Nuclear and mitochondrial forms of human uracil-DNA glycosylase are encoded by the same gene

Geir Slupphaug, Finn-Hugo Markussen¹, Lisbeth C.Olsen¹, Rein Aasland¹, Niels Aarsæther¹, Oddmund Bakke², Hans E.Krokan* and Dag E.Helland¹ UNIGEN, Center for Molecular Biology, University of Trondheim, N-7005 Trondheim, ¹Center of Biotechnology, University of Bergen, HiB, N-5020 Bergen and ²Molecular Cell Biology, Department of Biology, University of Oslo, N-0316 Oslo, Norway

Received March 26, 1993; Revised and Accepted May 6, 1993

ABSTRACT

Recent cloning of a cDNA (UNG15) encoding human uracil-DNA glycosylase (UDG), indicated that the gene product of M. = 33.800 contains an N-terminal sequence of 77 amino acids not present in the presumed mature form of $M_r = 25,800$. This led to the hypothesis that the N-terminal sequence might be involved in intracellular targeting. To examine this hypothesis, we analysed UDG from nuclei, mitochondria and cytosol by western blotting and high resolution gel filtration. An antibody that recognises a sequence in the mature form of the UNG protein detected all three forms, indicating that they are products of the same gene. The nuclear and mitochondrial form had an apparent $M_r = 27,500$ and the cytosolic form an apparent $M_r = 38,000$ by western blotting. Gel filtration gave essentially similar estimates. An antibody with specificity towards the presequence recognised the cytosolic form of $M_r = 38,000$ only, indicating that the difference in size is due to the presequence. Immunofluorescence studies of HeLa cells clearly demonstrated that the major part of the UDG activity was localised in the nuclei. Transfection experiments with plasmids carrying full-length UNG15 cDNA or a truncated form of UNG15 encoding the presumed mature UNG protein demonstrated that the UNG presequence mediated sorting to the mitochondria, whereas UNG lacking the presequence was translocated to the nuclei. We conclude that the same gene encodes nuclear and mitochondrial uracil-DNA glycosylase and that the signals for mitochondrial translocation resides in the presequence, whereas signals for nuclear import are within the mature protein.

INTRODUCTION

In mammalian cells, the nucleic acid metabolising enzymes of the mitochondria and the nucleus are encoded by nuclear genes, synthesised in the cytoplasm, and imported into their respective organelles by means of import-specific signals (1,2). For similar enzyme activities in the nuclear and mitochondrial compartments, this may theoretically be accomplished in two ways: Either different genes encode the enzymes for the nucleus and the mitochondria, respectively, or one gene encodes a precursor protein with signals for import to both organelles. Generally it has been assumed that the first alternative applies. Uracil-DNA glycosylase (UDG), which is found in all organisms investigated (3), catalyses the first step in a DNA repair pathway which removes uracil from DNA (4). In mammalian cells, UDG has been found both in the cell nucleus and in the mitochondria (5-7).

We have previously purified (8) and cloned (9) a uracil-DNA glycosylase from human placenta (UNG-protein). The gene, UNG, which is localised on human chromosome 12 (10) belongs to a group of highly conserved uracil-DNA glycosylase genes with members in *Escherichia coli*, *Streptococcus pneumoniae* (11), *Saccharomyces cerevisiae* and the herpes virus family (9 and refs. therein). The similarity between the *E. coli* and the human enzyme is extensive (55.7% amino acid identity), and transformation with the human gene complements *E. coli ung*⁻ mutants *in vivo* (12). Two other human genes possibly encoding nuclear uracil-DNA glycosylases have recently been reported (13,14). They are, however, unrelated to the group of homologous uracil-DNA glycosylases, as well as to each other.

