The EMBO Journal vol.8 no.10 pp.3121 - 3125, 1989

Molecular cloning of human uracil - DNA glycosylase, a highly conserved DNA repair enzyme*

Lisbeth C.Olsen, Rein Aasland, Christian U.Wittwer[1], Hans E.Krokan[2] and Dag E.Helland[3]

Laboratory of Biotechnology, University of Bergen, N-5029 Bergen, [1]Department of Biochemistry, Institute of Medical Biology, University of Tromsø, N-9000 Tromsø, and [2]UNIGEN, Institute of Biotechnology, University of Trondheim, N-7030 Trondheim, Norway

[3]Present address: Dana Farber Cancer Institute, Boston MA 02115, USA

*This paper is dedicated to the memory of Professor Kjell Kleppe

Communicated by T.Lindahl

Uracil - DNA glycosylase is the DNA repair enzyme responsible for the removal of uracil from DNA, and it is present in all organisms investigated. Here we report on the cloning and sequencing of a cDNA encoding the human uracil - DNA glycosylase. The sequences of uracil - DNA glycosylases from yeast, *Escherichia coli*, herpes simplex virus type 1 and 2, and homologous genes from varicella - zoster and Epstein - Barr viruses are known. It is shown in this report that the predicted amino acid sequence of the human uracil - DNA glycosylase shows a striking similarity to the other uracil - DNA glycosylases, ranging from 40.3 to 55.7% identical residues. The proteins of human and bacterial origin were unexpectedly found to be most closely related, 73.3% similarity when conservative amino acid substitutions were included. The similarity between the different uracil - DNA glycosylase genes is confined to several discrete boxes. These findings strongly indicate that uracil - DNA glycosylases from phylogenetically distant species are highly conserved.

Key words: DNA repair/evolutionary conservation/human uracil - DNA glycosylase/molecular cloning

Introduction

Uracil in DNA may result from deamination of cytosine (Lindahl and Nyberg, 1974; Shapiro, 1980) or misincorporation of dUMP residues (Tye *et al.*, 1977; Brynolf *et al.*, 1978; Wist *et al.*, 1978). The first process can give rise to transition mutations (Duncan and Weiss, 1982). Thus one likely biologically important function of uracil - DNA glycosylase is to prevent mutagenesis resulting from cytosine deamination. Uracil - DNA glycosylase eliminates uracil from DNA by cleaving the N-glycosylic bond and initiates the base-excision repair pathway. This presumed pathway is subsequently thought to involve the action of AP endonucleases, DNA polymerase and DNA ligase (reviewed by Lindahl, 1979). Experiments with monoclonal antibodies have indicated that uracil - DNA glycosylase and DNA polymerase α are physically associated. It was therefore suggested that these enzymes are parts of a functional complex of replication and repair enzymes (Seal and Sirover, 1986).

Uracil - DNA glycosylase activity was first detected in extracts of *Escherichia coli* (Lindahl, 1974), and has since been isolated from various other sources including yeast (Crosby *et al.*, 1981) and mammals (Sekiguchi *et al.*, 1976; Wist *et al.*, 1978; Krokan and Wittwer, 1981). Eukaryotic cells have both a nuclear and mitochondrial form of uracil - DNA glycosylase (Anderson and Friedberg, 1980; Wittwer and Krokan, 1985; Domena *et al.*, 1988). The *ung* and *UNG1* genes encoding the *E.coli* and yeast enzymes, respectively, have been isolated and recently sequenced (Duncan and Chambers, 1984; Varshney *et al.*, 1988; Percival *et al.*, 1989). The proteins encoded by these two genes show a high degree of similarity. A uracil - DNA glycosylase encoded by herpes simplex virus type 2 has been purified and cloned (Caradonna *et al.*, 1987). The cDNA clone was mapped to the region at about 0.07 map units at the HSV genome and derives from the UL-2 gene (Worrad and Caradonna, 1988; Mullaney *et al.*, 1989). Homologous genes are also found in other herpes virus such as the varicella - zoster (gene 59, Davison and Scott, 1986) and Epstein - Barr viruses (BKRF3, Baer *et al.*, 1984). Here we report on the cloning and sequencing of a cDNA encoding the human uracil - DNA glycosylase.

