The EMBO Journal vol.13 no.24 pp.6143–6151, 1994

A new class of DNA photolyases present in various organisms including aplacental mammals

Akira Yasui[1,2], Andries P.M.Eker[3], Shinji Yasuhira[1,4,5], Hirohiko Yajima[1,6], Takayasu Kobayashi[1], Masashi Takao[1] and Atsushi Oikawa[1]

[1]Institute of Development, Aging and Cancer, Tohoku University, Seiryomachi 4-1, Aoba-ku, Sendai 980, Japan, [3]Department of Cell Biology and Genetics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands and [4]Department of Zoology, Faculty of Science, Tokyo University, Hongo 7-3-1, Bunkyo-ku, Tokyo 113, Japan
[5]Present address: Howard Hughes Medical Institute Research Laboratories, University of California, Los Angeles, CA 90024-1662, USA
[6]Present address: Laboratory of Molecular Genetics, Mitsubishi Kasei Institute of Life Sciences, 11 Mina-miooya, Machida-shi, Tokyo 194, Japan
[2]Corresponding author

Communicated by D.Bootsma

DNA photolyase specifically repairs UV light-induced cyclobutane-type pyrimidine dimers in DNA through a light-dependent reaction mechanism. We have obtained photolyase genes from *Drosophila melanogaster* (fruit fly), *Oryzias latipes* (killifish) and the marsupial *Potorous tridactylis* (rat kangaroo), the first photolyase gene cloned from a mammalian species. The deduced amino acid sequences of these higher eukaryote genes show only limited homology with microbial photolyase genes. Together with the previously cloned *Carassius auratus* (goldfish) gene they form a separate group of photolyase genes. A new classification for photolyases comprising two distantly related groups is proposed. For functional analysis *P.tridactylis* photolyase was expressed and purified as glutathione S-transferase fusion protein from *Escherichia coli* cells. The biologically active protein contained FAD as light-absorbing cofactor, a property in common with the microbial class photolyases. Furthermore, we found in the archaebacterium *Methanobacterium thermoautotrophicum* a gene similar to the higher eukaryote photolyase genes, but we could not obtain evidence for the presence of a homologous gene in the human genome. Our results suggest a divergence of photolyase genes in early evolution.

Key words: aplacental mammal/DNA repair/*Drosophila melanogaster*/photolyase/*Potorous tridactylis*

Introduction

In the early stages of evolution, life on earth suffered from high doses of UV light, making repair of UV-induced damage in DNA of vital importance. Photoreactivation is an efficient and direct repair mechanism for cyclobutane pyrimidine dimers, a major type of DNA damage induced by UV light. It is mediated by a single enzyme, photolyase, which splits dimers using visible light as the source of energy (see Sancar, 1990, for a review). Photoreactivating activity is widely distributed amongst species, ranging from bacteria to plants and mammals. The ubiquity and efficient use of visible light emitted by the sun, the same source that induces DNA lesions, indicate that photolyase is an ancient protein which may have played an important role in evolution.

Photolyase absorbs visible or near-UV light due to the presence of intrinsic chromophores. Several microbial photolyases have been purified to homogeneity, permitting identification of the chromophores. All characterized photolyases contain (reduced) FAD, which is the photochemically active chromophore, and a second chromophore acting as an energy harvesting antenna pigment which transduces the absorbed energy to the FAD cofactor. Two types of microbial photolyase are distinguished on the basis of the structure of the second chromophore, which largely determines the absorption and action spectrum. Photolyases from the Gram-negative bacteria *Escherichia coli* and *Salmonella typhimurium*, the yeast *Saccharomyces cerevisiae* and the fungus *Neurospora crassa* contain 5,10-methenyltetrahydrofolate (MTHF) and have an absorption maximum at ~380 nm (Sancar *et al.*, 1987; Johnson *et al.*, 1988; Li and Sancar, 1991; Eker *et al.*, 1994), while photolyases from the cyanobacterium *Anacystis nidulans*, the Gram-positive bacterium *Streptomyces griseus* and the archaebacterium *Halobacterium cutirubrum* contain 8-hydroxy-5-deazaflavin (8-HDF) with maximum activity at ~440 nm (Eker *et al.*, 1981, 1988, 1990). Photolyase-encoding genes have been cloned from the seven microorganisms mentioned above. The deduced amino acid sequences are clearly homologous to each other and only minor differences were found between the two types of photolyase (Yasui *et al.*, 1988; Yajima *et al.*, 1991).

Photoreactivating activity has also been demonstrated in a number of higher organisms including vertebrata (see Friedberg, 1984, for a review). While the presence of photolyase activity was unequivocally demonstrated in aplacental mammals (marsupialia; Sabourin and Ley, 1988), the results for placental mammals, including human, are contradictory, leaving doubts on the existence of active photolyase in these species (De Gruijl and Roza, 1991; Li *et al.*, 1993).

Recently, we succeeded in cloning a photolyase gene from the goldfish *Carassius auratus* (Yasuhira and Yasui, 1992), the first example of a photolyase gene from a multicellular organism. To our surprise, the deduced amino acid sequence differed significantly from both bacterial and lower eukaryote photolyase genes. To confirm the presence of this new type of photolyase in higher eukaryotes, we report here the cloning of photolyase genes from another fish species *Oryzias latipes*, the fruit fly

© Oxford University Press

Drosophila melanogaster and the marsupial *Potorous tridactylis* (rat kangaroo), as well as the characterization of this mammalian photolyase obtained by expression in *E.coli* cells. Furthermore, we demonstrate the existence of a homologous gene in an archaebacterium, which suggests the divergence of photolyase genes into two groups early in evolution.

Results

Cloning of Oryzias, Drosophila, Potorous and Methanobacterium photolyase genes

cDNA libraries from the killifish *O.latipes* and *D.melanogaster* expressed under UV and visible light selection in *E.coli* SY2 cells, which are deficient in three DNA repair pathways including photoreactivation, yielded transformants that showed increased survival on illumination with visible light after UV irradiation. Plasmids, isolated from transformants and reintroduced, provided SY2 cells with photoreactivating activity (Figure 1). For *O.latipes*, a 1.9 kb insert was found to contain a 1512 bp open reading frame (ORF; 504 amino acids, mol. wt 57 517; EMBL accession no. D26022). The *D.melanogaster* 2.1 kb insert appeared to encode a 1920 bp ORF (640 amino acids, mol. wt 73 818; EMBL accession no. D26021), while two additional potential start codons are present resulting in polypeptides of 555 (mol. wt 63 768) and 535 amino acids (mol. wt 61 381). Both truncated ORFs provided *E.coli* SY2 host cells with photoreactivating activity. The actual start codon, therefore, has still to be determined. Southern analysis (Figure 2A) confirmed the genomic origin of both cloned photolyase genes.

The amino acid sequences of the *O.latipes* and *D.melanogaster* genes (Figure 3) show a high similarity (72 and 58%, respectively) with the previously cloned photolyase gene from the goldfish *C.auratus* (Yasuhira and Yasui, 1992). The similarity between fish and insect photolyase genes was utilized to isolate a homologous gene from the marsupial *P.tridactylis* using a PCR strategy based on oligonucleotide primers derived from the most conserved regions. A 1865 bp fragment of a cloned DNA was sequenced, which contained a 1596 bp ORF (532 amino acids, mol. wt 61 630; EMBL accession no. D26020). This fragment, introduced in a pBluescript expression vector, converted *E.coli* SY2 cells into photoreactivation-proficient cells (Figure 1). Southern analysis (Figure 2A) confirmed the genomic origin of the cloned cDNA.

