Nucleic Acids Research, Vol. 19, No. 19 5359–5362

Cloning and functional characterization of a eucaryotic DNA photolyase gene from *Neurospora crassa*

Hirohiko Yajima, Hirokazu Inoue[1], Atsushi Oikawa and Akira Yasui*
The Research Institute for Tuberculosis and Cancer, Tohoku University, Sendai 980 and [1]Laboratory of Genetics, Department of Regulation Biology, Faculty of Science, Saitama University, Urawa 338, Japan

Received June 7, 1991; Revised and Accepted September 5, 1991 EMBL accession no. X58713

ABSTRACT

We cloned a genomic fragment of a photolyase gene from *Neurospora crassa* by polymerase chain reaction using synthesized oligonucleotide primers designed from the most conserved amino acid sequences among photolyases of various organisms. Using the cloned fragment as a hybridization probe we isolated a genomic fragment and cDNA clones encoding the complete photolyase gene of this organism. The amino acid sequence of the photolyase deduced from the determined nucleotide sequence indicates a protein consisting of 615 amino acid residues (M_r 69,971), which is most similar to that of *Saccharomyces cerevisiae*. Like yeast photolyase it contains a protruding amino terminus which is missing in photolyases of bacterial origin. Comparison of amino acids sequences among six photolyases suggests that the *Neurospora crassa* photolyase is more similar to photolyases of pterin type than those of deazaflavin type.

INTRODUCTION

Photoreactivation is a defense mechanism in living cells against ultraviolet light in which photolyase (EC 4.1.99.3) monomerizes UV-induced pyrimidine dimers by using light energy of near-UV or visible light (see 1 for review). This DNA repair mechanism seems to be present already in the early period of evolution, because we found extensive similarities in the amino acid sequences deduced from the cloned photolyase genes of a gram-negative bacterium *Escherichia coli*, a cyanobacterium *Anacystis nidulans*, a bacterium of neutral type *Streptomyces griseus*, an archaebacterium *Halobacterium halobium* and a eucaryote *Saccharomyces cerevisiae* (2).

Photoreactivation in eucaryotic cells has been extensively analyzed and photolyase activity was found from yeast to mammalian cells (3). Although cells in internal organs of higher organisms like chicken, rat kangaroo, fish or insects possess active photolyases, functional meanings of this light dependent enzyme in those cells are not exactly understood.

A eucaryotic photolyase gene was cloned from *S.cerevisiae* (4,5) and the deduced amino acid sequence of the yeast gene contained a protruding amino terminus which is absent in all the above mentioned bacterial photolyases (6,7). In order to understand the functions of photolyase in eucaryotic cells, we tried to isolate genes from other eucaryotes. For this purpose we applied the polymerase chain reaction (8) by using synthesized oligonucleotide primers designed from the most conserved amino acid sequences among photolyases determined previously. Using this strategy we cloned a photolyase gene from a filamentous fungus, *Neurospora crassa*, as the second example from eucaryote, which enabled us to search for common characteristics of photolyase in eucaryotes and to determine the conserved amino acid sequences essential for enzymatic function.

MATERIALS AND METHODS

Enzymes and sequencing kits

Enzymes necessary for restriction and modification of DNA were obtained from TOYOBO, New England Biolabs, INc., TAKARA, and Boehringer Mannheim Biochemicals. The Sequenase and multiprime labeling kit were purchased from TOYOBO and Amersham Corp., respectively.

Oligonucleotides were synthesized by a DNA synthesizer 381A (Applied Biosystems). Among them two primers named PCR7 and PCR10 were successful for the cloning of the photolyase gene from *N.crassa*. The nucleotide sequences were:

PCR7 (32mer); 5'-GGGGAATTC AA(A/C) (A/G)A(T/C) (T/C)TN (A/C/T)TN (A/G)TN GA(T/C) TGG (A/C)G-3', and
PCR10 (26mer); 5'-GGGAAGCTT NN (A/C/G)(A/C)A NCC CCA NCC NCC-3',

where N means all the four nucleotides of A, G, C, and T. PCR7 corresponds to the amino acid sequence of K(N/D)L(I/L)(I/V) DWR and PCR10 to GGWG(F/W) found in the most conserved region among five photolyases.

