YEAST VOL. **13:** 777–781 (1997)

Yeast Sequencing Reports

Cloning and Characterization of the *KlDIM1* Gene from *Kluyveromyces lactis* Encoding the m_2^6A Dimethylase of the 18S rRNA

ISABELLE HOUSEN[1], DOMINIQUE DEMONTÉ[1], DENIS LAFONTAINE[2] AND JEAN VANDENHAUTE[1]*

[1]*Unité de Recherches en Biologie Moléculaire, Laboratoire de Génétique Moléculaire, Facultés Notre-Dame de la Paix, Rue de Bruxelles 61, B-5000 Namur, Belgium*
[2]*European Molecular Biology Laboratory (EMBL), D-69117 Heidelberg, Germany*

Received 25 October 1996; accepted 20 January 1997

The *KlDIM1* gene encoding the m_2^6A rRNA dimethylase was cloned from a *Kluyveromyces lactis* genomic library using a PCR amplicon from the *Saccharomyces cerevisiae ScDIM1* gene as probe. The *KlDIM1* gene encodes a 320-amino acid protein which shows 81% identity to ScDim1p from *S. cerevisiae* and 25% identity to ksgAp from *Escherichia coli*. Complementation of the kasugamycin-resistant *ksgA*-mutant of *E. coli* lacking dimethylase activity demonstrates that KlDim1p is the functional homologue of the bacterial enzyme. Multiple alignment of dimethylases from prokaryotes and yeasts shows that the two yeast enzymes display distinctive structural motives including a putative nuclear localization signal. © 1997 by John Wiley & Sons, Ltd.

Yeast 13: 777–781, 1997.
No. of Figures: 2. No. of Tables: 0. No. of References: 30.

KEY WORDS — dimethylase; ribosomal RNA; *Kluyveromcyes lactis*; kasugamycin

INTRODUCTION

During ribosome biogenesis, the pre-rRNA undergoes a series of cleavages and various nucleotide modifications (Venema and Tollervey, 1995). The only modification common to pro- and eukaryotes is the dimethylation of two adjacent adenosines ($m_2^6A_{1779}m_2^6A_{1780}$; yeast numbering) present at the 3′ end of the SSU-rRNA. The *Saccharomyces cerevisiae ScDIM1* gene encoding the ScDim1p enzyme responsible for this dimethylation was cloned (Lafontaine *et al.*, 1994) by complementation of the kasugamycin-resistant *Escherichia coli* *ksgA*$^-$ mutant lacking dimethylase activity (van Buul and van Knippenberg, 1985). In contrast with *E. coli*, where cells lacking dimethylase activity are viable, disruption of the yeast *ScDIM1* gene turned out to be lethal (Lafontaine *et al.*, 1994) and the protein was shown to be essential for the early steps of 18S rRNA maturation (Lafontaine *et al.*, 1995). In addition, despite the conservation between the ksgAp and ScDim1p amino-acid sequences (27% identity, 50% similarity), a *DIM1::ksgA* translational fusion, where the *DIM1* open reading frame (ORF) is substituted by *ksgA*, does not complement the growth defect on glucose of a *GAL10::DIM1* mutant (I. Housen *et al.*, in preparation). As part of our efforts to understand the molecular basis for the specificity of the yeast enzymes, we decided to clone the dimethylase gene from another yeast species. In

*Correspondence to: J. Vandenhaute.

CCC 0749–503X/97/080777–05 $17.50

this paper, we describe the cloning and sequencing of the *Kluyveromyces lactis KlDIM1* gene encoding the structural and functional homologue of m_2^6A dimethylases from both *E. coli* and *S. cerevisiae*.

