© 1996 Oxford University Press
Nucleic Acids Research, 1996, Vol. 24, No. 18 **3499–3506**

Transcription-coupled and global genome repair in the *Saccharomyces cerevisiae RPB2* gene at nucleotide resolution

Marcel Tijsterman, Judith G. Tasseron-de Jong, Pieter van de Putte and Jaap Brouwer*

Laboratory of Molecular Genetics, Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands

Received July 3, 1996; Revised and Accepted August 1, 1996

ABSTRACT

Repair of UV-induced cyclobutane pyrimidine dimers (CPDs) was examined at single nucleotide resolution in the yeast *Saccharomyces cerevisiae,* using an improved protocol for genomic end-labelling. To obtain the sensitivity required for adduct detection in yeast, an oligonucleotide-directed enrichment step was introduced into the current methodology developed for adduct detection in *Escherichia coli.* With this method, heterogeneous repair of CPDs within the *RPB2* locus is observed. Individual CPDs positioned in the transcribed strand are removed very efficiently with identical kinetics. This fast repair starts within 23 bases downstream of the transcription initiation site. The non-transcribed strand of the active gene exhibits slow repair without detectable repair variations between individual lesions. In contrast, CPDs positioned in the promoter region show profound repair heterogeneity. Here, CPDs at specific sites are removed very quickly, with comparable rates to CPDs positioned in the transcribed strand, while at other positions lesions are not repaired at all during the period studied. Interestingly, the fast repair in the promoter region is dependent on the *RAD7* and *RAD16* genes, as are the slowly repaired CPDs in this region and in the non-transcribed strand. This indicates that the global genome repair pathway is not intrinsically slow and at specific positions can be as efficient as the transcription-coupled repair pathway.

INTRODUCTION

When cells are subjected to UV light two major classes of lesions are introduced into the DNA (1), the cyclobutane pyrimidine dimer (CPD) and the pyrimidine (6–4) pyrimidone photoproduct (6–4PP). Both lesions are substrates for the nucleotide excision repair (NER) pathway. The CPD has been the most studied photoproduct, since its detection can be achieved by the phage enzyme T4 endonuclease V (T4endoV), which specifically recognizes CPDs and incises immediately 5′ of the lesion (2). Repair of CPD lesions is heterogeneous throughout the genome. Gene-specific repair analysis showed that lesions in active genes are more efficiently repaired than lesions in non-active DNA (3), primarily due to preferential repair of the transcribed strand over the non-transcribed strand. This phenomenon has been observed in mammalian cells (4), *Escherichia coli* (5) and *Saccharomyces cerevisiae* (6) and is dependent on transcription (7,8), indicating a role for the transcription process in efficient recognition of DNA adducts. In yeast, repair of UV-induced CPDs requires the 'core' NER enzymes Rad1, Rad2, Rad3, Rad4, Rad10, Rad14, Rad25 and Ssl1 (reviewed in 9). Besides these core enzymes, specific gene products are involved in the repair of different DNA sequences. In *rad26*Δ mutants, efficient repair of the transcribed strand is severely impaired (10), suggesting a specific function for Rad26p in the removal of CPDs from the transcribed strand of active genes. Other gene products are specifically involved in NER of non-transribed DNA. In *rad7*Δ and *rad16*Δ single and double mutants, repair of CPDs in non-transcribed strands of different active genes is completely abolished (11). However, efficient repair of the transcribed strand is unaffected in these mutants, indicating that the transcription-coupled repair pathway does not require these gene products. These observations led to the postulation of two subpathways of NER, namely transcription-coupled repair (TCR), which is dependent on transcription and stimulated by the *RAD26* gene product, and global genome repair (GGR), which requires the Rad7 and Rad16 proteins. Although it is clear that these proteins function in different subpathways (12), it is still unknown how these proteins act at the molecular level.

Recently, it has been shown that variations in repair rate are not confined to the gene-specific level. *In vivo* repair kinetics can vary even within a single DNA strand. CPDs are removed non-homogeneously from the *lacI* gene in *E.coli* (13) and from the *p53* and *PGK1* genes in human cells (14,15). Repair heterogeneity will have significant implications for mutagenesis, since slow repair of specific DNA damages might underlie the hotspots for mutation induction observed in various target genes in tissue culture (16) and in tumours (17). The objective of this study was to determine the kinetics of NER in *S.cerevisiae* at single nucleotide resolution. To obtain quantitative adduct detection in yeast cells, a purification and end-labelling procedure was

* To whom correspondence should be addressed

Table 1. Oligonucleotide primers

primer	sequence	T_m (°C)[a]	Enzyme
RR1TB	5'ttttttACCTAATTTTTCATAGACACAAAACAAAGTT-3'	57.6	*Ava*I
RD8TB	5'gggggAAAAGCATAATCCCAAATAATTC?-3'	51.7	*Dra*I
RH9TB	5'gggggCCACCTTAAAATAATTGGTTACCC-3'	50.3	*Mse*III
RN6TB	5'tttttCATGCAATAAGGAACTGACCAC-3'	53.5	*Nla*III
RR1NB	5'tttttACCCTAACTTTATAAAGCGGCTCCC-3'	56.6	*Ava*I
RP7NB	5'gggggTGCAGTAATTGGTGCACTTTCA-3'	55.5	*Pst*I
RE2TB	5'tttttAGCAATCCGGAAGTTTCTATGATTAGA-3'	56.7	*Eco*RV
RR2NB	5'tttttACCCAATGGTTTTTGAGGATAATACAGAAT-3'	58.8	*Ava*I
RP7N	5'TGCAGTAATTGGTGCACTTTCA-3'	55.5	-
RR2N	5'ACCCAATGGTTTTTGAGGATAATACAGAAT-3'	58.8	-

[a]T_m values were calculated with software using nearest-neighbour thermodynamic values (20)

developed partly based on methodology previously used to analyse repair in *E.coli* (13). This procedure allows the detection of *in vivo* DNA adduct incidence as well as the analysis of repair kinetics at the nucleotide level. The *RPB2* locus was chosen as a target because this gene has been extensively used in gene-specific repair analysis (8,10–12). In this report we show that fast repair of the *RPB2* locus starts near the transcription initiation site. The kinetics for this efficient repair are identical for differently positioned CPDs in the transcribed strand. In contrast, repair 5′ of the transcription start site is very heterogeneous on both DNA strands. These repair variations are not observed for CPDs in the non-transcribed strand of the transcription unit. However, both heterogeneous repair found in the promoter and slow repair of the non-transcribed strand require the *RAD7* and *RAD16* gene products. Therefore, all CPDs positioned within these regions are substrates for the global genome repair pathway.

