Cell, Vol. 77, 9–12, April 8, 1994, Copyright © 1994 by Cell Press

Where Transcription Meets Repair

Minireview

Ronny Drapkin,* Aziz Sancar,† and Danny Reinberg*
*Department of Biochemistry
Robert Wood Johnson Medical School
University of Medicine and Dentistry of New Jersey
Piscataway, New Jersey 08854
†Department of Biochemistry and Biophysics
University of North Carolina School of Medicine
Chapel Hill, North Carolina 27599

DNA damage can be caused by a broad range of agents. Both environmental and cellular damage-inducing factors threaten cellular integrity largely by impinging upon two vital processes: replication and transcription. The interference of replication by DNA damage and the deleterious consequences are relatively well understood (for details see Sancar and Sancar, 1988). Significantly, DNA damage has an inhibitory effect on transcription (Sauerbier and Hercules, 1978), and its consequences on cellular survival are 2-fold: in stationary, quiescent cells, the lack of transcription of essential genes can lead to cell death; in proliferating cells, a stalled RNA polymerase (RNAP) complex is not only likely to interfere with transcription but also with the replication fork. Although recent studies have reported that the T4 replication fork can bypass a stalled Escherichia coli RNAP ternary complex without displacing the nascent transcript (Liu et al., 1993), it remains to be seen whether such bypass can occur at ternary complexes stalled by lesions in the DNA. It is therefore likely that targeting repair enzymes to transcriptionally active genes would confer some selective advantages. In fact, though recent studies have shown a direct coupling of DNA repair with transcription, the precedence for such a functional link has existed for some time.

The first evidence that DNA repair and transcription might be coupled cellular processes came from the observations of Mayne and Lehmann (1982) showing that transcription inhibited by ultraviolet (UV) irradiation recovered faster than could be accounted for by the rate of damage removal from the entire genome. The authors suggested that perhaps transcribed regions of the genome were repaired more rapidly. Subsequent studies by Bohr et al. (1985) demonstrated that pyrimidine dimers (Pyr↔Pyr) in the efficiently transcribed *DHFR* gene of Chinese hamster ovary (CHO) cells were repaired 5-fold faster than the rest of the genome. These observations were extended by Mellon et al. (1987) to show that the repair was preferentially limited to the transcribed (template) strand. Moreover, though the initial findings were made in mammalian cells, the same observations of gene- and strand-specific repair were later observed in E. coli (Mellon and Hanawalt, 1989). Thus, it appears that the connection between transcription and repair is universal. Although the molecular mechanism of the transcription–repair connection is not completely understood, recent insights have been provided by studies in E. coli, yeast, and mammalian cells.

DNA Repair

Three basic mechanisms exist for eliminating modified bases (DNA lesions) from DNA: direct reversal, base excision, and nucleotide excision repair (reviewed by Sancar and Sancar, 1988). The former two are specific for a narrow range of lesions. In contrast, nucleotide excision repair (excision repair) removes virtually all lesions from DNA and therefore is the most important for the maintenance of genetic integrity. Excision repair in both prokaryotes and eukaryotes is accomplished by ATP-dependent multisubunit nucleases (excinucleases). In E. coli, the UvrA, UvrB, and UvrC proteins, collectively referred to as the (A)BC excinuclease, remove DNA damage by incising the eighth phosphodiester bond 5′ and the fifth phosphodiester bond 3′ to the lesion (Sancar and Sancar, 1988). The human excinuclease activity results from the coordinated action of 8–10 proteins, identified by complementation groups of the photosensitivity disease xeroderma pigmentosum (XP), and by UV-sensitive mutant CHO cell lines (Reardon et al., 1993; see below). The repair machinery removes DNA damage by hydrolyzing the fifth phosphodiester bond 3′ and the phosphodiesters 21–23 bond 5′ to the lesion (Huang et al., 1993). In both prokaryotic and eukaryotic excinuclease systems, damage recognition is the rate-limiting step, as these enzymes must probe the DNA to detect structural abnormalites regardless of the magnitude of the damage (Sancar and Sancar, 1988).

