*Proc. Natl. Acad. Sci. USA*
Vol. 88, pp. 8232–8236, September 1991
Biochemistry

# Gene- and strand-specific repair *in vitro*: Partial purification of a transcription-repair coupling factor

(*Escherichia coli*/UV/psoralen/cisplatin/repair synthesis)

CHRISTOPHER P. SELBY AND AZIZ SANCAR

Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, NC 27599

*Communicated by Mary Ellen Jones, June 17, 1991 (received for review May 22, 1991)*

ABSTRACT **In eukaryotic and prokaryotic cells, actively transcribed genes and, in some instances, the template strand of these genes have been found to be repaired 2–10 times more rapidly than nontranscribed genes or the coding strand of transcribed genes. We demonstrate here gene- and template strand-specific repair synthesis *in vitro* by using an *Escherichia coli* cell-free extract and a plasmid carrying a gene with the strong *tac* promoter. Strand-specific repair of UV, 4′-hydroxymethyl-4,5′,8-trimethylpsoralen, and *cis*-dicholorodiammine platinum(II) damage was dependent upon transcription and a functional nucleotide excision repair system and was stimulated by 6% (wt/vol) polyethylene glycol. A defined system consisting of the transcription and repair proteins in highly purified form did not perform strand-specific repair; however, active fractions of extract conferred strand specificity to the defined system. Transcription–repair coupling activity was partially purified from extract by successive DEAE-agarose and gel filtration chromatography. The coupling factor is heat-labile, with an estimated $M_r$ of 100,000.**

DNA repair enzymes in general and nucleotide-excision nuclease subunits in particular are not abundant (1, 2), and it would be advantageous if the cell's limited resources were channeled to the most threatening genetic lesions. In this regard, while replication-blocking lesions could be lethal in any location on the chromosome, lesions in genes pose the additional threats of causing deleterious mutations or blocking transcription. It now appears that eukaryotes and prokaryotes use specific targeting mechanisms for some types of lesions to direct nucleotide-excision repair to regions most crucial for survival.

The first and perhaps most influential report on the subject by Bohr *et al.* (3) showed that pyrimidine dimers (Pyr◇Pyr) in the dihydrofolate reductase (DHFR) gene of Chinese Hamster ovary (CHO) cells were repaired about five times more rapidly than those in flanking nontranscribed regions or in the genome overall. Consequently it was found that in both CHO and human cells, the template strand of the actively transcribed *DHFR* gene was repaired 2- to 10-fold more efficiently than the complementary strand, which was repaired at a rate comparable to that of nontranscribed regions (4). These observations, originally made in mammalian cells, have now been extended to other organisms including yeast (5, 6) and *Escherichia coli* (7).

An *in vitro* study on this subject attempted to examine the mechanism behind gene-specific repair using purified *E. coli* (A)BC excinuclease, RNA polymerase (RNA Pol), DNA polymerase I (pol I), DNA ligase, and a synthetic DNA or a plasmid with strong promoters as transcription–repair template–substrates. [(A)BC excinuclease comprises the products of *E. coli* genes *uvrA*/*uvrB*/*uvrC*. UvrA recognizes the damaged DNA but does not take part in the DNA incision, which is carried out by UvrB–UvrC.] This study yielded a paradoxical result: RNA Pol stalled at a thymine dimer (T◇T) in the template strand prevented access of (A)BC excinuclease (which incises the eighth phosphodiester bond 5′ and the fifth phosphodiester bond 3′ to the photodimer) to the lesion. Thus, a T◇T in the complementary (coding) strand (which did not block transcription) was repaired 2- to 4-fold more efficiently than a T◇T in the template strand (8). This result suggests that "transcription–repair coupling" (9) must accomplish two tasks: direct the repair enzyme (assembly) to the stalled complex and overcome the steric hindrance caused by the stalled RNA Pol. In this paper we show that a more complete but less-defined *in vitro* system utilizing *E. coli* cell-free extract carries out gene- and strand-specific repair. A coupling factor that conferred strand-specific repair function when added to a defined transcription–repair system was partially purified.

## MATERIALS AND METHODS

**Materials.** Extracts were made from *E. coli* strains AB1157 (wild type) and AB1886 (*uvrA*$^-$) by procedures of Wickner *et al.* (10) as modified by Lu *et al.* (11). One milliliter of culture yielded ≈1 ml of extract with 30–40 mg of protein per ml.

