Mechanisms of DNA Repair by Photolyase and Excision Nuclease

Nobel Lecture, December 8, 2015

by Aziz Sancar

Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA.

SUMMARY

The ultraviolet (UV) wavelengths in sunlight damage DNA by converting two adjacent thymines into a thymine dimer (T<>T) which is potentially mutagenic, carcinogenic, or lethal to the organism (Fig. 1). This damage is repaired by photolyase in *E. coli* and by the nucleotide excision repair system in *E. coli* and in humans. In this lecture I will present our work on photolyase and nucleotide excision repair, and I will conclude my talk by describing how our research on photolyase led to the discovery of an essential circadian clock protein, called cryptochrome, that links these two research subjects to one another and thus completes the circle.

PHOTOLYASE

Photolyase is a photon-powered nanomachine that uses blue light photons to repair thymine dimers that are induced in DNA by UV. Photolyase was discovered by my Ph.D. mentor Claud S. (Stan) Rupert (Fig. 1) in 1958, and this discovery marked the beginning of the field of DNA repair as a scientific discipline. Decades before the discovery of photolyase, it had been known that UV kills bacteria very efficiently (Fig. 1, right panel). In 1949 Albert Kelner, of Cold Spring Harbor, made the interesting observation that if bacteria killed by UV
The Nobel Prizes

were exposed to visible light, the dead bacteria were miraculously brought back to life [1, 2]. However, he had no explanation for this phenomenon, which was termed photoreactivation. Rupert analyzed this phenomenon further. He demonstrated that UV killed bacteria by damaging their DNA, and that there is an enzyme (photoreactivating enzyme = photolyase) that uses the blue light energy in visible light to repair DNA damage. Blue light thus brings dead cells back to life, demonstrating that this resurrection from the dead was not a miraculous phenomenon that needed a metaphysical explanation [3, 4], but could instead be explained by the laws of physics. The reaction mechanism that Dr. Rupert developed is as follows [5–7]: UV converts two adjacent pyrimidines, including thymines, to a CPD (cyclobutane pyrimidine dimer), and there is an enzyme called photolyase that uses blue light energy to break the two abnormal bonds joining

FIGURE 1. Photoreactivation and photolyase. Top panel (Left): Claud S. Rupert and Aziz Sancar at a function at the University of Texas at Dallas in 2009. (Right): Photoreactivation in E. coli. An E. coli strain defective in nucleotide excision and recombination repair and carrying the cloned photolyase gene was irradiated with the indicated UV dose and either plated directly (closed circles) or plated after exposure to a camera flash of 1 millisecond (open circles). Bottom panel: General Model for Photolyase based on the pioneering work of Rupert: UV induces the formation of a cyclobutane thymine dimer (T<>T), photolyase binds to the dimer, absorbs a blue light photon, and converts the dimer to two canonical thymines.
Mechanisms of DNA Repair by Photolyase and Excision Nuclease

the thymines and thus converts the thymine dimer to two normal thymines (Fig. 1, bottom panel). Photolyase therefore repairs DNA and eliminates the harmful effects of UV. While this was a satisfactory explanation of the photoreactivation phenomenon, it raised a physical question: Photolyase is a protein, and proteins do not absorb blue light. Therefore, for the next two decades Rupert and many other researchers attempted to identify the blue light-absorbing component of photolyase. They were unsuccessful because Rupert had determined that an *E. coli* cell contains only 10–20 molecules of photolyase, and this made it virtually impossible to purify and characterize the enzyme.

In 1974 when I joined Dr. Rupert’s lab, gene cloning had just been developed at Stanford University. As a fresh graduate student, I thought I could do anything I wanted, and therefore I proposed to Dr. Rupert to clone the photolyase gene, overproduce the enzyme, and purify it. He said, “Go ahead.” After months of work, I successfully cloned the gene [8, 9]. An electron micrograph of the plasmid containing the photolyase gene is shown in Fig. 2 (left). In subsequent years at the University of North Carolina, my colleagues and I used the cloned gene to purify the enzyme in gram quantities (Fig. 2, middle) [10–12], and while purifying it we found that it has a bright blue color (Fig. 2, right) [12]. That finding, without any chemical analysis, answered the physical question: It has a blue color which means it absorbs light. We proceeded to identify the light-absorbing component of the enzyme using analytical chemistry, and to our surprise we found that it contained not one, but two blue light-absorbing cofactors, which

![Electron micrograph of the plasmid containing *Phr*](image1.png)

**Figure 2.** Cloning and purification of photolyase. **Left:** Electron micrograph of the plasmid carrying the photolyase gene (*Phr*). **Middle:** Purification of photolyase from an *E. coli* strain overproducing the protein. **Right:** Purified photolyase protein has bright blue color.
The Nobel Prizes

are methenyltetrahydrofolate (folate) and two-electron reduced and deprotonated flavin adenine dinucleotide (FADH–) [13–28]. Moreover, we found that the enzyme exhibits colors ranging from purple to orange depending on the redox status of the flavin cofactor [29] (Fig. 3). We next determined the functions of the two cofactors by carrying out photochemical experiments. We found that the folate acts like a solar panel, absorbing light and transferring the excitation energy to FADH– [15–17, 24]. The flavin is the actual catalyst, and upon excitation by energy transfer from folate (and less efficiently by direct absorption of a photon) it carries out the repair reaction on the CPD by a radical mechanism through a cyclic redox reaction [24, 26].