Mechanism of Transcription–Repair Coupling in E. coli

Studies on the effect of DNA damage in E. coli led to the observation that pyrimidine dimers in the nontemplate strand had no overt effects on transcription but that dimers in the template strand completely blocked RNAP elongation and yielded stable ternary complexes (Sauerbier and Hercules, 1978; Selby and Sancar, 1990). When a purified in vitro system was used to investigate the effect of transcription on repair, an unexpected finding was obtained. Transcription specifically inhibited repair of the template strand without altering repair of the coding strand. As these findings were in sharp contrast with in vivo observations, where the template strand is preferentially repaired over the nontemplate strand, it was proposed that the in vitro system was lacking a crucial factor capable of coupling transcription and repair by overcoming the repair inhibitory effect of a stalled polymerase and promoting excision repair.

A transcription–repair coupling factor (TRCF) encoded by the *mfd* gene was identified using a cell-free transcription–repair system (Selby and Sancar, 1993) and its mechanism of action elucidated. TRCF is a protein of 130 kd with so-called helicase motifs but no helicase activity; it recognizes and binds a stalled RNAP ternary complex and the damage recognition subunit, UvrA, of the (A)BC excinuclease. It causes the release of the stalled RNAP and the truncated transcript as it escorts UvrA to the lesion site. UvrA is typically complexed with UvrB as A2:B1, and this association facilitates delivery of UvrB to the dam-

Table 1. Components of Mammalian Nucleotide Excision Repair Machinery and their Yeast Homologs

Complementation Groups of Human Mutants	Human Gene	MW (in kilodaltons)	Motif(s)	S. cerevisiae Gene	Function/Activity
XP					
XP-A	*XPA*	31 (40–45)[a]	Zinc finger	*RAD14*	DNA binding Damage recognition
XP-B	*XPB/ERCC3*	89	Nucleotide-binding fold Helicase Helix-loop-helix	*RAD25/SSL2*	Helicase/excision repair subunit of TFIIH Essential in yeast
XP-C	*XPC*	125	Hydrophilic protein	*RAD4*	
XP-D	*XPD/ERCC2*	87	Nucleotide-binding fold Helicase	*RAD3*	Helicase/excision repair subunit of factor b Essential in yeast
XP-E					
XP-F	*ERCC4*			*RAD1*[b]	DNA endonuclease[c]
XP-G	*XPG/ERCC5*	133	Helix-loop-helix	*RAD2*	Single-stranded DNA endonuclease[c]
CS					
CS-A					
CS-B	*CSB/ERCC6*	168	RNA/DNA helicase	*SNF2*[d] *RAD54* *RAD16* *MOT-1*	Preferential repair of transcribed strand Coupling factor?

[a] Native protein migrates at 40–45 kd.
[b] RAD1 exists in a complex with RAD10 and the RAD1–RAD10 complex constitutes a single-stranded DNA endonuclease.
[c] Activity has currently been detected only in the yeast homologs.
[d] Indicates only sequence homology. Functional homolog has not been isolated.

aged DNA. UvrA and TRCF then dissociate, leaving behind a stable UvrB–DNA complex. The ensuing association of UvrC with the UvrB–DNA complex induces UvrB to make the 3′ incision followed by the 5′ incision. The resulting 13-mer is released with the aid of UvrD (helicase II), and UvrB is displaced by DNA polymerase I as it fills in the gap. DNA ligase then seals the repair patch to complete the repair reaction.

