Transcription-coupled DNA repair: two decades of progress and surprises

Philip C. Hanawalt and Graciela Spivak

Abstract | Expressed genes are scanned by translocating RNA polymerases, which sensitively detect DNA damage and initiate transcription-coupled repair (TCR), a subpathway of nucleotide excision repair that removes lesions from the template DNA strands of actively transcribed genes. Human hereditary diseases that present a deficiency only in TCR are characterized by sunlight sensitivity without enhanced skin cancer. Although multiple gene products are implicated in TCR, we still lack an understanding of the precise signals that can trigger this pathway. Futile cycles of TCR at naturally occurring non-canonical DNA structures might contribute to genomic instability and genetic disease.

Translesion synthesis
The continuity of replication (or transcription) through a damaged site in the DNA template, in general with a high probability of errors that can lead to mutation.

Transcription-coupled repair
The specialized repair pathway that operates on lesions in the transcribed strands of expressed genes.

Nucleotide excision repair
The classic process by which a large range of structurally unrelated alterations in duplex DNA can be repaired. Following lesion recognition, incisions are made on both sides of the lesion and a short stretch of the DNA strand that contains the damage is excised. A repair patch is synthesized using the undamaged complementary DNA strand as a template.

Department of Biology, Stanford University, 371 Serra Mall, Stanford, California 94305-5020, USA. Correspondence to P.C.H. e-mail: hanawalt@stanford.edu
doi:10.1038/nrm2549

The maintenance of the genome and its accurate replication are prerequisites for life. Transcription of DNA-encoded instructions, which include RNA templates for protein synthesis and regulatory information to coordinate a myriad of chemical reactions, is also crucial for cell function and survival. These processes can be threatened by alterations in the DNA structure that interfere with the progression of DNA polymerases and RNA polymerases, or compromise the fidelity of replication and transcription. Damage to DNA arises from both environmental and endogenous sources, including genotoxic chemicals and radiation, reactive oxygen species and the intrinsic instability of the DNA molecule[1,2]. Naturally occurring non-canonical DNA structures can also interfere with both replication and transcription, giving rise to gross chromosomal aberrations and mutations[3–7]. Some types of obstruction are more serious than others: for example, one double-strand break or interstrand covalent crosslink in the cellular genome might be sufficient to preclude the generation of viable daughter cells. Other types of damage are equally disastrous if they pose insurmountable hurdles to genomic duplication. A compounded threat might arise when an advancing replication fork encounters an arrested transcription complex at a damaged site in the template DNA[8]. Thus, over the course of the past 3 billion years, a highly effective set of cellular responses has evolved to deal with life-threatening disruptions to replication and transcription.

The cellular 'toolbox' of responses to genomic stress includes checkpoints that result in arrested cell cycle progression, stabilization of DNA at blocked replication forks and recruitment of DNA repair enzymes, to facilitate progression of DNA and RNA synthesis[9–11]. One of the most surprising advances in the past decade has been the discovery of nearly a dozen novel mammalian DNA polymerases, of which many can perform translesion synthesis through various types of DNA lesions[12,13]. However, there has been no corresponding discovery of specialized RNA polymerases (RNAPs) that successfully overcome damage that arrests transcription. An outline of pathways for the processing of damaged or otherwise altered DNA structures is given in BOX 1.

Here, we focus on the mechanisms and function of transcription-coupled repair (TCR), a subpathway of nucleotide excision repair (NER) that targets DNA alterations that interfere with the translocation of RNAP through expressed genes. Phylogenomic evidence exists for the operation of TCR in most bacterial species and many eukaryotes, although it has not been established for archaea[14]. We consider the interface of TCR with the subpathway of global genomic repair (GGR), which deals with damage in both expressed and silent genomic domains. Although the mechanism of TCR is well understood in bacteria, research with eukaryotic cell-free systems has failed to fully validate *in vivo* TCR observations, and this has hindered the detailed biochemical analysis of TCR in higher organisms. Nevertheless, most of the implicated proteins have been identified and plausible models for TCR can be evaluated. We present the clinical features and genetic deficiencies in several human hereditary diseases that express TCR defects. We also approach the question of whether TCR might sometimes be deleterious if it were triggered by naturally occurring non-canonical DNA structures to engage in futile cycles of repair. This possibility is timely to consider as we learn that hot spots for mutations, translocations and chromosomal aberrations, localized to non-B form DNA structures, correlate with genetic disease and cancer[4,15].

Box 1 | **General categories of DNA repair and damage processing**

Direct reversal
Some DNA lesions can be reversed without incising the phosphodiester DNA backbone. The classic example of this is the splitting of ultraviolet (UV)-induced cyclobutane pyrimidine dimers *in situ* by light-activated photolyases. Another example is an alkyl transferase that removes a methyl group from the O-6 position in guanine. Single-strand breaks in which a 3′ hydroxyl abuts a 5′ phosphate with no missing nucleotide can be directly joined by ligase.

Excision repair
Most damage is nullified through excision repair, an error-correcting process in which the faulty segment of one DNA strand is excised and then replaced by repair replication, using the undamaged complementary strand as a template. The redundancy of genetic information inherent in duplex DNA is essential for excision repair as well as for semi-conservative genomic replication. The ubiquitous pathway of nucleotide excision repair (NER) deals with a large range of structurally unrelated and helix-distorting DNA lesions. NER was discovered 45 years ago in studies on the processing of UV-damaged DNA in bacteria, and the modes of base excision repair (BER) and mismatch repair (MMR) were revealed shortly thereafter. BER is initiated by the removal of an altered or inappropriate base (for example, uracil) from DNA, leaving an abasic site that can be excised and repaired by single-nucleotide insertion or short-patch excision repair. Nucleotide incision repair is a backup system for BER for certain types of oxidative base damage. MMR corrects mistakes in nucleotide incorporation during DNA replication and also removes small, unpaired single-strand loops.

Recombinational pathways
The resolution of double-strand breaks, interstrand crosslinks and covalently bound proteins might require several different repair pathways that act in concert: these include homologous recombination, single-strand annealing and non-homologous end-joining, in addition to NER. Recombination is similarly essential to restore genomic continuity from so-called collapsed replication forks that are generated when replication encounters a single- or double-strand break in the parental DNA template[2].

Translesion synthesis
The bypass of lesions by specialized DNA polymerases is a tolerance mode that postpones the requirement for NER, but which might also operate in conjunction with recombination and/or NER to deal with certain types of lesions, particularly those that affect both DNA strands at closely opposing sites.

A historical perspective of TCR

The first human NER-defective genetic disorder to be characterized was xeroderma pigmentosum (XP)[16]. Genomic heterogeneity in mammalian NER efficiency was documented by the finding that chemical damage in the highly repetitive α-DNA sequences in African green monkey kidney cells is resistant to repair, compared with that of bulk genomic DNA[17]. Intragenomic NER heterogeneity in human cells was established by the discovery that ultraviolet (UV)-induced cyclobutane pyrimidine dimers (CPDs) in some genomic domains were repaired in XPC mutants, a complementation group of XP that is defective in GGR[18]. Analysis of nucleic acid synthesis in cells from patients with the rare photosensitivity disease Cockayne syndrome (CS) revealed an unusual delay in recovery of RNA synthesis after UV irradiation, leading to the suggestion that CS cells might be deficient in repair of expressed genes[19]. It eventually became clear that TCR was responsible for the repaired domains in XPC cells[20,21], and that CS cells and those from the corresponding Chinese hamster ovary (CHO) cell mutant were defective in TCR[22–25].

It was fortunate that the first experiments documenting TCR were carried out with CHO cells, as rodent cells are generally deficient in GGR of CPDs (for a review, see REF. 26). Repair of CPDs in CHO cells was shown to be more efficient in the expressed dihydrofolate reductase (*Dhfr*) gene than in a transcriptionally silent downstream sequence[27]. Using strand-specific RNA probes to highlight DNA fragments in Southern blots[28], subsequent studies revealed that the proficient repair of CPDs in the *Dhfr* gene in CHO or human cells was primarily due to the preferential repair of the transcribed DNA strand[29]; this result became accepted as the operational definition of TCR. However, detection of TCR can be obscured for certain lesions, such as the 6–4 photoproduct, which is efficiently repaired by GGR. Therefore, the elimination of GGR in an XPC mutant might reveal TCR[30,31]. The most important function of TCR is probably to remove obstructions to RNAP translocation rather than simply to repair expressed genes more rapidly. Although the operation of TCR can result in a strand bias in mutagenesis[32], the absence of this subpathway seems to cause mild effects in humans (see below).

Studies in UV-irradiated *Escherichia coli* revealed that CPDs in the transcribed strand of the expressed *lacZ* gene are more rapidly repaired than those in the non-transcribed strand[33]. These studies rendered unlikely the hypothesis that TCR in mammalian cells is the consequence of an open chromosomal structure in actively transcribed domains. TCR was also documented in the budding yeast *Saccharomyces cerevisiae*, in which repair in an expressed gene on a centromeric plasmid was similar to that of the same gene expressed in the chromosome[34,35]. TCR in yeast was abolished at the non-permissive temperature in an RNAPII temperature-sensitive mutant[35] and by treatment with α-amanitin (an inhibitor of RNAPII translocation) in CHO cells, implicating RNAPII transcription elongation in the mechanism[36]. Increased interest in the connections between DNA repair and transcription arose from discovery of the dual role of the basal transcription factor TFIIH in these DNA transactions, and the reports of human diseases in which TCR was deficient[8,37,38] (TABLE 1).

Molecular mechanism of TCR in bacteria

The basic features of TCR in bacteria are well characterized[39] (FIG. 1). The arrested RNAP obscures immediate recognition of a CPD by the NER factors UvrA and UvrB. However, the RNAP is then bound by a transcription-repair coupling factor, Mfd, which releases RNAP and recruits UvrA. From then on, the pathway is identical to GGR. In addition to its RNAP-interaction domain, Mfd has a UvrB-homology module that enables its interaction with UvrA. Mfd also features tandem RecA-like domains that contain helicase motifs and are needed for DNA-dependent ATPase and DNA-translocation roles, although it has no strand-opening activity. This is a typical feature of the SNF2 superfamily of proteins, which are implicated in chromatin remodelling and the removal of other proteins from DNA[40].

The ATPase activity energizes Mfd translocation along the DNA, enabling it to push the RNAP forward. If the RNAP has backtracked, the nascent end of the transcript is extruded, such that when Mfd pushes it forward the 3′ end of the transcript becomes repositioned in the active site and transcription can resume[41]. However, this

Global genomic repair
DNA repair throughout the genome, without distinguishing between sequence, activity or chromatin structure.

Non-B form DNA structure
A DNA secondary structure that differs from the canonical Watson–Crick right-handed double helix of ten bases per turn of the helix. Examples include left-handed Z-DNA, triple-stranded H-DNA and four-stranded quadruplex DNA.

α-DNA
A highly repetitive DNA sequence that is found near centromeres in chromosomes from primates.