Results

Characterization of human uracil - DNA glycosylase cDNA

Uracil - DNA glycosylase has been purified from human placenta (Wittwer *et al.*, 1989). This tissue contains several

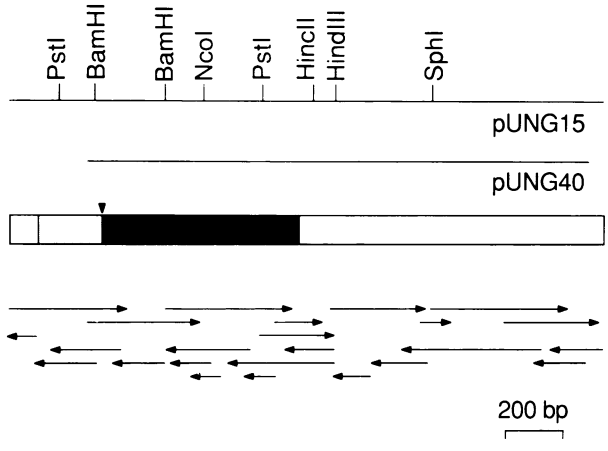

Fig. 1. Structure and sequencing strategy of pUNG-15 and pUNG-40. DNA sequencing was carried out using the dideoxy method (Sanger *et al.*, 1977). The arrows represent the extent and direction of sequencing using either SP6, T7 or other synthetic primers. The open reading frame is indicated by the shaded boxes, the darker box representing the predicted mature protein.

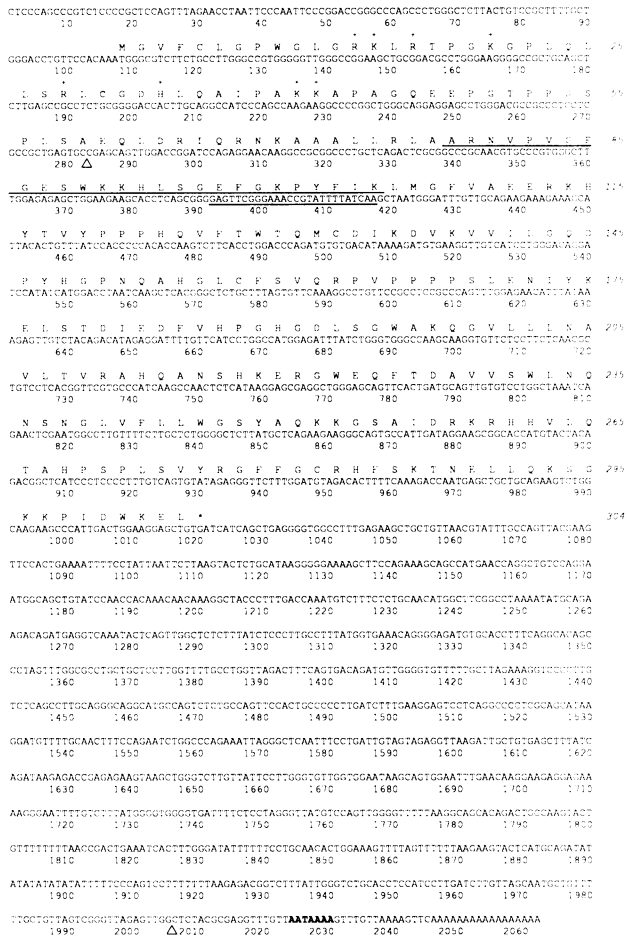

Fig. 2. Nucleotide sequence of human placental uracil−DNA glycosylase cDNA and the predicted amino acid sequence of the protein. The amino acid sequence corresponding to the N-terminal sequence of the placental enzyme and the oligonucleotide probe used to screen the library are underlined. The polyadenylation signal sequence is indicated by bold face letters. The 5′ and 3′-ends of the pUNG-40 clone are indicated by open triangles. The positively charged amino acid residues of the N-terminal segment resembling a mitochondrial presequence are indicated. The sequence data will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession number of X15653.

