Isolation of the *Pichia pastoris* Formaldehyde Dehydrogenenase Gene and Use of its Promoter for Expression of Foreign Genes in *P. pastoris*

Shigang Shen

B.S., Nanking University, 1987

M.D., Tianjin Medical College, 1992

A thesis submitted to the faculty of the

Oregon Graduate Institute of Science and Technology

in partial fulfillment of the

requirements for the degree

Master of Science

in

Biochemistry and Molecular Biology

October 2000

The thesis "Isolation of the *Pichia pastoris* formaldehyde dehydrogenase gene and use of its promoter for expression of foreign genes in *P. pastoris*" by Shigang Shen has been examined and approved by the following Examination Committee:

James M. Cregg, Thesis Advisor
Professor

Gebre Woldegiorgis
Associate Professor

Peter Zuber
Professor

ii

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

Isolation of the *Pichia pastoris* Formaldehyde Dehydrogenase Gene and Use of its Promoter for Expression of Foreign Genes in *P. pastoris*

Shigang Shen

M.S., Oregon Graduate Institute of Science and Technology

October 2000

Thesis Advisor: Dr. James M. Cregg

In methylotrophic yeasts, glutathione-dependent formaldehyde dehydrogenase (FLD) is a key enzyme 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. 40KDa) 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 β-lactamase as a reporter, we show that the *FLD1* promoter (P_{FLD1}) is storngly 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 β-lactamase produced under control of P_{FLD1} 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_{AOX} 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.

CHAPTER 1

INTRODUCTION*

1.1   Uses of Methylotrophic Yeasts in Applied and Basic Research

The existence of methylotrophic yeasts capable of utilizing methanol as sole carbon and energy source was first reported by Ogata and coworkers [Ogata et al., 1969]. During the past 30 years, methylotrophic yeasts have attracted the attention of industrial and academic investigators because of certain unique properties related to methanol utilization. During the 1970s, there was interest in their use for single cell protein (SCP) production, primarily as a source of high-protein animal feed [Cooney and Levine, 1972; Smith, 1981]. Beginning in the early 1980s, interest was generated in the use of methylotrophic yeasts as models for investigations on peroxisomal biosynthesis [Aitchison et al., 1992; Waterham et al., 1997] and in their use as hosts for the production of heterologous proteins [Romanos et al., 1992; Cereghino and Cregg, 2000].

1.1.1 Metabolism of single-carbon compounds in yeast

Single-carbon or C_1 compounds are defined as organic compounds that are more reduced than carbon dioxide and contain no carbon–carbon bonds. Examples of such compounds include methane, methanol, formaldehyde, formate, and methylamine. There are many microorganisms found in nature that are capable of

---

growth in C_1 compounds. Although most of these are prokaryotes and archaebacteria, some yeasts and fungi are able to metabolize certain C_1 compounds [Lee and Komagata, 1980]. Methylotrophs, as previously defined [Colby and Dalton, 1978], obtain their energy from the oxidation of C_1 growth substrates and assimilate carbon as formaldehyde or as a mixture of formaldehyde and CO_2. Most prefer methanol as a substrate, hence the name methylotrophs. Among fungi, certain species of yeast in the genera *Candida*, *Hansenula*, *Pichia* and *Torulopsis* are methylotrophic organisms [Veenhuis et al., 1983]. They also can use methylated amines as nitrogen source but not as sole source of carbon and energy. Below is summarized the metabolic pathway for utilization of methanol and methylamine by yeasts.

1.1.1.1 Methanol. The first methylotrophic yeasts were not discovered until 1969 [Ogata et al., 1969]. Beginning in the early 1970s, there was intense interest in exploiting these yeasts for the production of biomass or SCP from methanol, particularly in the petroleum industry. At that time, methane (natural gas) was virtually a waste product of the industry and could be inexpensively oxidized to methanol by chemical processes. Thus, methanol was viewed as a potential inexpensive feed stock for the large-scale production of microbial SCP, primarily for use as a high-protein animal feed. In parallel with the development of SCP production processes, academic laboratories elucidated the biochemical pathway used by yeasts to metabolize methanol [van Dijken et al., 1978]. It was discovered that the ability to use methanol as a carbon and energy source requires a unique pathway involving three peroxisomal enzymes: alcohol oxidase (AOX), catalase, and dihydroxyacetone synthase [Anthony, 1982; Veenhuis et al., 1983; Douma et al., 1985; Goodman, 1985]. As diagrammed in Fig. 1.1, the first step of the pathway involves the oxidation of methanol to formaldehyde by AOX, a peroxisomal enzyme. This reaction generates hydrogen peroxide, which is then degraded to water and oxygen by catalase, the classic peroxisomal marker enzyme. Two different pathways further metabolize formaldehyde. In one path, a portion of the formaldehyde leaves the peroxisome to the cytosol and is further oxidized to formate by formaldehyde dehydrogenase (FLD) and then to carbon dioxide by formate dehydrogenase (FDH).
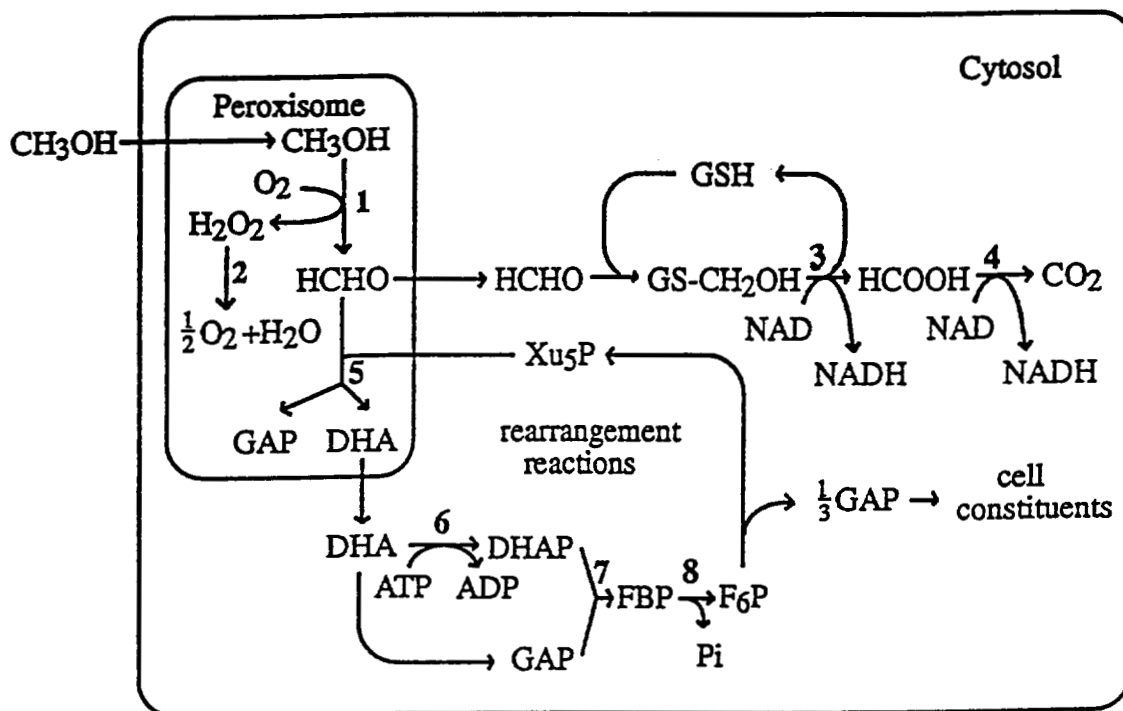
Figure 1.1 The methanol pathway in yeasts. (**1**) Alcohol oxidase; (**2**) catalase; (**3**) formaldehyde dehydrogenase; (**4**) formate dehydrogenase; (**5**) dihydroxyacetone synthase; (**6**) dihydroxacetone kinase; (**7**) fructose 1,6-bisphosphate aldolase; (**8**) fructose 1,6-bisphosphatase.

In the other pathway, a portion of the formaldehyde remains in the peroxisome and condenses with xylulose-5'-monophosphate in a reaction catalyzed by the third peroxisomal enzyme, dihydroxyacetone synthease. The products of this reaction, glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone, then leave the peroxisome and enter a cyclic pathway that regenerates xylulose-5'-monophosphate plus one net molecule of GAP for every three turns of the cycle. GAP is used for biosynthesis of carbon skeletons for cell growth.

In the energy generation pathway from methanol, formaldehyde is oxidized to formate in either a glutathione-independent or glutathione-dependent manner, depending upon the yeast species. The former reaction involves a methylformate synthase, and the latter involves a glutathione-dependent FLD. Both of these reactions produce reducing power in the form of NADH (nicotinamide adenine dinucleotide, reduced form). There are two models for the role of this oxidative pathway for the yeast. One model 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 of the energy for methanol-dependent growth comes from the oxidation of one or more of the xylulose-5'-monophosphate cycle intermediates by tricarboxylic acid cycle enzymes, and that the primary role of FLD and FDH is to protect the yeast from toxic formaldehyde that can accumulate in the cell with excess methanol in the medium [Sibirny et al., 1990].

1.1.1.2 Methylated amines. Many species of yeasts (with the notable exception of *Saccharomyces cerevisiae*) are able to utilize various alkylated amines as sole nitrogen source, but not as sole carbon source [van Dijken and Bos, 1981]. A key enzyme required for this metabolism is amine oxidase which *S. cerevisiae* does not have [Blaschko et al., 1937]. Many methylotrophic yeasts contain either or both of two types of amine oxidase: (i) methylamine oxidase, a heat-labile enzyme with high affinity for short-chain alkylamines, or (ii) benzylamine oxidase, a relatively heat-stable enzyme that has a high affinity for longer chain and aromatic alkylamines. Both enzymes are located in the peroxisome [Green and Large, 1984]. The substrate

specificity of the two enzymes overlaps considerably; they cannot be considered as complementary enzymes [Haywood and Large, 1981].

In the yeasts *Candida utilis*, *Candida boidinii*, and *Hansenula polymorpha*, the key enzyme in the metabolism of methylamine is amine oxidase [Zwart et al., 1980]. This peroxisomal enzyme catalyzes the oxidation of methylamine to formaldehyde, hydrogen peroxide, and ammonia as follows:

$$CH_3NH_2 + O_2 + H_2O \rightarrow HCHO + H_2O_2 + NH_3$$
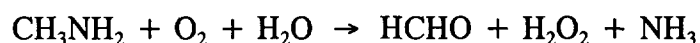
Formaldehyde, in the methylotrophic yeast *H. polymorpha*, can either be further oxidized via formate to CO_2 by FLD and FDH [Zwart et al., 1980], or assimilated into cellular materials via the xylulose monophosphate cycle, the same pathways used to metabolize formaldehyde generated from methanol [van Dijken et al., 1978]. In the non-methylotrophic yeast *C. utilis*, it is not possible to assimilate formaldehyde, but this compound can be oxidized to CO_2 by NAD-dependent FLD and FDH. Thus, metabolism of formaldehyde in this yeast can only yield metabolic energy. In all yeasts, the hydrogen peroxide generated by the oxidation of methylamine by amine oxidase is degraded to oxygen and water by peroxisomal catalase, as described for the methanol pathway. Ammonia is incorporated into 2-oxoglutarate in a reaction mediated by NADPH-dependent glutamate dehydrogenase [Zwart and Harder, 1983].

1.1.2 Peroxisome biogenesis mechanisms

Peroxisomes are single-membrane-enclosed organelles found in all eukaryotes [Waterham et al., 1997; Tabak et al., 1999]. They are unique among organelles in that they are highly diverse—varying dramatically in size, abundance, and enzyme content, depending upon the metabolic needs of the organism or tissue. Methylotrophic yeasts have been exploited for a combined morphological, biochemical, and molecular genetic investigation of peroxisome biogenesis. In *H. polymorpha* and *Pichia pastoris*, growth on methanol induces the synthesis of large amounts of three peroxisomal enzymes: AOX, catalase, and dihydroxyacetone

synthease. In methanol medium, up to 80% of the cytoplasmic volume is occupied by peroxisomes [Veenhuis and Harder, 1991]. However, in glucose medium, peroxisomes are small and few in number. Because fatty acid β-oxidation reactions occur exclusively in peroxisomes in yeasts [Tanaka et al., 1982], growth of these organisms on fatty acids (such as oleic acid) results in a substantial induction of peroxisomes containing enzymes of the peroxisomal β-oxidation pathway. *P. pastoris* is unique among yeasts in that it can grow on either of the peroxisome-inducing carbon sources, methanol or oleate.

