

**REGULATION OF *HYDRATASE-DEHYDROGENASE-EPIMERASE***

UPSTREAM SEQUENCES  
INVOLVED IN  
REGULATING THE  
*CANDIDA TROPICALIS* GENE ENCODING  
PEROXISOMAL TRIFUNCTIONAL ENZYME

By

JAMES ANDREW SLOOTS

A Thesis

Submitted to the School of Graduate Studies  
in Partial Fulfilment of the Requirements  
for the Degree  
Master of Science

McMaster University

(c) Copyright by James Andrew Sloots, March 1993

MASTER OF SCIENCE (1993)

(Biochemistry)

McMASTER UNIVERSITY

Hamilton, Ontario

TITLE: Upstream sequences involved in regulating the  
*Candida tropicalis* gene encoding peroxisomal  
trifunctional enzyme.

AUTHOR: James Andrew Sloots, B.Sc. (McMaster University)

SUPERVISOR: Dr. R. A. Rachubinski

NUMBER OF PAGES: viii, 105

## ABSTRACT

We have investigated the expression of the genes *hydratase-dehydrogenase-epimerase (HDE)*, *acyl-CoA oxidase (AOX)* and *catalase (CATL)* of the diploid yeast *Candida tropicalis*. These genes encode enzymes which are localized to the peroxisome. Expression of each gene was monitored by immunoblot analysis of yeast lysates using antibodies directed against each protein. We demonstrate that carbon sources influence expression of these genes, and do so in a coordinate fashion.

We expressed *C. tropicalis HDE* in *Saccharomyces cerevisiae* and demonstrate that this trifunctional enzyme can be regulated by *S. cerevisiae* in a fashion that closely resembles that of *C. tropicalis*.

Expression of constructs containing deletions in the upstream region of the *HDE* gene allowed us to localize regions responsible for regulating the expression of this gene. Regions were identified that are responsible for both repression by glucose and induction by oleic acid. A glucose-responsive region lies between nucleotides -466 and -334. An oleic acid-responsive region lies between nucleotides -333 and -281. An additional region controlling derepression by nonfermentable carbon sources is located downstream of nucleotide -281.

Comparison of the upstream nucleotide sequences of *HDE*, *AOX* and *CATL* both to each other, and to upstream regions of other oleic acid-responsive genes of *C. tropicalis* has identified possible consensus nucleotide sequences for glucose- and oleic acid-responsive upstream elements.

Since the regulation of the *HDE* gene in *S. cerevisiae* closely resembles that of *C. tropicalis*, this implies that similar mechanisms of transcriptional control operate in both yeasts.

## **ACKNOWLEDGEMENTS**

I want to recognize all those who made my years at McMaster not only educational, but also enjoyable. Those associated with my post-graduate experience, namely my fellow lab mates, John Aitchison, Andrea Bodnar, Tony Brade, John Glover, Wayne Murray, Bill Nuttley, and Baowei Zhang deserve special mention for their friendship and assistance. I am also indebted to both Shirley Jones and Jill Regoeczi for their excellent technical and moral support. I want to especially acknowledge the encouragement and support given to me by my parents, siblings and friends during times of both advancement and disappointment.

I would like to express my gratitude to my supervisor Rick Rachubinski for his patient direction, support and guidance throughout the past years. I also extend my appreciation to the committee members John Capone, Karl Freeman, Calvin Harley and Gerard Wright for taking time out from their busy schedules to assist in my training. My special thanks go to Wayne Murray, for his charitable friendship, assistance, and advice; and to Kirsten Helder, the kindred spirit who has opened and enriched my life.

The race is not to the swift  
or the battle to the strong  
nor does food come to the wise  
or wealth to the brilliant  
or favour to the learned;  
but time and chance happen to them all.

