ELSEVIER

Gene 216 (1998)93-102

GENE
AN INTERNATIONAL JOURNAL ON GENE AND GENOMES

A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*

Shigang Shen [a], Greitje Sulter [b], Thomas W. Jeffries [c], James M. Cregg [a,*]

[a] *Department of Biochemistry and Molecular Biology. Oregon Graduate Institute of Science and Technology, P. O. Box 91000. Portland, OR 97291-1000. USA*

[b] *Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Biological Centre, University of Groningen, Kerklaan 30, 9751 NN Haren. The Netherlands*

[c] *Institute Microbial and Biochemical Technology, Forest Products Laboratory, US Department of Agriculture, One Gifford Pinchot Drive, Madison, WI 53705-2398, USA*

Received 24 March 1998: received in revised form 31 May 1998: accepted 1 June 1998: Received by V. Larionov

Abtract

In methylotrophic yeasts, glutathione-dependent formaldehyde dehydrogenase (FLD) is a key enzyne required for the metabolism of methanol as a carbon source and certain alkylated amines such as methylamine as nitrogen sources. We describe the isolation and characterization of the *FLD1* gene from the yeast *Pichia pastoris.* The gene contains a single short intron with typical yeast-splicing signals near its 5′ end. the first intron to be demonstrated in this yeast. The predicted FLD1 product (Fld1p) is a protein of 379 amino acids (approx. 40 kDa) with 71% identity to the FLD protein sequence from the *n*-alkane-assimilating yeast *Candida maltosa* and 61-65% identity with dehydrogenase class III enzymes from humans and other higher eukaryotes. Using *b*-lactamase as a reporter, we show that the *FLD1* promoter (P_{FLD1}) is strongly and independently induced by either methanol as sole carbon source (with ammonium sulfate as nitrogen source) or methylamine as sole nitrogen source (with glucose as carbon source). Furthermore, with either methanol or methylamine induction, levels of *b*-lactamase produced under control of P_{FLDl} are comparable to those obtained with the commonly used alcohol oxidase I gene promoter (P_{AOX1}). Thus. P_{FLD1} is an attractive alternative to P_{AOX1} for expression of foreign genes in *P. pastoris,* allowing the investigator a choice of carbon (methanol) or nitrogen source (methylamine) regulation with the same expression strain.

Keywords: Heterologous gene expression; Expression vector; *b*-Lactamase reporter; Methylotrophic yeast; Methane; metabolism, Methylamine

1. Introduction

Pichia pastoris is a methylotrophic yeast that is widely used for the production of heterologous proteins of industrial and academic interest (Cregg, 1998; Higgins and Gregg. 1998). In this system, most foreign genes are expressed under the transcriptional control of the *P. pastoris* alcohol oxidase 1 gene promoter (P_{AOX1}). the regulatory characteristics of which are well suited for this purpose. The promoter is tightly repressed during growth of the yeast on most common carbon sources. such as glucose, glycerol, or ethanol, but is highly induced during growth on methanol (Tschopp et al., 1987). For production of foreign proteins, P_{AOX1}-controlled expression strains are grown initially on a repressing carbon source to generate biomass and then shifted to methanol as the sole carbon and energy source to induce expression of the foreign gene. One advantage of the P_{AOX1} regulatory system is that foreign genes whose products are toxic to the cells can be maintained by growing the yeast under repressing conditions. Although many proteins have been produced successfully using P_{AOX1}, this promoter may not be

* Corresponding author, Tel: -1 503 690-1217; Fax: -1 503 690-1464; e-mail; cregg@bmb.ogi.edu

Abbreviations; *amp, E. coli* gene encoding *b*-lactamase; AOX. alcohol oxidase; *AOX1, P. pastoris* gene encoding alcohol oxidase I: *bla.* modified *amp* gene; CAT, catalase; DAK, dihydroxyacetone kinase; DAS, dihydroxyacetone synthase: FDH, formate dehydrogenase : FLD, formaldehyde dehydrogenase: *FLD1,* formaldehyde dehydrogenase gene: Fld1p, product of *FLD1* gene; GAP. glyceraldehyde-3-phosphate: *H1S4, P. pastoris* gene encoding histidinol dehydrogenase: *b*-lac. β-lactamase; ORF, open reading frame; p. plasmid; *P.* promoter: RT. reverse transcriptase.

0378-1119 98 $19.00
PII:S0378-1119(98)00315-l

appropriate or convenient in some situations. For example, in shake-flask cultures, methanol rapidly evaporates, and it is inconvenient to monitor methanol concentrations and repeatedly add the compound to the medium. As a second example. the storage of large amounts of methanol needed for the growth and induction of P_{AOX1}-controlled expression strains in large-volume high-density fermentor cultures can be a potential fire hazard in certain situations. As an alternative to P_{AOX1}, it would be useful if a promoter was available that was transcriptionally efficient and could be regulated like P_{AOX1}, but was responsive to a less-volatile and flammable inducer. Here, we describe the isolation and characterization of the *P. pastoris* FLD gene (*FLD1*) whose promoter has these characteristics.

FLD is an important enzyme in the utilization of methanol as a carbon and energy source (Veenhuis et al., 1983). However, its role is the subject of some controversy. In methylotrophic yeasts. the methanol metabolic pathway is thought to be nearly the same. beginning with the oxidation of methanol to formaldehyde by AOX, a hydrogen peroxide-producing oxidase that is sequestered in an organelle called the peroxisome. Hydrogen peroxide is then degraded to oxygen and water by catalase, the classic peroxisomal marker enzyme. A portion of the resulting formaldehyde condenses with xylulose-5′-monophosphate in a reaction catalyzed by dihydroxyacetone synthase, the third peroxisomal methanol pathway enzyme. The products of this reaction, GAP and dihydroxyacetone, then leave the peroxisome and enter a cyclic pathway that regenerates xylulose-5′-monophosphate and also generates one net molecule of GAP for every three turns of the cycle. GAP is used for biosynthesis of carbon skeletons for cell growth. Another portion of the formaldehyde leaves the peroxisome and is oxidized to formate by FLD and then to carbon dioxide by FDH. Both of these reactions produce reducing power in the form of NADH. One model of FLD function is that the NADH generated by FLD and FDH serves as the primary source of energy during growth on methanol (Veenhuis et al., 1983). The second model proposes that most energy for methanol growth comes from the oxidation of one or more of the xylulose-5′-monophosphate cycle intermediates by tricarboxcylic acid cycle enzymes, and that the primary role of FLD is to protect the cell from toxic formaldehyde that accumulates with excess methanol in the medium (Sibirny et al., 1990).