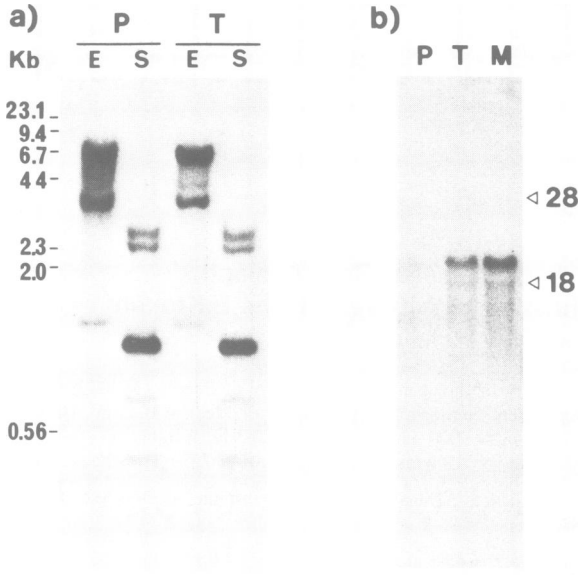

Fig. 3. Southern and Northern blot analysis with pUNG-40 cDNA. Hybridization was carried out as described in Materials and methods. (**a**) Southern analysis of human genomic DNA derived from placenta (P) and thyroid tissue (T) digested with *Eco*RI (E) and *Pst*I (S). (**b**) Northern analysis of total RNA derived from human placenta (P), human thyroid tumour (T) and the mcf-7 mammary carcinoma cell line (M). Positions of 18 and 28S ribosomal RNA are indicated.

biochemically closely related forms of uracil−DNA glycosylase. A mixed oligonucleotide was synthesized on the basis of the N-terminal amino acid sequence (residues 19−26) of the major form. The oligonucleotide was used as a probe to screen a λgt11 cDNA library from human placenta. Approximately 300 000 plaques were screened, and more than 20 positive clones were found. Four of these were subcloned in plasmids. Two of the positive clones, pUNG-15 and pUNG-40, were characterized by restriction enzyme mapping and DNA sequencing (Figure 1). These analyses revealed that pUNG-15 and pUNG-40 were overlapping and that pUNG-15 was 282 and 56 bp longer at the 5′ and 3′-ends, respectively. The insert in pUNG-15 terminates with a 17 bp poly(A)tail starting 16 bp downstream from a canonical polyadenylation signal sequence while pUNG-40 terminates 19 bp upstream from this polyadenylation signal (Figure 2).

DNA sequencing of pUNG-40 and pUNG-15 revealed an open reading frame of 981 nucleotides (positions 38−1018). The first ATG-codon downstream from the in-frame stop codon at position 35 is found at position 107 and is a possible initiation codon. The resulting reading frame encodes a translation product of 304 amino acids. The N-terminal amino acid sequence of the purified protein (Wittwer *et al.*, 1989) is found at residues 78−105 in the predicted protein sequence. This suggests that the initial translation product is processed to a mature protein which has a relative molecular mass of 25.8 kd. The N-terminal sequence of the 26.5 kd co-purifying polypeptide (see figure 4B of Wittwer *et al.*, 1989) is not present in the predicted protein sequence. The amino acid composition of the purified protein, as determined after acid hydrolysis was similar to that of the predicted mature protein sequence, with the exception of the values of glycine, tyrosine and arginine (Wittwer *et al.*, 1989). This discrepancy is most likely due to the limited accuracy of the applied method of amino acid analysis. The predicted mature protein sequence contains 40 positively and 20 negatively charged residues, suggesting that the protein has a net positive charge.