Ecclesiastes

## TABLE OF CONTENTS

	Page Number
Title Page . . . . .	i
Descriptive Note . . . . .	ii
Abstract . . . . .	iii
Acknowledgements . . . . .	v
Table of Contents . . . . .	vii
List of Illustrations, Tables, Figures . . . . .	ix
1) INTRODUCTION . . . . .	1
1.1) Peroxisomes and import . . . . .	1
1.2) Peroxisomal biogenesis . . . . .	2
1.3) $\beta$ -oxidation . . . . .	3
1.4) Prokaryotic gene regulation . . . . .	7
1.5) Eukaryotic gene regulation . . . . .	7
1.6) Regulatory schemes . . . . .	9
1.7) Gene regulation in yeast . . . . .	12
1.8) Techniques for transcription analysis . . . . .	14
2) MATERIALS AND METHODS . . . . .	16
2.1) List of Abbreviations . . . . .	16
2.2) Materials and Suppliers . . . . .	17
2.2.1) Common reagents . . . . .	17
2.2.2) Radiochemicals . . . . .	20
2.2.3) Enzymes . . . . .	20
2.3) Methods . . . . .	22
2.3.1) Nucleic Acids . . . . .	22
2.3.1.1) Plasmid Isolation . . . . .	22
2.3.1.2) DNA Modification Reactions . . . . .	24
2.3.1.3) Gel Electrophoresis of DNA . . . . .	25
2.3.1.4) Production of DNA with blunt ends . . . . .	26
2.3.1.4.1) Using Klenow . . . . .	27
2.3.1.4.2) Using T4 DNA Polymerase . . . . .	27
2.3.1.5) Random Primer Labelling of DNA . . . . .	28
2.3.1.6) Gel purification of DNA . . . . .	29
2.3.1.6.1) From agarose gels . . . . .	29
2.3.1.6.2) Over G-50 Sephadex spin columns . . . . .	29
2.3.1.7) Ligation of DNA Fragments . . . . .	30
2.3.1.8) Generation of HDE sequencing constructs . . . . .	31
2.3.1.8.1) Exo III generation of constructs for use in sequencing <i>HDE</i> in the 3'→5' direction . . . . .	31
2.3.1.8.2) Completion of 5'→3' sequencing of the <i>HDE</i> gene . . . . .	33
2.3.1.9) DNA Sequencing Reactions . . . . .	34
2.3.1.10) Sequencing Gels . . . . .	35
2.3.1.11) Nucleic Acids Quantification . . . . .	36
2.3.1.12) Colony Hybridization . . . . .	39



2.3.1.13)	RNA Isolation . . . . .	40
2.3.1.14)	Primer Extension . . . . .	41
2.3.2)	Proteins . . . . .	42
2.3.2.1)	Yeast Lysate Preparation . . . . .	42
2.3.2.2)	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) . . . . .	43
2.3.2.3)	Immunoblotting . . . . .	44
2.3.3)	Vectors and Constructs . . . . .	45
2.3.3.1)	Subcloning of <i>HDE</i> into YCp50 for Pilot Experiments . . . . .	45
2.3.3.2)	Subcloning of <i>HDE</i> into YCp50 for Bal31 Deletions . . . . .	46
2.3.3.3)	Bal 31 Deletion Reactions . . . . .	46
2.3.3.4)	Production of Positive and Negative Control Constructs . . . . .	48
2.3.4)	Strains and Media . . . . .	49
2.3.5)	Transformation Procedures . . . . .	49
2.3.5.1)	<i>S. cerevisiae</i> DL-1 . . . . .	49
2.3.5.2)	<i>Escherichia coli</i> DH5- $\alpha$ . . . . .	50
2.3.6)	Densitometric scanning . . . . .	50
3)	RESULTS AND DISCUSSION . . . . .	52
3.1)	Expression of genes encoding <i>HDE</i> , acyl-CoA oxidase and catalase in <i>Candida tropicalis</i> . . . . .	52
3.2)	Expression of <i>HDE</i> in <i>S. cerevisiae</i> . . . . .	62
3.3)	Primer extension of <i>HDE</i> mRNA . . . . .	67
3.4)	Sequencing of the <i>HDE</i> gene . . . . .	69
3.5)	Expression of <i>HDE</i> deletion constructs in <i>S. cerevisiae</i> . . . . .	73
3.6)	Future studies . . . . .	93
4)	REFERENCES . . . . .	95

## LIST OF FIGURES AND TABLES

### FIGURES

	Page Number
Figure 1. $\beta$ -oxidation pathway . . . . .	5
Figure 2. Effects of carbon source on <i>HDE</i> , AOX and <i>CATL</i> expression . . . . .	54
Figure 3. Plasmid YCp50H2 . . . . .	64
Figure 4. Expression and induction of <i>HDE</i> in <i>S. cerevisiae</i> . . . . .	66
Figure 5. Mapping of <i>HDE</i> transcription start point . .	68
Figure 6. Sequencing strategy of the gene encoding <i>HDE</i> . . . . .	70
Figure 7. Sequence of the <i>HDE</i> gene . . . . .	72
Figure 8. Plasmid YCp505H . . . . .	74
Figure 9. Bal 31 generated 5'-deletion constructs of <i>HDE</i> . . . . .	78
Figure 10. Expression of <i>HDE</i> 5'-deletion constructs . .	81
Figure 11. Expression of <i>HDE</i> 5'-deletion constructs 5WT, 5P and 5T in various carbon sources . . . . .	90