**Partial Purification of the "Coupling Factor."** Extract was fractionated on DEAE-agarose (Bio-Rad) and AcA 34 columns (IBF). About 2 ml was loaded onto an 8-ml DEAE-agarose column and eluted with 2–3 column volumes of 100 mM KCl in buffer B (50 mM Tris·HCl, pH 7.5/1.0 mM EDTA/20% glycerol/10 mM 2-mercaptoethanol) at about 0.1 ml/min. The column-bound fraction was eluted with 500 mM KCl in buffer B. The peak low-salt flow-through fractions were pooled, and 1–2 ml were loaded onto an AcA 34 column (1 cm × 34 cm) equilibrated with 300 mM KCl in buffer B; the column was developed with the same buffer at a rate of 0.1 ml/min. Fractions of 0.6 ml were collected and concentrated to 0.1–0.2 ml by using Amicon centricon microconcentrators.

**Substrates.** Supercoiled pDR3274 (12) was purified by CsCl density gradient centrifugation. Plasmid DNA (30 μg/ml in 10 mM Tris·HCl, pH 7.4/10 mM NaCl/1 mM EDTA) was damaged with (*i*) 225 J·m$^{-2}$ of 254-nm light from a Sylvania germicidal lamp, (*ii*) 3.3 μM 4′-hydroxymethyl-4,5′,8-trimethylpsoralen (psoralen) (HRI Associates, Emeryville, CA) with 6.7 × 10$^8$ J·m$^{-2}$ of 365-nm light from a Spectroline B-100 UV lamp, or (*iii*) 100 μM *cis*-dichlorodiammine platinum(II) (cisplatin) (Sigma) for 2.5 hr. These doses were optimized to give a strong and reproducible repair-synthesis signal without introducing measurable levels of crosslinking into the fragment of interest.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Pyr◇Pyr and T◇T, pyrimidine and thymine dimer(s), respectively; IPTG, isopropyl β-D-galactoside; RNA Pol, RNA polymerase; pol I, DNA polymerase I; psoralen, 4′-hydroxymethyl-4,5′,8-trimethylpsoralen; cisplatin, *cis*-dicholorodiammine platinum(II).

**Repair Synthesis with Extracts.** DNA damage-dependent repair synthesis was measured by the incorporation of radiolabeled dCMP into repair patches. Transcription–repair reaction mixtures included 40 mM Hepes (pH 7.8); 50 mM KCl; 8 mM $MgCl_2$; 4% (vol/vol) glycerol; 5 mM dithiothreitol; 2 mM ATP; 100 μg of bovine serum albumin per ml; 6% (wt/vol) polyethylene glycol 6000; 500 μM NAD; 200 μM each CTP, GTP, and UTP; 40 mM each dATP, dGTP, and dTTP; 4 μM unlabeled dCTP; 2.5–10 μCi (1 μCi = 37 kBq) of [α-$^{32}$P]dCTP (6000 Ci/mmol, New England Nuclear); and 1.3 nM DNA (expressed as moles of plasmid). Crude extracts were included at 1.2 mg of protein per ml. When included, we used 0.8–1 mM isopropyl β-D-galactoside (IPTG), 15–20 nM *lac* repressor, and 22 μg of rifampicin (Sigma) per ml. Reactions were at 37°C for 25 min, and were stopped by phenol addition. DNA was extracted twice with phenol, twice with ether, precipitated with ethanol, and resuspended in restriction enzyme reaction buffer (GIBCO/BRL).

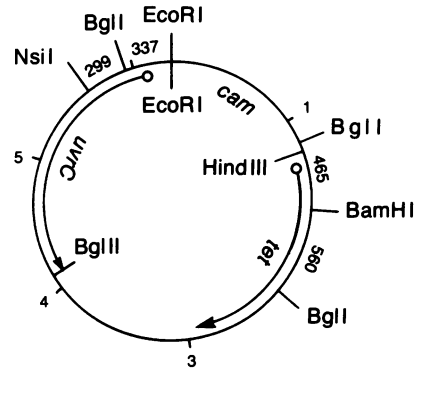
To examine gene-specific repair, the repaired DNA (Fig. 1) was digested with *Eco*RI, *Bgl* II, and *Hin*dIII, and separated by electrophoresis on a 1.2% agarose gel. Autoradiographs were made and radiolabel incorporation was quantitated by excising DNA bands and scintillation counting.

To examine strand-specific repair, the repaired plasmid was digested with *Nsi* I, *Bgl* I, *Eco*RI, and *Bam*HI. Products were separated by electrophoresis on a 3.6% sequencing gel. Autoradiographs of dried gels were made, and radiolabel incorporation in individual strands was quantitated by using a Biomed softlaser scanning densitometer. The template and coding strands of the *uvrC* gene fragments were identified based upon the *uvrC* sequence (13).

**Repair Synthesis with the Defined System.** Chromatographic fractions of UvrA$^-$ cell extract were tested for transcription–repair coupling activity upon addition to a reconstituted transcription–repair system. This system consisted of 1.2 units of RNA Pol (Promega) per ml, 4 nM UvrA, 100 nM UvrB, 70 nM UvrC (purified as in ref. 12), 5 nM helicase II, and 80 units of pol I (BRL) and 48 units of phage T4 DNA ligase (Promega) per ml. Reaction conditions were as in experiments with crude extracts using pDR3274 as the substrate.