In addition to methanol, FLD is also involved in the metabolism of certain alkylated amines (e.g. methylamine and choline) as sole nitrogen sources (Zwart et al., 1980). In this pathway, amine groups are first liberated by a peroxisomal amine oxidase, leaving formaldehyde which is further oxidized by FLD and FDH. When growing on methylamine as sole nitrogen source. high levels of FLD are induced even in the presence of excess glucose. Thus. the primary role of FLD in methylamine metabolism appears to be for protecting cells from the toxic effects of formaldehyde and not for generating carbon or enerey.

FLD synthesis is regulated independently in response to either methanol as sole carbon and energy source or to methylamine as sole nitrogen source. Thus. for example. only low levels of FLD are observed in cells growing on glucose- and ammonium ion-containing medium, whereas on either methanol-ammonium ion or glucose-methylamine media, FLD levels are high. Here we report the isolation and characterization of the *P. pastoris FLD1* gene and its promoter (P_{FLD1}). We show that *FLD1* expression in response to methanol or methylamine is controlled at the transcriptional level. Surprisingly, we also find that P_{FLD1} is similar to P_{AOX1} in strength in response to either methanol or methylamine induction.

2. Materials and methods

2.1. Strains and media

The wild-type *P. pastoris* strain was NRRL Y-11430, *P. pastoris fld1* mutant strains were generated using nitrosoguanidine and were a gift from Dr George Sperl of Phillips Petroleum Company (Bartlesville, OK). MS105, a *P. pastoris fld1 his4* strain, was constructed by crossing GS241 (*fld1-1*) with GS115 (*his4*). Complementation analysis and other classical genetic techniques were performed as described in Cregg and Russell (1998). The *Hansenula polymorpha* strain used was CBS4732. Bacterial recombinant DNA manipulations were performed in either *Escherichia coli* strain MC1061 or DH5*x*. Yeast strains were cultured in a rich YPD medium (1% yeast extract, 2% peptone, 0.4% glucose) or a minimal medium composed of 0.17% yeast nitrogen base without ammonium sulfate and amino acids. a carbon source (0.4% glucose or 0.5% methanol), and a nitrogen source (0.5% ammonium sulfate or 0.25% methylamine chloride). *E. coli* strains were cultured in Luria broth medium supplemented with either 100 μg ml ampicillin or 50 μg ml zeocin (Invitrogen Corporation, Carlsbad, CA) as required.

2.2. Plasmid and genomic library constructions

The *P. pastoris* genomic library in *P. pastoris* vector pYM8 was described in Liu et al. (1995). The *H. polymorpha* library in *P. pastoris* vector pYM8 was constructed in the same manner. Briefly, *H. polymorpha* genomic DNA was partially digested with *Sau*3AI and size-selected for fragments of 5-20 kb. These fragments were ligated into the *Bam*HI site of pYM8. The library was composed of approx. 100 000 independent *E. coli*

transformants with >90% containing an insert. The average size of insert DNA was approx. 10 kb. Assuming that the size of the *H. polymorpha* genome was 10 000 kb, the library contained approx. 100 genome equivalents of *H. polymorpha* genomic DNA. For expression of bacterial *b*-lac under the transcriptional control of P_{FLD1}, a 0.6 kb *Mun*I-*Bam*HI fragment composed of sequences beginning immediately 5′ of the methionine initiator ATG of *FLD* was generated by PCR using the *FDL1* plasmid pYG1 as template and primers composed of the following sequences: 5′-CGGGATCCG-CATGCAGGAATCTCTGGCA-3′ and 5′-CGCAA-TTGTGTGAATATCAAGAATTG-3′. The resulting fragment was cut with *Mun*I and *Bam*HI and ligated into *Eco*RI- and *Bg*/II-digested pHW018 and pK32l to create pSS050 and pSS040, respectively (Fig. 1), pSS050 and pSS040 contain the identical expression cassettes composed of P_{FLD1}, the *b*-lac gene, and the *AOX1* transcriptional terminator. They differ in that pSS050 contains a kanamycin-resistance gene for selection in *E. coli* and the *P. pastoris HIS4* gene for selection in *his4* strains of *P. pastoris.* whereas pSS040 contains the zeocin-resistance gene which serves as the selectable marker for both organisms.

2.3. Cloning of the P. pastoris *and* H. polymorpha FLDI *genes*

To isolate DNA framents containing the *P. pastoris* and *H. polymorpha FLDI* genes. 5-10 µg of each plasmid library were transformed into *P. pastoris* strain MS105 (*fld1-1 his4*) via the spheroplast method (Cregg et al., 1985), and approx. 50 000 His^- transformants were selected on YND medium agar. Transformants from each library were pooled and approx. 1×10^8 ceils from each were spread on YNM plates to select for growth on methanol (Mut^- phenotype). The resulting His^-Mut^- colonies (several hundred from the *P. pastoris* library and 20 from the *H. polymorpha* library transformations) were pooled. and total yeast DNAs were extracted. The yeast DKAs were then transformed into *E. coli* and plasmids contained in the resulting colonies examined. From the *P. pastoris* library transformation. one plasmid (pYG1) was recovered that contained a DNA insert and retransformed strain MS105 to both His^- and Mut^-. From the *H. polymorpha* library transformation. four piasmids were recovered that were able to confer both His^- and Mut^- phenotypes upon retransformation into MS105. One of these, pYG2, was used in these studies.

2.4. Biochemical methods

For enzyme assays. yeast strains were grown in shake flasks at 30°C in YNB (without amino acids and ammonium sulfate) medium using either 0.4% glucose or 0.5% methanol as carbon source and either 0.5% ammonium sulfate or 0.25% methylamine as nitrogen source. Cultures were harvested in the late logarithmic phase, and cell-free extracts were prepared using glass beads as described in Waterham et al. (1992). The protein concentrations in cell-free extracts were determined using either the method of Bradford (1976) or the Pierce BCA protein assay kit (Rockford, IL) with bovine serum albumin as standard. Alcohol oxidase (van der Klei et al., 1990), catalase (Lück, 1963), dihydroxyacetone synthase (Waites and Quayle, 1981), dihydroxyacetone kinase (van Dijken et al., 1978) and FDH (van Dijken, 1976) activities were determined by published methods. FLD activity was measured spectrophotometrically by following the rate of NADH formation at 340 nm in the presence of saturating amounts of formaldehyde, glutathione and NAD as described by Schutte et al. (1976). Reaction mixtures contained 33 mM sodium phosphate buffer (pH 7.9-8.0), 2 mM glutathione, 1 mM NAD, 1 mM formaldehyde and limiting amounts of enzyme in a final volume of 1.0 ml. The rate of absorbance change at 340 nm was followed for at least 2 min, and activities were calculated by using the constant $e = 6.22$ cm^2 nmol for NAD. Alcohol oxidase activities were expressed in µmol mg/min, and FLD activities were expressed in µmol/mg/min. *b*-Lactamase activity. expressed as nmol/mg/min, was assayed spectrophotometrically at 569 nm and 30°C in 25 mM Tris-HC1 (pH 7.5) using 11.1 mM PADAC as substrate (extinction coefficient 44.403 $cm^{-1}M^{-1}$).