MATERIALS AND METHODS

Strains, plasmids and procedure for E. coli *complementation*

Complementation in *E. coli* was performed using strain DL101 (*supE thi* Δ(*lacI-proAB*) F′[*traD36 proA*$^+$*B*$^+$ *lacI*^{q} lac ZΔM15] *ksgA*$^-$ *pyrF*$^-$ and the following plasmids: pUC8-ksgA7-URA3 carrying the *ksgA* gene and pGIDL31.42 carrying the *ScDIM1* gene as positive controls (Lafontaine, 1995), pFL44S carrying the *URA3* marker alone as negative control (Bonneaud *et al.*, 1991) and pGIIH23 carrying the *KlDIM1* gene isolated from a *K. lactis* genomic library (Bianchi *et al.*, 1987). The complementation test was conducted as previously described (Lafontaine *et al.*, 1994).

General DNA techniques

Standard DNA manipulations were performed as described by Sambrook *et al.* (1989) using the *E. coli XL1-blue* strain (Stratagene). The *K. lactis* genomic library (Bianchi *et al.*, 1987) was screened by filter hybridization using nylon filters washed at room temperature using 0·2 × SSC/0·1% SDS as final buffer. Sequencing was performed on denatured plasmid DNA with the T7 sequencing kit (Pharmacia) using synthetic oligonucleotides. PCR amplification conditions were standard. The two selected primers hybridize to the 5′ and 3′ ends of the *ScDIM1* ORF (coordinates 35–52 and 749–732, respectively; Lafontaine *et al.*, 1994). The 700-bp PCR amplicon obtained was radiolabelled using the multi-prime DNA labelling kit (Gibco BRL).

DNA and protein sequence analyses

DNA sequence data obtained from sequencing gels were compiled and analysed by the DNA Strider 1.2 program (Marck, 1988). For homology searches, the BLAST program (Altschul *et al.*, 1990) was used to screen the amino acid sequence version of the Non-Redundant database (NRDB; National Center for Biotechnology Information, NIH). The dimethylase sequences were analysed with the GCG 'Gap' software (Devereux *et al.*, 1984) for paired alignment and with the Match-Box software for the multiple alignment (Depiereux and Feytmans, 1991, 1992).

RESULTS AND DISCUSSION

Cloning of the KlDIM1 *gene*

A 0·7-kb DNA fragment was amplified by PCR from pGIDL31.42 using two primers from both ends of the *S. cerevisiae ScDIM1* ORF (see Materials and Methods). This fragment hybridized as a single band to *K. lactis* genomic DNA in a Southern blot (data not shown) and was used for screening a *K. lactis* CBS 2359 genomic library established in the shuttle vector Kep6 (Bianchi *et al.*, 1987). One positive clone was shown to restore kasugamycin sensitivity to the *ksgA*$^-$ *E. coli* mutant strain DL101, indicating that the *KlDIM1* gene is the functional homologue of the *ksgA* gene. In order to test whether the KlDim1p enzyme could also substitute for the ScDim1p dimethylase in *S. cerevisiae*, we transformed the *pGAL10::dim1* conditional lethal strain YDL150 (Lafontaine *et al.*, 1995) with a plasmid encoding *KlDIM1* and showed that growth was restored in glucose (data not shown).

Sequence of the KlDIM1 *gene*

The nucleotide sequence of the *KlDIM1* gene and the amino acid sequence of the encoded protein are shown in Figure 1. A 960-bp ORF (position 508–1468) was identified. There is a potential transcription initiation site matching to the consensus PuPuPyPuPu (Guarente, 1992) at positions 459 to 462 and a TATA box (Struhl, 1989) at residues 437 to 441. We have also identified an UAS_{rpg}-like sequence (Bergkamp-Steffens *et al.*, 1992) at positions 33 to 45. The initiation codon AUG is preceded by an A at −3 and followed by a G at +4, in good agreement with the consensus described by Kozak (1989). The deduced protein is 329 amino acids long with a calculated molecular weight of 36300.

