The yeast *CDC9* gene encodes both a nuclear and a mitochondrial form of DNA ligase I

M. Willer, M. Rainey, T. Pullen and C.J. Stirling

Background: The yeast *CDC9* gene encodes a DNA ligase I activity required during nuclear DNA replication to ligate the Okazaki fragments formed when the lagging DNA strand is synthesised. The only other DNA ligase predicted from the yeast genome sequence, *DNL4/LIG4*, is specifically involved in a non-homologous DNA end-joining reaction. What then is the source of the DNA ligase activity required for replication of the yeast mitochondrial genome?

Results: We report that *CDC9* encodes two distinct polypeptides expressed from consecutive in-frame AUG codons. Translational initiation at these two sites gives rise to polypeptides differing by a 23 residue amino-terminal extension, which corresponds to a functional mitochondrial pre-sequence sufficient to direct import into yeast mitochondria. Initiation at the first AUG codon results in a 755 amino-acid polypeptide that is imported into mitochondria, whereupon the pre-sequence is proteolytically removed to yield the mature mitochondrial form of Cdc9p. Initiation at the second AUG codon produces a 732 amino-acid polypeptide, which is localised to the nucleus. Cells expressing only the nuclear isoform were found to be specifically defective in the maintenance of the mitochondrial genome.

Conclusions: *CDC9* encodes two distinct forms of DNA ligase I. The first is targeted to the mitochondrion and is required for propagation and maintenance of mitochondrial DNA, the second localises to the nucleus and is sufficient for the essential cell-division function associated with this gene.

Address: School of Biological Sciences, 2.205 Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK.

Correspondence: C.J. Stirling
E-mail: Colin.Stirling@man.ac.uk

Received: 19 July 1999
Revised: 18 August 1999
Accepted: 19 August 1999

Published: 23 September 1999

Current Biology 1999, 9:1085–1094

0960-9822/99/$ – see front matter

Background

DNA ligases have important roles in several aspects of DNA metabolism including replication, recombination and repair. In mammalian nuclear extracts, four different classes of ligase have been identified. DNA ligase I functions in the ligation of Okazaki fragments, which are formed when the lagging strand of DNA is synthesised during replication [1,2], and has been further implicated in DNA repair [3]. DNA ligase III interacts with the XRCC1 protein, which is known to be involved in both recombination and repair, thus implicating DNA ligase III in these processes [4,5]. A testis-specific splice variant of DNA ligase III has also been identified, which does not interact with XRCC1 and whose specific role is currently unknown [6]. DNA ligase II is smaller than DNA ligase III but the degree of structural similarity between the two has led to the suggestion that DNA ligase II might correspond to either a splice variant or a proteolytic derivative of DNA ligase III [7,8]. Finally, DNA ligase IV has a different substrate specificity from DNA ligases I and III and is essential in mice, but its biological role has not yet been defined [9,10]. In *Saccharomyces cerevisiae*, the *CDC9* gene encodes a DNA ligase I that is essential for completion of DNA synthesis [11]. Conditional lethal *cdc9* mutant cells are defective in ligation of Okazaki fragments, are hypersensitive to DNA-damaging agents, and show enhanced frequency of mitotic recombination. Consistent with its role in DNA replication, the expression of *CDC9* has been shown to be cell cycle regulated in exponentially growing cells with its transcript level reaching a maximum at the G1–S boundary [12,13]. Nevertheless, expression of *CDC9* can also be induced in stationary-phase cells upon exposure to DNA-damaging agents, thereby satisfying the requirement for Cdc9p in DNA repair [11,14].

A second DNA ligase with DNA end-joining activity has been identified in yeast cell extracts [15] and seems likely to correspond to the product of an open reading frame (ORF) designated YOR005c, identified in the yeast genome as a homologue of mammalian DNA ligase IV. In three independent studies using reverse genetic analysis [16–18], it has been demonstrated that YOR005c does indeed encode a ligase activity; the ORF has been renamed *DNL4* [18] and *LIG4* [16,17]. The *DNL4/LIG4* gene predicts a 109 kDa polypeptide product that shares 25% sequence identity with mammalian DNA ligase IV but only 21% identity with yeast Cdc9p. Unlike *CDC9*, the *DNL4/LIG4* gene is not essential for growth; indeed null mutant cells exhibit no obvious defects in either DNA replication or excision-repair [16–18]. Nevertheless,

dnl4/lig4 mutant cells have been shown to sustain a substantially reduced level of non-homologous DNA end-joining, suggesting a role for Dnl4p/Lig4p in this process *in vivo* [16,18].

Given its role in nuclear DNA replication, it is not surprising that the mammalian DNA ligase I has been shown to be located in the nucleus [19]. The absence of any further ligase homologues in the yeast genome, however, raises an important question as to the source of DNA-ligase activity that would be required for the replication of the mitochondrial genome. The yeast mitochondrial genome is an 85.78 kb covalently closed circle encoding a small number of mitochondrial proteins, none of which show any similarity to known ligases. As *CDC9* encodes the only replication-specific ligase currently known in yeast, we undertook a detailed analysis of the expression and intracellular localisation of the *CDC9* gene product. We report that *CDC9* encodes two distinct polypeptide products that are translated from separate in-frame initiation codons. The longer form is targeted to the mitochondrion where it is required for mitochondrial DNA replication. The shorter polypeptide product is targeted to the nucleus and is essential for nuclear DNA replication and cell viability.

Results

Expression and localisation of Cdc9–GFP fusion proteins

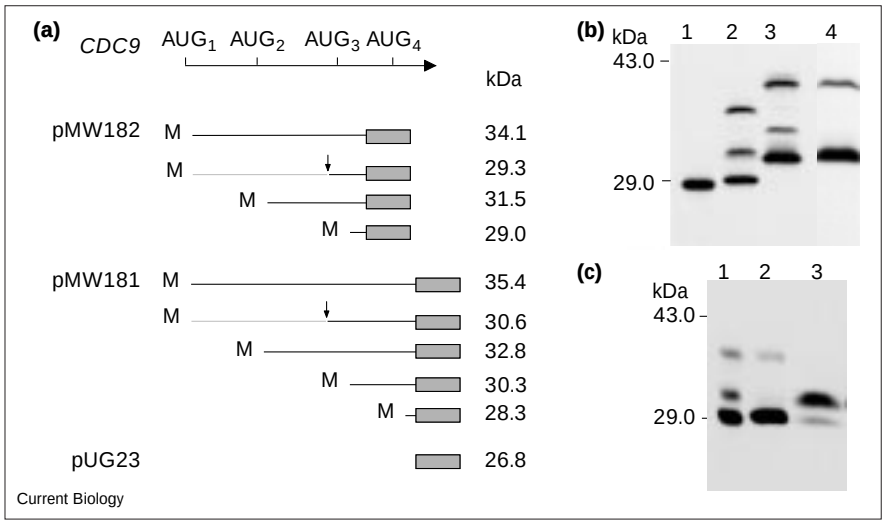
Previous studies have shown that transcription of *CDC9* initiates 39–41 bp upstream of the first AUG codon (AUG_1) in the ORF [11,14]. Translation initiation at AUG_1 would yield a 755 amino-acid polypeptide with a predicted molecular weight of 84.8 kDa [11]. Analysis of the *CDC9* sequence, however, revealed several further in-frame methionine codons including AUG_2, AUG_3 and AUG_4 located at codons +24, +48 and +66, respectively (Figure 1a). To determine which, if any, of these four potential start codons might initiate translation of Cdc9p *in vivo*, a series of GFP fusion proteins were generated and the products detected by western blotting with anti-GFP antibodies (Figure 2). Cells transformed with a vector expressing GFP alone (pUG23) contained a single immunoreactive band of the expected relative molecular weight of 26.8 kDa. Cells carrying a plasmid in which GFP was fused after AUG_3 (pMW182) contained three immunoreactive species of 34.1, 31.5 and 29.0 kDa, consistent with initiation at AUG_1, AUG_2 and AUG_3, respectively (Figure 2b). Similarly, three distinct species were observed when GFP was fused after codon 70 downstream of AUG_4 (in plasmid pMW181). The relative molecular weights of these three species (35.4, 32.8, 30.3 kDa) were again consistent with initiation at each of the first three AUG codons. No species was detected with the mobility expected from initiation at the fourth AUG.