2.5 Miscellaneous methods

Recombinant DNA methods were performed essentially as described in Sambrook et al. (1989). Oligonucleotides were synthesized and DNA sequencing was performed at the Oregon Regional Primate Research Center, Molecular Biology Core Facility (Beaverton, OR). PCR reactions were performed as described by Kramer and Coen (1995). Total *P. pastoris* RNA was isolated according to Schmitt et al. (1990). The RT-PCR reaction was performed as described previously (Frohman et al., 1988; Stewart et al., 1992) using the following oligonucleotide primers: 5′-CACAAT-GTCTACCGAAGGTC-3′ (5′ primer) and 5′-CCAGAAAGCGTGTAAGCATCAG-3′(3′ primer).

3. Results

3.1. Isolation of formaldehyde dehydrogenase-defective mutants of P. pastoris

As a first step in cloning the ***P. pastoris*** FLD gene *(FLD1),* mutants were sought that were specifically defective in FLD activity. Previous biochemical studies

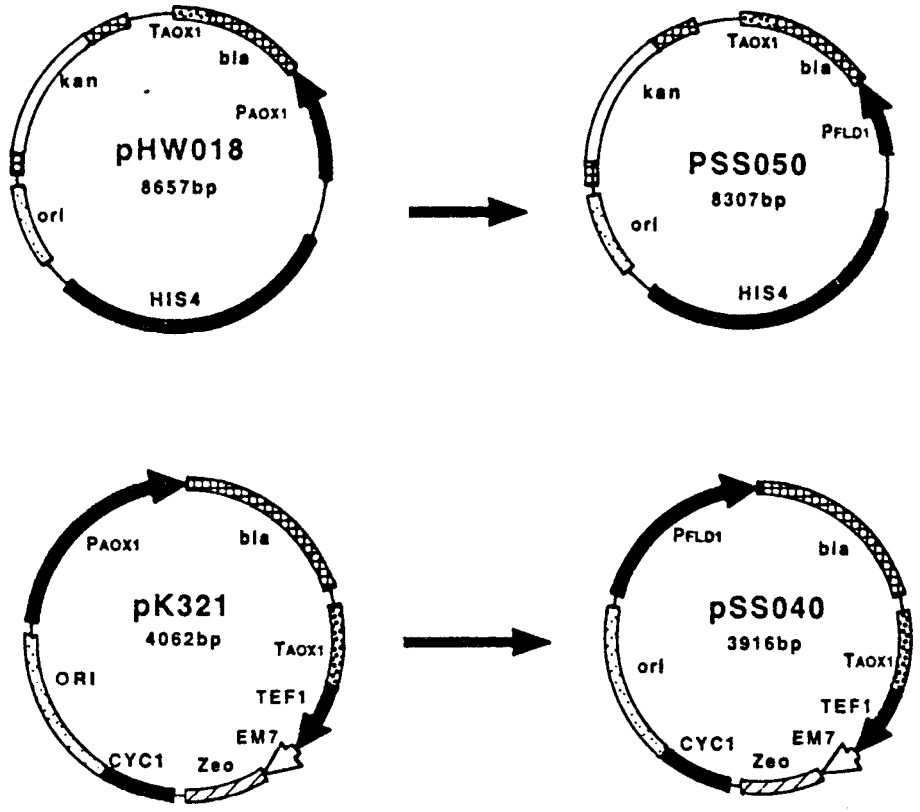

Fig. 1. Physical maps of selected vectors used in this study.

of methylotrophic yeasts indicated that FLD was involved in the metabolism of both methanol as carbon source and methylamine as nitrogen source (Zwart et al., 1980). This also appeared to be true for *P. pastoris*, since growth of the yeast on methanol as sole carbon source and or methylamine as sole nitrogen source specifically induced high levels of FLD activity (Table 1). To search for *P. pastoris fld1* mutants. nitrosoguanidine-mutagenized cultures were screened for strains that were unable to utilize methanol as carbon source and methylamine as nitrogen source. Five mutants belonging to a single complementation group were identified.

These five strains were examined further by measuring the levels of activity of key methanol pathway enzymes in extracts prepared from methanol-induced cultures of each strain, These enzymes included: AOX. catalase, dihydroxyacetone synthase, dihydroxyacetone kinase, FLD and FDH. Results were essentially the same for each of the five mutants and are shown in Table 1 for one of the mutant strains. GS241. Each mutant contained significant levels of activity for all enzymes assayed except FLD, which was undetectable. As controls, methanol-grown wild-type *P. pastoris* had normal levels of FLD activity. and methanol-induced cells of a *P. pastoris* strain that is deleted for its *AOX* genes and as a result cannot grow on methanol also contained substantial levels of FLD activity.

The phenotypic and biochemical characteristics of the mutants were consistent with the hypothesis that they were specifically defective in the *P. pastoris FLD1* gene.

Table 1
Relative enzyme activity levels in methanol-utlization-defective mutants of *P. pastoris*

	% Activity[4]					
Strain	AOX	CAT	FLD	FDH	DAS	DAK
WT (methanol)	100	100	100	100	100	100
WT (glucose)	0	2	1	0	3	53
KM7121 (*dox1 dox2*)	0	100	26	31	n.d.[b]	88
GS241 (*fld1*)	20	118	0	46	58	64

[4] Activity for each enzyme is expressed as a percentage of that observed in extracts prepared from methanol-grown cultures of wild-type *P. pastoris.*[b] Not determined.

further evidence in support of this will be provided later in this report. One putative *fld1* strain. GS241 (*fld1-1*), was selected for all further studies.

3.2. *Isolation and sequence of the* P. pastoris FLD1 *gene*

To clone the putative *FLD1* gene by functional complementation, strain GS241 was first crossed, to *P. pastoris* strain GS115 (*his4*) to obtain a derivative that was both methanol-utilization defective (Mut^-) and auxotrophic for histidine (His^-). One Mut^- His^- strain that resulted from this cross, MS105 (*fld1-1 his4*). was then transformed with a *P. pastoris* genomic DNA library constructed in the *P. pastoris–E. coli* shuttle vector pYM8 (Cregg et al., 1985; Liu et al., 1995). This plasmid is composed of the *Saccharomyces cerevisiae* histidinol dehydrogenase gene (*SHIS4*) and a *P. pastoris*-specific autonomous replication sequence (*PARS1*) inserted into *E. coli* plasmid pBR322. Library transformants were selected for His^+ prototrophy and then further selected for ones that were also Mut^+. Total DNA was extracted from a pool of several hundred His^- Mut^- colonies and transformed into *E. coli.* One plasmid recovered from this process, pYG1, was able to retransform strain MS105 to both His^+ and Mut^+ and was examined further.