MATERIALS AND METHODS

Yeast strains and media

The *S.cerevisiae* wild-type strain used for this study was W303-1B. The isogenic *rad* mutant strains used were MGSC104*rad7*Δ::*LEU2* and MGSC126*rad16*Δ::*LEU2* (11). All strains were kept on selective YNB medium (0.67% yeast nitrogen base, 2% glucose, 2% bacto-agar) supplemented with the appropriate markers. Cells were grown in complete medium (YEPD; 1% yeast extract, 2% bacto-peptone, 2% glucose) at 28°C under vigorous shaking.

UV irradiation and DNA isolation

Yeast cells diluted in chilled phosphate-buffered saline were irradiated with 254 nm UV light (Philips TUV 30W) at a rate of 3.5 J/m^2/s. Cells were collected by centrifugation, resuspended in complete medium and incubated for various times in the dark at 28°C prior to DNA isolation (18). DNA samples were purified on CsCl gradients (19).

Oligonucleotide-directed purification of a single DNA target

Samples of 20 µg DNA, containing ~1×10^9 copies of the yeast genome, were digested with an appropriate restriction endonuclease and precipitated according to standard procedures (22). Dynal M-280® streptavidin beads were used to enrich the desired chromosomal DNA target. After 3 min incubation at 93°C, 1 pmol biotinylated oligonucleotide (Table 1) complementary to the fragment of interest was annealed in 100 µl Beads-Binding buffer (BB buffer; 10 mM Tris–HCl, pH 9.0, 1.5 mM $MgCl_2$, 50 mM KCl and 0.1% Triton X-100) for 30 min. To increase specificity, the annealing temperature was chosen 2°C beneath the predicted T_m of the primers used. Subsequently, 10 µl (1 mg/µl) streptavidin-coated magnetic beads were added (pre-washed with BB buffer) and incubated for 15 min with occasional gentle agitation to avoid bead sedimentation. Using the Dynal magnetic partical concentrator, the immobilized templates were washed once with BW solution (2.0 M NaCl, 5.0 mM Tris–HCl, pH 7.5, 0.5 mM EDTA), three times with BB buffer and once with TE (10.0 mM Tris–HCl, pH 8.0, 1 mM EDTA). The captured DNA fragments were eluted from the beads by incubating for 3 min at room temperature in 10 µl 0.1 M NaOH.

Oligonucleotide-directed end-labelling

End-labelling conditions were used as described (13) with some modifications. An oligomer was designed to be complementary to the 3′-end of the desired DNA fragment with a six nucleotide non-complementary dGTP or dTTP stretch (Table 1). This nucleotide stretch is used as a template to extend the free 3′-OH end of the restriction fragment of interest with [α-^{32}P]dNTP (either dATP or dCTP depending on the primer used) and Taq DNA polymerase. The reaction mixture was generated by sequential addition of 10 µl 0.1 M NaOH containing the purified DNA fragment (see above), 37 µl BB buffer, 1.0 µl 1 M HCl, 1.0 pmol oligonucleotide, 0.2 µl [α-^{32}P]dA/CTP (3000 Ci/mmol) and 0.2 U SuperTaq polymerase (HT Biotechnology Ltd). Samples were denatured for 3 min at 93°C and subjected to four consecutive rounds of denaturing (30 s at 93°C), annealing (30 s at T_m) and extension (90 s at 72°C) to optimize end-labelling. To assure complete extension, 1 µl 10 mM dA/CTP was added followed by two additional cycles. Phenol/chloroform extraction was performed to exclude Taq activity in the later steps.

Cleavage at CPDs

CPDs were identified using T4endoV. Since incision is most efficient on a double-stranded (ds) DNA substrate, the end-labelled fragments were subjected to a hybridization protocol. A 200-fold molar excess of complementary strand, synthesized by linear amplification, was added, followed by 3 min incubation at 93°C and gradual cooling to room temperature. Native gel electrophoresis showed that all labelled DNA fragments were in the dsDNA configuration. The DNA was precipitated, redissolved and divided into two equal parts. One was incubated with T4endoV, while the other was mock treated. Samples were subjected to spin column chromatography and lyophilized to small volumes. Approximately equal amounts (measured as c.p.m.) were loaded on 6% denaturing acrylamide gels alongside Maxam–Gilbert sequencing reactions. After drying, autoradiograms were prepared from the gels.

Quantitation of repair rates

Autoradiograms were scanned using an LKB Ultrascan XL densitometer (Pharmacia) and analysed using ImageMaster™ software (Pharmacia). Background levels were subtracted and gel band intensities were corrected for loading variations. Serial dilutions of Maxam–Gilbert sequencing reactions were used to

determine the linear range of the autoradiograms. Quantification data were obtained from experiments carried out in triplicate. Repair plots were established for each CPD that gave a sufficient signal to background ratio and were within the linear range of Kodak X-OMAT™-AR scientific imaging films. The time at which 50% of the initial damage (signal at $t = 0$) was removed was calculated from these plots.

Oligonucleotide primers

Oligonucleotides specific for the *S.cerevisiae RPB2* locus were used to map CPDs along the *RPB2* promoter (oligonucleotides RR1TB and RR1NB), the transcription initiation site (oligonucleotides RD8TB, RH9TB and RP7NB) and within the transcription unit (oligonucleotides RE2TB and RR2NB). All oligonucleotide primers listed in Table 1 contain 5′-biotinylated ends and therefore could also be used in the oligo-directed purification protocol. All primers indicated were also synthesized without the six base non-complementary extension to generate PCR fragments for Maxam–Gilbert sequencing.

Maxam–Gilbert sequencing reactions

Maxam–Gilbert sequencing ladders were obtained according to standard procedures (21) using PCR fragments identical to the chromosomal DNA fragment under analysis. After the sequencing reactions the fragments were ^{32}P-labelled using the tailed oligonucleotides (as described earlier). In this way, a 3′-end-labelled product identical to the chromosomal DNA fragment used in the repair analysis was obtained.