Genomic DNA and cDNA libraries of *N.crassa*

Genomic DNA of *N.crassa* of the wild type strain C1-T10-37A was isolated as previously described (9). Genomic DNA was

* To whom correspondence should be addressed at The Research Institute for Tuberculosis and Cancer, Tohoku University, Seiryomachi 4-1, Aobaku, Sendai 980, Japan

partially digested with *Mbo*I and fractions between 10 kb and 20 kb were cloned into lambda EMBL4 vector by a standard procedure (10). Lambda ZAP cDNA library was a kind gift from Dr. M.Sachs (Stanford University).

Polymerase chain reaction

Polymerase chain reaction was performed under following conditions; 93°C 1 min, 55°C 2 min and 72°C 1.5 min.

Expression of cDNA in *E.coli*

In order to express the photolyase gene from the putative three initiation codons, ATG1, ATG2 and ATG3 (see RESULTS and DISCUSSION), an *Eco*RI site was introduced in front of each initiation codon by polymerase chain reaction and the amplified fragments were inserted behind a *tac* promoter in an expression vector pKK223-3 (Pharmacia LKB Biotechnology). The absence of error due to Taq polymerase in the amplified fragments was confirmed afterwards by sequencing.

Nucleotide sequencing

The cloned 3.6 kb fragment derived from the EMBL4 genomic library was subcloned at first into pUC18 vector. Then a number of subclones containing a part of the 3.6 kb fragment were created by further digestion with various restriction enzymes. Sequences of these subclones were determined by the dideoxy method (11) using Sequenase (United States Biochemicals). Similarly the nucleotide sequences of the cloned cDNA were determined.

Photoreactivation experiments

E.coli KY29 (genotype W(*lac pro*) *end* A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 F′ *tra*D36 *pro*AB *lac*IqZWM15 *rec*A56*srl*-C300::Tn(*tet*r *phr*19(Cmr)) cells harboring plasmids were cultured in LB medium containing ampicillin (50 mg/ml). Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the medium at the cell density of 0.5 absorbance unit (A_{600}) to a final concentration of 1 mM and cells were allowed to grow to the stationary phase. Cells were then diluted ten-fold with M9 buffer (50 mM Na_2HPO_4, 22 mM KH_2PO_4, 19 mM NH_4Cl) and irradiated with UV-light (Hitachi germicidal lamp GL10, UV-dose rate of 2.5×10^{-2} J m^{-2} s^{-1}). After UV-irradiation cells were subjected to photoreactivation by illumination with fluorescent lamps (Toshiba, FL20SS W/18) for 10 min. Cells were then plated on ampicillin-LB agar medium and incubated at 37°C until colony counting.

RESULTS AND DISCUSSION

Cloning and sequence determination of the *N.crassa* photolyase gene from genomic DNA library

Using primers, PCR7 and PCR10 (see MATERIALS AND METHODS), genomic DNA of *N.crassa* was amplified by polymerase chain reaction. The amplified DNA was digested with *Eco*RI and *Hin*dIII, at the primer sequences, and subsequently ligated to pUC18 vector plasmid that had been digested with both enzymes. Two out of ten transformants contained plasmids with an insert of 115 bp, a length of expected size predicted from photolyases of other organisms. The nucleotide sequence of the 115 bp was determined and the deduced amino acid sequence showed a similarity with those of other photolyases (sequence not shown).

Using this amplified DNA fragment as a probe, a genomic library of *N.crassa* made in EMBL4 vector was screened. Several positive clones were identified and the insert of a clone was subcloned into pUC18. The nucleotide sequence of a 3.6 kb long insert was determined, and has been registered in EMBL data library as X58713. In the determined sequence there is an open reading frame (ORF) of 2001 bp.