We found that the published N-terminal amino acid sequence of a photolyase purified from the archaebacterium *Methanobacterium thermoautotrophicum* (Kiener *et al.*, 1989) has no homology with the microbial photolyases, but is rather similar to metazoan photolyases (Figure 4). Using PCR we succeeded in cloning a part of the methanobacterial photolyase gene and then obtained a genomic fragment of 8 kb with a 1332 bp ORF (444 amino acids, mol. wt 51 761; EMBL accession no. D30752). The deduced amino acid sequence (Figure 3) has a clear homology to the metazoan photolyases and is similar to the N-terminus of *M.thermoautotrophicum* photolyase (Figure 4).

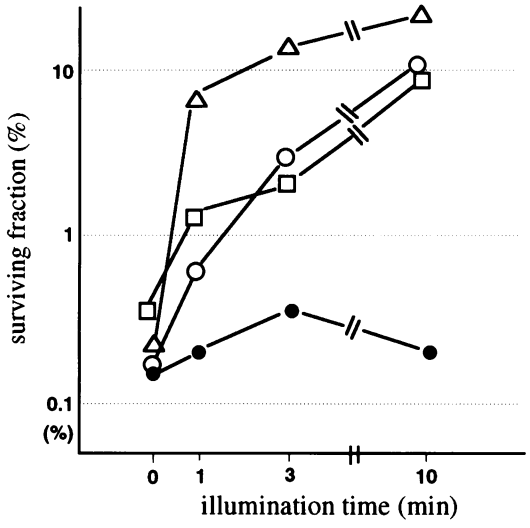

Fig. 1. Photoreactivating activity of cloned photolyase genes from higher eukaryotes expressed in *E.coli* SY2. The increase of colony-forming ability due to illumination with visible light of UV-irradiated cells harbouring photolyase genes of the fruit fly *D.melanogaster* (△), the killifish *O.latipes* (○) and the marsupial *P.tridactylis* (□) inserted in pBluescript or the vector plasmid alone (●) is shown.

Search for a photolyase gene in the human genome

The widespread distribution of photolyases amongst higher eukaryotes including marsupials prompted us to start a search for a human homologue. A *P.tridactylis* photolyase cDNA probe, however, did not hybridize to human genomic DNA under the conditions used (Figure 2B, lanes 3 and 5), whereas with *O.latipes* DNA a clear signal was obtained (Figure 2B, lanes 2 and 4). Attempts to amplify a human photolyase gene by PCR using primers PR4 and PR5 (Figure 3) were also unsuccessful.

Comparison of photolyase sequences

In Figure 3 the deduced amino acid sequences of the complete ORFs of cloned photolyase genes from *O.latipes*, *D.melanogaster*, *P.tridactylis* and *M.thermoautotrophicum* ΔH are compared with those from *C.auratus* and several microorganisms. Although the amino acid sequences of the newly cloned photolyases differ considerably from those of microorganisms, they can be aligned. The large differences in amino acid sequences indicate the existence of a new class of photolyases in addition to the microbial photolyases.

Besides the general difference in amino acid sequences the absence of two conserved tryptophan residues in the new class of photolyases is remarkable since they are thought to be important for DNA binding (W277, see Figure 3) and essential as hydrogen donors in the photoreduction of FAD semiquinone radical (W306; Li and Sancar, 1990; Li *et al.*, 1991).

Boxes I and P in Figure 3 indicate amino acid sequences specific for, respectively, MTHF- or 8-HDF-type microbial photolyases (Yasui, 1993). An I-P(Q) sequence (box I) is conserved in MTHF-type, but is also present in metazoan photolyases. A P-X-L-K(H)-F sequence (box P) is specific for 8-HDF-type photolyases, but an apparently similar sequence of P-W-L(I,F)-H(R)-F(T) is found in metazoan photolyases. Both boxes are also present in the methano-

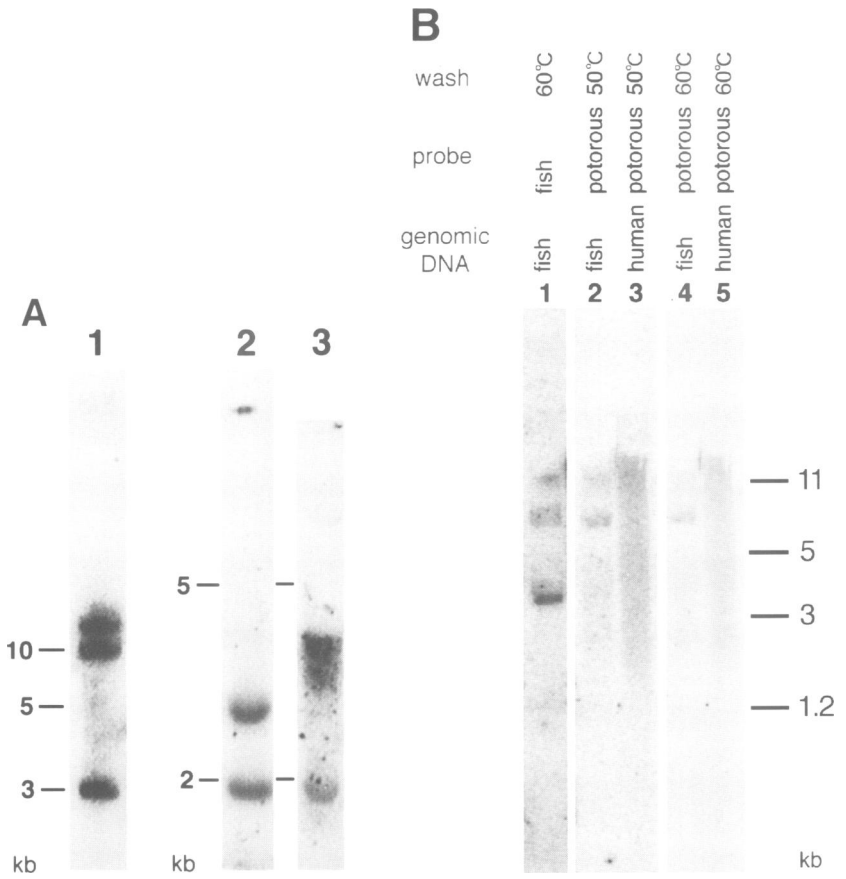

Fig. 2. (**A**) Southern analysis of cloned photolyase genes. Genomic DNA (15 μg), prepared from *D.melanogaster* DM2 (lane 1), *O.latipes* OL32 (lane 2) or *P.tridactylis* PtK-2 (lane 3) cells was digested with *Eco*RI (DM2 and PtK-2) or *Eco*RI plus *Hin*dIII (OL32) and hybridized at 65°C with cDNAs containing the whole coding sequence of cloned photolyase genes. (**B**) Southern analysis of human DNA. Genomic DNA from HeLa (lanes 3 and 5) or *O.latipes* (lanes 2 and 4) cells was digested with *Eco*RI and hybridized with cDNA encoding the *P.tridactylis* photolyase gene. Filters were washed at 50 (lanes 2 and 3) and 60°C (lanes 4 and 5). As a positive control the *O.latipes* probe was hybridized to its genomic DNA and washed at 60°C (lane 1). The sizes of the fragments in kb are indicated.

bacterial photolyase. Therefore, these photolyases may share some characteristics with both 8-HDF- and MTHF-type photolyases, although the actual type of chromophore cannot be predicted yet.

