical 'conic' shape (with the degeneracy point being the vertex), giving account of the term 'conical intersection' (CI) usually adopted to define the lowest energy structure of a crossing seam between two singlet states. A similar structure for crossing between states of different multiplicity the more general label of minimum energy crossing point (MECP) can be used. The presence of an easily accessible crossing region with S_0 on the PES of a bright excited state leads to a very fast and effective non-radiative ground state recovery, mirrored by an ultrashort excited state lifetime.

Strongly fluorescent molecules have QY close to 1 and radiative lifetime of several nanoseconds. The fluorescence QY of DNA and of its component is instead extremely low (10^{-3}~10^{-4}) and the lifetime of the bright excited states is of a few ps (at maximum). Actually, for nucleobases an almost barrierless path on the PES of the bright excited states leads to a crossing region with S_0 (see chapter 2). Moreover, for oligonucleotides, the bright excited state can decay to almost-dark excited states (for example transitions with Charge Transfer character) which then return to S_0, non radiatively, by charge recombination.

Figure 1.3. shows another possible outcome of the photoactivated dynamics, i.e. that the system reaches another stable arrangement of the nuclei, i.e. producing a 'new' molecule stable at room temperature. Examples of these 'photochemical reactions' are the photodimerizations discussed in Chapters 2,3 or the ionizations described in Chapter 4.

From the computational point of view, a full understanding of the photoactivated dynamics requires the complete characterization of the PES's of the electronic states,

of their crossings, stationary points, and the characterization of the spectral signatures of such stationary points (e.g. absorption and emission spectra). This 'static' picture can already give useful insights on the most important excited state processes, and it is nowadays quite routinely applied to many systems, including oligonucleotides. On the other hand, the simulation of the excited state dynamics requires also the determination of the non-adiabatic couplings between the different electronic states and the application of dynamical methods, either semiclassical (i.e. the motion of the nuclei is described according the classical law of mechanics) or quantum dynamical (i.e. describing also the motion of the nuclei at a QM level). Thanks to the development of more accurate computational methods, efficient software and the continuously increasing computing power, the possibility of studying, at the computation level, the excited state dynamics of molecular systems has known impressive advances in the last years, as witnessed also by many chapters of this book. [4].

1.3. Notion of spectrum in photobiology

All photochemical reactions and photobiological processes are strongly dependent on the wavelength of the incident photons. The notion of spectrum is fundamental and is used in several ways.

1.3.1 Emission spectrum of the light source

The first major parameter that needs to be properly assessed in real-life or laboratory experiments is the characterization of the light source to which the samples of interest are exposed. This emission spectrum described the intensity of the radiation over a

chosen range of wavelengths. This task is performed with more or less sophisticated radiometer. Obviously, values cannot be provided for the infinite number of wavelengths but on small wavelengths intervals the width of which defines the resolution of the measuring device, the radiometer. Not all light sources consist of mixture of radiation. For example, lasers emit photons at a single wavelength. They are referred to as monochromatic. A correct assessment of the emission spectra of sources used in photobiology is of paramount importance. For example, works on the effects of UVA can be completely misinterpreted if traces of contaminant UVB are present in the emitted radiation because photons in this wavelength can be as high as 2 or 3 orders of magnitude more biologically efficient than in the UVA.

1.3.2 Absorption spectrum

A first necessary step for a photochemical reaction or a photobiological process to be triggered is the absorption of a photon by a component often described as the chromophore. The ability of a molecule to absorb photons of different energies are shown by its absorption spectrum. They report the proportion of incident energy of a monochromatic radiation. The absorption is maximal when the absorbed energy corresponds to the energy of a transition between the ground state of the chromophore and one of its excited states. The transitions may for example, as anticipated in the previous section, correspond to be $\sigma \rightarrow \sigma^*$ (excitation of single bonds), $\pi \rightarrow \pi^*$ (excitation of double bonds) or $n \rightarrow \sigma^*$ (excitation of unpaired electrons). As a general trend, $\sigma \rightarrow \sigma^*$ are more energetic than $\pi \rightarrow \pi^*$ transitions. It should be stressed that the excitation energy decreases for conjugated double bonds and double bonds inserted in aromatic structures. The combination of these individual transitions constitute the complete

absorption spectrum of the chromophore. For chemists, absorption at a specific wavelength is characterized by the molecular absorption coefficient ε expressed in $cm^{-1}\ M^{-1}$ (with M in mole L^{-1}). ε is similar to the absorption cross section used by physicists. The absorption coefficient is used in the beer lambert law to determine the transmitted energy:

$$E_{transmitted} = E_{incident} \times 10^{\varepsilon \times l \times c}$$

where l is the optical path, ε the absorption coefficient and c the concentration of the solution.

