Neocallimastix frontalis enolase gene, enol: first report of an intron in an anaerobic fungus

R. Durand, M. Fischer, C. Rascle and M. Fèvre

A DNA clone containing a putative enolase gene was isolated from a genomic DNA library of the anaerobic fungus Neocallimastix frontalis. It was deduced from sequence comparisons that the enolase gene was interrupted by a large 331 bp intron. The enolase gene, termed enol, has an ORF of 1308 bp and encodes a predicted 436 amino acid protein. The deduced amino acid sequence shows high identity (71-71%) to those of enolases from the yeasts Saccharomyces cerevisiae and Candida albicans. The G+C content of the enolase coding sequence (43.8 mol%) is considerably higher than the G+C content of the intervening sequence (14.2 mol%) or the 5' and 3' non-translated flanking sequences (15.2 and 4.7 mol%, respectively). The codon usage of the N. frontalis enolase gene was very biased as has been found for the highly expressed genes of yeast and filamentous fungi. The gene has all the canonical features (polyadenylation signal, intron splicing boundaries) of genes isolated from aerobic filamentous fungi. Only one enolase gene could be detected in N. frontalis genomic DNA by Southern analysis with a homologous probe. RNA analysis detected a single enolase transcript of about 1.6 kb. When mycelium was grown on glucose, levels of enolase mRNA were markedly increased by comparison with enolase mRNA levels in mycelium grown on cellulose, suggesting that expression of the N. frontalis enolase gene was transcriptionally regulated by the carbon source.

Keywords: enolase, anaerobic fungus, rumen, Neocallimastix frontalis

INTRODUCTION

Although the main diet of herbivorous mammals is plant material, they are totally dependent on the microorganisms in their digestive tracts to break complex polysaccharides into small molecules prior to absorption. Rumen bacteria, protozoa and fungi are the main contributors to the process of digestion. The anaerobic fungus Neocallimastix frontalis, first isolated from the sheep rumen by Orpin (1975), has been intensively studied for its large production of a wide range of glycoside and polysaccharide hydrolases (Hébraud & Fèvre, 1988). In axenic culture, N. frontalis can utilize several monosaccharides and plant cell-wall polymers as carbon sources (Li & Heath, 1993). The fermentation of glucose by N. frontalis proceeds via the Embden–Meyerhof–Parnas pathway (Marvin-Sikkema et al., 1993). Enzyme activities leading to the formation of succinate, lactate, ethanol and formate are associated with the cytoplasmic fraction, while the enzymes leading to the formation of the main fermentation products H₂, CO₂, 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. A differential screening of a cDNA library lead to the isolation of several clones which were highly expressed during growth on cellulose. Several of these clones have been characterized by restriction analysis, partial sequencing and Northern blotting (Reymond et al., 1991). Three cDNAs have been identified as clones encoding enzymes of the glucose fermentation pathway: phosphoenolpyruvate carboxykinase (Reymond et al., 1991), hydrogenosomal β-succinyl-CoA synthetase (unpublished results) and enolase (this study).

Enolase (2-phospho-D-glycerate hydro-lyase, EC 4.2.1.11) is a ubiquitous enzyme that catalyses the conversion of 2-phosphoglycerate to phosphoenolpyruvate, the only dehydration step of the glycolytic
pathway. Enolase genes have been cloned from different sources: the yeasts Saccharomyces cerevisiae (Holland et al., 1981) and Candida albicans (Breit Mason et al., 1993), the human malaria parasite Plasmodium falciparum (Read et al., 1994), higher plants such as tomato and Arabidopsis (Van der Straeten et al., 1991), and several animals such as rat and chicken. In avian and mammalian tissues, three types of enolase isoenzyme (α, β, and γ) have been described, and the active form of the enzyme is a homodimer or a heterodimer as a consequence of the expression of genes in a tissue- and development-specific manner (Segil et al., 1988). The enzyme from S. cerevisiae is by far the best characterized and its crystal structure has been determined (Lebioda et al., 1989). Two structural genes, ENO1 and ENO2, are expressed differentially in yeast cells grown on glucose and in cells grown on gluconeogenic carbon sources. ENO2 is induced more than 20-fold in cells grown on glucose, whereas ENO1 expression is similar in cells grown on glucose or on gluconeogenic carbon sources (Cohen et al., 1986). Complex cis-acting regulatory sequences that bind distinct trans-acting factors have also been identified (Cohen et al., 1987; Carmen & Holland, 1994).