To determine the location of the putative *FLD1* gene on pYG1, the plasmid was restriction mapped. and selected fragments from the vector were subcloned and tested for the ability to complement strain MS105. The plasmid was found to be 14.5 kb in size and to contain an insert of 6.8 kb (Fig. 2). A 2.7 kb *Sph*I–*Bam*HI fragment was found to be sufficient to complement the Mut^- defect in MS105 and was sequenced. The DNA sequence identified a long ORF whose predicted product had strong similarity to other alcohol dehydrogenases (see below). The sequence also suggested the possible presence of an intron near the 5′ terminus of the gene.

To confirm the presence of an intron, this region of the ORF was amplified from mRNA by the RT-PCR method and the size and sequence of the product was compared to that obtained by PCR of the genomic fragment on plasmid pYG1(Fig. 3). Whereas the genomic product was 284 bp in length. the cDNA product was significantly shorter at 170 bp. Alignment of the cDNA and genomic sequences demonstrated that a segment of 114 bp that was present in the genomic DNA was absent from the cDNA. Furthermore, examination of the putative intron/exon junctions revealed typical yeast splice junctions (5′ junction. 5′-GTAAGT-3′ : 3′ junction. 5′-YAG-3′) and branch point (5′-TACTAAC-3′) (Domdey et al., 1984: Sasnauskas et al., 1992). We concluded that a single intron was present at this position in the ORF. Finally, Southern blots of selected restriction digests of wild-type genomic DNA. using a

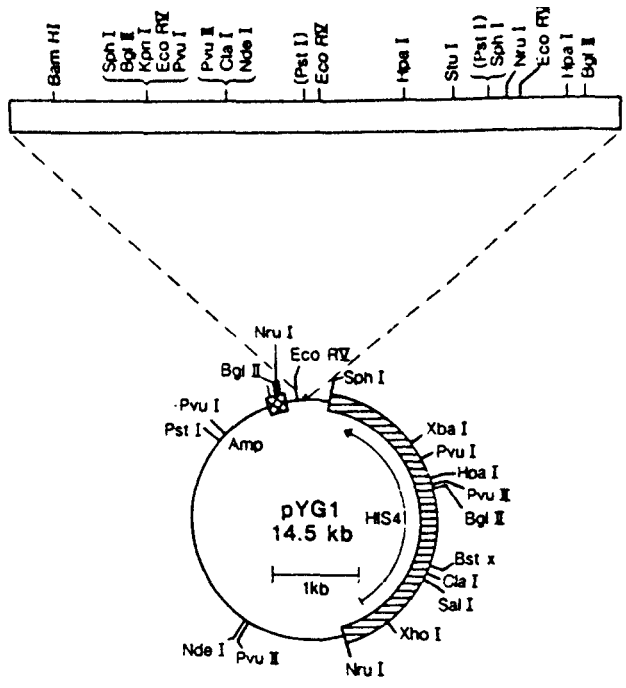

Fig. 2. Restriction enzyme map of the *FLD1*-containing vector pYG1

fragment from the ORF as hybridization probe. indicated that the *P. pastoris* genome contained only one copy of the gene (data not shown).

The DNA and predicted amino-acid sequences of the ORF are shown in Fig. 4. The ORF is 1137 bp long and is predicted to encode a protein of 379 amino acids with a calculated molecular mass of 39 870. The intron begins at a position 18 bp (six amino acids) 3′ of the A of the predicted methionine initiator ATG and is 114 bp in length. Northern blots of total RNA extracted from glucose- and methanol-grown wild-type *P. pastoris* cells. using a DNA fragment from the ORF region, showed a single mRNA species of approx. 1.3 kb that was present at high levels in methanol- but not glucose-grown cells (data not shown). Overall, the codon usage of the putative *FLD1* gene was typical of other highly expressed *P. pastoris* genes (Sreektishna. 1993).

The GenBank, NCBI database was searched for other proteins with amino-acid sequence similarity to the ORF product. The sequence of the putative FLD1 protein (Fld1p) showed the highest identity (71%) with that of glutathione-dependent FLD from the yeast *Candida maltosa* (Sasnauskas et al., 1992) (Fig. 5). *C. maltosa* is an *n*-alkane assimilating yeast and FLD is believed to be important in protecting the yeast from the toxic effects of formaldehyde (Sasnauskas et al., 1992). The close similarity of the predicted *C. maltosa* FLD product to that of the cloned ORF strongly supported our hypothesis that this ORF encodes *P. pastoris* Fld1p. The *P. pastoris* Fld1p sequence also showed high identity

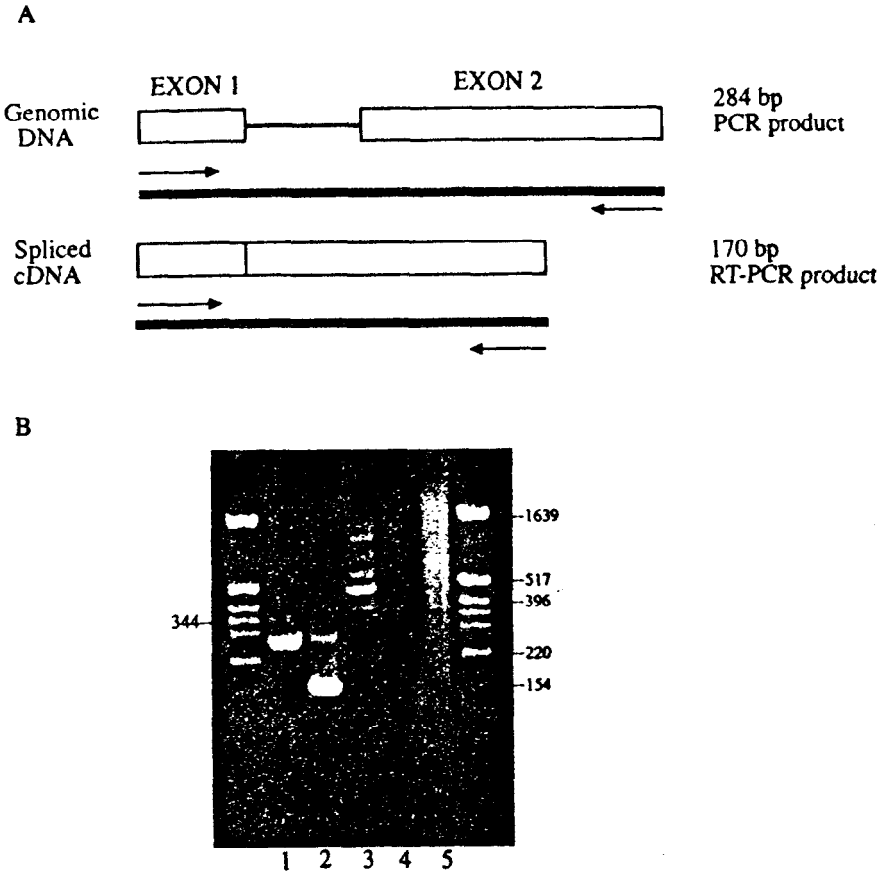

Fig. 3. Exon analysis of the *FLD1* gene. (A) Diagram of the expected products from PCR of unspliced (genomic) and spliced (cDNA) DNAs. Locations of the hybridized primers used in the PCR reactions are shown as convergent arrows. (B) Electrophoretogram of PCR and RT-PCR reaction products. PCR reactions were performed with the following: lane 1, genomic DNA template plus both primers: lane 2, cDNA template plus both primers: lane 3, cDNA template plus 5′ primer only; lane 4, cDNA template plus 3′ primer only; lane 5, both primers without DNA template. Flanking marker bands are denoted in base pairs.