Eukaryotic photolyases possess a protruding N-terminus in which two regions with clustered positively charged amino acid residues (boxes A and C) may contain signal sequences for nuclear location. Box A is present in both lower and higher eukaryote photolyases, while box C is conserved in higher eukaryote and partly in methanobacterial photolyases. The sequences in box B were proposed to be involved in the transport of photolyase into mitochondria in yeast and fungi (Yasui *et al.*, 1992). This sequence is not found in photolyases of higher eukaryotes.

Purification of P.tridactylis photolyase as GST fusion protein

For functional studies, *P.tridactylis* photolyase was overproduced as glutathione S-transferase (GST) fusion protein in *E.coli* cells and purified by consecutive chromatography on glutathione−Sepharose, heparin−Sepharose and phospho-ultrogel columns. The final preparation was estimated to be >98% pure from SDS−PAGE (Figure 5A), yielding a single band of 81 kDa molecular weight.

The purified protein was biologically active, showing light-dependent catalytic splitting of thymine dimers in the UV-pT_{20} assay (Figure 6). The fusion protein can therefore be considered as a true photolyase. Incubation with thrombin led to complete cleavage of the fusion protein (Figure 5B), yielding a 26 kDa GST band and a 58 kDa band which can be assigned to the photolyase part, comparing well with 61.6 kDa calculated from the deduced amino acid sequence. Removal of the GST part increased the photoreactivating activity by 30–40% (Figure 6).

Identification of P.tridactylis photolyase chromophore

The absorption spectra obtained for successive purification steps of GST−*Potorous* photolyase were different to some extent, but all showed a maximum at 583 and a shoulder at 620 nm, characteristic for a neutral (blue) flavin semiquinone radical (Figure 7). On denaturation under acidic conditions the absorption spectrum changed considerably (Figure 8). The long wavelength band disappeared and the resulting spectrum closely resembles a (oxidized) flavin absorption spectrum. Comparable results were obtained on denaturation with SDS at neutral pH (data not shown). Denatured photolyase showed a characteristic flavin fluorescence spectrum with maxima at 445 and 522 nm for excitation and emission spectrum, respectively

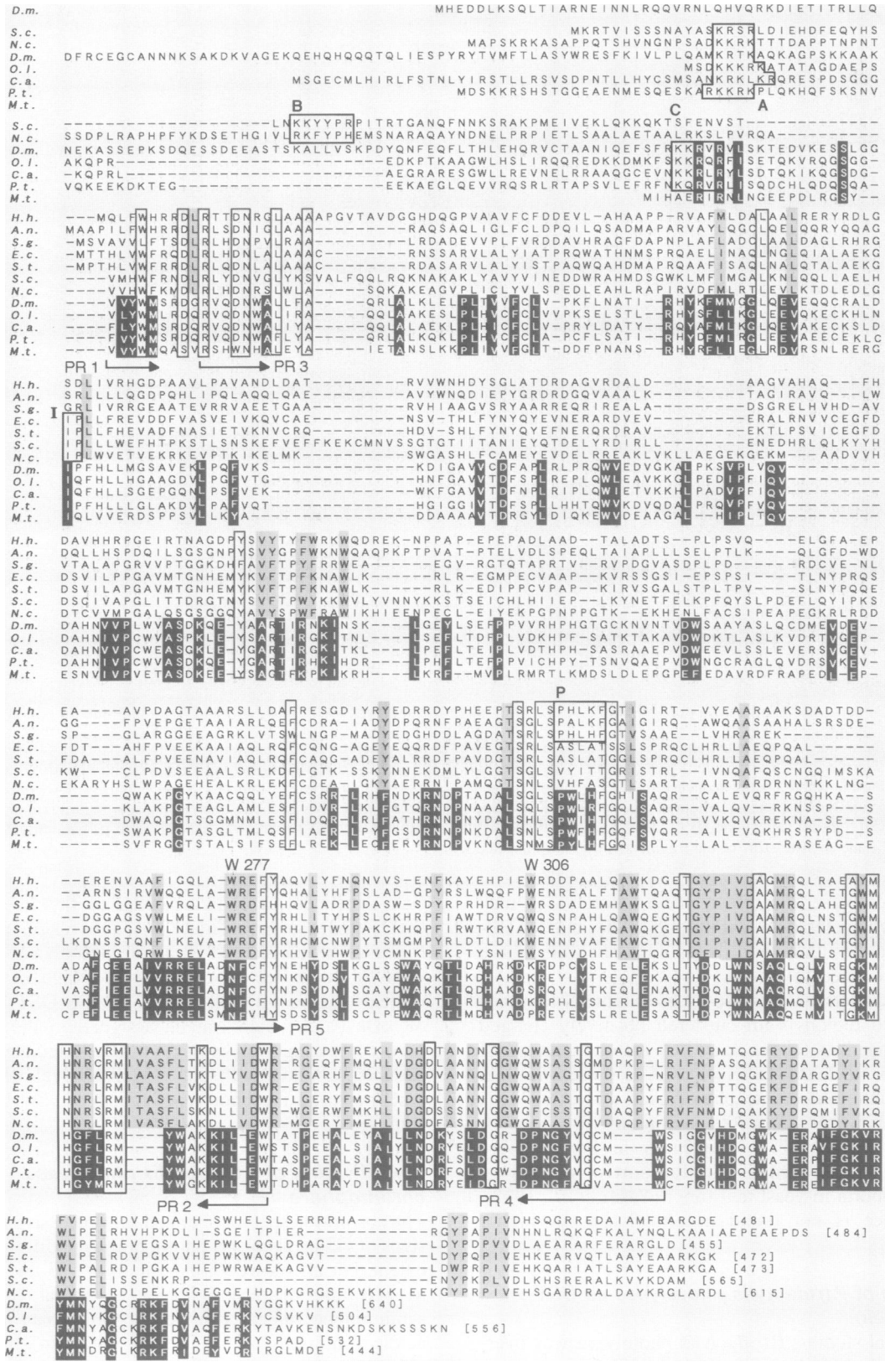

Fig. 3. Aligned amino acid sequences of photolyases. Abbreviations of species are: H.h., *H.halobium*; A.n., *A.nidulans*; S.g., *S.griseus*; E.c., *E.coli*; S.t., *S.typhimurium*; S.c., *S.cerevisiae*; N.c., *N.crassa*; D.m., *D.melanogaster*; O.l., *O.latipes*; C.a., *C.auratus*; and P.t., *P.tridactylis*. The upper three photolyases (H.h, S.g. and A.n.) contain 8-HDF and the next four (E.c., S.t., S.c. and N.c.) contain MTHF as the second chromophore. Amino acids conserved in almost all (at least 10 out of 13) photolyases are boxed, those conserved only in microorganisms (class I photolyases) are on a shadowed background and those conserved in higher eukaryotes (class II photolyases) are on a black background. PR1−PR5 indicate amino acid sequences used for PCR primers. Boxes A, B, C, I and P, as well as amino acid residues W277 and W306, are explained in the text.

```
          1                                   10
Marburg   : Met Ile His Asp Glu Arg Ile Arg Ser Leu Asn Thr Glu Lys
ΔH        : Met Ile His Ala Glu Arg Ile Arg Asn Leu Asn Gly Glu Glu

higher                        Arg Val Arg     Leu
eukaryotes


                              20
Marburg   : Pro Ala Arg Asp Gly Lys Tyr Val Ile Tyr Trp Met Gln Ala
ΔH        : Pro Asp Leu Arg Gly Ser Tyr Val Val Tyr Trp Met Gln Ala

higher                        Ser     Val Val Tyr Trp Met
eukaryotes                            Phe Leu
```

Fig. 4. Comparison of N-terminal amino acid sequences of *Methanobacterium* photolyases. The sequence of photolyase from *M.thermoautotrophicum* strain Marburg, obtained by amino acid sequencing (Kiener *et al.*, 1989), is compared with the deduced sequence of *M.thermoautotrophicum* strain ΔH photolyase as well as with the consensus sequence of photolyases from higher eukaryotes.