In spite of the isolation of various cDNAs, very little information is available on gene organization in rumen anaerobic fungi. A 3 kb genomic clone carrying the endoglucanase gene ce/B from the related species Neocallosiramia patriciarum was analysed by Zhou et al. (1994) and found to contain no intron. This paper describes the characterization of a genomic clone of the N. frontalis enolase gene. The primary structure of the enolase sequence of this filamentous fungus is compared with the yeast, plant and animal counterparts. To our knowledge, this is the first gene containing an intron isolated from an anaerobic fungus, so it was of interest to compare its organization to those of aerobic fungi. The overall organization of the enolase gene (intervening sequence, intron boundaries, polyadenylation sequence, codon usage) was found to be very similar to the organization of other genes from filamentous aerobic fungi.

**METHODS**

**Strains, plasmids, phage and media.** N. frontalis MCH3, isolated from sheep (Laboratoire de Microbiologie CRZV, INRA, Théix, France) was grown anaerobically at 39 °C for 2-5 d in culture flasks containing 500 ml synthetic liquid medium (Lowe et al., 1987) supplemented with 1-5 % (w/v) glucose or 0-6% (w/v) Avicel as carbon sources. S. cerevisiae strain MD4014C (α his-111 his-3-15 leu-2-3 leu-2-112 trp1 ura2 2u') provided by J. L. Souciet (IBMIC Strasbourg, France) was grown at 30 °C in liquid medium: 10 g yeast extract l-1, 20 g glucose or 0.6% (w/v) Avicel as carbon sources. The DNA was digested to completion by an overnight incubation with 10 U Xbal (Boehringer) and purified by phenol/chloroform extraction and ethanol precipitation. The genomic fragments were ligated in the λ Zap arms and the recombinant phages were packaged in vitro according to the manufacturer's instructions.

**Cloning and sequencing of the enolase gene.** The λ Zap and λ 1149 cDNA libraries that were previously constructed by Reymond et al. (1991) and the genomic library were screened by colony hybridization with the [α-32P]dCTP-labelled 1-3 kb EcoR1 incomplete enolase cDNA as a probe. This cDNA clone was previously isolated from a cDNA library (Reymond et al., 1991) and was identified by sequencing as a truncated cDNA enolase clone. The genomic inserts of positive plaques were isolated by in vivo excision using the helper phage R408 (Stratagene). After localization of the enol gene on the genomic clone by Southern hybridization, using the incomplete cDNA clone as a probe, subclones of this genomic fragment were constructed by conventional methods (Maniatis et al., 1982). Nucleotide sequences were determined by the dideoxy chain-termination method (Sanger et al., 1977) on different templates: double-stranded DNA, single-stranded DNA produced from the plBluescript plasmid (Stratagene), and deleted clones produced by unidirectional ExoIII digestion (Erase-a-Base, Promega) according to the manufacturer's instructions. When necessary, synthetic oligonucleotides were used to confirm the sequence. Sequence analyses were made by the Biasonic computer programs (Dessen et al., 1990).