RESULTS

Repair analysis at the nucleotide level

Saccharomyces cerevisiae cells were irradiated with UV light. DNA was isolated directly after irradiation and at several post-incubation time points. The DNA was digested with an appropriate restriction endonuclease. The experimental system to detect CPDs at nucleotide resolution as described in Materials and Methods is outlined in Figure 1. In brief, purification of the fragment of interest was obtained by annealing to a complementary biotinylated oligonucleotide. DNA hybrids were captured with paramagnetic streptavidin-coated beads. After extensive washing the target strand was eluted from the immobilized oligonucleotide. The DNA was labelled using an oligo-directed end-labelling procedure (13). Since incision of the DNA strand 5′ of the CPD by T4endoV is most efficient on dsDNA (2), the complementary strand was added in excess and annealed to the target. Samples were treated or mock treated with T4endoV, concentrated and subjected to denaturing PAGE alongside Maxam–Gilbert reactions of the corresponding sequence labelled in an identical manner. The positions of the individual CPDs are indicated by the length of the fragments, while the intensity of each signal is a measure of the frequency of the photoproduct at that particular position. The obtained distribution pattern reflects the *in vivo* CPD levels at single nucleotide resolution. This method obviates the need for PCR amplification, thereby circumventing disproportionate adduct distribution patterns due to site-to-site variations in amplification (22) and ligation efficiencies (23). After post-incubation, repair of CPDs at specific nucleotides will result in a decrease in the T4endoV-specific signal compared with the intensity detected directly after irradiation.

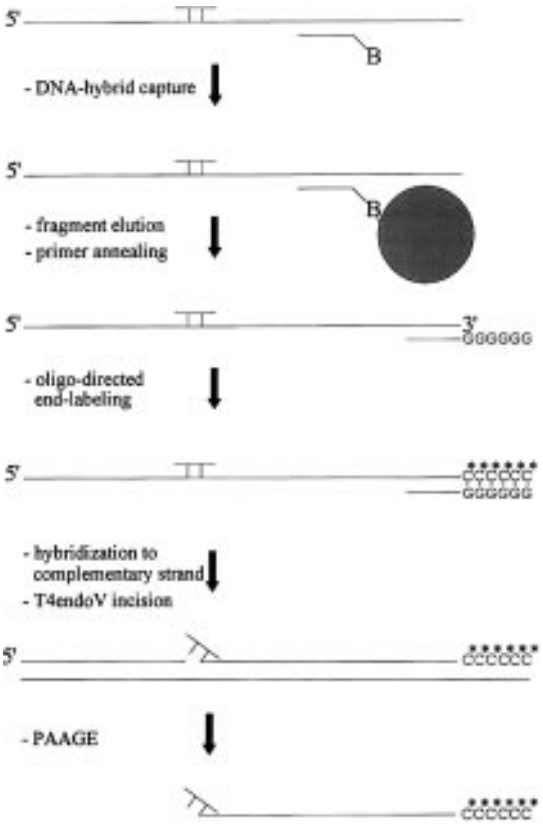

Figure 1. Schematic representation of the protocol to analyse repair with single nucleotide resolution in *S.cerevisiae*. The DNA fragment of interest is purified using magnetic particles covalently linked to biotinylated oligonucleotides. Subsequently, an oligonucleotide-directed end-labelling assay is used in which the 3′-end of the DNA fragment of interest is extended using an oligonucleotide as template. The single-stranded DNA target is annealed to an excess of added complementary strand. Detection of the damage occurs by virtue of CPD-specific DNA incision on the dsDNA substrate. Fragments of various length correspond to the various positions of CPDs, while the intensity of the corresponding gel band is a measure of the frequency of CPDs at that specific position.

Initial photoproduct frequency

The initial distribution pattern of CPDs after UV irradiation can be seen in Figure 2 and in the $t = 0$ lanes of Figures 3, 4, 5 and 7. As expected, lesions are exclusively found at adjacent pyrimidines. *In vivo* adduct levels are heterogeneous throughout the gene. On average, the order of preference for CPD induction is TT > TC ≈ CT > CC, and increased levels of induction are observed when one or more pyrimidines are positioned 5′ of the dinucleotide. These observations are consistent with experiments using cloned end-labelled DNA for irradiation (24,25). Not all potential dimer sites result in significant photoproduct formation when cells are irradiated *in vivo*. Previously, others have detected photofootprints in yeast (26,27) in which the absence of *in vivo* photoproduct induction was attributed to local protein–DNA interaction. When we compared the *in vivo* distribution pattern to the pattern obtained from DNA irradiated *in vitro*, clear differences were observed, e.g. formation of photoproducts at nt

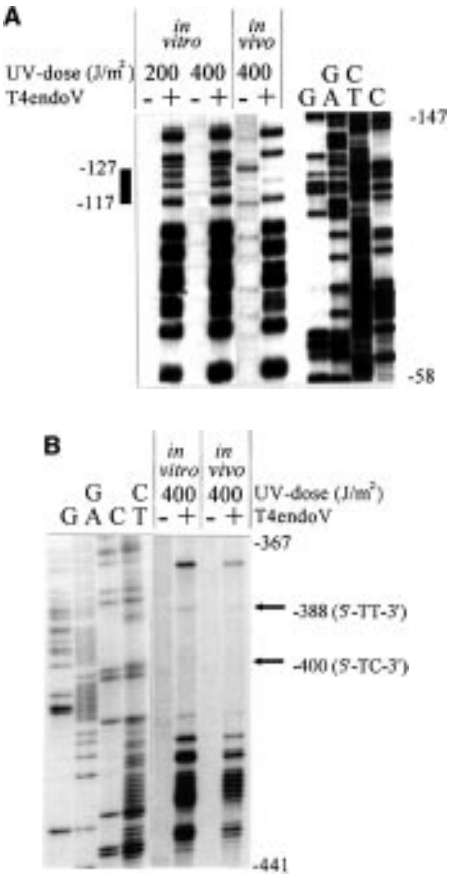

Figure 2. Determination of CPDs 5′ of the yeast *RPB2* locus at dinucleotide position –147 to –58 for the non-template strand (**A**) and at dinucleotide position –441 to –367 for the template strand (**B**). DNA was irradiated with 200 or 400 J/m^2 *in vitro* and with 400 J/m^2 *in vivo*. Samples mock-treated or treated with the dimer-specific enzyme T4endoV are denoted – and + respectively. Lanes G, G+A, T+C and C are Maxam–Gilbert sequencing reactions. Dinucleotide positions mentioned in the text are indicated. CPD bands migrate ~1 base slower than the corresponding 5′ nucleotide of the dipyrimidine pair in the Maxam–Gilbert lanes.