with alcohol dehydrogenase III (ADHIII) proteins of higher eukaryotes (65%, human; 63%, horse; 64%, rat) and a lower but significant identity with other higher eukaryotic ADHs (Holmquist and Vallee, 1991; Koivusalo et al., 1989; Giri et al., 1989). Finally, the Fld1p sequence showed little similarity with the predicted amino-acid sequences of *S. cerevisiae* ADHs. The closest, at 19% identity, was *S. cerevisiae* ADHI (Jornvall et al., 1987).

3.3. Comparison of the thermal stability of Fld1p from P. pastoris *and* H. polymorpha

Further evidence that the cloned *P. pastoris* gene actually encoded an FLD was obtained by comparing the thermal stability of its product to FLD from *H. polymorpha. H. polymorpha* is a related methylotrophic yeast that has a significantly higher optimal growth temperature than *P. pastoris* (42°C vs 30°C). We reasoned that. because of its higher growth temperature, FLD from *H. polymorpha* would display a significantly higher thermal stability than *P. pastoris* FLD. If so, a comparison of the thermal stability properties of the putative FLDs from the two yeasts would provide strong support for the identity of the gene product. Specifically, the putative *P. pastoris* and *H. polymorpha FLD1* genes would be expressed in methanol-grown cells of the *P. pastoris fld1-1 his4* strain MS105, and the thermal stability of FLD activity in each would be assessed by incubating extracts prepared from the strains at 60°C for selected periods of time and determining the rate of loss of FLD activity. If the genes actually encode Fld1p, the FLD inactivation rate for *H. polymorpha* Fld1p expressed in *P. pastoris* would be similar to that of wild-type *H. polymorpha* Fld1p, and the inactivation rate for the *P. pastoris* gene product would be similar to that of wild-type *P. pastoris* Fld1p.

To perform this comparison, it was first necessary to establish that the thermal stability of the *P. pastoris* and *H. polymorpha* FLDs were significantly different and to clone the putative *H. polymorpha FLD1* gene. Thermal stabilities were determined by preparing cell-free extracts from methanol-grown cultures of wild-type *P. pastoris* and *H. polymorpha* and incubating them at 60°C. At

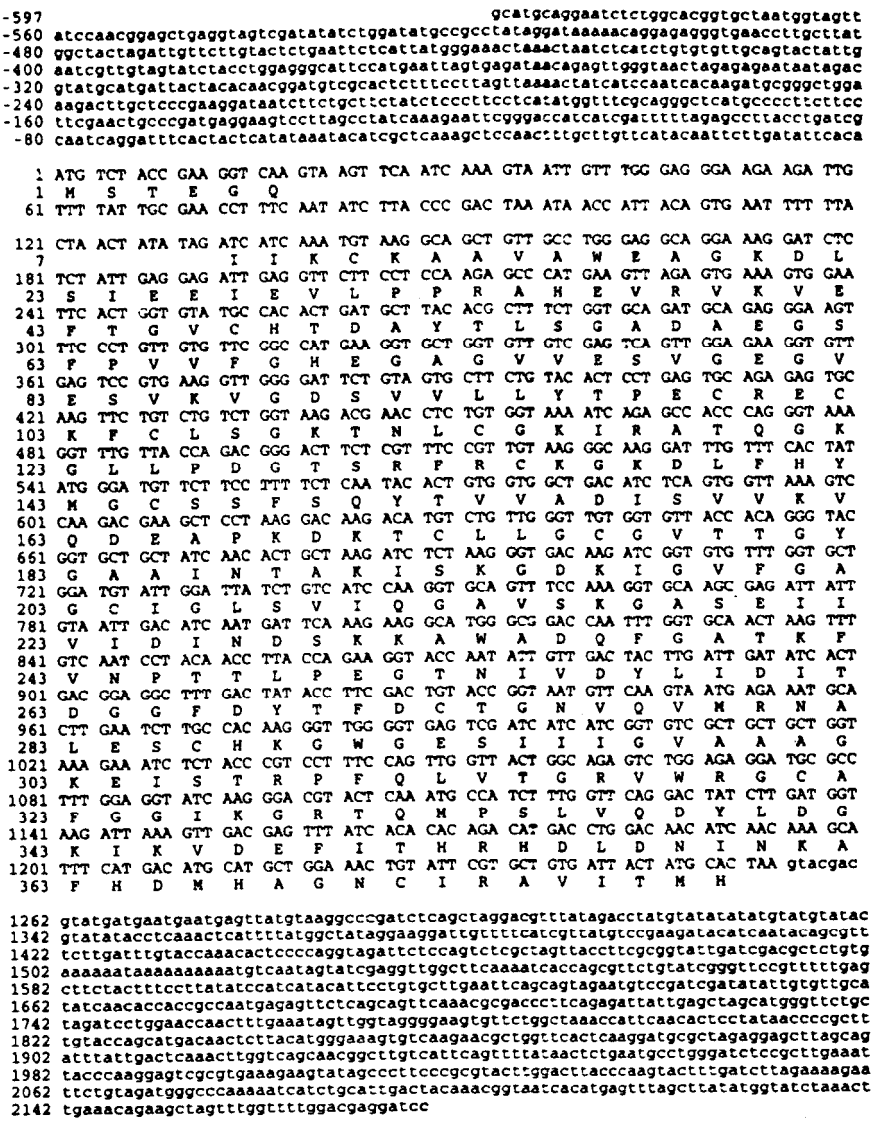

Fig. 4. Nucleotide and deduced amino-acid sequences of *P. pastoris FLD1* gene and its product. The sequence data are available from EMBL/GenBank/DDBJ under accession number AF066054.

selected times during incubation, samples of extract were removed and assayed for FLD activity. As shown in Fig. 6, *H. polymorpha* FLD activity was significantly more heat stable than *P. pastoris* activity.