**Southern analysis.** S. cerevisiae chromosomal DNA was isolated as described by Johnston (1988). N. frontalis and S. cerevisiae DNAs were digested to completion with the appropriate endonucleases and the resulting fragments were fractionated by electrophoresis on a 0-8% agarose gel and transferred to nitrocellulose membranes (Schleicher & Schuell). DNA probes were labelled with [α-32P]dCTP (111 TBq mmol-1, Amersham) by use of a random-primed labelling kit (Pharmacia). The DNA on the blots was hybridized with DNA probes at 42 °C in the presence of 0-5% (v/v) formamide as recommended by the manufacturer. After hybridization the membranes were washed twice for 20 min each time with 2 × SSC, 0-1% SDS and 0-2 × SSC, 0-1%
Northern analysis of *N. frontalis* RNA. Total RNA was isolated as described by Chirgwin et al. (1979). RNA (30 μg) were fractionated on a 1:2% (w/v) agarose gel in denaturing conditions (Maniatis et al., 1982), and transferred to nitrocellulose membrane using the Northern procedure (Maniatis et al., 1982). DNA probes were labelled with α[32P]dCTP by the random primer-labelling method (Feinberg & Vogelstein, 1983) using a Pharmacia kit. Hybridization was performed according to standard procedures (Maniatis et al., 1982): hybridization was done at 42 °C in the presence of 50% formamide, 5 × SSC, 1 × Denhardt’s solution, 20 mM sodium phosphate buffer, pH 6.8, 0.2% SDS and 200 μg yeast tRNA ml⁻¹; membranes were washed twice, for 20 min each time, at 42 °C with 2 × SSC, 0.1% SDS and 0.2 × SSC; 0.1% SDS, and exposed to autoradiographic film. The membranes were then washed in 0.1 × SSC, 0.1% SDS at 55 °C and re-hybridized with an rRNA genomic probe (R.D., unpublished results). The resultant autoradiograms were analysed by densitometry using the Gel Reader program (Appligene).

**RESULTS**

Cloning and sequencing of *N. frontalis* enolase gene

A 1.3 kb cDNA clone, pPR6.1, previously isolated by Reynolds et al. (1991), was sequenced and identified by sequence comparisons in the databases as an enolase clone. The deduced amino acid sequence of the cDNA clone had 69% identity with the amino acid sequence of a complete enolase cDNA clone. The partial cDNA clone was missing. Several screenings of the two genomic DNA libraries from *N. frontalis* indicated by shaded bars. Arrows indicate the positions and directions of sequencing. C, Clal; D, DraI; E, EcoRI; H, HindIII; K, KpnI; X, XbaI.

Fig. 1. Restriction map of the *N. frontalis* enolase clone and sequencing strategies. The coding region of the enol gene and the partial enolase cDNA are indicated by shaded bars. Arrows indicate the positions and directions of sequencing. C, Clal; D, DraI; E, EcoRI; H, HindIII; K, KpnI; X, XbaI.

RESULTS

Cloning and sequencing of *N. frontalis* enolase gene

A 1.3 kb cDNA clone, pPR6.1, previously isolated by Reynolds et al. (1991), was sequenced and identified by sequence comparisons in the databases as an enolase clone. The deduced amino acid sequence of the cDNA clone had 69% identity with the amino acid sequence of a complete enolase cDNA clone. The partial cDNA clone was missing. Several screenings of the two genomic DNA libraries from *N. frontalis* indicated by shaded bars. Arrows indicate the positions and directions of sequencing. C, Clal; D, DraI; E, EcoRI; H, HindIII; K, KpnI; X, XbaI.

Fig. 1. Restriction map of the *N. frontalis* enolase clone and sequencing strategies. The coding region of the enol gene and the partial enolase cDNA are indicated by shaded bars. Arrows indicate the positions and directions of sequencing. C, Clal; D, DraI; E, EcoRI; H, HindIII; K, KpnI; X, XbaI.