Pairwise and multiple alignments of m_2^6A *rRNA dimethylases*

Similarities between the KlDim1 protein and the sequences available in the databases were searched for with the BLAST algorithms (Altschul *et al.*, 1990). The KlDim1p sequence shows significant similarity (optimized score>100) to the *S. cerevisiae* ScDim1p sequence (Lafontaine *et al.*, 1994) and to the prokaryotic *ksgA* sequences from *E. coli*

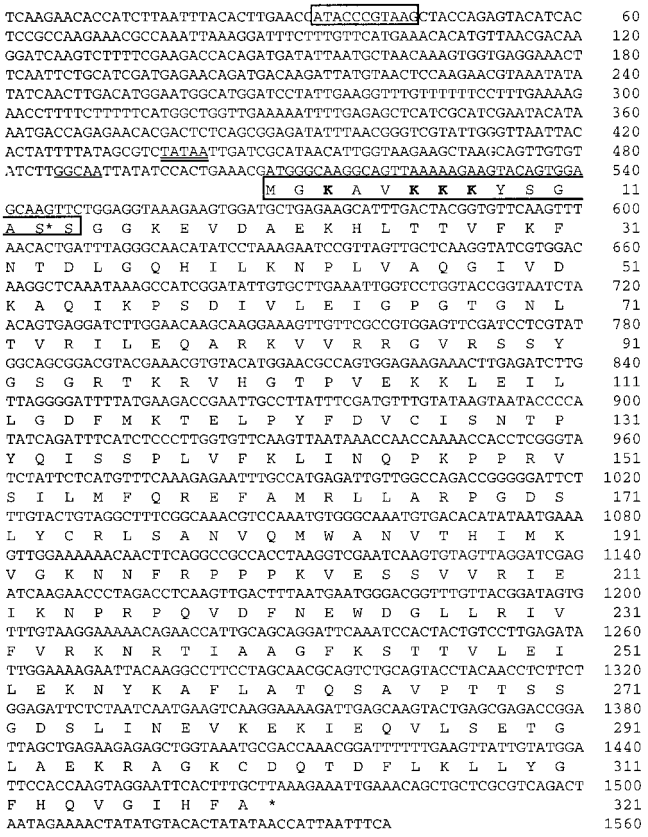

Figure 1. Sequences of the *KlDIM1* gene and its encoded protein. The putative UAS_{rpg} is boxed, the putative TATA box is double-underlined and the putative +1 transcription region is underlined. The lysine-rich-N-terminus is boxed with the putative NLS in bold. The asterisk denotes the termination codon. The *KlDIM1* nucleotide sequence is available from GenBank through accession number Z 68294.

(van Buul and van Knippenberg, 1985), *Bacillus subtilis* (Ogasawara *et al*., 1994), *Haemophilus influenzae* (Fleischmann *et al*., 1995), *Mycoplasma capricolum* (Miyata *et al*., 1993) and *Mycoplasma genitalium* (Fraser *et al*., 1995). The yeast Dim1 proteins are closely related (90% of similarity and 81% of identity), as are the ksgA proteins from *E. coli* and *H. influenzae* (82% similarity and 72% identity). The ScDim1p and KlDim1p aligned with ksgAp from *E. coli* present respectively 27% and 25% of identity. The other dimethylases aligned pairwise show approximately 50% similarity and 30% identity (GCG 'Gap' software; Devereux *et al*., 1984).

To investigate further the extent of conservation between KlDim1p and the other dimethylases, multiple alignment was performed using the Match-Box package (Depiereux and Feytmans, 1991, 1992). It can be observed (Figure 2) that eight open boxes (numbered I to VIII) corresponding to groups of identical or similar amino acids are predicted to be part of similar structural motives in pro- and eukaryotic enzymes. In all seven sequences, a glycine-rich motif preceded by a hydrophobic segment is conserved (box II); this motif is reported as the consensus for binding of the methyl-donor S-adenosyl-methionine UU[D/E]xGxGx[G/A]xU[S/T]xxUx (U being a bulky aliphatic amino acid and x representing any residue; Ingrosso *et al*., 1989). In box IV, the NxPY sequence is also strictly conserved and matches the motif [N/D/S]PP[F/Y/W] known to be critical for catalysing methyl transfer in N^6-adenine DNA methyltransferase (Willcock *et al*., 1994). The