The key to the application of molecular genetics to yeast peroxisome studies was the discovery of methods to isolate peroxisome biogenesis mutants of the organism [Erdman et al., 1989; Cregg et al., 1990]. Methylotrophic yeast mutants defective in the proper assembly of peroxisomes (*pex* mutants) were first isolated in *H. polymorpha* [Cregg et al., 1990] and subsequently in *P. pastoris* [Liu et al., 1992]. The *P. pastoris pex* mutants are able to grow at normal rates on most carbon sources, including glucose, glycerol, and ethanol, but are unable to grow on either methanol or oleic acid [Liu et al., 1992]. At present, 15 different *PEX* gene complementation groups have been identified in *P. pastoris* with a similar number in *H. polymorpha* [Tan et al., 1995; Johnson et al., 1999]. Most of these genes have been cloned from yeast genomic DNA libraries by functional complementation, and their products have been studied in detail. Thus, the ability to induce proliferation of peroxisomes and peroxisomal proteins by methanol (and oleic acid in *P. pastoris*) has made methylotrophic yeasts attractive systems for the study of peroxisome biogenesis.

1.1.3 Selective peroxisome autophagy

In methylotrophic yeasts such as *H. polymorpha* and *P. pastoris*, peroxisomes proliferate in size and number in response to certain changes in the cell's environment. As discussed in Section 1.2, peroxisomes contain key enzymes involved in the oxidative metabolism of carbon sources, such as methanol and oleic acid, and nitrogen sources, such as methylamine. As a result of the induction and import of these enzymes into peroxisomes, the organelles become large and numerous. However, upon shift of yeasts to conditions where peroxisomes are no

longer needed, such as from methanol medium to glucose medium, the organelles and their enzymatic contents are rapidly and selectively degraded by a process called autophagy [Veenhuis et al., 1978; Dunn, 1994; Klionsky, 1998]. In this process, peroxisomes are targeted to and sequestered within the yeast vacuole and degraded by the resident hydrolases. Autophagy is a general protein and organellar degradation process that takes place in all eukaryotic cells.

Two types of autophagic processes have been observed: microautophagy and macroautophagy [Tuttle et al., 1993; Tuttle and Dunn, 1995; Chiang et al., 1996]. Microautophagy is a process in which cellular components are surrounded by invaginations of the vacuolar membrane or by finger-like protrusions from the vacuole. The resulting intravacuolar vesicles containing the sequestered cellular components are eventually degraded by hydrolytic enzymes [Ahlberg et al., 1985; Mortimore et al., 1988]. Macroautophagy is a process in which cellular components are first sequestered within an autophagosome that arises from an invagination of the rough endoplasmic reticulum [Dunn, 1990]. The autophagosome then fuses with the vacuole, thereby delivering its contents to the vacuolar hydrolases.

Electron microscopy studies suggest that exposure of methanol-grown *H. polymorpha* to either glucose or ethanol induces peroxisomes to be degraded by macroautophagy [Veenhuis et al., 1983; Tuttle et al., 1993]. Similary, peroxisomes in methanol-grown *P. pastoris* cells exposed to ethanol are also subject to macroautophagy. However, peroxisomes in methanol-grown cells of *P. pastoris* appear to be degraded microautophagically when exposed to glucose [Tuttle and Dunn, 1995].

Methylotrophic yeasts have at least three significant advantages as model systems to investigate autophagic processes at the molecular level. The first is that methanol-induced peroxisomes are large, and their degradation by autophagy is easily visualized by standard microscopic techniques. The second is that autophagy can be induced synchronously in these yeasts by a simple shift of carbon sources. Moreover, in *P. pastoris*, the specific type of peroxisomal degradation process (i.e., micro- versus macroautophagy) can be selected via the appropriate carbon source. The third is that it is possible to combine a molecular genetic and biochemical attack toward

understanding autophagy in these yeasts. Through the study of peroxisome degradation mutants, two genes required for selective macroautophagy of peroxisomes in *H. polymorpha* and eight genes required for glucose-induced selective autophagy of peroxisomes in *P. pastoris* have been identified [Titorenko et al., 1995; Yuan et al., 1997, 1999].

1.2   Function and Structure of Eukaryotic Alcohol Dehydrogenases

In mammals, class III alcohol dehyrogenases (ADHs) and glutathione-dependent FLDs are the same enzymes [Koivusalo et al., 1989]. ADHs have been characterized in a wide variety of bacterial, yeast, plant, and animal species including humans. Most yeast ADHs are tetrameric enzymes containing zinc [Magonet et al., 1992]. Four ADHs have been characterized in the yeast *S. cerevisiae* [Bennetzen and Hall, 1982; Young and Pilgrim, 1985; Paquin and Williamson, 1986; Drewke and Ciriacy, 1988]. *S. cerevisiae* ADHI, encoded by the *ADH1* gene, is a fermentative enzyme that catalyzes the conversion of acetaldehyde to ethanol and is expressed in the presence of glucose in the culture medium [Denis et al., 1983]. ADHII is an oxidative isozyme of ADHI and is encoded by the *ADH2* gene. ADHII catalyzes the reverse reaction of ADHI and functions in the metabolism of ethanol as an energy and carbon source [Young and Pilgrim, 1985]. ADHII is highly repressed by glucose and derepressed in ethanol [Ciriacy, 1975]. ADHIII, encoded by the *ADH3* gene, is localized within the mitochondrion and is repressed by glucose [Young and Pilgrim, 1985]. The function of ADHIII is unknown [Passoth et al., 1998]. All three of these *S. cerevisiae* ADHs exhibit strong amino acid sequence similarity to each other [Young and Pilgrim, 1985]. ADHIV, whose function is unknown in yeast, has amino acid sequence similarity to iron-activated ADH2 from the bacterium *Zymomonas mobilis*, which is a major component of ADH activity in the fermentative bacterium [Williamson and Paquin, 1987; Reid and Fewson, 1994]. However, ADHIV does not share sequence similarity with the other *S. cerevisiae* ADHs [Drewke and Ciriacy, 1988].

On the basis of electrophoretic mobilities, kinetic properties, and inhibition by pyrazole, ADHs from the tissues of many vertebrates have been divided into five classes [Vallee and Bazzone, 1983]. Class I represents the classical liver zinc-containing enzyme, exhibiting considerable activity toward ethanol as a substrate [Jornvall et al., 1995]. Class III ADHs function *in vitro* as glutathione-dependent FLDs [Koivusalo et al., 1989]. Class IV ADHs are present only in stomach tissue and have the highest activity toward ethanol of all ADH classes [Pares et al., 1994]. Classes II and V are less well understood and have not been characterized in more than one species each [Jornvall et al., 1995].

1.2.1 Class III ADHs

As mentioned in the previous section, class III ADHs appear to be glutathione-dependent FLDs. These enzymes metabolize formaldehyde to which cells may be exposed from their environment or which may be generated as a product of normal metabolism [Koivusalo et al., 1989]. In contrast to the class I and II ADH enzymes, class III enzymes are insensitive to inhibition by pyrazole and have a low affinity for ethanol. In addition, unlike class I and II ADHs, they can catalyze the oxidation of long-chain aliphatic alcohols [Vallee and Bazzone, 1983]. Finally, class III ADHs are structurally distinct from the class I and II enzymes [Kaiser et al., 1988]. Accordingly, antibodies against human class III ADH do not cross-react with the class I and II enzymes [Montavon et al., 1989]. In contrast to class I ADH, which is absent in some animals including lower vertebrates and invertebrates [Kaiser et al., 1993], a class III ADH is found in all organisms from bacteria to humans [Schutte et al., 1976; Burnett and Felder, 1978; Pares and Vallee, 1981; Allais et al., 1983; Gutheil et al., 1992; Sasnauskas et al., 1992; Estonius et al., 1993; Wehner et al., 1993; Danielsson et al., 1994; Ras et al., 1995]. The mammalian class III enzymes are less variable in structure than those in class I. The sequences of the rat and horse class III ADHs are very similar to those of the human enzyme, showing differences in only 21 and 19 residues, respectively, out of a total of 373 amino acids [Julia et al., 1988; Kaiser et al., 1989]. In contrast, the class I enzymes from the same organisms differ by 84 amino acids. The similarity of class III enzyme sequences suggests that

the metabolic role of this class of ADHs has been highly conserved through evolution [Kaiser et al., 1989].

1.2.2 Yeast FLDs

FLDs catalyze a key step in the catabolism of methanol and methylated amines in yeasts. They have been purified from *S. cerevisiae* and the methylotrophic yeasts *C. boidinii*, *H. polymorpha* and *P. pastoris* [Kato, 1990]. FLD is NAD-linked and glutathione-dependent. $NADP^+$ cannot replace NAD^+ as a cofactor. The addition of reduced glutathione is required for the oxidation of formaldehyde by purified FLD enzymes. FLD's true substrate, however, is a hemimercaptal, S-hydroxymethyglutathione, spontaneously formed from formaldehyde and glutathione. Reduced glutathione cannot be replaced by other thiol compounds, such as cysteine, dithiothreitol, or thioglycerol [Patel et al., 1983]. FLD activity is completely inhibited by metal ions, such as Cd^{2+}, Hg^{2+}, Cu^{2+}, and Ag^+. As shown below, the reaction product is S-formylglutathione.

$$Formaldehyde + glutathione + NAD^+ = S\text{-formylglutathione} + NADH + H^+$$
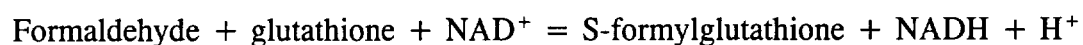
FLDs have a molecular weight of approximately 84,000 kDa as determined by gel filtration, and a subunit molecular weight of 41,000 kDa as determined by SDS-PAGE [Patel et al., 1983].

1.2.3 Evolution of the ADH family

The primary sequence of FLD from *S. cerevisiae* shares 71% identity with that of FLD from *Candida maltosa* and 63% identity with that of the human class III enzyme [Sharma et al., 1989; Sasnauskas et al., 1992]. The amino acid sequence homology and identical structural and kinetic properties indicate that FLDs and mammalian class III ADHs are closely related evolutionarily [Koivusalo et al., 1989].

The class III family of ADH enzymes appears to be functionally equivalent to FLDs in mammals [Koivusalo et al., 1989]. Functionally and structurally, class I ADHs are least related to the class III ADHs [Kaiser et al., 1993]. Structure and

activity studies on class III ADHs have been performed in prokaryotes [Gutheil et al., 1992], yeasts [Sasnauskas et al.,1992; Wehner et al., 1993], invertebrates (octopus) [Kaiser et al., 1993], vertebrates [Danielsson and Jornvall, 1992], and mammals [Kaiser et al., 1989]. Most of these studies were based on comparisons of class III and class I ADHs by amino acid sequence alignments. These studies suggested that the class III type with its glutathione-dependent FLD activity is the progenitor of both class I and III vertebrate ADHs. However, ethanol active class II ADHs of vertebrates do not appear to have arisen from an ethanol active class II yeast ADH. They appear to have arisen from independent duplicatory events and evolved towards similar substrate specificities by convergence [Danielsson and Jornvall, 1992].

1.3   Transcriptional Regulation by Methanol in Yeasts

1.3.1 General features of eukaryotic transcription initiation

Initiation of transcription represents a major control point for gene expression in eukaryotes. This step is mediated by RNA polymerase II (PolII) and a complex array of general initiation factors (GIFs) by a mechanism that is conserved from yeasts to man. In eukaryotes, PolII transcribes mRNAs and several small nuclear RNAs. PolII cannot recognize its target promoters directly. Genes transcribed by RNA PolII typically contain common core promoter elements that are recognized by one or more GIFs and gene-specific DNA elements that are recognized by transcriptional regulatory factors. These regulatory factors, in turn, modulate the function of the general initiation factors [Roeder, 1996].

Core-promoter elements are defined as minimal DNA elements that are necessary and sufficient for accurate transcription initiation by RNA PolII in reconstituted cell-free systems. The most common of these elements are the TATA box (consensus TATAa/tAa/t), located near position –25 relative to the transcription start site, and a pyrimidine-rich initiator located near the transcription start site [Nikolov and Burley, 1997]. The first step in transcription initiation is the recruitment of the GIF transcription factor IID (TFIID) to the TATA element. TFIID is the only GIF with components capable of sequence-specific binding to the

eukaryotic promoter DNA. These components include the TATA-binding protein (TBP), which directly recognizes the TATA element, and certain TBP-associated factors [Burley and Roeder, 1998]. Binding of TFIID is followed by the association of additional GIFs with the building complex and ultimately leads to the recruitment of PolII.