The purification protocol of human uracil−DNA glycosylase did not include a subcellular fractionation step. We therefore have to consider the possibility that the isolated cDNA clones may encode a nuclear or a mitochondrial enzyme. None of the unidentified open reading frames of the human mitochondrial genome (Anderson *et al.*, 1981) show any similarity to the pUNG cDNA clones. These cDNA clones must therefore be encoded by a nuclear gene. It should be noted, however, that the N-terminal sequence of the putative precursor protein contains a region rich in positively charged and hydrophobic amino acids reminiscent of a mitochondrial presequence (Roise and Schatz, 1988; Roise *et al.*, 1988).

Southern blot analysis of human genomic DNA showed the presence of both strongly and weakly hybridizing *Eco*RI and *Pst*I fragments adding up to less than 12 kb in each case

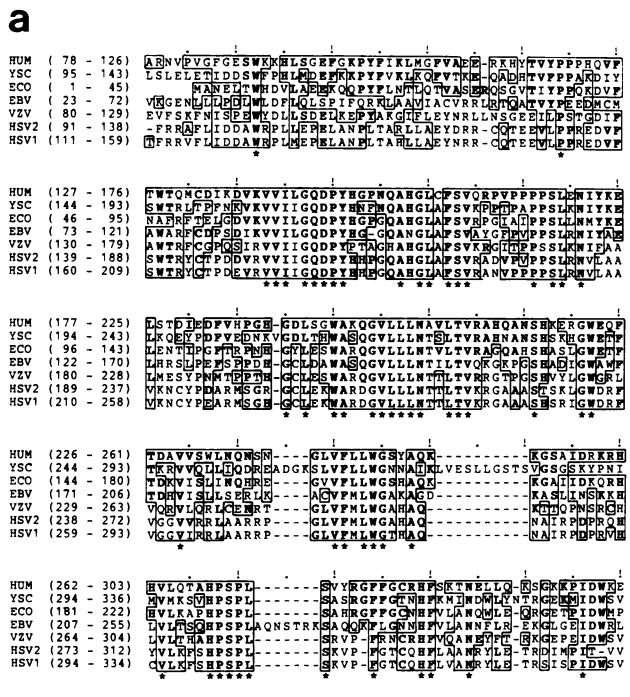

Fig. 4. Comparison of protein sequences of uracil – DNA glycosylases. (**a**) Alignment of uracil – DNA glycosylase from man (HUM), yeast (YSC, *UNG1* gene, Percival *et al.*, 1989), *E.coli* (ECO, *ung* gene, Varshney *et al.*, 1988), herpes simplex virus type 2 (HSV2, gene UL2, Worrad and Caradonna, 1988) and type 1 (HSV1, gene UL2, McGeoch *et al.*, 1988), and the homologous sequences from Epstein – Barr Virus (EBV, gene BKRF3, Baer *et al.*, 1984), and varicella – zoster Virus (VZV, gene 59, Davison and Scott, 1986). Gaps in the protein sequences have been introduced to yield maximal alignment. Similarity between the human sequence and at least one of the other sequences is indicated. Identical amino acids are indicated by bold face and conservative amino acid replacements (scores greater than zero in the mutation probability data matrix, Dayhoff, 1978) are boxed. Residues which are identical in all sequences are marked with asterisks. The numerical values from this analysis are presented in Table I. (**b**) Hydropathy profiles of human (HUM, residues 83 – 304) and *E.coli* (ECO, residues 1 – 223) uracil – DNA glycosylases (using the method of Kyte and Doolittle, 1982) with a window size of 15 as implemented in the Staden-Plus software (Amersham). Positive and negative values are indicated by arbitrary units.

(Figure 3a), indicating a small size of the gene encoding uracil – DNA glycosylase. At the high stringency used in this experiment, hybridization was also seen with mouse genomic DNA (not shown). A Northern blot analysis of RNA from several human tissues including placenta revealed a 2.4 kb transcript hybridizing to the pUNG-40 cDNA probe (Figure 3b).