Cloning, sequencing and characterization of cDNA

A cDNA library was screened with a 1.7 kb *Ban*II fragment of the cloned genomic DNA. Three cDNA clones were isolated from the library. The nucleotide sequence of a cDNA was determined and compared with the genomic sequence. The nucleotide sequence of the cDNA was exactly the same as that of the corresponding genomic sequence except a lack of 75 bp sequence at the middle of the ORF in the cDNA sequence. Fig1.B shows a comparison of genomic and cDNA sequences around the 75 bases. The 75 bases do not change the existing open reading frame, but its deduced amino acid sequence has no similarity to any part of the deduced sequences of other photolyases. On the other hand, consensus nucleotide sequences, GT at 5′- and AG at 3′-ends of the sequence, were found at the both end of the 75 bp. Furthermore, a sequence CCCTAAC, which is shown in Fig.1B with additional shade on the sequence, resembles the reported lariat formation sequence for intron splicing in *N.crassa* and other filamentus fungi, 5′-PyGCTAA-C-3′. In this consensus sequence only the fourth T seems to be well conserved (12).

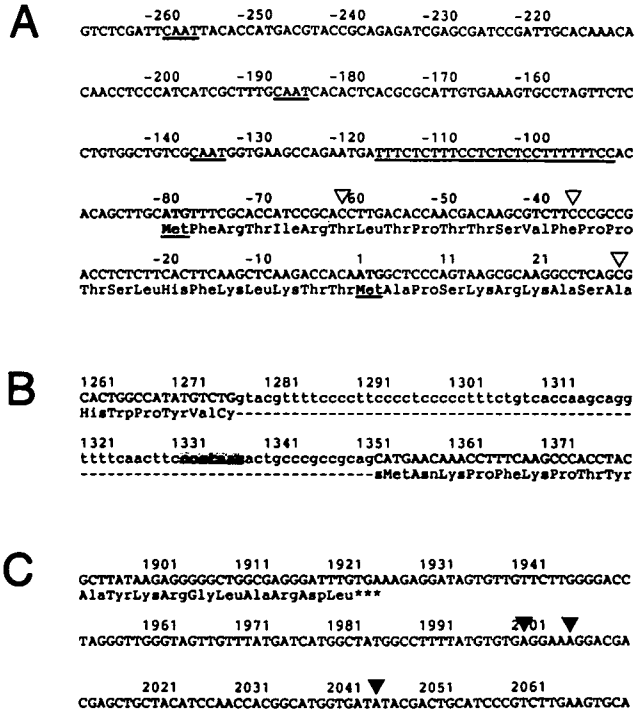

Fig.1. Three regions of the nucleotide and deduced amino acid sequences of the cloned genomic DNA. The adenine residue of the second methionine is numbered as 1. A. Sequences around possible translational initiation sites (underlined Mets) and in the upstream region. Open triangles indicate the 5′ termini of the cloned three cDNAs. Nucleotide sequences possibly influencing the transcription of the gene are underlined. B. Nucleotide sequence of the intron (written in small letters) and its surrounding. Putative lariat forming sequence within the intron is shaded. C. Sequences around carboxy terminus. Closed triangles indicate polyadenylylation sites found in the three cDNA clones.

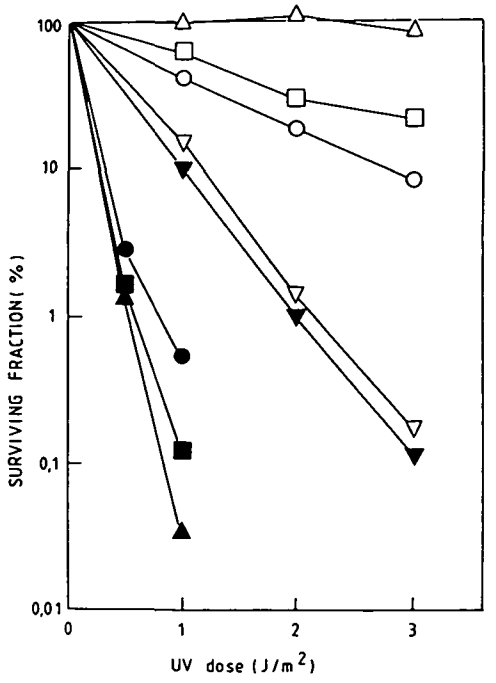

Fig.2. Survival rates of UV-irradiated *E.coli* KY29 cells. Cells were transformed with plasmids for expression of the photolyase gene from ATG1 (shown by circles), ATG2 (shown by triangles) or ATG3 (shown by squares). Survival curves of cells harboring vector plasmid are shown with inverted triangles. Open and closed symbols indicate survival rates with and without photoreactivation, respectively.