Most genes in eukaryotes show a regulated pattern of expression during development, during the course of the cell cycle, or in response to changes in the cellular environment [Bjorklund and Kim, 1996]. Transcriptional regulation of gene expression is typically mediated by multiple activator or repressor proteins that bind specifically to positively or negatively acting sequences in the promoter regions [Verdier, 1990]. Positive regulatory proteins, referred to as activators, often exert their effects through intermediary factors that convey signals from the activators to the transcription complex [Flanagan et al., 1991, 1992; Meisterernst et al., 1997]. Three types of intermediary factors have been described: (i) TBP-associated factors, which associate with TBP to form TFIID; (ii) a mediator which associates with PolII to form a holoenzyme; and (iii) a chromatorgraphic fraction from human cells, termed upstream stimulatory activity, whose resolution is still in progress [Bjorklund and Kim, 1996]. DNA-binding transcriptional repressors act by a variety of mechanisms. They can repress transcription by competing with activators or basal transcription factors for access to the DNA or recruiting inhibitory chromatin components to the promoter [Cowell, 1994; Johnson, 1995; Hanna-Rose and Hansen, 1996]. The primary mode of action of activators and repressors is to influence the rate of transcription initiation of each gene. The most common mechanism for this appears to be through the recruitment of PolII to the promoter site (or in the case of repressors, through blockage of the recruitment of PolII). A second mechanism appears to act through the acetylation and deacetylation of nucleosomes on DNA around the promoter site [Pazin and Kadonaga, 1997]. Acetylation neutralizes the strong basic character of nucleosomes, thus "loosening" their grip on DNA and providing more opportunity for the transcription apparatus to establish interaction with the genetic promoter.

1.3.2 Methanol regulation

As described in Section 1.2.1, methylotrophic yeasts (such as *H. polymorpha*, *P. pastoris* and *C. boidinii*) have the ability to utilize methanol as a carbon and energy source. The first step in methanol metabolism is performed by AOX [van der Klei et al., 1991]. The *AOX* genes are regulated at the transcriptional level [Ellis et al., 1985; Cregg et al., 1989] and are considered to be among the most strongly expressed and tightly regulated genes known in yeast [Tschopp et al., 1987].

Mechanisms of transcriptional regulation in methylotrophic yeasts have been investigated by analyzing the promoter sequences of the *H. polymorpha MOX* gene [Godecke et al., 1994], *P. pastoris AOX1* and *AOX2* genes [Cregg et al., 1989; Ohi et al., 1994], and *C. boidinii* dihydroxyacetone synthase gene (*DAS1*) [Sakai et al., 1998]. Analysis of the *H. polymorpha MOX* promoter (P_{MOX}) region revealed three *cis*-acting regulatory elements which are each occupied by a DNA-binding protein [Godecke et al., 1994]. The first upstream activating sequence element (*UAS1*) is located between nucleotides −507 and −430. When fused to a β-lactamase (β-lac) reporter gene, deletion of the region containing this element lowers β-lac activity 95-fold under derepressed (glycerol-grown) conditions and 35-fold under induced (methanol-grown) conditions compared with wild-type P_{MOX} levels. In methanol-grown cell extracts, a protein named MBF1 (MOX binding factor 1) was identified that specifically binds the *UAS1* region and appears to be a positive factor for P_{MOX} [Godecke et al., 1994]. *UAS2* is located between nucleotides −822 and −729 with respect to the translational start site or ATG. Reporter studies suggest an activating function for this sequence in derepressed cells but not in induced cells of *H. polymorpha*. *UAS2* is bound by MBF2, a homodimer, in a region of internal sequence symmetry. An upstream repression sequence (*URS1*), located between nucleotides −897 and −1008, exerts a repressive influence on expression from P_{MOX}. High β-lac expression is observed in glucose-grown cells after deletion of *URS1*. MBF3 was identified as a potential *URS1* binding factor that modulates transcription by interacting with the URS1 site [Godecke et al., 1994].

P_{MOX} function has also been investigated in *S. cerevisiae*, a species that lacks methanol metabolism [Pereira and Hollenberg, 1996]. In these studies, a 200-bp region (-361 to -112) of P_{MOX}, termed *MOX-B*, was found to regulate expression in *S. cerevisiae* strains carrying a *MOX-lacZ* reporter. Deletion of this region led to high levels of β-galactosidase in either *S. cerevisiae* or *H. polymorpha* cells growing in glucose. However, in *S. cerevisiae*, the three *cis*-acting elements described above do not affect *MOX-lacZ* expression. Thus, it is not clear that these deletion study results in *S. cerevisiae* are relevant to P_{MOX} function in *H. polymorpha*.

Interestingly, *MOX-lacZ* expression in *S. cerevisiae* is significantly reduced by disruption of *ADR1*, a gene encoding a well-known transcription factor required for induction of peroxisomal genes in *S. cerevisiae* [Vallari et al., 1992]. Conversely, overexpression of ADR1 strongly enhances *MOX-lacZ* expression. Only *MOX-lacZ* constructs that contain *MOX-B* are regulated by *ADR1*, and the *MOX-B* region appears to confer both Adr1p-dependent catabolite repression and activation of *lacZ* expression [Pereira and Hollenberg, 1996]. DNaseI footprint experiments demonstrated that Adr1p binds *MOX-B* at an inverted GGAGA sequence at position -204 [Pereira and Hollenberg, 1996]. In *H. polymorpha*, a potential Adr1p homologue, named FB2, binds to the *MOX-B* region at sequences located between -160 and -203 [Costanzo et al., 1995], a region that overlaps the Adr1p-protected sequence in *S. cerevisiae* (from -192 to -211). These results suggest that Adr1p may be a conserved transcription factor for genes encoding peroxisomal proteins in yeasts [Shain et al., 1992].

P. pastoris has two alcohol oxidase genes, *AOX1* and *AOX2* [Cregg et al., 1989; Koutz et al., 1989]. Transcriptional reporter studies indicate that regulation of the *AOX1* gene involves both positive and negative control [Tschopp et al., 1987; Cregg and Madden, 1988]. Studies of strains with disrupted *AOX* genes revealed that *AOX1* was the major source of methanol-oxidizing activity in methanol-grown *P. pastoris* [Cregg et al., 1989]. The *AOX2* gene is transcribed at a much lower level than the *AOX1* gene. Apart from this difference in expression levels, the two genes are regulated similarly [Ohi et al., 1994]. Three *cis*-acting regulatory elements were defined in the *AOX2* promoter. One positive *cis*-acting element, *AOX2-UAS*, located

between positions -337 and -313, is required for response to methanol. Artificial amplification of *AOX2-UAS* resulted in an increase in the transcriptional activity of the promoter. Sequences homologous to *AOX2-UAS* are found in the *P. pastoris AOX1* promoter [Koutz et al., 1989], the *C. boidinii AOX* gene promoter [Sakai and Tani, 1992], and the *H. polymorpha DAS1* promoter [Janowicz et al., 1985]. Two negative *cis*-acting elements, *AOX2-URS1* and *AOX2-URS2*, were also identified in the *AOX2* promoter, which were not found in the *AOX1* regulatory region. The transcriptional activity of the *AOX2* promoter was enhanced dramatically by deletion of *AOX2-URS1* or by a mutation in *AOX2-URS2*. In contrast to *AOX1*, which is highly expressed in methanol-grown cells, *AOX2* transcription through *AOX2-UAS* is completely repressed by the unique repression system consisting of *AOX2-URS1* and *AOX2-URS2*. Therefore, *AOX2-URS1-* and *AOX2-URS2*-mediated transcriptional effects are dominant over those at *AOX2-UAS*. The differential expression of the two *P. pastoris AOX* genes might be due to the presence or absence of the functional *URS*.

1.4   Regulation of FLD and FDH by Methylated Amines in Yeasts

The utilization of methylated amines as sole nitrogen source by yeasts was first described by van der Walt [1962]. As mentioned in Section 1.2.2, in the non-methylotrophic yeast *C. utilis* and methylotrophic yeasts *C. boidinii* and *H. polymorpha*, the complete oxidation of methylamine to carbon dioxide and ammonia is mediated by amine oxidase, FLD, and FDH [Zwart et al., 1980]. Of these enzymes, amine oxidase is peroxisomal. Synthesis of amine oxidase in *C. utilis* is regulated by a repression/derepression mechanism in which ammonia plays an essential role [Zwart and Harder, 1983]. An increase in the activities of amine oxidase in *C. utilis* and *H. polymorpha* was observed when alkylated amines instead of ammonium sulphate were used as sole nitrogen source. Repression of amine oxidase synthesis was observed in *H. polymorpha* after transfer of methanol/methylamine-grown cells into methanol/ammonium sulphate-containing media [Veenhuis et al., 1981].

When methylated amines were used as sole nitrogen source, a distinct increase was observed in the specific activities of FLD and FDH. FLD synthesis is also regulated independently in response to either methanol as sole carbon and energy source or to methylamine as sole nitrogen source, because their metabolism yields formaldehyde [Anthony, 1982; Mori et al., 1988]. 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. For cells of the non-methylotrophic yeast *C. utilis* and the methylotrophic yeast *H. polymorpha*, derepression of FLD is observed upon transfer from ammonium ion-containing medium into media containing methylamine as the nitrogen source. Thus, the presence of methylamine as sole nitrogen source can release the repression of FLD [Zwart et al., 1980].

Since the reaction catalyzed by FDH is located downstream of formaldehyde, it is also regulated by methylamine or choline, whose metabolism is supposed to yield formaldehyde and then formate. Regulation of FDH enzyme activity was extensively studied with various combinations of carbon and nitrogen sources [Sakai et al., 1997]. Expression of *FDH1* was found to be induced by choline or methylamine as well as by methanol. Induction of *FDH1* is not repressed in the presence of glucose when cells grow on methylamine or choline. Glucose or glycerol can repress FDH activity only when the nitrogen source is not methylamine or choline. FDH activity also can be induced by formate, supporting the conclusion that the physiological role of FDH is mainly in the detoxification of formate [Sakai et al., 1997]. Therefore, the activities of both FLD and FDH are strictly regulated in response to methanol or methylamine in order to protect cells from intracellular formaldehyde's toxicity.

1.5   Expression of Foreign Genes

Yeasts, as hosts for the production of eukaryotic proteins, have most of the advantages of *Escherichia coli* in terms of rapid growth rate and ease of genetic manipulation. In addition, they have the ability to perform many important post-translational protein modifications, such as folding, disulfide bridge formation, and

glycosylation characteristic of eukaryotes [Romanos et al., 1992]. Because of the accumulated knowledge available at the time, *S. cerevisiae* (baker's yeast) was the first yeast species to be exploited for this purpose [Hitzeman et al., 1981]. However, the expression level and the quality of the heterologous protein synthesized in *S. cerevisiae* was often limited for the following reasons: (i) a lack of strong, tightly regulated promoters; (ii) poor secretion efficiency, especially of larger (>30 kDa) proteins which remained trapped in the periplasm; (iii) instability of the production strain; and (iv) inability to attain high cell densities in fermenter cultures due to its production of extracellular ethanol [Faber et al., 1995].

The *P. pastoris* heterologous gene expression system has been utilized to produce a variety of intracellular and extracellular proteins of interest [Higgins and Cregg, 1998; Cereghino and Cregg, 2000]. Relative to *S. cerevisiae*, the methylotrophic yeast *P. pastoris* has two key advantages as a host for this purpose [Cregg et al., 1993]. First, *P. pastoris* possesses strong, tightly regulated promoter elements. *P. pastoris* gene expression is based on fusion of heterologous gene sequences to a strong methanol-inducible promoter. The promoter most often used to transcribe foreign proteins is derived from *AOX1*, which encodes a key enzyme of methanol metabolism [Ellis et al., 1985]. With the addition of methanol to the culture medium, this enzyme is rapidly induced to a high level (~30% of total protein) within the cell [Couderc and Baratti, 1980]. This induction is controlled primarily at the transcriptional level. Moreover, *AOX1* synthesis is also repressed in cells grown on glucose and most other carbon sources. Thus, the *AOX1* promoter provides high-level expression that is tightly controlled by a simple change in culture medium. Second, *P. pastoris* does not ferment as does *S. cerevisiae*. The product of fermentation, ethanol, can rapidly reach toxic levels in high-density cultures of *S. cerevisiae*. Moreover, well-developed fermentation technology exists for foreign protein production in *P. pastoris*, thanks to early work on the development of the yeast for SCP production from methanol [Cooney and Levine, 1972; Smith, 1981]. As a result of these and other advantages, *P. pastoris* expression strains characteristically generate high levels of heterologous intracellular and secreted

proteins. In addition, expression strains scale up readily from shake-flask to large-volume, high-density fermenter cultures without loss of yield and are genetically stable [Romanos et al., 1995; Cregg, 1999].

Although the *AOX1* promoter (P_{AOX1}) has been successfully used to control transcription of numerous foreign genes, the promoter has some limitations. 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. Additionally, the storage of large amounts of methanol needed for the growth and induction of P_{AOX1}-controlled expression strains in large-volume fermenter cultures can be a potential fire hazard in certain situations. As an alternative to P_{AOX1}, the *P. pastoris GAP*, *PEX8*, and *YPT1* promoters are available.