Table I. Pairwise comparisons between uracil – DNA glycosylase sequences from man, yeast, *E.coli* and the herpes viruses. The values are percentages identical (upper half) and conserved (lower half) residues resulting from the alignment shown in Figure 4a. The compared regions correspond to residues 83 – 303 (221 residues) of the human sequence and the percentages were calculated by division by a standard length of 221 residues

	% identical residues						
	HUM	YSC	ECO	EBV	VZV	HSV2	HSV1
HUM	—	53.9	55.7	43.4	40.3	40.7	41.2
YSC	70.1	—	52.5	40.7	41.2	43.4	43.4
ECO	73.3	70.6	—	47.1	45.2	43.9	44.8
EBV	61.5	58.8	68.8	—	46.2	40.3	41.6
VZV	59.3	60.2	63.4	64.3	—	51.1	52.9
HSV2	60.8	58.4	63.4	60.6	72.4	—	89.6
HSV1	61.5	59.7	63.8	61.5	74.7	94.6	—
	% conserved residues						

Sequence comparisons with the human uracil – DNA glycosylase

A comparison of the sequences of the uracil – DNA glycosylases from man, yeast, *E.coli* and herpes simplex virus type 1 and 2, and the homologous proteins of Epstein – Barr and varicella – zoster viruses, revealed a striking similarity between all these proteins (Figure 4a, Table I). The similarity spans about 230 amino acids and includes the entire bacterial and mature human proteins. The putative human precursor protein and the yeast and viral proteins have additional non-similar N-terminal regions of varying length. Unexpectedly, the human protein shows the highest similarity to the bacterial protein possessing 55.7% identical residues, as compared to 53.9 and 40.3 – 43.4% to the yeast and viral proteins, respectively (Table I). When conservative amino acid substitutions are included, the similarity increases to 73.3% for the human and the bacterial proteins. The significance of these similarities becomes even clearer when the degree of co-linearity is considered. Only one insertion in the human sequence is required to perfectly align the two sequences. Futhermore, the two mature proteins are almost equal in size. The alignment of the different uracil – DNA glycosylases (Figure 4a) exhibits the existence of particularly conserved regions of varying size. Hydropathy profiles (Figure 4b) of the human and the bacterial proteins are very similar, suggesting that these proteins have similar structures.

The sequence of the yeast uracil – DNA glycosylase has two major segments (residues 257 – 260 and 274 – 283) not present in the other compared sequences (Figure 4a). Additionally, the residues 288 – 294 in the yeast sequence show no relatedness to the corresponding regions of the other sequences. These differences account for the lower degree of similarity between the yeast and the human sequences as compared to the similarity between the bacterial and human sequences.

The human uracil – DNA glycosylase sequence was also compared with the contents of GenBank (release 57) in all six reading frames without finding any additional similarities. In particular, no similarity was seen between this protein and other sequenced DNA glycosylases including the *E.coli*

3-methyl-adenine glycosylases, *AlkA* and *TagA* (Nakabeppu *et al.*, 1984; Steinum and Seeberg, 1986) formamidopyrimidine−DNA glycosylase, *Fpg* (Boiteux *et al.*, 1987), and the cloned human and yeast DNA repair genes including the human *ERCC-1* gene (van Duin *et al.*, 1988) and the yeast *RAD* genes.

Discussion

In this paper we report the cloning and sequencing of cDNAs encoding the human uracil−DNA glycosylase. This is the first mammalian damage-specific DNA repair gene with a known function to be cloned.