## RESULTS

***In Vitro* System.** The system consisted of *E. coli* cell-free extract and a plasmid with a strong *tac* promoter. This extract can replicate single-stranded DNA phages (10) and *E. coli* minichromosomes (14) and can repair mismatched bases (11). Our preparation carried out transcription and nucleotide-excision repair efficiently, and in preliminary experiments we detected a weak transcription-dependent repair signal (data not shown). We attempted to improve the system by including polyethylene glycol, a hydrophilic polymer that increases the effective concentration of proteins by "molecular crowding" and was essential for replicating an *E. coli* minichromosome (14). PEG at 6% more than doubled the transcription-dependent repair signal, and was included in all reactions.




pDR3274

FIG. 1. Transcription–repair substrate. A partial restriction map of pDR3274 (6.24 kbp) is shown. Additional *Bgl* I sites exist but are omitted for clarity. The approximate locations of the *tet* and *uvrC* genes and the promotorless *cam* gene are indicated; promoters are indicated by circles, genes by lines, and the direction of transcription by arrows. Gene-specific repair was examined after digestion with *Eco*RI, *Bgl* II, and *Hin*dIII. Strand-specific repair was examined after digestion with *Nsi* I, *Bgl* I, *Eco*RI, and *Bam*HI, which generated fragments of 299, 337, 465, and 560 bp.

**Gene-Specific Repair.** Unirradiated or irradiated pDR3274 was incubated with crude extract in transcription–repair buffer with [α-$^{32}$P]dCTP to measure repair synthesis. Reactions were carried out under conditions optimal for transcription (identified as "+ transcription"); in the presence of rifampicin, which inhibits RNA Pol (identified as "− transcription"); and in the presence of the *lac* repressor, which binds to the *tac* promoter and specifically inhibits transcription of *uvrC*. After repair, the plasmid was digested with restriction enzymes to generate three fragments, which carry the *tet*, *uvrC*, and *cam* genes (Fig. 1), which were separated on an agarose gel.

Fig. 2 *Upper* shows an autoradiograph of such a gel and Fig. 2 *Lower* and Table 1 give a quantitative analysis of data from two experiments such as that in Fig. 2 *Upper*. Lane 1 is a measure of nonspecific incorporation in the absence of DNA damage; lane 2 represents the basal level of UV damage-dependent repair synthesis in the absence of transcription and corresponds to 0.3 repair patch per molecule. Repair synthesis in lane 3 was conducted under optimal transcription conditions; enhancement of repair in lane 3 versus lane 2 may be ascribed to an effect of transcription. After background levels of repair synthesis observed with the UvrA$^-$ extract (not shown) were subtracted, the ratios of damage-induced repair synthesis (lane 3 vs. lane 2) were 1.6 for *uvrC*, 1.1 for *cam*, and 0.9 for *tet* (Fig. 2 *Lower*).

That the enhanced repair in *uvrC* was due to transcription of the gene and not some secondary effects of RNA Pol, such as initiating replication from the origin of replication, was demonstrated by using *lac* repressor. Addition of *lac* repressor abolished the repair-enhancing effect of transcription in the *uvrC* fragment (Fig. 2 *Upper*, lane 4). The level of *tac–uvrC* repair in the presence of the repressor was comparable to that with rifampicin (Fig. 2 *Upper*, lane 4 vs. lane 2), while the repressor did not inhibit repair of the other two fragments (in fact, it mildly stimulated their repair by some unknown mechanism). Addition of the gratuitous *lac* inducer IPTG negated the effect of the *lac* repressor and brought the ratio of repair synthesis in transcribed vs. nontranscribed *uvrC* back up to 1.8 (Fig. 2 *Upper*, lane 5 vs. lane 2). Thus, transcription of *uvrC* specifically stimulates the repair synthesis of this gene.

**Strand-Specific Repair.** In most cases gene-specific repair observed *in vivo* appears to be due to an increased repair rate of the template strand in the actively transcribing gene (4, 7). Therefore, we developed an approach to examine the strand specificity of repair catalyzed *in vitro*. After transcription–repair synthesis reactions, the plasmid (Fig. 1) was digested with a set of restriction enzymes that was selected to produce two fragments (299 and 337 bp) from the very strongly transcribed *uvrC* gene, and two fragments (465 and 560 bp) from the weakly transcribed *tet* gene. Also, each fragment possessed heterogeneous 3′ and 5′ overhanging ends. Consequently, the template (t) and coding (c) strands of each