We noted that residues 1–23 of Cdc9p could form an amphipathic α helix with properties similar to those commonly found in mitochondrial pre-sequences (see Figure 1b) [20]. Should this constitute a functional pre-sequence, then full-length Cdc9p might be predicted to be targeted to the mitochondrial matrix where the pre-sequence might be removed by mitochondrial pre-sequence peptidase (MPP) [21]. Indeed, a consensus cleavage site for MPP cleavage exists in the sequence ARFFT (in the single-letter amino-acid code; amino-acid residues 42–46), predicting cleavage between residues F44 and F45 (Figure 1a; [22]). Such a cleavage product would be very similar in molecular weight to any *bona fide* translation product initiated from the nearby AUG_3. To distinguish these possibilities, we mutated both the AUG_2 and AUG_3 codons in the AUG_{1-4} construct to GCG (Ala) and then examined the expression profile *in vivo* ($AUG_{1,4}$–GFP; Figure 2b). Cells carrying this construct (pMW197) were found to express two major bands with gel mobilities indistinguishable from the 35.4 and 30.3 kDa forms observed with the wild-type AUG_{1-4}-GFP construct; the 32.8 kDa form predicted to initiate at AUG_2 was lacking, however (Figure 2b; lanes 3,4). The presence of the 30.3 kDa form in pMW197-containing cells

Figure 1

Structure of the *CDC9* ORF. (a) Predicted protein sequence. Asterisks, the four amino-terminal methionine codons (AUG_{1-4}); arrow, consensus site for possible cleavage by the MPP [22] (the location of the cleavage site between residues 44 and 45 suggests that the first 44 amino-terminal residues (boxed) might represent a mitochondrial pre-sequence); underlined sequences, nuclear localisation signals [22]; inverted black triangle, position of the triple HA epitope tag inserted in strain MDY26 (see Materials and methods). (b) Alpha-helical wheel of the first 18 residues (shown in single-letter amino acid code) of the predicted Cdc9 protein reveals an amphipathic structure characteristic of mitochondrial targeting sequences [20]. Black circles, hydrophobic residues; +, positively charged residues.

Figure 2



Cdc9–GFP constructs are translated from multiple start codons and proteolytically processed in mitochondria. (a) The 5′ end of the *CDC9* ORF is shown schematically with the positions of AUG_{1-4} indicated to scale. The structure of potential GFP fusion proteins expressed from AUG_{1-3} in pMW182 and AUG_{1-4} in pMW181 are shown, together with their predicted molecular weights. M, initiating methionine residue; grey box, GFP; arrow, putative site of mitochondrial pre-sequence processing. The AUG_2 and AUG_3 codons in pMW182 were both mutated to GCG to produce pMW197 ($AUG_{1,4}$) as described in the Materials and methods section. (b) Western blot of total cell extracts from yeast expressing Cdc9–GFP constructs, probed with anti-GFP antibody. Lane 1, pUG23 (no insert); lane 2, pMW182 (AUG_{1-3}); lane 3, pMW181 (AUG_{1-4}); lane 4, pMW197 ($AUG_{1,4}$). (c) Subcellular fractionation of yeast expressing the Cdc9–GFP fusion proteins encoded by pMW182 (AUG_{1-3}). Lane 1, total extract; lane 2, mitochondrial fraction; lane 3, post-mitochondrial supernatant.

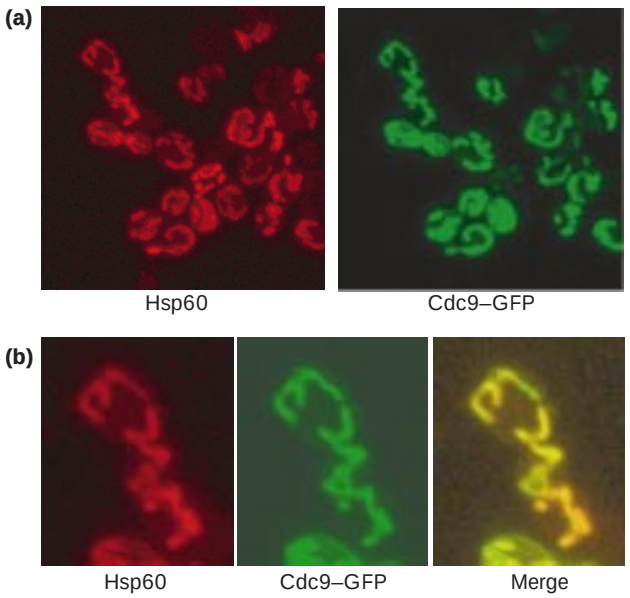
indicates that it cannot be a product of initiation at AUG_3, but rather suggests that it must be a proteolytic product of the full-length polypeptide expressed from AUG_1. If this processing was by MPP, then the 30.3 kDa form would be expected to fractionate with mitochondria. Therefore, purified mitochondria and a cytosolic fraction were prepared from cells carrying the AUG_{1-3}-GFP fusion construct (pMW182) and analysed by western blotting (Figure 2c). The 34.1 and 29.0 kDa forms were found to cofractionate with purified mitochondria, whereas the 31.5 kDa form was exclusively located in the cytosolic fraction. The small proportion of 29.0 kDa species present in the cytosol might have been released from broken mitochondria or might reflect very limited translational initiation from AUG_3.

To confirm that cofractionation of Cdc9–GFP with purified mitochondria reflects authentic localisation to this compartment, we next examined the patterns of GFP fluorescence in intact cells. Cells carrying pMW181 (AUG_{1-4}–GFP) were first compared, using conventional fluorescence microscopy, with those expressing GFP alone (pUG23). Cells carrying the GFP vector exhibited a strong generalised background staining consistent with a predominantly cytoplasmic localisation. In contrast, cells expressing Cdc9–GFP exhibited a strong pattern of green fluorescence localised to tubular and punctate structures, with a weak general background fluorescence (see Supplementary material published with this article on the internet). Control cells lacking GFP exhibited no fluorescence signal (data not shown). There was considerable correspondence between the tubular/punctate staining observed with Cdc9–GFP and extranuclear 4,6-diamidino-2-phenylindole (DAPI) staining of mitochondrial DNA, consistent with a mitochondrial localisation for Cdc9–GFP (see Supplementary material). This was confirmed by colocalisation of Cdc9–GFP with the mitochondrial heat-shock protein Hsp60, observed by confocal immunofluorescence microscopy (Figure 3a,b). The Hsp60 staining was strikingly similar to the pattern of Cdc9–GFP fluorescence observed in these cells (Figure 3a). Indeed, a merged image of Hsp60 and Cdc9–GFP staining in a dividing cell showed precise colocalisation of the two signals (Figure 3b). These data clearly indicate that the amino-terminal region of Cdc9p does indeed contain a functional mitochondrial pre-sequence. Moreover, the data indicate that both AUG_1 and AUG_2 in yeast *CDC9* can initiate translation of a GFP fusion protein *in vivo*, and suggest that the longer translation product is targeted to the mitochondrion where it is processed by MPP.

