



HAL
open science

Photoinduced processes in DNA: basic theoretic and experimental features

Roberto Improta, Thierry Douki

► **To cite this version:**

Roberto Improta, Thierry Douki. Photoinduced processes in DNA: basic theoretic and experimental features. Roberto Improta; Thierry Douki. DNA Photodamage: From Light Absorption to Cellular Responses and Skin Cancer, Royal Society of Chemistry, 2021, 978-1-83916-558-0. 10.1039/9781839165580-00001 . hal-03689530

HAL Id: hal-03689530

<https://hal.science/hal-03689530>

Submitted on 13 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

3

4 Roberto Improta ^{a*} and Thierry Douki ^{b*}.

5 ^aCNR--Consiglio Nazionale delle Ricerche, Istituto di Biostrutture e Bioimmagini (IBB-
6 CNR), via Mezzocannone 16,

7 ^bUniv. Grenoble Alpes, CEA, CNRS, IRIG, SyMMES, F-38000 Grenoble

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

9

10 ABSTRACT

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

19 **1.1. Introduction**

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

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

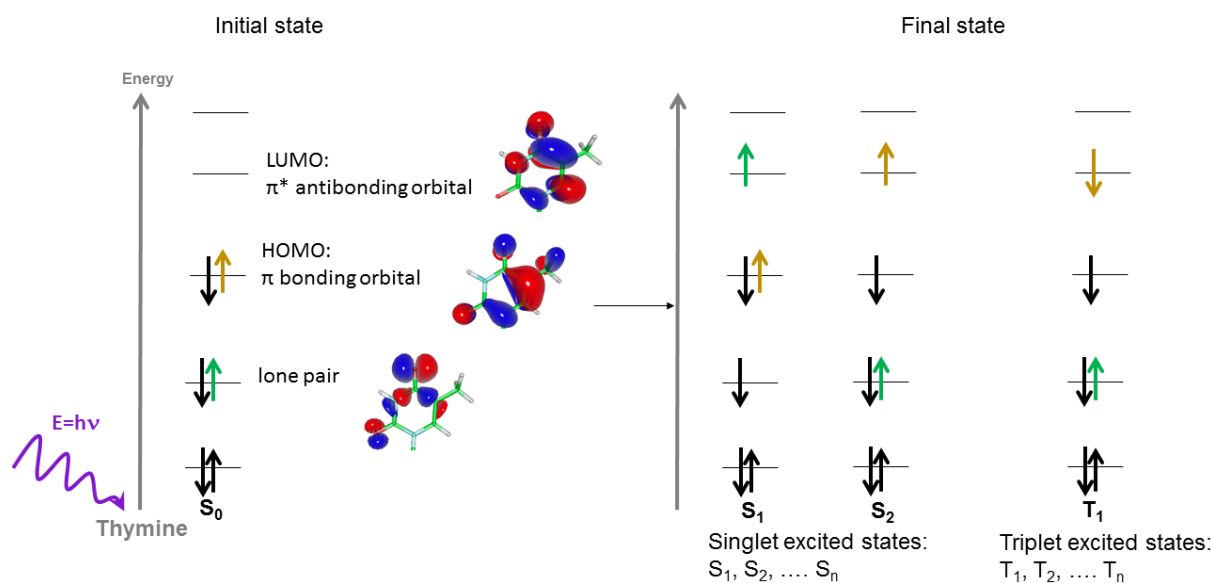
32

33 **1.2. Some theoretical basis**

34

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

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



50

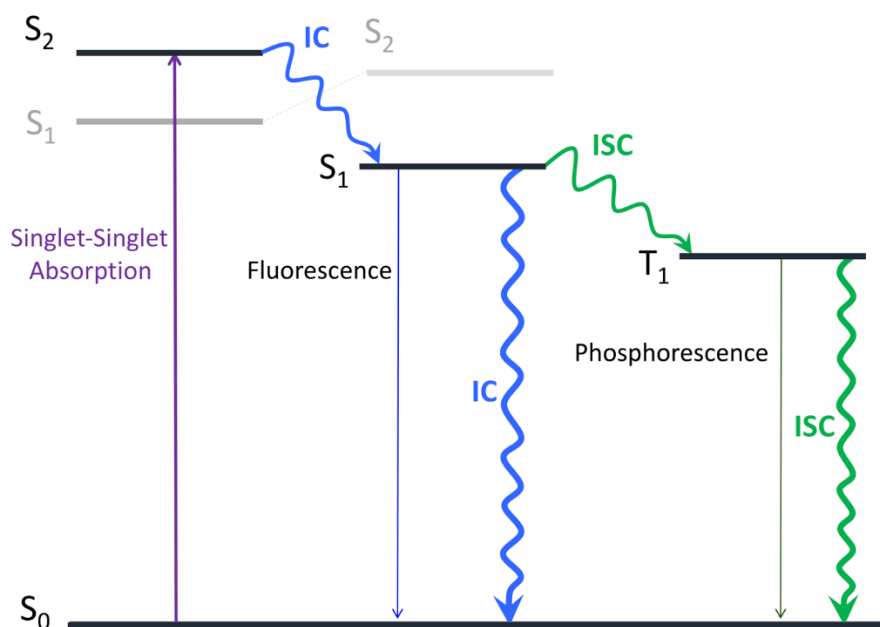
51 **Figure 1.1.** Schematic description of some of the changes in the electronic states
 52 population induced by light irradiation of a molecule (using thymine as example).

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

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

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

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