One clone pRDXX6.5 contained a 6.5 kb DNA insert and was used for further analysis. A number of overlapping restriction fragments from this clone hybridized to the cDNA pPR6.1 (data not shown) indicating that the genomic clone contained the enolase gene and large flanking sequences (Fig. 1). The nucleotide sequence of the 2.8 kb DraI-HindIII fragment of the pRDXX6.5 was determined (Fig. 2). The identification of the clone was confirmed by sequence analysis and comparisons with the NBRF database. The deduced amino acid sequence was identified as encoding an enolase on the basis of its high homology with the *S. cerevisiae* ENOl and EN02 sequences. By comparing the nucleotide and amino acid sequence with those from *S. cerevisiae*, one intron was assumed to be present in the coding region.

A large sequence (331 bp) from position +161 to +512 bp (relative to the first start codon) had intron-splicing boundaries matching the consensus sequences found in filamentous fungal genes (Gurr et al., 1987). The splicing signals were GTAAGT and TAG for the 5' and 3' intron boundaries, respectively, compared with the GTANGT and PyAG fungal consensus sequences. An internal sequence, TACTTAAA (+306 to +313 bp), presumed to be involved in lariat formation, showed some similarity with the sequence PyGCTAACN of other fungal genes (Gurr et al., 1987). Assuming this to be an intron, the deduced amino acid sequence of the coding region showed a high overall identity to those of *S. cerevisiae* ENOl and EN02 (71.5 and 71%, respectively). Therefore, we concluded that this gene does indeed encode an *N. frontalis* enolase; this gene was termed enol.

At the 3' end of the DNA sequence from bases +1742 to +1747, there is a putative AATAAA polyadenylation signal typical of higher eukaryotic genes (Fig. 2).

The G+C content of the coding sequence of enolase is 43.8 mol%, while the 5' and 3' flanking sequences and the intervening sequence have a much lower G+C content of 47, 15 and 14 mol%, respectively. This base composition is strongly divergent from that of an aerobic fungus as indicated by the data of statistical analysis of *Neurospora* genes, revealing that coding and non-coding sequences have a G+C content of 59 and 46 mol%, respectively (Edelmann & Staben, 1994).

The 1308 bp ORF encoded a polypeptide of 436 amino acids with a predicted molecular mass of 47 kDa. The deduced amino acid sequence of the pPR6.1 cDNA showed 100% identity to the corresponding part of the enolase gene (data not shown). The codon usage of the *N. frontalis* enol gene was highly biased, with only 36 codons used of the possible 61 sense codons (Table 1). The unused codons always end in A or G. If a purine is present in the third position, A is used in preference to G with the exception of lysine (100% AAG).

Comparisons of amino acid sequences

The alignment of the predicted enolase proteins from a representative range of organisms is shown in Fig. 3, and
the levels of amino acid identity are presented in Table 2. Homology has been retained throughout the protein and there are numerous regions of complete identity in the 11 sequences analysed. The 187 invariant amino acids are arranged in short tracts of highly conserved sequences.

The lowest level of identity is seen with the enolase of N. frontalis. The entire nucleotide sequence of the yeast enolase is shown with the corresponding amino acid sequence for comparison.

**Fig. 2.** Nucleotide and deduced amino acid sequences of the enol gene of N. frontalis. The entire nucleotide sequence of the enol gene is shown with the corresponding amino acid sequence printed below it. The nucleotide and amino acid sequences are numbered in the 5' to 3' direction from the respective line of the putative ATG initiation codon is designated +1, the stop codon is indicated by an asterisk. Lower case letters indicate nucleotides in non coding parts of the enol gene. The EMBL/GenBank accession number for the nucleotide sequence is X80474.

**Table 1.** Codon usage of the enol gene from N. frontalis.

<table>
<thead>
<tr>
<th>Codon</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG</td>
<td>0.0047</td>
</tr>
<tr>
<td>FT</td>
<td>0.0034</td>
</tr>
<tr>
<td>FT</td>
<td>0.0027</td>
</tr>
<tr>
<td>FT</td>
<td>0.0020</td>
</tr>
<tr>
<td>FT</td>
<td>0.0015</td>
</tr>
<tr>
<td>FT</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

animals (Drosophila, chicken, human) and plant enolases (tomato, maize, Arabidopsis). In the fungal family, the enolase sequence of N. frontalis is clearly divergent from the yeast enolases.