To provide structural basis for the proposed reaction mechanism, we collaborated with Johann Deisenhofer to crystallize photolyase and obtain the 3-D structure of the enzyme [30], which is shown in Fig. 4 in ribbon diagram and surface charge representations. As predicted from the biochemical experiments, the folate is like a solar panel, where it sits on the roof of the enzyme, absorbs light, and then transfers the light energy to the flavin cofactor within the core the enzyme to carry out catalysis. With this general structural view, then, the mechanism of photolyase was developed [31] (Fig. 5): Photolyase binds DNA containing a CPD because the T<>T distorts the backbone of the DNA. Upon

![Figure 3](image_url)
**Mechanisms of DNA Repair by Photolyase and Excision Nuclease**

**Figure 4.** Crystal structure of *E. coli* photolyase. **Left:** Ribbon diagram representation. **Right:** surface potential representation. Positively and negatively charged residues are highlighted in blue and red, respectively. The phosphodiester backbone of the damaged strand binds to the positively charged diagonal groove on the enzyme surface. The dashed box marks the hole leading to the FADH$^-$ catalytic cofactor.

**Figure 5.** Reaction mechanism of photolyase. The enzyme makes ionic bonds with the phosphate residues of the damaged DNA strand and flips out the thymine dimer dinucleotide into the active site cavity so that the T$>$T is within Van der Waals contact with FADH$^-$. The catalytic reaction is initiated by absorption of a photon (300–500 nm) by the folate (MTHF). The MTHF excited singlet state transfers the excitation energy to FADH$^-$ by Forster Resonance Energy Transfer (FRET). The excited FADH$^-$* splits the cyclobutane ring by cyclic redox reaction to convert T$>$T to T-T, and the repaired DNA dissociates from the enzyme. The inset shows the distances between the indicated atoms of FADH$^-$ and the cyclobutane pyrimidine dimer (CPD). Image courtesy of Dongping Zhong.
binding to damaged DNA, ionic interactions between the positively charged groove on the photolyase surface and negatively charged DNA phosphodiester backbone the enzyme pulls the T<>T out from within the helix and into the core of the enzyme so that the T<>T is within Van der Walls contact with FADH⁻. It makes a very stable complex, and nothing happens until folate absorbs a photon and transfers the excitation energy to the flavin cofactor. The excited state flavin, FADH⁻*, repairs the T<>T by a cyclic redox reaction, and then the enzyme dissociates from the DNA to go on in search of other damage sites to carry out the repair reactions again.

**Figure 6.** Microscopic rate constants for photolyase. The rate constants were determined by ultrafast time-resolved absorption and fluorescence up-conversion spectroscopy. The cleavage of the cyclobutane ring is by a concerted asynchronous mechanism which couples the cleavage of the C5-C5 bond in less than 10 ps to cleavage of the C6-C6 bond in 90 ps. The entire photochemical reaction is complete in 1.2 ns, with an overall quantum yield of ~0.9. The inner circle shows the relative locations of the photoantenna (MTHF), the catalyst (FADH⁻) and the thymine dimer substrate (CPD). Image courtesy of Dong-ping Zhong.

Over the past decade we have collaborated with Dongping Zhong of Ohio State University to determine the microscopic rate constants of DNA repair by photolyase. We have determined the rates of energy transfer, electron transfer, bond breakage, bond forming and electron return, in real time and at picosecond resolution [29–35] (Fig. 6). The entire catalytic cycle is complete in 1.2 ns, and the enzyme repairs \( T<>T \) with a quantum yield of 0.9 [29, 31, 34]. Photolyase is currently one of the best understood enzymes.

**NUCLEOTIDE EXCISION REPAIR**

**Excision repair in *E. coli***

The work on photolyase, in addition to its intrinsic value, contributed to the discovery of the other major DNA repair mechanism found in nearly all cellular organisms: Nucleotide excision repair (excision repair). In early work on photolyase, *E. coli* cells were irradiated with UV in a suspension in a buffer, and then one half was exposed to blue light while the other half was kept in the dark. It was found that the UV-induced \( T<>Ts \) disappeared from the genome of the blue light-exposed cells, but remained unchanged in the genome of the control cells kept in the dark. However, if the same experiment was carried out in a buffer containing glucose as an energy source, incubation of UV-irradiated *E. coli*, in either the dark or the light, resulted in the disappearance of \( T<>Ts \) from the genome [36, 37]. Nevertheless, there was still a fundamental difference between the two sets of cells. In light-exposed cells, the \( T<>Ts \) completely disappeared as expected. In contrast, in cells kept in the dark, even though the \( T<>Ts \) disappeared from the genomic DNA, they accumulated quantitatively in the cytosol [36–38]. This finding in 1964 by Paul Howard-Flanders [36] at Yale University and Richard Setlow [37] at Oak Ridge National Laboratory led to the concept of nucleotide excision repair. After these initial findings, research done in numerous labs led to the conclusions summarized in Fig. 7: \( T<>Ts \) are removed (excised) from the genome in both *E. coli* and humans [39, 40] in the form of 4–6 nucleotide-long oligomers [36, 37, 40–42] but remain within the cell and are not exported. The excision reaction is genetically controlled by the *uvr* genes in *E. coli* [43] and *XP* genes in humans [44, 45]. Following excision, the repair gap is filled in and ligated [39, 46]. The consensus model for nucleotide excision repair over the period of 1964–1982 was the so-called ‘cut-and-patch’ mechanism [47], whereby an endonuclease controlled by the Uvr proteins in *E. coli*, and XP proteins in humans, made an incision 5’ to the \( T<>T \), and then an exonuclease removed the \( T<>T \) in a reaction coupled with repair synthesis.
The Nobel Prizes

Figure 7. General model for nucleotide excision repair for *E. coli* and humans that was developed over the period of 1964–1982. It is referred to as the endonuclease/exonuclease coupled with repair synthesis or the cut-and-patch model.

