Volume 16 Number 8 1988 Nucleic Acids Research

Sequence and expression of *NUC1*, the gene encoding the mitochondrial nuclease in *Saccharomyces cerevisiae*

R.D.Vincent, T.J.Hofmann and H.P.Zassenhaus*

Department of Microbiology, St Louis University Medical Center, St Louis, MO 63104, USA

Received December 21, 1987; Revised and Accepted March 15, 1988

ABSTRACT

The DNA sequence and studies on the expression of the NUC1 gene from Saccharomyces cerevisiae are presented. The NUC1 locus is located in the distal portion of the left arm of Chromosome X and encodes the major nuclease found in mitochondria. The inferred amino acid sequence of NUC1 predicts that the nuclease is basic, rich in prolines, of average hydrophobicity, and has a molecular weight for the primary translation product of 37,209 daltons. NUC1 is very poorly expressed, consistent with the codon usage bias determined from the DNA sequence and our previous determination of the number of enzyme molecules per cell. Mapping of the 5' terminus of the NUC1 mRNA reveals that the mRNA has a long 400 base untranslated leader in which are found three open reading frames, each initiated by an AUG. The possibility that these upstream open reading frames contribute to the poor expression of the NUC1 gene is discussed.

INTRODUCTION

In yeast more than 50% of all cellular nuclease activity is due to a single enzyme located in mitochondria (1,2). This enzyme has both RNase and DNase activity, is bound to the mitochondrial inner membrane, and is encoded in the nucleus (3). In vitro the enzyme lacks site specificity for mitochondrial RNAs or DNA. In the accompanying report (4) we demonstrate that disruption of NUC1, the gene encoding this nuclease, produces a strain with no detectable mitochondrial DNase activity and little nonspecific RNase activity. Despite the absence of nuclease activity in this strain, mitochondrial function appears to be normal since the strain is rho^+. Moreover, a strain with 20-40 times the normal level of mitochondrial nuclease is also phenotypically rho^+.

Although the amount of enzymatic activity in the mitochondria of a wild-type yeast strain is prodigious, the number of enzyme molecules in the mitochondria is low, possibly as few as 100-300 per cell (3). This level does not vary more than two fold comparing cells grown in fermentable versus

© IRL Press Limited, Oxford, England.

nonfermentable carbon sources, cells in logarithmic versus stationary phase, or rho+ versus rho- cells (3).

Similar potent DNase and RNase activities have been observed in other eukaryotes (5,6), although only an enzyme from Neurospora crassa has been completely purified (7). A partially purified enzyme preparation from beef heart mitochondria has also recently been characterized (8,9). In both species, it appears that the characterized enzymes are at least responsible for the majority of all mitochondrial nuclease activity and may possibly be the only nonspecific nucleases present, as we have shown in yeast (3,4).

These observations suggest that the presence of a mitochondrial DNase/RNase may be widespread in eukaryotes. Although the function of this enzyme remains obscure, particularly since our yeast mutant lacking nuclease is phenotypically rho+ (3,4), the mitochondrial location and activity of this class of enzymes appears to have been conserved during evolution. Here we present the nucleotide sequence of the gene (NUC1) for mitochondrial DNase/RNase from yeast and an analysis of its expression. These data show that the gene is very poorly expressed, consistent with its codon usage bias and our previous determination of the amount of enzyme in mitochondria.

MATERIALS AND METHODS

DNA Sequencing

DNA sequencing was performed by the chemical method of Maxam and Gilbert (10). Restriction fragments were labeled (11) either by a fill in reaction for the 3' end or by end labeling the 5' end with T4 polynucleotide kinase and [γ-^{32}P]ATP after dephosphorylation with calf intestinal alkaline phosphatase. When necessary, labeled restriction fragments were isolated by gel electrophoresis and electroelution (12). All enzymes were purchased from New England Biolabs (Beverly, MA) and isotope from ICN (Irvine, CA).

Northern Blots

RNA preparation. Total cellular RNA was prepared from spheroplasts by the guanidine-HCl/CsCl method (13). Alternatively, spheroplasts were lysed in 100mM NaCl, 10mM Tris-HCl, pH 8.0, 10mM EDTA, 0.5% SDS at 65°C, extracted with phenol/chloroform (1:1), and RNA precipitated from total nucleic acid at 4°C for 24 hr after the addition of LiCl to 2M (14).

Preparation of poly(A)+-enriched RNA. A poly(A)+-enriched RNA fraction was prepared from total cellular RNA (isolated by the guanidine-HCl/CsCl method) by oligo(dT) cellulose column chromatography as described by Maniatis, et al. (11) with LiCl substituted for NaCl, except that the 0.1M salt wash was

omitted. All RNA still bound to the column after extensive washing with loading buffer was eluted with 10mM Tris-HCl, pH 7.5, 1mM EDTA, 0.05% SDS.

DNA probes. Double stranded DNA probes were labeled by random oligomer priming, using either electroeluted (12) restriction fragments or restriction fragments excised from low-melt agarose gels (15,16).