(data not shown). The fluorescence intensity increased 8.9-fold after incubation with snake venom phosphodiesterase, identifying the flavin as FAD. From denaturation experiments it was calculated that the final GST–photolyase preparation contained 0.57 mol FAD/mol protein and obviously is partly depleted in FAD.

The change in absorption spectra during purification (Figure 7) is apparently the result of a slow oxidation of enzyme-bound reduced FAD. During elution of the first (glutathione–Sepharose) column, GST–*Potorous* photolyase changed from colourless to blue. We conclude, therefore, that *in vivo* it probably contains fully reduced FAD, but in purified photolyase FAD is mainly in the oxidized form. The fine structure of the 450 nm absorption band, with discrete maxima at 443 and 467 nm, suggests a hydrophobic environment of photolyase-bound FAD, as found for *N.crassa* photolyase (Eker *et al.*, 1994).

The calculated absorption difference spectrum of denatured photolyase minus FAD (Figure 8) shows end absorption only, without recognizable absorption bands in the near-UV/visible region, indicating the absence of chromophoric compounds other than FAD. Attempts were made to reconstitute GST–*Potorous* photolyase with possible cofactors and related compounds. The addition of FAD increased photoreactivating activity by 27%, but incubation with known second chromophores of microbial photolyases (or related compounds), like MTHF, 8-hydroxy-5-deazariboflavin (F0) or SF420 [F0 with phosphate-lactyl-(glutamyl)$_n$ side chain], did not increase biological activity. Cleavage of the fusion protein by thrombin did not improve reconstitution.

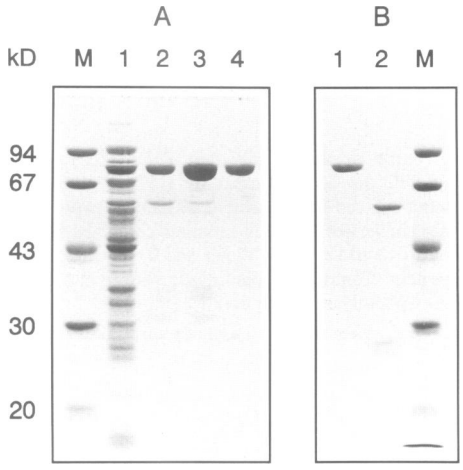

Fig. 5. (**A**) Purification of GST–*Potorous* photolyase. SDS–PAGE of cell extract (lane 1) and eluates of glutathione–Sepharose (lane 2), heparin–Sepharose (lane 3) and phospho-ultrogel (lane 4) columns. (**B**) SDS–PAGE of untreated (lane 1) and thrombin-cleaved (lane 2) GST–*Potorous* photolyase. Lanes M contain marker proteins. Gels are stained with Coomassie brilliant blue.

Discussion

Classification of photolyases

We have demonstrated by *in vivo* (Figure 1) and *in vitro* (Figure 6) experiments the existence of photolyase genes in higher eukaryotes. Table I shows the amino acid sequence similarities of photolyases from 13 species. Photolyases are currently divided into two types according to the structure of the second chromophore and spectral properties. These types of photolyase show a high degree of sequence homology (Figure 3). The work presented here unequivocally establishes the existence of a distinct class of photolyases with very little sequence resemblance to the previously known photolyases. Based on these findings we propose therefore an extended classification, as indicated in Table I, which incorporates the previously known types of photolyase (in class I; Table II).

The microbial photolyases in classes I and II show 25–43 and 38–72% mutual homology, respectively. Between the two classes a much lower homology of 10–17% is found, indicating a rather distant relationship. A recently cloned photolyase gene from the plant *Sinapis alba* (white mustard; Batschauer, 1993) is more closely related to class I (25–28% homology) than to class II (12–16%) photolyases, and obviously falls into class I.

The deduced amino acid sequences of 8-HDF- and MTHF-type microbial photolyases are highly homologous to each other (Yajima *et al.*, 1991). Only a few amino acid residues can be specifically assigned to one of these types, as indicated in Figure 3 (boxes I and P). The binding site for the second chromophore is quite specific since *A.nidulans* photolyase expressed in *E.coli*, containing only a reduced FAD cofactor with a vacant 8-HDF pocket (Takao *et al.*, 1989a), reconstitutes with 8-HDF but not with MTHF (Eker and Yasui, 1991; Miki *et al.*, 1993), while vice versa *E.coli* photolyase depleted from MTHF chromophore cannot be reconstituted with 8-HDF (Wang and Jorns, 1989; Payne *et al.*, 1990). Surprisingly, the

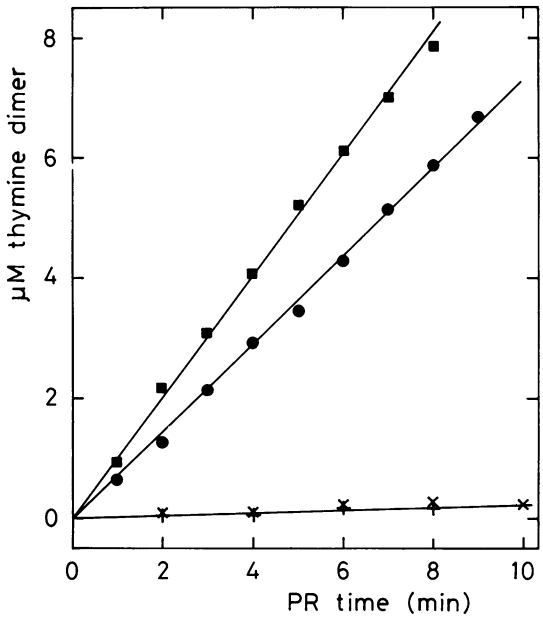

Fig. 6. Time courses of photoreactivation. Photoreactivating activity was measured with the UV-pT_{20} assay for GST–*Potorous* photolyase (●) or thrombin-cleaved fusion protein (■) and is expressed as µM thymine dimer converted. Reaction mixtures contained UV-pT_{20} corresponding to 18 µM thymine dimer and 0.6 µM (GST–) photolyase protein. Controls: no photolyase with light (×) and GST–*Potorous*-photolyase in the dark (+).