This exponential function reflects that most the energy is deposited in the upper layers of the irradiated medium. This observation is of paramount importance in some experiments performed with intense light and highly absorbing media where significant heating or undesired multiphotonic processes can take place.

1.3.3 Action spectrum

The onset of a physiological or cellular response to exposure to light is much more complex than photochemical processes in molecules. To overcome these limitations, photobiologists often use action spectra that report the variation of extent of the studied response over a specific range of wavelengths. From a practical point of view, data are collected within a series of narrow wavelengths ranges and data are fitted to provide a continuous spectrum. A widely used example is the erythema action spectra defined by the CIE (Commission International de l'Eclairage) shown in Figure 1.4

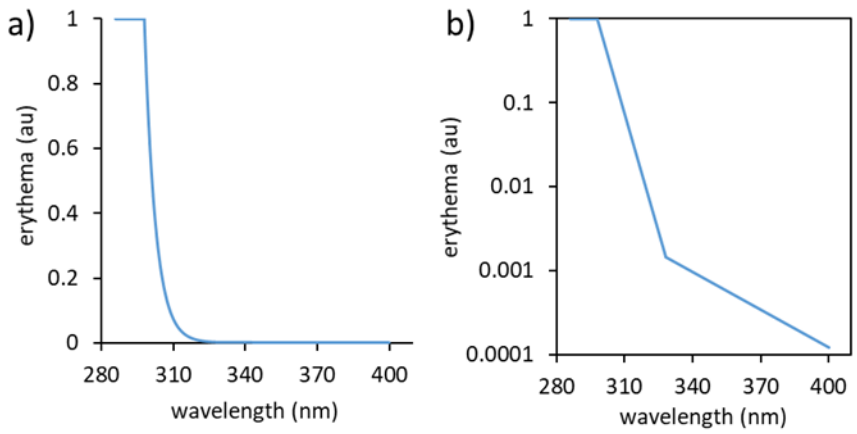


Figure 1.4: CIE action spectrum for induction of erythema shown in a) linear or b) logarithmic scale.

A first use of action spectra is the comparison and the prediction of the effect of different broadband sources. In this case, the emission spectrum is multiplied by the action spectrum of interest, thereby yielding a “weighted emission spectrum”. For example, the erythemal weighted emission spectrum is often used to convert physical doses into more biologically relevant units, which are fractions or multiples of erythemal doses. A second interest of action spectra is the identification of the cellular chromophores involved in photobiological responses. Indeed, the wavelength corresponding to the maximal amplitude of the studied effect is likely to correspond to a maximum of the absorption spectrum of the initial chromophore. This strategy is not always so straightforward since specific cellular features (attenuation of the incident light, local pH effects, structural constraints, etc.) may significantly affect the absorption properties of the chromophore by comparison with the pure molecule in solution.

1.4. A Basic introduction to pump-probe spectroscopies

Independently of the computational developments, our knowledge of the photoactivated dynamics of NAs has hugely profited (as well as many other fields) from the impressing advances made by Time Resolved (TR) spectroscopies in the last decades. Though steady state spectra can provide interesting insights on the main photophysical and photochemical paths, it is clear that the possibility of monitoring how spectral signals change with time, often with < 100 fs resolution, opens new perspectives to our understanding of these complex phenomena. A detailed discussion of each technique, though limiting our discussion to those more commonly used to investigate nucleic acids, is outside the scope of the present book. In the following chapters, the interested reader can find some concise explanations on specific techniques. We here simply sketch some principles of pump-probe spectroscopies, which allow obtaining time-resolved signals. [5-7]. In these techniques, at least two laser pulses interact with the system under investigation. After the first one, the pump, a certain percentage of the molecules populate an excited electronic state. Then a second pulse, the probe, is used to obtain information on the excited state dynamics. A transient spectrum is acquired by measuring the differential transmission, $\Delta T/T$, of the probe with and without the excitation from the pump at a series of different time delays between the two pulses. Just to make some examples, once the pump has excited some of the molecules on the excited state PES depicted in Figure 1.5, the probe can further excite some of them to a higher lying electronic state (S_n), according to a process usually named as Excited State Absorption, ESA). Consequently, in our spectra will have an absorption corresponding to the wavelength of the $S_1 \rightarrow S_n$ transition. On other hand, for some wavelengths, the excited state will absorb 'less' than the ground electronic state. In these regions, considering that part