Gels, blotting, and hybridization. RNAs were electrophoresed through gels containing 1.5% agarose and 2.2M formaldehyde (11). Gels were blotted overnight in 20X SSPE (3.0M NaCl, 0.2M NaH_2PO_4, 0.02M EDTA, pH 7.4) to either nitrocellulose or charged nylon membranes. The membranes were air dried, baked at 80°C for 2 hr, prehybridized at 42°C for 24-36 hr in 50% formamide, 5X SSPE, 1-10X Denhardt's solution (1X: 0.02% (w/v) each of Ficoll, polyvinylpyrrolidone, and bovine serum albumin), 200-400 µg/ml sonicated, denatured herring sperm DNA, and hybridized to probes in a fresh aliquot of the same solution for 36-48 hrs. After hybridization, membranes were washed twice for 15 minutes in 5X SSPE at 42°C, once in 1X SSPE, 0.1% SDS at 42°C for 30 min, and either once in 1X SSPE, 0.1% SDS at room temperature for 15 minutes or in 0.1X SSPE, 0.1% SDS at 55°C for 1 hr, depending upon the experimental protocol. Membranes were blotted dry and autoradiographed with Kodak XAR-5 film and Dupont Cronex Lightning Plus screens.

Nuclease S1 Mapping of the mRNA 5' Ends

A 5' end-labeled probe, preparatively electroeluted from an 8% polyacrylamide urea denaturing gel, was coprecipitated with an excess of poly(A)$^+$-enriched RNA prepared from a strain overproducing NUC1 mRNA. The pellet was dissolved in a microcentrifuge tube in 30 µl S1 hybridization buffer (40mM 1,4- piperazinediethanesulfonic acid, pH 6.4, 1mM EDTA, 0.4M NaCl, 80% (v/v) formamide), incubated at 72°C for 15 min, and then the tube was immediately transferred to a water bath at 50°C. Hybridization was continued as the temperature was decreased in 5°C increments to room temperature over 3 hrs and the tube was then placed on ice. A 10-fold excess (300µl) of ice cold S1 buffer (0.28M NaCl, 0.05M NaOAc, pH 5.2, 4.5mM $ZnSO_4$, 20 µg/ml tRNA) containing 0-100 units of nuclease S1 was then added, and the tube incubated at room temperature for 1 hr. Reactions were placed on ice and stopped by the addition of 50 µl 4M NH_4OAc, 0.1M EDTA. The reactions were extracted once with phenol/chloroform (1:1), 20 µg tRNA carrier was added, and then the reactions were precipitated with an equal volume of isopropanol at -70°C. Pellets were dissolved in formamide load buffer (11), heated to 90°C for 10 min, and the sample was electrophoresed through an acrylamide urea

sequencing gel (11) next to the Maxam-Gilbert sequencing reactions of the probe fragment.

Data Analysis

DNA sequences were analyzed using the Compugene DNALYSIS computer programs. The protein sequence derived from the NUC1 gene was compared to the Protein Identification Resource protein data base (17). The FASTP and ALIGN computer programs were used to determine homologies (18-21). Analysis of the derived amino acid sequence for hydrophobicity was accomplished using the values of Kyte and Doolittle (22). The hydrophobic moment was calculated according to Eisenberg *et al*. (23,24).

RESULTS AND DISCUSSION

The yeast chromosomal DNA contained in the YEp13 derivative pUZ2, described in the preceeding report (4), was used to sequence the NUC1 gene. In that report, we demonstrated that pUZ2 probably contained both NUC1 coding sequences and upstream sequences necessary for NUC1 expression because of the following observations. Transformation of wild-type yeast by pUZ2 yielded nuclease overproducing transformants. Homologous recombination of yeast chromosomal DNA and a gene fragment from pUZ2 containing an inserted copy of the LEU2 gene eliminated NUC1 expression. The strategy used to sequence both strands of the NUC1 gene from this plasmid is shown in Fig. 1, and the DNA sequence of the gene is presented in Fig. 2.

Starting at nucleotide +1 of the sequence is an ATG which initiates a single long open reading frame (ORF) extending to a TGA stop codon beginning at nucleotide +988. Including the other two possible reading frames on this strand and the three on the opposite strand no other open reading frame is of sufficient length to encode the nuclease. In the preceeding report (4) we showed that disruption of the chromosomal NUC1 allele at a unique BamHI site by insertion of LEU2 on a BglII cartridge led to the loss of the mitochondrial nuclease. Fig. 1 shows that the BamHI site is in the long ORF, suggesting the ORF encodes the nuclease. Transforming a yeast strain (4) with a plasmid containing the ORF plus upstream sequences yields progeny with higher than normal levels of nuclease (described in more detail below), also indicating that this ORF encodes nuclease.