Epitope tagging of full-length Cdc9

To detect the yeast *CDC9* gene product, we first introduced a triple HA tag using the transposon insertion/excision method developed by Ross-MacDonald *et al.* ([23]; Cdc9–HA). Several insertions were obtained that did not interfere with the normal growth rates of haploid strains, indicating that these tagged alleles retained full Cdc9 activity. Western blot analysis revealed that, in all cases, these cells contained a doublet of immunoreactive bands with relative gel mobilities of approximately 90–93 kDa. This is marginally smaller than the 94.9 kDa predicted from the sequence of the epitope-tagged Cdc9–HA. The results for one such insertion are shown in Figure 4a; this particular strain (MDY26) was used in the remaining experiments.

Figure 3



Cdc9–GFP localises to mitochondria in yeast. (a) Confocal immunofluorescence microscopy of cells carrying pMW181. The Cdc9–GFP signal shows a similar distribution to the staining found with indirect immunofluorescence after incubation with anti-Hsp60 antisera followed by a Cy3 (indocarbocyanine)-conjugated secondary antibody. Cells were fixed and stained as described in the Materials and methods section. The images shown correspond to an integration of 25 separate confocal sections recorded in either the GFP or Cy3 channels. (b) Enlarged area from (a) demonstrating precise overlap between the mitochondrial Hsp60 and Cdc9–GFP signals.

Intracellular distribution of Cdc9–HA

Next, we determined the localisation of the functional Cdc9–HA protein by sub-cellular fractionation. Mitochondria were purified by differential centrifugation as described [24]. The mitochondrial fraction was shown to be devoid of nuclear contamination by western blotting with antibodies against several nuclear proteins including Npl3 (Figure 4b), Swi3 and histone H3 (data not shown). As expected, the mitochondrial fraction was found to be enriched for a variety of mitochondrial proteins examined, namely Hsp60 (Figure 4b), Mdh1 and Porin (data not shown). Western blotting with anti-HA antibodies revealed that the mitochondrial fraction contained significant levels of Cdc9–HA (Figure 4b).

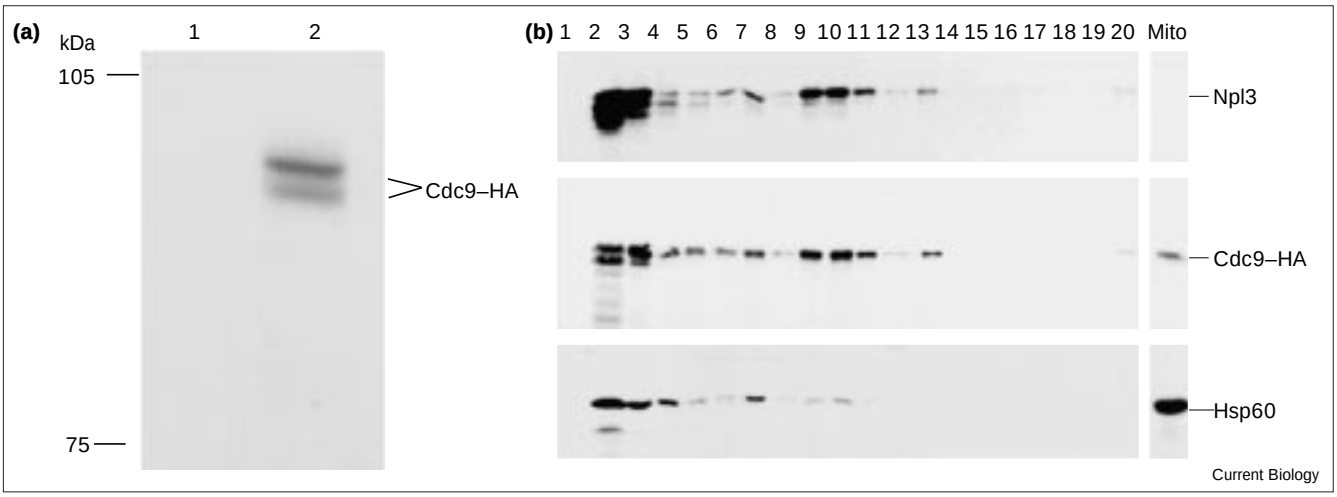
Enriched nuclear preparations were resolved by sucrose step-gradient centrifugation. As previously reported [25], we found that intact nuclei, observed microscopically, were present at the interface between 1.5–1.8 M sucrose, corresponding to a peak of Npl3 protein detected by western blotting (Figure 4b, lanes 9–11). A substantial amount of Npl3 was also detected at the top of the gradient (Figure 4b, lanes 2–3). This might reflect a cytosolic pool of Npl3, but might be largely due to a combination of nuclear breakage and the leakage of proteins through nuclear pores during the fractionation procedure. The distribution of other nuclear markers tested, Swi3 and histone H3, were identical to that of Npl3 (data not shown). As expected, there was a distinct peak of Cdc9–HA that precisely coincided with the peak of intact nuclei (Figure 4b, lanes 9–11), indicating that a proportion of Cdc9–HA is localised to the nucleus. The fractions containing intact nuclei showed marginal contamination with mitochondrial Hsp60, which was predominantly located towards the top of this gradient (Figure 4b, lanes 2–3). Identical staining profiles were also seen with other mitochondrial markers Mdh1 and Porin (data not shown).

The nuclear and mitochondrial forms of Cdc9–HA exhibit different properties

Copurification of Cdc9–HA with mitochondria is consistent with the localisation observed for Cdc9–GFP. To confirm that Cdc9–HA is imported into mitochondria, as opposed to being associated with the outer surface of the organelle, we performed protease-protection studies. Purified nuclei and mitochondria were separately incubated with proteinase K in the presence or absence of the detergent Triton X-100. Following incubation, extracts were analysed by western blotting with anti-HA antibody to detect any undigested Cdc9–HA protein. The mitochondrion-associated form of Cdc9–HA was found to be protected from proteolysis in the absence of detergent but was completely sensitive to proteinase K digestion when incubations were performed in the presence of detergent (Figure 5a). Identical results were obtained for the mitochondrial matrix protein Hsp60 (Figure 5b). As a control, the purified mitochondria were shown to be devoid of contamination with a nuclear marker (Figure 5c). These results indicate that the mitochondrion-associated form of Cdc9–HA is entirely sequestered within a membrane-bound compartment.