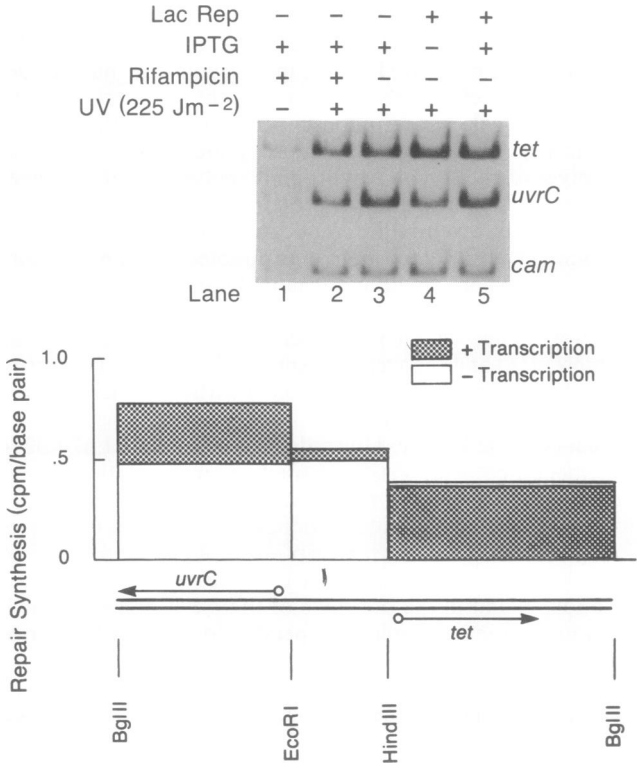



FIG. 2. Gene-specific nucleotide-excision repair. (*Upper*) Repair synthesis was performed with extracts from *E. coli* AB1157 (uvrA$^+$) in the presence or absence of UV photodamage to the DNA, rifampicin, IPTG, and *lac* repressor (Lac Rep) as indicated. Plasmids were then digested with *Eco*RI, *Bgl* II, and *Hin*dIII (see Fig. 1), and the three fragments were resolved on a 1.2% agarose gel. The fragments carry the strongly transcribed *uvrC* gene (2.1 kbp), the weakly transcribed *tet* gene (2.9 kbp), and the promotorless *cam* gene (1.2 kbp). (*Lower*) DNA bands from two separate experiments were excised, and repair synthesis was quantitated by scintillation counting. Background cpm values obtained with *E. coli* AB1886 (uvrA$^-$) extract for *tet*, *uvrC* and *cam*, respectively, were 1161, 907, and 542 (– transcription) and 1506, 1032, and 664 (+ transcription). These values were subtracted from the corresponding values obtained with AB1157 (*uvrA*$^+$) extract to obtain values that reflect nucleotide-excision repair synthesis. These values were divided by the fragment length in bp to give values for nucleotide-excision repair per unit length, which are plotted. The two dark lines below the *x* axis represent pDR3274 linearized at the *Bgl* II site; the adjacent circles, lines, and arrows represent the approximate sites of promoters, genes, and directions of transcription, respectively, for *uvrC* and *tet*. The promotorless *cam* gene is located between the *Eco*RI and *Hin*dIII sites. Restriction sites are indicated and demarcate the units of the abscissa against which repair synthesis values are plotted. Open bars represent repair synthesis in the presence of rifampicin (– transcription), and cross-hatched bars represent repair synthesis in the absence of rifampicin (+ transcription).

fragment varied in length and were separable on sequencing gels (Fig. 3 *Left*).

Extract from UvrA$^-$ cells gave only a weak UV-dependent signal (Fig. 3 *Left*, lanes 1–3), perhaps from repair of thymine glycols and pyrimidine hydrates by endonuclease III (2). Transcription did not appear to have any specific effect on this mode of repair. In contrast, wild-type extract gave a strong UV-dependent signal (cf. lanes 4 and 5): half of the signal was caused by Pyr◇Pyr and the other half by 6–4 pyrimidine–pyrimidone photoproducts (17). More importantly, there was a large effect of transcription on repair (cf. lanes 5 and 6). For the 299-bp fragment originating from the 5′-terminal half of *uvrC*, the ratio of repair synthesis under the transcription condition vs. nontranscription was 4.4:1 for the template strand and 1.2:1 for the complementary strand

Table 1. Effect of rifampicin and *lac* repressor on transcription-stimulated gene- and strand-specific repair in *tac–uvrC*

| Gene or fragment | Relative repair synthesis | | | |
|---|---|---|---|---|
| | + Rif | – Rif | + *lac* rep | + *lac* rep/IPTG |
| uvrC-2100* | 100 | 163 | 92 | 184 |
| uvrC-344c† | 100 | 107 | 141 | 121 |
| -337t | 114 | 403 | 186 | 351 |
| uvrC-300c‡ | 100 | 120 | 96 | 121 |
| -299t | 114 | 505 | 168 | 457 |

Quantitation of gene-specific repair was by scintillation counting of radioactivity in a fragment carrying *uvrC*; in the case of strand-specific repair, quantitation was by densitometric scanning of autoradiograms of sequencing gels. Rif, rifampicin; *lac* rep, *lac* repressor.