of the ground state population is on the excited state (ground state bleaching) we shall record ‘negative’ absorption with respect our reference. Analogously, for wavelength close to those of the fluorescence maximum of the excited state, we can have stimulated emission, which also translates in a negative signal. By measuring this broadband spectrum with different pump-probe delay, it is thus possible to obtain TR information on different excited state processes. This information can be crossed-checked from that obtained by other techniques, for example Fluorescence UP-conversion, monitoring the fluorescence of the sample, in order to more easily disentangle the excited state processes. Another important parameter is the energy of the pump and the probe. Since DNA absorbs in the UV, usually the pump falls in this spectral region. A probe in the UV will then provide information on electronic processes, as those described above. If the probe energy is sufficiently high, some of the molecules can be ionized (and we can measure a Time-resolved photoionization spectrum). At the other end of the spectrum, if the probe is in the IR region, it is possible to obtain indications on the vibrational spectra of the excited state minima, with positive/negative peaks depending if the excited state minimum absorbs more or less than the ground state minimum in a given region. In addition to these broadband spectra, it is possible to monitor the TR spectra at selected wavelengths, also obtaining very useful information. For example, monitoring the transient absorption spectrum at the wavelength where the GS strongly absorbs, we can measure the time necessary to recover the GS population and, therefore, on the excited state lifetime.

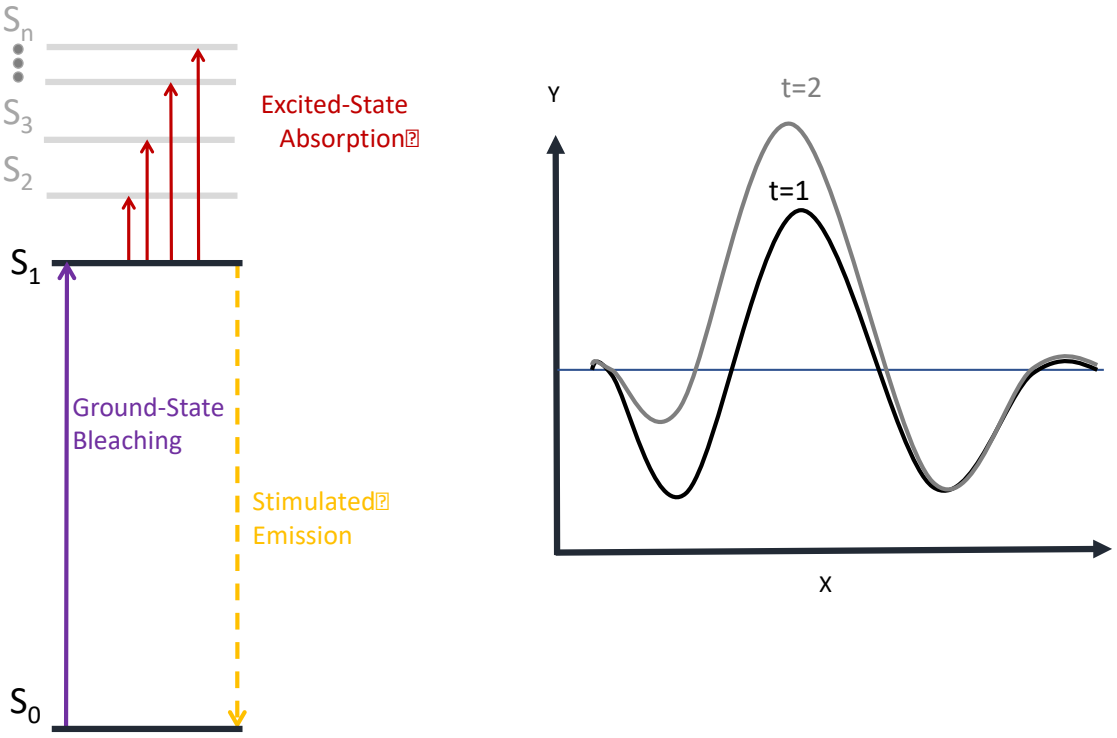


Figure 1.5 Pictorial Scheme of the possible processes observed in a pump-probe experiments (see text for details). The energy or the wavelength is on the x scale. A possible change in the signal with time from t=1 to t=2 is also pictorially shown.