The predicted translation product of the NUC1 gene sequence is a protein of 37,209 molecular weight. The apparent molecular weight of the mitochondrial nuclease is 38,000 daltons as determined by SDS-PAGE (3). Assuming that the mitochondrial signal sequence of the enzyme is of average

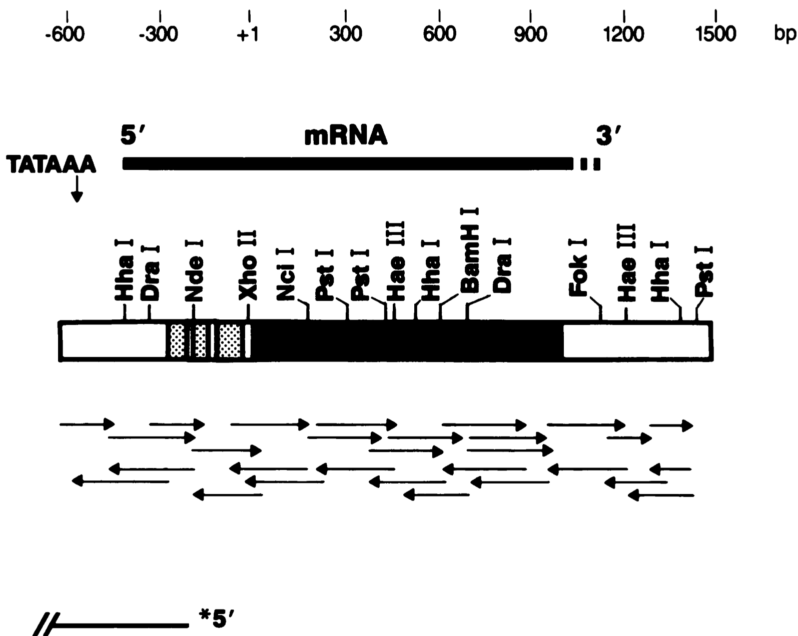

Fig. 1. Restriction map and strategy for sequencing the NUC1 gene. The middle portion of the figure displays the restriction map of the gene. The arrows in the lower portion indicate the lengths and directions of the sequences determined from end labeled restriction fragments by the Maxam-Gilbert method. The solid shading indicates the ORF corresponding to the NUC1 coding region, the dappled shading highlights the small ORFs included in the NUC1 mRNA. Aligned with and above the restriction map, the location of the NUC1 mRNA is indicated. The 5' ends of the message were determined by S1 protection mapping (Fig. 7) using the restriction fragment shown at the bottom of the figure (labeled end indicated). The 3' end of the NUC1 mRNA was determined by Northern mapping (Fig. 6).

size (3-5 kilodaltons) and is cleaved, the measured molecular weight appears to be a few thousand daltons greater than the gene sequence would predict. Considering that molecular weight estimations by SDS-PAGE of membrane bound proteins are often somewhat inaccurate (25), this discrepancy seems slight. Taken together these data show that the long ORF contains the NUC1 coding sequences.

The amino acid composition of the protein predicted from the NUC1 DNA sequence indicates that the nuclease is basic (50 Lys+Arg+His vs 36 Glu+Asp) and rich in proline residues (8.8 mole percent) relative to the amino acid composition of an average protein (26). The proline residues are not clustered but rather are scattered throughout the protein. Only three

Fig. 2. Nucleotide sequence of the NUC1 gene. The nucleotide sequence of the nontranscribed strand of the NUC1 gene is displayed. The amino acid sequences inferred from the DNA sequence for the four ORFs in the NUC1 message are shown below the sequence. Arrows indicate the presence of the 5' ends of the NUC1 mRNA as determined by S1 protection experiments (Fig. 7). The putative TATA box is underlined.

methionines are found in the NUC1 protein, including the initiating one.

We have previously shown that the mitochondrial nuclease is bound to the inner membrane and, as isolated, is a dimer requiring non-ionic detergents for solubility and activity (3). The hydrophobicity plot of the predicted amino acid sequence (Fig. 3), however, indicates that overall the protein is of only average hydrophobicity; it contains two potential regions that have a higher than average probability of potential membrane interactions. The first region

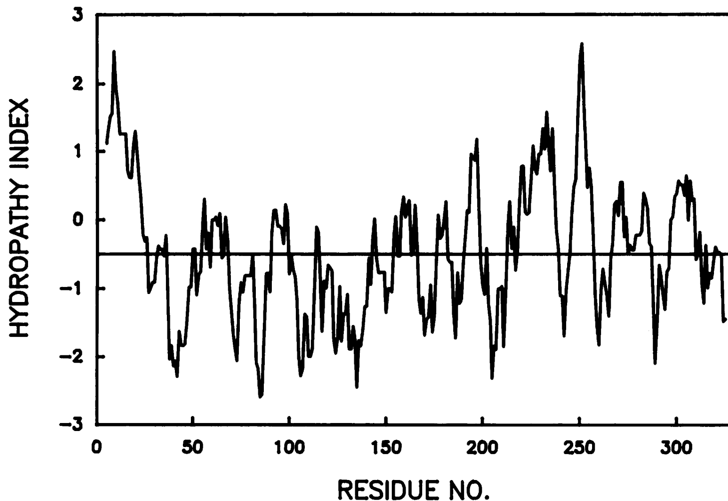

Fig. 3. Hydropathy plot for predicted amino acid sequence of NUC1. The hydropathy index was calculated according to Kyte and Doolittle (22) using a window of 9 amino acid residues.

