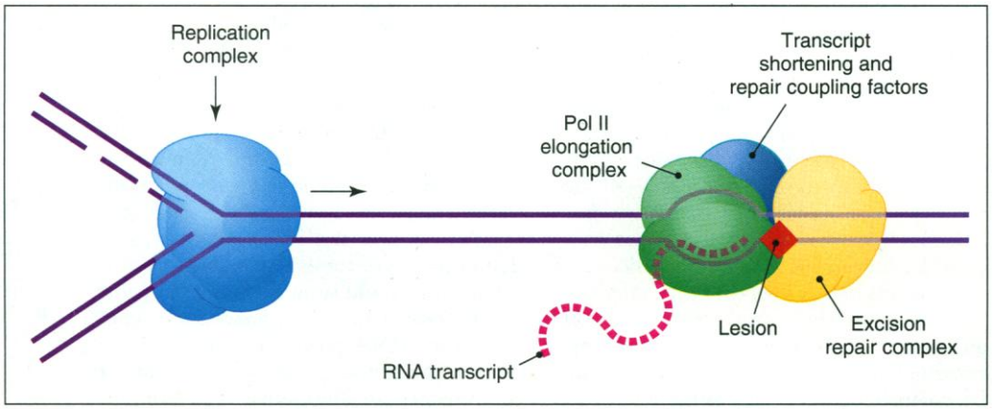
Transcription-Coupled Repair and Human Disease

Philip C. Hanawalt

Several recent discoveries have challenged the view that DNA repair and transcription are fully separable processes. Central to this conceptual shift are the findings that there is preferential repair of the transcribed DNA strand in expressed genes (*1*) and that components of a factor essential for transcription initiation, TFIIH, are also required for nucleotide excision repair (NER) throughout the genome as well as in expressed genes (*2*). These discoveries have led to the notion that some, but not all, NER may be coupled to transcription. Adding to the excitement and clinical relevance of these discoveries is the fact that several gene products implicated in this "transcription-coupled repair" are defective in three rare human hereditary disorders: xeroderma pigmentosum (XP), Cockayne's syndrome (CS), and trichothiodystrophy (TTD).

Patients with XP suffer a predisposition to cancer in sun-exposed skin as a consequence of their DNA-repair defect. Patients with CS or TTD are not predisposed to cancer, but exhibit developmental problems. In CS, these problems include growth retardation, neurological deficiencies, and skeletal abnormalities, whereas in TTD they include brittle (sulfur-deficient) hair, short stature, scaly skin, and mental underdevelopment. With the molecular mechanics of NER and transcription-coupled repair now unfolding, the question has been raised: Are these developmental problems due to defective transcription or can they be attributed to defects in repair?

The ubiquitous process of NER was discovered three decades ago through basic research on the response of the bacterium *Escherichia coli* to ultraviolet (UV) radiation. UV light causes the formation of cyclobutane pyrimidine dimers, DNA lesions that obstruct transcription and DNA replication. UV-induced cyclobutane pyrimidine dimers were shown to be excised from the DNA in the irradiated bacteria, and short stretches of new DNA synthesis were detected (*3*). Although this "cut and patch" mechanism for repair was soon confirmed in essentially all types of cells and for a multitude of diverse lesions, it has been regarded in the broader scientific community as an extraordinary scheme, of value primarily to organisms that are exposed to sunlight. Seemingly in support of this view was the remarkable discovery by Cleaver that UV-irradiated cells from several XP patients were defective in NER (*4*). Attesting to the biochemical complexity of NER is the fact that XP patients have been assigned to seven different genetic complementation groups (XP-A through XP-G, each carrying a mutation in a different gene) that are characterized by varying levels of UV sensitivity and corresponding deficiencies in repair (*5*).



Model for transcription-coupled DNA repair in mammalian cells. Arrested RNA polymerase II undergoes a conformational change to initiate transcript shortening and reannealing of DNA strands at the lesion site. The coupling factor or factors recruit TFIIH and the additional elements required for incisions, excision, and repair synthesis. After repair, transcription resumes, thereby completing the RNA transcript in progress and facilitating the passage of an advancing replication fork.

NER proceeds in stepwise fashion (*6*), beginning with recognition of the DNA lesion, followed by enzymatic incisions in the damaged strand on both sides of the lesion, removal of the damaged single-stranded segment, repair synthesis to fill in the resultant gapped DNA duplex, and ligation of the repair patch to the existing DNA strand. Several DNA-unwinding enzymes termed helicases participate in lesion recognition and in removal of the damaged segment. Interestingly, the helicases encoded by the *XPB* and *XPD* genes are components of the transcription initiation factor TFIIH, so they evidently act not only in NER but in opening up the DNA to initiate transcription (*2*). A deficiency in one of these helicases can also produce the symptoms of CS and, in some cases, TTD as well. That finding has led to the suggestion that CS and TTD might be "transcription syndromes," in which malfunctioning of the *XPB*- or *XPD*-encoded helicases in transcription disrupts early development of neuronal or ectodermal tissues. The *haywire* gene of the fruit fly *Drosophila* is a homolog of *XPB*; viable mutants of *haywire* are UV-sensitive, sterile, and display neurologic abnormalities (*7*). However, no gene-specific defects in transcription have yet been reported for XP, CS, TTD, or even *haywire*. The deficiencies in an essential transcription initiation factor would have to be very subtle indeed to preclude more serious and probably lethal developmental problems.