The protease accessibility of nucleus-associated Cdc9–HA was found to be very different. In this case, Cdc9–HA remained stable during mock incubations in the absence of protease, but was digested upon incubation with protease either in the presence or absence of detergent (Figure 5d). Identical results were obtained for Npl3 (Figure 5f), consistent with free diffusion of small proteins such as proteinase K into the nuclear compartment through nuclear pores. As discussed previously, we found minor contamination of the nuclear fraction with mitochondrial Hsp60 and, once again, we found that this was protected from proteolysis in the absence of detergent (Figure 5e). These results indicate that the nucleus-associated and mitochondrion-associated forms of Cdc9-HA are topologically distinct. Moreover, we have noted that the two forms of Cdc9–HA exhibit different

Figure 4



Expression of Cdc9–HA in yeast. (a) Western blot with anti-HA antibodies on total extracts prepared from a wild-type haploid yeast strain (W303-3c, lane 1) and from the MDY26 strain (see Figure 1a) in which sequences encoding the triple HA tag have been inserted into the *CDC9* gene (lane 2). MDY26 cells express an immunoreactive doublet of apparent molecular weights 90 and 93 kDa; these are only very marginally smaller than the size predicted for full-length Cdc9–HA (94.9 kDa). (b) Cdc9–HA is found in both nuclear and mitochondrial fractions. Intact nuclei were prepared from MDY26 and purified by centrifugation on a sucrose step-gradient as described in Materials and methods. Equal volumes of 20 individual fractions (lanes 1–20) were subjected to western blotting with antibodies specific to Npl3, HA, or Hsp60. Mitochondria (mito) were purified from the same cells (see the Materials and methods section) and equivalent amounts of these were then analysed by western blotting using the same three antibodies.

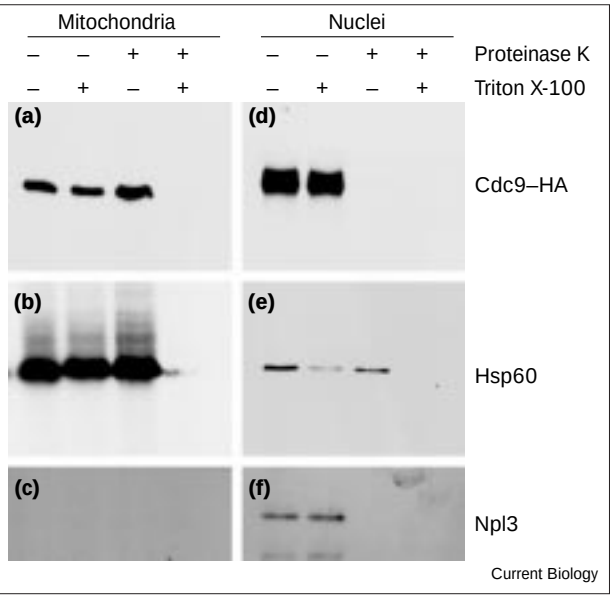
gel mobilities. The nuclear form migrated more slowly in SDS–PAGE than did the mitochondrial form and these two distinct species corresponded closely to the bands of the Cdc9–HA doublet observed in the total cell extract (Figure 6). Taken together, these data indicate the presence of two distinct forms of Cdc9–HA expressed in MDY26 cells. The larger form is localised in the nuclear compartment, whereas the smaller form is localised in mitochondria. The difference observed in relative molecular weights would be consistent with the nuclear form being the primary translation product initiated from AUG_2, the mitochondrial form being initiated at AUG_1, and the pre-sequence processed as predicted between residues 44 and 45.

AUG_1 of *CDC9* is required for the maintenance of the mitochondrial genome

Yeast cells require mitochondria for growth, but the mitochondrial genome itself is dispensable. Nevertheless, mutants lacking some or all of the mitochondrial genome lack enzymes required for respiratory growth. Such mutants ('petite' mutants) are therefore unable to grow on non-fermentable carbon sources such as glycerol. If the mitochondrial and nuclear forms of Cdc9 are expressed from AUG_1 and AUG_2, respectively, then a mutation in AUG_1 might be predicted to disrupt the replication of the mitochondrial genome without impacting on the essential nuclear function of DNA ligase. To test this hypothesis, both AUG_1 and AUG_2 were individually mutated and the effects on functional expression of *CDC9–HA* determined. AUG_1 was mutated to a UAG stop codon, and AUG_2 to GCG (Ala). The individual mutations were integrated into a diploid strain, and the heterozygous diploids dissected and grown on glucose. In either case, viable haploids were recovered carrying the mutant *cdc9* allele. A haploid strain carrying the $AUG_1 \rightarrow$UAG mutation grew well on glucose medium but was unable to grow on glycerol, demonstrating that these cells were petite (Figure 7). This compares with otherwise isogenic wild-type and $AUG_2 \rightarrow$GCG strains, which grew normally on glycerol (Figure 7).

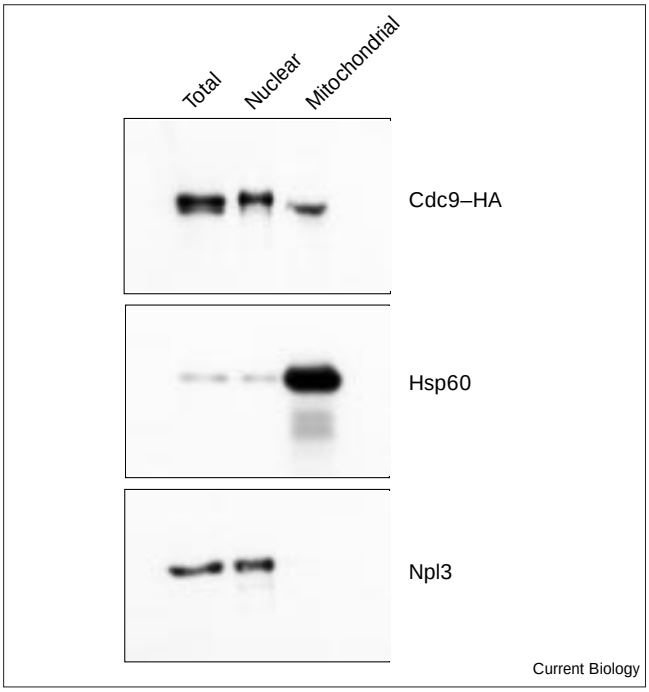
Western blot analysis clearly demonstrated that the upper form of the Cdc9p doublet was absent in $AUG_2 \rightarrow$GCG mutant cells, whereas the lower form of the doublet was lacking in the $AUG_1 \rightarrow$UAG mutant, thus confirming that these forms are indeed expressed from different AUG codons (data not shown). As *CDC9* is an essential gene, we were initially surprised by the finding that $AUG_2 \rightarrow$GCG cells were viable. This anomaly would be most easily explained if, in the absence of AUG_2, a scanning ribosome were to now initiate at AUG_3 to yield a functional nuclear form of Cdc9p. Western blot analysis of the $AUG_2 \rightarrow$GCG mutant strain suggested that this was indeed the case (data not shown). From these data, we conclude that the mitochondrial form of Cdc9p is required for the stable maintenance of mitochondrial DNA but not for cell division.

Figure 5



Cdc9–HA is protected against proteinase K in mitochondrial fractions. Nuclei and mitochondria were purified from MDY26 and treated with proteinase K in the absence or presence of Triton X-100 as indicated. After precipitation with tricarboxylic acid (TCA), the samples were analysed by western blotting with anti-HA, anti-Hsp60 and anti-Npl3 antibodies. Both (a–c) mitochondria and (d–f) nuclei were subjected to the four different combinations of proteinase K and/or detergent.

Figure 6



Two distinct forms of Cdc9–HA are found in nuclei and mitochondria. Western blot analysis of total cell extract, and mitochondrial and nuclear preparations from MDY26 probed with antibodies against HA, Hsp60 and Npl3.