93

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

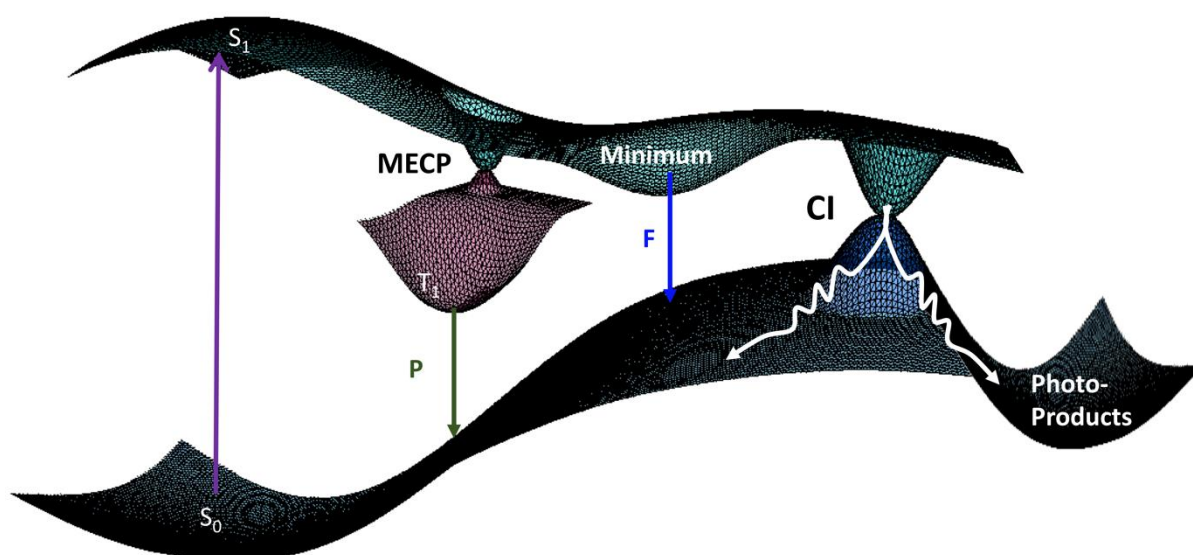
96

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

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

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

112



113

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

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

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

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

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

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

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

159

160 **1.3. Notion of spectrum in photobiology**

161

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

165

166 **1.3.1 Emission spectrum of the light source**

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

170 chosen range of wavelengths. This task is performed with more or less sophisticated
171 radiometer. Obviously, values cannot be provided for the infinite number of
172 wavelengths but on small wavelengths intervals the width of which defines the
173 resolution of the measuring device, the radiometer. Not all light sources consist of
174 mixture of radiation. For example, lasers emit photons at a single wavelength. They
175 are referred to as monochromatic. A correct assessment of the emission spectra of
176 sources used in photobiology is of paramount importance. For example, works on the
177 effects of UVA can be completely misinterpreted if traces of contaminant UVB are
178 present in the emitted radiation because photons in this wavelength can be as high as
179 2 or 3 orders of magnitude more biologically efficient than in the UVA.