The predicted reading frame of the cDNA clone pUNG-15 encodes a protein of 304 amino acids with a calculated molecular weight of 33.8 kd. The sequence data suggest that the initial translation product is processed since the N-terminal amino acid sequence is found 77 amino acids downstream from the first methionine. The molecular weight of the predicted mature protein (25.8 kd) is different from that reported for the purified protein (29 kd, see Wittwer *et al.*, 1989). Possible explanations of this discrepancy include post-translational modification of the protein and anomalous electrophoretic mobility. The size of the predicted mature protein is also different from the 37 kd polypeptide immunoprecipitated with anti-placental uracil−DNA glycosylase monoclonal antibodies (Vollberg *et al.*, 1987). Using these antibodies, processing of this 37 kd polypeptide was not observed (Vollberg *et al.*, 1987). The purification procedure used by Wittwer *et al.* (1989) may, however, have resulted in an artificial cleavage of the enzyme. This is apparently the case in yeast, where the 39.5 kd primary translation product is converted *in vitro* to a 28 kd species during a partial purification procedure (Percival *et al.*, 1989). Uracil−DNA glycosylase activity was detected after *in vitro* transcription/translation of the pUNG-15 cDNA clone (unpublished results), indicating that this cDNA has a coding capacity for a functional enzyme. An alternative explanation for the conflicting results regarding the size of the human enzyme is that human placenta contains several different uracil−DNA glycosylases.

Since the probe used to isolate the pUNG cDNA clones was based on the N-terminal amino acid sequence of the major form of the enzyme, these clones most likely encode a nuclear enzyme. The observation that human uracil−DNA glycosylase contains a region with features of a mitochondrial presequence (Roise and Schatz, 1988; Roise *et al.*, 1988), suggests that this gene may encode both a mitochondrial and a nuclear enzyme. Further analyses will be required to test this hypothesis.

Southern blot analysis with human DNA revealed the presence of both strongly and weakly hybridizing fragments. This may suggest that one or more related genes exists or, alternatively, that the gene may contain several small exons. An abundant 2.4 kb transcript was detected by Northern blot analysis. Additionally, the human placenta cDNA library contained many uracil−DNA glycosylase clones. These findings show that the uracil-DNA glycosylase gene is highly transcribed in mammalian cells.

In order to evaluate the relatedness of the human enzyme to other reported uracil-DNA glycosylases, the predicted protein sequences of all these genes were compared. From this study it is evident that uracil−DNA glycosylase is highly conserved between man, yeast, *E.coli* and the herpes viruses, and that the proteins of human and bacterial origin were most similar. In addition these proteins exhibit the existence of particularly conserved regions of varying size. It is possible that some of these regions may be part of the catalytic site. Site directed mutagenesis will be used to identify the functional regions of this protein (work in progress). The unexpected higher similarity between the bacterial and human uracil−DNA glycosylase sequences than between the yeast and human sequences is largely due to a segment of seven unrelated residues in the yeast sequence.

Interestingly, DNA polymerases from the herpes viruses and other animal DNA viruses, bacteriophages and mammals also show extensive similarities (Spicer *et al.*, 1988; Wong *et al.*, 1988). This supports the view that these organisms employ similar enzymes in DNA repair and replication.

The uracil−DNA glycosylase sequences were also compared to other sequenced DNA glycosylases including the *tagA* (Steinum and Seeberg, 1986), *alkA* (Nakabeppu *et al.*, 1984) and *fpg* (Boiteux *et al.*, 1987) gene products. Despite their related enzymatic activities, their amino acid sequences do not show any extensive similarities. This may reflect the narrow substrate specificities of this group of enzymes.

The high degree of conservation between the uracil−DNA glycosylases from phylogenetically distant species indicates that these proteins have evolved from a common ancestor. No other protein is known to be so strongly conserved from bacteria to man. The degree of conservation is greater than that reported for the heat shock proteins (Chappell *et al.*, 1986) and the human P-glycoprotien/*E.coli* haemolysin (*HlyB*) (Gerlach *et al.*, 1986). The 230 C-terminal amino acids of the latter pair show 74% similarity when conservative amino acid substitutions are considered, but have only 46% identical residues. This contrasts with the similarity and colinearity of the entire human and bacterial uracil−DNA glycosylases throughout their sequences.

Materials and methods

Isolation of cDNA clones

A human placenta cDNA λgt11 library (Clontech) was screened using a ^{32}P-labelled oligonucleotide as a probe. Hybridization was carried out at 42°C overnight. The hybridization solution contained 5 × SSPE, 5 × Denhardt's solution, 0.1% SDS, 200 μg/ml heat denatured salmon sperm DNA. Filters were washed in 2 × SSPE/0.1% SDS and 0.2 × SSPE/0.1% SDS at 42°C and autoradiographed.