Discussion

Mitochondrial DNA replication ought to require an activity similar to DNA ligase I for the ligation of Okazaki fragments. Analysis of the coding sequence of the yeast *CDC9* gene led us to speculate that this gene might encode both a nuclear and a mitochondrial form of DNA ligase I. The ORF has several in-frame AUG codons towards its 5′ end as originally noted by Barker *et al.* ([11]; see Figure 1). Translation from AUG_1 would lead to a 755 amino-acid primary translation product of which the amino-terminal region has several properties characteristic of a mitochondrial pre-sequence, most notably, a putative amphipathic α-helical region (residues 1–18) with a positively charged face, a general lack of acidic residues and a preponderance of hydroxylated residues. Moreover, this putative matrix-targeting signal is followed, between residues 44 and 45, by a potential site for specific cleavage by the MPP located within the mitochondrial matrix [20,22].

Cleavage at this site would predict a mature mitochondrial form with a molecular weight of 80.1 kDa, compared with a predicted primary translation product of 84.8 kDa. Initiation of translation solely at the AUG_1 codon would endow every Cdc9 molecule with this putative pre-sequence. Any model for Cdc9 expression would have to account, therefore, for the required expression of a nuclear form of this enzyme. Translational initiation at AUG_2 would accommodate this as it would result in a 732 amino-acid primary translation product (82.3 kDa) lacking the amphipathic region discussed above, but which would still contain the three putative nuclear localisation signals indicated in Figure 1a. This model therefore predicts two forms of Cdc9p in wild-type yeast cells and, indeed, Tomkinson *et al.* [26] reported the purification of wild-type Cdc9p as a tight doublet of approximately 80 kDa with an additional 70 kDa proteolytic fragment corresponding to an enzymatically active carboxy-terminal portion of Cdc9p.

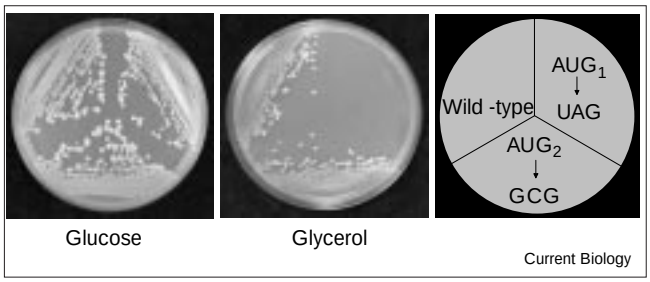
To test this model, we first engineered several *CDC9–GFP* gene fusions and, by western blot analysis, determined the nature of the fusion proteins expressed *in vivo*. When GFP was fused downstream of either AUG_3 or AUG_4, translation products were observed in both cases with relative molecular weights very closely comparable to those expected from initiation at both AUG_1 and AUG_2. In addition, each construct led to the expression of a third major form whose mobility would be consistent either with initiation at AUG_3 or with a mitochondrial processing product (Figure 2). When the products from $AUG_{1–4}$–GFP were compared with those of a mutant derivative in which both AUG_2 and AUG_3 were replaced, then a clearer picture

emerged. Firstly, the 32.8 kDa product predicted to arise from initiation at AUG_2 was now absent, whereas the smaller form remained unaffected (Figure 2b). From this, it is clear that this smaller fusion protein cannot arise through initiation at AUG_3. Fractionation studies of the three products expressed from AUG_{1-3}–GFP revealed that the intermediate species (the size of which was consistent with initiation at AUG_2) remained in the soluble cytosolic phase whereas the other two forms fractionated with mitochondria.

These results would be consistent with translational initiation of Cdc9–GFP at both AUG_1 and AUG_2. The fusion protein observed with an intermediate gel mobility would correspond to the product of initiation at AUG_2. The largest species would correspond to the primary translation product from AUG_1, with the smallest form being derived from this by removal of the pre-sequence on import into mitochondria. Our fractionation studies indicated that both the largest and smallest forms of Cdc9–GFP were mitochondrially associated. It is not clear why a small proportion of the fusion should remain unprocessed although it may be due to some overexpression from the multicopy plasmid constructs. The cofractionation of Cdc9–GFP with mitochondria was confirmed by confocal microscopy in intact cells where it was found that the majority of the Cdc9–GFP signal was precisely colocalised with the mitochondrial marker, Hsp60 (Figure 3). These results indicate that Cdc9–GFP can be expressed from each of two consecutive AUG codons corresponding to AUG_1 and AUG_2 of *CDC9*, and that the first of these incorporates a functional mitochondrial pre-sequence which is cleaved upon import into mitochondria.

We went on to examine the cellular distribution of an HA-tagged derivative of Cdc9p. The tagged *CDC9–HA* gene expressed functional ligase activity as it could support wild-type growth rates in a haploid strain background. Moreover, the tagged Cdc9–HA protein migrated as a doublet in SDS–PAGE as did native Cdc9p (Figure 4a; [26]). Nuclei and mitochondria were purified from cells expressing Cdc9–HA, and western blotting studies revealed substantial levels of Cdc9–HA in both organelles (Figure 4b). These two pools of Cdc9–HA were further distinguished by their protease accessibility in which it was evident that the mitochondrion-associated form was entirely sequestered within a sealed membrane-bound compartment. In contrast, the nuclear form remained characteristically accessible to protease (Figure 5). Moreover, the two forms could be resolved by SDS–PAGE and their different mobilities would appear to account for the Cdc9–HA doublet observed in whole cell extracts (Figure 6), the higher molecular weight band of the doublet apparently corresponding to the nuclear form and the faster-migrating species being indistinguishable from the mitochondrial form. Overall, these data indicate that yeast Cdc9 is expressed in two topologically distinct isoforms from adjacent AUG codons in the *CDC9* ORF.

Figure 7



Functional distinction between Cdc9 encoded from AUG_1 and AUG_2. After tetrad analysis, haploid leucine auxotrophs were tested for growth on YPD (glucose) and YPEG (glycerol) media. The yeast strains were MWY21 (wild type), MWY32 and MWY51 (heterozygous diploids in which one copy of *CDC9* was engineered to alter either AUG_1 or AUG_2, respectively). The specific mutations were $AUG_1 \rightarrow UAG$ and $AUG_2 \rightarrow GCG$.

The biological significance of these observations was confirmed by our findings of the phenotypes of AUG mutants engineered into the otherwise wild-type *CDC9* gene. A mutant lacking AUG_1 remained viable but became petite (Figure 7). From this, it is evident that initiation at AUG_1 is required to produce the DNA ligase necessary to maintain mitochondrial DNA, but is not required for replication of the nuclear genome. Our data further demonstrate that the nuclear form of Cdc9p is normally expressed from AUG_2, but that initiation at this codon is not essential for viability because of the close proximity of AUG_3, only 24 codons downstream. In the absence of AUG_2, a functional form of Cdc9p is initiated at AUG_3.