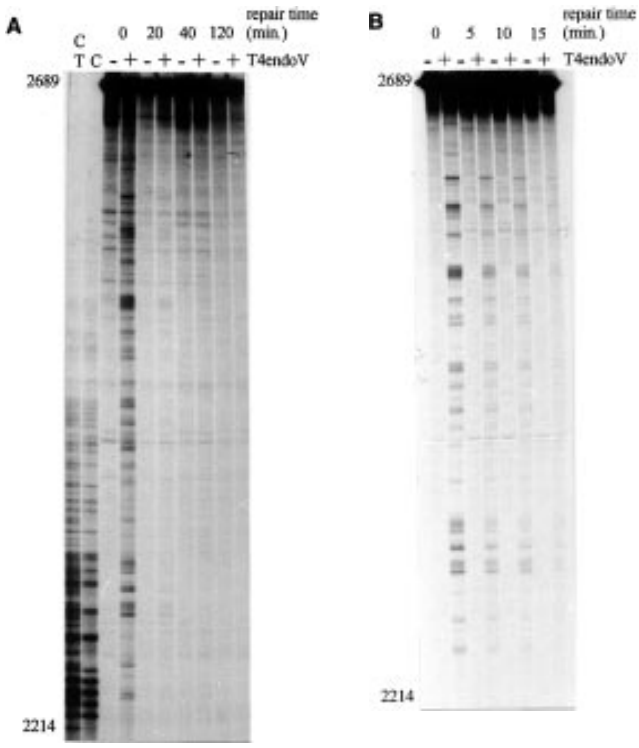

Figure 3. Repair of UV-induced CPDs at single nucleotide resolution along downstream sequences of the yeast *RPB2* gene. Data are for the transcribed DNA strand for nt +2214 to +2689. Cells were irradiated with 70 J/m^2 and repair was allowed for (**A**) standard post-incubation periods (0, 20, 40 and 120 min) and (**B**) short post-incubation periods (0, 5, 10 and 15 min). *Eco*RV and *Rsa*I were used as endonucleases. Purification and end-labelling utilized primer RE2TB. Samples mock-treated or treated with the dimer-specific enzyme T4endoV are denoted – and + respectively. Lanes T+C and C are Maxam–Gilbert sequencing reactions

–127 to –117 is significantly lower *in vivo*. (Fig. 2A). However, other potential dimer sites, e.g. nt –388 (5′-TT-3′) and –400 (5′-TC-3′) do not show detectable photoproduct levels either after *in vivo* or after *in vitro* irradiation at 70 J/m^2 (data not shown). At 400 J/m^2 low levels of CPDs are induced at these sites both *in vitro* and *in vivo* (Fig. 2B), demonstrating that the lack of CPD induction at 70 J/m^2 is not the result of DNA protection *in vivo*. Therefore, cold spots for DNA damage induction are not only influenced by DNA-interacting proteins but also by the sequence context.

Repair of CPDs in the coding region of *RPB2*

To determine the repair rates of UV-induced CPDs *in vivo* at single nucleotide resolution, cells were irradiated at a UV dose of 70 J/m^2 and incubated to allow DNA repair. Figure 3 shows repair of CPDs along the transcribed strand of the *RPB2* locus in an ORF fragment at position +2214 to +2689. DNA adduct levels were determined directly after UV irradiation and following 20, 40 and 120 min incubation. After 20 min incubation, over 80% of the signal present at $t = 0$ is removed for all CPDs (Fig. 3A). To study in more detail whether repair rates vary between different CPDs shorter intervals were used. Figure 3B shows CPD levels at 5, 10 and 15 min after UV irradiation. Repair plots were produced for each individual CPD to determine the $t_{1/2}$ value as the time at which 50% of the signal present at $t = 0$ has disappeared. $t_{1/2}$ values of 8 ± 1 min were found for dinucleotides in the transcribed strand. No significant variations in repair rate were observed for differently positioned CPDs, nor for different dipyrimidine combinations (e.g. TT, CT, TC and CC).

The removal of CPDs from the non-transcribed strand of the *RPB2* locus at this region is distinct. $t_{1/2}$ values are 120 ± 3 min for each CPD examined in this fragment, but for this strand also no significant variation in repair rate between individual dinucleotides could be observed (data not shown). Thus, individual dinucleotides in the transcribed strand are repaired 15 times more efficiently compared with lesions positioned in the non-transcribed strand, but no significant positional repair variations were observed between different dinucleotides in both DNA strands.

Repair of CPDs near the transcription initiation site

Fast repair of CPDs was observed in the transcribed strand of the *RPB2* locus compared with repair in the non-transcribed strand. To determine whether the start of this fast repair coincides with

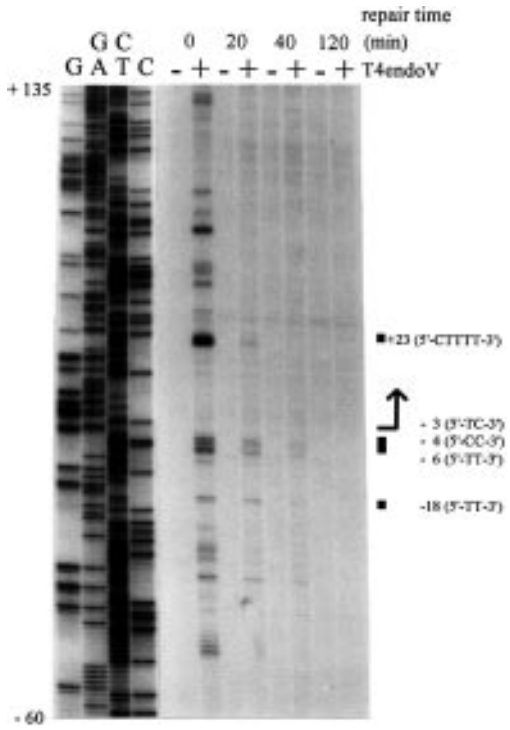

Figure 4. Repair of UV-induced CPDs near the transcription initiation site along the template strand of the *RPB2* gene. Sequences between nt –60 and +135 with respect to the transcription initiation site are shown. *Hsp*92II and *Pst*I were used as endonucleases. Purification and end-labelling utilized primer RH9TB. The large arrow indicates the major transcription initiation site and the direction of transcription. CPD-specific positions described in the text are indicated.