180

181 **1.3.2 Absorption spectrum**

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

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

$$199 \quad E_{transmitted} = E_{incident} \times 10^{\epsilon \times l \times c}$$

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

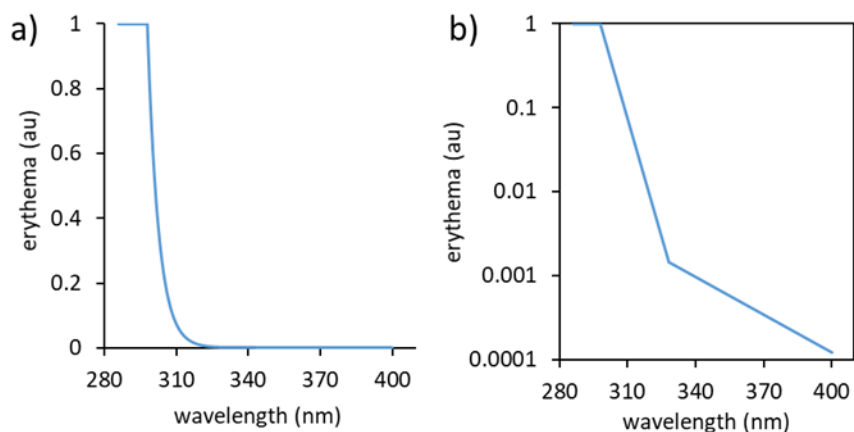
202 This exponential function reflects that most the energy is deposited in the upper layers
203 of the irradiated medium. This observation is of paramount importance in some
204 experiments performed with intense light and highly absorbing media where significant
205 heating or undesired multiphotonic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