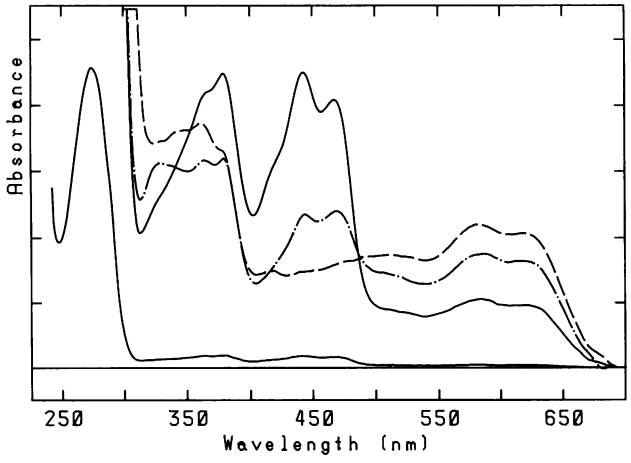

Fig. 7. Absorption spectra of different purification stages of GST–*Potorous* photolyase after glutathione–Sepharose (– – –), heparin–Sepharose (– · –) and phospho-ultrogel (—) chromatography. The latter is also shown at 20-fold reduced sensitivity. The spectrum of glutathione–Sepharose eluate is corrected for scatter by λ^n correction (Jagger, 1967).

amino acid sequences of class II photolyases share characteristics with both 8-HDF- and MTHF-type photolyases, but from sequence comparison it cannot be decided whether they are MTHF- or 8-HDF-type.

Absence of second chromophore in P.tridactylis photolyase

P.tridactylis photolyase was purified until apparent homogeneity as GST fusion protein from *E.coli* cells. Since no other chromophore could be detected, we conclude that reduced FAD is the only active chromophore of *P.tridactylis* photolyase. This was confirmed by the increase of photoreactivating activity on reconstitution

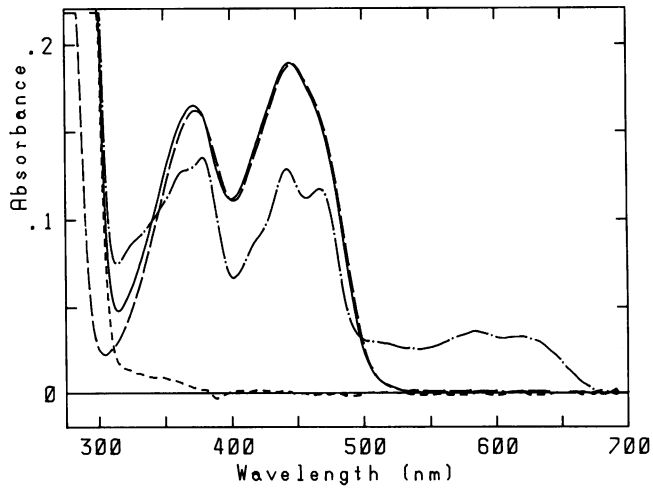

Fig. 8. Absorption spectra of native (– · –) and acid-denatured (—) GST–*Potorous* photolyase, compared with the absorption spectrum of FAD at pH 2 (– –). Also shown is the difference spectrum obtained by subtracting the FAD from the denatured photolyase spectrum (----).

with FAD. This finding unifies the basic action mechanism of both classes of photolyase, as in all characterized microbial photolyases the presence of (reduced) FAD has been demonstrated. Reconstitution of *P.tridactylis* photolyase with known second chromophores from microbial photolyases was unsuccessful. An obvious explanation is that the correct chromophore is not present in *E.coli* cells used for overexpression, as was found for *A.nidulans* photolyase (Takao *et al.*, 1989a), but other explanations are possible. Synthetic MTHF used for reconstitution contained a single glutamate residue, while photolyases from yeast and *E.coli* are known to contain MTHF with a polyglutamate tail (Johnson *et al.*, 1988). Alternatively, *P.tridactylis* photolyase might lack the binding site for a second chromophore: our experiments (Figure 6) indicate that the presence of FAD alone is sufficient for enzymatic activity. An action spectrum for photoreactivation of *P.tridactylis* cells shows a maximum at ~350 nm (Ishizaki and Takebe, 1985), which is not contradictory to the presence of a single reduced FAD chromophore. Identification of a possible second chromophore must await further studies using photolyases obtained directly from higher eukaryotes, which are now under way.

The biological activity of GST–*Potorous* photolyase increased on cleavage with thrombin. Some (weak) interaction between GST and photolyase may cause steric hindrance in substrate binding. From kinetic analysis of the UV-pT_{20} assay (Eker *et al.*, 1994) we estimate that the quantum yield increases on removal of the GST part, possibly due to a more efficient electron/energy transfer from reduced FAD to the dimer substrate.

Photolyase in placental mammals

The existence of a mammalian photolyase, demonstrated by cloning of the gene from an established marsupial cell line and analysis of the encoded protein, confirms previous reports describing photoreactivating activity in this cell line (Cook and Regan, 1969; Buhl *et al.*, 1974; Ishizaki and Takebe, 1985). The cloned marsupial cDNA, however, did not hybridize to human genomic DNA under the conditions used. Neither could we amplify a human

Table I. Similarity of amino acid sequences of photolyases

	A.n.	*S.g.*	*S.t.*	*E.c.*	*S.c.*	*N.c.*	*S.a.*	*D.m.*	*C.a.*	*O.l.*	*P.t.*	*M.t.*
H.h.	43	38	37	36	29	37	28	14	13	13	12	14
A.n.		31	38	39	31	36	27	15	13	15	13	14
S.g.			32	33	27	32	27	17	13	14	15	17
S.t.				79	33	41	27	15	14	14	12	12
E.c.		class I			34	39	27	14	13	13	12	14
S.c.						35	25	11	10	10	11	13
N.c.							26	12	13	13	14	14
S.a.								12	15	13	13	16
D.m.									58	58	61	37
C.a.										72	66	36
O.l.								class II			65	38
P.t.												38

Abbreviations are the same as in the legend to Figure 3 except that *S.alba* (S.a., white mustard) is also included. Similarities of photolyases are given as the percentage of identical amino acid residues after alignment of the sequences, without the protruding N-terminus in eukaryote photolyases. The partition into two classes is indicated.

Table II. DNA photolyases

	Class I (microbial photolyases)	Class II (higher eukaryote photolyases)
Occurrence	lower eukaryotes halobacteria eubacteria	higher eukaryotes methanobacteria eubacteria
Second chromophore	type 1: MTHF type 2: 8-HDF	?

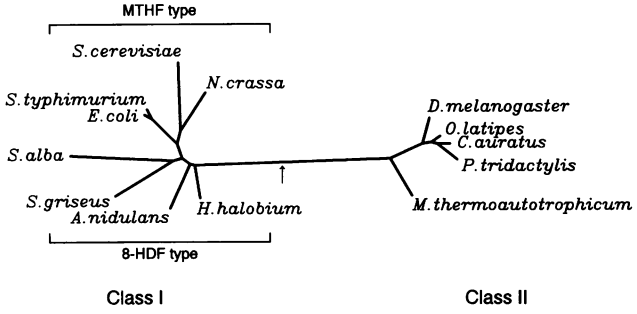

Fig. 9. Unrooted phylogenetic tree obtained by the neighbour-joining method (PHYLIP software package, Felsenstein, 1993) using pairwise distances calculated from photolyase amino acid sequences with the Dayhoff PAM 001 matrix. A possible root (obtained with the UPGMA method) is indicated by an arrow.

homologue by PCR. This is in accordance with the reported absence of photoreactivating activity in human cell extracts (Chao, 1993; Li *et al.*, 1993). It may indicate that if there exists a human gene related to the photolyase genes of higher eukaryotes, its sequence (and possibly also function) may be different from photolyases splitting cyclobutane-type pyrimidine dimers in UV-irradiated DNA. It has been reported that photolyase enhances excision repair in *E.coli* (Yamamoto *et al.*, 1983; Sancar *et al.*, 1984a) and it is conceivable that during evolution this second function became more important, culminating in a complete loss of photoreactivating activity in placental mammals. Another possible second function for photolyase has been reported recently (Ahmad and Cashmore, 1993). A blue light response (inhibition of hypocotyl elongation) in the plant *Arabidopsis thaliana* is mediated by a photolyase-like protein. Although no auxiliary function is known for photolyase in higher eukaryotes, a large change in the sequence may be the consequence of the development of additional functions.