The putative *H. polymorpha FLD1* gene was isolated using the same functional complementation strategy described above for the *P. pastoris* gene. Briefly, a *H. polymorpha* genomic DNA library was transformed into *P. pastoris* strain MS105 (*fld1-1 his4*) and His^+ Mut^+ colonies were selected. Plasmids were recovered and analyzed for those capable of simultaneously retransforming MS105 to both His^+ and Mut^+ phenotypes. One plasmid that met these criteria, pYG2, was selected for use in these studies. This plasmid contained a *H. polymorpha* DNA insert of 7.2 kb and the Mut complementing activity was found to reside within a 2.4 kb *Sph*1 fragment. Southern blot studies demonstrated that DNA fragments containing the 2.4 kb *H. polymorpha* fragment hybridized to DNA fragments containing the putative *P. pastoris FLD1* gene (data not shown). indicating that the fragments contained homologous genes.

Thermal stability of FLD expressed from *H. polymorpha* vector pYG2 was then compared to that of FLD from the *P. pastoris* vector pYG1. As shown in Fig. 6, FLD in MS105(pYG2) had a thermal inactivation rate similar to that of wild-type *H. polymorpha*, while MS105(pYG1) had a rate similar to that of *P. pastoris*. From these results, along with those demonstrating the specific absence of FLD activity in *P. pastoris* strain GS241 (and MS105) and the close similarity of the primary amino acid sequences of the cloned *P. pastoris* ORF and *C. maltosa* FLD, we concluded that the cloned ORF encoded *P. pastoris* Fld1p.

```
P.pFLD  - MS--TEGQIIKCKAAVAWEAGKDLSIEEIEVLPPRAHEVRVKVEFTGVCH -48
C.mFLD  - MSESTVGKPITCKAAVAWEAAKPLSIEDVTVAPPKRHEVRIKLYDTGVCH -50
P.pFLD  - TDAYTLSGADAEGSFPVVFGHEGAGVVESVGEGVESVKVGDSVVLLYTPE -98
C.mFLD  - TDAYTLSGVDPEGAFPVILGHEGAGIVESIGEGVTNVKVGDHVIALYTPE -100
P.pFLD  - CRECKFCLSGKTNLCGKIRATQGKGLLPDGTSRFRCKGKDLFHYMGCSSF -148
C.mFLD  - CGECKFCKSGKTNLCGKIRATQGKGVMPDGTSRFTCKGKEILHFMGCSTF -150
P.pFLD  - SQYTVVADISVVKVQDEAPKDKTCLLGCGVTTGYGAAINTAKISKGDKIG -198
C.mFLD  - SQYTVVADISVVAINPKAEFDKACLLGCGITTGYGAATITANVQKGDNVA -200
P.pFLD  - VFGAGCIGLSVIQGAVSKGASEIIVIDINDSKKAWADQFGATKFVNPTTL -248
C.mFLD  - VFGGGIVGLSVIQGCAERGAAQIILVDISDKKEEWGQKLGATAFVNPTKL -250
P.pFLD  - PEGTNIVDYLIDITDGGFDYTFDCTGNVQVMRNALESCHKGWGESIIIGV -298
C.mFLD  - PEGTTIVDKLIEMTDGGCDFTFDCTGNVGVMRNALEACHKGWGTSVIIGV -300
P.pFLD  - AAAGKEISTRPFQLVTGRVWRGCAFGGIKGRTQMPSLVQDYLDGKIKVDE -348
C.mFLD  - AAAGKEISTRPFQLVTGRTWKGAAFGGVKGRSQLPGIVNNYLDGKLKVEE -350
P.pFLD  - FITHRHDLDNINKAFHDMHAGNCIRAVITMH -379
C.mFLD  - FITHREPLAAINKAFEEMHAGDCIRAVVDLS -381
```

Fig. 5. Comparison of the predicted ammo-acid sequences of *P. pastoris* and *C. maltosa* FLD proteins. Sequences were aligned using PC gene software. The character '*' between sequences indicates residues that are identical. The character '.' indicates similar residues. Similar residues are defined as: A,S,T; D,E; N,Q; R,K; I,L,M,V; F,Y,W.

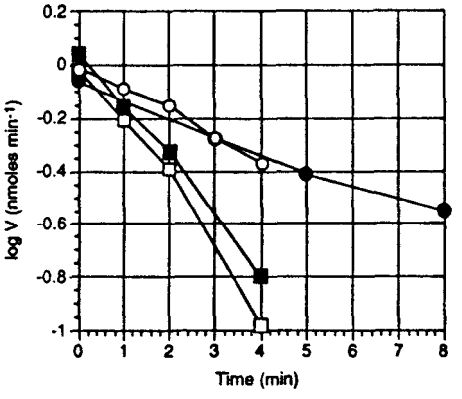

Fig. 6. Thermal stability of formaldehyde dehydrogenase activities in *P. pastoris* strains transformed with putative *FLD1* genes from *P. pastoris* and *H. polymorpha.* Strains shown are: wild-type *P. pastoris* (■); wild-type *H. polymorpha* (●); *P. pastoris* MS105 (pYG1) (•); and *P. pastoris* MS105 (pYG2) (O).

3.4. *Analysis of* P_{FLDl} *and comparison to* P_{AOX1}

As described in the Introduction, a major interest in the *P. pastoris FLD1* gene is in the use of its promoter (P_{FLD1}) as an alternative to the methanol-regulated alcohol oxidase I gene promoter (P_{AOX1}) for controlling expression of foreign genes in *P. pastoris.* To examine gene expression under the transcriptional control of P_{FLD1}, two vectors were constructed (Fig. 1). Both vectors contained identical expression cassettes composed of a 0.6 kb *Mun*I-*Bam*HI fragment with sequences originating from just 5′ of the methionine initiator ATG codon of ***FLD1*** fused to the bacterial ***bla*** gene encoding *b*-lac. followed by a fragment containing the *AOX1* transcriptional terminator. One vector, pSS040, contained a unique *Nsi*I restriction site within the P_{FLD1} fragment. When cut at this site and transformed into *P. pastoris,* the vector efficiently integrated at the P_{FLD1} locus. The result of this integration event was a P_{FLD1}-*bla* expression cassette that also included native *FLD1* sequences upstream of the P_{FLD1} fragment (*WT-*P_{FLD1}*-bla*). Assuming that all sequences required for transcriptional control of *FLD1* are located 5′ of the *FLD1* ORF, regulation of *bla* and *FLD1* expression in this strain should be nearly identical. As shown in Table 2, this appeared to be true in that the relative levels of *b*-lac and FLD activity in the strain were similar in cells grown in four expression test media. These four media contained as carbon and nitrogen sources, respectively: (1) glucose and ammonium sulfate (G/NH_4^+), (2) glucose and methylamine (G/MA), (3) methanol and ammonium sulfate (M/NH_4^+), and (4) methanol and methylamine (M/MA). As expected, *b*-lac and FLD activities were highly (although not totally) repressed in cells grown on G/NH_4^+ medium. Cells grown on either G/MA or M/NH_4^+ media contained at least 10-fold more *b*-lac and FLD, with the highest level of both enzymes observed in cells grown in M/MA medium.