206

207 **1.3.3 Action spectrum**

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

215



216

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

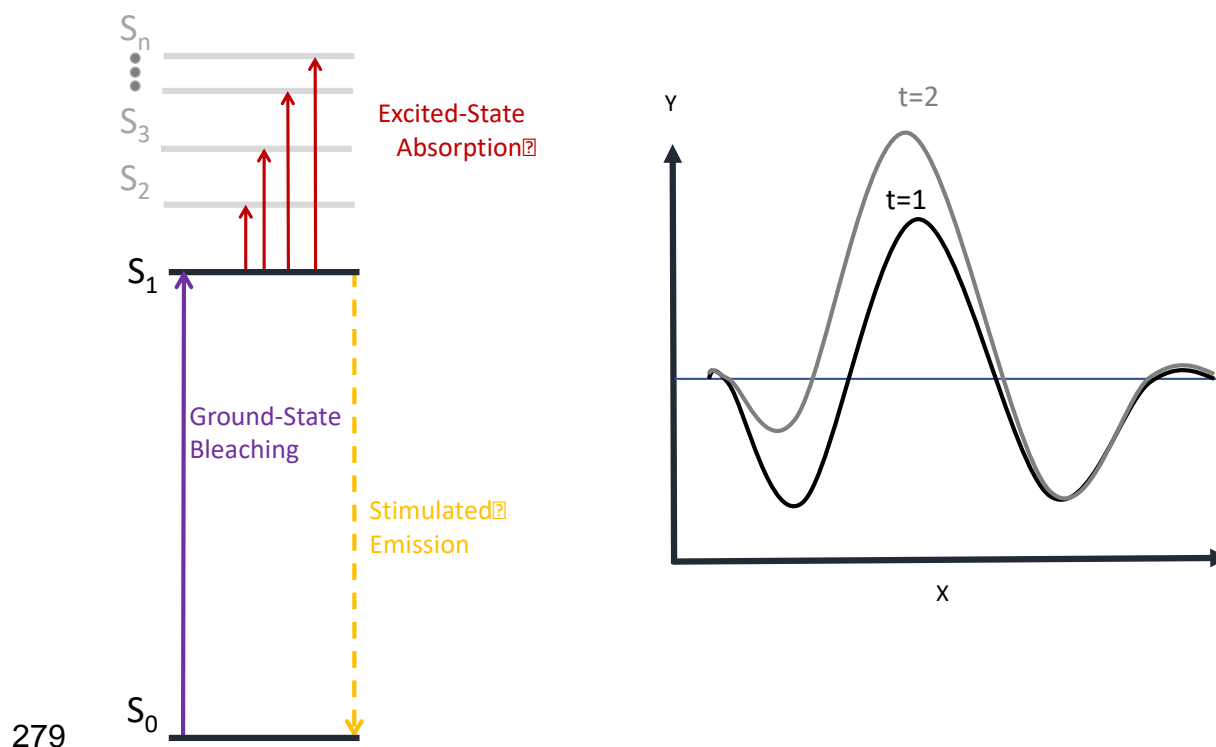
219

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

233 1.4. A Basic introduction to pump-probe spectroscopies

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

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



279

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

283

284 **1.5. Analytical tools for the detection of DNA damage**

285

286 A large fraction of the scientific advances discussed in the present volume relies on
 287 the accurate quantification of DNA damage. Those are either dimeric photoproducts
 288 (cyclobutane pyrimidine dimers, CPD; pyrimidine (6-4) pyrimidone photoproducts,
 289 64PP; Dewar valence isomers, Dew) or oxidation products (single strand breaks, SSB;
 290 8-oxo-7,8-dihydroguanine, 8-oxoGua). The following section describes the most
 291 widely applied techniques, including immunological detection, electrophoresis-based
 292 assays and chromatography combined with mass spectrometry. Some techniques will

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

297 **1.5.1 Antibody-based assays**

298 Immunological detection of dimeric photoproducts is a widely used approach in
299 photobiology. This class of techniques relies on the production of polyclonal or
300 monoclonal antibodies raised against UV-irradiated DNA or oligonucleotides ^{8, 9}.
301 Precise characterization works allowed several groups to produce antibodies specific
302 against each of the three main classes of bipyrimidine dimers, namely CPDs, 64PPs
303 and Dews. These antibodies are then used in different types of assays to reveal the
304 presence of photoproducts. These techniques, when applied to isolated or extracted
305 DNA, include among others radioimmunoassays, enzyme-linked immunosorbent
306 assay and immuno-dot-blot. In whole cells, flow cytometry has also been used. A
307 unique opportunity afforded by immunological assays is the possibility to localize the
308 photoproducts in tissues, for instance in skin, by using immunohistochemistry. A
309 limitation of immunological assays is that results are often obtained in arbitrary units
310 rather than in absolute frequency of damage. Another drawback is that, for a given
311 type of dimeric photoproduct, no difference is made between the TT, TC, CT and CC
312 derivatives. Antibodies are also used for the quantification of oxidized bases such as
313 8-oxoGua by ELISA. However, the accuracy and the specificity of the results has been
314 challenged.

315 **1.5.2 Electrophoresis-based assays**

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

331 **1.5.3 Chromatography-based assays**

332 Chromatographic approaches have also been designed to quantify DNA damage.
333 These techniques involve extraction of DNA, hydrolysis aimed at releasing damage
334 and separation of the hydrolysed DNA mixture on a chromatographic system with
335 appropriate detection ¹². This strategy is very specific and provide information on
336 individual damage, often simultaneously in a single analysis. In addition, because the
337 detection can be easily calibrated, results are quantitative. They are for example
338 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.