Phylogenetic implications

The photolyase gene we obtained from the archaebacterium *M.thermoautotrophicum* encodes a protein which is ~37% homologous to higher eukaryote and only 15% to microbial photolyases (Table I). This and the amino acid sequence alignment (Figure 3) clearly classifies the methanobacterial photolyase as a class II photolyase. The difference in amino acid sequence of photolyases from *M.thermoautotrophicum* strain Marburg (Kiener *et al.*, 1989) and ΔH (this work, Figure 4) could be attributed to the relatively large difference between these strains, which should possibly be classified as different species (Touzel *et al.*, 1992; Nölling *et al.*, 1993). Moreover, a class II photolyase gene has also been found in the bacterium *Myxococcus xanthus* (K.O'Connor, M.McBride and D.Zusman, unpublished results). Therefore it appears that class II photolyases occur in all three primary kingdoms (domains) of life, as was found for class I photolyases.

A phylogenetic tree obtained from amino acid sequence comparison (Figure 9) shows the clear separation of the two classes, as well as the subdivision of microbial photolyases according to 8-HDF and MTHF second chromophores. Together with the presence of reduced FAD as intrinsic chromophore in both classes of photolyase, this suggests that the two classes originate from an ur-photolyase by gene duplication in early evolution, followed by a divergent evolution yielding the photolyases found in extant organisms.

Materials and methods

E.coli strains

E.coli strain SY2 (JM107 *phr::Cmr uvrA::Kmr recA::Tetr*; Yasuhira and Yasui, 1992), deficient in excision repair, recombination and photoreactivation, was used as host cell for the expression and selection of cDNA libraries as well as for the analysis of photoreactivating potency induced by cloned genes. *E.coli* strains DH10B (BRL, USA) and XL1-Blue (Clonetech) were used for the construction of expression libraries of *O.latipes* and *D.melanogaster*, respectively.

Selection of E.coli clones expressing a foreign photolyase gene

Samples (100 µl) of *E.coli* SY2 harbouring cDNA libraries, grown overnight in LB medium supplemented with 50 µg/ml ampicillin, 30 µg/ml chloramphenicol, 50 µg/ml kanamycin and 12.5 µg/ml tetracycline, were plated on LB agar containing four antibiotics and irradiated with 0.12 J/m^2 UV light (254 nm, Hitachi germicidal lamp GL10, dose rate 0.01 J/m^{2}s). Irradiation was followed immediately by illumination with white fluorescent lamps (Toshiba FL20SS) for 5 min. To eliminate short wavelength light as well as heat produced by the lamps, a glass plate and a transparent plastic box filled with water were inserted between white lamps and Petri dishes. After incubation at 37°C for 12 h, surviving colonies were recovered and subjected to the next round of UV and visible light treatment. After the third selection round, plasmids from surviving colonies were isolated and analysed.

Cloning of O.latipes and D.melanogaster photolyase genes

Poly(A)$^+$ RNA was isolated from 1×10^8 confluently grown cells of the established cell line OL32 from *O.latipes* (Komura *et al.*, 1988) with a Fast-Track mRNA isolation kit (Invitrogen). Double-stranded cDNA was synthesized from 2–5 µg of poly(A)$^+$ RNA using cDNA Synthesis Plus (Amersham). cDNA was purified by agarose gel electrophoresis, inserted into pBluescript II (Stratagene) and introduced in DH10B cells by electroporation. Plasmid DNA was isolated from transformants and used as an *O.latipes* cDNA expression library. A cDNA expression library of *D.melanogaster* was constructed by *in vivo* excision of a lambda library in Uni-ZAPXR (Stratagene) made from mRNA of 2–14 h old insect embryos of the Canton-S strain. After three rounds of selection with UV and visible light, surviving clones were analysed, yielding a 1.9 kb insert for *O.latipes* and a 2.1 kb insert for *D.melanogaster*.

PCR for cloning of P.tridactylis photolyase gene

The following oligonucleotides derived from highly conserved amino acid sequences (indicated in brackets) of *C.auratus*, *O.latipes* and *D.melanogaster* photolyase (Figure 3) were used as PCR primers: PR3, 5′-GIGTICA(A/G)GA(C/T)AA(C/T)TGGGCI(C/T)T-3′ [RVQDNWAL]; PR4, 5′-CCACAT(A/G)CAICCIAC(A/G)(A/T)ATCC(A/G)TTIGG(A/G)TC-3′ [DPNG(F/Y)VGCMW]; and PR5, 5′-GA(C/T)AA(C/T)TT(C/T)TG(C/T)TT(C/T)TA(C/T)AA(C/T)-3′ [DNFCFYN]. (I denotes inosine; degenerated sequences are shown in parentheses and the corresponding amino acid sequence between brackets.)

From 1×10^6 PtK-2 cells, an established cell line of the marsupial *P.tridactylis* (Walen and Brown, 1962), 2 µg poly(A)$^+$ RNA were isolated and converted into cDNA with reverse transcriptase (Life Science). DNA was amplified in 35 PCR cycles of 1 min 94°C, 1 min 50°C and 1 min 72°C. Amplified DNA was purified by agarose gel electrophoresis (Gene Clean II, Bio101) and cloned into pCR-Script (Stratagene) for sequence determination. PCR with primers PR4 and PR5 yielded a 335 bp DNA fragment. Subsequent PCR with PR3 and a primer derived from the 336 bp fragment yielded a 1.1 kb DNA fragment. Additional sequences at the 5′ and 3′ ends were obtained by RACE amplification of cDNA ends (Frohman *et al.*, 1988). Using primers around the putative start and stop codons, cDNA from PtK-2 cells was once more amplified, yielding a 1.6 kb fragment.

Cloning of a M.thermoautotrophicum photolyase gene by PCR

Oligonucleotides derived from the N-terminal amino acid sequence of *M.thermoautotrophicum* strain Marburg photolyase (Kiener *et al.*, 1989) {PR1, 5′-GGIAA(A/G)TA(T/C)GTI(A/C)TITA(T/C)TGGATG-3′ [GKYV(L/I)YWM]} and from conserved parts of photolyases from higher eukaryotes [PR2, 5′-CCA(C/T)TCIA(A/G)(A/G/T/)AT(C/T)-TT(C/T)TT-3′ (KKILEW)], as well as PR4 (Figure 3), were used as PCR primers. DNA from *M.thermoautotrophicum* strain ΔH (DSM1053) was prepared by sonication of cells, phenol extraction and ethanol precipitation. An amplified fragment was used to isolate the whole photolyase gene on an 8 kb insert by screening a DNA library made by complete digestion with *Hin*dIII and ligation into pUC118.