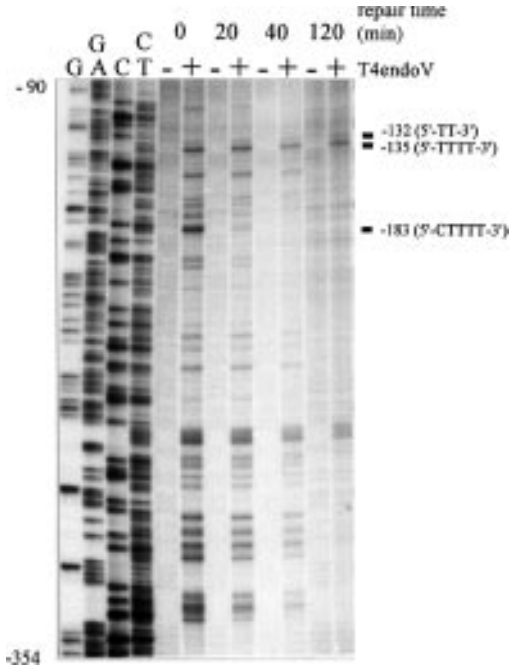

Figure 5. Repair of UV-induced CPDs along the promoter region of the *RPB2* gene for the template strand nt –354 to –90 (primer RR1TB). *Rsa*I was used as restriction endonuclease. CPDs discussed in the text are indicated.

the start of transcription, repair analysis was performed on the transcribed strand around the transcription initiation site. The initiation site was previously designated at nt 269 (± 25 bp) 5′ of the ATG using S1 nuclease digestion (28). To allow more accurate correlation of CPD repair with transcription initiation, we used primer extension to refine the mapping of the major transcription initiation site at 278 bp 5′ of the ATG (data not shown). All sequence positions mentioned are calculated according to this position (nt +1). Figure 4 shows induction and repair of CPDs along the template for transcription of *RPB2* from position –60 to +135. CPDs induced immediately 5′ of the transcription initiation site show moderate repair rates. $t_{1/2}$ values calculated for lesions at nt –3 (5′-TC-3′), –4 (5′-CC-3′), –6 (5′-TT-3′) and –18 (5′-TT-3′) were 26, 24, 24 and 27 min respectively. However, fast repair of the template strand is observed for CPDs at dinucleotide +23 and for all CPDs which are 3′ of this position ($t_{1/2}$ = 8 min). Nucleotide +23 is the first position 3′ of the transcription initiation site with detectable adduct formation. Although potential dimer sites are present at DNA positions +1 (5′-TC-3′) and +17 (5′-TT-3′), these did not result in detectable CPD incidence *in vivo* or *in vitro*. Thus, fast repair of the transcribed strand starts within 23 bases from the transcription initiation site and continues downstream into the transcribed strand. The CPDs analysed downstream of this dimer site are repaired with equal efficiency and exhibit identical repair rates to the CPDs within the analysed ORF fragment further downstream.

Repair of CPDs in the promoter region of the *RPB2* gene

Removal of CPDs from the promoter region of the *RPB2* locus was analysed in both DNA strands. For both strands a considerable repair heterogeneity is observed. For example, in the template strand, repair at positions –132 (5′-TT-3′) and –135 (5′-TTTT-3′) is not detectable after 120 min of repair, whereas $t_{1/2}$ for nt –183 (5-CTTTT-3′), separated from the latter by <50 nt, is 14 min (Fig. 5). Hence, at least a 10-fold variation in repair between specific nucleotides can be observed in the upstream region of the *RPB2* locus. Figure 6 shows a schematic representation of dimer removal along the promoter and the transcription initiation site for both strands of the *RPB2* gene. An interesting observation is that specific dimer sites positioned outside the transcribed regions of the *RPB2* gene are repaired with comparable rates to CPDs in the transcribed strand. For example, CPDs between nt –70 and +15 in the non-template strand are removed with $t_{1/2}$ values of the order of 10 min (Fig. 6).

Repair analysis in mutants deficient in global genome repair

Repair rates within the promoter region were shown to be heterogeneous, in contrast to the slow and uniform repair of the non-transcribed strand of the active gene. Neither DNA sequence was transcribed and therefore repair of CPDs within these sequences should be dependent on the global genome repair pathway (12). This suggests that at specific positions, the global genome repair pathway can be very fast. To investigate whether quickly repaired CPDs positioned within the promoter are indeed substrates for this pathway, repair analysis was performed in a *rad7*Δ disruption mutant, which is disturbed in global genome

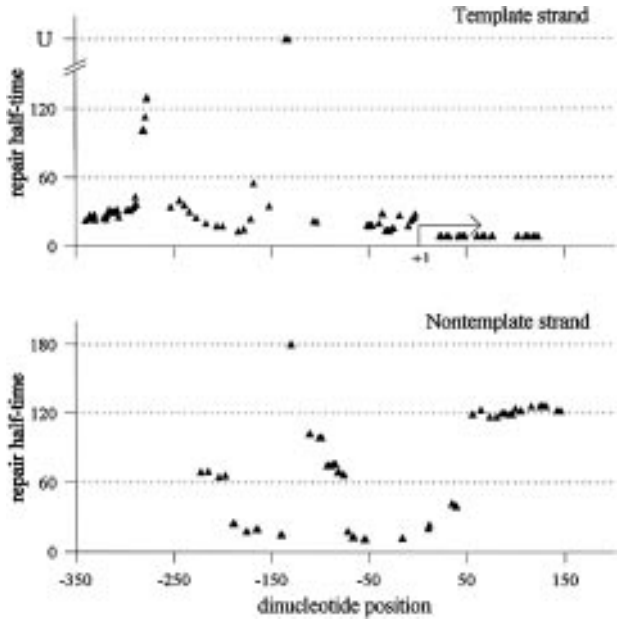

Figure 6. Graphic representation of quantified CPD repair rates along the promoter and transcription initiation site of the *RPB2* locus. The major initiation site is indicated by the arrow at position +1. Repair half-time values, determined as the time at which 50% of the initial CPD signal was removed, were calculated for each individual CPD position with a sufficient signal-to-noise ratio and are depicted for the transcribed and non-transcribed strand above their corresponding dipyrimidine position. Repair $t_{1/2}$ = U indicates that CPDs were unrepaired after 2 h post-incubation.