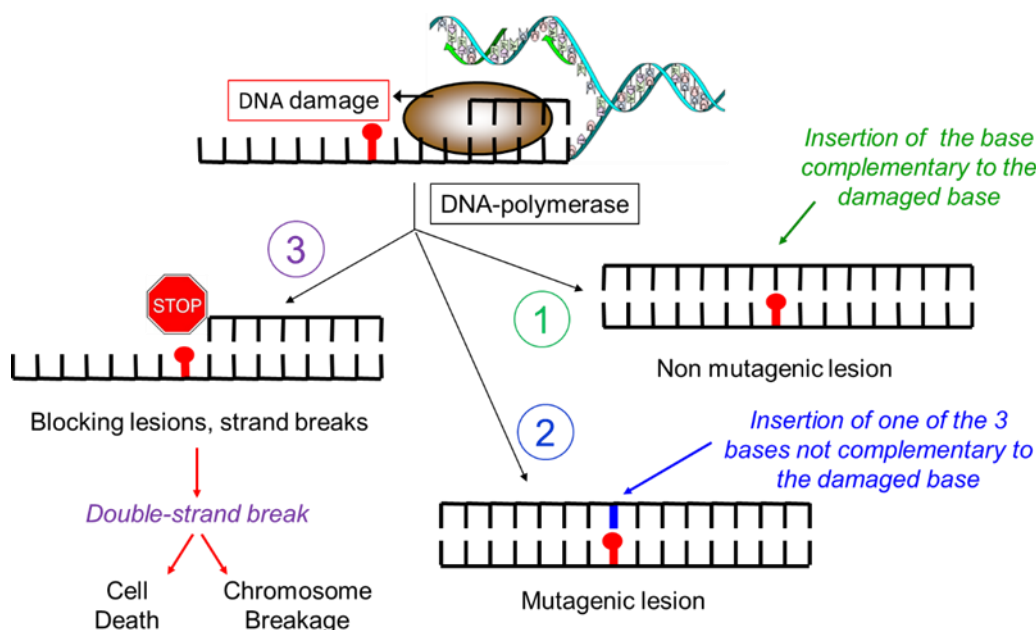
357

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

363 transcription¹³. Replication involves a complex series of enzymes that use each strand
364 of DNA as a template for the synthesis of a new double-strand. The synthesis of DNA
365 occurs in the replication fork, a structure where the two original strands are separated
366 and the replication machinery is bound to synthesize the new strands. Because DNA
367 synthesis always proceed from the 5'-end to the 3'-end, one of the original strand
368 (leading strand) is replicated in a continuous way by DNA-polymerases. The other
369 strand (lagging strand) has to be synthesized by fragments. Those are first initiated by
370 small transient pieces of RNA that are then extended into DNA Okasaki fragments.
371 RNA is removed and the DNA fragments are ligated to yield a continuous DNA strand.
372 The replication of these DNA strands rely on the formation of base pairs, and thus on
373 the complementarity between adenine and thymine on the one hand, and guanine and
374 cytosine on the other hand. Consequently, any modification of the chemical structure
375 of the bases, namely a damage, can interfere with this process. A first consequence
376 may be errors in the incorporated base opposite the damage upon replication and a
377 change in the local sequence of the DNA (Fig. 1.6). This constitutes a mutation. A
378 second possible effect is the blockage of the DNA-polymerases at the site of the
379 damage. Stalled replication fork are converted by cells into double strand breaks
380 (DSB) that can either be lethal or lead to chromosomal damage. It should be stressed
381 that all the following generations of cells will bear the same mutations. Indeed, even
382 at a mutated site, DNA is constituted of normal bases in correct base pairing and does
383 not represent an abnormal chemical structure that could be detected by the repair
384 machinery.