Table 1 | **Human hereditary diseases exhibiting TCR, GGR or TCR and GGR deficiency**

Disease	Implicated gene	HUGO nomenclature	Overlap with other diseases*	TCR	GGR	Cancer prone
XP	*XPA*	*XPA*	DSC	–	–	+
	XPB	*ERCC3*	CS and TTD	–	–	+
	XPC	*XPC*	Unknown	+	–	+
	XPD	*ERCC2*	CS, TTD and COFS	–	–	+
	XPE	*DDB2*	Unknown	+	–	+
	XPF	*ERCC4*	Unknown	–	–	+
	XPG	*ERCC5*	CS and COFS	–	–	+
CS	*CSA*	*ERCC8*	Unknown	–	+	–
	CSB	*ERCC6*	$UV^{S}S$, COFS and DSC	–	+	–
	XPB	*ERCC3*	XP and TTD	–	–	+
	XPD	*ERCC2*	XP and TTD	–	–	+
	XPG	*ERCC5*	XP	–	–	+
	ERCC1	*ERCC1*	COFS	–	–	Unknown
$UV^{S}S$	*CSB*	*ERCC6*	CS, COFS and DSC	–	+	–
	Unknown	Unknown	Unknown	–	+	–
TTD	*XPB*	*ERCC3*	XP and CS	–	–	–
	XPD	*ERCC2*	XP and CS	–	–	–
	TTDA	*GTF2H5*	Unknown	–	–	–
	TTDN1	*C7orf11*	Unknown	+	+	Unknown
	Others	Unknown	Unknown	+	+	Unknown

*Mutations in the gene implicated in the primary disease, listed in the first column, can also cause other diseases with defective DNA repair. There is no agreement as to whether COFS and DSC (and even XP–CS) should be classified as separate diseases or as the most severe manifestations of CS and/or XP. Similarly, $UV^{S}S$ might be considered as the mildest presentation of CS. *C7orf11*, chromosome 7 open reading frame-11; COFS, cerebro-oculo-facio-skeletal syndrome; CS, Cockayne syndrome; *DDB2*, DNA damage-binding-2; DSC, De Sanctis–Cacchione syndrome; *ERCC*, excision repair cross-complementing; GGR, global genomic repair; *GTF2H5*, general transcription factor-IIH polypeptide-5; HUGO, Human Genome Organisation; TCR, transcription-coupled repair; TTD, trichothiodystrophy; $UV^{S}S$, ultraviolet-sensitive syndrome; XP, xeroderma pigmentosum.

Cyclobutane pyrimidine dimer
The principal photoproduct that is formed in DNA following absorption of short-wavelength ultraviolet light. Adjacent pyrimidines are covalently linked through their C-5 and C-6 carbon atoms.

6–4 photoproduct
An ultraviolet-induced DNA lesion formed between the C-4 position of a 5′ pyrimidine and the C-6 position of an adjacent pyrimidine. This occurs approximately one-third as frequently as the cyclobutane pyrimidine dimer, but causes greater structural distortion.

SOS genomic stress response
A bacterial response to replication arrest, regulated by RecA and LexA proteins, to upregulate over 40 genes that control nucleotide excision repair, translesion DNA synthesis, recombination and other functions, such as cell division.

Mfd 'pressure' might apply a twisting force on the DNA if the RNAP is blocked at a lesion, thereby eliminating the RNA–DNA hybrid in the transcription bubble as DNA strands re-anneal, and also forcing release of the RNAP from the DNA[42]. The interaction of RNAP with Mfd depends on three amino acids in the β-subunit of the RNAP[43]. However, another amino acid that is not required for this interaction also results in loss of TCR when substituted, suggesting that the Mfd–RNAP interaction is necessary, but not sufficient, to facilitate TCR[44].

The 'decision' of whether repair should be initiated is evidently made by UvrA and UvrB after the RNAP has been released. Although the offending lesion is typically in the transcribed DNA strand, this information might be lost when the strands re-anneal, so the selection of the damaged strand would need to be re-determined by the UvrB helicase. Of course, the selective repair of the transcribed strand is already ensured by the fact that the RNAP-arresting lesion is in that strand. If a lesion in the non-transcribed strand were to block transcription, then that lesion would probably be preferentially repaired, although this prediction has not been amenable to testing. The truncated RNA product is presumably destroyed after the ternary complex (RNAP–RNA–DNA) is disrupted at the site of RNAP arrest, so transcription must be reinitiated from the promoter after repair.

The entire transcriptome should be subject to TCR, as all transcription is carried out by one type of RNAP in *E. coli*. The fact that *Mfd*-deficient bacteria grow normally and are only mildly sensitive to UV indicates that although TCR contributes to survival, it is not essential for viability in undamaged cells. It should be re-emphasized that the identification of damage in the case of bacterial TCR is essentially the same as that for GGR: the more rapid repair is simply due to the targeted recruitment of lesion-recognition factors to the site of an arrested RNAP. Although a CPD can thus be sensitively detected for TCR, it is only weakly recognized by GGR owing to the low cellular concentrations of UvrA and UvrB. The activation of the SOS genomic stress response by DNA replication arrest in UV-irradiated *E. coli* results in upregulation of *uvrA* and *uvrB* gene expression, leading to increased concentrations of UvrA and UvrB and proficient GGR of CPDs[45,46].

In a simplistic model of TCR, one could suppose that the recruitment of UvrA by Mfd leads to an increase in concentration of UvrA in the region of transcription arrest. The demonstration of NER-based incisions in undamaged DNA by the reconstituted UvrA–UvrB–UvrC complex[47] provides an evolutionary rationale for the low constitutive concentrations of UvrA and UvrB, and it prompts the suggestion that the enrichment of these factors at an arrested RNAP might result in

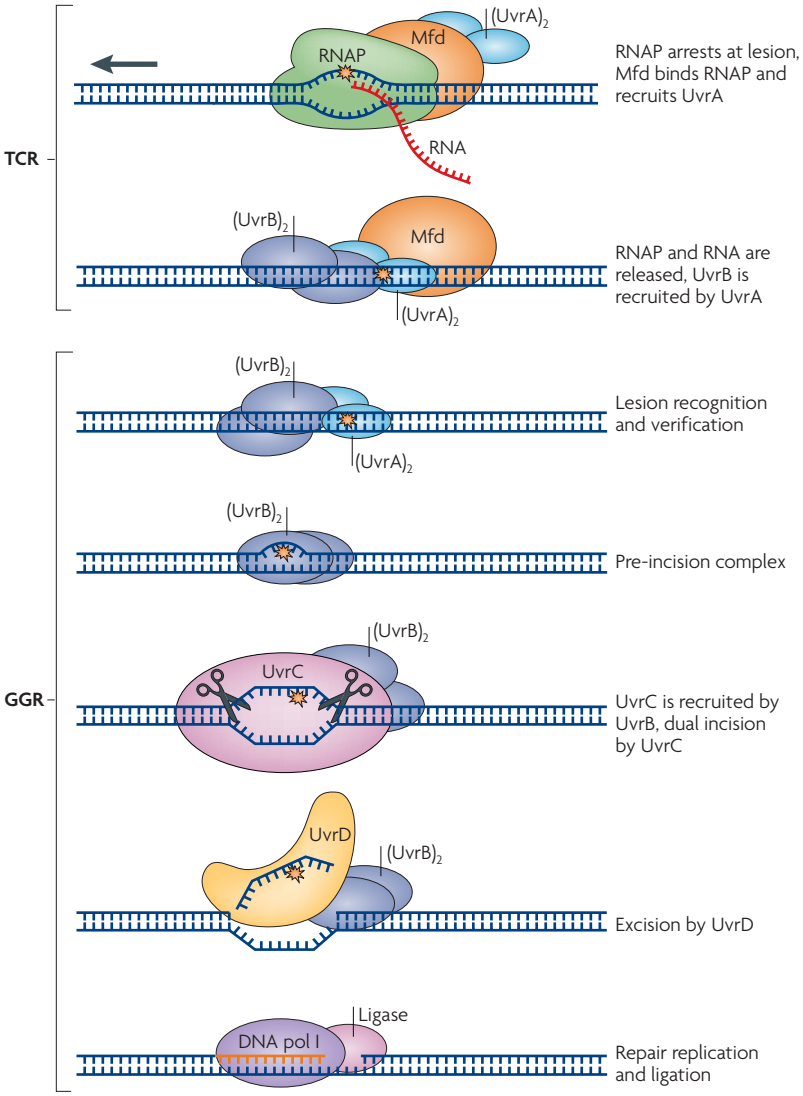

Figure 1 | **The stepwise pathway for bacterial TCR and GGR.** The arrest of *Escherichia coli* RNA polymerase (RNAP) provides the initial step to highlight a potential lesion: the arrested polymerase recruits Mfd, which then recruits UvrA. Once the UvrA homodimer is bound, it recruits the UvrB helicase, which also operates as a homodimer[146], to verify the presence of a lesion and to determine which strand is damaged. Mfd is released, and the transcription-coupled repair (TCR) reaction proceeds as for global genomic repair (GGR), in which UvrA is the primary element for lesion recognition. The cutting enzyme UvrC is recruited and activated by UvrB to make incisions on each side of the lesion ~12 nucleotides apart. The damaged oligonucleotide is then removed by the UvrD helicase. DNA polymerase-I (DNA pol I) performs repair replication and ligase joins the completed repair patch to the contiguous DNA strand. The entire GGR pathway requires only six proteins and has been reconstituted *in vitro*[147].

gratuitous repair events at naturally occurring non-B form DNA structures, if they arrested transcription[48].

Models for TCR in mammalian cells

TCR in eukaryotes is seemingly more complex than in bacteria, even though it also feeds into a common pathway that is shared with GGR (FIG. 2). A major impediment for GGR in eukaryotes is the packing of DNA in chromatin, although this is less of an impediment for TCR as the nucleosome structure is transiently opened during RNAP translocation. We focus on TCR in mammalian cells, although much of our basic understanding has also derived from studies in yeast (for a review, see REF. 49).

In addition to the involvement of different RNAPs, the bacterial and mammalian systems differ in that the initial recognition of an obstruction by RNAPII removes the need for GGR lesion-recognition elements in mammals. A transcription-coupling factor is required and would be expected to have similar characteristics to those of Mfd. The obvious candidate, CSB (also known as ERCC6), is also a member of the SNF2 family of DNA-dependent ATPases[40]; it interacts loosely with the elongating RNAPII and stimulates transcription[50] but becomes more tightly bound following transcription arrest[51]. CSB might also push the RNAP forward, such that an additional nucleotide is incorporated opposite a CPD, but dissociation of RNAP from the DNA was not shown *in vitro*[52]. Because XPC generates a small bubble following the recognition of a lesion[53], it could be reasoned that XPC is not necessary for TCR because the translocating RNAP has already displaced the non-transcribed strand to form the transcription bubble. DNA damage-binding-2 (DDB2), which is encoded by *XPE* and is implicated in lesion recognition for GGR and recruitment of XPC[54,55], is also dispensable for TCR. DDB2 in association with DDB1 accomplishes for CPD recognition in GGR what the RNAP does in TCR; it serves as an antenna to sense an otherwise poorly recognized lesion, whereas the more structure-distorting 6–4 photoproduct can be directly recognized by XPC[56]. Once the gap-opening helicase factor TFIIH and other factors have been recruited, the TCR pathway is identical to that of GGR (FIG. 2) and the corresponding repair patch sizes are similar[57].

Possible scenarios following RNAP arrest. A number of outcomes are possible after RNAP arrest and tighter binding by CSB (FIG. 3). The half-life of an arrested RNAPII at a CPD *in vitro* is ~20 hours[52], and its footprint extends 10 nucleotides ahead of the CPD and 25 nucleotides behind[58]. This creates an impasse: transcription cannot continue and the blocking lesion cannot be accessed by repair factors. Prolonged transcription arrest, as in CS cells, leads to a strong signal for cellular apoptosis[11,59,60]. For some types of lesions, the RNAPII might be able to carry out translesion transcription but with the likelihood of transcriptional mutagenesis, a process that is also documented in bacteria[61].