- Thymine dimers are removed from the genome in both *E. coli* and humans.
- Excised thymine dimers were reported to exist in oligonucleotides 4-6 nt in length.
- Excision is genetically controlled by *uvr* genes in *E. coli* and *XP* genes in humans.
- Following excision, the repair gap is filled in and ligated.
- Excised dimers remain within the cell.

**Figure 8.** Identification of *E. coli* Uvr proteins. In the maxicell method, a *recA*-*uvrA* mutant strain containing a plasmid carrying the gene of interest is irradiated with a moderate UV dose that hits the chromosomal DNA at multiple sites but not the much smaller (typically 500–1000 fold) plasmid. This causes total degradation of the chromosomal DNA in 6–12 hours leaving cells (maxicells) with only plasmid DNA. At this point, addition of the 35S-methionine radiolabel to the medium labels only the plasmid encoded proteins which can be detected by autoradiography. This is an autoradiogram of three *E. coli* maxicells expressing UvrA, UvrB, and UvrC, respectively. The lower molecular weight bands are proteins encoded by the drug resistance genes, tetracycline and ampicillin.
that filled in the single-stranded gap, followed by ligation of the repair patch to complete the repair process.

In 1977, I joined the laboratory of Dean Rupp at Yale University to work on the mechanism of nucleotide excision repair in *E. coli*. I invented the maxicell method to specifically radiolabel plasmid-encoded proteins [48], which enabled me to identify and clone the three genes implicated in excision repair: *uvrA*, *uvrB*, and *uvrC* [49–51]. Nothing was known about the specific functions of these genes at the time. I found that the three genes encoded proteins of 100 kDa, 85 kDa, and 66 kDa, respectively (Fig. 8). With the aid of the maxicell method, I then purified the three proteins in milligram quantities (Fig. 9) and investigated their effects on UV-damaged DNA. I found that the UvrA, B, C proteins repaired DNA by a mechanism different from the classic endonuclease/exonuclease (cut-and-paste) model: The three proteins instead act together to carry out concerted dual incisions at precise distances from the photoproduct [52], seven nucleotides 5' and three nucleotides 3' from the damage to generate a dodecamer (12-mer) carrying the T<>T photoproduct. The excised oligomer is then removed from the duplex, and the resulting gap is filled and ligated (Fig. 10). Later on at the University of North Carolina, my colleagues and I investigated the roles of the three proteins in the repair reaction. We found that UvrA and UvrB are ATPases and that UvrC is a nuclease, and we carried out detailed biochemical studies to develop the reaction mechanism shown in Fig. 11 [53–68]: UvrA recognizes the damage and recruits UvrB to the damage site, which promotes the formation of

![Figure 9](image_url)
a stable UvrB-DNA complex in an ATP hydrolysis-dependent reaction. UvrA then disassociates from the complex, and UvrB recruits UvrC to the damage site. UvrC has two nuclease active sites, which make the 5’ and 3’ incisions in a concerted manner [69, 70]. UvrC and the excised oligomer are then released from the duplex by the action of the UvrD helicase [61]. Finally, DNA Polymerase I displaces UvrB and fills in the gap, and the repair patch is sealed by DNA ligase [67].

Transcription-coupled repair

While we were characterizing the reaction mechanism of the *E. coli* excision nuclease (excinuclease), Philip Hanawalt of Stanford University reported that transcription strongly stimulated repair *in vivo* in both mammalian cells and in *E. coli* [71–73]. It was proposed that this was the consequence of RNA polymerase accelerating the rate of the damage recognition, which is the rate limiting step in nucleotide excision repair. We therefore initiated a project to study the mechanism of *E. coli* transcription-coupled repair in a defined system using a labeled DNA damage substrate, purified RNA polymerase, and the UvrA, B, and C proteins. We found that DNA damage blocked the progression of RNA polymerase, as predicted, and led to the formation of a very stable RNA polymerase

![Figure 10](image1.png)

**Figure 10.** Excision by dual incisions in *E. coli*. UvrA + UvrB + UvrC proteins in the presence of ATP + Mg²⁺ incise 7 nucleotides 5’ and 3–4 nucleotides 3’ to the thymine dimer (T<>T) as shown in the 3-D model (left) and line diagram representation (right). The 12–13-nucleotide gap is filled in by polymerases and ligated.
Mechanisms of DNA Repair by Photolyase and Excision Nuclease

elongation complex at the damage site. However, contrary to expectations that this stalled complex would accelerate repair by constituting a large target for the repair enzyme, we found instead that stalled RNA polymerase did not stimulate but actually inhibited repair, presumably by interfering with the access of UvrA, B, C to the damage site [74]. We reasoned that there must be a missing factor that performs two functions. First, it overcomes the repair inhibitory effect of stalled RNA polymerase, and second, it helps in recruiting UvrA, B, and C to the damage site to accelerate the repair rate. We developed an in vitro biochemical assay to purify this coupling factor. We succeeded in purifying a protein of 130 kDa that performed both functions [75–83]. We named the protein transcription-repair coupling factor (TRCF). Using an entirely defined system, we elucidated the reaction mechanism of transcription-coupled repair in E. coli (Fig. 12): TRCF is a translocase that recognizes stalled RNA polymerase and displaces it from the damage site while concomitantly recruiting UvrA to accelerate the repair rate.