An alternative view is that the clinical features of CS result from the DNA repair defect that is unique in CS, namely, a deficiency in the repair of lesions in the transcribed strands of expressed genes (*1*, *8*). Several XP-G patients have CS symptoms, but the *XPG* gene product has not been implicated in transcription. Two of the complementation groups of CS (CS-A and CS-B) exhibit no defect in overall genomic NER, but they are severely deficient in transcription-coupled repair. The *CSA* and *CSB* genes are not mutated in XP, and their products have not been shown to play a direct role in transcription. How then might a defect in the repair of expressed genes result in a rather specialized set of developmental abnormalities?

One possible answer is provided by the recent report that CS-A and CS-B patients are also deficient in the transcription-coupled repair of certain lesions produced by ionizing radiation (*9*). Ionizing radiation produces lesions largely through oxidation by free radicals that attack DNA. The same

The author is in the Department of Biological Sciences, Stanford University, Stanford, CA 94305–5020, USA.

types of free radicals are generated as by-products of oxidative metabolism in nonirradiated cells. Cells with unusually high metabolic activity, such as neurons and cells that proliferate rapidly during early development, might be expected to produce more of this endogenous DNA-damaging agent, and therefore might be particularly vulnerable when transcription-coupled repair is defective. If so, the characteristic demyelination of neurons in CS could be due to excessive neuronal cell death during early development.

Many of the lesions produced by free radicals are repaired by base excision repair. In this process, which is distinct from NER, the damaged base is first cleaved from its deoxyribose moiety by a glycosylase and the DNA backbone is then incised [see (*6*)]. Transcription-coupled repair has been observed in γ-irradiated XP-A cells, with the important implication that repair pathways other than NER (such as those initiated by some glycosylases) may sometimes be coupled to transcription (*9*). The reason that XP-A patients do not also have CS may be traceable to that unique class of oxidative damage that is amenable to transcription-coupled repair in XP-A cells but not in CS-A, CS-B, XP-B/CS, XP-D/CS, or XP-G/CS cells. The accelerated neurodegeneration that accompanies the most severe XP-A cases could be the consequence of another type of free radical damage in neurons that is not repaired in XP-A, since NER has been shown to operate on at least one type of lesion induced by oxidative free radicals (*10*).

RNA polymerase II is highly processive and remains tightly bound to the DNA and to the nascent RNA transcript at the site of an obstruction such as a cyclobutane pyrimidine dimer (*11*). The transcript cannot be completed until the obstruction is removed; this could result in a selective attenuation of gene function during critical phases of early development. Mayne and Lehmann observed that although cells from CS patients appear to carry out normal levels of DNA-repair synthesis, they are severely deficient in the recovery of RNA synthesis after UV irradiation and thus could be defective in some special repair mechanism for expressed genes (*12*). These workers had made an earlier observation that UV-irradiated CS cells are defective in recovery of DNA synthesis (*13*). How could this be explained by our current understanding of the repair defect in CS?

A likely possibility is that the stalled ternary transcription complex at a cyclobutane pyrimidine dimer poses a formidable barrier to DNA replication, even though it has been reported that a DNA replication fork can pass a stalled RNA polymerase in lesion-free DNA without displacing the nascent RNA transcript (*14*). Also possible is the idea that replication cannot resume because requisite gene products cannot be made from the incomplete transcripts. The transcription-coupled repair pathway could have evolved to facilitate the expression of essential, active genes as needed to maintain viability. Another rationale for the existence of transcription-coupled repair is simply that the stalled RNA polymerase interferes with access of repair enzymes to the blocking lesion (*1, 11, 15*). Thus, some extraordinary scheme is required to enable the repair of lesions that block transcription.