Alternatively, the RNAPII might be displaced to reveal the blocking lesion, either through reverse translocation or by release from the DNA. Regression of RNAPII from arrest at a CPD was documented *in vitro*[62]. In a proof-of-principle experiment, it was shown that backtracking could proceed far enough that a small DNA-repair enzyme, photolyase, could bind and directly reverse the CPD[63]. If the transcription bubble collapses as the polymerase is pushed forward by CSB (as in *E. coli* by Mfd), or is eliminated as the RNAP regresses, the binding of the single-stranded DNA (ssDNA)-binding complex replication protein A (RPA) to the lesion site might prevent complete renaturation

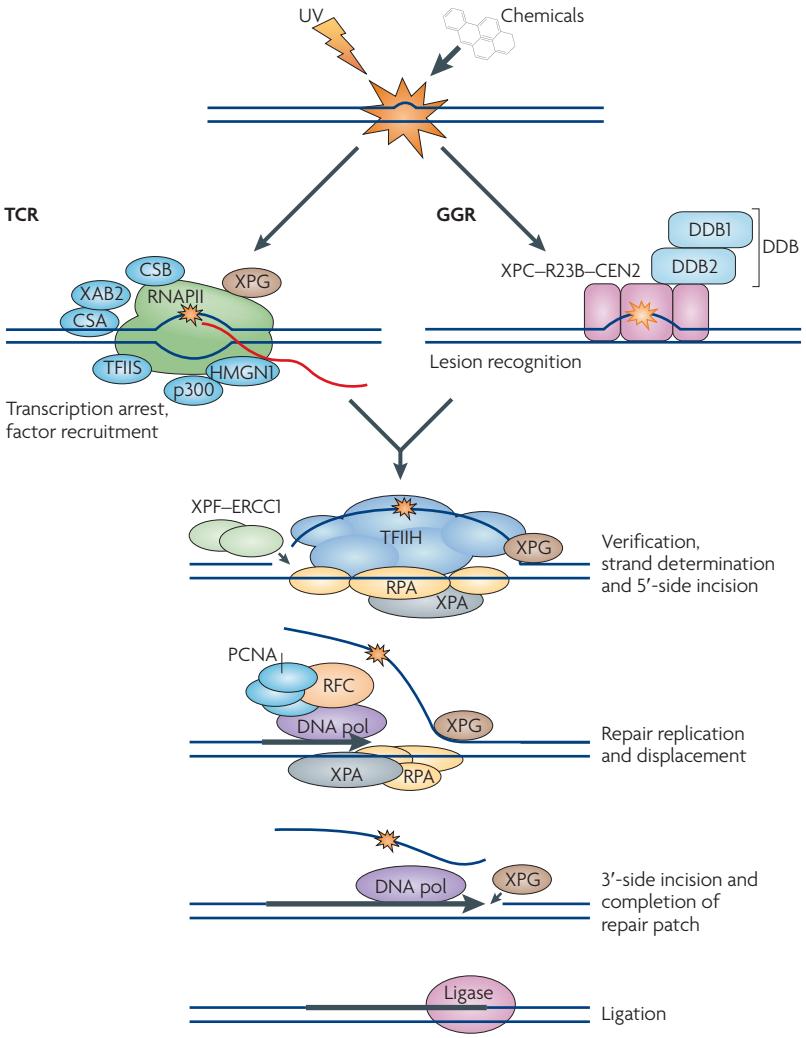

Figure 2 | **Probable sequence of events in mammalian TCR and GGR.** Lesions are initially recognized, either by a translocating RNA polymerase (RNAP) (for transcription-coupled repair (TCR); left) or through the binding of the lesion sensor DNA damage-binding-2 (DDB2; the product of xeroderma pigmentosum complementation group E (*XPE*)) — which forms a heterodimer with DDB1 to constitute the DDB complex — and/or XPC in complex with RAD23B and centrin-2 (CEN2) (for global genomic repair (GGR); right). The subpathways converge to the following steps for nucleotide excision repair (NER): transcription factor TFIIH is recruited (along with XPG, which stabilizes TFIIH); and the helicase and ATPase activities of its subunits XPD and XPB, respectively, are stimulated for further opening of the damaged DNA (not shown). TTDA, another subunit of TFIIH, is required for NER, but its role has not been clarified. Replication protein A (RPA) and XPA might be present before and/or after the appearance of TFIIH, as they have lesion-verification roles and protect the single-stranded DNA in the denatured bubble and stabilize the pre-incision complex. The XPF–ERCC1 (excision repair cross complementing-1) endonuclease complex is recruited and incises the damaged DNA strand at the 5′ side of the bubble, whereas XPG incises on the 3′ side. Notably, recent *in vivo* studies have provided strong evidence that once lesion verification has occurred, the 5′-side incision can be made, and then repair replication begins before the 3′-side incision is produced by XPG, thus minimizing the amount of exposed single-stranded DNA during repair[148]. Replication factor C (RFC) loads the processivity factor proliferating cell nuclear antigen (PCNA) to accommodate DNA polymerases (DNA pol) δ, ε and/or κ[149] that have been implicated in repair replication. Unlike the situation in *Escherichia coli*, in which UvrD 'peels off' the damaged oligonucleotide, this is probably accomplished in mammals by the translocating DNA pol. The final ligation step can be carried out by ligase-I and flap endonuclease-1 or by the ligase-III–XRCC1 complex[2,150,151]. CS, Cockayne syndrome; HMGN1, high-mobility group nucleosome-binding domain-containing protein-1; UV; ultraviolet; XAB2, XPA-binding protein-2.

and mark the site for access by other NER factors, such as XPA. If there were no lesion or an unusual secondary DNA structure, the strands should be able to perfectly re-anneal and RPA would be displaced. Thus, RPA might be a crucial factor for testing whether arrest is due to a lesion. Another possibility is that TFIIH might bind to the transcription bubble when the RNAP is displaced, thus leading directly to the subsequent steps in NER.

Although there is no functional difference between the options of RNAPII regression and RNAPII removal (followed by NEDD4-dependent ubiquitylation[64], which leads to proteosomal degradation of RPB1, the largest subunit of RNAPII[65]), it is intuitive that the regression mode might be preferable for mammalian genes, which are generally much larger than those of prokaryotes. It would be surprising if the partially completed RNA product were aborted each time RNAPII encounters damage in the largest human gene, dystrophin, which is 2.8 Mb long and requires about 16 hours to transcribe[66]. The RNAPII regression model is similar, in principle, to that proposed for NER of a CPD at an arrested replication fork in *E. coli*[67]. Regression of polymerases from an obstacle might be a conserved process from the early stages of evolution, with respect to both replication and transcription.

In principle, TCR might be initiated by remodelling the RNAPII without removal from the arrest site, as elaborated in a recent review[68]. Experiments *in vitro* have suggested that ATP-dependent remodelling of the arrested RNAPII can occur in association with XPG and TFIIH, such that the DNA template strand for transcription in the arrested RNAPII becomes accessible to nicking by XPG[69]. Crystallographic analyses[70–72] have localized a CPD or a cisplatin intrastrand crosslink within or in front of the active site of RNAPII, respectively, and have suggested that no structural change occurs in the polymerase.

In the *in vitro* studies mentioned above, there was no requirement for CSB or lesion-recognition factors other than the RNAP itself, and the nicking reaction was shown to occur at a stalled RNAPII owing to nucleotide depletion, which should not be a substrate for TCR *in vivo*[69]. Although it is possible that the pre-incision complex and the NER endonucleases could assemble on the DNA without displacing RNAPII, there are a number of conceptual difficulties with the idea that repair could be completed inside such a structure; these difficulties include the obstruction that is caused by the transcribed DNA strand–RNA hybrid, the necessary re-annealing of DNA strands (which occupy separate regions in the RNAPII) to allow the non-transcribed strand to serve as template for repair synthesis, the problem of removing the damaged oligonucleotide, and steric hindrance of the associated replication factors and ligase needed to complete repair. However, the documentation of the presence of XPG along with CSB and their interaction[73] indicates a primary role for XPG in the early stages of TCR.

Factors and proteins implicated in TCR. Chromatin immunoprecipitation analyses have revealed proteins, including CSA, CSB, XPA-binding protein-2 (XAB2), TFIIS, high-mobility group nucleosome-binding

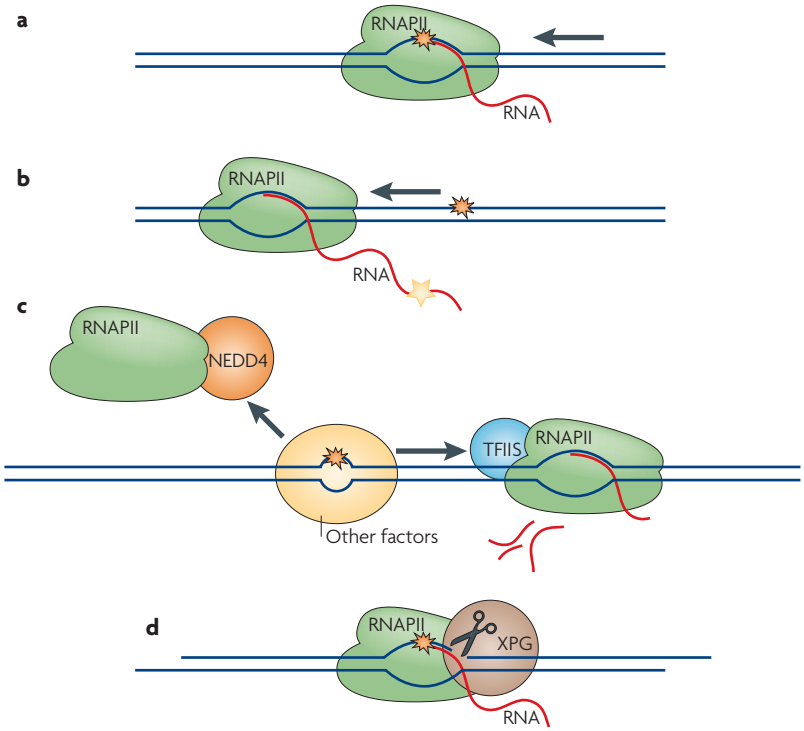

Figure 3 | **Potential outcomes following RNAPII arrest. a** | Prolonged arrest of RNA polymerase-II (RNAPII) at a lesion activates cell cycle checkpoints and might also lead to apoptosis. **b** | For some lesions, translesion transcription is possible but might result in transcriptional mutagenesis. **c** | There are two options for displacing the RNAPII to access the lesion for repair: first, reverse translocation with transcription factor TFIIS-stimulated partial degradation of the nascent transcript; and second, removal of the RNAPII from the DNA (through NEDD4-dependent ubiquitylation and proteosomal degradation, or recycling for re-initiation of transcription) and abortion of the incomplete transcript. **d** | Initiation of the repair process by XPG incision without displacing the RNAPII.

domain-containing protein-1 (HMGN1) and p300, that are associated with an arrested RNAPII[74] (FIG. 4; TABLE 2). CSB recruits CSA to the nuclear matrix where both transcription and repair take place[75,76]. CSA is part of a multiprotein complex with E3 ubiquitin-ligase activity. XAB2 seems to have a pivotal role in TCR as well as in transcription, and it is also important for mRNA splicing. Its specific interaction with XPA, a protein that is essential for both GGR and TCR, suggests that it might be the key protein at the juncture of these two subpathways, for which all subsequent steps are shared[77] (FIG. 2).

The transcription elongation factor TFIIS might be an important player in TCR[8,25], as it activates a cryptic nuclease activity of RNAPII that, following backtracking, cleaves the nascent transcript and allows its 3′ end to become re-established in the active site, thereby allowing resumed transcription. Although yeast deficient in TFIIS are not UV sensitive, those with an additional deficiency in GGR are very sensitive[78]. HMGN1 is a nucleosome-binding protein that increases acetylation of residue Lys14 in histone H3 and reduces compaction of chromatin[79]. *Hmgn1*-knockout mice exhibit reduced repair of CPDs in chromatin, implying a role in TCR[80]. HMGN1 could conceivably function to help displace nucleosomes that have become re-established behind the translocating transcription complex, so that the RNAPII can regress from the lesion (FIG. 4). p300 is implicated in chromatin remodelling during transcription and repair[81,82]. Another protein, chromatin assembly factor-1 (CAF1), is eventually required for re-establishing nucleosome structure after completion of repair[83].