**Copy number and expression of enol**

Multiple enolase genes have been demonstrated in several eukaryotic species; evidence of additional enolase genes was sought in N. frontalis by Southern analysis. A 24 kb Xbal–HindIII fragment from the clone pRDXX6.5 (Fig. 1), which includes the 5' flanking sequence and overlaps the 5' part of the N. frontalis enol gene, was hybridized against genomic DNA digested with BamHI, XbaI, EcoRI and HindIII, (Fig. 5). As XbaI and BamHI have no cleavage site within the enol gene, each hybridization band should represent a unique enolase gene. HindIII and EcoRI have a single cleavage site in the 5' part of the enolase gene and the flanking sequence overlapped by the probe (Fig. 1). The size of the DNA fragments hybridizing to the probe were 25 and 27 kb, respectively, indicating that the HindIII and EcoRI sites are close to the XbaI site of the clone pSKXXX6.5 in an adjacent genomic clone. The presence of only one band strongly suggested that there is only a single enolase gene in the N. frontalis genome.

The partial cDNA clone for N. frontalis enolase was used to probe the genomic DNA from S. cerevisiae digested with BamHI and XbaI. Neither of these enzymes has a cleavage site within the S. cerevisiae enol gene, two fragments of the S. cerevisiae enolase were recognized by the probe, confirming the high degree of similarity between the N. frontalis enol and the two S. cerevisiae enolases (Fig. 5).

The effects of different carbon sources of the culture medium on expression of the N. frontalis enol gene were investigated by growing the fungus in the presence of glucose, crystalline cellulose (Avicel) or methylglucose (known as high-level inducers of secreted polyclac-
Enolase from Neocallimastix frontalis

Fig 3. Amino acid alignment of N. frontalis enolase with eukaryotic enolases. Asterisks indicate identical residues in the 11 sequences. Dashes indicate gaps created by the TREEALIGN program of Bisance (Dessen et al., 1990).

Table 2. Percentage amino acid identity between different eukaryotic enolases

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N. frontalis</td>
<td>70</td>
<td>71</td>
<td>70</td>
<td>70</td>
<td>68</td>
<td>73</td>
<td>70</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>C. albicans</td>
<td>78</td>
<td>76</td>
<td>63</td>
<td>64</td>
<td>63</td>
<td>62</td>
<td>66</td>
<td>65</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>S. cerevisiae 1</td>
<td>95</td>
<td>61</td>
<td>61</td>
<td>60</td>
<td>61</td>
<td>65</td>
<td>64</td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>S. cerevisiae 2</td>
<td>60</td>
<td>61</td>
<td>59</td>
<td>59</td>
<td>63</td>
<td>63</td>
<td>61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Tomato</td>
<td>87</td>
<td>90</td>
<td>66</td>
<td>70</td>
<td>71</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Maize</td>
<td>86</td>
<td>65</td>
<td>67</td>
<td>69</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Arabidopsis</td>
<td>66</td>
<td>69</td>
<td>70</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Chicken</td>
<td>70</td>
<td>85</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Drosophila</td>
<td>72</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Human β</td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Human γ</td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The NBRF accession numbers for enolases are indicated in brackets. 1. N. frontalis; 2. C. albicans (A45241); 3. S. cerevisiae ENO1 (NOBY); 4. S. cerevisiae ENO2 (NOBY2); 5. tomat (JQ1185); 6. maize (S16257); 7. Arabidopsis thaliana (JQ1187); 8. chicken (A23850); 9. Drosophila melanogaster (S07586); 10. human β form (S06756); 11. human γ form (NOHUG).

Fig 4. Phylogenetic relationships between eukaryotic enolases. The dendrogram was established from the multiple alignment seen in Fig. 3 using the parsimony method (Hein, 1990) from the TREEALIGN program of Bisance (Dessen et al., 1990).