The nuclear form of uracil-DNA glycosylase is far more abundant than the mitochondrial form (7,15,16). Transcription of the UNG gene is up-regulated in the late G1-phase, accounting for more than 90% of the uracil-DNA glycosylase activity at the early S-phase (17). This agrees with the assumption that the UNG gene encodes a nuclear enzyme. Analysis of the UNG gene showed that it encodes a 77-amino acid N-terminal sequence not found in the purified protein (9). There are several possible explanations to this observation: 1) the active enzyme is formed by *in vivo* processing of a precursor, 2) the shorter form was a result of proteolysis during purification, or 3) the formation of different forms through alternative splicing. The latter explanation seems unlikely, since only one transcript was detected (9,17). Intriguingly, inspection of the N-terminal amino acid

^{*} To whom correspondence should be addressed

sequence revealed several clusters of positively charged residues reminiscent of a mitochondrial presequence (1). These considerations led us to hypothesise that this cDNA encodes both a nuclear and a mitochondrial form of uracil-DNA glycosylase (9). No obvious nuclear localisation signal (NLS) can be deduced from the sequence of the human UNG protein, although a cluster of positively charged amino acids at position 258–262 (Arg-Lys-Arg-His-His) may be part of such a signal.

To investigate the subcellular localisation of the UNG protein, rabbit polyclonal antibodies were raised against synthetic peptides derived from parts of the N-terminal region and the putative mature UNG protein, respectively. These antibodies were used in biochemical analysis of uracil-DNA glycosylase in different subcellular compartments of HeLa cells, in immunofluorescence analysis of endogenous uracil-DNA glycosylase in HeLa cells, and in the analysis of CV1-cells transiently expressing full-length and truncated *UNG* gene constructs.

Here we demonstrate that a uracil-DNA glycosylase encoded by a single nuclear gene in human cells, is translocated both to the nuclear and to the mitochondrial compartments. The nuclear import is facilitated by removal of a presequence carrying the signal for import to the mitochondria.

MATERIALS AND METHODS

Cell fractionation and purification of uracil-DNA glycosylase

HeLa S₃ $(3 \cdot 10^8$ cells) fractionation and DEAE-cellulose chromatography was as described by Wittwer and Krokan (7). CM Sephadex G50 chromatography was as described by Wittwer et al. (8). Prior to each gel filtration experiment, 0.4 ml of the extracts from the CM Sephadex fractionation was concentrated $10 \times$ in SpinX ultrafiltration tubes (cut-off 10 kD, MSI, Ma) according to the manufacturer's protocol, and 50 μ l of the ultrafiltered extracts were loaded onto a prepacked Superdex 75 gel filtration column (32×300 mm, Pharmacia) previously equilibrated in the same buffer as in the CM Sephadex samples. Flow was maintained at 40 μ l/min, and fractions of 80 μ l were collected using the Smart System (Pharmacia). The molecular weight markers (Sigma) used were bovine serum albumin (BSA) $(M_r=66,000)$, carbonic anhydrase $(M_r=29,000)$ and cytochrome c ($M_r = 12,700$). All steps were performed at 4°C except gel filtration, which was performed at room temperature.

Uracil-DNA glycosylase assay

UDG activity was assayed using [³H]dUMP-containing DNA (sp. act. 0.5 mCi/ μ mole) as described by Krokan and Wittwer (19), except that 1 μ l extract was used in each assay. The NaCl concentration in the reactions was generally maintained at 40 mM. One unit of UDG is the amount of enzyme that releases 1 nmole uracil from standard substrate per min at 30°C. The protein content of the extracts was measured using the Bio-Rad protein assay (Bio-Rad).

Preparation of peptide-specific antisera

Peptides corresponding to amino acids 1-21 (designated 309) and 110-129 (designated 337) of the predicted amino acid sequence encoded by UNG15 were conjugated to keyhole limpet hemocyanine using 1-ethyl-3-(dimethylaminopropyl) carbodiimide hydrochloride (EDC, Pierce) and purified by gel filtration. The conjugates were emulsified in Freund's complete adjuvant and injected into Chinchilla rabbits (Dr Carl Thomae, Germany). 200 μ g conjugate was administered subcutaneously four times at about three weeks intervals. The antisera were collected 14 days after the last injection and the serum lipids were removed using trichlorotrifluoroethane (Frigen, Hoechst). The IgG fractions were purified by affinity chromatography on protein A-agarose (Bio-Rad), and designated PU309 and PU337, respectively.