P_{GAP} is a strong constitutive promoter derived from the glyceraldehyde-3-phosphate dehydrogenase gene of *P. pastoris*. In glucose- and glycerol-grown cultures, expression levels provided by P_{GAP} are higher than those of the commonly used *AOX1p* in shake-flask methanol-grown cultures [Waterham et al., 1996]. An advantage of using P_{GAP} is that methanol is not required for induction. In addition, since it is constitutive, it is not necessary to shift cultures from one carbon source to another to induce expression. However, because P_{GAP} is constitutive, it is not a good choice for the production of proteins that are toxic to the yeast.

Both *P. pastoris PEX8* and *YPT1* are expressed at a low level on either glucose or methanol medium. The *PEX8* gene encodes a peroxisomal matrix protein that is essential for peroxisome biogenesis [Liu et al., 1995], and the *YPT1* gene encodes a GTPase involved in secretion [Sears et al., 1998]. P_{PEX8} and P_{YPT1} are useful for expressing foreign genes whose products might otherwise challenge the post-translational machinery of the cell, causing a significant proportion of foreign protein to be misfolded, unprocessed, or mislocalized [Thill et al., 1990; Brierley, 1998].

FLD synthesis is independently regulated 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. Herein we have reported the isolation and characterization of the *P. pastoris FLD* gene and its promoter (P_{FLD}). We have shown that FLD expression in response to methanol or methylamine is controlled at the transcriptional level. P_{FLD} can be an alternative to P_{AOX1} in strength in response to either methanol or methylamine induction.

CHAPTER 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 *H. polymorpha* strain used was CBS4732. Bacterial recombinant DNA manipulations were performed in either *E. coli* strain MC1061 or DH5α. 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

approximately 100,000 independent *E. coli* transformants with >90% containing an insert. The average size of insert DNA was approximately 10 kb. Assuming that the size of the *H. polymorpha* genome was 10,000 kb, the library contained approximately 100 genome equivalents of *H. polymorpha* genomic DNA. For expression of bacterial β-lactamase (β-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 *FLD1* plasmid pYG1 as template and primers composed of the following sequences: 5'-CGGGAT-CCGCATGCAGGAATCTCTGGCA-3' and 5'-CGCAATTGTGTGAATATCAAGAA-TTG-3'. The resulting fragment was cut with *Mun*I and *Bam*HI and ligated into *Eco*RI- and *Bgl*II-digested pHW018 and pK321 to create pSS050 and pSS040, respectively (Fig. 2.1). pSS050 and pSS040 contain the identical expression cassettes composed of P_{FLD1}, the β-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 FLD1* Genes

To isolate DNA fragments containing the *P. pastoris* and *H. polymorpha FLD1* 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 approximately 50,000 His⁺ transformants were selected on YND medium agar. Transformants from each library were pooled and approximately 1×10^8 cells 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 DNAs were then transformed into *E. coli* and plasmids contained in the resulting colonies examined. From the *P. pastoris* library
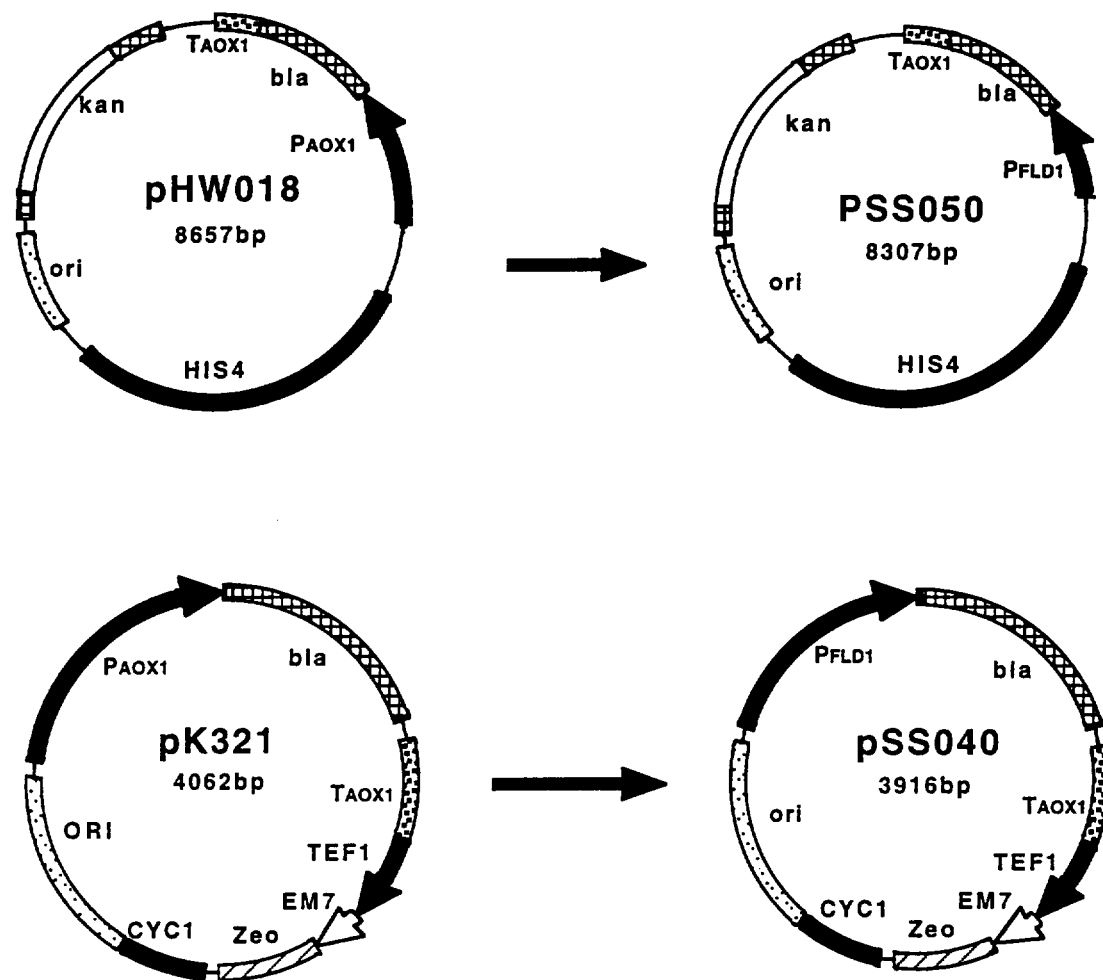
Figure 2.1 Physical map of selected vectors used in this study. [From Shen et al. (1998) *Gene* **216,** 93-102, and used with permission from Elsevier Science.]

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 plasmids 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 $\epsilon = 6.22$ cm^2/nmol for NAD. Alcohol oxidase activities were expressed in μmol/mg/min, and FLD activities were expressed in μmol/mg/min. β-Lactamase activity, expressed as nmol/mg/min, was assayed spectrophotometrically at 569 nm and 30°C in 25 mM Tris-HCl (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'-CACAATGTCTACCGAAGGTC-3' (5' primer) and 5'-CCAGAAAGCGTGTAAG-CATCAG-3' (3' primer).

CHAPTER 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 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 3.1). 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 3.1 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.

25

Table 3.1

**Relative Enzyme Activity Levels in
Methanol-Utilization-Defective Mutants of *P. pastoris***

Strain	% Activity[a]					
	AOX	CAT	FLD	FDH	DAS	DAK
WT (methanol)	100	100	100	100	100	100
WT (glucose)	0	2	1	0	3	53
KM7121 (*aox1 aox2*)	0	100	26	31	ND[b]	88
GS241 (*fld1*)	20	178	0	46	58	64

[a] 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.

Abbreviations are: AOX, alcohol oxidase; CAT, catalase; FLD, formaldehyde dehydrogenase; FDH, formate dehydrogenase; DAS, dihydroxyacetone synthase; DAK, dihydroxyacetone kinase; WT, wild-type.

From Shen et al. (1998) *Gene* **216**, 93-102, and used with permission from Elsevier Science.

Further evidence in support of this is provided below. 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. 3.1). 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 open reading frame (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 reverse transcription-polymerase chain reaction (RT-PCR) method, and the size and sequence of the product was compared to that obtained by
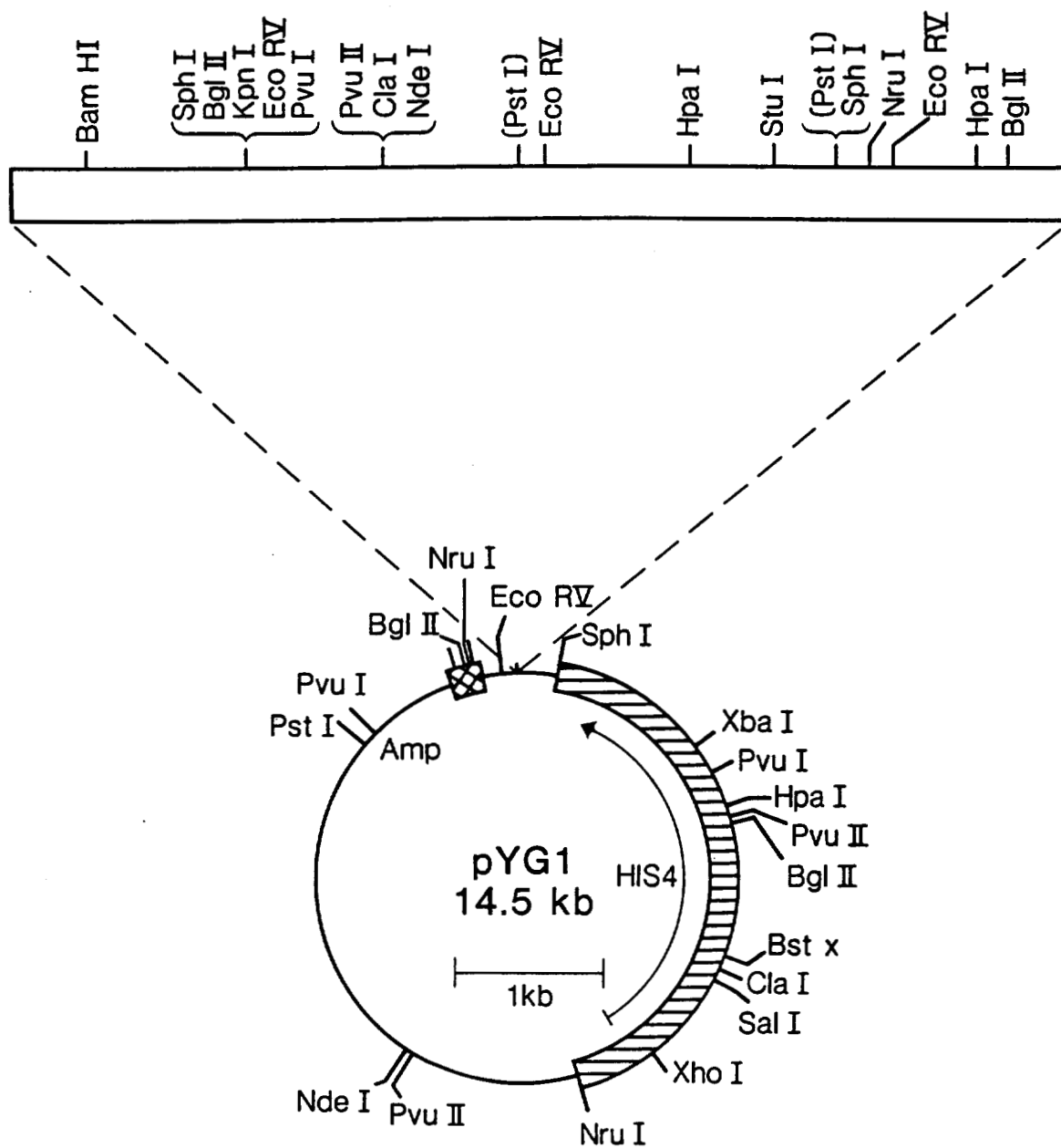
Figure 3.1 Restriction enzyme map of the *FLD1*-containing vector pYG1. [From Shen et al. (1998) *Gene* **216,** 93-102, and used with permission from Elsevier Science.]