1.5. Analytical tools for the detection of DNA damage

A large fraction of the scientific advances discussed in the present volume relies on the accurate quantification of DNA damage. Those are either dimeric photoproducts (cyclobutane pyrimidine dimers, CPD; pyrimidine (6-4) pyrimidone photoproducts, 64PP; Dewar valence isomers, Dew) or oxidation products (single strand breaks, SSB; 8-oxo-7,8-dihydroguanine, 8-oxoGua). The following section describes the most widely applied techniques, including immunological detection, electrophoresis-based assays and chromatography combined with mass spectrometry. Some techniques will

not be described here. Early assays such a use of radiolabelled DNA are still used in specific cases. ^{32}P-post labelling also received some recent applications. Very recent approaches such as the use of next generation sequencing will be discussed in specific chapters of the volume.

1.5.1 Antibody-based assays

Immunological detection of dimeric photoproducts is a widely used approach in photobiology. This class of techniques relies on the production of polyclonal or monoclonal antibodies raised against UV-irradiated DNA or oligonucleotides [8, 9]. Precise characterization works allowed several groups to produce antibodies specific against each of the three main classes of bipyrimidine dimers, namely CPDs, 64PPs and Dews. These antibodies are then used in different types of assays to reveal the presence of photoproducts. These techniques, when applied to isolated or extracted DNA, include among others radioimmunoassays, enzyme-linked immunosorbent assay and immuno-dot-blot. In whole cells, flow cytometry has also been used. A unique opportunity afforded by immunological assays is the possibility to localize the photoproducts in tissues, for instance in skin, by using immunohistochemistry. A limitation of immunological assays is that results are often obtained in arbitrary units rather than in absolute frequency of damage. Another drawback is that, for a given type of dimeric photoproduct, no difference is made between the TT, TC, CT and CC derivatives. Antibodies are also used for the quantification of oxidized bases such as 8-oxoGua by ELISA. However, the accuracy and the specificity of the results has been challenged.

1.5.2 Electrophoresis-based assays

Electrophoretic separation of biomolecules such as DNA and proteins is a powerful tool relying on large differences in molecular weight and electrical charge. Consequently, electrophoretic techniques have been extensively used for the quantification of strand breaks in DNA. Pulsed field gel electrophoresis has long been the reference technique for double-strand breaks while the Comet assay has become a standard for the detection of SSB and alkali-labile sites [10]. Valuable data have also been obtained by alkaline gel elution. Such methods cannot be directly applied to dimeric photoproducts or oxidized bases, which are stable and do not induce DNA cleavage. Yet, this type of damage can be converted into additional strand breaks by using purified repair enzymes as biochemical tools [11]. The most commonly used for detecting CPDs is the endonuclease V of phage T4. Detection of oxidized purine bases can be performed with the bacterial formamidopyrimidine DNA glycosylase (Fpg) and less frequently with the human oxidized guanine glycosylase 1 (OGG1), which is more specific than Fpg for 8-oxoGua. Oxidized pyrimidines are most frequently visualized by using the bacterial endonuclease III (EndoIII).

1.5.3 Chromatography-based assays

Chromatographic approaches have also been designed to quantify DNA damage. These techniques involve extraction of DNA, hydrolysis aimed at releasing damage and separation of the hydrolysed DNA mixture on a chromatographic system with appropriate detection [12]. This strategy is very specific and provide information on individual damage, often simultaneously in a single analysis. In addition, because the detection can be easily calibrated, results are quantitative. They are for example expressed in number of damage per million normal bases. The hydrolysis step can be

339 performed under strong acidic conditions and releases damage as modified bases. A
340 milder approach that prevents degradation of brittle lesions involves the combination
341 of enzymes such as endonucleases, phosphodiesterases and phosphatase. Under
342 these conditions, monomeric damage are released as nucleosides (base + 2-
343 deoxyribose) and dimeric photoproducts as dinucleoside monophosphates. The
344 chromatographic system used for the separation of the hydrolysed mixture is most
345 often high performance liquid chromatography, although gas chromatography
346 following derivatisation into volatile compounds can be used. The most commonly
347 used detection is mass spectrometry, and in particular tandem mass spectrometry
348 used in the “multiple reaction monitoring” mode. This detection method provides two
349 levels of specificity, one on the molecular mass and the other on the mass to charge
350 ratio of specific fragments. These features, combined with chromatographic
351 properties, insure unique specificity and make possible the simultaneous detection of
352 several damage. The assay has been successfully applied to the three types of
353 photoproducts for which TT, TC, CT and CC derivatives can be quantified individually.
354 Oxidized bases such as 8-oxoGua have also been detected by HPLC-MS/MS. In this
355 case, a limitation comes from the possible induction of artifactual oxidation of DNA
356 extraction step.
357