Can TCR be initiated by other eukaryotic RNAPs? Whereas RNAPII transcribes the genes for mRNAs and microRNAs, RNAPI transcribes rDNAs — genes that encode ribosomal RNAs — and RNAPIII transcribes genes for transfer RNAs (tRNAs). There is no evidence for TCR of the small RNAPIII-transcribed genes[84]. The multiple and tandem rDNA genes are also small, so it is possible that RNAPI removal and degradation is not as much of a problem as it would be for large transcripts, although we do not know the half-life of RNAPI at a lesion. Early studies suggested that UV-induced lesions resulted in premature termination of transcription by mouse RNAPI[85]. Further research with CHO cells implicated both CSB and XPC in repair of rDNA[86]. TCR has been documented for RNAPI-transcribed genes in yeast, and CSB has been found in mammalian nucleoli as a component of RNAPI transcription[87–89]. As TCR in yeast requires the *RAD4* gene, which is needed for lesion recognition in GGR, the model for TCR of RNAPI- and RNAPII-transcribed DNA in yeast might be similar to that in *E. coli*; after targeting to an arrested RNAP, the same recognition factors that are needed for GGR might be recruited to complete NER. In fact, the same scenario might operate for transcribed rDNA in mammals, for which XPC might be recruited to the arrested RNAPI to initiate repair.

The initiating signal for TCR

Although RNAP is an exquisitely sensitive detector of DNA lesions[90], we still do not know the precise nature of signals that can elicit TCR. In addition to DNA lesions, nucleotide precursor depletion and intrinsic arrest sites (for example, the histone H3.3 pause site[91]) can stall or arrest the translocating RNAPII. What is the difference, with respect to TCR, between a lesion and a pause site or a naturally occurring non-canonical DNA structure?

Transcription arrest *in vitro* correlates with TCR *in vivo* at many different lesions[58,92]. In each case, the responsible lesion is on the transcribed strand. However, in the case of G4 quadruplex DNA, the obstruction to transcription can be found in the guanine-rich non-transcribed strand[6]. Prokaryotic RNAPs can bypass strand breaks and gaps on the transcribed strand, whereas those lesions in the non-transcribed strand pose a block, thereby implicating the non-transcribed strand in RNAP tracking along the DNA[93,94].

Whereas some lesions, such as CPDs, arrest transcription as they occupy the active site in the RNAP[71], others, such as bound proteins or interstrand crosslinks[95] that prevent strand separation, are likely to block transcription by steric hindrance before the lesion reaches the active site. Therefore, one might wonder at what point the translocating RNAPII senses lesions, such as strand breaks that relax superhelix tension in the template DNA and might provide early warning that danger lies ahead?

G4 quadruplex DNA
Planar guanine-quartet structures that are formed by DNA strands with repeats of three or more consecutive guanines. In the mammalian genome, these repeats are found in telomeres, ribosomal DNA and in genes for the immunoglobulin heavy chain segments.

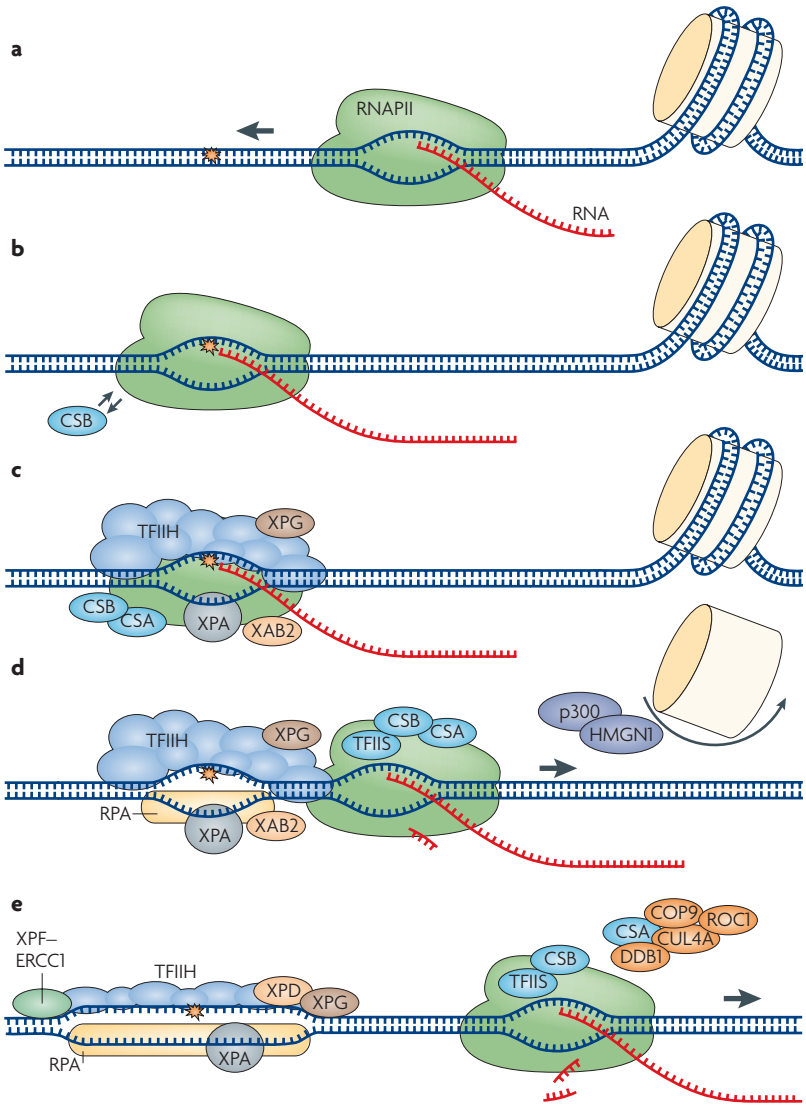

Figure 4 | **A suggested scenario for initiation of repair through TCR. a** | As RNA polymerase-II (RNAPII) translocates along the DNA, nucleosomes are dislodged in front of the polymerase and reassembled behind it. **b** | Transcription is arrested when RNAPII encounters an obstacle. **c** | Cockayne syndrome type B protein (CSB) becomes tightly bound to the arrested RNAPII and recruits factors that are needed to accomplish transcription-coupled repair (see the main text). TFIIH localizes to the arrested elongation complex with xeroderma pigmentosum complementation group G (XPG) and XPA (which is possibly brought to the scene by XPA-binding protein-2 (XAB2)); replication protein A (RPA) arrives simultaneously or shortly thereafter. **d** | The chromatin-remodelling factors high-mobility group nucleosome-binding domain-containing protein-1 (HMGN1) and p300 loosen the nucleosome structure behind the polymerase; RNAPII reverses direction, backtracking from the obstacle and degrading the nascent RNA product through its cryptic 3′–5′ exonuclease activity, which is activated by TFIIS. TFIIH with associated XPG, XPA and RPA remain at the site of the obstacle, possibly maintaining the bubble of denatured DNA, but without the RNA–DNA hybrid. XPA and RPA bind the single-stranded DNA in the vicinity of the obstruction, providing lesion verification and strand specificity before the next steps. **e** | Once RNAPII has backtracked, TFIIH extends the denatured region around the lesion to ~30 nucleotides, thus setting up the substrate for the subsequent DNA nicking by the structure-specific endonucleases XPG and the XPF–ERRC1 complex. CSA, as a component of a cullin-containing ubiquitylation E3 ligase complex, might facilitate resumption of transcription (once the repair process has been completed) by removing or deactivating factors, including CSB. The sizes of the respective proteins and complexes, and the expected distance of RNAPII backtracking are not drawn to scale, nor do they indicate their respective footprints on the DNA substrate. CUL4A, cullin-4A; DDB1, DNA damage-binding-1.

Oxidative lesions. There is current research on the role of endogenous DNA damage in genomic instability and ageing, with particular focus on the question of whether oxidized bases are substrates for TCR. Oxidative base damage is recognized by glycosylases to initiate base excision repair (BER). Early studies showed the enhanced sensitivity of CS cells to agents that induce oxidative DNA damage, such as γ-rays[96], hydrogen peroxide[97] and potassium bromate[98]. These studies suggested the possibility of the coupling of BER to transcription, a mechanism that requires that transcription be arrested by oxidized bases and requires a link between the BER and NER pathways. To date, there is no direct evidence in support of this hypothesis, and a number of papers containing data from the laboratory of A.S. Leadon have had to be retracted. Various laboratories, using different transcription systems and different nucleotide sequence contexts, have reported that oxidized bases might, or might not, be an obstruction for RNAPII[99]. We found that thymine glycol[63] or 8-oxo-guanine (8-oxoG)[100] did not arrest RNAPII. Both thymine glycol and 8-oxoG are rapidly removed through BER, and there is no compelling biochemical evidence for coupling their repair to transcription. No strand specificity for repair was observed for photosensitizer Ro19-8022-induced 8-oxoG, or for dimethyl sulphate-induced *N*-methylpurines in the hypoxanthine–guanine phosphoribosyltransferase (*Hprt*) gene of CHO cells[101,102].

However, host-cell reactivation studies with plasmids that contain these oxidative base lesions have indirectly implicated TCR, or the possible CSA- and CSB-dependent transcriptional bypass of oxidized bases[97]. Moreover, the enhanced transcriptional mutagenesis observed in *mfd*-deficient *E. coli* also provides indirect evidence for TCR of oxidative lesions[103]. It has been suggested that a decreased elongation rate or brief pausing of RNAPII at one of these lesions might be sufficient to trigger TCR[61]. It is also possible that there are particular nucleotide sequence contexts within which an oxidized base poses a block to transcription[99,104].

NER-based TCR might also occur in oxidatively damaged DNA for some lesions, such as malondialdehyde adducts[105] and 8,5′-cyclopurine-2′-deoxynucleotides[106]. Malondialdehyde is a by-product of lipid peroxidation that forms an adduct to guanine, and is the first example of an endogenous oxidative lesion that is an NER substrate. Malondialdehyde-guanine and its stable, exocyclic analogue 1,N^2-propanodeoxyguanine in the transcribed DNA strand are potent blocks to transcription by T7 RNAP and mammalian RNAPII, and are therefore likely targets for TCR[107].

Threats to transcription can also be generated from intermediates in the BER pathway. Thus, an abasic site in the transcribed DNA strand is a strong block to RNAPII transcription[108]. Complete blockage of RNAPII transcription by 2′-deoxyribonolactone, an oxidized abasic residue produced by free radical attack on the nucleotidyl C-1′ carbon following base removal, has also been documented[109]. Although this lesion is efficiently incised by the human abasic endonuclease APE1, which implicates the BER pathway, subsequent

Table 2 | **Genes involved in NER**

Human gene*	Role	Comments	Rodent‡	*Saccharomyces cerevisiae*‡	*Escherichia coli*‡
GGR genes					
XPE (also known as *DDB2*)	Lesion recognition	Recruits XPC and is p53 inducible	*Ddb2* (also known as *Xpe*)	Unknown	Unknown
DDB1	Lesion recognition	Forms a complex with DDB2	*Ddb1*	Unknown	Unknown
XPC	Lesion recognition	Opens DNA and is p53 inducible	*Xpc*	*RAD4*	*UvrA*
RAD23B	Lesion recognition	Forms a complex with XPC	*Rad23b*	*RAD23*	Unknown
Centrin-2	Lesion recognition	Forms a complex with XPC	Unknown	Unknown	Unknown
GGR and TCR genes					
XPB	Helicase and ATPase	TFIIH subunit§	*Ercc3* (also known as *Xpb*)	*RAD25*	*UvrB*
XPD	Helicase and ATPase	TFIIH subunit§	*Ercc2* (also known as *Xpd*)	*RAD3*	*UvrB*
XPA	Lesion verification	Stabilizes pre-incision complex for GGR and TCR	*Xpa*	*RAD14*	Unknown
RPA p70, *p32* and *p14*	ssDNA binding	Binds to XPA	*Rpa*	*RFA1*, *RFA2* and *RFA3*	*Ssb*?
XPF	Structure-specific endonuclease	3′ incision	*Ercc4* (also known as *Xpf*)	*RAD1*	*UvrC*
ERCC1	Forms a complex with XPF	3′ incision	*Ercc1*	*RAD10*	Unknown
XPG	Structure-specific endonuclease	5′ incision and stabilization of TFIIH	*Ercc5* (also known as *Xpg*)	*RAD2*	*UvrC*
PCNA	DNA replication sliding clamp	Three subunits; contains docking sites for DNA pol	*Pcna*	*PCNA*	β-clamp
RFC1	Loads PCNA on DNA	RFC large subunit	*Rfc1*	*CDC44*	Unknown
Unknown	Removal of incised oligo	None	Unknown	Unknown	*UvrD*‖
DNA pol δ or ε	DNA replication and repair	None	DNA pol δ or ε	DNA pol δ or ε	DNA pol I
DNA pol κ	Bypass polymerase	None	DNA pol κ	DNA pol κ	*Dinb1*
DNA ligase-I	Ligase	None	Unknown	*CDC9*	Ligase
DNA ligase-III	Ligase complex	Might be associated with dividing or non-dividing cells	Unknown	Unknown	Ligase
XRCC1	Ligase complex	Might be associated with dividing or non-dividing cells	*Xrcc1*	Unknown	Unknown
TCR genes					
CSA	Ubiquitin-ligase complex	WD repeat	*Ercc8* (also known as *Csa*)	*RAD28*	Unknown
CSB	TCR coupling factor and chromatin remodelling	Transcription elongation factor	*Ercc6* (also known as *Csb*)	*RAD26*	*Mfd*
XAB2	Transcription factor	Link between XPA and RNAPII	Unknown	Unknown	Unknown
TFIIS	RNAPII elongation factor	Stimulation of transcript cleavage by RNAPII	Unknown	Unknown	*GreA* and *GreB*
HMGN1	Chromatin relaxation	Nucleosome removal?	Unknown	Unknown	Unknown
p300	Chromatin remodelling	Nucleosome removal?	Unknown	Unknown	Unknown