**Figure 11.** Reaction mechanism of excision repair in *E. coli*. The damage is recognized by the (UvrA)$_2$ homodimer which functions as a molecular matchmaker to recruit UvrB to the damage site. An ATP hydrolysis-dependent reaction then promotes the formation of a very stable UvrB-DNA complex. This complex recruits UvrC, which incises 5’ and 3’ to the damage due to active site nucleases within the N-terminal and C-terminal halves of the protein. UvrC and the excised dodecamer (12-mer) are then displaced by the UvrD helicase, and UvrB is displaced by DNA Polymerase I as it fills in the gap. The gap is then sealed by ligase.
We also showed that TRCF was encoded by the \textit{mfd} gene (\textit{mutation frequency decline}) that was discovered by Evelyn M. Witkin in 1966 as a gene responsible for preventing a specific type of UV-induced mutagenesis. The discovery of the equivalency of TRCF and Mfd provided a mechanistic explanation for the \textit{mfd} \textsuperscript{-} phenotype that had remained mysterious for 25 years [84].

\textbf{Excision repair in humans}

Having described the \textit{E. coli} excision repair mechanism in some detail in 1987, we started to work on human excision repair. At that time, it was presumed that
human excision repair proceeded by an 5’ endonuclease/exonuclease mechanism (cut-and-patch), as in the classical *E. coli* model [47] (Fig. 7). Importantly, James Cleaver of the University of California, San Francisco had discovered in 1968 that patients with the hereditary disease xeroderma pigmentosum (XP) were defective in excision repair [44]. These patients are extremely sensitive to sunlight and exhibit ~5,000-fold increase in sunlight-induced skin cancer relative to individuals with normal excision repair (Fig. 13). Genetic analysis of XP patients revealed that seven genes, termed XPA through XPG, were responsible for the removal of UV-induced photoproducts [45]. Using our expertise from working on *E. coli* excision repair, we proceeded to characterize human excision repair [85]. From the very beginning, we found that, as in *E. coli*, nucleotide excision repair was carried out in humans by dual incisions and not by the conventional model [86–91]. Beyond that, the human excision repair mechanism turned out to be very different from *E. coli* excision repair [87–126]. To begin with, in humans, not just three proteins (UvrA, B, C), but 16 proteins in six repair factors [93, 97] were necessary for making the dual incisions (Fig. 14). Furthermore, these human proteins are not evolutionarily related to the *E. coli* excision repair proteins. Secondly, even though in principle both *E. coli* and humans carry out nucleotide excision repair by dual incisions, the dual incision
Human excision repair factors. Six repair factors encompassing 16 proteins are needed for making the dual incisions in humans. Note that RPA also functions in replication and recombination. TFIIH, which contains eight other subunits in addition to the XPB and XPD helicases encoded by the respective XP genes, is a general transcription factor for the initiation of transcription by RNA Polymerase II. These human excision repair proteins are not evolutionarily related to prokaryotic excision repair proteins. The figure shows purified repair factors separated by SDS-PAGE and silver stained.

Mechanisms are quite different. Whereas in *E. coli* the 5’ incision is seven nucleotides away from the damage, and the 3’ incision is three nucleotides away from the damage, the human excision repair system incises the damaged strand 20–22 nucleotides 5’ and five nucleotides 3’ to the damage to release an excised oligomer of 27–30 nt in length [86], in contrast to the 12–13 nt oligomer generated by *E. coli* dual incisions (Fig. 15). Finally, the actual damage recognition and processing is also different: Whereas damage is recognized by UvrA in *E. coli*, damage is instead recognized in humans by RPA, XPA, and XPC, followed by recruitment of TFIIH, which contains the XPB and XPD helicases that unwind the helix and recruit the XPG and XPF nucleases to make the 3’ and 5’ incisions. The dual incision event is followed by the release of the 30-nucleotide excised oligomer, gap filling and ligation by DNA polymerase and ligase to produce a 30-nucleotide repair patch [102]. Fig. 16 summarizes our current model for the mechanism of human nucleotide excision repair. Most recently, this work led us to study other cellular responses to DNA damage including the DNA damage checkpoints [127–130].
**Figure 15.** Excision by dual incisions in humans. In humans, thymine dimers (T<>T) and other bulky base adducts are removed by dual incisions located 20±5 phosphodiester bonds 5’ and 6±3 phosphodiester bonds 3’ to the damage, which releases an oligonucleotide 24–32 nt in length (referred to as nominal 30-mers). **Left panel:** Dual incision sites on a 3-D representation of DNA. **Right panel:** Schematic of human dual incisions followed by repair synthesis and ligation.

**Figure 16.** Reaction mechanism of the human excision nuclease system. The damage is recognized by cooperative interactions of RPA, XPA, and XPC followed by recruitment of TFIIH. The helicase activity of TFIIH provides the major specificity by kinetic proofreading and results in formation of a tight complex from which XPC is ejected. Note that XPC plays a role similar to that of *E. coli* UvrA by playing an essential role in damage recognition, leaving the complex to allow subsequent steps to proceed (molecular matchmaker). Concomitant with the dissociation of XPC, the XPG and XPF nucleases are recruited to make the 3’ and 5’ incisions in a concerted reaction. The excised “30-mer” is released in a tight complex with TFIIH. The excision gap is filled in by DNA polymerases and ligated to produce a 30 nucleotide-long repair patch.
Our discovery of excision of a nominal 30-mer by the human nucleotide excision repair system in cell extracts and with the reconstituted enzyme system was confirmed by other groups. However, these findings were at odds with numerous reports that the excised T<>T’s were in the form of oligomers 4–6 nucleotides in length in human cells [40–42]. This discrepancy between the *in vivo* and *in vitro* data remained unresolved for two decades. The solution to the problem came from the analysis of the fate of the excised 30-mer *in vitro* using cell-free extracts and reconstituted repair reactions. We found that the excised 30-mer released from the duplex is in a tight complex with TFIIH *in vitro* reactions [131, 132]. We reasoned that this may be the case *in vivo* as well. To test this prediction, we irradiated human cells and after incubating for 1–6 hours to allow for repair, we lysed the cells, immunoprecipitated TFIIH, and analyzed the DNA fragments associated with TFIIH. We found that the excision products generated in vivo were in fact 30-nucleotides in length, as in the case of the *in vitro* reaction [133–137]. Upon longer incubation, the primary excision product is degraded to smaller fragments less than 10 nucleotides in length, which explained the previous *in vivo* studies in which the excised oligonucleotides were typically isolated from the cells 24 hours after irradiation.