The longest ORF determined from the cDNA was 1926 bp. There are three potential initiation codons in the ORF designated as ATG1, ATG2 (Fig.1A, underlined Mets) and ATG3 (Fig.3, the 70th Methionine). Only the second initiation codon ATG2 has a nucleotide sequence, CCACAATGGC, in its vicinity, which matches very well to the consensus sequence around reported initiation codons of *N.crassa*, (T/C/A)CA(C/A)(A/C) ATGGC, where the underlined ATG is initiation codon (12). Biological results shown below also support that the second ATG is the major, at least, initiation codon resulting in a protein with 615 amino acids (M_r, 69,971).

There are several nucleotide sequences upstream to the first ATG, which may infulence the transcription of this gene. They are CAAT and pyrimidine rich sequences (12), which are underlined in Fig.1A. Translation termination site and polyadenylylation sites of each cDNA clones are shown in Fig.1C. A number of *N.crassa* genes do not possess any clear polyadenylylation signal (12). This is also true for photolyase gene.

Expression of *N.crassa* photolyase gene in *E.coli*

We introduced a *tac* promoter in front of ATG1, ATG2 and ATG3, and expressed the *N.crassa* photolyase cDNA in *E.coli* cells of *rec*A and *phr*. UV-survival curves with and without photoreactivation light illumination are shown in Fig.2. All plasmids provided the cells with photolyase activity, but the expression from ATG2 gave the highest activity to the host cells as judged from the increase in the survival rate after photoreactivation. Thus, the gene cloned from *N.crassa* encodes