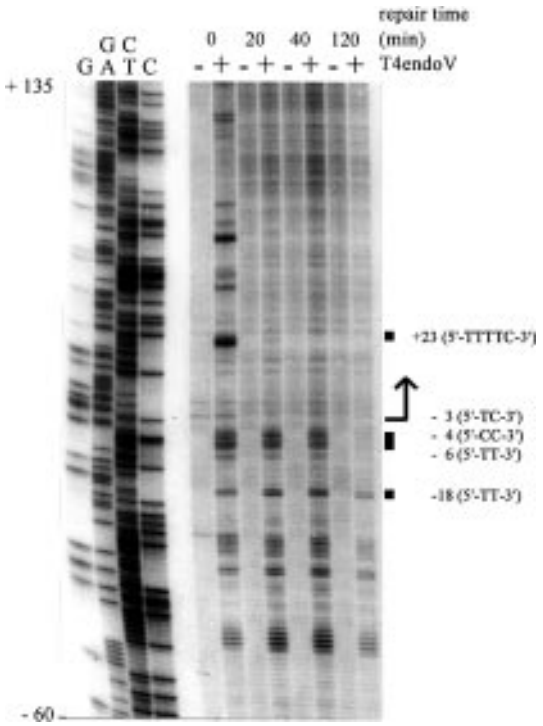

Figure 7. Repair of UV-induced CPDs in a *S.cerevisiae rad7*Δ background. Data for the transcribed strand of the *RPB2* gene near the transcription initiation site. Sequences between nt –146 and +358 with respect to the transcription initiation site are shown. *Hsp*92II and *Pst*I were used as endonucleases. Purification and end-labelling utilized primer RH9TB. The large arrow indicates the major transcription initiation site and the direction of transcription. CPD-specific positions described in the text are shown. Less DNA was loaded in lane 120 min.

repair (11). In this mutant, no repair could be observed for sequences upstream of the transcription initiation site in the template strand (Figs 7 and 8). Also, CPDs positioned in the non-template strand were not repaired, neither near the transcription initiation site (Fig. 8) nor in the ORF fragment (data not shown). This indicates that repair of each CPD within these sequences is completely dependent on the *RAD7* gene product. Furthermore, a repair gradient could be observed near the transcription initiation site (Fig. 7). CPDs positioned 5′ of nt –18 (5′-TT-3′) were not repaired at all in *rad7*Δ, whereas moderate repair was observed at positions –3 (5′-TC-3′), –4 (5′-CC-3′) and –6 (5′-TT-3′), which were completely repaired after 120 min. Subsequent fast transcription-coupled repair was observed from nt +23 onwards. This repair gradient at the transcription initiation site is only seen for the template strand. A schematic representation of the repair $t_{1/2}$ values is depicted in Figure 8. Identical results were obtained for a *rad16*Δ mutant (data not shown), which is also deficient in the global genome repair pathway.

DISCUSSION

Intragenic repair variation is of considerable interest with regard to the mutagenic potential of carcinogens. In both human tumours (17) and tissue culture (16), hotspots for mutation induction have been found in different target genes. Recent data suggest that slow repair of DNA damage at specific sites underlies the observed mutation hotspots (13,14), although mutation spectra are also biased by phenotypic selection. Another parameter that is influenced by the sequence context is the initial distribution of DNA lesions (24,25; this study). Since mutations are produced when DNA lesions are by-passed by DNA polymerase during DNA replication, the mutation frequency at any nucleotide position depends on, besides the mutagenic potential of the lesion itself, the product of both parameters.

We have developed an assay to detect DNA lesions in the yeast *S.cerevisiae*. With this method, we have analysed repair of UV-induced CPDs along the *RPB2* locus. Since different laboratories have used this locus to study repair at the gene-specific level, data obtained in both assays could be compared. CPDs within the transcribed strand are repaired with a $t_{1/2}$ of ~8 min, whereas CPDs positioned in the non-transcribed strand are repaired with a $t_{1/2}$ of ~120 min. Repair rates of CPDs positioned in the transcribed strand appear higher than repair rates observed in gene-specific repair analysis using this locus (8,11,12). Since in those data CPD frequencies were averaged over kilobase-length DNA fragments of which only part was transcribed, repair kinetics were influenced by slowly repaired CPDs in non-transcribed regions of the DNA fragment.

Fast repair of the transcribed strand is observed downstream of the transcription initiation site from dinucleotide position +23 onwards. However, since no significant number of lesions were detected between this position and the transcription initiation site, transcription-coupled repair might start farther upstream. An indication of fast repair prior to dinucleotide 23 is the observation that CPDs at positions –3 and –4 are repaired in a *rad7*Δ mutant, albeit with reduced efficiency. Positions for the onset of

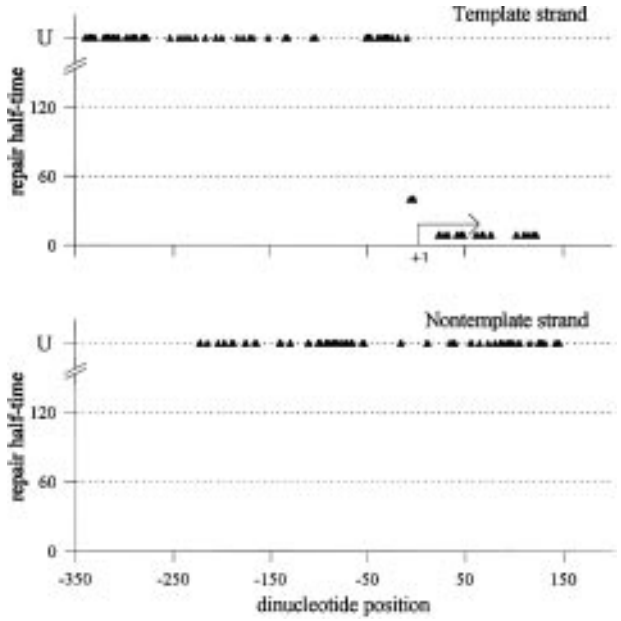

Figure 8. Graphic representation of quantified CPD repair rates along the promoter, transcription initiation site and downstream sequences of the *RPB2* locus in a yeast *S.cerevisiae rad7*Δ mutant strain. The major initiation site is indicated by the arrow at nt +1. $t_{1/2}$ values of individual CPDs are depicted above their corresponding nucleotide position for the transcribed and non-transcribed strand. Repair $t_{1/2}$ = U indicates that CPDs were unrepaired after 2 h post-incubation.