358 **1.6. Basic biological phenomena associated with DNA damage**

359 **1.6.1 From DNA damage to abnormal proteins**

360 DNA has two main functions in cells, namely the storage of the genetic information
361 required for protein synthesis and the transmission of this information to daughter cells
362 during division. These two functions rely on basic biochemical events: replication and

transcription [13]. Replication involves a complex series of enzymes that use each strand of DNA as a template for the synthesis of a new double-strand. The synthesis of DNA occurs in the replication fork, a structure where the two original strands are separated and the replication machinery is bound to synthesize the new strands. Because DNA synthesis always proceed from the 5'-end to the 3'-end, one of the original strand (leading strand) is replicated in a continuous way by DNA-polymerases. The other strand (lagging strand) has to be synthesized by fragments. Those are first initiated by small transient pieces of RNA that are then extended into DNA Okasaki fragments. RNA is removed and the DNA fragments are ligated to yield a continuous DNA strand. The replication of these DNA strands rely on the formation of base pairs, and thus on the complementarity between adenine and thymine on the one hand, and guanine and cytosine on the other hand. Consequently, any modification of the chemical structure of the bases, namely a damage, can interfere with this process. A first consequence may be errors in the incorporated base opposite the damage upon replication and a change in the local sequence of the DNA (Fig. 1.6). This constitutes a mutation. A second possible effect is the blockage of the DNA-polymerases at the site of the damage. Stalled replication fork are converted by cells into double strand breaks (DSB) that can either be lethal or lead to chromosomal damage. It should be stressed that all the following generations of cells will bear the same mutations. Indeed, even at a mutated site, DNA is constituted of normal bases in correct base pairing and does not represent an abnormal chemical structure that could be detected by the repair machinery.

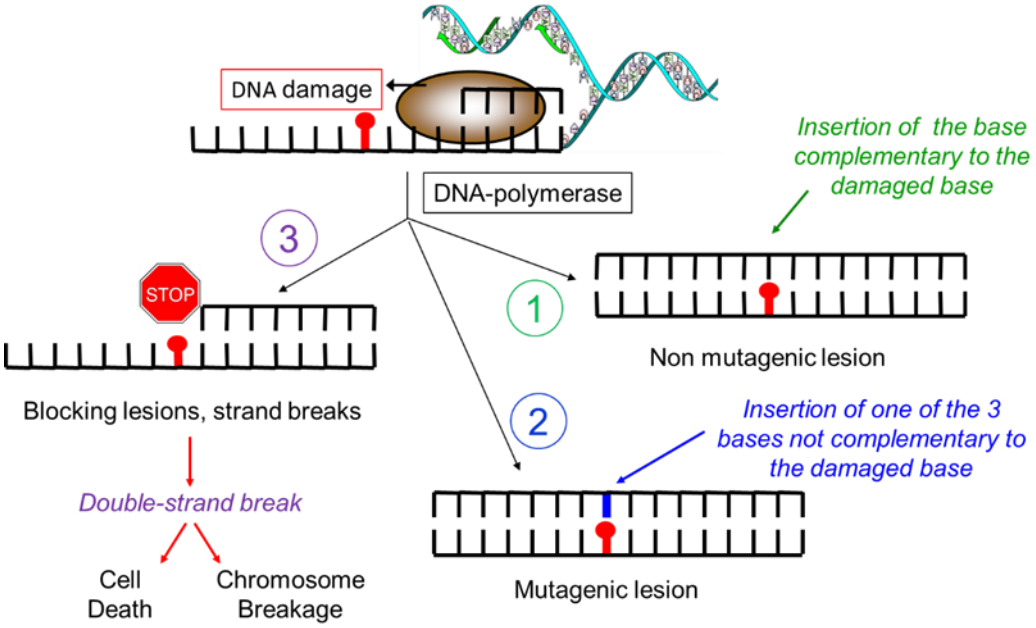


Figure 1.6: Mutagenic events at a damaged site in DNA during replication by a DNA polymerase. Three main scenarii are possible. 1) The modified base exhibits the same coding properties than the normal base it originates from. The replicated sequence is the same as in the absence of damage. 2) Opposite the lesion, the polymerase introduces one of the 3 bases that were not complementary to the initially present normal base. The replicated sequence is thus modified. This mutation is then transferred without modification to the next generation of cells. 3) The damage is too bulk or distorting, which leads to the blockage of the polymerase. The replication fork is stalled which leads ultimately to the induction of DBS when the cell attempts to resume replication. If the cell does not die, DSB can be at the origin of chromosomal aberrations.

