

## Molecular characterization of genes encoding enzymes of the metabolic pathways in *Neocallimastix frontalis*. An insight into the structure and organization of genes from an anaerobic fungus

R Durand, M Fischer, M Fèvre

Laboratoire de Biologie Cellulaire Fongique, Centre de Génétique Moléculaire et Cellulaire, CNRS, URM 106, Université Lyon I, Bâtiment 405, 69622 Villeurbanne Cédex, France

In axenic culture, the rumen anaerobic fungus, *N. frontalis* can utilize several monosaccharides and plant cell-wall polymers as carbon sources. The fermentation of glucose proceeds via the Embden-Meyerhof-Parnas pathway. Enzyme activities leading to the formation of succinate, lactate, formate and ethanol are associated with the cytoplasmic fraction, while the enzymes leading to the formation of the main fermentation products H<sub>2</sub>, CO<sub>2</sub>, acetate and the production of ATP are localized in microbodies identified as hydrogenosomes. *N. frontalis* has been used as a model to clone the genes expressed during the hydrolysis of complex cell wall polysaccharides and genes encoding enzymes of the metabolic pathways.

Three cDNAs have been characterized and identified as clones encoding enzymes of the glucose fermentation pathway: phosphoenolpyruvate carboxykinase (PEPCK) (Reymond et al, 1992, Gene, 110, 57-63), enolase (ENO) (Durand et al, 1995, Microbiology, 141, 1301-1308), hydro-genosomal  $\beta$ -succinyl-CoA synthetase ( $\beta$ SCS). The enolase gene (*enol*) and 5' flanking sequences have also been characterized. Studied of DNA base composition from *N. frontalis* and other anaerobic fungi have demonstrated these organisms to have the most A+T-rich genome of any organisms identified so far (15 mol% G+C) and it could be pointed out that this base composition might affect the amino acid composition of the encoded proteins. Clearly our results showed that this is not the case as sequence comparisons of the *N. frontalis* proteins with

aerobic counterparts revealed high degree of identity.

The codon usage observed in the *N. frontalis* genes was highly biased since only 33 codons were used in the  $\beta$ scs gene, 36 in the *enol* gene and 38 in the *pepck* gene with a strong preference for T in the third position.

The enolase gene (*enol*) was the first gene containing an intron isolated from an anaerobic fungus. This large intervening sequence (331 bp) showed a very low G+C content (14.2 mol%) but splicing boundaries followed of *enol* gene was found to be very similar to the organization of other genes from aerobic fungi. The G+C content of the coding sequences of *enol*, *pepck* and  $\beta$ scs genes (respectively 43.8, 44.6 and 43.8 mol%) was considerably higher than the G+C content of the 5' and 3' non-translated flanking sequences (10 mol % on the average) indicating that in the *N. frontalis* genome A+T rich «linker» sequences systematically punctuate regions that contain coding sequences.

The functional organization of the *N. frontalis enol* gene promoter was investigated by heterologous transient expression assays. The promoter could not drive the expression of the *lacZ* gene in Ascomycetes. However transient expression of the reporter gene was observed in the Oomycete *Saprolegnia monoica* and the *enol* gene promoter was shown to drive the expression of the *uidA* gene in *N. frontalis* may differ significantly from that of the Ascomycetes but that enough conservation exists within the lower fungi to allow a transient driven expression in an Oomycete fungus.