### TABLES

Table 1. Fluorescence of nucleic acids. . . . .	38
Table 2. Densitometric signals of bands observed in Figure 2. . . . .	55
Table 3. Plasmid loss of YCp50 based plasmids. . . . .	76
Table 4. Densitometric signals of bands observed in Figure 10. . . . .	82
Table 5. Conserved upstream elements found in genes encoding peroxisomal proteins. . . . .	86
Table 6. Densitometric signals of bands observed in Figure 11. . . . .	91

## INTRODUCTION

### 1.1) *Peroxisomes and import*

Peroxisomes, which are organelles found in virtually all eukaryotic cells, are characterized as containing a hydrogen peroxide-producing oxidase and a catalase (de Duve and Baudhuin, 1966). They range in size from 0.5-1.5  $\mu\text{m}$  (Tolbert, 1981) and are bounded by a single unit membrane, which allows for the formation of a microenvironment ideally suited for the metabolism that occurs inside (Borst, 1989). Free radicals of oxygen, such as the superoxide  $\text{O}_2^-$ , which can form the extremely reactive hydroxyl ( $\cdot\text{OH}$ ) radical, are produced in the peroxisome under oxidative stress. The encapsulation of these oxygen radical-producing enzymes in the peroxisome protects the rest of the cell from these harmful agents (del Rio et al., 1990).

Peroxisomal proteins are genetically encoded in the nucleus and synthesized on free polysomes in the cytosol. They are posttranslationally imported into the peroxisome in their mature form. Import of proteins into peroxisomes is dependent on the peroxisome's stage of development and energy state (van der Klei et al., 1991).

A general characteristic of proteins that are imported into unique subcellular compartments is the presence of targeting motifs recognized by the organellar import

machinery. One motif sufficient for targeting mammalian proteins to peroxisomes is the C-terminal tripeptide SKL or conservative variations thereof (Gould *et al.*, 1990). A chemically similar terminal tripeptide AKI and variations thereof are sufficient for targeting trifunctional enzyme (HDE) to the peroxisomes of both *Candida albicans* and *Saccharomyces cerevisiae* (Aitchison *et al.*, 1991b). This is by no means an exhaustive list of peroxisomal targeting signals, since many imported peroxisomal proteins do not exhibit these C-terminal sequences.

A difference in pH of 1.1-1.3 units below that of the cytosol has been inferred to exist in the peroxisome, demonstrating that the membrane must be impervious to free ion flow and that if energy is required for transport, it would be required to overcome such electrochemical gradients (Douma *et al.*, 1990).

### 1.2) *Peroxisomal biogenesis*

The yeast *Candida tropicalis* has been useful in the study of peroxisomes, especially peroxisome biogenesis, because of its ability to grow on alkane carbon sources (C<sub>10</sub>-C<sub>13</sub>) (Tanaka *et al.*, 1982). Growth on alkane sources or fatty acids causes the peroxisomes to proliferate by what is believed to be fission of pre-existing peroxisomes (Lazarow and Fujiki, 1985). Not only do the peroxisomes divide, but they also increase in size (Veenhuis and Goodman, 1990). When

cells containing many peroxisomes are placed in media containing glucose or ethanol, and so no longer need to enlist the metabolic functions of the peroxisomes, these organelles are quickly degraded by an autophagic process involving fusion with vacuolar organelles (van der Klei *et al.*, 1991). Interestingly, this active removal of peroxisomes is not observed when cells are re-cultured in media containing other carbon sources. Rather, a simple dilution of peroxisomes occurs as the cells divide (Veenhuis *et al.*, 1983). Concomitant with these changes in peroxisome number, one also sees coordinate variations in the amounts and hence activities of the  $\beta$ -oxidation enzymes (Dommes *et al.*, 1983). These variations are at least partially due to changes in the level of transcription (Nuttley *et al.*, 1988); and for activities of peroxisomal catalase, they are dependent on concentrations of glucose, heme, and oxygen tension (Skoneczny *et al.*, 1988). Our laboratory is interested in elucidating the genetic elements responsible for regulating the  $\beta$ -oxidation enzymes at the level of transcription.