Consequences of mutations are the most drastic in genes coding for proteins but the effects in non-coding portions of DNA such as those involved in the formation of miRNA should be fully evaluated. The impact of a mutation in a gene is determined at

two levels. It first should be briefly reminded that the genes have first to be translated into RNA, in a process that is based like replication on the complementarity between bases. A main difference is the presence of uracil instead of thymine and of ribose instead of 2-deoxyribose in RNA. The RNA strand is then spliced and some fragments (exons) are ligated and excreted from the nucleus to the cytoplasm. There, these messenger RNAs (mRNA) are used as templates for the synthesis of the proteins. The mRNA is read by triplet of bases (codon), each codon driving the incorporation of a specific amino acid as defined in the genetic code. In case of induction of a mutation in a gene, the sequence of the corresponding mRNA is not that expected and some codons do not exhibit the proper triplet of bases. Because some amino acids are coded by more than one codon, this may correspond to a silent mutation, which do not change the sequence of the protein. More often, the mutation changes the codon into one coding for another amino acid. Alternatively, the abnormal codon can be a stop codon, in which case the synthesized protein is truncated. When the protein is mutated or truncated in portions important for its structure or its enzymatic activity, the abnormal protein may lose or gain functions. If such events take place in proteins involved in the regulation of division, cells may gain a replication advantage. This process actually represent initiation, the first step of tumour development [14, 15].

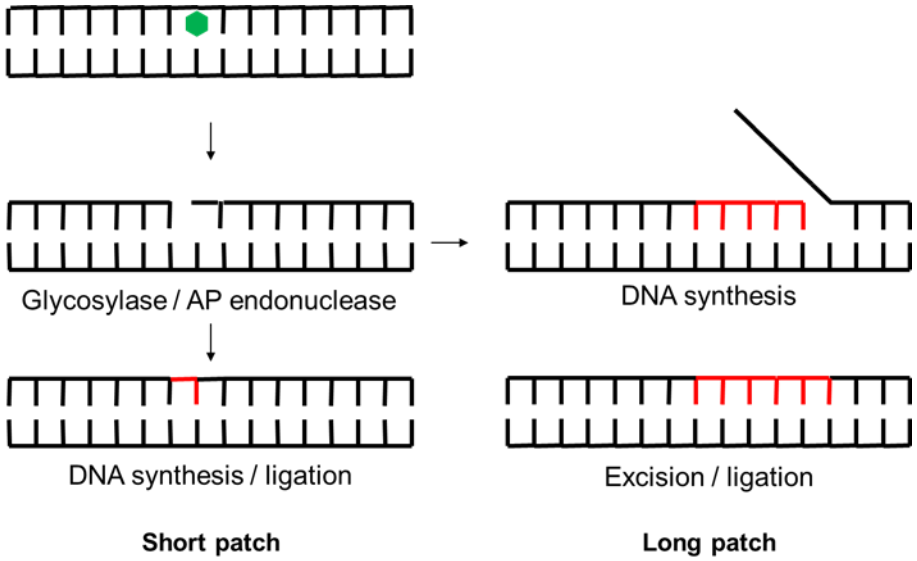
1.6.2 Principle of DNA repair

Because of the drastic deleterious effects of DNA damage, all cell types are equipped with sets of proteins capable of restoring the original structure of DNA and most frequently the initial sequence in error-free repair events. DNA repair gathers very diverse biochemical processes. Some repair enzyme simply revert the damage back into the original bases. This is the case of photolyases encountered in bacteria, plants

and some non-placental mammals. These enzymes use the energy of UVA or visible light to cleave the covalent bonds created by UVB in CPDs or 64PPs and restore unmodified pyrimidines. In contrast to this rather simple reaction, repair of double-strand breaks requires, in non-homologous end-joining, the cleaning of the extremities of the broken DNA strands followed by their ligation in a process involving many proteins. NHEJ is thus error-prone because genetic information is lost. Alternatively, DSB can be repaired, when cells are in the S phase, by homologous recombination. This pathway involves the invasion of a homologous sequence by the strand under repair. A correct fragment is thus synthesized and the complex junction resolved to yield two complete and repaired strands exhibiting the correct sequence.