Our ability to isolate the primary excision product not only solved the apparent discrepancy between the *in vivo* and *in vitro* excision reactions, it also provided a means for generating a repair map of the entire human genome [138] (Fig. 17). Following irradiation of cells with UV and incubation for a period of time to allow for repair, we then lysed cells and immunoprecipitated TFIIH to isolate the associated excised oligomers. The excised oligomers are then sequenced using Next Generation Sequencing (NGS). In a typical experiment we obtain 15–20 million reads. We align these reads to the human genome to place all of the excision products at specific locations, thus generating a repair map. Fig. 18 shows the repair map of the 22 somatic and 2 sex chromosomes of a male individual. The black tracks represent transcription and the green tracks represent the repair tracks of the two photoproducts for the two strands of the entire genome. This figure is a screenshot of the repair map of the entire genome, and is meant to illustrate the coverage of repair over the whole genome. However, it does not reveal much information about the determinants of repair mode and rate at a given locus. By concentrating on one specific chromosome at various resolutions, the information contained within this map becomes apparent. Fig. 19 shows the repair map of chromosome 17. This chromosome is 83 megabases in length and carries the \(p53\) gene, which is mutated in about 50% of human cells.
The map shows the transcripts along the entire chromosome in both strands in black and the excision repair (XR-seq) tracks for both strands in green. The repair map is a map in the true, geographic sense of the word. It has mountains, it has valleys, and it has canyons, meaning there are regions of high repair, low repair and no repair at all. Importantly, with such a map, we can answer the question of the repair mode and efficiency at any given nucleotide in the genome. As an example, Fig. 19 shows the p53 transcription and repair maps at kilobase resolution (middle). Finally, at single nucleotide resolution (bottom) the map shows the repair efficiency and the mode of repair of a p53 mutation hotspot.
Figure 18. Excision repair map of the entire human genome. The locations of the XR-seq signals for CPD and (6-4)PP in both strands of the duplex across all chromosomes of the human NHF1 cell genome (male) are indicated by green tracks. The ENCODE total stranded RNA-seq tracks in black are plotted on top of the XR-seq tracks for comparison. Chromosome 17, which is boxed in red, carries the p53 gene which is mutated in nearly 50% of cancers.

at T-T (7,577,150-7,577,151) dinucleotide position. The thymine dimer at this position is removed by incision 20 nucleotides 5’ and 4 nucleotides 3’ to the photoproduct. It is evident that much more information can be gathered from this map regarding the determinants of repair of UV damage at any given location of the genome. More importantly, nucleotide excision repair also repairs the DNA damage caused by the major anticancer drug, cisplatin. We are currently generating a cisplatin damage repair map of the genomes of normal and cancerous human cells that we hope will have some implications for cancer treatment.

To summarize (Fig. 20) our work on nucleotide excision repair [67, 68, 82, 83, 120, 121, 124]: Repair is initiated by dual incisions both in E. coli and in humans, which generates 12-13-mers in E. coli and ~30-mers in humans. The dual incisions require UvrA, B, and C proteins in E. coli and six repair factors encompassing 16 proteins, including the proteins encoded by the XPA through the XPG genes, in humans. Following excision, the gap is filled in by DNA polymerases and ligated to generate repair patches of 12-13 and ~30 nucleotides in
Mechanisms of DNA Repair by Photolyase and Excision Nuclease

Figure 19. Excision repair at single nucleotide resolution. The transcription and repair maps of chromosome 17 are shown in an XP-C mutant cell line, which can only carry out transcription-coupled repair, to illustrate the dramatic effect of transcription on repair. The red line indicates the position of the mutation hotspot in the p53 gene. **Top:** Transcription and repair map at megabase resolution. **Middle:** Transcription and repair maps at kilobase resolution. Note the strong repair signal in the transcribed strands of the p53 and WRAP53 genes with a nearly absolute lack of repair in the non-transcribed strand. **Bottom:** The repair pattern of T<>T dinucleotide at a mutagenic hotspot, position 7,577,150-7,577,151. The photodimer is removed in the form of a 26-mer by dual incisions 20 nucleotides 5' and 4 nucleotides 3' to the dimer.

- Nucleotide excision repair is initiated by dual incisions in both *E. coli* and humans.
- Excision is genetically controlled by the evolutionarily unrelated *uvr* genes in *E. coli* and XP genes in humans.
- Dual incisions remove an oligomer of ~12 nucleotides in *E. coli* and ~30 nucleotides in humans.
- Following excision, the repair gap is filled in and ligated.
- By capturing the excised oligomers, we have generated an excision repair map of the whole human genome.