The Repair Machinery in Eukaryotes

In the human system, most of what is known about DNA repair and repair-coupled transcription has come from the studies of two clinical syndromes: XP and Cockayne's syndrome (CS). XP belongs to a small family of autosomal recessive conditions, collectively referred to as chromosomal or DNA instability syndromes, that are characterized by some defect in DNA repair mechanisms and a predisposition to cancer. The fully expressed condition of XP includes a strong predisposition to sunlight-induced melanocarcinomas and basal cell and squamous cell carcinomas of the skin, in addition to other nonneoplastic cutaneous and ocular abnormalities (Cleaver and Kraemer, 1989). Cell lines from XP patients are defective in repair of pyrimidine dimers as well as other DNA lesions. Somatic cell genetics has identified seven complementation groups (XP-A to XP-G) defective in excision repair (Vermeulen et al., 1991). Analyses of these naturally occurring mutants has been extended to include a number of UV-sensitive CHO cell lines that fall into 11 complementation groups. Several of the human genes involved in excision repair have been cloned by their ability to correct the excision defect in these mutant rodent cell lines and are thus designated excision repair cross complement (*ERCC*) genes (see Table 1).

Characterization of human excision repair genes has revealed that these factors are evolutionarily conserved. The yeast Saccharomyces cerevisiae contains more than 100 genetic loci in which mutations confer abnormal sensitivity to UV and ionizing radiation. The majority of these genes have been categorized into three major repair pathways or epistasis groups: the excision repair *RAD3* group, the postreplication repair *RAD6* group, and the recombination repair group *RAD52*. Of importance here is the *RAD3* epistasis group, which consists of 12 proteins involved in excision repair. A number of these yeast excision repair components display significant homology with human excision repair factors (see Table 1), implying that excision repair is a highly conserved process.

Transcription–Repair Coupling and CS

There is no known yeast mutant with a defect in gene- and strand-specific repair. However, in humans, cell lines from patients afflicted with CS, a malady characterized by UV hypersensitivity, growth retardation, and neurological deterioration, yet with no predisposition to cancer, are unable to perform gene- and strand-specific repair (Venema et al., 1990). Cell fusion experiments have revealed the presence of at least two complementation groups (CS-A and CS-B). The gene complementing CS-B, *ERCC6*, has been cloned and sequenced (Troelstra et al., 1992). The *CSB/ERCC6* gene encodes a protein of 160 kd with helicase motifs and extensive sequence homology to the yeast SNF2 class of proteins (see Table 1). The phenotypic similarities between E. coli *mfd* mutants and CS cells and the limited homology between Mfd and CSB/ERCC6 proteins suggests that the CSB/ERCC6 protein might be the human TRCF analogous to the bacterial Mfd protein. Though attractive, this model has been complicated by

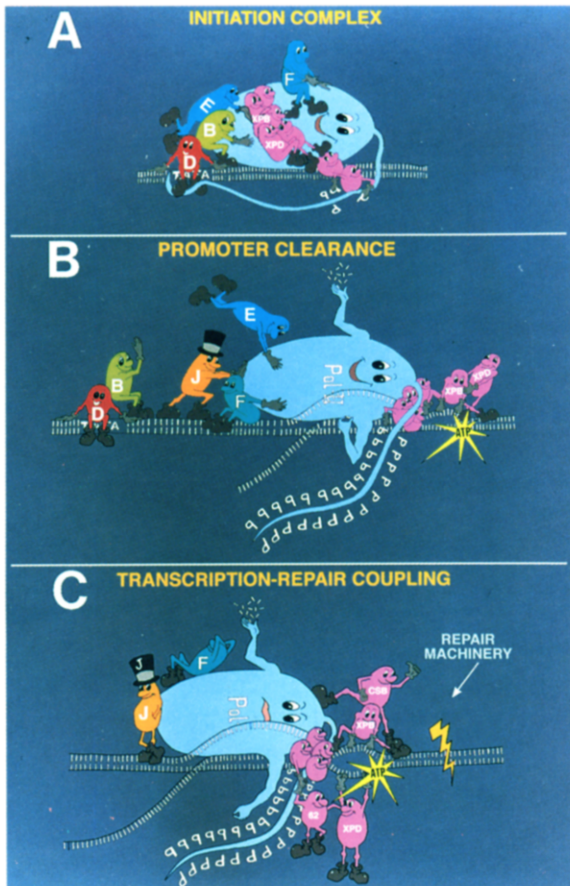

Figure 1. The Role of Repair Proteins in Transcription and Transcription-Repair Coupling

recent findings of a more intimate connection between subunits of RNAPII general transcription factor TFIIH and subunits of human excinuclease.