encompasses the first 25 amino-terminal amino acids and probably serves as a mitochondrial targeting sequence (27). Such sequences have been shown to be enriched for Arg, Leu, and Ser, have few Asp, Glu, Val, and Ile, and many of these sequences form amphiphilic helices (28). Although the amino acid composition of the first 25 residues of NUC1 fits that pattern (11 Arg+Leu+Ser vs 2 Asp+Val+Ile), calculations of the hydrophobic moment using the rules described by Eisenberg (23,24) and von Heinje (28) do not suggest that the amino terminal residues form an amphiphilic helix. As Fig. 4 shows, the potential NUC1 signal sequence lacks the pronounced peak of hydrophobicity at an inter-residue angle of 90°-105° characteristic of amphiphilic helices; rather it has shallow maxima at 80° and 180° which make such a conformation unlikely (28).

The second region of elevated hydrophobicity in the protein occurs between amino acid residues 245 and 259 near the carboxy terminus. Calculations show that this sequence may serve as a membrane spanning region if paired with another like itself (23,24). Since the enzyme appears to be a homodimer by sedimentation analysis (3), such a dimer may be anchored in the membrane by that hydrophobic region and, if so, that the carboxy terminus of the enzyme may extend into the mitochondrial inter-membrane space.

The codon usage in the NUC1 gene is presented in Table I and calculation

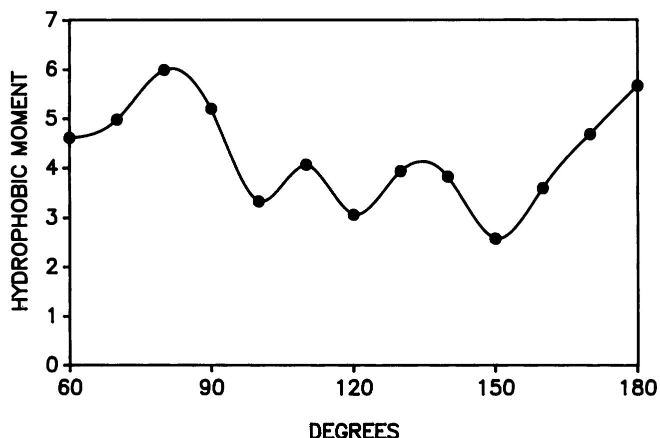

Fig. 4. Hydrophobic moment of the first 30 amino acid residues of NUC1 vs the inter-residue angle. The hydrophobic moment was caluclated according to Eisenberg et al. (23,24) for inter-residue angles between 60° and 180°.

of the bias in codon usage predicts that NUC1 is poorly expressed. The codon adaptation index for NUC1 is 0.139 which is among the lowest observed for sequenced yeast genes (29). This index has been shown to correlate well with the level of gene expression in yeast, in which highly expressed genes have indices > 0.5 whereas lowly expressed ones (e.g. regulatory genes) have

Table I

Codon Usage in NUC1

	T	C	A	G	
T	13	5	10	4	T
	6	4	3	2	C
	6	4	0	1	A
	10	2	0	2	G
C	3	12	5	0	T
	2	3	2	1	C
	3	12	10	1	A
	3	2	3	0	G
A	9	6	14	5	T
	4	2	6	0	C
	5	3	17	9	A
	3	4	9	6	G
G	5	8	12	8	T
	4	3	5	4	C
	7	7	12	3	A
	3	4	7	2	G

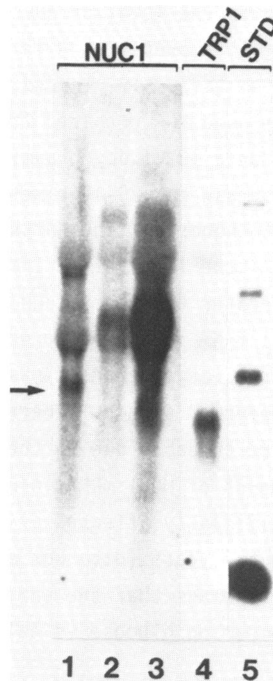

Fig. 5. Expression of NUC1 relative to TRP1. Total cellular RNAs of a nuclease overproducing strain or a wild-type strain were electrophoresed through a gel containing 1.5% agarose and 2.2M formaldehyde, blotted to nitrocellulose, and hybridized to NUC1 or TRP1 probes (see Methods). NUC1 message is not detectable in wild-type strains using an amount of total cellular RNA sufficient to see TRP1 sequences (compare lanes 2 and 4). However, a nuclease overproducing transformant gives a NUC1 signal (arrow) at a blotted level of RNA much lower than that needed to detect NUC1 message in a wild-type strain (compare lanes 1 and 3). Lane 1: 24 µg RNA from a nuclease overproducing transformant, NUC1 probe; Lane 2: 32 µg RNA from a wild-type strain, NUC1 probe; Lane 3: 128 µg wild-type RNA, NUC1 probe; Lane 4: 32 µg wild-type RNA, TRP1 probe; Lane 5: RNA standards, molecular weights top to bottom: 4.4 kb, 2.37 kb, 1.37 kb, 0.24 kb. The NUC1 probe is the 210 bp PstI fragment of the NUC1 ORF pictured as probe 2 in Fig. 6. The TRP1 probe is a 1.45 kb EcoRI fragment containing the entire TRP1 gene. Due to large amounts of total cellular RNAs blotted and the low stringencies of the washes of this Northern, additional bands not related to NUC1 mRNA are observed.

indices close to 0.1 (29). Previously we reported that the number of enzyme molecules per cell was low, possibly as few as 100 (3) which is consistent with that prediction.