### 1.3) $\beta$ -oxidation

Fatty acids play an important role in the normal day-to-day functioning of cells. They are integral components of membrane lipids and triacylglycerols. In mammalian cells, the breakdown of fatty acids can be performed by the  $\beta$ -oxidation systems of both peroxisomes and mitochondria. These two

pathways vary in both substrate specificity and the mechanism of fatty acid oxidation (Fahimi and Sies, 1987). Yeasts, on the other hand, contain only one contingent of  $\beta$ -oxidation enzymes, which are found solely in the peroxisome (Kawamoto *et al.*, 1978).

$\beta$ -oxidation processes the carboxyl terminus of fatty acids two carbon units at a time, producing acetyl-CoA molecules (see Fig. 1). In the yeast peroxisome, this process begins with the activation of a fatty acid by an ATP-dependant acyl-CoA synthetase to produce an acyl-CoA. Acyl-CoA is subsequently oxidized by acyl-CoA oxidase with the assistance of the coenzyme FAD, to *trans*-2-enoyl-CoA by transferring two electrons from acyl-CoA to molecular oxygen to produce hydrogen peroxide. In peroxisomes, unlike mitochondria, there is no direct coupling to an electron transport chain, so that the chemical energy of this step is released as heat. The trifunctional enzyme hydratase-dehydrogenase-epimerase (HDE) (Moreno de la Garza *et al.*, 1985) stereospecifically adds H<sub>2</sub>O across the double bond to produce L-3-hydroxylacyl-CoA. The epimerase function of HDE, when required, converts D-3-hydroxylacyl-CoA, the hydration product of *cis*-unsaturated fatty acids, to the metabolizable L form (Bremer and Osmundesen, 1984). L-3-hydroxylacyl-CoA is subsequently dehydrogenated by the dehydrogenase function of HDE, with the assistance of the cofactor NAD<sup>+</sup>, to 3-ketoacyl-CoA. The energy of this step is conserved in the cofactor NADH

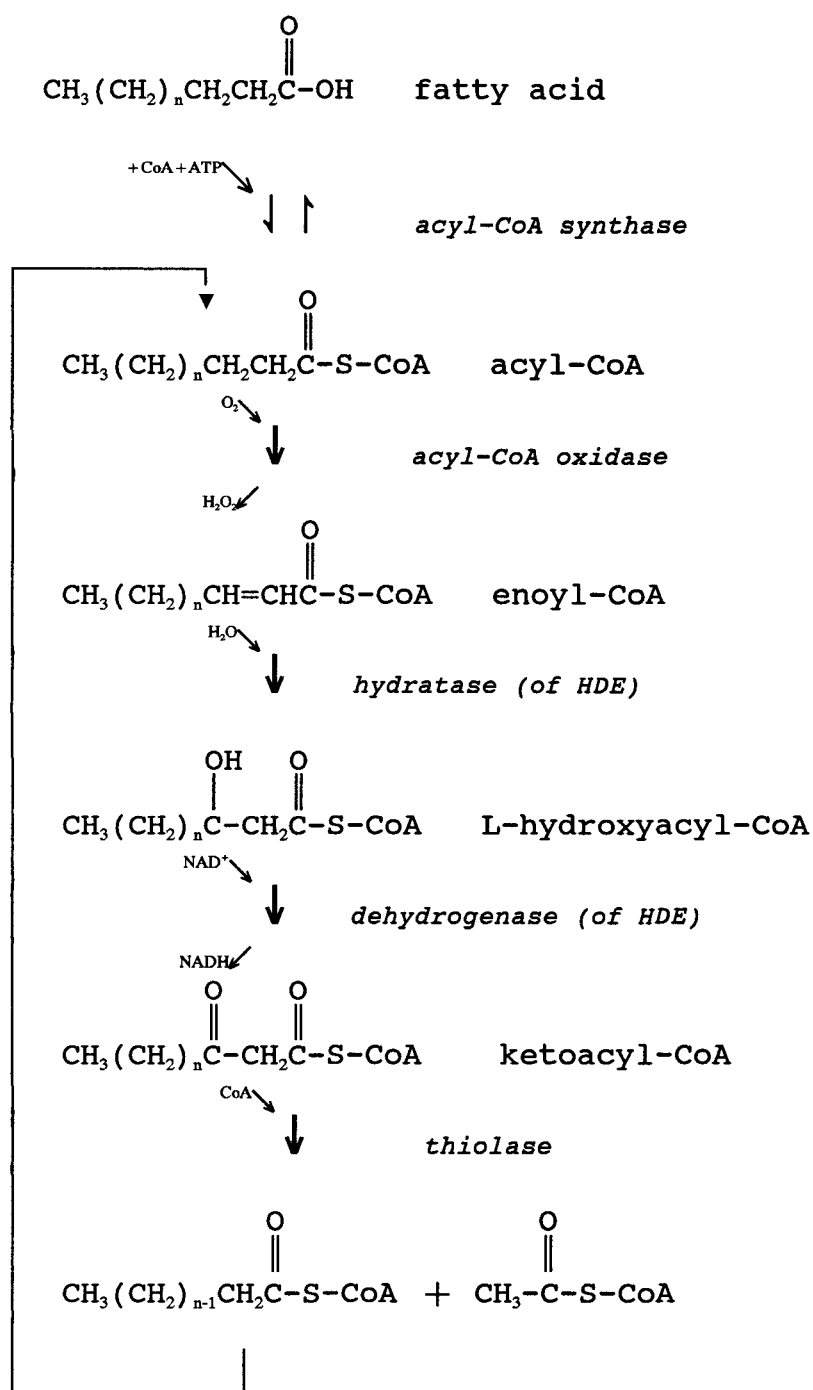
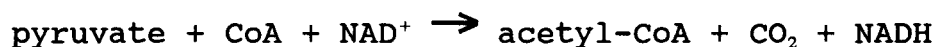