Polyacrylamide gel electrophoresis and Western analysis

SDS-PAGE was performed on the Pharmacia Phast System in 10-15% gradient gels. The molecular weight markers used were low range prestained SDS-standards (Bio-Rad) containing phosphorylase B (M_r=106,000), BSA (M_r=80,000), ovalbumin $(M_r=49,500)$, carbonic anhydrase $(M_r=32,500)$, soybean trypsin inhibitor ($M_r = 27,500$) and lysozyme ($M_r = 18,500$). Gels were contact-blotted onto Trans-Blot nitrocellulose membranes (Bio-Rad, 0.45 mm pore size) in the presence of 10 mM NaHCO₃, 3 mM Na₂CO₃, 20% methanol (pH 9.9) for 1 h at 65°C. Membranes were processed for Western analysis by blocking for 1 h in 3% BSA (Sigma, fatty-acid free). After 5 min washing in phosphate buffered saline, 0.1% Tween 20 (PBST), primary antibodies (20 µg/ml IgG in PBST) were added and incubated for 1 h. Membranes were washed for 3×10 min in PBST, and then incubated with secondary antibodies (horse radish peroxidase-conjugated swine \times rabbit, Dako, 1:1000 in PBST) for 30 min. After 3×15 min washing in PBST, membranes were incubated in ECL Western blotting detection reagent solution (Amersham) according to the manufacturers instructions and exposed to Hyperfilm-ECL (Amersham). All steps were performed at room temperature.

Plasmid constructions and cell transfection

Expression vector pSV51 is a late replacement vector with a polylinker behind the simian virus 40 (SV40) promoter (20). The vector encodes a large T-antigen. In pSVL the polylinker of pSV51 has been replaced by XbaI, SalI, SmaI, SacI and BamHI sites (U. Hamann and K. Stanley, personal communication). pSVUNG: The cDNA encoding the UNG-protein was isolated after EcoRI and HindIII digestion of pUNG15 (9). The 1144 bp fragment was blunted with Klenow DNA polymerase and ligated into the Smal site of pSVL. pSVUNG APre: The 1144 bp EcoRI-HindIII fragment of pUNG15 was subcloned in pGEM3. From the resulting vector both a 797 bp MvnI-HindIII fragment and a modified 110 bp EcoRI-AcyI fragment was isolated and inserted into the *Eco*RI-*Hind*III sites of pGEM3 yielding UNG Δ Pre in plasmid pG3UNGAPre-3. DNA sequencing revealed that in addition to the mediated deletion of Val³ through Leu⁷⁶, Ala⁷⁷ had also been removed. UNG APre was inserted into pSVL as described for UNG. The orientation of the inserts in pSVUNG and pSVUNG APre was confirmed by BamHI and by BamHI plus Xbal restriction fragment analysis of the vectors. Transient expression in CV1 simian fibroblasts were performed essentially as previously described (21), using DEAE-Dextran $(M_r = 500,000, Pharmacia)$. Briefly, following transfection, expression was permitted for 16 to 20 hours after approximately 12 hours of butyrate treatment.

Immunofluorescence analysis of uracil-DNA glycosylase localisation

HeLa cells: HeLa S3 cells were grown on glass coverslips in Dulbecco's modified Eagle's medium without sodium pyruvate, supplemented with 4500 mg/l glucose, 10% newborn calf serum (NCS), 0.3mg/ml glutamine and 0.1 mg/ml gentamicin. All

media components for cell culture were from GIBCO. Cells on coverslips were briefly washed in PBS and fixed in methanol for 15 min at -20° C. Cells were then washed in PBST for 3×5 min, blocked for 10 min in 10% normal goat serum (Dako) and incubated in $50 \mu g/ml$ PU337 in PBS + 0.5% BSA for 1 h. Cells were also incubated in parallel with equal concentrations of preimmune IgG from the same rabbit. After washing for 3×5 min in PBST, cells were incubated with biotinylated goat × rabbit secondary antibodies (Dako, 1:300 in PBST) for 30 min. Cells were then washed 3×5 min in PBST, and incubated with fluorescein-conjugated streptavidin (Dako, 1:1000 in PBST). After a final 3×5 min wash in PBST, coverslips were mounted in Citifluor (Department of Chemistry, The City University of London), and photographed using a standard Nikon immunofluorescence microscope equipped for fluorescein.