Studies on the expression of NUC1 show directly that the steady state level of the mRNA is also low. In either total RNA or poly$(A)^+$-enriched RNA

prepared from a wild-type strain the level of NUC1 mRNA is barely detectable in a Northern analysis, using an amount of RNA in which URA3 or TRP1 transcripts are easily detected (Fig. 5). Indeed, in order to map the 5' and 3' limits of the NUC1 mRNA we used RNA prepared from a strain in which the entire NUC1 gene was present on a multicopy plasmid. As shown in the preceeding report (4), this strain has 20-40 times more NUC1 protein and enzymatic activity in its mitochondria. The Northern analysis in Fig. 5 demonstrates that the steady state level of NUC1 mRNA in that overproducing strain is at least 25 fold greater than that found in a wild-type strain. Furthermore, the size of NUC1 mRNA in the overproducer is identical to that seen in the wild-type, although larger transcripts containing NUC1 sequences are also present in the transformed strain. Therefore, the increased enzyme levels in the overproducer are probably due to the elevated level of NUC1 mRNA. In addition, the relative amount of NUC1 mRNA in either the overproducer or a wild-type strain is not significantly affected by the carbon source (galactose vs glucose) used for growing the yeast (data not shown). This is also in agreement with our previous findings that the enzyme level within mitochondria varies less than two-fold comparing glucose vs galactose grown yeast cells (3).

Transcripts larger than the NUC1 mRNA are present in the Northern analysis of Fig. 5 due to the large amounts of total cellular RNAs blotted and the low stringency of the washes used to maximize the NUC1 signal, especially for wild-type RNA. These larger transcript signals include nonspecific hybridization to ribosomal RNA (seen in all blotted RNAs with all probes) and additional bands hybridizing to NUC1 region probes, but which are not NUC1 message (see below).

The location of transcriptional initiation of NUC1 was suggested by our previous results demonstrating that transformation of yeast by a plasmid (pUZ3, see Fig. 2 of the preceeding report, ref. 4) containing a 5.2 kb genomic insert starting from the XhoII site at nucleotide -22 (Fig. 1) and extending beyond the NUC1 ORF failed to increase nuclease levels, whereas transformation by a plasmid (pUZ2) containing the same sequences plus 1.5 kb of DNA upstream of this XhoII site led to a marked elevation in nuclease levels (4). This indicates that although both plasmids contain NUC1 coding sequences only the latter plasmid contains the NUC1 promoter region and can produce functional transcripts of NUC1 in transformants.

To locate the termini of the NUC1 mRNA we first localized their approximate positions by mapping the transcript for homology to restriction