*Data are for the *Bgl* II–*Eco*RI fragment (2100 bp), which carries nearly the entire *uvrC* coding region. The radioactivity incorporated is expressed relative to that obtained in the absence of transcription (+Rif) after subtracting background levels obtained with the *uvrA*$^-$ strain.

†These represent the 337-base template (t) and the 344-base coding (c) strands generated from the 5′ terminus (including the promoter) of *tac-uvrC*. Repair synthesis is expressed relative to that of the complementary (coding) strand under nontranscription conditions after subtracting background synthesis.

‡These are the two strands of the fragment generated by *Bgl* I and *Nsi* I enzymes. The fragment is adjacent to and downstream from the 337-bp fragment carrying the promoter and 5′ terminus of *tac-uvrC*. Repair synthesis is expressed relative to that of the complementary (coding) strand under nontranscription conditions after subtracting background synthesis.

(quantitative values are given in Fig. 3 *Right* and Table 1). The same trend held for the other transcribed fragment from *uvrC* (337 bp). However, only a marginal change was observed in the fragments from within the weakly transcribed *tet* gene, and this change was not specific for either strand (Fig. 3 *Right*). The strand-specific repair observed in *uvrC* transcribed from the *tac* promoter was specifically inhibited by the *lac* repressor (Fig. 3 *Left*, lane 7; Table 1), and it could be restored by including IPTG in the reaction mixture (lane 8). All of these data combined suggests that our *in vitro* system reflects the *in vivo* condition faithfully with regard to gene- and strand-specific repair.

**Strand-Specific Repair of Psoralen and Cisplatin Damage.** To find out if other adducts are repaired in a strand-specific manner, we conducted repair synthesis experiments with psoralen- or cisplatin-damaged DNAs and analyzed the data in terms of strand specificity as we did for UV-damaged DNA. The results are summarized in Table 2. Clearly, these two adducts also were preferentially removed from the template strand during transcription. Thus, this phenomenon may be universal for the many different substrates of (A)BC excinuclease.

**Partial Purification of the "Coupling Factor."** Addition of purified UvrA protein to extract made from UvrA$^-$ cells restored its strand-specific repair activity (data not shown). This suggested that strand-specific repair is not due to a preexistent supramolecular assembly of repair and transcription proteins and raised the possibility that a coupling protein enhances the repair of lesions where a stable elongation complex has formed. We reasoned that if such a factor existed, then upon addition to a defined nucleotide-excision repair–transcription system, in which the transcribed strand is not repaired preferentially (8), the coupling factor should restore strand selectivity.

Toward this goal we fractionated extract from UvrA$^-$ cells on the DEAE anion-exchange column. The fraction eluted with a low-salt eluant was incapable of repair synthesis even upon addition of UvrA protein. However, this fraction (but not the high-salt eluate), when added to a defined transcrip-