CV1 cells: Simian CV1-cells (ATTC number CCL 70) were maintained in Dulbecco's modified Eagle's medium supplemented with 10% foetal calf serum (FCS). Cells on coverslips were carefully washed in PBS and fixed for 10 min in freshly prepared 3% (weight/volume) paraformaldehyde (PFA) with 0.05% Tween 20. The cells were again washed in PBS and incubated in methanol for 10 min at -20° C. Cells were then incubated in 50 mM glycine for 5 min and washed in PBS. Fixed cell were blocked for 20 min in 10% normal sheep serum. The following steps were essentially as described for HeLa cells, except that anti-mitok (Chemicon International, Inc.) and anti-large T (Oncogene Science, Inc.) were added with the primary rabbit antibodies, and Texas Red-conjugated anti-mouse IgG (Amersham) was added with the fluorescein-conjugated secondary antibodies as required. Coverslips were mounted and photographed as described, using additional equipment for texasred fluorescence. All steps were performed at room temperature unless otherwise stated.

RESULTS

Subcellular fractionation and partial purification of uracil-DNA glycosylase

Subcellular fractionation of HeLa cells was performed in hypotonic buffers under conditions that minimises possible leakage from the nuclei (7). UDG assays indicated that 70% of the cellular UDG activity resided in the nuclei, while 21% and 9% was observed in both mitochondria and cytosol, respectively (Table I). It cannot, however, be excluded that the nuclear content of UDG may be underestimated due to leakage during fractionation. Contamination of nuclear and cytosol fractions by mitochondria was less than 2.1% and 0.1%, respectively, as estimated by cytochrome c oxidase activity in the fractions. Contamination of nuclear and mitochondrial fractions by cytosol was less than 1.8% and 1.2%, respectively, as estimated from NAD(P)H dehydrogenase activity. No nuclei or nuclear fragments were observed in the mitochondrial fraction using phase contrast microscopy, indicating that possible contamination of the mitochondrial fraction by nuclei was very low.

Extracts from mitochondria and nuclei, as well as cytosol, were subjected to chromatography on DEAE-cellulose and CM-Sephadex. To avoid formation of complexes of UDG with other proteins, the extracts were adjusted to final concentrations of 1 M NaCl and 10% glycerol prior to separation by gel filtration on Superdex 75 to maintain UDG in a disaggregated state (7). Nuclear and mitochondrial activities eluted after identical

 Table I. Partial purification of uracil-DNA glycosylase from different subcellular compartments^a

Fraction	Total activity (units)	Specific act. (units/mg prot.)	Yield (%)	Purification (fold)
Nuclei				
Crude extract	23.4	0.69	100	-
DEAE-cellulose	18.3	1.06	78	1.5
CM-Sephadex	17. 9	13.9	76	20.1
Superdex 75	3.7	324	16	470
Mitochondria				
Crude extract	7.1	0.45	100	-
DEAE-cellulose	5.6	1.46	79	3.2
CM-Sephadex	4.9	12.3	69	27.3
Superdex 75	2.8	225	39	500
Cytosol				
Crude extract	3.0	0.07	100	-
DEAE-cellulose	2.8	0.14	93	2.0
CM-Sephadex	1.9	1.9	63	27.1
Superdex 75	0.3	27.6	10	394

^aOnly parts of the fractions eluted from CM-sephadex were subjected to gel filtration on Superdex 75. Values given in the table correspond to the total volumes.