*Complexes are indicated by the same adjacent background colours. Other gene products, such as FEN1, MMS19L and p53, might participate in NER or induce expression of NER genes. ‡Genetic and functional homologues. §Although only XPB and XPD (and TTDA, which is not listed) have NER functions, all ten TFIIH subunits are essential for NER. ‖There are no mammalian homologues of UvrD; its role in removing the damaged oligonucleotide might be carried out by DNA polymerases. *ERCC*, excision repair cross-complementing; CDC, cell-division cycle; GGR, global genomic repair; NER, nucleotide excision repair; *PCNA*, proliferating cell nuclear antigen; Pol, polymerase; *Ssb*, single-stranded DNA-binding protein; ssDNA, single-stranded DNA; *RF*, replication factor; RNAP, RNA polymerase; *RPA*, replication protein A; TCR, transcription-coupled repair; TF, transcription factor; XAB2, XPA-binding protein-2; XP, xeroderma pigmentosum.

excision or repair by DNA polymerase-β is precluded by the formation of a covalent polymerase–DNA crosslink[110]. This strongly implies that NER is required for 2′-deoxyribonolactone removal, and that TCR (through NER rather than BER) might operate *in vivo* for removal of 2′-deoxyribonolactone lesions, as well as abasic sites. However, expression levels of a reporter gene containing abasic sites (or strand breaks) is similar in wild-type and CSB cells, indicating that these lesions might not be repaired by TCR[97].

Human diseases exhibiting TCR deficiency

An understanding of the mechanism of TCR might have profound consequences for human health, particularly in relation to cancer and ageing. Several autosomal recessive hereditary diseases that are characterized by sun sensitivity involve deficiencies in NER, including the classic form of XP, CS, trichothiodystrophy (TTD) and UV-sensitive syndrome (UV^SS)[2] (TABLE 1).

XP is characterized by severe photosensitivity, dry parchment-like skin and high cancer incidence in sun-exposed areas, such as the skin, tongue and eyes. The incidence of internal tumours is also increased; for example, there is a tenfold increased risk of nervous system tumours in XP[111]. Some patients present progressive neurological degeneration, which was proposed to be caused by poor NER of certain oxidative DNA lesions in the brain[112]. Gene products from XP complementation groups (A–G) have roles in NER (FIG. 2; TABLE 2).

CS is characterized by sun sensitivity, microcephaly, developmental delay, contractures, cold extremities, gait ataxia, short stature and leukodystrophy. Various other manifestations, such as deafness, optic atrophy, cataracts, calcium deposits in the brain and cerebellar tissue, typical facies with sunken eyes, beaked nose and large ears, osteoporosis and dental caries appear in some patients. Most patients with CS belong to complementation groups A and B; in three additional CS complementation groups with certain mutations in *XPB*, *XPD* or *XPG*, the patients manifest symptoms of both CS and XP (XP–CS).

TTD is characterized by sulphur-deficient hair, ichthyosis, mental and physical retardation, and photosensitivity in some cases. Among the ~50% of patients with TTD who are sun sensitive, a defect in NER has been ascribed to mutations in TFIIH subunits, XPD, XPB or TTDA. Mutations in *TTDN1* or in other (unknown) genes without NER roles account for the disease in patients who are not photosensitive[113]. The cause for the particular set of symptoms of TTD is not known, although several theories have been proposed[114].

Patients with UV^SS are sensitive to sunlight and exhibit mild clinical manifestations, including acute sunburn, skin dryness, freckles in some cases, pigmental anomalies and telangiectasia. Patients present neither developmental nor neurological deficiencies. The six patients identified to date belong to two complementation groups, with mutations in the *CSB* gene or in a gene that has not been characterized[115]. The responses of UV^SS cells to UV are indistinguishable from those of CS cells: they exhibit normal levels of genomic repair of both principal UV photoproducts, CPDs and 6–4 photoproducts, but they are much more sensitive to UV than normal cells and they are deficient in the recovery of RNA synthesis after UV exposure. In addition, UV^SS cells are as sensitive as CS cells to accumulation of the tumour suppressor protein p53 after low doses of UV. We have shown that UV^SS cells are defective in TCR of CPD, and have postulated that the main difference between this syndrome and CS resides in their respective abilities to process certain oxidative base damage in transcriptionally active DNA[97,116].

Molecular models to explain disease aetiology. It is important to emphasize that the syndromes that are caused by a deficiency solely in TCR, CS and UV^SS, do not express enhanced susceptibility to cancer; in fact, it might be claimed that they are protected from cancer. The high levels of apoptosis in sun-damaged tissues of patients with CS or UV^SS are accompanied by proficient GGR in the surviving cells, minimizing mutagenesis and carcinogenesis. By contrast, in the case of XPC, with a near-normal apoptotic response, surviving cells are deficient in GGR and would therefore be expected to accumulate many mutations that could account for the observed enhanced carcinogenesis. Although it is generally believed that the high incidence of cancer in patients with XP is a consequence of defective GGR, patients with TTD with GGR defects are not cancer prone. Additional causes for the prevalent cancer rate in patients with XP might include deficient immune responses[117] or low catalase activity[118].

The aetiology of classic XP can be attributed to defects in NER (perhaps in combination with other deficiencies; see above), either GGR alone (owing to mutations in XPC or XPE) or both GGR and TCR. Likewise, the defect in TCR of UV-induced damage is causative for UV^SS. UV sensitivity and photophobia in patients with CS are also due to the TCR defect, but there is no biochemical explanation for the severe characteristics of CS. The mere absence of NER, such as in patients with XPA, does not result in CS; patients with XP–CS with deficiencies in both GGR and TCR exhibit the premature ageing and the neurological and developmental complications seen in patients with CSA and CSB in addition to symptoms typical of XP. Thus, CS is not caused by defective TCR of bulky DNA lesions.

Models proposed to explain CS. One hypothesis to explain CS proposes that CS cells are defective in TCR of certain oxidative lesions. The hypothesis postulates that active metabolism during early development generates high levels of oxidative damage in DNA; in patients with CS , transcription arrest at the consequent lesions leads to apoptosis, causing tissue degeneration. Arguments in support of and against this model are presented above.

An alternative hypothesis focuses on the possible roles of CSA, CSB and TFIIH proteins in transcription initiation and elongation. Studies *in vivo* and *in vitro* have yielded contradictory results regarding basal transcription efficiency by RNA polymerases I and II

Base excision repair
An excision repair pathway that is initiated by a glycosylase to remove a damaged or inappropriate base, leaving an abasic site. In the second step, the phosphodiester backbone is cleaved and a short patch (one nucleotide or more) is synthesized and ligated to the contiguous DNA.

Adduct
A covalently bound chemical residue, such as a methyl group or cisplatin, to a base or nucleotide in DNA. In some cases, the chemical binds more than one nucleotide, forming intrastrand or interstrand crosslinks.

Abasic site
A common form of DNA damage that is caused by the spontaneous loss of a base or an intermediate in the base excision repair pathway. This follows the removal of a base by glycosylase, while leaving the phosphodiester bond intact.

in CS cells[88]. The study of Rockx *et al.*[119] has shown unequivocally that CS cells cannot re-initiate transcription by RNAPII after UV irradiation, and that the cells do not recover from the UV-induced depletion of the hypophosphorylated form of the polymerase, RNAPIIa. Thus, CS has been considered a transcription disease.

A third model focuses on the highly specific depletion of certain tissues and organs in CS. Thus, defects in the transactivation of hormone receptors and nuclear receptors that are essential for controlling the activation of hormone-dependent genes that are involved in, for example, lipid metabolism, have been documented in NER-deficient mice[120–122].

An interesting dilemma concerns the case of two unrelated UV^{S}S individuals carrying the same mutation in the *CSB* gene, in whose cells no CSB protein has been detected. This finding suggested that the presence of mutated or truncated CSB caused the dramatic effects seen in patients with CS, whereas the complete absence of the CSB protein only resulted in sun sensitivity[115,123]. However, two additional patients with severe CS have been found to have mutations that abolish expression of CSB[124] and thus an alternative explanation for the mild symptoms of the individuals with UV^{S}S–CSB is needed. Newman *et al.*[125] suggested that unbalanced expression of CSB and the product of alternative splicing, a CSB–PiggyBac fusion protein, could be the underlying cause of CS, but this model does not explain cases in which the CSB truncation occurs upstream of the splice signal (thus no fusion protein is made), or CS owing to mutated CSA. CSB mouse models do not exhibit the severe phenotypes of their human counterparts, unless they are also deficient in XPA or XPC: interestingly, this observation supports the concept that CSB has important functions in addition to those in TCR.

Does TCR process unusual DNA structures?

The traditional view of transcription of an mRNA-encoding gene is that the transcript is synthesized at a constant rate following initiation at a promoter, until the termination signal is reached. We now know that initiation from the promoter, although essential for gene expression, is only an early step in the regulation of transcription, and the rate-limiting step for completion of the RNA product might involve promoter-proximal pausing[126]. Furthermore, there might be adventitious initiations of transcription downstream from the promoter, resulting in transcripts that are missing the early sequences and are usually terminated before reaching the late sequences. Little is known about the nature of the signals that determine these unusual and unexpected events: are they mostly accidents, or do they serve some useful (or indeed, essential) purpose? Also, what is the nature of the DNA sequences that regulate these events: could they be non-canonical DNA secondary structures, such as Z-DNA? Guanine-rich DNA sequences are remarkably frequent in the human genome: they are found in biologically relevant regions, such as telomeres, centromeres, immunoglobulin hypermutation hot spots and switch regions, and in the triplet nucleotide repeat expansions that are responsible for certain hereditary diseases[127]. A number of studies have documented the arrest of RNAPII, T7 RNAP or *E. coli* RNAP *in vitro* caused by non-canonical DNA structures[5,6,128–130].

Reports of an association of the RecQ-family helicase RecQ5 with RNAPII[131,132] are of considerable interest and potential relevance in light of the documented roles of the RecQ-family helicases in processing non-canonical DNA structures. Thus, the RQC domain, which is found only in RecQ-family helicases, confers a high affinity for G4 quadruplex DNA formations and contributes to the unwinding of those structures[133]. The association of RecQ helicases with RNAPII could serve in a 'cowcatcher' role to 'clear the tracks' during transcription elongation, thereby minimizing the disruption from non-canonical structures. Otherwise, the arrest of RNAPII at natural pause sites or sequence-dependent DNA secondary structures might occasionally result in 'gratuitous' TCR. The resulting futile cycles of repair replication might then cause significantly enhanced levels of mutagenesis in a frequently transcribed gene, possibly because of the likelihood that the short patches of nascent DNA might not be subject to the replication-related process of strand-directed mismatch repair[48]. The fidelity of repair replication might also be affected by the unusual DNA sequence itself, because of DNA polymerase 'slippage' or other problems with the structure. The RecQ helicase deficiencies that cause genetic instabilities, cancer predisposition and/or premature ageing in hereditary diseases, such as Bloom syndrome, Werner syndrome and Rothmund–Thomson syndrome, might involve deleterious consequences of gratuitous TCR.