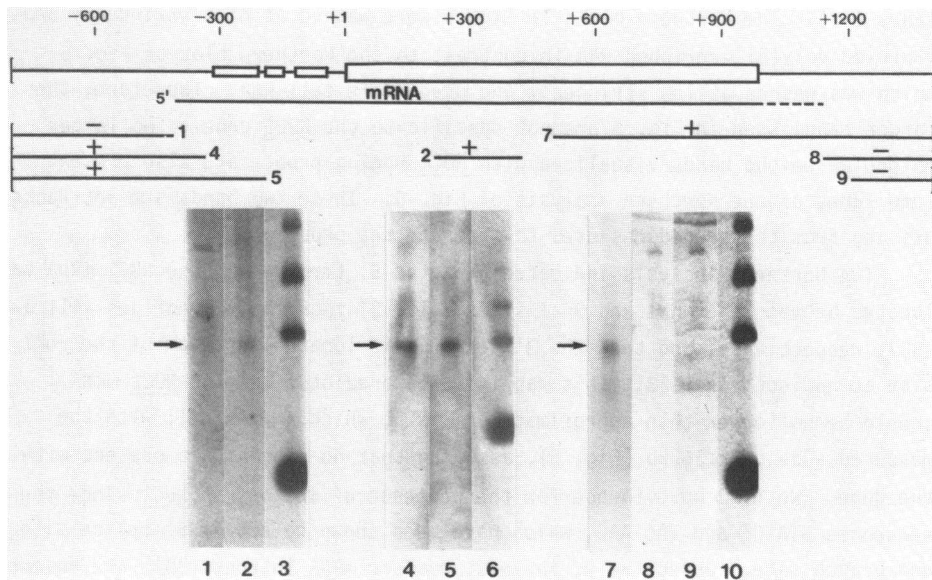

Fig. 6. Localization of NUC1 mRNA 5' and 3' ends: Northern Mapping. A poly(A)$^{+}$-enriched RNA fraction from a NUC1 overproducing transformant was electrophoresed through gels containing 1.5% agarose and 2.2M formaldehyde and blotted to charged nylon membranes (see Methods). DNA fragments spanning the NUC1 ORF region were labeled by random oligomer priming and hybridized to individual lanes of the Northern blots (lower portion of figure). The location of the NUC1 signal is indicated by the arrow. A schematic representation of the NUC1 region is displayed above the blots. The locations of the DNA probes relative to this region are indicated and numbered to correspond to the lanes in the lower portion of the figure. A + above the probe indicates hybridization to NUC1 message, a - indicates no hybridization. The position of the NUC1 message is displayed, as determined from the pattern of hybridization. Because probe 1 does not hybridize to NUC1 mRNAs, but probes 4 and 5 do, the 5' end of the NUC1 message must be located downstream of the HhaI site defined by the 3' end of probe 1. The lack of hybridization of probes 8 and 9 indicate the 3' end of NUC1 message (indicated by the interrupted portion of the line) is between the end of the large ORF and the FokI site defined by the 5' end of probe 8. Lane 1: BglII-HhaI probe, 3' end at nt -408 (Fig. 2); Lane 2: PstI-PstI probe, nt +215 to nt +425; Lane 4: BglII-DraI probe, 3' end at nt -355; Lane 5: BglII-NdeI probe, 3' end at nt -189; Lane 7: PstI-PstI probe, 5' end at nt +425; Lane 8: FokI-PstI probe, 5' end at nt +1136; Lane 9: HaeIII-PstI probe, 5' end at nt +1210; Lanes 3,6,10: RNA standards. Probes 1,4,5,7,8, and 9 extend beyond the limits of the NUC1 region displayed.

fragments derived from either the 5' or 3' portions of the gene (Fig. 6). Note that in this Northern analysis we detect the NUC1 mRNA but the additional larger molecular weight bands observed in Fig. 5 are absent, although in both analyses the same NUC1 probe was used (compare Fig. 6, lane 2 with Fig. 5,

lanes 1-3). The Northern blots in Fig. 6 were washed at high stringency and employed poly(A)$^+$-enriched RNA in contrast to the Northern blot of Fig. 5 which was washed at low stringency and used whole cell RNA. Therefore, the larger bands seen in Fig. 5 are not specific to the NUC1 gene. Two larger molecular weight bands visualized with NUC1 region probes are also present in some lanes of the Northern analysis of Fig. 6. These two bands are artifacts arising from the procedures used to generate the probes.

The Northern analysis indicated that the 5' terminus of the NUC1 mRNA was located between the HhaI and DraI sites (Fig. 2) found at nucleotides -411 and -337, respectively, and that the 3' terminus is located upstream of the FokI site at nucleotide +1123. This mapping data predicts that the NUC1 mRNA should be no longer than approximately 1.5 kb, which agrees well with the measured size of 1.35 kb (Fig. 5), assuming that no introns are present within the gene. We find no evidence for the presence of introns in NUC1 since the sequences GTATCT and TACTAAC, which have been shown to act as 5' splice site and branch site, respectively, in yeast nuclear mRNA introns (30), are absent from the NUC1 region.

The 5' terminus of NUC1 mRNA was mapped by an S1 protection experiment (Fig. 7). This analysis showed that the 5' terminus is slightly heterogeneous, typical of yeast mRNAs (31), and that the 5' termini map to the region between nucleotides -420 to -405 (Fig. 2). The S1 experiment described in Fig. 7 utilized a probe whose 5' end was approximately 230 nucleotides upstream from the initiating AUG of the NUC1 mRNA. To rule out that we might therefore have missed transcripts beginning within those 230 nucleotides we also performed an S1 protection experiment using as a probe a 1.6 kb NciI fragment in which the labeled 5' end is 176 nucleotides into the NUC1 reading frame. With this probe only one S1 protected fragment of about 600 nucleotides was observed (data not shown) indicating that the major detectable 5' termini of the NUC1 mRNA map to the region around nucleotide -410. The results of the S1 analysis were confirmed by an analysis of cDNA's synthesized using an oligonucleotide primer homologous to the mRNA sequences located from nucleotides -249 to -231. The major primer extension products were a pair of bands approximately 165-170 nucleotides in length (data not shown). This analysis places the 5' termini of the major NUC1 mRNAs to the vicinity of nucleotides -395 to -400, in close agreement with the S1 protection analysis.