The expression of two distinct isoforms of DNA ligase I might not be restricted to *S. cerevisiae*. We have found strong candidates for amino-terminal mitochondrial pre-sequences encoded by both the *Schizosaccharomyces pombe cdc17* gene [27] and a DNA ligase I homologue from *Arabidopsis thaliana* [28]. In each case, the putative pre-sequence is closely followed by a second in-frame AUG codon. Interestingly, the sequence of the cloned human DNA ligase I cDNA does not encode any recognisable mitochondrial pre-sequence. There could of course be some cryptic mitochondrial import signal, or perhaps an alternatively spliced mRNA that might incorporate a mitochondrial pre-sequence. Alternatively, a second DNA ligase I homologue in the human genome might encode a mitochondrial form of this enzyme. Intriguingly, expression of the human cDNA complements the growth defect in a yeast *cdc9-1* mutant [29]. It would be very interesting to examine the respiratory competence of such cells to test for the human enzyme's ability to complement any mitochondrial defect. Interestingly, a DNA ligase III activity

has been identified in mitochondrial extracts from *Xenopus laevis* [30]. More recently, it has been shown that human DNA ligase III is expressed in both nuclear and mitochondrial forms from tandem AUG codons [31] in a manner which appears similar to that described here for yeast *CDC9*. DNA ligase III has been implicated primarily in DNA repair and recombination but might also perform the replication role required in the human mitochondrion. Although antisense inhibition studies indicate that DNA ligase III is the major source of ligase activity in the human mitochondrion, they do not, however, exclude the possibility that other forms of DNA ligase may also function in this compartment [31].

In eukaryotes, the majority of proteins studied have been shown to be accurately and efficiently sorted to a single cellular compartment. Nevertheless, several examples are now known where an enzyme can be expressed in two or more isoforms from a single genetic locus (for a review, see [32]). These isoforms can carry distinct targeting signals leading to their being sorted to different compartments. A variety of mechanisms can account for the expression of distinct isoforms but, in yeast, the majority of cases rely on either altered transcriptional or translational initiation to produce isoforms which differ at their amino termini (see [32] and references therein). From our current data, we cannot determine whether the mechanism responsible for expression of distinct isoforms of Cdc9p operates at the transcriptional or the translational level. Previous studies have reported a single transcriptional initiation site upstream of the first AUG codon of *CDC9* [11,14]. It would therefore follow that some post-transcriptional mechanism must serve to ensure the expression of the nuclear form of yeast DNA ligase I. The nature of this control mechanism is currently under investigation.

Materials and methods

Strains and culture conditions

All yeast strains used in this study are derived from W303 (*MAT*a/*MAT*α *ura3-1/ura3-1 leu2-3,112/leu2-3,112 trp1-1/trp1-1 his3-11/his3-11 ade2-1/ade2-1 can1-100/can1-100*) including the haploid wild-type derivative W303-3c (*MAT*a, *ura3-1-1 leu2-3,112 trp1-1-1 his3-11 ade2-1 can1-100*), MDY26 (as W303-3c but *CDC9–HA*), MWY21 (as W303 but *CDC9/cdc9–HA*::pMW185), MWY32 (as W303 but *CDC9/cdc9–HA*::pMW186), and MWY51 (as W303 but *CDC9/cdc9–HA*::pMW190). Yeast cultures were grown in either YPD (2% peptone, 1% yeast extract, 2% glucose) or, to maintain selection in YNB (0.675% yeast nitrogen base, 2% glucose), supplemented with appropriate amino-acid supplements. Mitochondrial function was assessed by plating cells on YPEG (2% peptone, 1% yeast extract, 3% glycerol, 3% ethanol). Sporulation and tetrad dissection were performed as described previously [33].

Generation of CDC9–GFP fusions

Expression vector pUG23 for carboxy-terminal GFP fusions was kindly provided by J.H. Hegemann. Two plasmids containing various lengths of the amino terminus of *CDC9* fused in frame to GFP under the control of a *MET25* promoter were constructed. *CDC9* was amplified by PCR with primers A (5′-CATCTAGAACATGCGCAGATTACTG-3′) and B (5′-GGAGATCTAGTGAATTCTGTTGGCTTATTTTTC-3′) and subcloned into the *Eco*RV site of pBluescript KS+ (pMW178); pMW178 was subsequently digested with *Xba*I and *Eco*RI and the 172 bp *CDC9* fragment was cloned into the *Xba*I–*Eco*RI sites of pUG23. The resulting plasmid (pMW182) contains GFP fused in frame after residue 54 of *CDC9*. A similar construct (pMW181) was made by cloning a 223 bp *Xba*I–*Bam*HI fragment of the PCR product with primers A and C (5′-CTATCGATATGCATTCTATCTTCCAGCATAT-3′) into the *Xba*I–*Bam*HI sites of pUG23. In this construct, GFP is fused to *CDC9* after residue 70. The entire inserts of both plasmids were sequenced by automated sequencing.

Subcellular fractionation

Yeast mitochondria and a post-mitochondrial supernatant were prepared as described previously [24]. Yeast nuclei were prepared as follows. Cells were grown in YPD to mid-exponential phase, harvested by centrifugation and converted into sphaeroplasts [24], which were pelleted, washed twice with 1.2 M sorbitol, 20 mM potassium phosphate pH 6.5 and resuspended in cold 18% Ficoll 400 (w/v), 20 mM potassium phosphate pH 6.5, 0.5 mM $MgCl_2$. Sphaeroplasts were then homogenised by two strokes with a Dounce homogeniser on ice. An equal volume of 2.4 M sorbitol, 20 mM potassium phosphate pH 6.5, 0.5 mM $MgCl_2$ was added and the homogenate was spun at $1000 \times g$ for 10 min at 4ºC to remove unbroken cells and cell debris. The supernatant was transferred to a fresh tube and nuclei were pelleted by centrifugation at $12000 \times g$ for 25 min at 4ºC. The pellet was resuspended in a small volume of 1.2 M sorbitol, 20 mM potassium phosphate pH 6.5, 0.5 mM $MgCl_2$ and layered on top of a three-step glycerol gradient as described by Kalinich and Douglas [34]. The nuclear fraction was then further purified by centrifugation on a sucrose gradient as described by Hurt *et al.* [25]. Proteinase K treatment of isolated mitochondria and nuclei was carried out by adding an equal volume of 500 μg/ml proteinase K in 0.6 M mannitol, 10 mM Tris-HCl pH 7.4. Triton X-100 was added to a final concentration of 0.5%. After incubation at 0ºC for 30 min, proteins were TCA precipitated. SDS-containing sample buffer was added and the samples were immediately heated to 95ºC for 4 min and analysed by SDS–PAGE as described below.

Western blotting

Methods for the preparation of whole cell extracts and for SDS–PAGE and immunoblotting were as described previously [33]. Antiserum directed against yeast Hsp60 was used at 1:10,000 dilution (K. Tokatlidis, personal communication), anti-Npl3 antiserum at 1:3000 dilution (A.H. Corbett, personal communication), anti-GFP antiserum at 1:1000 dilution and anti-HA monoclonal antibody (12CA5/Boehringer Mannheim) at 1:1000 dilution. Secondary HRP-conjugated anti-rabbit antibody (Sigma) and HRP-conjugated anti-mouse antibody (DAKO) were both used at a 1:10,000 dilution before detection with Renaissance ECL (NEN).