The second vector, pSS050, contained the *P. pastoris HIS4* gene as the selectable marker. When cut at a unique *Sal*I site within *HIS4* and transformed in *P. pastoris,* this vector efficiently integrated at the *P. pastoris HIS4* locus. The result of this integration event was a P_{FLD1}-*bla* expression cassette with sequences from pBR322 just 5′ of the 0.6 kb P_{FLD1} fragment (*pB-*P_{FLD1}*-bla*). Comparison of *b*-lac activity levels in this strain with those observed in the *WT-*P_{FLD1}*-bla* strain allowed us to evaluate whether the 0.6 kb fragment contained all upstream regulatory sequences required for normal regulation. Table 2 shows that *b*-lac activity levels in the *pB-*P_{FLD1}*-bla* strain were approx. 2-fold higher than those observed in the *WT-*P_{FLD1}*-bla* strain when grown in each of the four expression test media. These results indicated that most sequences required for normal regulation were present within the P_{FLD1} fragment, but that sequences that constitutively repress P_{FLD1} by a factor of about 2-fold existed somewhere 5′ of the P_{FLD1} fragment and were missing from the 0.6 kb fragment.

Lastly, we compared levels of *b*-lac activity produced under control of P_{FLD1} with those of a strain in which *bla* expression was under the transcriptional control of P_{AOX1} (Waterham et al., 1997). As previously reported, P_{AOX1} expression is strongly repressed in the glucose-containing media and was highly and specifically induced in methanol-containing media (Tschopp et al., 1987; Waterham et al., 1997) (Table 2). Comparable levels of *b*-lac were present in cells of the *WT-*P_{FLD1}*-bla* strain grown in either M/NH_4^+ or M/MA media, whereas cells

Table 2
Comparison of *b*-lactamase activity in extracts of *P. pastoris* strains expressing *bla* under control of P_{FLD1} and P_{AOX1}

Strain	source of:[a]		Enzyme activity[b]			
	C	N	*b*-Lactamase		FLD	
WT-P_{FLD1}*-bla* (at *FLD1* locus)	G	NH_4^+	14	(4%)	0.13	(6%)
	G	MA	168	(48%)	1.50	(69%)
	M	NH_4^+	310	(88%)	1.69	(78%)
	M	MA	352	(100%)	2.16	(100%)
pB-P_{FLD1}*-bla* (at *HIS4* locus)	G	NH_4^+	19	(5%)	0.11	(5%)
	G	MA	357	(102%)	0.82	(38%)
	M	NH_4^+	529	(150%)	1.48	(69%)
	M	MA	530	(151%)	1.75	(81%)
P_{AOX1}*-bla*	G	NH_4^+	0.3	(0.1%)	0.12	(6%)
	G	MA	0.5	(0.1%)	0.65	(30%)
	M	NH_4^+	241	(68%)	1.40	(65%)
	M	MA	254	(72%)	2.06	(95%)

[a] Each strain was grown in media containing either glucose (G) or methanol (M) as carbon source and ammomum sulfate (NH_4^+) or methylamine (MA) as nitrogen source.
[b] *b*-Lactamase activities are expressed as nmol/mg/min and, in parentheses, as a percentage of activity seen in the *WT-*P_{FLD1}*-bla* strain grown on methanol and methylamine. Activities represent the mean of three experiments using two independently transformed strains.

of the *pB-*P_{FLD1}*-bla* strain contained levels of *b*-lac that were significantly higher than those in the P_{AOX1}*-bla* strain. Especially noteworthy were the levels of *b*-lac in the *pB-*P_{FLD1}*-bla* strain on M/NH_4^+ and M/MA media which were consistently about twice those observed in the P_{AOX1}*-bla* strain on the same media.

4. Discussion

In this report, we describe the isolation and partial characterization of the *FLD1* gene from *P. pastoris* and its promoter. Glutathione-dependent FLD is a key enzyme involved in the metabolism of methanol as a carbon source and certain alkylated amines such as methylamine and choline as nitrogen sources (Veenhuis et al., 1983; Zwart et al., 1980). Its primary role appears to be the protection of cells from the toxic effects of formaldehyde, with a side benefit of yielding net reducing power in the form of NADH (Sibirny et al., 1990). Our observation that *fld1* mutants of *P. pastoris* are defective specifically in the ability to grow on methanol or methylamine demonstrates the importance of FLD in these metabolic pathways.

We provide three independent sources of evidence, each indicating that the *P. pastoris* gene described here is, in fact, *FLD1*. The first is the phenotype of *P. pastoris fld1* mutant strains. As mentioned above, all are specifically defective in the ability to utilize methanol as carbon source and methylamine as nitrogen source. These pathways are known to share only two enzymes, FLD and FDH (see Fig. 1). Furthermore, analysis of activities for the major methanol metabolic pathway enzymes (AOX, catalase, dihydroxyacetone synthase, dihydroxyacetone kinase, FLD and FDH) in methanol-induced cells of the *fld1* strains showed that only FLD was absent. The second piece of evidence is provided by the predicted primary sequence of the cloned gene which shows highest identity (71%) to a yeast (*C. maltosa*) FLD (Sasnauskas et al., 1992). The sequence shows lesser but significant similarity to other higher eukaryotic alcohol dehydrogenases. Third, we show that cloned DNA fragments from both *P. pastoris* and the related yeast *H. polymorpha* restore FLD activity to a *P. pastoris fld1* mutant. Most importantly, the FLD activity restored by the cloned *H. polymorpha* fragment has thermal stability properties similar to FLD from wild-type *H. polymorpha* and that the cloned *P. pastoris* fragment conferred thermal stability properties similar to FLD from wild-type *P. pastoris.*

A major objective of these studies was to evaluate the *P. pastoris FLD1* promoter, P_{FLD1}, as a potential alternative to the commonly used P_{AOX1} for expression of foreign genes in *P. pastoris.* Our results with *P*$_{FLD1}$*-bla* expression strains demonstrate that P_{FLD1} is a highly regulatable promoter that is capable of producing a heterologous protein at levels equal to or higher than those produced by P_{AOX1}. Moreover, use of P_{FLD1} provides a wider choice of conditions for expression than P_{AOX1}. In addition to methanol, expression from P_{FLD1} can also be induced by methylamine as nitrogen source (with glucose, glycerol or other as carbon sources). The ability to induce foreign gene expression in *P. pastoris* without methanol may be useful for certain applications of the system where methanol induction is inconvenient, such as in shake-flask studies where the rapid rate of methanol evaporation makes it difficult to know how much methanol is actually present in cultures or in certain large-scale fermentor culture processes where the

large amounts of methanol needed for growth and induction may be a potential fire hazard.