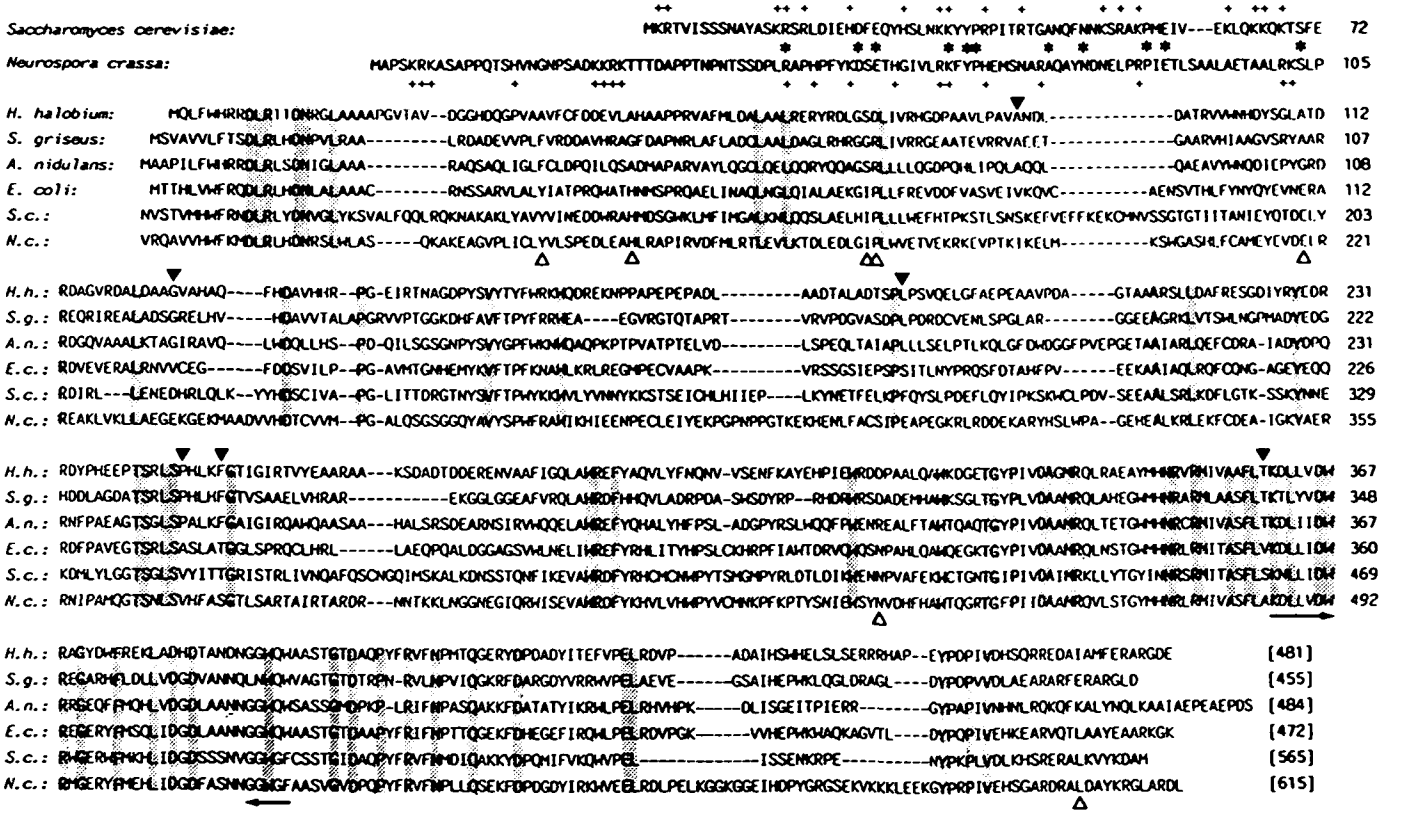

Fig.3. Amino acid sequence alignment among six photolyases. Residues conserved among all the six photolyases are shaded. Closed triangles show residues conserved exclusively in photolyases of *H.halobium*, *S.griseus*, and *A.nidulans*. Open triangles show those conserved exclusively in photolyases of *E.coli*, *S.cerevisiae*, and *N.crassa*. Two arrows indicate the sites corresponding to the primer nucleotides (PCR7 and PCR10) used for amplification of genomic DNA of *N.crassa*. Amino acid sequence of the *N.crassa* photolyase is shown from the second methionine among the three possible initiation codons. In the protruding amino terminal regions of *S.serevisiae* and *N.crassa*, positively charged residues are indicated with +. Identical residues between both sequences are indicated with asterisks.

a photolyase, and designated as *phr*, the second example of eucaryotic photolyase gene and the first DNA repair gene from *N.crassa* ever cloned.

In *E.coli* cells excision repair activity is stimulated by own photolyase (13, 14). All the plasmids introduced in *E.coli* cells for expression of *N.crassa* photolyase gene reduced the survival after UV-irradiation without light illumination as compared with the survival of cells harboring only vector plasmids (Fig.2). This suggests that the produced *Neurospora* photolyase molecules disturb the excision repair of *E.coli* host cells as was the case for other foreign photolyases genes introduced into the same host cells (15). It has been reported that some UV-sensitive *N.crassa* mutant strains showed reduced levels of photore-activation activity (16), suggesting that *N.crassa* photolyase cooperate with its own excision repair systems.

Alignment of deduced amino acids sequences among six photolyases

Fig.3 depicts an alignment of deduced amino acid sequences of all the six photolyases ever cloned. Photolyases analyzed until now contain two different cofactors in a molecule as chromophores. Photolyases from *E.coli* and *S.cerevisiae* exhibit an absorption maximum at 380 nm and contain a reduced flavin adenine dinucleotide ($FADH_2$) and a pterin, whereas those from *A.nidulans*, *S.griseus* and *H.halobium* harbor derivatives of deazaflavin besides $FADH_2$, which give the maximum absorption and optimal enzyme activity around 435 nm (17,18,19).

Cofactors in photolyase of *N.crassa* has not been determined, but the most effective wavelength of light for photoreactivation in this organism was reported as around 400 nm by in vivo and in vitro experiments (20,21), an intermediate optimum wavelength for photoreactivation between pterin (380nm) and deazaflavin (435nm) type photolyases. The alignment of deduced amino acid sequences shown in Fig.3 suggests what kind of cofactors photolyase of *N.crassa* may have. Three photolyases harboring a deazaflavin as the second cofactor possess a unique amino acid sequence of P(H or A)L(K or H)F about in the middle of the whole sequence, which *Neurospora* photolyase lacks. On the other hand, this photolyase contains an IP sequence found near amino terminal region of *E.coli* and yeast photolyases as indicated in Fig.3. These sequence comparisons suggest that *N.crassa* possesses a pterin like element as the second cofactor, which may be identified by analysis of purified photolyase of this organism.