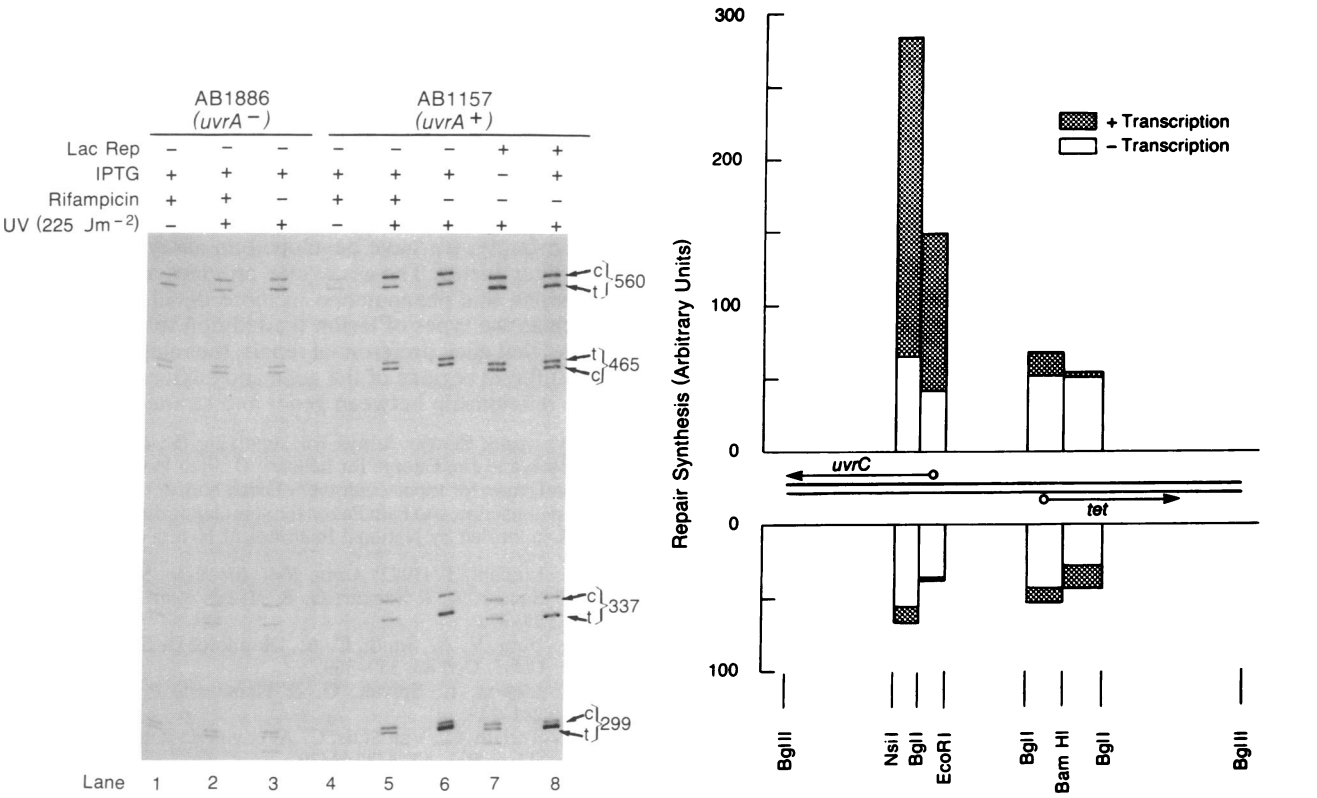



FIG. 3. Strand-specific nucleotide excision repair. (*Left*) Repair synthesis was performed with extracts from *E. coli* AB1886 (uvrA$^-$) and AB1157 (uvrA$^+$) and in the presence and absence of UV photodamage to the DNA, rifampicin, IPTG, and *lac* repressor (Lac Rep) as indicated. Plasmids were then digested with *Nsi* I, *Bgl* I, *Eco*RI, and *Bam*HI to generate two fragments from *uvrC* (299 and 337 bp) and two fragments (465 and 560 bp) from the weakly transcribed *tet* gene (Fig. 1). Fragments were resolved on a 3.6% sequencing gel; the template (t) and coding (c) strands of each fragment are indicated. The extra bands seen in lanes 3, 6, 7, and 8 resulted from replication that took place in the absence of rifampicin. (*Right*) Intensities of the DNA bands were quantified by scanning densitometry, and values for nucleotide-excision repair per unit length of DNA (obtained as described in the legend to Fig. 2) are plotted. The *uvrC* and *tet* genes and pDR3274 linearized at the *Bgl* II site are represented as described for Fig. 2; the arrows indicate the directions of template DNA strands. *Right Upper* shows repair in the template strand of *uvrC* and the coding strand of *tet*. *Right Lower* shows repair in the coding strand of *uvrC* and the template strand of *tet*. Restriction sites are indicated and demarcate the units of the abscissa against which repair synthesis values are plotted; (−) transcription values are those obtained in the presence of rifampicin.

tion–repair system containing UvrA, UvrB, UvrC, and UvrD proteins; pol I; DNA ligase, RNA Pol, dNTPs, and rNTPs, enhanced the repair of the template strand (Fig. 4 *Upper*, lanes 1 and 2). We reasoned that the low-salt eluate contained the coupling factor and proceeded to purify it further. The active fractions from the DEAE column were loaded onto an AcA 34 gel permeation column, and the fractions were tested for coupling activity in the defined system. The result is shown in Fig. 4 *Lower* (see also lane 3 in Fig. 4 *Upper*). Activity was eluted in a wide peak over the range of 40–120 kDa with a midpoint at 80 kDa. The coupling activity at this stage of purification was totally lost after heating for 5 min at 65°C (not shown), suggesting that the coupling factor is a protein of $M_r \approx 100{,}000$.

Table 2. Strand-specific repair of UV, cisplatin, and psoralen lesions

| | Ratio of template/coding strand repair | | |
|---|---|---|---|
| Rifampicin | UV | Cisplatin | Psoralen |
| With | 0.85 | 1.25 | 0.98 |
| Without | 3.32 | 3.34 | 2.38 |

pDR3274 was treated with UV (254 nm), cisplatin, or psoralen with UV (365 nm), and repair by AB1157 extract was assayed as in Fig. 3 *Left*. Repair synthesis in the template and coding strands of the 299-bp *Bgl* I–*Nsi* I fragment were quantitated by scanning densitometry. The numbers are averages from two experiments. The presence of rifampicin inhibits transcription.