Charidases), and maltose or lactose as carbon sources. The size of the enol transcript predicted from the knowledge of the coding sequence is about 1600 nt. Equal amounts of total cellular RNAs from the different cultures were analysed by Northern blotting using a probe for the enol coding sequence. The N. frontalis enolase mRNA was detected as an abundant single transcript of 1.6 kb (Fig. 6). The expression levels of enol were quite similar when the fungus was grown on glucose, methylglucose, maltose or lactose. However, the amount of enol mRNA from mycelium grown on crystalline cellulose (Avicel) was tenfold lower. In the presence of cellulose, the fungal biomass was quite similar to the biomass obtained on other carbon sources. Furthermore, the levels of rRNA detected with the 16S rRNA probe were similar in the different culture conditions (data not shown).

DISCUSSION

In this paper we report the cloning and characterization of the first enolase gene (enol) from a filamentous fungus, N. frontalis. A cDNA clone was first isolated and identified by the high degree of similarity to the enolase genes from other organisms, but this 1.3 kb clone was found to be
were isolated from several rounds of hybridization of cDNA libraries, but all the analysed clones were the same size. This feature may be common to enolase cDNAs, as clones isolated from other organisms appeared to be incomplete. In \textit{C. albicans}, Franklin et al. (1990) have isolated a partial cDNA clone, and in \textit{P. falciparum}, Read et al. (1994) have isolated a 1·3 kb enolase cDNA and used the 5' rapid amplification of cDNA ends (RACE) technique to isolate a complete cDNA clone. It is possible that in these different organisms the primary structure of the enolase genes leads to the formation of strong secondary structures in the mRNA, producing a pause in the reverse transcription reaction during cDNA synthesis.

Comparison of the predicted amino acid sequence of the \textit{N. frontalis} enolase with enolase sequences from other eukaryotes shows a high degree of evolutionary conservation (Fig. 3). Homology has been retained throughout the protein and 187 amino acids arranged in short tracts are conserved in the eleven enolase sequences analysed. Enolase sequences from \textit{Bacillus} and \textit{P. falciparum}, sharing 53·5 and 63·% homology, respectively, with \textit{N. frontalis} enolase, were excluded from the alignment studies.

\textit{N. frontalis} possesses a single enolase gene as confirmed by Southern blotting (Fig. 5), and the catalytically active enzyme must therefore be a homodimer of the enolase gene product. In higher eukaryotes the expression of enolase isoenzymes is specific for a cell type or a developmental stage of the organism (Segil et al., 1988). In \textit{S. cerevisiae}, the expression of \textit{ENO1} ensures that glycolysis continues when the environmental conditions do not favour expression of \textit{ENO2} (McAlister & Holland, 1982). However, \textit{C. albicans} (Brett Mason et al., 1993) and \textit{N. frontalis} appear to have only a single enolase gene. These two organisms are adapted to special ecological niches, the human body and the rumen, where the environmental conditions are relatively constant and this may explain the pattern of expression of the enolase gene.

Studies of DNA base composition from \textit{N. frontalis} and other anaerobic fungi have demonstrated these organisms to have the most A+T-rich genome of any organisms identified so far (Billon-Grand et al., 1991; Brownlee, 1994). Evidence was presented that the A+T-rich sequences are widely spread in the genome as repeated sequences and spacer regions of DNA flanking the highly repeated RNA coding regions. From our present study on enolase gene organization and from the characterization of several cDNA clones of \textit{N. frontalis}, phosphoenolpyruvate carboxykinase (Reymond et al., 1992), \(\beta\)-succinyl-CoA synthetase (T. H. C. Brondijk, personal communication), xylanase 1 and xylanase 2 (R. D., unpublished results), it could be postulated that the coding sequences of the genes from \textit{N. frontalis} are separated in the genome by long stretches of very A+T-rich flanking sequences (less than 15mol\% G+C). Moreau et al. (1982) reported a general pattern of organization of DNA in eukaryotic genomes whereby A+T-rich ‘linker’ sequences systematically punctuate regions that contain coding sequences and might be