Photoreactivating activity of E.coli cells expressing cloned photolyase genes

E.coli SY2 cells harbouring cloned photolyase genes of *O.latipes*, *D.melanogaster* and *P.tridactylis*, inserted behind a *lac* promoter in the multicloning site of the vector plasmid pBluescript, were incubated overnight in LB medium containing four antibiotics. Cells were diluted with M9 buffer (50 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NH_4Cl) to 1×10^6 cells/ml; 3 ml were irradiated in a Petri dish with 0.2 J/m^2 UV light, followed by illumination with fluorescent lamps as described above. Samples were plated on LB agar containing four antibiotics; after 18 h colonies were counted. All irradiation/illumination experiments were performed under yellow light.

Determination and analysis of nucleotide sequences of photolyase genes

Nucleotide sequences of isolated cDNA fragments were determined for both DNA strands by the dideoxy method (Sanger *et al.*, 1977) using Sequenase Version 2.0 (US Biochemicals). Various deletions of the cloned genes in pBluescript II vector, created by incubation with mung bean nuclease, were isolated from agarose gel after electrophoresis, religated in the same vector and used for the production of single-stranded DNA with helper phage for sequence determination. Additionally, a number of synthesized oligonucleotides were used to determine and confirm the sequence. Nucleotide and deduced amino acid sequences were analysed with the ODEN software package (Ina, 1994) available on niguts@ddbj.nig.ac.jp (National Institute of Genetics, Japan). Amino acid sequences of photolyases were deduced from the published nucleotide sequences for *E.coli* (Sancar *et al.*, 1984b), *S.cerevisiae* (Sancar, 1985; Yasui and Langeveld, 1985), *A.nidulans* (Yasui *et al.*, 1988), *S.griseus* (Kobayashi *et al.*, 1989), *Halobacterium halobium* (Takao *et al.*, 1989b), *N.crassa* (Yajima *et al.*, 1991), *S.typhimurium* (Li and Sancar, 1991), *C.auratus* (Yasuhira and Yasui, 1992) and *S.alba* (Batschauer, 1993).

Isolation of genomic DNA and Southern hybridization

Genomic DNA (10–20 µg), isolated from confluent cultures of *O.latipes* OL32 cells, *D.melanogaster* Dm2 and *P.tridactylis* PtK-2 cells, was digested with various restriction enzymes. After electrophoresis, DNA was transferred to nylon membranes (Hybond-N$^+$, Amersham). Hybridization was performed in Rapid-hyb buffer (Amersham) at 65°C with labelled DNA probes containing whole coding sequences using the Random Primer Extension Labeling System (NEN). Membranes were washed twice in 2× SSPE (1× SSPE; 180 mM NaCl, 10 mM sodium phosphate, pH 7.4, 1 mM Na_2EDTA), 0.1% (w/v) SDS for 10 min at room temperature, once in 1× SSPE, 0.1% SDS for 65°C and twice in 0.7× SSPE, 0.1% SDS for 15 min at 65°C.

For the analysis of human DNA, genomic DNA (10 µg) from HeLa cells or from *O.latipes* (5 µg) was digested with *Eco*RI, blotted to a nylon membrane and hybridized with the *P.tridactylis* probe in 25% formamide, 6× SSC, 10% dextran sulfate, 10× Denhardt and 50 µg/ml salmon DNA at 65°C overnight. The membrane was washed in 0.5× SSC and 0.1% SDS at 50°C for 1 h, followed by incubation at 60°C. As a positive control, *O.latipes* DNA was hybridized to the *O.latipes* probe and washed at 60°C. Hybridized membranes were analysed with a Fujix BAS2000 (Fuji Photo Film) image analyser.

Purification of a GST–Potorous photolyase fusion protein

cDNA of the cloned *P.tridactylis* photolyase gene in pBluescript was amplified with two 27 base PCR primers around the putative start and stop codons containing a *Bcl*I site at the 5′ end of the sequences. After digestion with restriction enzyme and purification by agarose gel electrophoresis, the DNA fragment was introduced at the *Bam*HI site of pGEX-2T (Pharmacia, Sweden) behind the glutathione S-transferase gene yielding pGEX–Ptphr1. This fusion gene provided *E.coli* SY2 cells with photoreactivating activity.

E.coli JM109 transformed with pGEX–Ptphr1 was grown at 37°C in medium containing per litre: 10 g tryptone, 5 g yeast extract, 7 g Na_2HPO_4, 0.5 g KH_2PO_4 and 4 g glucose supplemented with 200 mg/l ampicillin, until an A_{600} of 0.9–1.0 was reached. Expression was induced by the addition of 15 mg/l IPTG, and growth was continued at 28°C for 3 h. During cultivation, pH was maintained at 7.1 by the addition of Na_3PO_4 solution. Cell extract was prepared by sonication of 36 g cells suspended in 60 ml 100 mM NaCl, 10 mM K-phosphate, 10 mM 2-mercaptoethanol, 1 mM PMSF pH 7.0, followed by centrifugation (twice) at 43 000 *g* for 40 min at 4°C. GST–photolyase fusion protein was purified from the supernatant by consecutive column chromatography on glutathione–Sepharose 4B (Pharmacia, 1.0×10.2 cm, flow rate 10 ml/h, elution with 15 mM glutathione in 50 mM Tris–HCl, final pH 7.9), heparin–Sepharose CL-6B (Pharmacia, 1.0×7.5 cm, flow rate 16 ml/h, linear gradient 0.1–0.75 M NaCl in 10 mM K-phosphate, pH 7.0, elution at 0.4 M NaCl) and phospho-ultrogel A6B (IBF Biotechnics, 0.66×11.50 cm, flow rate 10.5 ml/h, linear gradient 0.1–0.5 M NaCl in 10 mM K-phosphate, pH 7.0, elution at 0.32 M NaCl). Fractions were selected according to their absorption spectrum and

biological activity. All solutions contained 10 mM 2-mercaptoethanol. The final preparation was stored frozen at −70°C after the addition of 15% (v/v) glycerol. Fusion protein was cleaved with thrombin (Sigma, from human plasma) by incubation overnight at 4°C.

Characterization of GST–Potorous photolyase

Photoreactivating activity was measured with the UV-pT_{20} assay. Reaction mixtures, containing UV-irradiated (254 nm) pT_{20} corresponding to ~18 μM thymine dimer, photolyase and buffer (100 mM NaCl, 10 mM K-phosphate, 10 mM 2-mercaptoethanol, pH 7.0) until a final volume of 220 μl, were incubated in a thermostatted (35°C) quartz micro-cuvette for 10 min in the dark prior to illumination with white light from a Xe-arc (Osram HBO 150 W). Low wavelength radiation was blocked with window glass and plexiglas filters. At regular time intervals the increase of A_{265} was measured, which was converted into the amount of split thymine dimer using $\varepsilon = 8790$ for thymine in oligo(dT) (Eker *et al.*, 1994). Protein concentrations were determined with the Coomassie brilliant blue method (Bradford, 1976).

Absorption spectra were measured with an Ultrospec II (Pharmacia/LKB) spectrophotometer using a scanning software package. Fluorescence measurements were performed with an LS50B (Perkin Elmer) luminescence spectrometer.