Figure 20. Excision repair in *E. coli* and humans. In both organisms, excision is by dual incisions. However, the proteins required for the dual incisions, the mechanisms for damage recognition, and the dual incision patterns are entirely different.
E. coli and humans, respectively. Finally, by capturing the excised 30-mer generated by human nucleotide excision repair in vivo, we have generated the excision repair map of UV damage of the whole human genome.

**CRYPTOCHROME, CIRCADIAN CLOCK, AND CLOSING THE CIRCLE**

I wish to conclude this presentation by explaining how our work on photolyase led to the discovery of cryptochrome as an essential component of the mammalian circadian clock, and how the circadian clock regulates nucleotide excision repair in mammals, thus linking the two subjects of our long-term research projects, photolyase and nucleotide excision repair.

**Discovery of mammalian cryptochrome**

I have discussed the excision repair mechanisms in both humans and E. coli, although I only presented photolyase data for E. coli. This is because humans do not have photolyase [139]. In fact, for 30 years after the discovery of photolyase in E. coli, its presence in humans was a matter of controversy. Some investigators reported that photolyase was not detectable in human cell lines, while others reported robust photolyase activity in human cells and reported purification of the enzyme to homogeneity from human white blood cells [140]. Having developed very sensitive and specific assays for photolyase in the 1980s, we decided to resolve this controversy. We conducted a comprehensive search for photolyase in freshly isolated human white blood cells. We detected no photolyase activity and then published a paper in 1993 [139] categorically stating that humans do not have photolyase (Fig. 21). However, two years later, in one of the first public releases of the human genome project, one of the ESTs (expressed sequence tags) was listed as the photolyase homolog [141]. Reasoning that we may have missed the photolyase activity in our earlier work, we decided to investigate the function of the gene by obtaining the entire cDNA clone, expressing it, and analyzing its function. While this work was in progress we discovered a second photolyase “homolog” in the human genome. We obtained the entire cDNAs of both genes. They are remarkably similar to E. coli photolyase at the sequence level and equally remarkable at the 3-D structure level (Fig. 21). We expressed and purified the proteins encoded by these genes and established that they had no photolyase activity and concluded that they were photolyase paralogs. This work was completed in April 1996, and not knowing what functions of these paralogs might be, we were reluctant to publish our findings (Fig. 22). In May
1996, I made my annual pilgrimage to Turkey to visit my family. On my return trip I read an article on jetlag in the airline magazine entitled “Internal Time-keeping,” by William Schwartz [142]. I believe this was the first time I learned the meaning of the phrase “circadian clock.” The article, among other things, noted that the circadian clock was synchronized to the physical clock by light, and was particularly sensitive to blue light. After reading this article, I suspected that the human photolyase paralogs might be clock proteins that sense blue light. Upon returning to the lab, I suggested to my coworkers that we publish our data and that we name the human photolyase paralogs cryptochromes in analogy with the plant blue light photoreceptors with sequence similarity to photolyase [143]. I also proposed that we suggest that the human cryptochromes (CRY1 and CRY2) are circadian clock proteins [144]. We wrote the paper and submitted it for publication in August, and it was published in November 1996 [144]. We then proceeded to test this claim [145] by knocking out the CRY genes in mice and testing them for circadian clock function.

**FIGURE 21.** Photolyase-cryptochrome connection. **Top:** Key papers leading to the discovery of the human cryptochromes. **Middle:** Sequence similarities among E. coli photolyase and human cryptochromes 1 and 2. **Bottom:** Three-dimensional structures of E. coli photolyase and Arabidopsis cryptochrome 1 showing the similarities in the photolyase homology domain. **Side panel:** Purified E. coli photolyase and hCRY1 and hCRY2 analyzed by SDS-PAGE followed by Coomassie Blue staining.
Before presenting the data on the mouse CRY knockouts, I will briefly summarize what the circadian clock is (Fig. 23) [146–149]: The clock, in general, is a timekeeping object/system. The circadian clock is similar to the clocks we

- **Clock is a Time Keeping Object/System**
  - Mechanic
  - Electronic
  - Molecular (Circadian Clock)

- **Circadian Clock** is an innate timekeeping molecular mechanism that maintains daily rhythmicity in biochemical, physiological and behavioral functions independent of external input.

**FIGURE 22.** Sequence of events leading to discovery of human CRY as a circadian clock protein. Although the prediction was that CRY was the blue light sensor of the mammalian clock, genetic analysis revealed it to be an essential cog in the core clock machinery.

**FIGURE 23.** Clock and circadian clock. The circadian clock, like mechanical and electronic clocks, measure time independent of external stimuli, but it is instead made up of molecules rather than gears and levers or electronic circuits.
Mechanisms of DNA Repair by Photolyase and Excision Nuclease

are familiar with, including mechanical and electrical clocks that are based on mechanical and electronic principles. The circadian clock has the same kind of design except that the components that make up the clock are molecules, and the function of the clock is to inform us of the time of the day. By doing so, the circadian clock maintains daily rhythmicity in biochemical, physiological, and behavioral functions of the organism even in the absence of external input.