## DISCUSSION

Both prokaryotes (18) and eukaryotes (19) can survive and replicate with a large number of Pyr◇Pyr in their genomes. In contrast, a single Pyr◇Pyr in the template strand is an absolute block for transcription (8, 20, 21); therefore, such a lesion located in an essential gene could be lethal. The coupling of transcription to repair helps to prevent such a catastrophe. Such coupling, although not universal for all adducts (22) or all genes (23), has been documented *in vivo* in several systems (4–7, 23), and it also has been inferred from studies showing selective mutagenesis (24, 25). Models describing this coupling may be divided into two general and not necessarily exclusive categories: coupling by protein–protein interaction or induction of preferential repair by structurally unique features of the template strand during transcription.

In one structural model, the open chromatin conformation of transcribed DNA is more "accessible" to repair, and RNA Pol stalled at photoproducts in the template strand (but not the coding strand) retains the open conformation (6). These factors may contribute to strand specificity; however, RNA Pol stalled at a T◇T in the template inhibits repair of the T◇T *in vitro* (8). A second structural model is the targeting of the repair enzyme to the template strand by the RNA·DNA hybrid. An RNA·DNA hybrid is a very poor substrate for

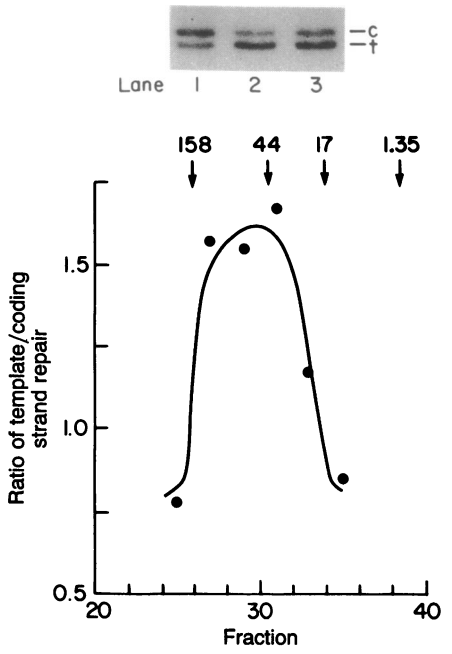



FIG. 4. Partial purification of a coupling factor that enhances repair of template DNA when added to a defined transcription–repair system. The defined system consisted of pDR3274 irradiated with 225 $J \cdot m^{-2}$ of UV as the template–substrate for purified RNA Pol, UvrA, UvrB, UvrC, UvrD, pol I, and DNA ligase. After transcription–repair synthesis, the plasmid was processed as in Fig. 3 *Left*, and the ratio of repair in the template (t) and coding (c) strands of the 299-bp *Bgl* I–*Nsi* I fragment from the *uvrC* gene (Fig. 1) was measured by densitometry of autoradiographs. (*Upper*) Effect of the coupling factor on the defined system. Lanes: 1, repair synthesis with the defined system alone; 2, defined system plus DEAE active fraction; 3, defined system plus AcA34 fraction 29. The t/c intensity ratios were 0.8, 2.1, and 1.6, respectively, for the three lanes. (*Lower*) Elution profile of the coupling factor on an AcA34 column (1 cm × 34 cm). Active DEAE fraction (2 ml) was loaded onto the column; 0.6-ml fractions were collected and tested for activity by adding to the defined transcription–repair system. Arrows indicate the positions of the size markers: from left to right, gamma globulin, ovalbumin, myoglobin, and vitamin $B_{12}$.

(A)BC excinuclease (A. M. Phillips and A.S., unpublished data), so this model can be eliminated. In a third structural model, the major, asymmetric conformational changes occurring during transcription, such as local unwinding, positive and negative supercoiling preceding and following the transcription complex (26, 27), and topoisomerases that modulate the topology of DNA, could be contributing factors for strand specificity. However, topological changes do occur in the defined system in which repair is not strand-specific (Fig. 4 and ref. 8). Therefore, we favor direct coupling by protein–protein interactions.

Proteins that may be considered candidate coupling factors because of their known functions in transcription and repair include RNA Pol itself, photolyase [which stimulates (A)BC excinuclease action on T◇T in the absence of transcription] and its potential mammalian analog (28), Rho protein (which releases a stalled RNA Pol, thus removing the steric hindrance), helicase II (product of *uvrD* that is essential for repair synthesis), and topoisomerase I (which affects the topology of transcribed DNA). A previous study ruled out RNA Pol and *E. coli* DNA photolyase as coupling factors (8), and this study either directly (helicase II) or in preliminary experiments with the defined system (not shown) rules out Rho protein and topoisomerase I. An intriguing candidate for coupling factor in human cells is the protein defective in Cockayne syndrome and its possible *E. coli* analog. Cockayne syndrome cells are unable to carry out gene-specific repair but are normal in total DNA repair (29–31).