TFIIH/Factor b

In mammalian cells, transcription by RNAPII requires the concerted action of seven accessory proteins, referred to as general transcription factors, for the accurate initiation of transcription from class II genes. Of the seven general transcription factors, TFIIH is the most complex and the only one with numerous enzymatic activities (for details see Buratowski, 1994 [this issue of *Cell*]). Current studies indicate that TFIIH is composed of 8–10 polypeptides (Drapkin et al., 1994), its rat homolog is composed of 8 (Conaway and Conaway, 1993), and its yeast homolog, factor b, of 5 (Feaver et al., 1993). TFIIH copurifies with a DNA-dependent ATPase, a DNA helicase, and a kinase activity specific for the carboxy-terminal domain (CTD) of the largest subunit of RNAPII. Unlike eukaryotic RNAPI and RNAPIII, transcription initiation (Bunick et al., 1982) or promoter clearance (Goodrich and Tjian, 1994 [this issue of *Cell*]) by RNAPII requires the coupling of energy derived from hydrolysis of the $\beta\gamma$ bond of ATP. CTD phosphorylation cannot account for the ATP energy requirement during the initial phases of transcription (initiation or promoter clearance) as some promoters can support efficient transcription in the absence of the CTD yet still exhibit the ATP energy requirement for initation or promoter clearance and as GTP can faithfully substitute for ATP in the CTD phosphorylation reaction but does not obviate the ATP requirement. Therefore, it remains possible that other catalytic properties of TFIIH or another, currently unidentified factor account for the ATP hydrolysis requirement.

TFIIH is composed of numerous polypeptides, and several have been cloned. Schaeffer et al. (1993) characterized the largest subunit of TFIIH, p89, and discovered that it is identical to the XPB/ERCC3 protein of excision repair (Table 1). XPB/ERCC3 has helicase motifs and a DNA unwinding activity that is ATP dependent. Localized melting of DNA at the promoter by XPB/ERCC3 may facilitate open complex formation or promoter clearance and trigger RNAPII chain elongation. A helicase activity has also been observed in the TFIIH homologs in yeast (Feaver et al., 1991) and rat (Serizawa et al., 1993). The yeast homolog of XPB/ERCC3 is RAD25, also called SSL2. RAD25/SSL2 has helicase activity and is required for transcription by RNAPII as demonstrated by transcriptional shutoff in a temperature-sensitive mutant (Qui et al., 1993), lending support to the above model of XPB/ERCC3 activity.

Additional connections between TFIIH and excision repair emerged from analysis of the components of yeast factor b (p85, p75, p55, p50, p38). The 75 kd subunit of factor b, named TFB1, was cloned (Gileadi et al., 1992) and is the homolog of the p62 subunit of human TFIIH (Fischer et al., 1992). The sequence of this gene does not suggest an activity for the gene product. Feaver et al. (1993) used a histidine-tagged TFB1 to purify factor b to apparent homogeneity. The p50 subunit of factor b is encoded by the *SSL1* gene, which is known to be involved in excision repair in yeast (Yoon et al., 1992). Cloning of the largest (p85) subunit of factor b revealed it was *RAD3*, an excision repair gene, and the yeast homolog of human *XPD/ERCC2* (Feaver et al., 1993). *RAD3* is also essential for transcription by RNAPII (Guzder et al., 1994). Thus, three (p50/SSL1, p75/TFB1, and p85/RAD3) of the five subunits of factor b are involved in excision repair, and both RAD25 and RAD3 are required for RNAPII transcription.

The finding that the largest subunit of yeast factor b is *RAD3* differs from observations in the human system that indicate that the largest subunit of TFIIH is *XPB/ERCC3*, the homolog of yeast *RAD25* (Schaeffer et al., 1993). Feaver et al. (1993) suggest that *RAD25/SSL2* is not a subunit of factor b, but rather that it interacts with the factor b complex. This interpretation appears to contradict recent evidence that *RAD25* is essential for transcription by RNAPII. Whether the differences in composition are merely a consequence of the purification scheme or indicate fundamental differences between yeast and human factors remains to be seen.