Protruding amino terminal region in eucaryotic photolyases

Yeast and *N.crassa* photolyases possess protruding amino terminal sequences which are missing in bacterial ones. Between these two terminal sequences there are some similarities in amino acids sequence as well as in distribution of positively charged amino acids residues (Fig.3). It is interesting that the sequence KKYYPR in photolyase of yeast is similar to the sequence RKFYPH in that of *N.crassa*, which resembles the sequence in front of the cleavage site for transport into mitochondria of the mitochondrial leucyl-tRNA synthetase of *N.crassa* (22). Since yeast photolyase repairs pyrimidine dimers not only in nuclei but also in mitochondria (23,24), the above mentioned sequence may be one of the signal sequences necessary for translocation of photolyase molecules.

ACKNOWLEDGEMENT

We wish to thank Dr.M.Sacks for his kindly gift of *N.crassa* cDNA library.

REFERENCES

1. Sancar,G.B. (1990) Mutation Res., 236, 147–160.
2. Takao,M., Kobayashi,T., Oikawa,A., and Yasui,A. (1989) J.Bacteriol., 171, 6323–6329.
3. Friedberg,E.C. (1985) DNA Repair, W.H.Freeman and Company, New York.
4. Yasui,A., and Chevallier,M.R. (1983) Curr.Genet., 7, 191–194.
5. Schild,D., Johnston,J., Chang,C., and Mortimer,R.K. (1984) Mol.Cell.Biol., 4, 1864–1870.
6. Yasui,A., and Langeveld,S.A. (1985) Gene, 36, 349–355.
7. Sancar,G. (1985) Nucleic Acid Res., 13, 8231–8246.
8. Saiki,R.K., Scharf,S., Faloona,F., Mullis,K.B., Horn,G.T., Erlich,H.A., and Arnheim,N. (1985) Science, 230, 1350–1354.
9. Orbach,M.J., Porro,E.B., and Yanofsky,C. (1986) Mol.Cell.Biol., 6, 2452–2461.
10. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning.A Laboratory Manual.Second Edition. Cold Spring Harbor Laboratory.
11. Sanger,F., Nicklen,S., and Coulson,A.R. (1977) Proc.Natl.Acad.Sci. USA, 74, 5463–5467.
12. Gurr,S.J., Unkles,S.E., and Kinghorn,J.R. (1987) In Kinghorn, J.R. (ed.), Gene Structure in Eucaryotic Microbes. IRL Press, pp.93–139.
13. Yamamoto,K., Satake,M., Shinagawa,H., and Fujiwara,Y. (1983) Mol.Gen.Genet., 190, 511–515.
14. Sancar,A., Franklin,K.A., and Sancar,G.B. (1984) Proc.Natl.Acad.Sci. (USA), 81, 7397–7401.
15. Kobayashi,T., Takao,M., Oikawa,A., and Yasui,A. (1990) Mutation Res., 236, 27–34.
16. Ishii,C., and Inoue,H. (1989) Mut. Res., 218, 95–103.
17. Johnson,J.QL., Hamm-Alvares,S., Payne,G., Sancar,G.B., Rajagopalan,K.V., and Sancar,A. (1988) Proc. Natl. Acad. Sci. (USA), 85, 2046–2050.
18. Eker,A.P.M., Hessels,J.K.C., and Dekler,R.H. (1986) Photochem. and Photobiol., 44, 197–205.
19. Eker,A.P.M., Kooiman,P., Hessels,J.K.C., and Yasui,A. (1990) J. Biol. Chem., 265, 8009–8015.
20. Terry,C.E., kilbey,B.J., and Howe,H.B.,Jr. (1967) Radiation Res., 30, 739–747.
21. Terry,C.E. and Setlow,J.K. (1967) Photochem. Photobiol., 6, 799–803.
22. Chow,C.M., Metzenberg,R.L., and RajBhandary,U.L. (1989) Mol. Cell. Biol., 9, 4631–4644.
23. Prakash,L. (1975) J.Mol.Biol., 98, 781–795.
24. Waters,R., and Moustacchi,E. (1974) Biochim. Biophys. Acta, 366, 241–250.