For reconstitution experiments, photolyase was incubated in the dark with an excess of FAD (Merck), F0 (Eker *et al.*, 1989), natural 8-HDF cofactor SF420 from *S.griseus* (Eker *et al.*, 1980) or MTHF (Eker *et al.*, 1994) prior to the UV-pT_{20} assay. For MTHF a correction was made for the conversion into UV-absorbing product at neutral pH by measuring the absorbance increase in the dark.

Acknowledgements

The technical assistance of Miss Junko Kikuchi and Miss Izumi Chiba is acknowledged. We would like to thank Drs J.H.J.Hoeijmakers and D.Bootsma for valuable discussions and continuous support, Y.Koga for *M.thermoautotrophicum* cells and K.O'Connor for sharing *M.xanthus* sequence information before publication. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas from The Ministry of Education, Science and Culture of Japan (no. 06270101) to A.Y.

References

Ahmad,M. and Cashmore,A.R. (1993) *Nature*, **366**, 162–166.
Batschauer,A. (1993) *Plant J.*, **4**, 705–709.
Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
Buhl,S.N., Setlow,R.B. and Regan,J.D. (1974) *Biophys. J.*, **14**, 791–803.
Chao,C.C.-K. (1993) *FEBS Lett.*, **336**, 411–416.
Cook,J.S. and Regan,J.D. (1969) *Nature*, **223**, 1066–1067.
De Gruijl,F.R. and Roza,L. (1991) *J. Photochem. Photobiol. B: Biol.*, **10**, 367–371.
Eker,A.P.M. and Yasui,A. (1991) *Photochem. Photobiol.*, **53**, 17S.
Eker,A.P.M., Pol,A., Van der Meyden,P. and Vogels,G.D. (1980) *FEMS Microbiol. Lett.*, **8**, 161–165.
Eker,A.P.M., Dekker,R.H. and Berends,W. (1981) *Photochem. Photobiol.*, **33**, 65–72.
Eker,A.P.M., Hessels,J.K.C. and Van de Velde,J. (1988) *Biochemistry*, **27**, 1758–1765.
Eker,A.P.M., Hessels,J.K.C. and Meerwaldt,R. (1989) *Biochim. Biophys. Acta*, **990**, 80–86.
Eker,A.P.M., Kooiman,P., Hessels,J.K.C. and Yasui,A. (1990) *J. Biol. Chem.*, **265**, 8009–8015.
Eker,A.P.M., Yajima,H. and Yasui,A. (1994) *Photochem. Photobiol.*, **60**, 125–133.
Felsenstein,J. (1993) *PHYLIP*. Department of Genetics, University of Washington, Seattle, WA, version 3.5c.
Friedberg,E.C. (1984) *DNA Repair*. W.H.Freeman and Co., New York, p. 35.
Frohman,M.A., Dush,M.K. and Martin,G.R. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 8998–9002.
Ina,Y. (1994) *Comput. Appl. Biosci.*, **10**, 11–12.
Ishizaki,K. and Takebe,H. (1985) *Mutat. Res.*, **150**, 91–97.
Jagger,J. (1967) *Introduction to Research in Ultraviolet Photobiology*. Prentice-Hall, Englewood Cliffs, NJ, 53 pp.
Johnson,J.L., Hamm-Alvarez,S., Payne,G., Sancar,G.B., Rajagopalan, K.V. and Sancar,A. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 2046–2050.
Kiener,A., Husain,I., Sancar,A. and Walsh,C. (1989) *J. Biol. Chem.*, **264**, 13880–13887.
Kobayashi,T., Takao,M., Oikawa,A. and Yasui,A. (1989) *Nucleic Acids Res.*, **17**, 4731–4744.
Komura,J., Mitani,H. and Shima,A. (1988) *In Vitro Cell. Dev. Biol.*, **24**, 294–298.
Li,Y.F. and Sancar,A. (1990) *Biochemistry*, **29**, 5698–5706.
Li,Y.F. and Sancar,A. (1991) *Nucleic Acids Res.*, **19**, 4885–4890.
Li,Y.F., Heelis,P.F. and Sancar,A. (1991) *Biochemistry*, **30**, 6322–6329.
Li,Y.F., Kim,S.-T. and Sancar,A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 4389–4393.
Miki,K., Tamada,T., Nishida,H., Inaka,K., Yasui,A., de Ruiter,P.E. and Eker,A.P.M. (1993) *J. Mol. Biol.*, **233**, 167–169.
Nölling,J., Hahn,D., Ludwig,W. and de Vos,W.M. (1993) *System. Appl. Microbiol.*, **16**, 208–215.
Payne,G., Wills,M., Walsh,C. and Sancar,A. (1990) *Biochemistry*, **29**, 5706–5711.
Sabourin,C.L.K. and Ley,R.D. (1988) *Photochem. Photobiol.*, **47**, 719–723.
Sancar,G.B. (1985) *Nucleic Acids Res.*, **13**, 8231–8246.
Sancar,G.B. (1990) *Mutat. Res.*, **236**, 147–160.
Sancar,A., Franklin,K.A. and Sancar,G.B. (1984a) *Proc. Natl Acad. Sci. USA*, **81**, 7397–7401.
Sancar,G.B., Smith,F.W., Lorence,M.C., Rupert,C.S. and Sancar,A. (1984b) *J. Biol. Chem.*, **259**, 6033–6038.
Sancar,G.B., Smith,F.W. and Heelis,P.F. (1987) *J. Biol. Chem.*, **262**, 15457–15465.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
Takao,M., Oikawa,A., Eker,A.P.M. and Yasui,A. (1989a) *Photochem. Photobiol.*, **50**, 633–637.
Takao,M., Kobayashi,T., Oikawa,A. and Yasui,A. (1989b) *J. Bacteriol.*, **171**, 6323–6329.
Touzel,J.P., Conway de Macario,E., Nölling,J., de Vos,W.M., Zhilina,T. and Lysenko,A.M. (1992) *Int. J. System. Bacteriol.*, **42**, 408–411.
Walen,K.H. and Brown,S.W. (1962) *Nature*, **194**, 406.
Wang,B. and Jorns,M.S. (1989) *Biochemistry*, **28**, 1148–1152.
Yajima,H., Inoue,H., Oikawa,A. and Yasui,A. (1991) *Nucleic Acids Res.*, **19**, 5359–5362.
Yamamoto,K., Satake,M., Shinagawa,H. and Fujiwara,Y. (1983) *Mol. Gen. Genet.*, **190**, 511–515.
Yasuhira,S. and Yasui,A. (1992) *J. Biol. Chem.*, **267**, 25644–25647.
Yasui,A. (1993) In Shima,A., Ichahashi,M., Fujiwara,Y. and Takebe,H. (eds), *Frontiers of Photobiology*. Elsevier Science Publishers BV, Amsterdam, The Netherlands, pp. 355–356.
Yasui,A. and Langeveld,S.A. (1985) *Gene*, **36**, 349–355.
Yasui,A., Takao,M., Oikawa,A., Kiener,A., Walsh,C.T. and Eker,A.P.M. (1988) *Nucleic Acids Res.*, **16**, 4447–4463.
Yasui,A., Yajima,H., Kobayashi,T., Eker,A.P.M. and Oikawa,A. (1992) *Mutat. Res.*, **273**, 231–236.

Received on August 15, 1994; revised on September 28, 1994