Human TFIIH was purified to apparent homogeneity and tested for transcription, repair activity, and protein content. In contrast with the study in yeast, pure, transcriptionally active TFIIH contains both XPB/ERCC3 and XPD/ERCC2 proteins as integral subunits. This is in agreement with predictions made by Reardon et al. (1993) that XPB and XPD must be present in a tight complex. Antibodies directed against XPB and XPD inhibit transcription and,

moreover, highly purified TFIIH can complement cell-free extracts of XPB/ERCC3 and XPD/ERCC2 mutants in an in vitro excision assay (Drapkin et al., 1994). Thus, TFIIH can directly function in both transcription and DNA repair.

Model for Transcription-Repair Coupling in Humans

Two models can be put forth for transcription–repair coupling in humans. In the first model, XPB and XPD travel, as part of TFIIH, with the elongating polymerase, and upon encountering a lesion are delivered to the damage site by RNAPII. The RNAP backs up as if at any other pause site, and the TFIIH (XPB–XPD complex) serves as a nucleation site for the other excision repair subunits (XPA, XPC, XPE, XPF, XPG, and ERCC1). Excision and resynthesis takes place, and RNAPII, perhaps aided by TFIIS, continues its transcription from where it left off. This model has two limitations. First, it assumes that TFIIH (XPB–XPD) travels with RNAPII. Currently, there is no available evidence to suggest that this is the case. Second, the model does not assign a role to the *CSB/ERCC6* gene, whose only known biochemical mutant phenotype is the lack of transcription–repair coupling (Venema et al., 1990). With these facts in mind, we advance the following model. XPB and XPD have independent roles in transcription initiation and repair. Transcription initiation is nucleated by the ordered assembly of the general transcription factors on core promoter elements (for details see Buratowski, 1994; see Figure 1A). Once the complete initiation complex is formed, TFIIH (XPB–XPD), in the presence of each of four ribonucleosides triphosphates, locally unwinds the DNA near the transcription start site in an ATP-dependent fashion. RNAPII can then clear the promoter region as it begins synthesis of the nascent RNA chain (see Figure 1B). Elongation ceases when RNAPII encounters a DNA lesion. The stalled RNAPII has high affinity for TFIIH (XPB–XPD), and as a result this complex is recruited to the damage site; CSB/ERCC6 also has high affinity for the stalled RNAPII as well as some of the excinuclease subunits involved in damage recognition. The CSB protein displaces the RNAPII as it delivers one or more of the repair proteins to the lesion site, promoting formation of a stable DNA–protein complex in an ATP-dependent reaction (see Figure 1C). CSB subsequently dissociates from the complex, conceivably together with some of the XP proteins involved in the assembly of the excinuclease but not in the actual incision. Following dissociation of CSB/ERCC6, excision takes place.

If transcription-directed repair involves a coupling factor as we have described, theoretically three types of mutations should cause uncoupling of transcription from repair, resulting in a Mfd$^-$ phenotype in E. coli and CS in humans. These include mutations in the coupling factor itself (Mfd in E. coli and CSB in humans), mutations in RNAP, and mutations in the subunits that are recruited to the stalled complex by the coupling factor or by their association with transcription factors (XPB–XPD). It is noteworthy that some patients in the XP-B, XP-D, and XP-G complementation groups suffer from both XP and CS. It is conceivable that one or more of these subunits interact with CSB/ERCC6 and that in the mixed phenotype syndromes, the mutations are at the interface of these proteins with the CSB coupling factor. As *CSB/ERCC6* only complements one of the two CS complementation groups, it is possible that two factors may be required for coupling repair to transcription. Present studies in yeast and human systems will provide additional clues in our attempt to decipher the mechanism of transcription-coupled DNA repair.