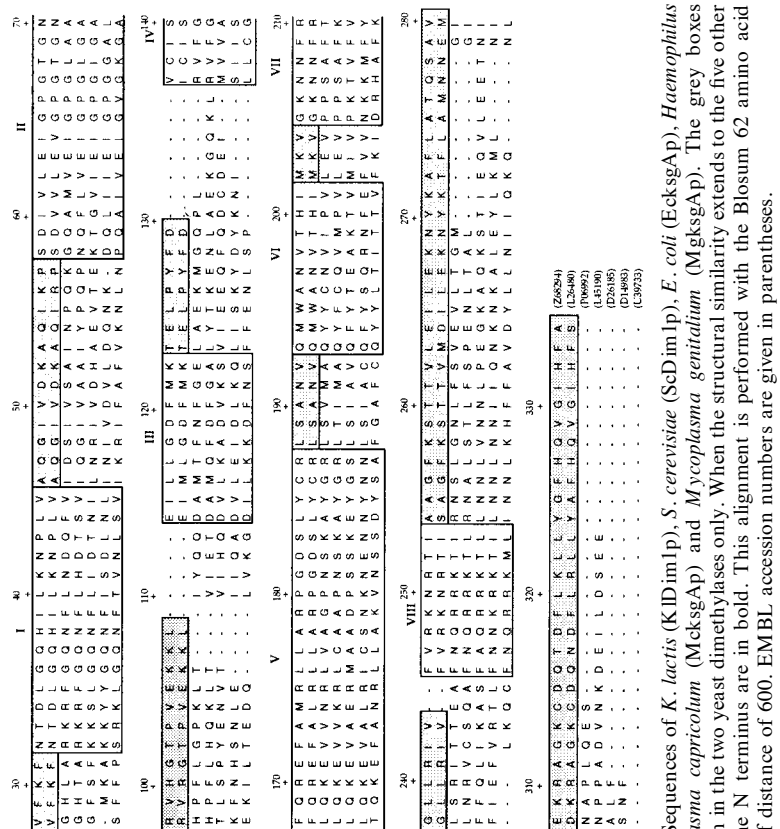

Figure 2. Multiple alignment of m_2^6A rRNA dimethylases. Sequences of *K. lactis* (KlDim1p), *S. cerevisiae* (ScDim1p), *E. coli* (EcksgAp), *Haemophilus influenzae* (HisksgAp), *Bacillus subtilis* (BsksgAp), *Mycoplasma capricolum* (McksgAp) and *Mycoplasma genitalium* (MgksgAp). The grey boxes correspond to regions predicted to be in a similar conformation in the two yeast dimethylases only. When the structural similarity extends to the five other sequences, the boxes are unshaded. The lysine residues at the N terminus are in bold. This alignment is performed with the Blosum 62 amino acid substitution matrix (Henikoff and Henikoff, 1992), at a cutoff distance of 600. EMBL accession numbers are given in parentheses.

function of the seven other conserved regions is unknown.

Several regions turn out to be structurally and perhaps functionally different in pro- and eukaryotic enzymes (Figure 2, grey boxes). It should be noted that serine 13 or 14 respectively in the *K. lactis* and *S. cerevisiae* sequences appears as a potential protein-kinase-C-phosphorylation site (Woodgett *et al.*, 1986). The sequence KxxKKK (residues 3 to 8) found only in the yeast N-termini seems in agreement with the general NLS structure, *i.e.* NLS is usually short, contains a high proportion of basic residues and is not located at a specific site within the protein (Hicks and Raikhel, 1995). No typical acidic-serine-rich stretches, GAR domains (Lapeyre *et al.*, 1987; Shaw and Jordan, 1995; Yan and Mélèse, 1993) nor RNA recognition motives (RRM; Kenan *et al.*, 1991), known as nucleolar protein signatures (Shaw and Jordan, 1995), are found.

Since it was shown recently (Lafontaine *et al.*, 1995) that ScDim1 protein is required for nucleolar cleavage A_1 and A_2 of the rRNA precursors, it is tempting to speculate that the sequence KxxKKK addresses ScDim1p to the nucleus.