transcription-coupled repair have been determined in different organisms and different genes and do not seem to coincide exactly with the transcription initiation site. In *E.coli*, fast repair of CPDs in the *lacI* and *lacZ* genes starts 10 and 32 bp respectively downstream of the transcription initiation site (29). In the human *PGK1* gene, fast repair starts in a region 140 bases downstream of the transcription initiation site (15). Recently, however, repair analysis along the UV-inducible human *JUN* promoter showed that for this gene fast repair of the transcribed strand starts upstream of the transcription initiation site (30). These authors suggested that the presence of the general transcription factor TFIIH, which has a dual role in transcription and NER (31), results in locally increased repair efficiency. Although this explanation seems plausible, it might only be true for promoters with very high transcriptional activity, since this phenomenon is absent in the *PGK1* gene in human cells and the *RPB2* gene in *S.cerevisiae* (15; this work).

Fast repair of the transcribed strand of the *RPB2* gene continues from position +23 downstream into the gene, with uniform repair rates for differently positioned CPDs. This observation strengthens the hypothesis that the elongating RNA polymerase has a role in efficient repair of DNA lesions from transcribed regions (32). It has been shown that elongating RNA polymerase is blocked by CPDs in the transcribed strand *in vitro* (33,34). The uniform rates at which individual sites are repaired in the transcribed strand can be explained assuming that recognition of the damage by the RNA polymerase determines the repair rate of individual lesions. Once transcription is initiated, lesions are recognized by the RNA polymerase with equal probability under conditions where little variation in transcription rate exists throughout the gene. Provided that each CPD blocks transcription to the same extent, this results in identical recognition rates and therefore uniform repair rates. However, uniform repair within the transcribed strand of the *RPB2* gene is in contrast to the repair heterogeneity observed in the transcribed strand of the human *p53* gene (14). Slow repair was observed at dinucleotide positions frequently mutated in skin cancer, suggesting that repair variation strongly influences mutation induction in the *p53* gene. This profound repair heterogeneity in the transcribed strand of the *p53* gene is not a general rule in human genes, since only moderate repair variations are observed in the human housekeeping gene *PGK1* and the UV-inducible *JUN* locus. Also, in the *E.coli lacI* gene, repair variation in the transcribed strand is confined to one slowly repaired dinucleotide (13). Furthermore, it has been shown for this gene (29) that modest repair heterogeneity observed immediately downstream of the transcription initiation site converts to fast and uniform repair upon induction of the gene with isopropylthiogalactosepyranoside, suggesting a more uniform repair pattern when transcription activity is increased. These latter observations suggest that the transcribed strands of active genes are repaired in general with little or no repair heterogeneity. Repair variations in the transcribed strand of the *p53* gene might be explained by a low transcription rate. Reduced transcription-coupled repair probably leads to a more prominent role of global genome repair, with possible heterogeneity, in repair of CPDs from the transcribed strand of the *p53* gene, especially since in human cells repair of the transcribed strand is only 2-fold more efficient compared with the non-transcribed strand. In support of this hypothesis, repair of CPDs from the transcribed strand is more efficient in *JUN* and *PGK1* compared with *p53* when repair rates are averaged. Although repair efficiency clearly influences mutation induction in both strands of a target gene (35,36), the observed uniform repair rates of CPDs within the transcribed strand imply a more prominent role for CPD induction levels in the distribution of UV-induced mutations, since CPD induction is heterogeneous and dependent on the sequence context. Further support for this suggestion awaits analysis of DNA damage incidence, repair and mutation spectra in a yeast locus (in progress).

Repair rates for the non-transcribed strand also do not exhibit significant positional variations, with $t_{1/2}$ values of the order of 120 min. This is in contrast to the profound heterogeneity in repair of CPDs located upstream of the transcription initiation site, where at least a 10-fold variation can be observed between individual lesions depending on the dinucleotide position. We suggest two possible explanations for the observed differences in repair of these distinct non-transcribed DNA regions. One possibility is that repair of non-transcribed DNA exhibits uniform slow repair rates throughout the genome and behaves like the non-transcribed strand of the *RPB2* gene, except at positions in the genome with a more open or disturbed chromatin structure. Chromatin perturbations at promoter sequences might render the DNA more accessible to repair proteins, as they do for the transcription initiation machinery.

The other possibility is that repair heterogeneity is an intrinsic feature of global genome repair. CPDs are repaired with profound variations depending on the chromatin organization and accessibility to DNA repair proteins. Heterogeneous repair of non-transcribed DNA turns to uniform repair only when transcription on the opposite strand disturbs or randomizes the local chromatin organization (37). In this hypothesis, transcription leads not only

to uniformity of repair rates in the transcribed strand, but also in the non-transcribed strand, although with reduced efficiency, since recognition of the damage still depends on different factors. Also, a combination of both possibilities could underlie the repair characteristics observed.

Fast repair patches within the promoter are not the consequence of abberrant transcription, since this repair heterogeneity is totally dependent on the Rad7 and Rad16 proteins. This indicates that global genome repair at specific positions can be very efficient and even comparable with repair observed for CPDs in the transcribed strand. Thus, global genome repair is not necessarily inefficient. This observation suggests that slow repair of specific CPDs is due to inhibition rather than to the previously assumed intrinsic slow repair rate of the global genome repair pathway for CPDs. Bulky chemical adducts and 6–4PPs are repaired more efficiently by the global genome repair pathway than are CPDs (38). It has been suggested that the more profound disturbance of the DNA conformation at the site of damage underlies this difference. CPDs are minor distorting lesions compared with bulky adducts (39,40) and probably therefore less well recognized. One can envisage that dimers positioned at dinucleotides which are arranged in a nucleosomal structure are not accessible to DNA repair proteins unless specific gene products rearrange the DNA structure. However, all CPDs examined in the non-transcribed DNA of the *RPB2* locus, i.e. the fast and slowly repaired lesions, require the RAD7 and RAD16 proteins, indicating that differences in repair rates for individual CPDs do not result from the action of these proteins at specific positions.