References

Bohr, V. A., Smith, C. A., Okumoto, D. S., and Hanawalt, P. C. (1985). Cell *40*, 359–369.

Bunick, D., Zandomeni, R., Ackerman, S., and Weinmann, R. (1982). Cell *29*, 877–886.

Buratowski, S. (1994). Cell *77*, this issue.

Cleaver, J. E., and Kraemer, K. H. (1989). In The Metabolic Basis of Inherited Disease, Volume 2. C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle (New York: Mcgraw Hill), pp. 2949–2971.

Conaway, R. C., and Conaway, J. W. (1993). Annu. Rev. Biochem. *62*, 161–170.

Drapkin, R., Reardon, J., Ansari, A., Huang, J.-C., Zawel, L., Ahn, K., Sancar, A., and Reinberg, D. (1994). Nature, in press.

Feaver, W., Gileadi, O., Li, Y., and Kornberg, R. D. (1991). Cell *67*, 1223–1230.

Feaver, W., Svejstrup, J. Q., Bardwell, L., Bardwell, A. J., Buratowski, S., Gulyas, K. D., Donahue, T. F., Friedberg, E. C., and Kornberg, R. D. (1993). Cell *75*, 1379–1387.

Fischer, L., Gerard, M., Chalut, C., Lutz, Y., Humbert, S., Kanno, M., Chambon, P., and Egly, J.-M. (1992). Science *257*, 1392–1395.

Gileadi, O., Feaver, W. J., and Kornberg, R. D. (1992). Science *257*, 1389–1392.

Goodrich, J. A., and Tjian, R. (1994). Cell *77*, this issue.

Guzder, S. N., Qui, H., Sommers, C. H., Sung, P., Prakash, L., and Prakash, S. (1994). Nature *367*, 91–94.

Huang, J. C., Svoboda, D. L., Reardon, J. T., and Sancar, A. (1992). Proc. Natl. Acad. Sci. USA *89*, 3664–3668.

Liu, B., Wong, M. L., Tinker, R. L., Geiduschek, E. P., and Alberts, B. M. (1993). Nature *366*, 33–39.

Mayne, L. V., and Lehmann, A. R. (1982). Cancer Res. *42*, 1473–1478.

Mellon, I., and Hanawalt, P. C. (1989). Nature *342*, 95–98.

Mellon, I., Spivak, G., and Hanawalt, P. C. (1987). Cell *51*, 241–249.

Qui, H., Park, E., Prakash, L., and Prakash, S. (1993). Genes Dev. *7*, 2161–2171.

Reardon, J. T., Thompson, L. H., and Sancar, A. (1993) Cold Spring Harbor Symp. Quant. Biol. *58*, 605–617.

Sancar, A., and Sancar, G. B. (1988) Annu. Rev. Biochem. *57*, 29–67.

Sauerbier, W., and Hercules, K. (1978). Annu. Rev. Genet. *12*, 329–363.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulin, W., Hoeijmakers, J. H. J., Chambon, P., and Egly, J.-M. (1993). Science *260*, 58–63.

Selby, C. P., and Sancar, A. (1990). J. Biol. Chem. *265*, 21330–21336.

Selby, C. P., and Sancar, A. (1993). Science *260*, 53–58.

Serizawa, H., Conaway, R. C., and Conaway, J. W. (1993). J. Biol. Chem. *268*, 17300–17308.

Troelstra, C., van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D., and Hoeijmakers, J. H. J. (1992). Cell *71*, 939–953.

Venema, J., Mullenders, L. H. F., Natarajan, A. T., van Zeeland, A. A., and Mayne, L. V. (1990). Proc. Natl. Acad. Sci USA *87*, 4707–4711.

Vermeulen, W., Stefanini, M., Giliani, S., Hoeijmakers, J. H. J., and Bootsma, D. (1991). Mutat. Res. *255*, 201–208.

Yoon, H., Miller, S. P., Pabich, E. K., and Donahue, T. F. (1992). Genes Dev. *6*, 2463–2477.