PCR of the genomic fragment on plasmid pYG1 (Fig. 3.2). 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 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. 3.3. The ORF is 1,139 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 approximately 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 [Sreekrishna, 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]. *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 with alcohol dehydrogenase III (ADHIII) proteins of higher eukaryotes (65%, human; 63%, horse; 64%, rat) (Fig. 3.4) and a lower but

A

Genomic
DNA

EXON 1   EXON 2   284 bp
PCR product

Spliced
cDNA

170 bp
RT-PCR product

B
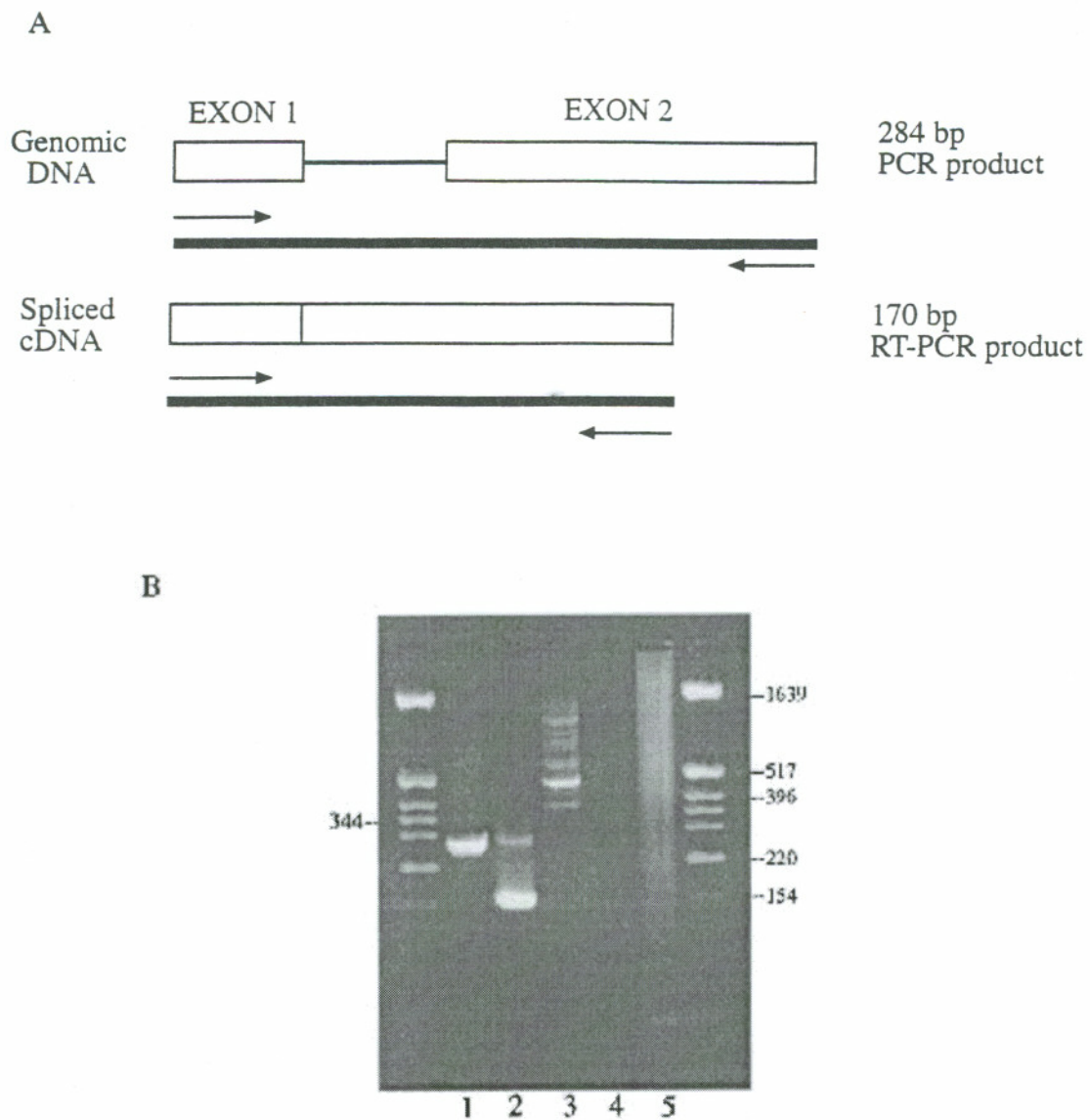
Figure 3.2 Exon analysis of the *FLD* gene. (**A**) Diagram of the expected products from PCR of unspliced (genomic) and spliced (cDNA) DNAs. Locations of the hybridized primers used in the PCR reactions are shown as convergent arrows. (**B**) Electrophoretogram of PCR and RT-PCR reaction products. PCR 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′ primers only; *lane 4*, cDNA template plus 3′ primer only; *lane 5*, both primers without DNA template. Flanking marker bands are denoted in base pairs. [From Shen et al. (1998) *Gene* **216**, 93-102, and used with permission from Elsevier Science.]

```
-597                                          gcatgcaggaatctctggcacggtgctaatggtagtt
-560 atccaacggagctgaggtagtcgatatatctggatatgccgcctataggataaaaacaggagagggtgaaccttgcttat
-480 ggctactagattgttcttgtactctgaattctcattatgggaaactaaactaatctcatctgtgtgttgcagtactattg
-400 aatcgttgtagtatctacctggagggcattccatgaattagtgagataacagagttgggtaactagagagaataatagac
-320 gtatgcatgattactacacaacggatgtcgcactctttccttagttaaaactatcatccaatcacaagatgcgggctgga
-240 aagacttgctcccgaaggataatcttctgcttctatctcccttcctcatatggtttcgcagggctcatgcccccttcttcc
-160 ttcgaactgcccgatgaggaagtccttagcctatcaaagaattcgggaccatcatcgattttttagagccttacctgatcg
 -80 caatcaggatttcactactcatataaatacatcgctcaaagctccaactttgcttgttcatacaattcttgatattcaca
```

```
   1 ATG TCT ACC GAA GGT CAA GTA AGT TCA ATC AAA GTA ATT GTT TGG GAG GGA AGA AGA TTG
   1  M   S   T   E   G   Q
  61 TTT TAT TGC GAA CCT TTC AAT ATC TTA CCC GAC TAA ATA ACC ATT ACA GTG AAT TTT TTA

 121 CTA ACT ATA TAG ATC ATC AAA TGT AAG GCA GCT GTT GCC TGG GAG GCA GGA AAG GAT CTC
   7           '  I   I   K   C   K   A   A   A   V   A   W   E   A   G   K   D   L
 181 TCT ATT GAG GAG ATT GAG GTT CTT CCT CCA AGA GCC CAT GAA GTT AGA GTG AAA GTG GAA
  23  S   I   E   E   I   E   V   L   P   P   R   A   H   E   V   R   V   K   V   E
 241 TTC ACT GGT GTA TGC CAC ACT GAT GCT TAC ACG CTT TCT GGT GCA GAT GCA GAG GGA AGT
  43  F   T   G   V   C   H   T   D   A   Y   T   L   S   G   A   D   A   E   G   S
 301 TTC CCT GTT GTG TTC GGC CAT GAA GGT GCT GGT GTT GTC GAG TCA GTT GGA GAA GGT GTT
  63  F   P   V   V   F   G   H   E   G   A   G   V   V   E   S   V   G   E   G   V
 361 GAG TCC GTG AAG GTT GGG GAT TCT GTA GTG CTT CTG TAC ACT CCT GAG TGC AGA GAG TGC
  83  E   S   V   K   V   G   D   S   V   L   L   Y   T   P   E   C   R   E   C
 421 AAG TTC TGT CTG TCT GGT AAG ACG AAC CTC TGT GGT AAA ATC AGA GCC ACC CAG GGT AAA
 103  K   F   C   L   S   G   K   T   N   L   C   G   K   I   R   A   T   Q   G   K
 481 GGT TTG TTA CCA GAC GGG ACT TCT CGT TTC CGT TGT AAG GGC AAG GAT TTG TTT CAC TAT
 123  G   L   L   P   D   G   T   S   R   F   R   C   K   G   K   D   L   F   H   Y
 541 ATG GGA TGT TCT TCC TTT TCT CAA TAC ACT GTG GTG GCT GAC ATC TCA GTG GTT AAA GTC
 143  M   G   C   S   S   F   S   Q   Y   T   V   V   A   D   I   S   V   V   K   V
 601 CAA GAC GAA GCT CCT AAG GAC AAG ACA TGT CTG TTG GGT TGT GGT GTT ACC ACA GGG TAC
 163  Q   D   E   A   P   K   D   K   T   C   L   L   G   C   G   V   T   T   G   Y
 661 GGT GCT GCT ATC AAC ACT GCT AAG ATC TCT AAG GGT GAC AAG ATC GGT GTG TTT GGT GCT
 183  G   A   A   I   N   T   A   K   I   S   K   G   D   K   I   G   V   F   G   A
 721 GGA TGT ATT GGA TTA TCT GTC ATC CAA GGT GCA GTT TCC AAA GGT GCA AGC GAG ATT ATT
 203  G   C   I   G   L   S   V   I   Q   G   A   V   S   K   G   A   S   E   I   I
 781 GTA ATT GAC ATC AAT GAT TCA AAG AAG GCA TGG GCG GAC CAA TTT GGT GCA ACT AAG TTT
 223  V   I   D   I   N   D   S   K   K   A   W   A   D   Q   F   G   A   T   K   F
 841 GTC AAT CCT ACA ACC TTA CCA GAA GGT ACC AAT ATT GTT GAC TAC TTG ATT GAT ATC ACT
 243  V   N   P   T   T   L   P   E   G   T   N   I   V   D   Y   L   I   D   I   T
 901 GAC GGA GGC TTT GAC TAT ACC TTC GAC TGT ACC GGT AAT GTT CAA GTA ATG AGA AAT GCA
 263  D   G   G   F   D   Y   T   F   D   C   T   G   N   V   Q   V   M   R   N   A
 961 CTT GAA TCT TGC CAC AAG GGT TGG GGT GAG TCG ATC ATC ATC GGT GTC GCT GCT GCT GGT
 283  L   E   S   C   H   K   G   W   G   E   S   I   I   I   G   V   A   A   A   G
1021 AAA GAA ATC TCT ACC CGT CCT TTC CAG TTG GTT ACT GGC AGA GTC TGG AGA GGA TGC GCC
 303  K   E   I   S   T   R   P   F   Q   L   V   T   G   R   V   W   R   G   C   A
1081 TTT GGA GGT ATC AAG GGA CGT ACT CAA ATG CCA TCT TTG GTT CAG GAC TAT CTT GAT GGT
 323  F   G   G   I   K   G   R   T   Q   M   P   S   L   V   Q   D   Y   L   D   G
1141 AAG ATT AAA GTT GAC GAG TTT ATC ACA CAC AGA CAT GAC CTG GAC AAC ATC AAC AAA GCA
 343  K   I   K   V   D   E   F   I   T   H   R   H   D   L   D   N   I   N   K   A
1201 TTT CAT GAC ATG CAT GCT GGA AAC TGT ATT CGT GCT GTG ATT ACT ATG CAC TAA gtacgac
 363  F   H   D   M   H   A   G   N   C   I   R   A   V   I   T   M   H
```

```
1262 gtatgatgaatgaatgagttatgtaaggcccgatctcagctaggacgtttatagacctatgtatatatatgtatgtatac
1342 gtatatacctcaaactcattttatggctataggaaggattgtttttcatcgttatgtccgaagatacatcaatacagcgtt
1422 tcttgatttgtaccaaacactccccaggtagattctccagtctcgctagttaccttcgcggtattgatcgacgctctgtg
1502 aaaaaataaaaaaaaaatgtcaatagtatcgaggttggcttcaaaatcaccagcgttctgtatcgggttccgtttttgag
1582 cttctactttccttatatccatcatacattcctgtgcttgaattcagcagtagaatgtccgatcgatatattgtgttgca
1662 tatcaacaccaccgccaatgagagttctcagcagttcaaacgcgaccccttcagagattattgagctagcatggggttctgc
1742 tagatcctggaaccaactttgaaatagttggtagggaagtgttctggctaaaccattcaacactcctataaccccgctt
1822 tgtaccagcatgacaactcttacatgggaaagtgtcaagaacgctggttcactcaaggatgcgctagaggagcttagcag
1902 atttattgactcaaacttggtcagcaacggcttgtcattcagtttttataactctgaatgcctgggatctccgcttgaaat
1982 tacccaaggagtcgcgtgaaagaagtatagcccttcccgcgtacttggacttacccaagtactttgatcttagaaaagaa
2062 ttctgtagatgggcccaaaaatcatctgcattgactacaaacggtaatcacatgagtttagcttatatggtatctaaact
2142 tgaaacagaagctagtttggttttggacgaggatcc
```

