Yeast Sequencing Reports

Cloning and Characterization of the KlDIM1 Gene from Kluyveromyces lactis Encoding the m$_2^6$A Dimethylase of the 18S rRNA

ISABELLE HOUSEN$^1$, DOMINIQUE DEMONTÉ$^1$, DENIS LAFONTAINE$^2$ AND JEAN VANDENHAUTE$^*$

$^1$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$_2^6$A 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.

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$_2^6$A$_{1779}$m$_2^6$A$_{1780}$; 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 of E. coli 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
© 1997 by John Wiley & Sons, Ltd
this paper, we describe the cloning and sequencing of the \textit{Kluyveromyces lactis} \textit{KIDIM1} gene encoding the structural and functional homologue of \textit{m}_2^6\textit{A} dimethylases from both \textit{E. coli} and \textit{S. cerevisiae}.

**MATERIALS AND METHODS**

\textbf{Strains, plasmids and procedure for E. coli complementation}

Complementation in \textit{E. coli} was performed using strain DL101 (\textit{supE} \textit{thi} \textit{Δ(lacI-proAB)} F[\textit{traD36 proA}^{+}\textit{B}^{+} \textit{lac}^Z \textit{ΔAM15}] \textit{ksgA} \textit{pyrF}^−) and the following plasmids: pUC8-\textit{ksgA7-URA3} carrying the \textit{ksgA} gene and pGIDL31.42 carrying the \textit{ScDIM1} gene as positive controls (Lafontaine, 1995), pFL44S carrying the \textit{URA3} marker alone as negative control (Bonneaud et al., 1991) and pGIIH23 carrying the \textit{KIDIM1} gene isolated from a \textit{K. lactis} genomic library (Bianchi et al., 1987). The complementation test was conducted as previously described (Lafontaine et al., 1994).

\textbf{General DNA techniques}

Standard DNA manipulations were performed as described by Sambrook \textit{et al.} (1989) using the \textit{E. coli} \textit{XL1-blue} strain (Stratagene). The \textit{K. lactis} genomic library (Bianchi \textit{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 \textit{ScDIM1} ORF (coordinates 35–52 and 749–732, respectively; Lafontaine \textit{et al.}, 1994). The 700-bp PCR amplicon obtained was radiolabelled using the multi-prime DNA labelling kit (Gibco BRL).

\textbf{DNA and protein sequence analyses}

DNA sequence data obtained from sequencing gels were compiled and analysed by the DNA Strider 1.2 program (March, 1988). For homology searches, the BLAST program (Altschul \textit{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 \textit{et al.}, 1984) for paired alignment and with the Match-Box software for the multiple alignment (Depireux and Feytmans, 1991, 1992).

\section*{RESULTS AND DISCUSSION}

\textbf{Cloning of the KIDIM1 gene}

A 0.7 kb DNA fragment was amplified by PCR from pGIDL31.42 using two primers from both ends of the \textit{S. cerevisiae} \textit{ScDIM1} ORF (see Materials and Methods). This fragment hybridized as a single band to \textit{K. lactis} genomic DNA in a Southern blot (data not shown) and was used for screening a \textit{K. lactis} CBS 2359 genomic library established in the shuttle vector Kep6 (Bianchi \textit{et al.}, 1987). One positive clone was shown to restore kasugamycin sensitivity to the \textit{ksgA} \textit{−} \textit{E. coli} mutant strain DL101, indicating that the \textit{KIDIM1} gene is the functional homologue of the \textit{ksgA} gene. In order to test whether the \textit{KIDIM1p} enzyme could also substitute for the \textit{ScDIM1p} dimethylase in \textit{S. cerevisiae}, we transformed the \textit{pGAL10::dim1} conditional lethal strain YDL150 (Lafontaine \textit{et al.}, 1995) with a plasmid encoding \textit{KIDIM1} and showed that growth was restored in glucose (data not shown).

\textbf{Sequence of the KIDIM1 gene}

The nucleotide sequence of the \textit{KIDIM1} 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 \textit{PuPuPuPuPuPuPu} (Guarente, 1992) at positions 459 to 462 and a TATA box (Struhl, 1989) at residues 437 to 441. We have also identified an \textit{UAS}_{	ext{rpg}}-like sequence (Bergkamp-Steffens \textit{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.

\textbf{Pairwise and multiple alignments of \textit{m}_2^6\textit{A} rRNA dimethylases}

Similarities between the \textit{KIDIM1p} protein and the sequences available in the databases were searched for with the BLAST algorithms (Altschul \textit{et al.}, 1990). The \textit{KIDIM1p} sequence shows significant similarity (optimized score >100) to the \textit{S. cerevisiae} \textit{ScDIM1p} sequence (Lafontaine \textit{et al.}, 1994) and to the prokaryotic \textit{ksgA} sequences from \textit{E. coli}.
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 N6-adenine DNA methyltransferase (Willcock et al., 1994). The
Figure 2. Multiple alignment of m₅A 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 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 nuclear cleavage A, and A, 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. Węsolowski-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


© 1997 by John Wiley & Sons, Ltd