These mapping studies show that the NUC1 mRNA has an unusually long 5' leader sequence of nearly 400 bases. Inspection of the DNA sequence 5' to the transcriptional start sites shows that the nearest "TATA" sequence occurs some

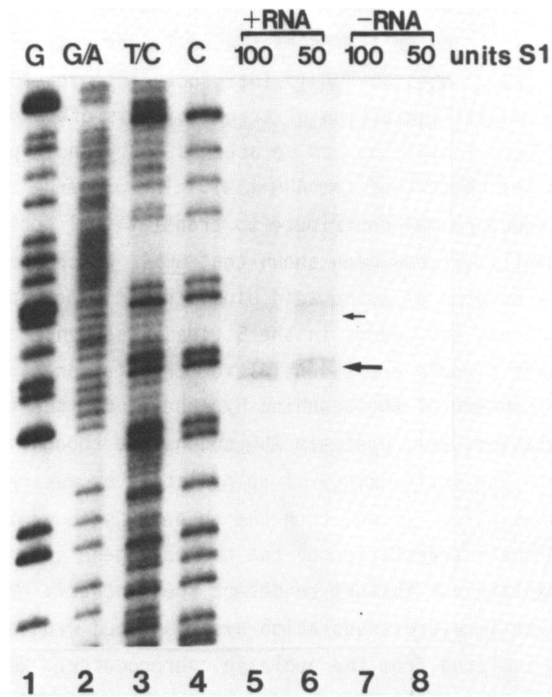

Fig. 7. Mapping of NUC1 mRNA 5' ends: S1 nuclease analysis. A 5' end labeled NdeI fragment from the NUC1 ORF region was hybridized to poly(A)$^+$ RNA from a nuclease overproducing transformant and digested with 50-100 units of S1 nuclease (see Methods). Nuclease resistant bands in the digests were visualized by autoradiography after electrophoresis through an 8% polyacrylamide-urea sequencing gel next to Maxam-Gilbert sequencing reactions of the same NdeI fragment. Five fragments protected from the activity of S1 can be observed (arrows). The two smallest fragments (large arrow) are the most abundant species. The locations of the 5' termini of NUC1 mRNA calculated from the sizes of the fragments are indicated by arrows in Fig. 2. Lanes 1-4: Maxam-Gilbert sequencing reactions of a 5' end labeled NdeI fragment (the location of the NdeI site is displayed in Fig. 1); Lanes 5 and 6: S1-protected fragments using poly(A)$^+$ RNA digested with either 100 units or 50 units of nuclese S1, respectively. Lanes 7 and 8: tRNA control (no poly(A)$^+$ RNA), digested with either 100 units or 50 units of nuclease S1, respectively.

170 nucleotides upstream at nucleotide -576 at which begins a canonical TATAAA sequence (Fig. 2). In yeast, as in other eukaryotes, such a sequence motif is involved in mRNA start-site selection (32). Generally in yeast it is located within 30-150 nucleotides from the point of transcriptional initiation (32).

Scanning the DNA sequence encoding the 5' leader of the NUC1 mRNA reveals the presence of three open reading frames initiated by an ATG and

terminated by a TAA for ORF1 and ORF2 and by a TAG for ORF3 (Figs. 1 and 2). AUGs are usually not found in the untranslated leader region of most eukaryotic mRNAs (33). When an AUG is introduced into yeast leader sequences by mutation, substantial inhibition of translation is often observed (34). In mammalian cells, such inhibition can be attenuated by an in phase termination codon as well as the context of the AUG (39). The presence of these ORF's in the NUC1 leader sequence may contribute to translational control of NUC1 expression. Recently, it has been shown that GCN4, which encodes a transcriptional activator of amino acid biosynthetic genes in yeast (35), contains four upstream AUG codons in the 5' leader sequence and that these translational start signals are essential for translational repression of GCN4 (36,37). In the context of the scanning hypothesis for the initiation of eukaryotic translation (38), upstream AUG codons are thought to inhibit mRNA translation due to the inefficiency of reinitiation by eukaryotic ribosomes at internal AUG codons (37). If so, then the three AUGs found in the long NUC1 leader may well impair translation of the nuclease gene sequences. This hypothesis may explain our failure to detect the synthesis of nuclease in a rabbit reticulocyte *in vitro* translation system primed with a saturating level of poly(A)$^+$ mRNA isolated from the nuclease overproducer. We note, however, that each of the ORF's in the NUC1 leader has a codon adaptation index (0.176, 0.171 and 0.244 for ORF1, ORF2, and ORF3, respectively) higher than that of NUC1 itself (0.139). Given the low abundance of NUC1 mRNA and of NUC1 enzyme in mitochondria, it may be that one or more of the encoded peptides specified by those ORF's are functional, possibly in terms of regulating mitochondrial nuclease function. As mentioned in the preceeding report (4) the overproducing strain, in which the entire NUC1 gene plus flanking sequences is present on a multicopy plasmid, is phenotypically normal despite 20-40 times more nuclease in its mitochondria compared to a wild-type. If the ORF peptides are stoichiometric to the nuclease then they may similarly be increased in that strain. We are testing whether the upstream ORF's contribute to regulation of the mitochondrial nuclease by constructing a NUC1 gene missing those sequences.

A search of a protein data base (17) for proteins homologous to the NUC1 gene product (and also to each of the leader ORF's) failed to detect any with significant homology. Furthermore, low stringency hybridization of NUC1 gene probes to yeast genomic blots indicated the absence of any other sequences in yeast with homology to NUC1 (data not shown). Given the very similar nucleases also present in the mitochondria of *Neurospora* and mammals, it is

possible that sequences homologous to NUC1 may be found in their nuclear genomes. Preliminary results, however, have failed to detect such sequences. The location of the NUC1 gene in the yeast genome is on the distal portion of the left arm of Chromosome X (M. Olson, personal communication).