Figure 3.3 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. [From Shen et al. (1998) *Gene* **216**, 93-102, and used with permission from Elsevier Science.]

```
                                                                      50
1.human  ADH III      -AN    EV-------------P--------A--K-----I-IIA-A---
2.horse  ADH III      -A     EV-------------PV----V--A--K-----I-IIA-A---
3.rat    ADH III      AN     -V-R----------P--------A--Q-----I-IIA-A---
4.C.maltosa FDH     --ES-V-KP-T---------A-P----DVT-A--KR----I-LYD-----
5.P.pastoris FDH     MS   TEGQIIKCKAAVAWEAGKDLSIEEIEVLPPRAHEVRVKVEFTGVCH
                                                                     100
1 ----------      P--C---IL-    -----I----    ----TKL-A-    -T-IP--I-Q
2 ----------      P------IL-    -----I----    ----TKL-A-    -T-IP--I-Q
3 ----------      P--C---IL-    -----I----    ----TKL-A-    -T-IP--I-Q
4 --------V-      P--A---IL-    -----I---I    ----TNV-V-    -H-IA-----
5 TDAYTLSGAD      AEGSFPVVFG    HEGAGVVESV    GEGVESVKVG    DSVVLLYTPE
                                                                     150
1 -G-----LN       -----Q---V    ------M---    ----T----T    IL----T-T-
2 -G-----LNP      Q----Q---T    ------M---    ----T----T    IL----T-T-
3 -G-----LN       -----Q---V    ------M---    ----T----P    IL-F--T-T-
4 -G-----K--      ----------    -----VM---    ----T----E    IL-F----T-
5 CRECKFCLSG      KTNLCGKIRA    TQGKGLLPDG    TSRFRCKGKD    LFHYMGCSSF
                                                                     200
1 -E--------      -A-IDPL--L    --V------I    S------V--    --LEP-SVCA
2 -E--------      -A-IDP---L    --V-------    S------V--    --VEP-STCA
3 -E--------      -A-IDPS--L    --V------I    S------V--    --VEP-STCA
4 -Q--------      --AINPK-EF    --A------I    -------TI-    -NVQ---NVA
5 SQYTVVADIS      VVKVQDEAPK    DKTCLLGCGV    TTGYGAAINT    AKISKGDKIG
                                                                     250
1 ---L-GV--A      --M-CKVA--    -R--GV---K    D-F-R-KE--    --ECI--QDF
2 I--L-GV--A      --M-CKVA--    -R--GV---K    D-F-K-KE--    -SECI--QDF
3 ---L-GV--A      --M-CKVA--    -R--G----K    D-F-K-KE--    --ECI--QDF
4 ---G-IV---      ----CAER--    AQ--LV--S-    K-EE-GQKL-    --A-----K-
5 VFGAGCIGLS      VIQGAVSKGA    SEIIVIDIND    SKKAWADQFG    ATKFVNPTTL
                                                                     300
1 SK   P-QEV-     -EM----V--    S-E-I---K-    --A---A---    ---V-VVV--
2 SK   P-QEV-     -EM----V--    S-E-I---K-    --A---A---    ---V-VVV--
3 SK   S-QEV-     -EM----V-F    S-E-I---K-    --S---AA--    ---V-VVV--
4 ----T---K-      -EM----C-F    --------G-    ------A---    ---T-V----
5 PEGTNIVDYL      IDITDGGFDY    TFDCTGNVQV    MRNALESCHK    GWGESIIIGV
                                                                     350
1 --S-E-IA--      --------T-    K-T----W-S    VESV-K--SE    -MSK------
2 --S-E--A--      --------T-    K-T----W-S    VESI-K--SE    -MSK------
3 --S-E-----      --------T-    K-T----W-S    VESV-K--SE    -MSK------
4 ----------      --------T-    K-A----V--    -S-L-G--NN    -----L--E-
5 AAAGKEISTR      PFQLVTGRVW    RGCAFGGIKG    RTQMPSLVQD    YLDGKIKVDE
                                 381
1 -V--NLSF-E      -----EL--S    -KS--T-VKI                     64.97%
2 -V--SLSF-Q      -----EL--A    -KS--T-VKL                     63.81%
3 -V-GNLSF-Q      -----DL--S    --S--T-LKL                     64.34%
4 -----EP-AA      -----EE---    -D-----VDLS                    71.24%
5 FITHRHDLDN      INKAFHDMHA    GNCIRAVITMH
```

Figure 3.4 Alignment of amino acid sequences of subunits of three ADHs—(1) human ADHIII, (2) horse ADHIII, (3) rat ADHIII]—and one FLD—(4) C. maltosa FLD. The predicted amino acid sequence from the cloned gene of P. pastoris and the percentages of amino acid homology to the P. pastoris are shown. The dashes indicate homologous amino acids. The numbering system is applied to the C. maltosa FLD.

but significant identity with other higher eukaryotic ADHs [Giri et al., 1989; Koivusalo et al., 1989; Holmquist and Vallee, 1991]. 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 versus 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 selected times during incubation, samples of extract were removed and assayed for FLD activity. As

shown in Fig. 3.5, *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*I 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. 3.5, 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.

3.4   Analysis of P_{FLD1} and comparison to P_{AOX1}

As described in Chapter 1, 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. 2.1). Both vectors contained identical expression
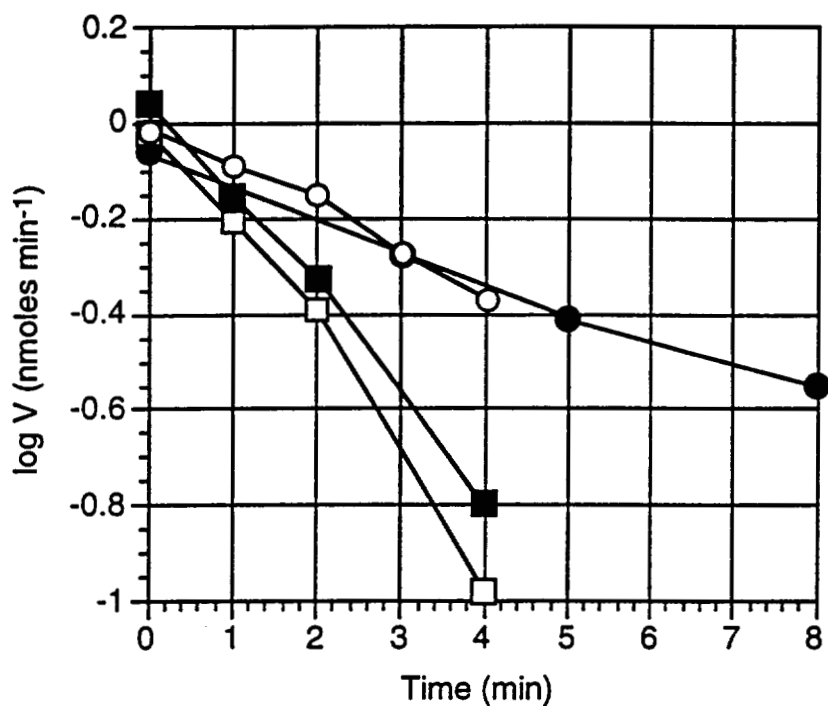
Figure 3.5 Thermal stability of formaldehyde dehydrogenase activities in *P. pastoris* strains transformed with putative *FLD1* genes from *P. pastoris* and *H. polymorpha*. Strains shown are: wild-type *P. pastoris* (■), wild-type *H. polymorpha* (●), *P. pastoris* MS105 (pYG1) (□), and *P. pastoris* MS105 (pYG2) (○). [From Shen et al. (1998) *Gene* **216,** 93-102, and used with permission from Elsevier Science.]

cassettes composed of a 0.6-kb MunI–BamHI fragment with sequences originating from just 5' of the methionine initiator ATG codon of FLD1 fused to the bacterial bla gene encoding β-lac, followed by a fragment containing the AOX1 transcriptional terminator. One vector, pSS040, contained a unique NsiI 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 3.2, this appeared to be true in that the relative levels of β-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: (i) glucose and ammonium sulfate (G/NH$_4^+$), (ii) glucose and methylamine (G/MA), (iii) methanol and ammonium sulfate (M/NH$_4^+$), and (iv) methanol and methylamine (M/MA). As expected, β-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 ten-fold more β-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 SalI 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 β-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 3.2 shows that β-lac activity levels in the pB-P_{FLD1}-bla strain were approximately two-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

Table 3.2

Comparison of β-Lactamase Activity in Extracts of *P. pastoris* Strains Expressing *bla* under Control of P_{FLD} and P_{AOX1}

Strain	Source of:[a]		Enzyme activity[b]			
	C	N	β-lactamase		FLD	
$WT\text{-}P_{FLD1}\text{-}bla$ (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%)
$pB\text{-}P_{FLD1}\text{-}bla$ (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%)
$P_{AOX1}\text{-}bla$	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 ammonium sulfate (NH_4^+) or methylamine (MA) as nitrogen source.

[b] β-Lactamase activities are expressed as nmol/mg/min and, in parentheses, as a percentage of activity seen in the $WT\text{-}P_{FLD1}\text{-}bla$ strain grown on methanol and methylamine. Activities represent the mean of three experiments using two independently transformed strains.

From Shen et al. (1998) *Gene* 216, 93-102, and used with permission from Elsevier Science.

P_{FLD1} fragment, but that sequences that constitutively repress P_{FLD1} by a factor of about two-fold existed somewhere 5' of the P_{FLD1} fragment and were missing from the 0.6-kb fragment.

Lastly, we compared levels of β-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 3.2). Comparable levels of β-lac were present in cells of the *WT-P_{FLD1}-bla* strain grown in either M/NH_4^+ or M/MA media, whereas cells of the *pB-P_{FLD1}-bla* strain contained levels of β-lac that were significantly higher than those in the P_{AOX1}-*bla* strain. Especially noteworthy were the levels of β-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.

CHAPTER 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 [Zwart et al., 1980; Veenhuis et al., 1983]. 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. 2.1). 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

39

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.

LITERATURE CITED

Ahlberg, J., Berkenstam, A., Henell, F., and Glaumann, H. (1985) Degradation of short and long lived proteins in isolated rat liver lysosomes. Effects of pH, temperature, and proteolytic inhibitors. *J. Biol. Chem.* **260**, 5847–5854.

Aitchison, J. D., Nuttley, W. M., Szilard, R. K., Brade, A. M., Glover, J. R., and Rachubinski, R. A. (1992) Peroxisome biogenesis in yeast. *Mol. Microbiol.* **6**, 3455–3460.

Allais, J. J., Louktibi, A., and Baratti, J. (1983) Oxidation of methanol by the yeast *Pichia pastoris*. Purification and properties of the formate dehydrogenase. *Agric. Biol. Chem.* **47**, 1509–1516.

Anthony, C. (1982) *The Biochemistry of Methylotrophs*. Academic Press, New York.

Bennetzen, J. L. and Hall, B. D. (1982) The primary structure of the *Saccharomyces cerevisiae* gene for alcohol dehydrogenase. *J. Biol. Chem.* **257**, 3018–3025.

Bjorklund, S. and Kim, Y.-J. (1996) Mediator of transcriptional regulation. *Trends Biochem. Sci.* **21**, 335–337.

Blaschko, H., Richter, D., and Schlossmann, H. (1937) The oxidation of adrenaline and other amines. *Biochem. J.* **31**, 2187.

Bradford, M. 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.

Brierley, R. A. (1998) Secretion of recombinant human insulin-like growth factor I (IGF-1). *Methods Mol. Biol.* **103**, 149–177.

Burley, S. K. and Roeder, R. G. (1998) TATA box mimicry by TFIID: autoinhibition of pol II transcription. *Cell* **94**, 551–553.

Burnett, K. G. and Felder, M. R. (1978) Genetic regulation of liver alcohol dehydrogenase in *Peromyscus*. *Biochem. Genet.* **16**, 443–454.

41

Cereghino, J. L. and Cregg, J. M. (2000) Heterologous protein expression in the methylotophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* **24**, 45–66.

Chiang, H. L., Schekman, R., and Hamamoto, S. (1996) Selective uptake of cytosolic, peroxisomal, and plasma membrane proteins into the yeast lysosome for degradation. *J. Biol. Chem.* **271**, 9934–9941.

Ciriacy, M. (1975) Genetics of alcohol dehydrogenase in *Sacchromyces cerevisiae*. II. Two loci controlling synthesis of the glucose-repressible ADH II. *Mol. Gen. Genet.* **138**, 157–164.

Colby, J. and Dalton, H. (1978) Resolution of the methane mono-oxygenase of *Methylococcus capsulatus* (Bath) into three components. Purification and properties of component C, a flavoprotein. *Biochem. J.* **171**, 461–468.

Costanzo, G., Di Mauro. E., Negri, R., Pereira, G., and Hollenberg, C. (1995) Multiple overlapping positions of nucleosomes with single *in vivo* rotational setting in the *Hansenula polymorpha* RNA polymerase II MOX promoter. *J. Biol. Chem.* **270**, 11091–11097.