The experiments reported here demonstrate transcription-driven preferential repair of template DNA *in vitro* and give direct evidence for a protein that is distinct from the known transcription and repair proteins and functions to couple the two processes. We have eliminated some possible mechanisms and proteins as candidate coupling factors. More importantly, we have developed an assay for purifying the coupling factor. The assay also provides a simple system to examine this phenomenon in more detail regarding factors such as the types of lesion (in addition to the ones reported here) that elicit preferential repair, the relative rates of repair in different regions of the gene and adjacent sequences, and the relationship between gene- and strand-specific repair.

We thank Sankar Adhya for supplying the *lac* repressor, Steve Matson and Jim George for helicase II, Pam Pavco for Rho protein, John Hearst for topoisomerase I, Gwen Sancar for critical reading of the manuscript, and Pam Pavco for stimulating discussions. This work was supported by National Institutes of Health Grant GM32833.

1. Lindahl, T. (1982) *Annu. Rev. Biochem.* **51**, 61–87.
2. Sancar, A. & Sancar, G. B. (1988) *Annu. Rev. Biochem.* **57**, 29–67.
3. Bohr, V. A., Smith, C. A., Okumoto, D. S. & Hanawalt, P. C. (1985) *Cell* **40**, 359–369.
4. Mellon, I., Spivak, G. & Hanawalt, P. C. (1987) *Cell* **51**, 241–249.
5. Terleth, C., Van Sluis, C. A. & van de Putte, P. (1989) *Nucleic Acids Res.* **17**, 4433–4439.
6. Smerdon, M. J. & Thoma, F. (1990) *Cell* **61**, 675–684.
7. Mellon, I. & Hanawalt, P. C. (1989) *Nature (London)* **342**, 95–98.
8. Selby, C. P. & Sancar, A. (1990) *J. Biol. Chem.* **265**, 21330–21336.
9. Hanawalt, P. C. (1989) *Environ. Mol. Mutagen.* **14**, Suppl. 16, 90–98.
10. Wickner, W., Brutlag, D., Schekman, R. & Kornberg, A. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 965–969.
11. Lu, A.-L., Clark, S. & Modrich, P. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4639–4643.
12. Thomas, D. C., Levy, M. & Sancar, A. (1985) *J. Biol. Chem.* **260**, 9875–9883.
13. Sancar, G. B., Sancar, A. & Rupp, W. D. (1984) *Nucleic Acids Res.* **12**, 4593–4608.
14. Fuller, R. S., Kaguni, J. M. & Kornberg, A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7370–7374.
15. Sancar, A. & Rupp, W. D. (1983) *Cell* **33**, 249–260.
16. Amann, E., Brosius, J. & Ptashne, M. (1983) *Gene* **25**, 167–178.
17. Sibghat-Ullah & Sancar, A. (1990) *Biochemistry* **29**, 5711–5718.
18. Rupp, W. D., Wilde, C. E., Reno, D. L. & Howard-Flanders, P. (1971) *J. Mol. Biol.* **61**, 25–44.
19. Meneghini, R., Menck, C. F. M. & Schumacher, R. I. (1981) *Q. Rev. Biophys.* **14**, 381–432.
20. Michalke, H. & Bremer, H. (1969) *J. Mol. Biol.* **41**, 1–23.
21. Sauerbier, W. & Hercules, K. (1978) *Annu. Rev. Genet.* **12**, 329–363.
22. Tang, M. S., Bohr, V. A., Zhang, X.-S., Pierce, J. & Hanawalt, P. C. (1989) *J. Biol. Chem.* **264**, 14455–14462.
23. Bohr, V. A. (1990) *J. Cancer Res. Clin. Oncol.* **116**, 384–391.
24. Chen, R.-H., Maher, V. M. & McCormick, J. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8680–8684.
25. McGregor, W. G., Chen, R.-H., Lukash, L., Maher, V. M. & McCormick, J. J. (1991) *Mol. Cell. Biol.* **11**, 1927–1934.
26. Gamper, H. B. & Hearst, J. E. (1982) *Cell* **29**, 81–90.
27. Liu, L. F. & Wang, J. C. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7024–7027.
28. Patterson, M. & Chu, G. (1989) *Mol. Cell. Biol.* **9**, 5105–5112.
29. Venema, J., Mullenders, L. H. F., Natarajan, A. T., van Zeeland, A. A. & Mayne, L. V. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4707–4711.
30. Weeda, G., van Ham, R. C. A., Vermeulen, W., Bootsma, D., van der Eb, A. J. & Hoeijmakers, J. H. J. (1990) *Cell* **62**, 777–791.
31. Wood, R. D. (1991) *Nature (London)* **350**, 190.