385

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

397

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

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

419

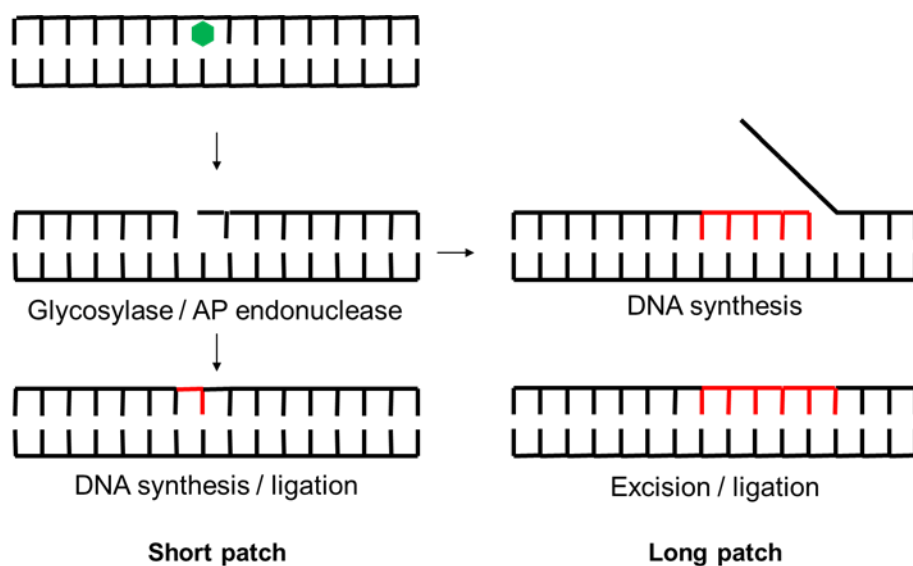
420 **1.6.2 Principle of DNA repair**

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

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

436 As shown in the following chapters, UV-induced DNA damage mostly involve
437 modification of the bases. In human cells, which do not contain photolyase activities,
438 those are handled by excision-resynthesis pathways. Two main pathways are
439 involved, namely base excision repair (BER) and nucleotide excision repair (NER).
440 BER is mostly active against small modified bases such as oxidation products¹⁶. The
441 key enzymes in BER are DNA N-glycosylases (Fig. 1.7). Each of these proteins
442 recognize a specific set of damaged bases (oxidized purines, oxidized pyrimidines,
443 etc.). They cleave the covalent bond between the base and the 2-deoxyribe unit (the
444 N-glycosidic bond) and leave an abasic site (AP sites). DNA is then cleaved at AP
445 sites by AP endonucleases or the initial DNA glycosylase when it is bifunctional. The
446 gap left in the double strand can be either filled by incorporation of one nucleotide
447 based on the complementarity with the undamaged strand. This is the short patch
448 pathway of BER. Alternatively, a small strand of DNA can be synthesized under the
449 damaged portion, which is then cleaved. This is known as the long patch pathway of
450 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

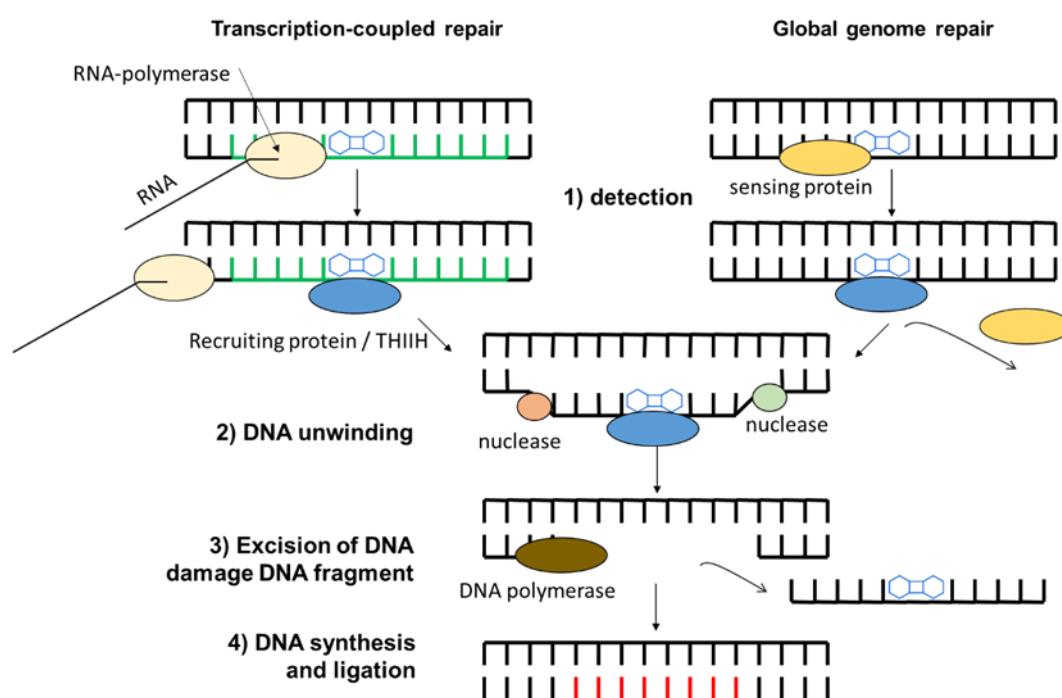
454

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 xeroderma 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

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



474
 475 *Figure 1.8: Basic steps in nucleotide excision repair. Transcription-coupled repaired*
 476 *handles photoproducts present in active genes and is activated by stalled RNA-*
 477 *polymerase. Global genome repair is initiated by the detection of damage by sensing*
 478 *proteins. In both pathways, NER then involves DNA unwinding, cleavage of a short*
 479 *DNA fragment bearing the photoproduct, and last synthesis and ligation of new DNA.*