In summary, we have analysed repair at the nucleotide level in the yeast *S.cerevisiae*. This report presents the methodology to study nucleotide excision repair *in vivo* at single nucleotide resolution in yeast. Since no amplification steps are used, the *in vivo* damage distribution levels are measured quantitatively. We have shown that heterogeneity in repair of CPDs is observed within the *RPB2* locus. Fast repair of the transcribed strand starts at or directly downstream of the transcription initiation site and exhibits uniform kinetics. Also, no significant variations in the repair rate are observed for differently positioned CPDs in the non-transcribed strand. However, profound variations are observed in the promoter region of this gene. Both heterogeneous repair within both strands of the promoter and slow repair of CPDs in the non-transcribed strand are totally dependent on the *RAD7* and *RAD16* gene products, which indicates that repair of CPDs by the global genome repair pathway can be efficient for non-transcribed DNA.

ACKNOWLEDGEMENTS

We thank Drs R.A.Verhage for discussion and critical reading of the manuscript, H.den Dulk for advice regarding the methodology and M.van Nierop and E.E.A.Verhoeven for technical assistance.

REFERENCES

1 Sage,E. (1993) *Photochem. Photobiol.*, **57**, 163–174.
2 Gordon,L.K. and Haseltine,W.A. (1980) *J. Biol. Chem.*, **255**, 12047–12050.
3 Bohr,V.A., Smith,C.A., Okumoto,D.S. and Hanawalt,P.C. (1985) *Cell*, **40**, 359–369.
4 Mellon,I.M., Spivak,G.S. and Hanawalt,P.C. (1987) *Cell*, **51**, 241–249.
5 Mellon,I. and Hanawalt,P.C. (1989) *Nature*, **342**, 95–98.
6 Smerdon,M.J. and Thoma,F. (1990) *Cell*, **61**, 675–684.
7 Leadon,S.A. and Lawrence,D.A. (1992) *J. Biol. Chem.*, **267**, 23175–23182.
8 Sweder,K.S. and Hanawalt,P.C. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 10696–10700.
9 Friedberg,E.C., Walker,G.C. and Siede,W. (1995) *DNA Repair and Mutagenesis.* American Society for Microbiology, Washington, DC.
10 Van Gool,A.J., Verhage,R., Swagemakers,S.M.A., van de Putte,P., Brouwer,J., Troelstra,C., Bootsma,D. and Hoeijmakers,J.H.J. (1994) *EMBO J.*, **13**, 5361–5369.
11 Verhage,R., Zeeman,A.-M., de Groot,N., Gleig,F., Bang,D.d., van de Putte,P. and Brouwer,J. (1994) *Mol. Cell. Biol.*, **14**, 6135–6142.
12 Verhage,R.A., van Gool,A.J., de Groot,N., Hoeijmakers,J.H.J., van de Putte,P. and Brouwer,J. (1996) *Mol. Cell. Biol.*, **16**, 496–502.
13 Kunala,S. and Brash,D.E. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 11031–11035.
14 Tornaletti,S. and Pfeifer,G.P. (1994) *Science*, **263**, 1436–1438.
15 Gao,S., Drouin,R. and Holmquist,G.P. (1994) *Science*, **263**, 1438–1440.
16 Brash,D.E., Seetharam,S.S., Kraemer,K.H., Seidman,M.M. and Bredberg,A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 3782–3786.
17 Ziegler,A. Leffell,D.J., Kunala,S., Sharma,H.W., Gailani,M., Simon,J.A., Halperin,A.J., Baden,H.P., Shapiro,P.E., Bale,A.E. and Brash,D.E. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 4216–4220.
18 Sherman,F., Fink,G.R. and Hicks,J.B. (1986) *Methods in Yeast Genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20 Rychlik,W. and Rhoads,R.E. (1989) *Nucleic Acids Res.*, **17**, 8543–8551.
21 Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560
22 Pfeifer,G.P., Steigerwald,S.D., Hansen,R.S., Gartler,S.M. and Riggs,A.D. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8252–8256.
23 Steigerwald,S.D., Pfeifer,G.P. and Riggs,A.D. (1990) *Nucleic Acids Res.*, **18**, 1435–1439.
24 Brash,D.E. and Haseltine,W.A. (1982) *Nature*, **298**, 189–192.
25 Bourre,F., Renault,G. and Sarasin,A. (1987) *Nucleic Acid Res.*, **15**, 8861–8875.
26 Selleck,S.B. and Majors,J. (1987) *Nature*, **325**, 173–177.
27 Axelrod,J.D. and Majors,J. (1989) *Nucleic Acids Res.*, **17**, 171–183.
28 Sweetser,D., Non-et,M. and Young,R.A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1192–1196.
29 Kunala,S. and Brash,D.E. (1995) *J. Mol. Biol.*, **246**, 264–272.
30 Tu,Y., Tornaletti,S. and Pfeifer,G.P. (1996) *EMBO J.*, **15**, 675–683.
31 Drapkin,R., Reardon,J.T., Ansari,A., Huang,J.H., Zawel,L., Ahn,K., Sancar,A. and Reinberg,D. (1994) *Nature*, **368**, 769–772.
32 Hanawalt,P.C., Donahue,B.A. and Sweder,K.S. (1994) *Curr. Biol.*, **4**, 518–521
33 Donahue,B.A., Yin,S., Taylor,J.S., Reines,D. and Hanawalt,P.C. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 8502–8506.
34 Selby,C.P and Sancar,A. (1990) *J. Biol. Chem.*, **265**, 21330–21336.
35 Vrieling,H. van Rooijen,M.L., Groen,N.A., Zdzienicka,M.Z., Simons,J.W.I.M., Lohman,P.H.M. and van Zeeland,A.A. (1989) *Mol. Cell. Biol.*, **9**, 1277–1283.
36 McGregor,W.G., Chen,R.H., Lukash,L., Maher,V.M. and McCormick,J.J. (1991) *Mol. Cell. Biol.*, **11**, 1927–1934.
37 Cavalli,G. and Thoma,F. (1993) *EMBO J.*, **12**, 4603–4613.
38 Mitchell,D.L. and Nairn,R.S. (1989) *Photochem. Photobiol.*, **49**, 805–819.
39 Taylor,J.S., Garrett,D.S. and Cohrs,M.P. (1988) *Biochemistry*, **27**, 7206–7215.
40 Taylor,J.S., Garrett,D.S., Brockie,I.R., Svoboda,D.L. and Tesler,J. (1990) *Biochemistry*, **29**, 8858–8866.