\textbf{Fig 5.} Southern hybridization of \textit{N. frontalis} enolase to restricted genomic DNAs. (a) \textit{N. frontalis} genomic DNA was digested with \textit{Bam}HI (1), \textit{Xba}I (2), \textit{Eco}RI (3), \textit{Hind}III (4) or was undigested (5). The probe was the 2·4 kb \textit{Xba}I–\textit{Hind}III fragment overlapping the coding part and the 5' non-coding part of the genomic clone. (b) \textit{S. cerevisiae} genomic DNA was digested with \textit{Bam}HI (1), \textit{Xba}I (2) or was undigested (3). The probe was the 1·3 kb partial enolase cDNA clone pPR6.1. The sizes of marker DNA fragments are indicated at the sides.

\textbf{Fig 6.} Analysis of \textit{N. frontalis} enolase expression after growth on different carbon sources. (a) Total RNAs (30 \mu g per lane) from \textit{N. frontalis} grown on glucose (1), \textit{Avicel} (2), methylglucopyranoside (3), maltose (4) and lactose (5) were separated on agarose gel. (b) Identification of the \textit{enol} gene transcript was obtained by probing total RNAs with the partial cDNA enolase clone pPR6.1. The numbers under each lane indicate the level of expression obtained by densitometric analysis after re-probing the blot with a rRNA probe. The ratio of the signal for \textit{enol} mRNA to that for rRNA was taken as 100\% for glucose-grown mycelium.

incomplete: the part of the sequence encoding the N-terminal amino acids of the expected protein was missing (Reymond et al., 1992). A large number of positive clones

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Carbon Source} & \textbf{Expression Level} \\
\hline
Glucose & 100 \% \\
Avicel & 90 \% \\
methylglucopyranoside & 90 \% \\
maltose & 84 \% \\
lactose & 72 \% \\
\hline
\end{tabular}
\caption{Comparison of \textit{N. frontalis} enolase expression on different carbon sources.}
\end{table}
involved in DNA functional organization at the chromo-
some level. Brownlee (1994) pointed out that the very
A + T-rich genome of \textit{N. frontalis} might be expected to
affect the amino acid composition of the encoded proteins.
Clearly the amino acid sequence comparisons with
eukaryotic enolases (Fig. 3) demonstrate that this is not
the case, as 269 residues (63% of the total amino acids) are
invariant or similar.

The codon usage observed in \textit{N. frontalis} enol gene was
biased since only 36 codons out of the 61 sense codons
were used. This codon preference (Table 1) is very close
to that of the enolase from \textit{C. albicans} (Sundstrom &
Aliaga, 1992), where 35 codons were used. The only
minor differences were that the \textit{N. frontalis} gene utilized
AAC for Asn, GAT for Asp, CAT for His, TTT for Phe
and TAT for Tyr more frequently (Table 1), whereas in \textit{C.
albicans}, the last T was replaced by a C in the same
codons except for Leu (TTG). However, the \textit{P. falciparum}
enolase showed a lower codon usage bias as 46 codons
were used. The codons always ended in A, reflecting the
36 mol% G+C content of the coding sequence. In the
\textit{p190} gene of \textit{P. falciparum} the coding area had an A + T
content of 74 mol% (Mackay, 1987). In the filamentous
fungi \textit{N. crassa} and \textit{Aspergillus nidulans}, highly expressed
genes tend on the whole to show more marked codon bias
than those expressed at a low level (Gurr et al., 1987). The
high degree of codon bias for \textit{N. frontalis} enol is consistent
with it being a highly expressed gene using the most
abundant isoacceptor tRNAs.