Role of cryptochrome in the circadian clock

To test whether CRYs are circadian clock proteins, we generated mice with mutations in either CRY1 or CRY2, or both, and analyzed their circadian clock by

FIGURE 24. Behavioral analyses of cryptochrome knockout mice. Mice of the indicated genotypes were kept in cages with running wheels for 28 days and their activity profiles were recorded (actogram). The rpm of the running wheel is plotted on the y axis, and the time of the day is plotted on the x axis. The bar on top indicates the dark and light phases of the day. On the 7th day, indicated by arrows, the mice were switched from 12 hr light: 12 hr dark (LD) cycle into constant darkness (DD). Note that under LD all 4 mice exhibit similar activity and rest phases with a 24 hr periodicity. In DD, the mice exhibit activity/rest phases with periodicities imposed by their intrinsic clock: wild type 23.7 hr; Cry1–/– 22.7 hr; Cry2–/– 24.7 hr. The Cry1–/– Cry2–/– double knockout is arrhythmic because it has no functional circadian clock.
recording their daily wheel running activity for 28 days [150, 151]. For the first week the mice were kept under 12 hours of light and 12 hours of dark (LD12:12), and for the final 3 weeks they were kept in constant darkness. The results are shown in Fig. 24. Under these light:dark conditions, wild-type (WT) mice and mutant mice were active during the dark and rested during the day, as would be expected because mice are nocturnal animals. However, under conditions of continuous darkness with no external stimuli, the mice behaved differently. Whereas the WT mouse exhibited an activity-rest rhythm with a periodicity of 23.7 hours, the CRY2 mutant exhibited a longer period of 24.7 hours and the CRY1 mutant had a rhythmicity with a period of 22.7 hours. More strikingly, the CRY1–/–CRY2–/– double knockout totally lost rhythmic behavior in constant darkness. These findings established cryptochromes as core clock proteins.

While this work was going on in our lab [152–170], there was a great deal of scientific discoveries in the circadian field over the period of 1996–2000 that led to the identification of the four classes of proteins (Fig. 25) that are essential for controlling the circadian clock in humans [146–149, 171, 172]. The following model was developed for the molecular clock: CLOCK and BMAL1 activate the transcription of CRY and PER, which after a time delay, enter the nucleus and inhibit their own transcription resulting in rise and fall of CRY and PER levels with a periodicity of about 24 hours. In addition, these core clock proteins control the expression of about 30% of all genes in a given tissue to confer this cyclic expression pattern and thus a daily rhythmicity of functions (Fig. 26).

**CIRCADIAN CLOCK, CONTROL OF EXCISION REPAIR, AND CARCINOGENESIS**

Among the genes regulated by the circadian clock, we found that the excision repair protein XPA is controlled by the biological clock, and we therefore asked whether the entire nucleotide excision repair oscillates with daily periodicity. As shown in Fig. 27, XPA transcription and protein levels are at a maximum at

---

**FIGURE 25.** Mammalian clock genes and proteins. The four genes and their paralogs were cloned, and the proteins were characterized over the course of five years.
around 5 pm and at a minimum at around 5 am. Importantly, the entire excision repair activity shows the same pattern [173–175]. This led to the prediction that mice would be more sensitive to UV light when exposed at 5 am (when repair is low), compared to 5 pm (when repair is high). We proceeded to test this prediction. We irradiated two groups of mice with UV at 5 am and 5 pm, respectively, and found that the group irradiated at 5 am exhibited 4–5 fold higher incidence of invasive skin carcinoma than the group irradiated at 5 pm [176]. Currently, we are investigating whether this rhythmicity of excision repair exists in humans,

**Figure 26.** Molecular mechanism of the mammalian circadian clock. CLOCK and BMAL1 are transcriptional activators, which form a CLOCK-BMAL1 heterodimer that binds to the E-box sequence (CACGTG) in the promoters of Cry and Per genes to activate their transcription. CRY and PER are transcriptional repressors, and after an appropriate time delay following protein synthesis and nuclear entry, they inhibit their own transcription, thus causing the rise and fall of CRY and PER levels with circa 24 hour periodicity (core clock). The core clock proteins also act on other genes that have E-boxes in their regulatory regions. As a consequence, about 30% of all genes are clock-controlled genes (CCG) in a given tissue, and hence exhibit daily rhythmicity. Among these genes, the Xpa gene, which is essential for nucleotide excision repair, is also controlled by the clock.
The core circadian clock machinery controls the rhythmic expression of XPA, such that XPA RNA and protein levels are at a minimum at 5 am and at a maximum at 5 pm. The entire excision repair system therefore exhibits the same type of daily periodicity. As a consequence, when mice are irradiated with UVB at 5 am they develop invasive skin carcinoma at about 5-fold higher frequency compared to mice irradiated at 5 pm when repair is at its maximum. The mouse in the picture belongs to the 5 am group with multiple invasive skin carcinomas at the conclusion of the experiment.

and if it does, whether it can be used to make public health recommendations to prevent skin cancer [170, 177]. Equally important, excision repair is also the repair mechanism for the DNA damage caused by the anticancer drug cisplatin. We are thus also investigating whether this periodicity of excision repair can be used to improve cisplatin treatment in cancer.

SUMMARY

To conclude, Fig. 28 is the summary of my 40 years of work on photolyase, 35 years of work on nucleotide excision repair in E. coli and humans, and 20 years of work on photolyase-related cryptochrome that links these two repair pathways that I have worked with all my career.
Mechanisms of DNA Repair by Photolyase and Excision Nuclease

Figure 28. Photorepair, cryptochrome, and nucleotide excision repair. I have worked for over 40 years on photolyase, whose photocycle is shown here. I have spent 35 years on nucleotide excision repair and discovered two different dual incision patterns in *E. coli* and humans as shown. The photolyase work eventually led to the discovery of cryptochrome as a core clock protein, as exemplified by these actograms. The core clock controls nucleotide excision repair in mice which revealed that a photolyase-like protein (cryptochrome) with no repair activity nevertheless controls repair through the circadian clock.