TCR has been implicated in triplet nucleotide repeat contraction in human cells. Lin *et al.*[134] showed that small interfering RNA-mediated knockdown of ERCC1 or XPG, required for incisions, stabilized CAG tract lengths in the *HPRT* gene, whereas knockdown of XPC (which is required for lesion recognition in GGR) did not. Repeats were also stabilized by knockdown of CSB (which is required for TCR) or TFIIS[134,135]. Parsons *et al.*[136] had reported RNAPII pausing in triplet repeat-containing DNA (CTG and CGG, respectively) from human myotonic dystrophy and fragile X loci. Wang *et al.*[137] showed that oligonucleotides bound to DNA double helices — so-called triplex-forming oligonucleotides (TFOs) — induced mutations in a simian virus 40 vector in mammalian cells. This required both XPA and CSB, and thereby implicated TCR. It was then shown that repair of TFOs in transcribed genes also requires CSA, ribonucleotides and the transcription factors TFIID, TFIIH and RNAPII, although, unexpectedly, the reaction required XPC as well[138]. In a study of repair of site-specific mitomycin-C interstrand crosslinks in mammalian CSA or CSB cells, Zheng *et al.*[139] have implicated the involvement of TCR, with initial incisions in the transcribed DNA strand. There is much more to learn about the role of TCR in transcription through unusual structures, its role in processing lesions in these structures and, in general, how the expression of genes that contain such structures might be affected by TCR.

Z-DNA
A non-canonical form of DNA that can occur in stretches of alternating guanine and cytosine bases, in which the strands are wound in a left-handed manner, to form a zigzag backbone structure.

Triplet nucleotide repeat expansions
The phenomenon that is characteristic in over a dozen human genetic diseases, in which massive expansions occur in a particular genomic region of repetitive nucleotide triplets, such as CAG in the case of Friedreich's ataxia.

RecQ-family helicase
One of a family of evolutionarily conserved helicases that were first discovered in *Escherichia coli*. Mutations in three of the five RecQ homologues in humans can cause cancer predisposition and/or premature ageing, such as in Werner's syndrome.

Conclusions and future perspectives

What could be the essential role of TCR? Although CS cells in culture are UV-sensitive and undergo apoptosis at low UV doses, normal or near-normal embryonic development is possible in humans with deficiencies in CSA or CSB. However, the severe type II form of CS expresses multiple birth defects, including intrauterine growth restriction and congenital cataracts[140]. The low birth weight and eventual developmental deficiencies that are typical of CS cannot be a consequence of a defect in TCR processing of bulky lesions that arrest transcription, as the TCR defect in UV^{S}S presents no developmental problems. It is therefore important to learn the nature of the difference between CS and UV^{S}S. Another challenging problem is to understand the reason for the lack of skin cancer in diseases that express defects solely in TCR.

The fact that transcription has been shown to increase DNA damage-induced mutagenesis in mouse embryonic stem cells is also relevant[141]. Detailed analysis indicated an enhanced frequency of nucleotide substitutions and intragenic deletions, and this is suggested to be due to error-prone double-strand breaks that are generated when replication forks encounter stalled transcription complexes at lesions. This relates to the likely importance of TCR in preventing collisions between replication forks and arrested transcription complexes (see above)[8].

Some types of lesions, such as those caused by the anti-tumour drug irofulven, have been reported to be subject to TCR but not to GGR[142]. Irofulven forms covalent adducts with guanine and adenine[143] and has been shown to be effective in the treatment of prostate cancer; its cytotoxicity in solid tumour cells evidently depends on TCR[144]. The anticancer drug ET743 interacts with the TCR pathway to produce lethal strand breaks[145]. Thus, inhibition of TCR in combination with irofulven or with ET743 might be used to control gene expression or to eliminate cells in which transcription is severely impaired. It is timely to explore the translational opportunities that arise from an increased mechanistic understanding of TCR.

Our knowledge of TCR has benefited from cooperative interactions between many research groups, as the specialties of DNA repair and transcription are synergistically combined. However, not surprisingly, the more we have learnt, the more questions have been raised about the intricate details of TCR and its relevance to human disease. The primary tools for analysis continue to be derived from the genetic and biochemical investigation of specific gene defects in human genetic diseases, such as CS and UV^{S}S. Our continuing goal is to understand molecular mechanisms, to thereby enable the development of effective therapies for some sets of human health problems.

1. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).
2. Friedberg, E. *et al. DNA Repair and Mutagenesis* (ASM Press, Washington DC, 2006).
A comprehensive text that covers all aspects of DNA repair, mutagenesis and human disease related to these topics.
3. Wang, G. & Vasquez, K. M. Non-B DNA structure-induced genetic instability. *Mutat. Res.* **598**, 103–119 (2006).
4. Wells, R. D. Non-B DNA conformations, mutagenesis and disease. *Trends Biochem. Sci.* **32**, 271–278 (2007).
5. Ditlevson, J. V. *et al.* Inhibitory effect of a short Z-DNA forming sequence on transcription elongation by T7 RNA polymerase. *Nucleic Acids Res.* **36**, 3163–3170 (2008).
6. Tornaletti, S., Park-Snyder, S. & Hanawalt, P. C. G4-forming sequences in the non-transcribed DNA strand pose blocks to T7 RNA polymerase and mammalian RNA polymerase II. *J. Biol. Chem.* **283**, 12756–12762 (2008).
7. Voineagu, I., Narayanan, V., Lobachev, K. & Mirkin, S. Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. *Proc. Natl Acad. Sci. USA* **105**, 9936–9941 (2008).
8. Hanawalt, P. C. Transcription-coupled repair and human disease. *Science* **266**, 1957–1958 (1994).
9. Harper, J. W. & Elledge, S. J. The DNA damage response: ten years after. *Mol. Cell* **28**, 739–745 (2007).
10. Jiang, G. & Sancar, A. Recruitment of DNA damage checkpoint proteins to damage in transcribed and nontranscribed sequences. *Mol. Cell. Biol.* **26**, 39–49 (2006).
11. Ljungman, M. Activation of DNA damage signaling. *Mutat. Res.* **577**, 203–216 (2005).
12. Hanawalt, P. C. Paradigms for the three Rs: DNA replication, recombination, and repair. *Mol. Cell* **28**, 702–707 (2007).
13. Yang, W. & Woodgate, R. What a difference a decade makes: insights into translesion DNA synthesis. *Proc. Natl Acad. Sci. USA* **104**, 15591–15598 (2007).
14. Eisen, J. A. & Hanawalt, P. C. A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat. Res.* **435**, 171–213 (1999).
15. Raghavan, S. C. & Lieber, M. R. DNA structure and human diseases. *Front. Biosci.* **12**, 4402–4408 (2007).
16. Cleaver, J. E. Defective repair replication of DNA in xeroderma pigmentosum. *Nature* **218**, 652–656 (1968).
17. Zolan, M. E., Cortopassi, G. A., Smith, C. A. & Hanawalt, P. C. Deficient repair of chemical adducts in α DNA of monkey cells. *Cell* **28**, 613–619 (1982).
18. Mansbridge, J. N. & Hanawalt, P. C. in *Cellular Responses to DNA Damage, UCLA Symposium on Molecular and Cellular Biology* (eds Friedberg, E. & Bridges, B.) 195–207 (Alan R. Liss, New York, 1983).
19. Mayne, L. V. & Lehmann, A. R. Failure of RNA synthesis to recover after UV irradiation: an early defect in cells from individuals with Cockayne's syndrome and xeroderma pigmentosum. *Cancer Res.* **42**, 1473–1478 (1982).
Suggests that CS cells might be defective in repairing expressed genes.
20. Kantor, G. J., Barsalou, L. S. & Hanawalt, P. C. Selective repair of specific chromatin domains in UV-irradiated cells from xeroderma pigmentosum complementation group C. *Mutat. Res.* **235**, 171–180 (1990).
21. Venema, J., van Hoffen, A., Natarajan, A. T., van Zeeland, A. A. & Mullenders, L. H. F. The residual repair capacity of xeroderma pigmentosum complementation group C fibroblasts is highly specific for transcriptionally active DNA. *Proc. Natl Acad. Sci. USA* **18**, 443–448 (1990).
Shows that XPC cells are deficient in GGR but proficient in TCR.
22. Venema, J., Mullenders, L. H. F., Natarajan, A. T., van Zeeland, A. A. & Mayne, L. V. The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. *Proc. Natl Acad. Sci. USA* **87**, 4707–4711 (1991).
Documents the TCR defect in CS.
23. Lommel, L. & Hanawalt, P. C. The genetic defect in the Chinese hamster ovary cell mutant UV61 permits moderate selective repair of cyclobutane pyrimidine dimers in an expressed gene. *Mutat. Res.* **255**, 183–191 (1991).
24. Troelstra, C. *et al.* ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. *Cell* **71**, 939–953 (1992).
Identifies the gene that causes CSB.
25. Hanawalt, P. in *DNA Repair Mechanisms, Alfred Benzon Symposium 35* (eds Bohr, V. A., Wassermann, K. & Kraemer, K. H.) 231–242 (Munksgaard, Copenhagen,1992).
26. Hanawalt, P. C. Revisiting the rodent repairadox. *Environ. Mol. Mutagen* **38**, 89–96 (2001).
27. Bohr, V. A., Smith, C. A., Okumoto, D. S. & Hanawalt, P. C. DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient that in the genome overall. *Cell* **40**, 359–369 (1985).
Reports the preferential repair of an expressed gene in mammalian cells.
28. Spivak, G., Pfeifer, G. P. & Hanawalt, P. C. in *Methods in Enzymology: DNA Repair* (eds Campbell, J. C. & Modrich, P.) 223–246 (Elsevier, New York, 2006).
29. Mellon, I., Spivak, G. & Hanawalt, P. C. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**, 241–249 (1987).
Shows that preferential repair of an active gene in mammalian cells is targeted to the transcribed DNA strand.
30. van Hoffen, A., Venema, J., Meschini, R., van Zeeland, A. A. & Mullenders, L. H. Transcription-coupled repair removes both cyclobutane pyrimidine dimers and 6–4 photoproducts with equal efficiency and in a sequential way from transcribed DNA in xeroderma pigmentosum group C fibroblasts. *EMBO J.* **14**, 360–367 (1995).
31. van Oosterwijk, M. F., Filon, R., Kalle, W. H., Mullenders, L. H. & van Zeeland, A. A. The sensitivity of human fibroblasts to *N*-acetoxy-2-acetylaminofluorene is determined by the extent of transcription-coupled repair, and/or their capability to counteract RNA synthesis inhibition. *Nucleic Acids Res.* **24**, 4653–4659 (1996).
32. Vrieling, H. *et al.* Strand specificity for UV-induced DNA repair and mutations in the Chinese hamster HPRT gene. *Nucleic Acids Res.* **19**, 2411–2415 (1991).
33. Mellon, I. & Hanawalt, P. C. Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* **342**, 95–98 (1989).
Documents transcription-coupled repair in bacteria.
34. Smerdon, M. J. & Thoma, F. Site-specific DNA-repair at the nucleosome level in a yeast minichromosome. *Cell* **61**, 675–684 (1990).
35. Sweder, K. S. & Hanawalt, P. C. Preferential repair of cyclobutane pyrimidine dimers in the transcribed strand of a gene in yeast chromosomes and plasmids is dependent on transcription. *Proc. Natl Acad. Sci. USA* **89**, 10696–10700 (1992).