As shown in the following chapters, UV-induced DNA damage mostly involve modification of the bases. In human cells, which do not contain photolyase activities, those are handled by excision-resynthesis pathways. Two main pathways are involved, namely base excision repair (BER) and nucleotide excision repair (NER). BER is mostly active against small modified bases such as oxidation products [16]. The key enzymes in BER are DNA N-glycosylases (Fig. 1.7). Each of these proteins recognize a specific set of damaged bases (oxidized purines, oxidized pyrimidines, etc.). They cleave the covalent bond between the base and the 2-deoxyribe unit (the N-glycosidic bond) and leave an abasic site (AP sites). DNA is then cleaved at AP sites by AP endonucleases or the initial DNA glycosylase when it is bifunctional. The gap left in the double strand can be either filled by incorporation of one nucleotide based on the complementarity with the undamaged strand. This is the short patch pathway of BER. Alternatively, a small strand of DNA can be synthetized under the damaged portion, which is then cleaved. This is known as the long patch pathway of BER. Interestingly, the mechanism of BER is associated with the induction of strand

451 breaks. Therefore, BER is also the main repair mechanism of this class of lesions. It
452 may be added that other scaffold proteins are present during the whole repair event.



453

454 *Figure 1.7 Basic steps in base excision repair*

455

456 The other excision/resynthesis repair pathway involved in the removal of UV-induced
457 damage is nucleotide excision repair (NER) [17, 18]. NER is mostly active against bulky
458 lesions. It is sensitive to modifications of the DNA conformation rather than the
459 chemical structure of the damage. Most of the proteins involved in NER were found
460 inactive in the skin cancer prone xerodema pigmentosum syndrome (XP-A to G)
461 together with the replication factor TFIIH and RPA. NER involves four main steps (Fig.
462 1.8): i) detection of the damage, ii) unfolding of the double stranded DNA by XPA, RPA
463 and TFIIH, iii) cleavage of a single-stranded fragment bearing the damage by XPG
464 and XPF, and iv) filling of the resulting gap by a polymerase and a ligase. Two sub-
465 pathways have been identified depending on initial detection event. When the RNA-
466 polymerase involved in DNA transcription is blocked by the presence of a
467 photoproduct in an active gene, a very fast repair process is triggered by binding of

the proteins CSB and XPG. It is known as transcription coupled repair (TCR). TCR removes all types of photoproducts at a similar rate. The other NER pathway, the global genome repair, is active in the bulk of the DNA. The recognition of the damage depends on sensor proteins like XPC or XPE, which scan DNA for abnormal local conformation. GGR is less efficient than TCR, especially towards CPDs. Following the detection of the damage, TCR and GGR share the same steps.