Much progress has been made in understanding the mechanism of transcription-coupled repair. It was originally suggested that "the arrest of transcription at lesions and release of RNA polymerase from the template could serve as a specific signal to accelerate repair in active domains" (*16*). Selby and Sancar (*17*) isolated a transcription-repair coupling factor from *E. coli* extracts and showed that it binds to and releases the RNA polymerase blocked at a lesion. This factor may then interact with the excision-repair complex to remove the offending lesion. In human cells, the product of the *CSB* gene, ERCC6, has been implicated in the coupling process (*18*). However, the mechanism may be more complicated than that in bacteria. Mammalian genes can be much longer and are transcribed more slowly than genes in *E. coli* (transcription of the 2.5-megabase human dystrophin gene, for example, requires over 8 hours). It would seem inefficient to abort nearly completed transcripts of such genes every time RNA polymerase encounters a lesion. The transcription elongation factor SII provides an alternative scenario; this factor catalyzes nascent transcript cleavage by RNA polymerase II at natural pause sites, enabling the polymerase to "back off" and try again without aborting the incomplete transcript. A similar reaction has been demonstrated at the site of a cyclobutane pyrimidine dimer in a model DNA template in vitro (*11*), suggesting that this cleavage activity may be a key feature of transcription-coupled repair (see figure). Perhaps factor SII will prove to be the *CSA* gene product.

An important question about the mechanism of transcription-coupled repair is whether the stalling of RNA polymerase II is sufficient to initiate a repair event or whether the repair complex still has an opportunity to distinguish between a natural sequence-dependent pause site and a bona fide lesion. If the former is true, then the system may sometimes mistakenly trigger a repair reaction at a pause site. The effect of this "gratuitous transcription-coupled repair" would be the reiterative generation of repair patches at lesion-free sites, at a frequency related to the frequency of transcription through that region. This in turn could lead to higher levels of spontaneous mutagenesis in a frequently transcribed gene because of the natural error frequency of the DNA-repair polymerase. There is indeed evidence that the rate of spontaneous mutagenesis in yeast increases when the transcription rate is increased (*19*).

More surprises and complexities may be in store before we fully understand the mechanism and biological implications of transcription-coupled repair. Many of the answers will likely come from basic research on bacteria and cell-free systems for transcription and repair. Max Delbrück, one of the founders of the field of molecular biology, reflected long ago that "... any living cell carries with it the experiences of a billion years of experimentation by its ancestors. You cannot expect to explain so wise an old bird in a few simple words" (*20*).

References

1. I. Mellon, G. Spivak, P. Hanawalt, *Cell* **51**, 241 (1987); I. Mellon and P. Hanawalt, *Nature* **342**, 95 (1989); P. Hanawalt and I. Mellon, *Curr. Biol.* **3**, 67 (1993); P. Hanawalt, B. Donahue, K. Sweder, *ibid.* **4**, 518 (1994); V. Bohr *et al.*, *Cell* **40**, 359 (1985).
2. D. Bootsma and J. Hoeijmakers, *Nature* **363**, 114 (1993); R. Drapkin, A. Sancar, D. Reinberg, *Cell* **77**, 9 (1994); W. Feaver *et al.*, *ibid.* **75**, 1379 (1993); L. Schaeffer *et al.*, *Science* **260**, 58 (1993); H. Qui *et al.*, *Genes Dev.* **7**, 2161 (1993); K. Sweder and P. Hanawalt, *J. Biol. Chem.* **269**, 1852 (1994).
3. R. Setlow and W. Carrier, *Proc. Natl. Acad. Sci. U.S.A.* **51**, 226 (1964); R. Boyce and P. Howard-Flanders, *ibid.* **51**, 293 (1964); D. Pettijohn and P. Hanawalt, *J. Mol. Biol.* **9**, 395 (1964).
4. J. Cleaver, *Nature* **218**, 652 (1968).
5. ——— and K. Kraemer, in *The Metabolic Basis of Inherited Disease*, C. R. Scriver *et al.*, Eds. (McGraw-Hill, New York, ed. 6, 1989), pp. 2949–2971.
6. A. Sancar, *Science* **266**, 1954 (1994).
7. L. Mounkes *et al.*, *Cell* **71**, 925 (1992).
8. J. Venema *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4707 (1990).
9. S. Leadon and P. Cooper, *ibid.* **90**, 10499 (1993).
10. M. Satoh *et al.*, *ibid.*, p. 6335.
11. B. Donahue *et al*, *ibid.* **91**, 8502 (1994).
12. L. Mayne and A. Lehmann, *Cancer Res.* **42**, 473 (1982).
13. A. Lehmann, S. Kirk-Bell, L. Mayne, *ibid.* **39**, 4237 (1979).
14. B. Liu *et al.*, *Nature* **366**, 33 (1993).
15. C. P. Selby and A. Sancar, *J. Biol. Chem.* **265**, 21330 (1990).
16. I. Mellon *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8878 (1986).
17. C. P. Selby and A. Sancar, *Science* **260**, 53 (1993).
18. C. Troelstra *et al.*, *Cell* **71**, 939 (1992).
19. Y. Korogodin *et al.*, *Yeast* **7**, 105 (1991).
20. M. Delbrück, *Trans. Conn. Acad. Arts Sci.* **38**, 173 (1949).