ACKNOWLEDGEMENTS

We thank L. Guarente for providing the strain BWG1-7A, M. Wésolowski-Louvel for providing the *K. lactis* library, and P. Briffeuil, X. De Bolle and J. Delcour for fruitful discussions. We are also grateful to M. Dewez for valuable technical assistance.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). *J. Mol. Biol.* **215,** 403–410.

Bergkamp-Steffens, G. K., Hoekstra, R. and Planta, R. J. (1992). *Yeast* **8,** 903–922.

Bianchi, M. M., Falcone, C., Jie, C. X., Wésolowski-Louvel, M., Frontali, L. and Fukuhara, H. (1987). *Curr. Genet.* **12,** 185–192.

Bonneaud, N., Ozier-Kalogeropoulos, O., Li, G., Labouesse, M., Minvielle-Sebastia, L. and Lacroute, F. (1991). *Yeast* **7,** 609–615.

Depiereux, E. and Feytmans, E. (1991). *Protein Eng.* **4,** 603–613.

Depiereux, E. and Feytmans, E. (1992). *CABIOS* **8,** 501–509.

Devereux, J., Haeberli, P. and Smithies, O. (1984). *Nucl. Acids Res.* **12,** 387–395.

Fleischmann, R. D., *et al.* (1995). *Science* **269,** 496–512.

Fraser, C. M., *et al.* (1995). *Science* **270,** 397–403.

Guarente, L. (1992). In (Broach, J. R., Pringle, J. R. and Jones, E. (Eds), *The Molecular and Cellular Biology of the Yeast S. cerevisiae*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 49–98.

Henikoff, S. and Henikoff, J. G. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 10915–10919.

Hicks, G. R. and Raikhel, N. V. (1995). *Annu. Rev. Cell Dev. Biol.* **11,** 155–188.

Ingrosso, D., Flower, A. V., Bleibaum, J. and Clarke, S. (1989). *J. Biol. Chem.* **264,** 20131–20139.

Kenan, D. J., Query, C. C. and Keene, J. D. (1991). *Trends Biochem. Sci.* **16,** 214–220.

Kozak, M. (1989). *J. Cell Biol.* **108,** 229–241.

Lafontaine, D., Delcour, J., Glasser, A.-L., Desgrès, J. and Vandenhaute, J. (1994). *J. Mol. Biol.* **241,** 492–497.

Lafontaine, D., Vandenhaute, J. and Tollervey, D. (1995). *Genes Dev.* **9,** 2470–2481.

Lafontaine, D. (1995). PhD Thesis, Facultés Universitaires, Namur (Belgium).

Lapeyre, B., Bourbon, H. and Amalric, F. (1987). *Proc. Natl. Acad. Sci. USA* **84,** 1472–1476.

Marck, C. (1988). *Nucl. Acids Res.* **16,** 1829–1836.

Miyata, K., Sano, R. and Fukumura, T. (1993). *Nucl. Acids Res.* **21,** 4816–4823.

Ogasawara, N., Nakai S. and Yoshikawa, H. (1994). *DNA Res.* **1,** 1–14.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual,* 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shaw, P. J. and Jordan, E. G. (1995). *Annu. Rev. Cell Dev. Biol.* **11,** 93–121.

Struhl, K. (1989). *Annu. Rev. Biochem.* **58,** 1051–1077.

van Buul, C. P. and van Knippenberg, P. H. (1985). *Gene* **38,** 65–72.

Venema, J. and Tollervey, D. (1995). *Yeast* **11,** 1629–1650.

Willcock, D. F., Dryden, D. T. F. and Murray, N. E. (1994). *EMBO J* **13,** 3902–3908.

Woodgett, J. R., Gould, K. L. and Hunter, T. (1986). *Eur. J. Biochem.* **161,** 2987–2993.

Yan, C. and Mélèse, T. (1993). *J. Cell Biol.* **123,** 1081–1091.