Acknowledgements

I have had the good fortune of having worked with outstanding students and postdocs over the course of my career who have conducted most of the experiments I described here (Fig. 29). I am grateful to my internal medicine professor, Muzaffer Aksoy, who encouraged me to go to the United States to do research. My mentor, Dr. Rupert, discovered DNA repair in the modern sense, and he has been my role model throughout my scientific career. W. Dean Rupp and Paul Howard-Flanders introduced me to nucleotide excision repair and helped lay the scientific foundation for my research on DNA repair (Fig. 30).

I have been very fortunate to have had collaborators who were leaders in the field of flavin photochemistry, analytical chemistry, crystallography, ultrafast chemistry, and the mammalian circadian clock (Fig. 30). Finally, I wish to acknowledge the scientists who have made important contributions in the fields of photolyase, excision repair, and circadian clock research. In this lecture I was not able to fully reference all contributions to these fields, but have tried to place
Figure 29. Sancar lab members. I have been fortunate to have had outstanding postdoctoral fellows, students, and technicians who have carried out the experimental work described in this lecture.

Sancar Lab Members

Adar, Sheera  
Ahn, KyuJeong  
Akan, Zafar  
Annayev, Yunus  
Arujo, Francisco  
Arat, Nezahat  
Arnet, Robin  
Asimgil, Hanife  
Bereketoğlu, Sadır  
Berrocal, Gloria  
Bessho, Tadayoshi  
Bondo, Eddie  
Boyuer, James  
Branum, Mark  
Cakt, Ceylan  
Cantürk, Fazile  
Capp, Christophor  
Carton, Wendi  
Chiou, Yi-Ying  
Choi, Jun-Hyuk  
Croteau, Deborah  
Dawut, Lale  
Denaro, Tracy  
DeRocco, Vanessa  
Ersin-Simon, Ingrid  
Erkmen, Gulnai Kulakzsi  
Gaddameedi, Shobhan  
Gauger, Michele  
Han, Chih-Chiang (Eric)  
Hara, Ryujiro  
Hassan, Bachar  
Heenan, Erin  
Hsu, Shiao-Wen (David)  
Huang, Juch-Chin (JC)  
Husain, Intesar  
Hutrell, Stephanie  
Jiang, Gouchun  
Kang, Tae-Hong  
Karakan, Muhammet  
Kavakli, Ibrahim (Halli)  
Kawara, Hiroaki  
Kazantsev, Aleksey  
Kemp, Michael  
Kim, Sang-Tae  
Lee, Jin-Hyup  
Levy, Michael  
Li, Wentao  
Li, Ywan-Feng  
Lin, Jing-Jer  
Lindsey-Boltz, Laura  
Malhotra, Kshubbeer  
Matsunaga, Tsukasa  
McCullough-Buchanan, Carla  
Meganck, Rita  
Miyanoto, Yasuhide  
Mo, Jinyao  
Morrison, Lydia  
Mu, David  
Myles, Gary  
Nichols, Anne  
Öğrenc, Muge  
Orren, David  
Özer, Zahide  
Özgür, Bezgin  
Ozkan-Daghiyan, Irem  
Öztürk, Nuri  
Park, Chi-Hyun  
Partch, Carrie  
Payne, Gillian  
Payne, Nicola  
Petit, Claude  
Phillips, A. Meleah  
Rastogi, Promila  
Reardon, Joyce  
Sar, Funda  
Selby, Christopher  
Serin, Ozdemirhan  
Shields, Katie  
Sibghat-Ullah  
Smith, Frances  
Song, Sang-Hun  
Swoboda, Daniel  
Thomas, David  
Thompson, Carol  
Thresher, Randy  
Ünsal-Kaçmaz, Keziban  
Vagas, Elif  
Van Houten, Ben  
Wakasugi, Mitsu  
Worthington, Erin (Nikki)  
Yang, Yanan  
Ye, Rui  
Yılmaz, Seçil  
Zhao, Xiaodong (Jerry)  
Zhao, Shaying

Mentors

RUPERT, CLAUD S.  
Aksoy, Muzaffer  
Rupp, W. Dean  
Howard-Flanders, Paul

Collaborators

Bambara, Robert  
Linn, Stuart  
Lippard, Stephen  
Modrich, Paul  
Rajagopalan, K.V.  
Reinberg, Danny  
Sancar, Gwendolyn  
Smithies, Oliver  
Takahashi, Joseph  
Taylor, John-Stephen  
Thompson, Larry  
Van Gelder, Russel  
Wold, Marc  
Zhong, Dongping

Funding

NIH National Institute of General Medical Sciences

Figure 30. Mentors and collaborators. Professor Muzaffer Aksoy, my internal medicine professor at the Istanbul University School of Medicine, taught me the scientific method. Professor Claud S. Rupert, my doctoral advisor at the University of Texas at Dallas, has had the most impact on my development as a scientist. He is my role model. Professors W. Dean Rupp and Paul Howard-Flanders introduced me to the field of nucleotide excision repair and helped create an exciting DNA repair research environment at Yale University. My collaborators, which are listed here, have been instrumental to the accomplishment of all three major research projects in our lab. My research programs have been supported by the National Institutes of Health (NIH) for the past 33 years.
our work in historical context. Science is not done in a vacuum. We have greatly benefited from work done on these topics by our predecessors as well as our contemporaries who carried out similar work on these topics (Fig. 31). Their ideas, findings, and shared reagents have been critical to the success of my laboratory. I wish to acknowledge Laura Lindsey-Boltz, Michael Kemp, and Rita Meganck for their assistance preparing this lecture and manuscript.

REFERENCES


cell lines in complementation groups 6–10 have normal general excision repair activity. *Nucleic Acids Research* 25, 1015–1021.