Figure 1. High resolution gel filtration of UDG extracts on Superdex 75 after subcellular fractionation. A) Nuclear extract, B) mitochondrial extract and C) cytosplasmic extract. Fractions containing UDG activity are indicated by shading. Molecular weight markers: a) Bovine serum albumin (M_r =66,000); b) carbonic anhydrase (M_r =29,000) and c) cytochrome c (M_r =12,700).

volumes, corresponding to an apparent $M_r = 24,000$ as calculated from a standard curve using molecular weight markers. In contrast, the cytosol activity eluted corresponding to an apparent $M_r = 36,000$ (Figure 1). Molecular weights are close to the molecular weights of the putative mature form and the preform of uracil-DNA glycosylase, respectively, as predicted from their amino acid sequences (9).

Maximum activity of the nuclear and mitochondrial UDG fractions was observed in the presence of 40 mM NaCl (data not shown), while the corresponding cytoplasmic UDG activity peaked in the presence of 60 mM NaCl. Furthermore, *in vitro*-transcription/translation experiments indicated that the specific activity of the UNG-protein closely corresponding to the mature form, is 4-5 times higher than the corresponding full-length protein as determined from densitometric scanning of bands



Figure 2. Shematic illustration of the full-length UNG-protein and the corresponding UNG Δ Pre-protein lacking amino acids Val³ through Ala⁷⁷. The positioning of the synthetic peptides 309 and 337 are as indicated.



Figure 3. Western analysis of subcellular fractions of UDG. N: nuclear extracts; M: mitochondrial extracts; C: cytosolic extracts. Pre337 and Pre309: Pre-immune IgGs from the same rabbits that produced PU337 and PU309, respectively. Molecular weights of markers are as indicated.

from [³⁵S]methionine-labelled translation products after polyacrylamide gel electrophoresis and fluorography (data not shown).

Identification of subcellular forms of uracil-DNA glycosylases using peptide-specific antibodies

To verify the nature of the uracil-DNA glycosylases in the different subcellular compartments, rabbit polyclonal antisera were raised against a panel of 5 synthetic peptides derived from the putative mature- and full length uracil-DNA glycosylases. Peptides were chosen on basis of their calculated hydrophilicity and their antigenic index (22). Two of these sera, having titers above 1:15,000, were chosen from their specificity in enzymelinked immunosorbent assays (ELISA). No specific binding was observed when the sera were tested against four control peptides of identical length (data not shown). Purified IgG fractions of the antisera were designated PU337 and PU309. The sequences of the peptides used for immunisation and their positions in the amino acid sequence of UNG is shown in Figure 2. The antibodies did not recognise native uracil-DNA glycosylase in immunoprecipitation experiments, suggesting that the epitopes in the native UNG sequence are inaccessible or have conformations different from those in the corresponding synthetic peptides.



Figure 4. a) Immunocytochemical preparation of non-confluent HeLa cells showing specific staining of the nuclei by the polyclonal antibody PU337. b) HeLa cells treated as in a), but labelled with pre-immune IgGs from the same rabbit. Note that the nuclei are not stained. Immunofluorescence microscopy ×400.

Western analysis of the different UDG preparations using PU337 indicated an $M_r=27,500$ for both the nuclear and mitochondrial form, whereas UDG from the cytosol had an apparent $M_r=38,000$. These experiments are in agreement with the results from the gel filtration experiments (Figure 1) and indicate that the cytosol contains a preform of the nuclear and mitochondrial uracil-DNA glycosylases. Western analysis using PU309 directed against a sequence in the putative N-terminal presequence recognised the cytosol form of UDG only, strongly suggesting that a preform of UDG is found in the cytosol. Using PU337, two weaker bands of lower molecular weight could also be observed in the cytosol fraction, possibly resulting from degradation of the glycosylase.

Functional analysis of the N-terminal sequence of uracil-DNA glycosylase

To study the function of the N-terminal sequence further, immunocytochemical analyses of HeLa cells and transfection of CV1 cells were performed. Strong fluorescence over the nuclei in HeLa cells was observed using PU337 (Figure 4a) as compared to the weak nuclear fluorescence observed using pre-immune serum from the same rabbit (Figure 4b). Some fluorescence was also observed over the cytoplasm, but most of this is believed to be unspecific, as it also appeared in the control cells. We did not observe a distinct mitochondrial staining in these expreiments, although some extranuclear staining is evident. The lack of mitochondrial staining may be due to smaller amounts of UNG protein present in the mitochondria. However, the experiments strongly indicate that the UDG activity located in the nucleus is encoded by the UNG gene.