Cooney, C. L. and Levine, D. W. (1972) Microbial utilization of methanol. *Adv. Appl. Microbiol.* **15**, 337–365.

Couderc, R. and Baratti, J. (1980) Oxidation of methanol by the yeast *Pichia pastoris*. Purification and properties of alcohol oxidase. *Agric. Biol. Chem.* **44**, 2279–2289.

Cowell, I. G. (1994) Repression versus activation in the control of gene transcription. *Trends Biochem. Sci.* **19**, 38–42.

Cregg, J. M. (1999) Expression in the methylotrophic yeast *Pichia pastoris*. In: *Nature: The Palette for the Art of Expression*, Chapter 6 (Hoeffler, J. and Fernandez, J., Eds.). Academic Press, San Diego, CA, pp. 157–191.

Cregg, J. M. and Madden, K. R. (1988) Development of the methylotrophic yeast, *Pichia pastoris*, as a host system for the production of foreign proteins. *Dev. Ind. Microbiol.* **29**, 33–41.

Cregg, J. M. and Russell, K. A. (1998) Transformation. *Methods Mol. Biol.* **103**, 27–39.

Cregg, J. M., Barringer, K. J., Hessler, A. Y., and Madden, K. R. (1985) *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.* **5**, 3376–3385.

Cregg, J. M., and Madden, K. R., Barringer, K. J., Thill, G. P., and Stillman, C. A. (1989) Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris. Mol. Cell. Biol.* **9,** 1316–1323.

Cregg, J. M., van der Klei, I. J., Sulter, G. J., Veenhuis, M., and Harder, W. (1990) Peroxisome-deficient mutants of *Hansenula polymorpha. Yeast* **6,** 87–97.

Cregg, J. M., Vedvick, T. S., and Raschke, W. C. (1993) Recent advances in the expression of foreign genes in *Pichia pastoris. Biotechnology (NY)* **11,** 905–910.

Danielsson, O. and Jornvall, H. (1992) "Enzymogenesis": classical liver alcohol dehydrogenase origin from the glutathione-dependent formaldehyde dehydrogenase line. *Proc. Natl. Acad. Sci. U.S.A.* **89,** 9247–9251.

Danielsson, O., Shafqat, J., Estonius, M., and Jornvall, H. (1994) Alcohol dehydrogenase class III contrasted to class I. Characterization of the cyclostome enzyme, the existence of multiple forms as for the human enzyme, and distant cross-species hybridization. *Eur. J. Biochem.* **225,** 1081–1088.

Denis, C. L., Ferguson, J., and Young, E. T. (1983) mRNA level for the fermentative alcohol dehydrogenase of *Saccharomyces cerevisiae* decrease upon growth on a nonfermentable carbon source. *J. Biol. Chem.* **258,** 1165–1171.

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

Douma, A. C., Veenhuis, M., de Koning, W., Evers, M., and Harder, W. (1985) Dihydroxacetone synthase is localized in the peroxisomal matrix of methanol-grown *Hansenula polymorpha. Arch. Microbiol.* **143,** 237–243.

Drewke, C. and Ciriacy, M. (1988) Overexpression, purification and properties of alcohol dehydrogenase IV from *Saccharomyces cerevisiae. Biochim. Biophys. Acta* **950,** 54–60.

Dunn, W. A., Jr. (1990) Studies on the mechanism of autophagy: formation of the autophagic vacuole. *J. Cell Biol.* **110,** 1923–1933.

Dunn, W. A., Jr. (1994) Autophagy and related mechanisms of lysosome-mediated protein degradation. *Trends Cell Biol.* **4,** 139–143.

Ellis, S. B., Brust, P. F., Koutz, P. J., Waters, A. F., Harpold, M. M., and Gingeras, T. R. (1985) Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris. Mol. Cell. Biol.* **5,** 1111–1121.

Erdman, R., Veenhuis, M., Mertens, D., and Kunau, W. H. (1989) Isolation of peroxisome-deficient mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5419–5423.

Estonius, M., Danielsson, O., Hoog, J. O., Persson, H., and Jornvall, H. (1993) Tissue distribution of alcohol and sorbitol dehydrogenase mRNAs. *Adv. Exp. Med. Biol.* **328**, 545–553.

Faber, K. N., Westra, S., Waterham, H. R., Keizer-Gunnink, I., Harder, W., and Veenhuis, G. A. (1995) Foreign gene expression in *Hansenula polymorpha*. A system for the synthesis of small functional peptides. *Appl. Microbiol. Biotechnol.* **45**, 72–79.

Flanagan, P. M., Kelleher, R. J., III, Sayre, M. H., Tschochner, H., and Kornberg, R. D. (1991) A mediator required for activation of RNA polymerase II transcription *in vitro*. *Nature* **350**, 436–438.

Flanagan, P. M., Kelleher, R. J., III, Tschochner, H., Sayre, M. H., and Kornberg, R. D. (1992) Simple derivation of TFIID-dependent RNA polymerase II transcription systems from *Schizosaccharomyces pombe* and other organisms, and factors required for transcriptional activation. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7659–7663.

Frohman, M. A., Dush, M. K., and 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. U.S.A.* **85**, 8998–9002.

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

Godecke, S., Eckart, M., Janowicz, Z. A., and Hollenberg, C. P. (1994) Identification of sequences responsible for transcriptional regulation of the strongly expressed methanol oxidase-encoding gene in *Hansenula polymorpha*. *Gene* **139**, 35–42.

Goodman, J. M. (1985) Dihydroxacetone synthase is an abundant constituent of the methanol-induced peroxisome of *Candida biodinii*. *J. Biol. Chem.* **260**, 7108–7113.

Green, J. and Large, P. J. (1984) Subcellular localization and properties of partially purified dimethylamine and trimethylamine mono-oxygenase activities in *Candida utilis*. *J. Gen. Microbiol.* **130**, 2577–2588.

Gutheil, W. G., Holmquist, B., and Vallee, B. L. (1992) Purification, characterization, and partial sequence of the glutathione-dependent formaldehyde dehydrogenase from *Escherichia coli*: a class III alcohol dehydrogenase. *Biochemistry* **31**, 475–481.

Hanna-Rose, W. and Hansen, U. (1996) Active repression mechanisms of eukaryotic transcription repressors. *Trends Genet.* **12**, 229-234.

Haywood, G. W. and Large, P. J. (1981) Microbial oxidation of amines. Distribution, purification and properties of two primary-amine oxidases from the yeast *Candida boidinii* grown on amines as sole nitrogen source. *Biochem. J.* **199**, 187-201.

Higgins, D. R. and Cregg, J. M. (1998) Introduction to *Pichia pastoris*. *Methods Mol. Biol.* **103**, 1-15.

Hitzeman, R. A., Hagie, F. E., Levine, H. L., Goeddel, D. V., Ammerer, G., and Hall, B. D. (1981) Expression of a human gene for interferon in yeast. *Nature* **293**, 717-722.

Holmquist, B. and 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.

Janowicz, Z. A., Eckart, M. R., Drewke, C., Roggenkamp, R. O., Hollenberg, C. P., Maat, J., Ledeboer, A. M., Visser, C., and Verrips, C. T. (1985) Cloning and characterization of the *DAS* gene encoding the major methanol assimilatory enzyme from the methylotrophic yeast *Hansenula polymorpha*. *Nucleic Acids Res.* **13**, 3043-3062.

Johnson, A. D. (1995) The price of repression. *Cell* **81**, 655-658.

Johnson, M. A., Waterham, H. R., Ksheminska, G. P., Fayura, L. R., Cereghino, J. L., Stasyk, O. V., Veenhuis, M., Kulachkovsky, A. R., Sibirny, A. A., and Cregg, J. M. (1999) Positive selection of novel peroxisome biogenesis-defective mutants of the *Pichia pastoris*. *Genetics* **151**, 1379-1391.

Jornvall, H., Persson, B., and Jeffery, J. (1987) Characteristics of alcohol/polyol dehydrogenases. The zinc-containing long-chain alcohol dehydrogenases. *Eur. J. Biochem.* **167**, 195-201.

Jornvall, H., Danielsson, O., Hjelmqvist, L., Persson, B., and Shafqat, J. (1995) The alcohol dehydrogenase system. *Adv. Exp. Med. Biol.* **372**, 281-294.

Julia, P., Pares, X., and Jornvall, H. (1988) Rat liver alcohol dehydrogenase of class III. Primary structure, functional consequences and relationships to other alcohol dehydrogenases. *Eur. J. Biochem.* **172**, 73-83.

Kaiser, R., Holmquist, B., Hempel, J., Vallee, B. L., and Jornvall, H. (1988) Class III human liver alcohol dehydrogenase: a novel structural type equidistantly related to the class I and class II enzymes. *Biochemistry* **27**, 1132–1140.

Kaiser, R., Holmquist, B., Vallee, B. L., and Jornvall, H. (1989) Characteristics of mammalian class III alcohol dehydrogenases, an enzyme less variable than the traditional liver enzyme of class I. *Biochemistry* **28**, 8432–8438.

Kaiser, R., Fernandez, M. R., Pares, X., and Jornvall, H. (1993) Origin of the human alcohol dehydrogenase system: implications from the structure and properties of the octopus protein. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11222–11226.

Kato, N. (1990) Formaldehyde dehydrogenase from methylotrophic yeasts. *Methods Enzymol.* **188**, 455–459.

Klionsky, D. J. (1998) Nonclassical protein sorting to the yeast vacuole. *J. Biol. Chem.* **273**, 10807–10810.

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

Koutz, P., Davis, G. R., Stillman, C., Barringer, K., Cregg, J., and Thill, G. (1989) Structural comparison of the *Pichia pastoris* alcohol oxidase genes. *Yeast* **5**, 167–177.

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

Lee, J. D. and Komagata, K. (1980) Taxonomic study of methanol assimilating yeasts. *J. Genet. Appl. Microbiol.* **26**, 133.

Liu, H., Tan, X., Veenhuis, M., McCollum, D., and Cregg, J. M. (1992) An efficient screen for peroxisome-deficient mutants of *Pichia pastoris*. *J. Bacteriol.* **174**, 4943–4951.

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: *Methods of Enzymatic Analysis* (Bergmeyer, H. U., Ed.). Academic Press, New York, London, pp. 885–894.

Magonet, E., Hayen, P., Delforge, D., Delaive, E., and Remacle, J. (1992) Importance of the structural zinc atom for the stability of yeast alcohol dehydrogenase. *Biochem. J.* **287**, 361–365.

Meisterernst, M., Stelzer, G., and Roeder, R. G. (1997) Poly(ADP-ribose) polymerase enhances activator-dependent transcription *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2261–2265.

Montavon, P., Felber, J. P., Holmquist, B., and Vallee, B. L. (1989) A human liver alcohol dehydrogenase enzyme-linked immunosorbent assay method specific for class I, II, and III isozymes. *Anal. Biochem.* **176**, 48–56.

Mori, N., Shirakawa, K., Uzura, K. S., Kitamoto, Y., and Ichikawa, Y. (1988) Formation of ethylene glycol and trimethylamine from choline by *Candida tropicalis*. *FEMS Microbiol. Lett.* **51**, 41–44.

Mortimore, G. E., Lardeux, B. R., and Adams, C. E. (1988) Regulation of microautophagy and basal protein turnover in rat liver. Effects of short-term starvation. *J. Biol. Chem.* **263**, 2506–2512.

Nikolov, D. B. and Burley, S. K. (1997) RNA polymerase II transcription initiation: a structural view. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 15–22.

Ogata, K., Nishikawa, H., and Ohsugi, M. (1969) A yeast capable of utilizing methanol. *Agric. Biol. Chem.* **33**, 1519–1520.

Ohi, H., Miura, M., Hiramatsu, R., and Ohmura, T. (1994) The positive and negative *cis*-acting elements for methanol regulation in the *Pichia pastoris AOX2* gene. *Mol. Gen. Genet.* **243**, 489–499.

Paquin, C. E. and Williamson, V. M. (1986) Ty insertions at two loci account for most of the spontaneous antimycin A resistance mutations during growth at 15 degrees C of *Saccharomyces cerevisiae* strains lacking ADH1. *Mol. Cell. Biol.* **6**, 70–79.

Pares, X. and Vallee, B. L. (1981) New human liver alcohol dehydrogenase forms with unique kinetic characteristics. *Biochem. Biophys. Res. Commun.* **98**, 122–130.

Pares, X., Cederlund, E., Moreno, A., Hjelmqvist, L., Farres, J., and Jornvall, H. (1994) Mammalian class IV alcohol dehydrogenase (stomach alcohol dehydrogenase): structure, origin and correlation with enzymology. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1893–1897.