Site-directed mutagenesis of CDC9

A 1136 bp *Pst*I–*Xho*II fragment from pR12ScLig2 [35] containing the promoter region and part of the CDC9 ORF was cloned into the *Pst*I–*Bam*HI sites of YIp351 [36]. The resulting plasmid, pMW185 was subjected to site-directed mutagenesis performed with Quickchange site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. The first ATG of the *CDC9* ORF in pMW185 was changed to TAG (stop codon) using primers D (5′-CATTGTTCATCAATTACTAGCGCAGATTACTGACC-3′) and E (5′-GGTCAGTAATCTGCGCTAGTAATTGATGAACAATG-3′) to produce pMW186. In pMW190, the second ATG of the *CDC9* ORF in pMW185 was changed to GCG (Ala) with primers F (5′-CCTTGAAATCAAGATTGCCATTATTGGCGTCATCCTCATTACCTTCC-3′) and G (5′-GGAAGGTAATGAGGATGACGCCAATAATGGCAATCTTGATTTCAAGG-3′). The inserts of pMW186 and pMW190 were sequenced to verify the changes of the DNA sequence. Subsequently, the two mutant plasmids and a wild-type control (pMW185) were linearised with *Nde*I and transformed into MDY26 (*CDC9/CDC9–HA*) selecting for leucine auxotrophy. All integrations were confirmed by PCR of genomic DNA using

primers J (5′-CCGGATCTATGCGGTGTGAAATAC-3′) and K (5′-ATGCACCTGAGCCCTTTCACCATC-3′). The diploid strains containing the integrated constructs were named: MWY21 (pMW185/wild-type), MWY32 (pMW186/$AUG_1 \rightarrow$UAG) and MWY51 (pMW190/$AUG_2 \rightarrow$GCG). All three strains were transformed with pR12ScLig2 selecting for leucine/uracil auxotrophy and tetrad analysis revealed a 2:2 segregation of the leucine auxotrophy, suggesting a single integration event in each of the strains MWY21, MWY32 and MWY51. Finally, the effects of the individual ATG mutants were assessed by tetrad analysis of the three diploid strains. Mutagenesis of the second and third ATG in pMW181 (*CDC9–GFP*, described above) was performed in two steps: the third ATG of *CDC9* in pMW181 was changed to GCG with primers H (5′-GGCTA-GATTCTTCACTTCCGCGAAAAATAAGCCAACAGAAGG-3′) and I (5′-CCTTCTGTTGGCTTATTTTTCGCGGAAGTGAAGAATCTAGCC-3′), resulting in pMW197 in which the second ATG was changed to GCG with primers F and G (pMW198). The correct changes were verified by DNA sequencing.

Immunofluorescence microscopy

Cells were grown to mid-logarithmic phase in minimal media selecting for the presence of plasmid, as appropriate, and fixed by addition of formaldehyde to a final concentration of 4.4% and subsequently treated as described [37]. Cells were incubated for 60 min with 1:1000 dilution of anti-Hsp60 antisera in PBS + 1 mg/ml BSA. Cy3-coupled anti-rabbit secondary antibody was used at 1:1000 dilution for 60 min. Hsp60 staining was visualised by immunofluorescence microscopy using a Leica TCS NT confocal microscope and associated software supplied by the manufacturer. Simultaneous recording of GFP signal was performed and 25 individual sections throughout each sample were integrated.

Epitope tagging with transposon

As a target for mTn3-3xHA/lacZ transposon insertion, a 1886 bp *Bgl*II–*Xba*I fragment containing the amino-terminal part of CDC9 was cloned from pR12ScLig2 into the *Bam*HI–*Xba*I sites of the vector pHSS6. This plasmid (pHSS6-*CDC9*) was subjected to transposon mutagenesis as described [23]. One haploid viable strain, MDY26, expressing a functional *CDC9–HA* was chosen for further analysis. Genomic DNA from MDY26 was amplified by PCR with primers L (5′-CACCTTAACGCGAAAACGCGTGAAAGTG-3′) and M (5′-TTTGACCCACAAAGGTGGGAAAAGTGCC-3′) and sequenced to determine the position of the 279 bp 3 × HA insertion.

Supplementary material

A supplementary figure showing Cdc9–GFP localisation in intact yeast cells is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

We thank Kostas Tokatlidis (University of Manchester) for kindly providing us with antibodies against Hsp60, Mdh1 and Porin; Craig Peterson (UMass Medical School, Worchester) for anti-Swi3 antisera; Anita Corbett (Emory University School of Medicine, Atlanta) for anti-Npl3 antisera; Ken Sawin (ICRF, London) for anti-GFP antisera; Johannes Hegemann (University of Dusseldorf) for providing us with plasmids for GFP fusion constructs; Lee Johnston (NIMR, Mill Hill) for providing a clone (pR12ScLig2) of CDC9 and for stimulating discussions; and Usha Lad (University of Manchester) for help with confocal microscopy. This work was supported by the BBSRC and the Lister Institute for Preventive Medicine.

References

1. Waga S, Bauer G, Stillman B: Reconstitution of complete sv40 DNA replication with purified replication factors. *J Biol Chem* 1994, 269:10923-10934.
2. Petrini JJ, Xiao YH, Weaver DT: DNA ligase I mediates essential functions in mammalian cells. *Mol Cell Biol* 1995, 15:4303-4308.
3. Prigent C, Satoh MS, Daly G, Barnes DE, Lindahl T: Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I. *Mol Cell Biol* 1994, 14:310-317.
4. Thompson LH, Brookman KW, Jones NJ, Allen SA, Carrano AV: Molecular cloning of the human XRCC1 gene, which corrects defective DNA strand break repair and sister chromatid exchange. *Mol Cell Biol* 1990, 10:6160-6171.
5. Kubota Y, Nash RA, Klungland A, Schar P, Barnes DE, Lindahl T: Reconstitution of DNA base excision repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. *EMBO J* 1996, 15:6662-6670.
6. Chen JW, Tomkinson AE, Ramos W, Mackey ZB, Danehower S, Walter CA, *et al.*: Mammalian DNA ligase III — molecular cloning, chromosomal localization, and expression in spermatocytes undergoing meiotic recombination. *Mol Cell Biol* 1995, 15:5412-5422.
7. Roberts E, Nash RA, Robins P, Lindahl T: Different active sites of mammalian DNA ligase I and ligase II. *J Biol Chem* 1994, 269:3789-3792.
8. Wang YJ, Burkhart WA, Mackey ZB, Moyer MB, Ramos W, Husain I, *et al.*: Mammalian DNA ligase II is highly homologous with vaccinia DNA ligase — identification of the DNA ligase II active site for enzyme adenylate formation. *J Biol Chem* 1994, 269:31923-31928.
9. Robins P, Lindahl T: DNA ligase IV from HeLa cell nuclei. *J Biol Chem* 1996, 271:24257-24261.
10. Barnes DE, Stamp G, Rosewell I, Denzel A, Lindahl T: Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. *Curr Biol* 1998, 8:1395-1398.
11. Barker DG, White JM, Johnston LH: The nucleotide sequence of the DNA ligase gene (***CDC9***) from ***Saccharomyces cerevisiae*** — a gene which is cell cycle regulated and induced in response to DNA damage. *Nucleic Acids Res* 1985, 13:8323-8337.
12. Peterson TA, Prakash L, Prakash S, Osley MA, Reed SI: Regulation of ***CDC9***, the ***Saccharomyces cerevisiae*** gene that encodes DNA ligase. *Mol Cell Biol* 1985, 5:226-235.
13. White JM, Barker DG, Nurse P, Johnston LH: Periodic transcription as a means of regulating gene expression during the cell cycle — contrasting modes of expression of DNA ligase genes in budding and fission yeast. *EMBO J* 1986, 5:1705-1709.
14. Johnson AL, Barker DG, Johnston LH: Induction of yeast DNA ligase genes in exponential and stationary phase cultures in response to DNA damaging agents. *Curr Genet* 1986, 11:107-112.
15. Ramos W, Tappe N, Talamantez J, Friedberg EC, Tomkinson AE: Two distinct DNA ligase activities in mitotic extracts of the yeast ***Saccharomyces cerevisiae***. *Nucleic Acids Res* 1997, 25:1485-1492.
16. Schar P, Herrmann G, Daly G, Lindahl T: A newly identified DNA ligase of ***Saccharomyces cerevisiae*** involved in ***RAD52*** independent repair of DNA double-strand breaks. *Genes Dev* 1997, 11:1912-1924.
17. Teo SH, Jackson SP: Identification of ***Saccharomyces cerevisiae*** DNA ligase 4. Involvement in DNA double-strand break repair. *EMBO J* 1997, 16:4788-4795.
18. Wilson TE, Grawunder U, Lieber MR: Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* 1997, 388:495-498.
19. Lasko DD, Tomkinson AE, Lindahl T: Mammalian DNA ligases — biosynthesis and intracellular localization of DNA ligase I. *J Biol Chem* 1990, 265:12618-12622.
20. Von Heijne G: Mitochondrial targeting sequences may form amphiphilic helices. *EMBO J* 1986, 5:1335-1342.
21. Yang M, Jensen RE, Yaffe MP, Oppliger W, Schatz G: Import of proteins into yeast mitochondria — the purified matrix processing protease contains 2 subunits which are encoded by the nuclear ***MAS1*** and ***MAS2*** genes. *EMBO J* 1988, 7:3857-3862.
22. Nakai K, Kanehisa M: A knowledge base for predicting protein localisation sites in eukaryotic cells. *Genomics* 1992, 14:897-911.
23. Ross-Macdonald P, Sheehan A, Friddle C, Roeder GS, Snyder M: Transposon tagging I: a novel system for monitoring protein production, function and localization. *Methods Microbiol* 1998, 26:161-179.
24. Daum G, Bohni PC, Schatz G: Import of proteins into mitochondria — cytochrome-b2 and cytochrome-c peroxidase are located in the intermembrane space of yeast mitochondria. *J Biol Chem* 1982, 257:3028-3033.
25. Hurt EC, McDowall A, Schimmang T: Nucleolar and nuclear envelope proteins of the yeast ***Saccharomyces cerevisiae***. *Eur J Cell Biol* 1988, 46:554-563.