To examine the function of the presequence more directly, CV1-cells were transfected with plasmids encoding full-length and partially deleted UNG proteins. For the expression of UNG, we constructed the vectors pSVUNG and pSVUNG Δ Pre using pSVL, a modification of the late replacement vector pSV51,

Figure 5. Full length UNG-protein is localised in the mitochondria and UNG Δ Pre-protein in the nucleus. **a**) Staining of the mitochondria in a cell transiently expressing UNG and labelled with PU337 and fluorescein-conjugated antibodies. **b**) Same field as in a) labelled with monoclonal antibodies against mitochondria (anti-mitok) and monoclonals angainst SV40 large T-antigen (anti-large T), both visualised by Texas red-conjugated anti-mouse secondary antibodies. Compare the mitochondrial pattern of the transfected cell with the pattern in a). Note that non-transfected cells (see arrowhead) adjacent to the transfected cell show no largeT-antigen nuclear staining. **c**) Staining of double nucleus in a cell transiently expressing UNG Δ Pre. The area close to the nuclei is also stained. Adjacent cells were not transfected (they were not stained by anti-large T, data not shown) and consequently were not stained specifically by PU337. Immunofluorescence microscopy×600.

Figure 6. Structural similarity with N-terminal signal sequences of eukaryotic nucleic acid metabolizing enzymes encoded by one nuclear gene, and shared by the mitochondria and the nucleus—the yeast UNG1 being putatively so. (+) denotes a basic residue. Yeast PHR1: (29); human UNG: (9).

which has been shown to give high transient expression of several proteins in simian cells (20). pSVUNG contained the complete open reading frame including the 5' non-coding region of the UNG cDNA, while in pSVUNG Δ Pre, the codons for amino acids Val³ through Ala⁷⁷ of the presequence were deleted (Figure 2). Immunolabelling with PU337 IgG of CV1 cells transiently expressing full-length UNG protein demonstrated that the protein was translocated to the mitochondria (Figure 5a and b). However, the results from western analysis with PU337 and PU309 IgGs (Figure 3) indicated that the presequence is removed in the translocation process, but do not demonstrate where this removal takes place. By contrast, the UNG Δ Pre protein became concentrated in the nuclei (Figure 5c). The data presented in Figure 5c represent cells with a relatively low expression of the UNG Δ Pre protein. In cells with a higher level of expression, the nuclei as well as the cytoplasm were stained (data not shown), suggesting that the nuclear sorting apparatus was saturated and/or not properly regulated. This was not seen for the pSVL-encoded simian virus 40 large T-antigen which carries a highly efficient nuclear signal (23), as it was localised solely in the nucleus.

Finally, the specificity of PU337 in immunocytochemical labelling of methanol-fixed cells was confirmed by the specific staining of CV1-cells transiently expressing full length human UNG-protein, leaving adjacent non-transfected cells unstained (Figure 5).

In summary, the immunocytochemical labelling experiments strongly support the findings from the cell fractionation and western analyses in that the human UNG gene encodes a precursor protein giving rise to both nuclear and mitochondrial protein forms. Furthermore, the signals for mitochondrial translocation resides in the presequence, whereas signals for nuclear import are within the mature protein.

DISCUSSION

Signals for mitochondrial localisation are generally located at the N-terminus of mitochondrial proteins encoded by nuclear genes (1), while nuclear signals (NLS) can be located virtually anywhere in a protein (2). Mitochondrial import of proteins involves removal of the N-terminal targeting sequence by a mitochondrial

processing protease in the matrix (1) or in the inner mitochondrial membrane (24). Hence, a mitochondrial UNG-protein of molecular weight lower than the $M_r=33,800$ full-length protein (9) is expected. The apparent molecular weights of the two glycosylases as determined by SDS-PAGE and western analysis are somewhat larger than predicted by their amino acid sequences specified by *UNG*. This was also observed for the mature placental UNG protein after purification to near homogeneity (8,9). However, the low molecular weight amino acids are rather highly represented along the UNG amino acid sequence, being especially prominent in the N-terminal presequence, and this may give rise to larger apparent M_r in SDS-PAGE. This has also recently been observed for another small basic DNA-repair protein (25).