Subcloning and DNA sequencing

Positive λ clones were purified using LambdaSorb Immunaffinity Adsorbent (Promega). The DNA was digested with *Eco*RI and subcloned in the vector pGEM-7Zf(+) (Promega). DNA subclones of varying lengths were generated by subcloning of restriction fragments and by using the Erase-a-Base system (Promega). DNA sequencing was carried out using the dideoxy terminator method (Sanger *et al.*, 1977) employing the modified T7 polymerase (Sequenase, USB), and SP6, T7 and other synthetic primers. To resolve compressions, selected regions were sequenced with the addition of dITP to the reaction mixture according to the manufacturer's recommendations.

Southern and Northern blot analyses

Human genomic DNA from different tissues (15 μg) was digested with *Eco*RI and *Pst*I, separated on 0.8% agarose gel, blotted onto nylon filter (Hybond-N, Amersham) and hybridized to ^{32}P-labelled insert of pUNG-40 in the presence of 45% formamide, 5 × SSC, 200 μg/ml heat denatured salmon sperm DNA, 0.1% SDS, 25 mM sodium phosphate, pH 6.5, 8.25% dextran

sulphate at 42°C for 15 h. Filters were washed in 0.2 × SSC, 0.1% SDS at 65°C and autoradiographed. Total RNA isolated (Chomczynski and Sacchi, 1987) from different tissues was separated on 1% agarose in the presence of 2.2% formaldehyde, blotted onto nylon filter and hybridized to the pUNG-40 probe as described above.

Acknowledgements

We thank Anders Molven, Dr Rune Male, Sigve Håvardstein, Karl-Henning Kalland, Dr Ian F.Pryme and Dr Dietmar Kamp for advice and helpful discussions. This work was supported by grants from the Norwegian Cancer Society, The Norwegian Society for Science and the Humanities, the Norwegian Society for the Agricultural Sciences, R.A. is a fellow of the Norwegian Cancer Society.

References

Anderson,C.T.M. and Friedberg,E.C. (1980) *Nucleic Acids Res.*, **8**, 875–888.

Anderson,S., Bankier,A.T., Barrell,B.G., de Bruijn,M.H.L., Coulson,A.R., Drouin,J., Eperon,I.C., Nierlich,D.P., Roe,B.A., Sanger,F., Schreier,P.H., Smith,A.J.H., Staden,R. and Young,I.G. (1981) *Nature*, **290**, 457–465.

Baer,R., Bankier,A.T., Biggin,M.D., Deininger,P.L., Farrell,P.J., Gibson,T.J., Hatfull,G., Hudson,G.S., Satchwell,S.C., Seguin,C., Tuffnell,P.S. and Barrell,B.G. (1984) *Nature*, **310**, 207–211.

Boiteux,S., O'Connor,T.R. and Laval,J. (1987) *EMBO J.*, **6**, 3177=3183.

Brynolf,K., Eliasson,R. and Reichard,P. (1978) *Cell*, **13**, 573–580.

Caradonna,S. Worrad,D. and Lirette,R. (1987) *J. Virol.*, **61**, 3040–3047.

Chappell,T.G., Welch,W.J., Schlossman,D.M., Palter,K.B., Schlesinger,M.J. and Rothman,J.E. (1986) *Cell*, **45**, 3–13.

Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.

Crosby,B., Prakash,L., Davis,H. and Hinkle,D.C. (1981) *Nucleic Acids Res.*, **9**, 5797–5809.

Davison,A.J. and Scott,J.E. (1986) *J. Gen. Virol.*, **67**, 1759–1816.

Dayhoff,M.O. (1978) *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation, Washington, Vol. 5, Suppl. 3.

Domena,J.D., Timmer,R.T., Dicharry,S.A. and Mosbaugh,D.W. (1988) *Biochemistry*, **27**, 6742–6751.