26. Tomkinson AE, Tappe NJ, Friedberg EC: DNA ligase I from ***Saccharomyces cerevisiae*** — physical and biochemical characterization of the ***CDC9*** gene product. *Biochemistry* 1992, 31:11762-11771.
27. Johnston LH, Barker DG, Nurse P: Cloning and characterization of the ***Schizosaccharomyces pombe*** DNA ligase gene ***cdc17***. *Gene* 1986, 41:321-325.
28. Taylor RM, Hamer MJ, Rosamond J, Bray AM: Molecular cloning and functional analysis of the ***Arabidopsis thaliana*** DNA ligase I homologue. *Plant J* 1999, 14:75-81.
29. Barnes DE, Johnston LH, Kodama K, Tomkinson AE, Lasko DD, Lindahl T: Human DNA ligase I cDNA — cloning and functional expression in ***Saccharomyces cerevisiae***. *Proc Natl Acad Sci USA* 1990, 87:6679-6683.
30. Pinz KG, Bogenhagen DF: Efficient repair of abasic sites in DNA by mitochondrial enzymes. *Mol Cell Biol* 1998, 18:1257-1265.
31. Lakshmipathy U, Campbell C: The human DNA ligase III gene encodes nuclear and mitochondrial proteins. *Mol Cell Biol* 1999, 19:3869-3876.
32. Danpure CJ: How can the products of a single gene be localized to more than one intracellular compartment. *Trends Cell Biol* 1995, 5:230-238.
33. Craven RA, Egerton M, Stirling CJ: A novel Hsp70 of the yeast ER lumen is required for the efficient translocation of a number of protein precursors. *EMBO J* 1996, 15:2640-2650.
34. Kalinich JF, Douglas MG: ***In vitro*** translocation through the yeast nuclear envelope — signal dependent transport requires ATP and calcium. *J Biol Chem* 1989, 264:17979-17989.
35. Barker DG, Johnston LH: ***Saccharomyces cerevisiae CDC9***, a structural gene for yeast DNA ligase which complements ***Schizosaccharomyces pombe cdc17***. *Eur J Biochem* 1983, 134:315-319.
36. Hill JE, Myers AM, Koerner TJ, Tzagoloff A: Yeast/***E. coli*** shuttle vectors with multiple unique restriction sites. *Yeast* 1986, 2:163-168.
37. Pringle JR, Adams AEM, Drubin DG, Haarer BK: Immunofluorescence methods in yeast. *Methods Enzymol* 1991, 194:565-602.

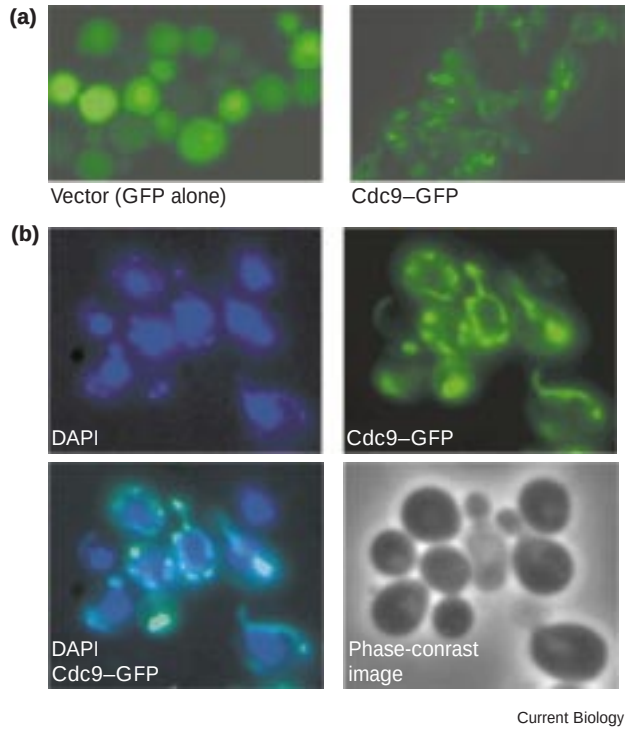
Supplementary material

The yeast *CDC9* gene encodes both a nuclear and a mitochondrial form of DNA ligase I

M. Willer, M. Rainey, T. Pullen and C.J. Stirling

Current Biology 23 September 1999, 9:1085–1094

Figure S1



Cdc9–GFP localisation in intact yeast cells. (a) Fluorescence microscopy of wild-type cells expressing the Cdc9–GFP fusion protein from pMW181 (AUG_{1-4}) revealed a punctuate/tubular pattern compared with the vector control (pUG23). (b) Cdc9–GFP staining overlapped substantially with mitochondrial DNA in DAPI-stained cells. Cells were fixed and stained as described in the Materials and methods section. For DAPI staining, fixed cells were washed twice with PBS and then stained with 1 μg/ml DAPI for 5 min. Stained cells were then washed twice with PBS and then mounted in 1 mg/ml *p*-phenylenediamine in 90% glycerol before examination on a Zeiss Axiophot fluorescence microscope.