The \textit{enol} gene presents other features characteristic of
filamentous fungal genes. The enolase gene was inter-
rupted by a 331 bp long intervening sequence. The introns
of filamentous fungi are generally short, on average less
than 100 bp (Gurr et al., 1987), although in \textit{Neurospora
crassa} some are as large as 691 bp (Edelmann & Staben,
1994). The \textit{enol} gene intron boundaries followed the
consensus sequences GTANGT and PyAG determined for
filamentous fungal genes (Gurr et al., 1987). All these
results show that the remarkably low overall G + C
content in an anaerobic fungus does not affect the gene
organization which is similar to that in aerobic fungi.

\textit{N. frontalis} enolase mRNA was detected as an abundant
transcript in mycelia grown on glucose or on the
disaccharides maltose or lactose (Fig. 6). However, when
the fungus was grown on glucose the level of the enolase
transcript was tenfold higher than that detected in
cellulose-grown mycelia. These results are reminiscent of
the 20-fold glucose-mediated induction of the \textit{S. cerevisiae}
\textit{ENO2} gene (McAlister & Holland, 1992). In \textit{C. albicans},
the level of enolase transcripts was sixfold higher in
glucose-grown cells than in cells grown on a glucone-
ogenic carbon source (Brett Mason et al., 1993)

Concluding remarks

The cloning and characterization of the \textit{enol} gene from a
genomic clone containing 5'flanking sequences greater
than 1.5 kb will allow the characterization of the enolase
gene promoter. Promoters of highly expressed genes such
as 3-phosphoglycerate kinase (Takaya et al., 1994) have
been useful for the construction of transformation vectors.
Preliminary experiments have shown that the transformation of oomycetes or \textit{N. frontalis} are unsuc-
sessful using regulatory sequences from higher fungi. The
presence of homologous promoting sequences in the
transformation vectors should allow the efficient trans-
formation of anaerobic fungi.

REFERENCES

Billon-Grand, G., Fiol, J. B., Breton, A., Bruyère, A. & Oulhaj, Z.
(1991). DNA of some anaerobic fungi: G+C-content determina-

cloning and characterization of the \textit{Candida albicans} enolase gene. \textit{J
Bacteriol} 175, 2632–2639.


Carmen, A. A. & Holland, M. J. (1994). The upstream repression
sequence from the yeast enolase gene \textit{ENO1} is a complex regulatory
element that binds multiple \textit{trans}-acting factors including \textit{REB1}. \textit{J
Biol Chem} 269, 9790–9797.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. S.
(1979). Isolation of biologically active ribonucleic acid from sources
enriched in ribonuclease. \textit{Biochemistry} 18, 5294–5297.

Identification of a regulatory region that mediates glucose-
dependent induction of the \textit{Saccharomyces cerevisiae} enolase gene
\textit{ENO2}. \textit{Mol Cell Bio} 6, 2287–2297.

Cohen, R., Yokoi, T., Holland, J. P., Pepper, A. E. & Holland, M. J.
(1987). Transcription of the constitutively expressed yeast enolase
gene \textit{ENO1} is mediated by positive and negative \textit{cis}-acting regulatory

A French service for access to biomolecular sequence databases.

sequence features within genes from \textit{Neurospora crassa}. \textit{Exp Mol
Yol} 18, 70–81.

labelling DNA restriction endonuclease fragments to high specific

(1990). An immunodominant antigen of \textit{Candida albicans} shows

Gurr, S. J., Unkles, S. E. & Kinghorn, J. R. (1987). The structure and
organization of nuclear genes of filamentous fungi. In \textit{Gene Structure
Oxford: IRL Press.

polysaccharide hydrolases secreted by the rumen anaerobic fungi
\textit{Neocallimastix frontalis}, \textit{Sphaeromonas communis} and \textit{Piromonas
communis}. \textit{J Gen Microbiol} 134, 1123–1129.

\textit{Methods Enzymol} 183, 626–645.

The primary structures of two yeast enolase genes. \textit{J Biol Chem} 256,
1385–1395.

Oxford: IRL Press.


Received 19 December 1994; revised 27 February 1995; accepted 10 March 1995.