Duncan,B.K. and Chambers,J.A. (1984) *Gene*, **28**, 211–219.

Duncan,B.K. and Weiss,B. (1982) *J. Bacteriol.*, **151**, 750–755.

Gerlach,J.H., Endicott,J.A., Juranka,P.F., Henderson,G., Sarangi,F., Deuchars,K.L. and Ling,V. (1986) *Nature*, **324**, 485–489.

Krokan,H. and Wittwer,C.U. (1981) *Nucleic Acids Res.*, **9**, 2599–2613.

Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.

Lindahl,T. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 3649–3653.

Lindahl,T. (1979) *Prog. Nucleic Acid Res. Mol. Biol.*, **22**, 135–192.

Lindahl,T. and Nyberg,B. (1974) *Biochemistry*, **13**, 3405–3410.

McGeoch,D.J., Dalrymple,M.A., Davison,A.J., Dolan,A., Frame,M.C., McNab,D., Perry,L.J., Scott,J.E. and Taylor,P. (1988) *J. Gen. Virol.*, **69**, 1531–1574.

Mullaney,J., Moss,H.W.M. and McGeoch,D.J. (1989) *J. Gen. Virol.*, **70**, 449–454.

Nakabeppu,Y., Miyata,T., Kondo,H., Iwanaga,S. and Sekiguchi,M. (1984) *J. Biol. Chem.*, **259**, 13730–13736.

Percival,K.J., Klein,M.B. and Burgers,P.M.J. (1989) *J. Biol. Chem.*, **264**, 2593–2598.

Roise,D. and Schatz,G. (1988) *J. Biol. Chem.*, **263**, 4509–4511.

Roise,D., Theiler,F., Horvath,S.J., Tomich,J.M., Richards,J.H., Allison,D.S. and Schatz,G. (1988) *EMBO J.*, **7**, 649–653.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.

Seal,G. and Sirover,M.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7608–7612.

Sekiguchi,M., Hayakawa,H., Makino,F., Tanaka,K. and Okada,Y. (1976) *Biochem. Biophys. Res. Commun.*, **73**, 293–299.

Shapiro,R. (1980) In Seeberg,E. and Kleppe,K. (eds), *Chromosome Damage and Repair*. Plenum Press, New York, pp. 3–18.

Spicer,E.K., Rush,J., Fung,C., Reha-Krantz,L.J., Karam,J.D., Konigsberg,W.H. (1988) *J. Biol. Chem.*, **263**, 7478–7486.

Steinum,A.-L. and Seeberg,E. (1986) *Nucleic Acids Res.*, **14**, 3763–3772.

Tye,B.K., Nyman,P.O., Lehman,I.R., Hochhauser,S. and Weiss,B. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 154–157.

Van Duin,M., van den Tol,J., Warmerdam,P., Odijk,H., Meijer,D., Westerveld,A., Bootsma,D. and Hoeijmakers,J.H.J. (1988) *Nucleic Acids Res.*, **16**, 5305–5322.

Varshney,U., Hutcheon,T. and van de Sande,J.H. (1988) *J. Biol. Chem.*, **263**, 7776–7784.

Vollberg,T.M., Cool,B.L. and Sirover,M.A. (1987) *Cancer Res.*, **47**, 123–128.

Wist,E., Unhjem,O. and Krokan,H. (1978) *Biochim. Biophys. Acta*, **520**, 253–270.

Wittwer,C.U. and Krokan,H. (1985) *Biochim. Biophys. Acta*, **832**, 308–318.

Wittwer,C.U., Bauw,G. and Krokan,H.E. (1989) *Biochemistry*, **28**, 780–784.

Wong,S.W., Wahl,A.F., Yuan,P.-M., Arai,N., Pearson,B.E., Arai,K., Korn,D., Hunkapiller,M.W. and Wang,T.S.-F. (1988) *EMBO J.*, **7**, 37–47.

Worrad,D.M. and Caradonna,S. (1988) *J. Virol.*, **62**, 4774–4777.

Received on March 3, 1989; revised on June 13, 1989