Compartmentalization of proteins carrying both nuclear and mitochondrial targeting signals has been studied using artificial hybrid proteins. When the SV40 large T-antigen NLS was fused to the N-terminus of yeast cytochrome c₁, which carries a typical mitochondrial signal, the hybrid protein was directed to the nucleus (26). By contrast, the yeast nuclear protein HAP2 fused to the C-terminus of the mitochondrial protein HEM15 was localised to mitochondria (M. Haldi and L. Guarente, personal communication) indicating that the most N-terminal signal dominates. Intramolecular masking of the NLS in the precursor protein is also possible. This has recently been shown for the p110 precursor of the NF-xB subunit, in which the NLS is masked by a C-terminal domain, preventing nuclear uptake of p110 (27). It is therefore likely that the mature nuclear UNGprotein is localised to the nucleus after processing of the dominating/masking presequence by a protease. As judged from the N-terminus of the purified UNG-protein (8), processing most likely takes place between Ala⁷⁷ and Ala⁷⁸. It should be noted, that full-length UNG-protein transiently expressed in simian CV1-cells was not detected in the nucleus (Figure 5a), although it must carry both signals initially, as does the endogenous UNGprotein of HeLa cells which was localised in the nucleus (Figure 4). The reason for this is not clear, but one explanation may be that processing of the presequence was not efficient after transient expression in the simian cells, as all the UNG-protein was detected in the mitochondria. Search for a proteolytic activity specifically cleaving the UNG-protein is currently being undertaken in our laboratory.

The yeast photolyase, encoded by the nuclear PHR1 gene, is also apparently localised to both the mitochondria and the nucleus (28). Alignment of the PHR1 gene and the E. coli phr gene revealed that the former gene encodes a N-terminal presequence of 74 amino acids (29) reminiscent of a mitochondrial signal (28) not found in the bacterial protein. This feature is also similar to the uracil-DNA glycosylase of the yeast nuclear UNG1 gene, which has a 99 amino acid N-terminal presequence not found in the E. coli UNG protein (30). It contains features of mitochondrial signals; clusters of basic amino acids and enrichment of hydroxylated residues (1). Interestingly, the yeast UNG1 gene encodes a Mr=39,500 protein which is degraded to a protein of $M_r = 33,000$ which was not a product of in vitro proteolytic degradation (30). These features are similar to those of the human UNG-protein, which was synthesised as a $M_r = 34,000$ protein in vitro (9,17), but was purified as a $M_r = 29,000$ protein from placenta (8). This strongly suggests, the putative mitochondrial signal of the yeast UNG1 gene product taken into account, that also the UNGI gene encodes a protein shared by mitochondria and nuclei. Common to these three

eukaryotic proteins are high levels of similarity to prokaryotic homologues, and the extended N-termini harbouring mitochondrial signals (putative in the case of UNG1, Figure 6). In addition, the yeast nuclear *TRM1* gene encodes a tRNA methyltransferase sorted to both mitochondria and nuclei. The 213 N-terminal amino acids of TRM1 are capable of targeting heterologous proteins to mitochondria (residues 1-48) (31) and to the nucleus (residues 70-213) (32). A prokaryotic gene with similarities to *TRM1* has to our knowledge not been cloned.

ACKNOWLEDGEMENTS

We thank Dr Rune Male for his help with the preparation of the photographs and Dr Jon A.Eriksen for supplying purified peptides for immunizations. This work was supported by grants from the Norwegian Cancer Society and the Norwegian Society for Science and the Humanities.