Acknowledgement

We thank Jessica Gettemy for advice with the RT-PCR reactions and Terrie Hadfield for help in preparing the manuscript. This work was supported by a contract from Phillips Petroleum Company (Bartlesville, OK) and grants from the US National Institutes of Health (DK43698) and the US National Science Foundation (MCB-9514289) to J.M.C.

References

Bradford, M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Anal. Biochem. 72, 248–254.

Cregg, J.M., 1998. Expression in the methylotrophic yeast *Pichia pastoris*. In: Fernandez J., Hoeffler, J. (Eds.), Nature: The Palette for the Art of Expression, ch. 10. Academic Press, San Diego, in press.

Cregg, J.M., Barringer, K.J., Hessler, A.Y., Madden, K.R., 1985. *Pichia pastoris* as a host system for transformations. Mol. Cell. Biol. 5, 3376

Cregg, J.M., Russell, K.A., 1998. Transformation. In: Higgins, D.R., Cregg, J.M. (Eds.), Methods in Molecular Biology: *Pichia* Protocols, ch. 3. Humana Press, Totowa, NJ, pp. 27–39.

Domdey, H., Apostol, B., Lin, R.J., Newman, A., Brody, E., Abelson, J., 1984. Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. Cell 39, 611–621.

Frohman, M.A., Dush, M.K., Martin, G.R., 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA 85, 8998–9002.

Giri, R., Krug, J.F., Kozak, C., Moretti, T., O'Brien, S.J., Seuanez, H.N., Goldman, D., 1989. Cloning and comparative mapping of human class III alcohol dehydrogenase cDNA. Biochem. Biophys. Res. Commun. 164, 453–460.

Higgins, D.R., Cregg, J.M., 1998. Methods in Molecular Biology: *Pichia* Protocols. Humana Press, Totowa, NJ, pp. 1–15.

Holmquist, B., Vallee, B.L., 1991. Human liver class III alcohol and glutathione dependent formaldehyde dehydrogenase are the same enzyme. Biochem. Biophys. Res. Commun. 178, 1371–1377.

Jornvall, H., Persson, B., Jeffery, J., 1987. Characteristics of alcohol/polyol dehydrogenases: the zinc-containing long-chain alcohol dehydrogenases. Eur. J. Biochem. 167, 195–201.

Koivusalo, M., Barmann, M., Uotila, L., 1989. Evidence for the identity of glutathione-dependent formaldehyde dehydrogenase and class III alcohol dehydrogenase. FEBS Lett. 257, 105–109.

Kramer, M.F., Coen, D.M., 1995. The polymerase chain reaction. In: Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G . Smith, J.A., Struhl, K. (Eds.), Current Protocols in Molecular Biology, vol. 2, ch. 15. John Wiley, New York, pp. 15.1.1–15.1.9.

Liu, H., Tan, X., Russell, K.A., Veenhuis, M., Cregg, J.M., 1995. *PER3*, a gene required for peroxisome biogenesis in *Pichia pastoris*, encodes a peroxisomal membrane protein involved in protein import. J. Biol. Chem. 270, 10940–10951.

Lück, H., 1963. Catalase. In: Bergmeyer, H.U. (Ed.), Methods of Enzymatic Analysis. Academic Press, San Diego, pp. 885–894.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sasnauskas, K., Jomantiene, R., Januska, A., Lebediene, E., Lebedys, J., Janulaitis, A., 1992. Cloning and sequencing analysis of a *Candida maltosa* gene which confers resistance to formaldehyde in *Saccharomyces cerevisiae*. Gene 122, 207–211.

Schmitt, M., Brown, T.A., Trumpower, B.L., 1990. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. Nucleic Acids Res. 18, 3091–3092.

Schutte, H., Flossdorf, J., Sahm, H., Kula, M.R., 1976. Purification and properties of formaldehyde dehydrogenase and formate dehydrogenase from *Candida boidinii*. Eur. J. Biochem. 62, 151–160.

Sibirny, A.A., Ubiyvovk, V.M., Gonchar, M.V., Titorenko, V.I., Voronovsky, A.Y., Kapultsevich, Y.G., Bliznik, K.M., 1990. Reactions of direct formaldehyde oxidation to CO_2 are non-essential for energy supply of yeast methylotrophic growth. Arch. Microbiol 154, 566–575.

Sreekrishna, K., 1993. Strategies for optimizing protein expression and secretion in the methylotrophic yeast *Pichia pastoris*. In: Baltz, R.H., Hegeman, G.D., Skatrud, P.L. (Eds.), Industrial Microorganisms: Basic and Applied Molecular Genetics. American Society for Microbiology, Washington, DC, pp. 119–126.

Stewart, P., Kersten, P., Wymelenberg, A.V., Gaskell, J., Cullen, D., 1992. Lignin peroxidase gene family of *Phanerochaete chrysosporium*: complex regulation by carbon and nitrogen limitation and identification of a second dimorphic chromosome. J. Bacteriol. 174, 5036–5042.

Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C.A., Gingeras, T.R., 1987. Expression of the *lacZ* gene from two methanol-regulated promoters in *Pichia pastoris*. Nucleic Acids Res. 15, 3859–3876.

van Dijken, J.P., Harder, W., Beardsmore, A.J., Quayle, J.R., 1978. Dihydroxyacetone: an intermediate in the assimilation of methanol by yeasts? FEMS Microbiol. Lett. 4, 97–102.

van Dijken, J.P., 1976. Oxidation of methanol by *Hansenula polymorpha*: purification and kinetic properties of methanol oxidase. Ph.D. Thesis, University of Groningen, pp. 30–43.

van der Klei, I.J., Bystrykh, L.V., Harder, W., 1990. Alcohol oxidase from *Hansenula polymorpha* CBS 4732. Methods Enzymol. 188, 420–422.

Veenhuis, M., van Dijken, J.P., Harder, W., 1983. The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts. Adv. Microb. Physiol. 24, 1–82.

Waites, M.J., Quayle, J.R., 1981. The interrelation between transketolase and dihydroxyacetone synthase activities in the methylotrophic yeast *Candida boidinii*. J. Gen. Microbiol. 124, 309–316.

Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V., Cregg, J.M., 1997. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene 186, 37–44.

Waterham, H.R., Keizer-Gunnink, I., Goodman, J.M., Harder, W., Veenhuis, M., 1992. Development of multi-purpose peroxisomes in *Candida boidinii* grown in oleic acid–methanol limited continuous cultures. J. Bacteriol. 174, 4057–4063.

Zwart, K., Veenhuis, M., van Dijken, J.P., Harder, W., 1980. Development of amine oxidase-containing peroxisomes in yeast during growth on glucose in the presence of methylamine as the sole source of nitrogen. Arch. Microbiol. 126, 117–126.