36. Christians, F. C. & Hanawalt, P. C. Inhibition of transcription and strand-specific DNA repair by α-amanitin in Chinese hamster ovary cells. *Mutat. Res.* **274**, 93–101 (1992).
37. Bootsma, D. & Hoeijmakers, J. H. DNA repair. Engagement with transcription. *Nature* **363**, 114–115 (1993).
38. Schaeffer, L. *et al*. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* **260**, 58–63 (1993).
Identifies the NER proteins XPB and XPD as subunits of the basal transcription factor TFIIH.
39. Selby, C. P. & Sancar, A. Mechanisms of transcription-repair coupling and mutation frequency decline. *Microbiol. Rev.* **58**, 317–329 (1994).
An excellent account of the mechanism of TCR in bacteria.
40. Eisen, J. A., Sweder, K. S. & Hanawalt, P. C. Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* **23**, 2715–2723 (1995).
41. Park, J. S., Marr, M. T. & Roberts, J. W. *E. coli* transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* **109**, 757–767 (2002).
42. Park, J. S. & Roberts, J. W. Role of DNA bubble rewinding in enzymatic transcription termination. *Proc. Natl Acad. Sci. USA* **103**, 4870–4875 (2006).
43. Savery, N. J. The molecular mechanism of transcription-coupled DNA repair. *Trends Microbiol.* **15**, 326–333 (2007).
44. Ganesan, A. K., Smith, A. J., Savery, N. J., Zamos, P. & Hanawalt, P. C. Transcription coupled nucleotide excision repair in *Escherichia coli* can be affected by changing the arginine at position 529 of the β subunit of RNA polymerase. *DNA Repair* **6**, 1434–1440 (2007).
45. Crowley, D. J. & Hanawalt, P. C. Induction of the SOS response increases the efficiency of global nucleotide excision repair of cyclobutane pyrimidine dimers, but not 6–4 photoproducts, in UV-irradiated *Escherichia coli*. *J. Bacteriol.* **180**, 3345–3352 (1998).
46. Courcelle, J., Khodursky, A., Peter, B., Brown, P. O. & Hanawalt, P. C. Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* **158**, 41–64 (2001).
47. Branum, M. E., Reardon, J. T. & Sancar, A. DNA repair excision nuclease attacks undamaged DNA. A potential source of spontaneous mutations. *J. Biol. Chem.* **276**, 25421–25426 (2001).
48. Hanawalt, P. C. Subpathways of nucleotide excision repair and their regulation. *Oncogene* **21**, 8949–8956 (2002).
49. Svejstrup, J. Q. Contending with transcriptional arrest during RNAPII transcript elongation. *Trends Biochem. Sci.* **32**, 165–171 (2007).
A good review of TCR in yeast and in eukaryotes in general.
50. Selby, C. P. & Sancar, A. Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *J. Biol. Chem.* **272**, 1885–1890 (1997).
51. van Gool, A. J. *et al*. The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerase II-containing complex. *EMBO J.* **16**, 5955–5965 (1997).
52. Selby, C. P., Drapkin, R., Reinberg, D. & Sancar, A. RNA polymerase II stalled at a thymine dimer: footprint and effect on excision repair. *Nucleic Acids Res.* **25**, 787–793 (1997).
53. Mu, D. & Sancar, A. Model for XPC-independent transcription-coupled repair of pyrimidine dimers in humans. *J. Biol. Chem.* **272**, 7570–7573 (1997).
54. Fitch, M. E., Nakajima, S., Yasui, A. & Ford, J. M. *In vivo* recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *J. Biol. Chem.* **278**, 46906–46910 (2003).
55. Ford, J. M. Regulation of DNA damage recognition and nucleotide excision repair: another role for p53. *Mutat. Res.* **577**, 195–202 (2005).
56. Sugasawa, K. *et al*. Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol. Cell* **2**, 223–232 (1998).
57. Bowman, K. K., Smith, C. A. & Hanawalt, P. C. Excision-repair patch lengths are similar for transcription-coupled repair and global genome repair in UV-irradiated human cells. *Mutat. Res.* **385**, 95–105 (1997).
58. Tornaletti, S. & Hanawalt, P. C. Effect of DNA lesions on transcription elongation. *Biochimie* **81**, 139–146 (1999).
59. Ljungman, M. & Zhang, F. Blockage of RNA polymerase as a possible trigger for u.v. light-induced apoptosis. *Oncogene* **13**, 823–831 (1996).
60. Yamaizumi, M. & Sugano, T. U.v-induced nuclear accumulation of p53 is evoked through DNA damage of actively transcribed genes independent of the cell cycle. *Oncogene* **9**, 2775–2784 (1994).
61. Saxowsky, T. & Doetsch, P. W. RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis? *Chem. Rev.* **106**, 474–488 (2006).
62. Donahue, B. A., Yin, S., Taylor, J.-S., Reines, D. & Hanawalt, P. Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc. Natl Acad. Sci. USA* **91**, 8502–8506 (1994).
Shows that a CPD is a strong block to RNAPII elongation, and describes the TFIIS role in transcript cleavage related to RNAPII backtracking from a lesion.
63. Tornaletti, S., Maeda, L. S., Lloyd, D. R., Reines, D. & Hanawalt, P. C. Effect of thymine glycol on transcription elongation by T7 RNA polymerase and mammalian RNA polymerase II. *J. Biol. Chem.* **276**, 45367–45371 (2001).
64. Anindya, R., Aygun, O. & Svejstrup, J. Q. Damage-induced ubiquitylation of human RNA polymerase II by the ubiquitin ligase Nedd4, but not Cockayne syndrome proteins or BRCA1. *Mol. Cell* **28**, 386–397 (2007).
65. Malik, S., Bagla, S., Chaurasia, P., Duan, Z. & Bhaumik, S. R. Elongating RNA polymerase II is disassembled through specific degradation of its largest but not other subunits in response to DNA damage *in vivo*. *J. Biol. Chem.* **283**, 6897–6905 (2008).
66. Tennyson, C. N., Klamut, H. J. & Worton, R. G. The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nature Genet.* **9**, 184–190 (1995).
67. Courcelle, J. & Hanawalt, P. C. RecA-dependent recovery of arrested DNA replication forks. *Annu. Rev. Genet.* **37**, 611–646 (2003).
68. Fousteri, M. & Mullenders, L. H. Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell Res.* **18**, 73–84 (2008).
69. Sarker, A. *et al*. Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH; insights for transcription-coupled repair and Cockayne syndrome. *Mol. Cell* **20**, 187–198 (2005).
70. Brueckner, F. & Cramer, P. DNA photodamage recognition by RNA polymerase II. *FEBS Lett.* **581**, 2757–2760 (2007).
71. Brueckner, F., Hennecke, U., Carell, T. & Cramer, P. CPD damage recognition by transcribing RNA polymerase II. *Science* **315**, 859–862 (2007).
Reports the crystal structure of a CPD in the RNAP active site.
72. Damsma, G. E., Alt, A., Brueckner, F., Carell, T. & Cramer, P. Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nature Struct. Mol. Biol.* **14**, 1127–1133 (2007).
73. Iyer, N., Reagan, M. S., Wu, K. J., Canagarajah, B. & Friedberg, E. C. Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochemistry* **35**, 2157–2167 (1996).
74. Fousteri, M., Vermeulen, W., van Zeeland, A. A. & Mullenders, L. H. Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II *in vivo*. *Mol. Cell* **23**, 471–482 (2006).
Identifies the roles of CSA and CSB in recruiting proteins and factors to arrested RNAPII.
75. Kamiuchi, S. *et al*. Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proc. Natl Acad. Sci. USA* **99**, 201–206 (2002).
76. Koehler, D. R. & Hanawalt, P. C. Recruitment of damaged DNA to the nuclear matrix in hamster cells following ultraviolet irradiation. *Nucleic Acids Res.* **24**, 2877–2884 (1996).
77. Kuraoka, I. *et al*. RNA polymerase II bypasses 8-oxoguanine in the presence of transcription elongation factor TFIIS. *DNA Repair* **6**, 841–851 (2007).
78. Wong, J. & Ingles, C. A compromised yeast RNA polymerase II enhances UV sensitivity in the absence of global genome nucleotide excision repair. *Mol. Gen. Genet.* **264**, 842–851 (2001).
79. Lim, J. H. *et al*. Chromosomal protein HMGN1 enhances the acetylation of lysine 14 in histone H3. *EMBO J.* **24**, 3038–3048 (2005).
80. Birger, Y. *et al*. Chromosomal protein HMGN1 enhances the rate of DNA repair in chromatin. *EMBO J.* **22**, 1665–1675 (2003).
81. Hasan, S. *et al*. Regulation of human flap endonuclease-1 activity by acetylation through the transcriptional coactivator p300. *Mol. Cell* **7**, 1221–1231 (2001).
82. Cazzalini, O. *et al*. Interaction of p21(CDKN1A) with PCNA regulates the histone acetyltransferase activity of p300 in nucleotide excision repair. *Nucleic Acids Res.* **36**, 1713–1722 (2008).
83. Green, C. M. & Almouzni, G. Local action of the chromatin assembly factor CAF-1 at sites of nucleotide excision repair *in vivo*. *EMBO J.* **22**, 5163–5174 (2003).
84. Dammann, R. & Pfeifer, G. P. Lack of gene- and strand-specific DNA repair in RNA polymerase III-transcribed human tRNA genes. *Mol. Cell. Biol.* **17**, 219–229 (1997).
85. Hackett, P. B. & Sauerbier, W. The transcriptional organization of the ribosomal RNA genes in mouse L cells. *J. Mol. Biol.* **91**, 235–256 (1975).
86. Christians, F. C. & Hanawalt, P. C. Lack of transcription-coupled repair in mammalian ribosomal RNA genes. *Biochemistry* **32**, 10512–10518 (1993).
87. Conconi, A., Bespalov, V. & Smerdon, M. Transcription-coupled repair in RNA polymerase I-transcribed genes of yeast. *Proc. Natl Acad. Sci. USA* **99**, 649–654 (2002).
88. Bradsher, J. *et al*. CSB is a component of RNA pol I transcription. *Mol. Cell* **10**, 819–829 (2002).
89. Meier, A., Livingstone-Zatchej, M. & Thoma, F. Repair of active and silenced rDNA in yeast: the contributions of photolyase and transcription-coupled nucleotide excision repair. *J. Biol. Chem.* **277**, 11845–11852 (2002).
90. Lindsey-Boltz, L. A. & Sancar, A. RNA polymerase: the most specific damage recognition protein in cellular responses to DNA damage? *Proc. Natl Acad. Sci. USA* **104**, 12778–12783 (2007).
91. Reines, D. Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J. Biol. Chem.* **267**, 3795–3800 (1992).
92. Tornaletti, S. Transcription arrest at DNA damage sites. *Mutat. Res.* **577**, 131–145 (2005).
93. Zhou, W., Reines, D. & Doetsch, P. W. T7 RNA polymerase bypass of large gaps on the template strand reveals a critical role of the nontemplate strand in elongation. *Cell* **82**, 577–585 (1995).
94. Liu, J. & Doetsch, P. W. Template strand gap bypass is a general property of prokaryotic RNA polymerases: implications for elongation mechanisms. *Biochemistry* **35**, 14999–15008 (1996).
95. Islas, A. L., Vos, J. M. & Hanawalt, P. C. Differential introduction and repair of psoralen photoadducts to DNA in specific human genes. *Cancer Res.* **51**, 2867–2873 (1991).
96. Cooper, P. K. & Leadon, S. A. Defective repair of ionizing radiation damage in Cockayne's syndrome and xeroderma pigmentosum group G. *Ann. NY Acad. Sci.* **726**, 330–332 (1994).
97. Spivak, G. & Hanawalt, P. C. Host cell reactivation of plasmids containing oxidative DNA lesions is defective in Cockayne syndrome but normal in UV-sensitive syndrome cells. *DNA Repair* **5**, 13–22 (2006).
98. Ropolo, M. *et al*. Complementation of the oxidatively damaged DNA repair defect in Cockayne syndrome A and B cells by *Escherichia coli* formamidopyrimidine DNA glycosylase. *Free Radic. Biol. Med.* **42**, 1807–1817 (2007).
99. Pastoriza-Gallego, M., Armier, J. & Sarasin, A. Transcription through 8-oxoguanine in DNA repair-proficient and Csb^{-}/$Ogg1^{-}$ DNA repair-deficient mouse embryonic fibroblasts is dependent upon promoter strength and sequence context. *Mutagenesis* **22**, 343–351 (2007).
100. Tornaletti, S., Maeda, L. S., Kolodner, R. D. & Hanawalt, P. C. Effect of 8-oxoguanine on transcription elongation by T7 RNA polymerase and mammalian RNA polymerase II. *DNA Repair* **3**, 483–494 (2004).