REFERENCES

- 1. Hartl, F.U and Neupert, W. (1990) Science, 247, 930-938.
- Garcia-Bustos, J., Heitman, J. and Hall, M.N. (1991) Biochim. Biophys. Acta, 1071, 83-101.
- 3. Tomilin, N.V. and Aprelikova, O.N. (1989) Int. Rev. Cyt., 114, 125-179.
- 4. Dianov, G., Price, A. and Lindahl, T. (1992) Mol. Cell. Biol., 12, 1605-1612.
- Anderson, C.T.M. and Friedberg, E.C. (1980) Nucleic Acids Res., 8, 875-888.
- 6. Domena, J.D. and Mosbaugh, D.W. (1985) Biochemistry, 24, 7320-7328.
- 7. Wittwer, C.U. and Krokan, H. (1985) Biochim. Biophys. Acta, 832, 308-318.
- Wittwer, C.U., Bauw, G. and Krokan, H.E. (1989). Biochemisty, 28, 780-784.
- Olsen, L.C., Aasland, R., Wittwer, C.U., Krokan, H.E. and Helland, D.E. (1989) EMBO J., 8, 3121-3125.
- Aasland, R., Olsen, L.C., Spurr, N.K., Krokan, H.E. and Helland, D.E. (1990) Genomics, 7, 139-141.
- 11. Méjean, V., Rives, I. and Claverys, J.-P. (1990) Nucleic Acids Res., 18, 6693.
- Olsen, L.C., Aasland, R., Krokan, H.E. and Helland, D.E. (1991) Nucleic Acids Res., 19, 4473-4478.
- Meyer-Siegler, K., Mauro, D.J., Seal, G., Wurzer, J., deRiel, J.K., and Sirover, M. (1991) Proc. Natl. Acad. Sci. USA, 88, 8460-8464.
- 14. Muller, S.J. and Caradonna, S. (1991) Biochim. Biophys. Acta, 1088, 197-202.
- 15. Gupta, P.K. and Sirover, M.A. (1981) Cancer Res., 41, 3133-3136.
- 16. Colson, P and Verly, W.G. (1983) Eur. J. Biochem., 134, 415-420.
- Slupphaug,G., Olsen,L.C., Helland,D., Aasland,R. and Krokan,H.E. (1991) Nucleic Acids Res., 19, 5131-5137.
- 18. Krokan, K., Bjørklid, E. and Prydz, H. (1975) Biochemistry, 14, 4227-4232.
- 19. Krokan, H.E. and Wittwer, C.U. (1981) Nucleic Acids Res., 9, 2599-2613.
- Huylebroek, D. Maertens, G., Verhoeyen, M., Lopez, C., Raeymakers, A. and Fiers, W, Gene, 66, 163-181.
- 21. Bakke, O. and Dobberstein, B. (1990) Cell, 63, 707-716.
- 22. Jameson, B. and Wolf, H. (1988) CABIOS, 4, 181-186.
- Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984) Cell, 39, 499-509.
- Braun,H.-O., Emmermann,M., Kruft,V. and Schmitz,K. (1992) EMBO J., 11, 3219-3227.
- 25. Rydberg, B., Spurr, N. and Karran, K. (1990) J. Biol. Chem., 265, 9563-9569.
- Sadler, I., Chiang, A., Kurihara, T., Rothblatt, J., Way, J. and Silver, P. (1989) J. Cell Biol., 109, 2665-2675.
- Henkel, T., Zabel, U., van Zee, K., Müller, J.M., Fanning, E. and Baeuerle, P.A. (1992) Cell, 68, 1121-1133.
- Yasui, A., Yajima, H., Kobayashi, T., Eker, A.P.M. and Oikawa, A. (1992) Mutation Res., 273, 231-236.
- 29. Sancar, G.B. (1985) Nucleic Acids Res., 13, 8231-8246.
- Percival, K.J., Klein, M.B. and Burgers, P.M.J (1989) J. Biol. Chem., 264, 2593-2598.
- 31. Ellis, S.R., Hopper, A.K. and Martin, N.C. (1989) Mol. Cell. Biol., 9, 1611-1620.
- 32. Li,J.M., Hopper,A.K. and Martin,N.C. (1989) J. Cell. Biol., 109, 1411-1419.