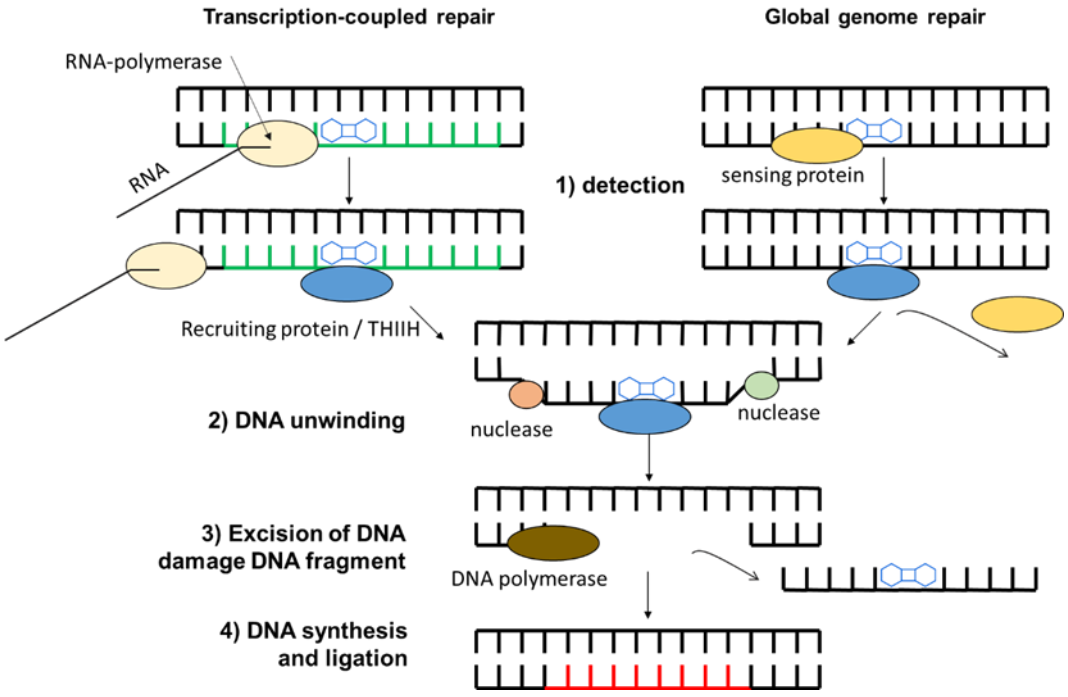


Figure 1.8: Basic steps in nucleotide excision repair. Transcription-coupled repaired handles photoproducts present in active genes and is activated by stalled RNA-polymerase. Global genome repair is initiated by the detection of damage by sensing proteins. In both pathways, NER then involves DNA unwinding, cleavage of a short DNA fragment bearing the photoproduct, and last synthesis and ligation of new DNA.

Acknowledgments

RI thanks Dr. Lara Martìnez-Fernàndez for helpful discussions and help with the figures

ABBREVIATIONS

Abbreviations must be defined at first mention in the chapter and abbreviated thereafter. A list of abbreviations may be provided at the end of the chapter if necessary.

REFERENCES

1. P. W. Atkins and R. S. Friedman, *Molecular Quantum*, Oxford University Press, Oxford, 2010.
2. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, Garden City, NY 11530 1996.
3. N. J. Turro, *Modern Molecular Photochemistry*, Univ Science Books,, Herndon, VA 1991.
4. R. Improta, F. Santoro and L. Blancafort, *Chem. Rev.*, 2016, **116**, 3540-3593.
5. M. K. Shukla and J. Leszcynski, *Radiation Induced Molecular Phenomena in Nucleic Acids: A Comprehensive Theoretical and Experimental Analysis*, Springer Nature, Switzerland 2008.
6. E. H. Egelman, *Comprehensive Biophysics*, Elsevier, 2011.
7. G. Cerullo, C. Manzoni, L. Luer and D. Polli, *Photochem. Photobiol. Sci.*, 2007, **6**, 135-144.
8. N. Kobayashi, S. Katsumi, K. Imoto, A. Nakagawa, S. Miyagawa, M. Furumura and T. Mori, *Pigment Cell Res.*, 2001, **14**, 94-102.
9. D. Mitchell and B. Brooks, *Photochem. Photobiol.*, 2010, **86**, 2-17.
10. E. Cordelli, M. Bignami and F. Pacchierotti, *Toxicol. Res.*, 2021, **10**, 68-78.
11. D. Muruzabal, A. Collins and A. Azqueta, *Food and Chemical Toxicology*, 2021, **147**.
12. T. Douki, *Photochemical and Photobiological Sciences*, 2013, **12**, 1286-1302.
13. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell. 4th edition.*, Garland Science, New York, 2002.

14. L. B. Alexandrov, J. Kim, N. J. Haradhvala, M. N. Huang, A. W. Tian Ng, Y. Wu, A. Boot, K. R. Covington, D. A. Gordenin, E. N. Bergstrom, S. M. A. Islam, N. Lopez-Bigas, L. J. Klimczak, J. R. McPherson, S. Morganella, R. Sabarinathan, D. A. Wheeler, V. Mustonen, P. M. S. W. Group, G. Getz, S. G. Rozen, M. R. Stratton and P. Consortium, *Nature*, 2020, **578**, 94-101.
15. E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J. Humphray, C. D. Greenman, I. Varela, M. L. Lin, G. R. Ordonez, G. R. Bignell, K. Ye, J. Alipaz, M. J. Bauer, D. Beare, A. Butler, R. J. Carter, L. Chen, A. J. Cox, S. Edkins, P. I. Kokko-Gonzales, N. A. Gormley, R. J. Grocock, C. D. Haudenschild, M. M. Hims, T. James, M. Jia, Z. Kingsbury, C. Leroy, J. Marshall, A. Menzies, L. J. Mudie, Z. Ning, T. Royce, O. B. Schulz-Trieglaff, A. Spiridou, L. A. Stebbings, L. Szajkowski, J. Teague, D. Williamson, L. Chin, M. T. Ross, P. J. Campbell, D. R. Bentley, P. A. Futreal and M. R. Stratton, *Nature*, 2010, **463**, 191-196.
16. D. O. Zharkov, *Cellular and Molecular Life Sciences*, 2008, **65**, 1544-1565.
17. M. Fousteri and L. H. Mullenders, *Cell Research*, 2008, **18**, 73-84.
18. J. E. Cleaver, *Advances in Experimental Medicine and Biology*, 2008, **637**, 1-9.