101. Thorslund, T., Sunesen, M., Bohr, V. A. & Stevnsner, T. Repair of 8-oxoG is slower in endogenous nuclear genes than in mitochondrial DNA and is without strand bias. *DNA Repair* **1**, 261–273 (2002).
102. Scicchitano, D. A. & Hanawalt, P. C. Lack of sequence-specific removal of *N*-methylpurines from cellular DNA. *Mutat. Res.* **233**, 31–37 (1990).
103. Brāegeon, D., Doddridge, Z. A., You, H. J., Weiss, B. & Doetsch, P. W. Transcriptional mutagenesis induced by uracil and 8-oxoguanine in *Escherichia coli*. *Mol. Cell* **12**, 959–970 (2003).
104. Charlet-Berguerand, N. *et al*. RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *EMBO J.* **25**, 5481–5491 (2006).
105. Marnett, L. J. Lipid peroxidation–DNA damage by malondialdehyde. *Mutat. Res.* **424**, 83–95 (1999).
106. Brooks, P. J. The case for 8,5′-cyclopurine-2′-deoxynucleosides as endogenous DNA lesions that cause neurodegeneration in xeroderma pigmentosum. *Neuroscience* **145**, 1407–1417 (2007).
107. Cline, S. D., Riggins, J. N., Tornaletti, S., Marnett, L. J. & Hanawalt, P. C. Malondialdehyde adducts in DNA arrest transcription by T7 RNA polymerase and mammalian RNA polymerase II. *Proc. Natl Acad. Sci. USA* **101**, 7275–7280 (2004).
108. Tornaletti, S., Maeda, L. S. & Hanawalt, P. C. Transcription arrest at an abasic site in the transcribed strand of template DNA. *Chem. Res. Toxicol.* **19**, 1215–1220 (2006).
109. Wang, Y., Sheppard, T. L., Tornaletti, S., Maeda, L. S. & Hanawalt, P. C. Transcriptional inhibition by an oxidized abasic site in DNA. *Chem. Res. Toxicol.* **19**, 234–241 (2006).
110. Demple, B. & DeMott, M. S. Dynamics and diversions in base excision DNA repair of oxidized abasic lesions. *Oncogene* **21**, 8926–8934 (2002).
111. Kraemer, K. H., Lee, M. M. & Scotto, J. Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases. *Arch. Dermatol.* **123**, 241–250 (1987).
112. Satoh, M. S., Jones, C. J., Wood, R. D. & Lindahl, T. DNA excision-repair defect of xeroderma pigmentosum prevents removal of a class of oxygen free radical-induced base lesions. *Proc. Natl Acad. Sci. USA* **90**, 6335–6339 (1993).
113. Botta, E. *et al*. Mutations in the *C7orf11* (*TTDN1*) gene in six nonphotosensitive trichothiodystrophy patients: no obvious genotype–phenotype relationships. *Hum. Mutat.* **28**, 92–96 (2007).
114. de Boer, J. & Hoeijmakers, J. H. Nucleotide excision repair and human syndromes. *Carcinogenesis* **21**, 453–460 (2000).
115. Spivak, G. UV-sensitive syndrome. *Mutat. Res.* **577**, 162–169 (2005).
116. Spivak, G. *et al*. Ultraviolet-sensitive syndrome cells are defective in transcription-coupled repair of cyclobutane pyrimidine dimers. *DNA Repair* **1**, 629–643 (2002).
117. Gaspari, A. A., Fleisher, T. A. & Kraemer, K. H. Impaired interferon production and natural killer cell activation in patients with the skin cancer-prone disorder, xeroderma pigmentosum. *J. Clin. Invest.* **92**, 1135–1142 (1993).
118. Quilliet, X. *et al*. Retroviral-mediated correction of DNA repair defect in xeroderma pigmentosum cells is associated with recovery of catalase activity. *Mutat. Res.* **385**, 235–242 (1997).
119. Rockx, A. *et al*. UV-induced inhibition of transcription involves repression of transcription initiation and phosphorylation of RNA polymerase II. *Proc. Natl Acad. Sci. USA* **97**, 10503–10508 (2000).
120. Rapin, I., Lindenbaum, Y., Dickson, D. W., Kraemer, K. H. & Robbins, J. H. Cockayne syndrome and xeroderma pigmentosum. *Neurology* **55**, 1442–1449 (2000).
121. Keriel, A., Stary, A., Sarasin, A., Rochette-Egly, C. & Egly, J. M. *XPD* mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARα. *Cell* **109**, 125–135 (2002).
122. Compe, E. *et al*. Dysregulation of the peroxisome proliferator-activated receptor target genes by *XPD* mutations. *Mol. Cell. Biol.* **25**, 6065–6076 (2005).
123. Horibata, K. *et al*. Complete absence of Cockayne syndrome group B gene product gives rise to UV-sensitive syndrome but not Cockayne syndrome. *Proc. Natl Acad. Sci. USA* **101**, 15410–15415 (2004).
124. Laugel, V. *et al*. Deletion of 5′ sequences of the CSB gene provides insight into the pathophysiology of Cockayne syndrome. *Eur. J. Hum. Genet.* **16**, 320–327 (2008).
125. Newman, J. C., Bailey, A. D., Fan, H. Y., Pavelitz, T. & Weiner, A. M. An abundant evolutionarily conserved CSB–PiggyBac fusion protein expressed in Cockayne syndrome. *PLoS Genet.* **4**, e1000031 (2008).
126. Core, L. J. & Lis, J. T. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319**, 1791–1792 (2008).
127. Phan, A. T., Kuryavyi, V. & Patel, D. J. DNA architecture: from G to Z. *Curr. Opin. Struct. Biol.* **16**, 288–298 (2006).
128. Duquette, M. L., Handa, P., Vincent, J. A., Taylor, A. F. & Maizels, N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* **18**, 1618–1629 (2004).
129. Belotserkovskii, B. P. *et al*. A triplex-forming sequence from the human *c-MYC* promoter interferes with DNA transcription. *J. Biol. Chem.* **282**, 32433–32441 (2007).
130. Peck, L. J. & Wang, J. C. Transcriptional block caused by a negative supercoiling induced structural change in an alternating CG sequence. *Cell* **40**, 129–137 (1985).
131. Aygün, O., Svejstrup, J. & Liu, Y. A RECQ5–RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proc. Natl Acad. Sci. USA* **105**, 8580–8584 (2008).
132. Izumikawa, K. *et al*. Association of human DNA helicase RecQ5β with RNA polymerase II and its possible role in transcription. *Biochem. J.* **413**, 505–516 (2008).
133. Huber, M. D., Duquette, M. L., Shiels, J. C. & Maizels, N. A conserved G4 DNA binding domain in RecQ family helicases. *J. Mol. Biol.* **358**, 1071–1080 (2006).
134. Lin, Y., Dion, V. & Wilson, J. H. Transcription promotes contraction of CAG repeat tracts in human cells. *Nature Struct. Mol. Biol.* **13**, 179–180 (2006).
135. Lin, Y. & Wilson, J. H. Transcription-induced CAG repeat contraction in human cells is mediated in part by trancription-coupled nucleotide excision repair. *Mol. Cell. Biol.* **27**, 6209–6217 (2007).
136. Parsons, M. A., Sinden, R. R. & Izban, M. G. Transcriptional properties of RNA polymerase II within triplet repeat-containing DNA from the human myotonic dystrophy and fragile X loci. *J. Biol. Chem.* **273**, 26998–27008 (1998).
137. Wang, G., Seidman, M. M. & Glazer, P. M. Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science* **271**, 802–805 (1996).
138. Wang, G., Chen, Z., Zhang, S., Wilson, G. L. & Jing, K. Detection and determination of oligonucleotide triplex formation-mediated transcription-coupled DNA repair in HeLa nuclear extracts. *Nucleic Acids Res.* **29**, 1801–1807 (2001).
139. Zheng, H. *et al*. Nucleotide excision repair- and polymerase η-mediated error-prone removal of mitomycin C interstrand cross-links. *Mol. Cell. Biol.* **23**, 754–761 (2003).
140. Nance, M. A. & Berry, S. A. Cockayne syndrome: review of 140 cases. *Am. J. Med. Genet.* **42**, 68–84 (1992).
141. Hendriks, G. *et al*. Gene transcription increases DNA damage-induced mutagenesis in mammalian stem cells. *DNA Repair* **7**, 1330–1339 (2008).
142. Escargueil, A. E. *et al*. Influence of irofulven, a transcription-coupled repair-specific antitumor agent, on RNA polymerase activity, stability and dynamics in living mammalian cells. *J. Cell Sci.* **121**, 1275–1283 (2008).
143. Neels, J. F., Gong, J., Yu, X. & Sturla, S. J. Quantitative correlation of drug bioactivation and deoxyadenosine alkylation by acylfulvene. *Chem. Res. Toxicol.* **20**, 1513–1519 (2007).
144. Koeppel, F. *et al*. Irofulven cytotoxicity depends on transcription-coupled nucleotide excision repair and is correlated with XPG expression in solid tumor cells. *Clin. Cancer Res.* **10**, 5604–5613 (2004).
145. Takebayashi, Y. *et al*. Antiproliferative activity of ecteinascidin 743 is dependent upon transcription-coupled nucleotide-excision repair. *Nature Med.* **7**, 961–966 (2001).
146. Verhoeven, E. E., Wyman, C., Moolenaar, G. F. & Goosen, N. The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands. *EMBO J.* **21**, 4196–4205 (2002).
147. Sancar, A. DNA excision repair. *Ann. Rev. Biochem.* **65**, 43–81 (1996).
148. Staresincic, L. *et al*. Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J.* **22**, 5293–5303 (2008).
149. Ogi, T. & Lehmann, A. R. The Y-family DNA polymerase κ (pol κ) functions in mammalian nucleotide-excision repair. *Nature Cell Biol.* **8**, 640–642 (2006).
150. Mocquet, V. *et al*. Sequential recruitment of the repair factors during NER: the role of XPG in initiating the resynthesis step. *EMBO J.* **27**, 155–167 (2008).
151. Moser, J. *et al*. Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III α in a cell-cycle-specific manner. *Mol. Cell* **27**, 311–323 (2007).

Acknowledgements
The authors thank A. Ganesan and C. A. Smith for helpful discussions, and the National Cancer Institute for research support.

DATABASES

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
Dhfr | *XPB* | *XPD* | *XPG*
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
Bloom syndrome | Cockayne syndrome | Rothmund–Thomson syndrome | trichothiodystrophy | UV-sensitive syndrome | Werner syndrome
UniProtKB: http://www.uniprot.org
CSB | DDB2 | HMGN1 | Mfd | NEDD4 | p300 | UvrA | UvrB | XAB2 | XPC

FURTHER INFORMATION

Philip C. Hanawalt's laboratory: http://www.stanford.edu/~hanawalt
DNA Repair Interest Group: http://tango01.cit.nih.gov/sig/home.taf?_function=main&SIGInfo_SIGID=32
HUGO: http://www.hugo-international.org/index.php
Human DNA repair genes:
http://www.cgal.icnet.uk/DNA_Repair_Genes.html
NIA Mouse Gene Index:
http://lgsun.grc.nia.nih.gov/geneindex5

ALL LINKS ARE ACTIVE IN THE ONLINE PDF