480

481 Acknowledgments

482 RI thanks Dr. Lara Martinez-Fernández for helpful discussions and help with the
483 figures

484

485 ABBREVIATIONS

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

489

490

491 REFERENCES

492

- 493 1. P. W. Atkins and R. S. Friedman, *Molecular Quantum*, Oxford University Press,
494 Oxford, 2010.
- 495 2. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to*
496 *Advanced Electronic Structure Theory*, Dover Publications, Garden City, NY
497 11530 1996.
- 498 3. N. J. Turro, *Modern Molecular Photochemistry*, Univ Science Books,, Herndon,
499 VA 1991.
- 500 4. R. Improta, F. Santoro and L. Blancafort, *Chem. Rev.*, 2016, **116**, 3540-3593.
- 501 5. M. K. Shukla and J. Leszczynski, *Radiation Induced Molecular Phenomena in*
502 *Nucleic Acids: A Comprehensive Theoretical and Experimental Analysis*,
503 Springer Nature, Switzerland 2008.
- 504 6. E. H. Egelman, *Comprehensive Biophysics*, Elsevier, 2011.
- 505 7. G. Cerullo, C. Manzoni, L. Luer and D. Polli, *Photochem. Photobiol. Sci.*, 2007,
506 **6**, 135-144.
- 507 8. N. Kobayashi, S. Katsumi, K. Imoto, A. Nakagawa, S. Miyagawa, M. Furumura
508 and T. Mori, *Pigment Cell Res.*, 2001, **14**, 94-102.
- 509 9. D. Mitchell and B. Brooks, *Photochem. Photobiol.*, 2010, **86**, 2-17.
- 510 10. E. Cordelli, M. Bignami and F. Pacchierotti, *Toxicol. Res.*, 2021, **10**, 68-78.
- 511 11. D. Muruzabal, A. Collins and A. Azqueta, *Food and Chemical Toxicology*, 2021,
512 **147**.
- 513 12. T. Douki, *Photochemical and Photobiological Sciences*, 2013, **12**, 1286-1302.
- 514 13. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular*
515 *Biology of the Cell. 4th edition.*, Garland Science, New York, 2002.

- 516 14. L. B. Alexandrov, J. Kim, N. J. Haradhvala, M. N. Huang, A. W. Tian Ng, Y. Wu,
517 A. Boot, K. R. Covington, D. A. Gordenin, E. N. Bergstrom, S. M. A. Islam, N.
518 Lopez-Bigas, L. J. Klimczak, J. R. McPherson, S. Morganella, R. Sabarinathan,
519 D. A. Wheeler, V. Mustonen, P. M. S. W. Group, G. Getz, S. G. Rozen, M. R.
520 Stratton and P. Consortium, *Nature*, 2020, **578**, 94-101.
- 521 15. E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J.
522 Humphray, C. D. Greenman, I. Varela, M. L. Lin, G. R. Ordonez, G. R. Bignell,
523 K. Ye, J. Alipaz, M. J. Bauer, D. Beare, A. Butler, R. J. Carter, L. Chen, A. J.
524 Cox, S. Edkins, P. I. Kokko-Gonzales, N. A. Gormley, R. J. Grocock, C. D.
525 Haudenschild, M. M. Hims, T. James, M. Jia, Z. Kingsbury, C. Leroy, J.
526 Marshall, A. Menzies, L. J. Mudie, Z. Ning, T. Royce, O. B. Schulz-Trieglaff, A.
527 Spiridou, L. A. Stebbings, L. Szajkowski, J. Teague, D. Williamson, L. Chin, M.
528 T. Ross, P. J. Campbell, D. R. Bentley, P. A. Futreal and M. R. Stratton, *Nature*,
529 2010, **463**, 191-196.
- 530 16. D. O. Zharkov, *Cellular and Molecular Life Sciences*, 2008, **65**, 1544-1565.
- 531 17. M. Fousteri and L. H. Mullenders, *Cell Research*, 2008, **18**, 73-84.
- 532 18. J. E. Cleaver, *Advances in Experimental Medicine and Biology*, 2008, **637**, 1-
533 9.
- 534