ACKNOWLEDGEMENTS

We gratefully acknowledge the technical assistance of G. Denniger and the assistance of S. McKenzie in the preparation of this manuscript. This work was supported in part by USPHS GM 34096 and GM 37740.

*To whom correspondence should be addressed

REFERENCES

1. Rosamond, J. (1981) Eur. J. Biochem. 120, 541-546.
2. von Tigerstrom, R. (1982) Biochem. 21, 6397-6403.
3. Dake, E., Hofmann, T.J., McIntire, S., Hudson, A., and Zassenhaus, H.P. (1988) J. Biol. Chem., in press.
4. Zassenhaus, H.P., Hofmann, T.J., Uthayashanker, R., Vincent, R.D., and Zona, M. (1988) Nucleic Acids Research, manuscript submitted.
5. Linn, S.M. and Roberts, R.J. (1985) in Nucleases, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
6. Durphy, M., Manley, P.N., and Friedberg, E.C. (1974) J. Cell. Biol. 62, 695-706.
7. Chow, T.Y.K. and Fraser, M. (1983) J. Biol. Chem. 258, 12010-12018.
8. Cummings, O.W., King, T.C., Holden, J.A., and Low, R.L. (1987) J. Biol. Chem. 262, 2005-2015.
9. Low, R.L., Cummings, O.W., and King, T.C. (1987) J. Biol. Chem. 262, 16164-16170.
10. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
11. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) in Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
12. Zassenhaus, H.P., Butow, R.A., and Hannon, Y. (1982) Anal. Biochem. 125, 125-130.
13. Krawetz, S.A. and Anwar, R.A. (1984) BioTechniques 2, 342-347.
14. Nevins, J.R. (1980) Methods Enzymol. 65, 768-785.
15. Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
16. Feinberg, A.P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266-267.
17. Barker, W.C., Hunt, L.T., George, D.G., Yeh, L.S., Chen, H.R., Blomquist, M.C., Seibel-Ross, E.I., Elzanowski, A., Hong, M.H., Ferrick, D.A., Blair, J.K., Chen, S.L., and Ledley, R.S. (1987) Protein Sequence Database of the Protein Identification Resource, Release 12.0 (4253 Sequences), National Biomedical Research Foundation, Washington, D.C.
18. Lipman, D.J. and Peason, W.R. (1985) Science 227, 1435-1441.
19. Wilbur, W.J. and Lipman, D.J. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 726-731.
20. Needleman, S.B. and Wunsch, C.D. (1970) J. Mol. Biol. 48, 443-453.
21. Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1979) Atlas of Protein Sequence and Structure (Dayhoff, M.O., ed.) Vol. 5, Suppl. 3, pp. 345-362, National Biomedical Research Foundation, Washington, D.C.

22. Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105-132.
23. Eisenberg, D., Schwartz, E., Komaromy, M., and Wall, R. (1984) J. Mol. Biol. 179, 125-142.
24. Eisenberg, D., Weiss, R.M., and Terwilliger, T.C. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 140-144.
25. Reynolds, J.A. and Tanford, C. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 4467-4471.
26. Dayhoff, M.O., Hunt, T.L., and Hurst-Calderone, S. (1979) Atlas of Protein Sequence and Structure (Dayhoff, M.O., ed.) Vol. 5, Suppl. 3, pp. 363-369, National Biomedical Research Foundation, Washington, D.C.
27. Hurt, E.C. and van Loon, A.P.G.M. (1986) Trends Biochem. Sci. 11, 204-207.
28. von Heijne, G. (1986) EMBO J. 5, 1335-1342.
29. Sharp, P.M. and Li, W.-H. (1987) Nucl. Acids. Res. 15, 1281-1295.
30. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. (1986) Ann. Rev. Biochem. 55, 1119-1150.
31. Faye, G. Leung, D.W., Tatchell, K., Hall, B.D., and Smith, M. (1981) Proc. Natl. Acad. Sci. U.S.A. 81, 7860-7864.
32. Sentenac, A. and Hall, B. (1982) In Molecular Biology of the Yeast Saccharomyces cerevisiae: Metabolism and Gene Expression (Strathern, J.N., Jones, E.W., and Broach, J.R., eds.), pp 561-606, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. Kozak, M. (1983) Microbiol. Rev. 47, 1-45.
34. Sherman, F., and Stewart, J.W. (1983). In Molecular Biology of the Yeast Saccharomyces cerevisiae: Metabolism and Gene Expression (Strathrn, J.N., Jones, E.W. and Broach, J.R., eds.) pp. 301-334, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
35. Hinnebusch, A.G. and Fink, G.R. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 5374-5378.
36. Thireos, B., Driscoll-Penn, M., and Greer, H. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 5096-5100.
37. Mueller, P.P. and Hinnebusch, A.G. (1986) Cell 45, 201-207.
38. Kozak, M. (1978) Cell 15, 1109-1123.
39. Kozak, M. (1984) Nucleic Acids Research 9:3873-3893.