Passoth, V., Schafer, B., Liebel, B., Weierstall, T., and Klinner, U. (1998) Molecular cloning of alcohol dehydrogenase genes of the yeast *Pichia stipitis* and idenification of the fermentative ADH. *Yeast* **14**, 1311–1325.

Patel, R. N., Hou, C. T., and Derelanko, P. (1983) Microbial oxidation of methanol: purification and properties of formaldehyde dehydrogenase from a *Pichia* sp. NRRL-Y-11328. *Arch. Biochem. Biophys.* **221**, 135–142.

Pazin, M. J. and Kadonaga, J. T. (1997) What's up and down with histone deacetylation and transcription? *Cell* **89**, 325–328.

Pereira, G. G. and Hollenberg, C. P. (1996) Conserved regulation of *the Hansenula polymorpha MOX* promoter in *Saccharomyces cerevisiae* reveals insights in the transcriptional activation by Adr1p. *Eur. J. Biochem.* **238**, 181–191.

Ras, J., Van Ophem, P. W., Reijnders, W. N., Van Spanning, R. J., Duine, J. A., Stouthamer, A. H., and Harms, N. (1995) Isolation, sequencing, and mutagenesis of the gene encoding NAD- and glutathione-dependent formaldehyde dehydrogenase (*GD-FALDH*) from *Paracoccus denitrificans*, in which *GD-FALDH* is essential for methylotrophic growth. *J. Bacteriol.* **177**, 247–251.

Reid, M. F. and Fewson, C. A. (1994) Molecular characterization of microbial alcohol dehydrogenases. *Crit. Rev. Microbiol.* **20**, 13–56.

Roeder, R. G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* **21**, 327–335.

Romanos, M. A., Scorer, C. A., and Clare, J. J. (1992) Foreign gene expression in yeast: a review. *Yeast* **8**, 423–488.

Romanos, M. A., Hughes, F. J., Comerford, S. A., and Scorer, C.A. (1995) Production of a phosphorylated GST::HPV-6 E7 fusion protein using a yeast expression vector and glutathione S-transferase fusions. *Gene* **152**, 137–138.

Sakai, Y. and Tani, Y. (1992) Cloning and sequencing of the alcohol oxidase-encoding gene (*AOD1*) from the formaldehyde-producing asporogeneous methylotrophic yeast, *Candida boidinii* S2. *Gene* **114**, 67–73.

Sakai, Y., Murdanoto, A. P., Konishi, T., Iwamatsu, A., and Kato, N. (1997) Regulation of the formate dehydrogenase gene, *FDH1*, in the methylotrophic yeast *Candida boidinii* and growth characteristics of an *FDH1*-disrupted strain on methanol, methylamine, and choline. *J. Bacteriol.* **179**, 4480–4485.

Sakai, Y., Nakagawa, T., Shimase, M., and Kato, N. (1998) Regulation and physiological role of the *DAS1* gene, encoding dihydroxyacetone synthase, in the methylotrophic yeast *Candida boidinii*. *J. Bacteriol.* **180**, 5885–5890.

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.

Sasnauskas, K., Jomantiene, R., Januska, A., Lebediene, E., Lebedys, J., and Janulaitis, A. (1992) Coning and analysis of a *Canadida maltosa* gene which confers resistance to formaldehyde in *Saccharomyces cerevisiae*. *Gene* **122**, 207-211.

Schmitt, M., Brown, T.A., and 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., and Kula, M. R. (1976) Purification and properties of formaldehyde dehydrogenase and formate dehydrogenase from *Candida boidinii*. *Eur. J. Biochem.* **62**, 151-160.

Sears, I. B., O'Connor, J., Rossanese, O. W., and Glick, B. S. (1998) A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*. *Yeast* **14**, 783-790.

Shain, D. H., Salvadore, C., and Denis, C. L. (1992) Evolution of the alcohol dehydrogenase (*ADH*) genes in yeast: characterization of a fourth *ADH* in *Kluyveromyces lactis*. *Mol. Gen. Genet.* **232**, 479-488.

Sharma, C. P., Fox, E. A., Holmquist, B., Jornvall, H., and Vallee, B. L. (1989) cDNA sequence of human class III alcohol dehydrogenase. *Biochem. Biophys. Res. Commun.* **164**, 631-637.

Sibirny, A. A., Ubiyvovk, V. M., Gonchar, M. V., Titorenko, V. I., Voronovsky, A. Y., Kapultsevich, Y. G., and 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.

Smith, S. R. L. (1981) Some aspects of ICI single cell protein process. In: *Microbial Growth on C_1 Compounds: Proceedings of the 3rd International Symposium* (Dalton, H., Ed.). Heyden, London, pp. 342-348.

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

Stewart, P., Kersten, P., Vanden Wymelenberg, A., Gaskell, J., and 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.

Tabak, H. F., Braakman, I., and Distel, B. (1999) Peroxisomes: simple in function but complex in maintenance. *Trends Cell Biol.* **9**, 447–453.

Tan, X., Waterham, H. R., Veenhuis, M., and Cregg, J.M. (1995) The *Hansenula polymorpha PER8* gene encodes a novel peroxisomal integral membrane protein involved in proliferation. *J. Cell Biol.* **128**, 307–319.

Tanaka, A., Osumi, M., and Fukui, S. (1982) Peroxisomes of alkane-grown yeast: fundamental and practical aspects. *Ann. N.Y. Acad. Sci.* **386**, 183–199.

Thill, G. P., Davis, G. R., Stillman, C., Holtz, G., Brierly, R., Engel, M., Buckholtz, R., Kenney, J., Provow, S., Vedvick, T., and Siegel, R. S. (1990) Positive and negative effects of multicopy integrated expression vectors on protein expression in *Pichia pastoris*. In: *Proceedings of the Sixth International Symposium on the Genetics of Industrial Microorganisms*, Vol. 2 (Heslot, H., Davies, J., Florent, J., Bobichon, L., Durand, G., and Penasse, L., Eds.). Société Française de Microbiologie, Paris, pp. 477–490.

Titorenko, V. I., Keizer, I., Harder, W., and Veenhuis, M. (1995) Isolation and characterization of mutants impaired in the selective degradation of peroxisomes in the yeast *Hansenula polymorpha*. *J. Bacteriol.* **177**, 357–363.

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

Tuttle, D. L. and Dunn, W. A., Jr. (1995) Divergent modes of autophagy in the methylotrophic yeast *Pichia pastoris*. *J. Cell Sci.* **108**, 25–35.

Tuttle, D. L., Lewin, A. S., and Dunn, W. A., Jr. (1993) Selective autophagy of peroxisomes in methylotrophic yeasts. *Eur. J. Cell Biol.* **60**, 283–290.

Vallari, R. C., Cook, W. J., Audino, D. C., Morgan, M. J., Jensen, D. E., Laudano, A. P., and Denis, C. L. (1992) Glucose repression of the yeast *ADH2* gene occurs through multiple mechanisms, including control of the protein synthesis of its transcriptional activator, ADR1. *Mol. Cell. Biol.* **12**, 1663–1673.

Vallee, B. L. and Bazzone, T. J. (1983) Isozymes of human liver alcohol dehydrogenase. *Isozymes Curr. Top. Biol. Med. Res.* **8**, 219–244.

van Dijken, J. P. (1976) *Oxidation of methanol by* Hansenula polymorpha. *Purification and kinetic properties of methanol oxidase*. Ph.D. Dissertation, University of Groningen, The Netherlands.

van Dijken, J. P. and Bos, P. (1981) Utilization of amines by yeasts. *Arch. Microbiol.* **128**, 320–324.

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

van der Klei, I. J., Bystrykh, L. V., and Harder, W. (1990) Alcohol oxidase from *Hansenula polymorpha* CBS 4732. *Methods Enzymol.* **188**, 420–427.

van der Klei, I. J., Harder, W., and Veenhuis, M. (1991) Biosynthesis and assembly of alcohol oxidase, a peroxisomal matrix protein in methylotrophic yeasts: a review. *Yeast* **7**, 195–209.

van der Walt, J. P. (1962) Utilization of ethylamine by yeasts. *Antonie van Leeuwenhoek J. Microbiol. Serol.* **28**, 91–96.

Veenhuis, M. and Harder, W. (1991) Microbodies. In *The Yeasts*, Vol. 4, 2nd Edn. (Rose, A. H. and Harrison, J. S., Eds.). Academic Press, London, pp. 601–653.

Veenhuis, M., Zwart, K., and Harder, W. (1978) Degradation of peroxisomes after transfer of methanol-grown *Hansenula polymorpha* into glucose-containing media. *FEMS Microbiol. Lett.* **4**, 283–286.

Veenhuis, M., Zwart, K., and Harder, W. (1981) Biogenesis and turnover of peroxisomes involved in the concurrent oxidation of methanol and methylamine in *Hansenula polymorpha*. *Arch. Microbiol.* **129**, 35–41.

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

Verdier, J. M. (1990) Regulatory DNA-binding proteins in yeast: an overview. *Yeast* **6**, 271–297.

Waites, M. J. and Quale, 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., Keizer-Gunnink, I., Goodman, J. M., Harder, W., and Veenhuis, M. (1992) Development of multi-purpose peroxisomes in *Candida boidinii* grown in oleic acid-methanol limited continuous cultures. *J. Bacteriol.* **174**, 4057-4063.

Waterham, H. R., de Vries, Y., Russell, K. A., Xie, W., Veenhuis, M., and Cregg, J. M. (1996) The *Pichia pastoris PER6* gene product is a peroxisomal integral membrane protein essential for peroxisome biogenesis and has sequence similarity to the Zellweger syndrome protein PAF-1. *Mol. Cell. Biol.* **16**, 2527-2536.

Waterham, H. R., Digan, M. E., Koutz, P. J., Lair, S. V., and 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.

Wehner, E. P., Rao, E., and Brendel, M. (1993) Molecular structure and genetic regulation of *SFA*, a gene responsible for resistance to formaldehyde in *Saccharomyces cerevisiae*, and characterization of its protein product. *Mol. Gen. Genet.* **237**, 351-358.

Williamson, V. M. and Paquin, C. E. (1987) Homology of *Saccharomyces cerevisiae ADH4* to an iron-activated alcohol dehydrogenase from *Zymomonas mobilis. Mol. Gen. Genet.* **209**, 374-381.

Young, E. T. and Pilgrim, D. (1985) Isolation and DNA sequence of *ADH3*, a nuclear gene encoding the mitochondrial isozyme of alcohol dehydrogenase in *Saccharomyces cerevisiae. Mol. Cell. Biol.* **5**, 3024-3034.

Yuan, W., Tuttle, D. L., Shi, Y. J., Ralph, G. S., and Dunn, W. A., Jr. (1997) Glucose-induced microautophagy in *Pichia pastoris* requires the α-subunit of phosphofructokinase. *J. Cell Sci.* **110**, 1935-1945.

Yuan, W., Stromhaug, P. E., and Dunn, W. A., Jr. (1999) Glucose-induced autophagy of peroxisomes in *Pichia pastoris* requires a unique E1-like protein. *Mol. Biol. Cell* **10**, 1353-1366.

Zwart, K. B. and Harder, W. (1983) Regulation of the metabolism of some alkylated amines in the yeasts *Candida utilis* and *Hansenula polymorpha. J. Gen. Microbiol.* **129**, 3157-3169.

Zwart, K., Veenhuis, M., van Dijken, J. P., and 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.

BIOGRAPHICAL SKETCH

Shigang Shen was born on April 28, 1965, in Tianjin, the People's Republic of China. In 1984, he began his college education in the Department of Biology, Nankai University of China, graduating in three years. In 1987, he entered the Tianjin Medical College; he received his M.D. degree in 1992. He then worked in the Heart Surgery Department of the Second Affiliated Hospital of Tianjin Medical College as a heart surgeon.

At the end of 1994, he came to the United States to reunite with his wife who was a Ph.D student in the Department of Biochemistry and Molecular Biology of the Oregon Graduate Institute of Science and Technology. Early in 1996, he joined the laboratory of Dr. James Cregg as a Ph.D. candidate at the same institution. He passed the department Qualifying Examination after only six months of study. During his four years of work in Dr. Cregg's lab, he completed all course work towards a Ph.D. degree. However, his wife, who completing her Ph.D. at in 1999, was offered an excellent position in the Boston area. Therefore, in order to keep his family together, he decided to conclude his research with a M.S. degree.

Publications:

Cregg, J. M., Shen, S., Johnson, M., and Waterham, H. R. (1998) Classical genetic manipulation. In: *Methods in Molecular Biology: Pichia Protocols*, Chapter 3 (Higgins, D. R. and Cregg, J. M., Eds.), Humana Press, Totowa, NJ, pp. 17–27.

Shen, S., Sulter, G., Jeffries, T. W., and Cregg, J. M. (1998) A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **216**, 93–102.