Figure 1.  $\beta$ -Oxidation of a Saturated Fatty Acid. Sequence of reactions performed on a saturated fatty acid carried out by the  $\beta$ -oxidation system of peroxisomes of yeasts. Each cycle reduces the carbon content of the fatty acid by two carbons by producing an acetyl unit covalently bonded to coenzyme A.

and shuttled to the mitochondria, where the reduced nucleotide enters the oxidative-phosphorylation pathway, resulting in ATP production. 3-ketoacyl-CoA is then cleaved by thiolase, upon the addition of CoA, to produce acetyl-CoA and the newly formed acyl-CoA, which re-enters the cycle (Fahimi and Sies, 1987) two carbon units shorter in length.

In order for *C. tropicalis* to utilize fatty acids as a sole carbon and energy source, the acetyl-CoA produced by the peroxisome is transported to the mitochondria for further oxidation by the tricarboxylic acid cycle (Tolbert, 1981) to produce ATP. This transport is achieved by an acetylcarnitine transferase enzyme (recycling the CoA molecule for further reaction with fatty acids) producing acetylcarnitine which is shuttled to the mitochondria. *C. tropicalis* also requires a functional glyoxylate pathway, which is partially resident in the peroxisome as the enzymes malate synthetase and isocitrate lyase. This pathway, which does not exist in animal cells, allows cells the versatility of producing all the carbon compounds they need solely from fatty acids. This is possible because the glyoxylate pathway bypasses the irreversible decarboxylation reaction:



and so allows these cells the anabolic ability to use acetyl-CoA molecules for net synthesis of glucose (gluconeogenesis). This gluconeogenic pathway converts acetyl-CoA into succinate molecules, which can then enter the tricarboxylic cycle



(Kispal and Srere, 1991; Kindl, 1984).

#### 1.4) *Prokaryotic gene regulation*

Prokaryotes regulate message levels by the ability of RNA polymerase to bind with two promoter elements, namely the Pribnow box at -10 nucleotides (nt) and an upstream consensus at -35 nt from the transcription start site (Schibler and Sierra, 1987). An element's level of conformity to a consensus sequence that is known to bind the polymerase tightly influences the affinity of the polymerase for that site and, consequently, the level of message produced. RNA polymerase binds directly to the promoter elements along with environmentally controlled sigma factors, which are required for chain elongation, promoter recognition, and DNA strand separation (Lillie and Green, 1989). When the environmental cue is removed, transcription stops, and the fast RNA turnover rate removes the message from the cell (Helmann and Chamberlin, 1988).

#### 1.5) *Eukaryotic gene regulation*

Eukaryotes, unlike prokaryotes, have three RNA polymerases, all of which recognize a protein-DNA complex rather than just promoter sequences alone. RNA polymerase I and III synthesize structural RNA's, i.e. rRNAs and tRNAs, respectively. RNA polymerase II is responsible for the production of mRNAs. These polymerases require a host of

specific and ubiquitous transcription factors in order to initiate transcription efficiently (Newport and Forbes, 1987; Schibler et Sierra, 1987). Some of these factors bind directly to promoter sequences, e.g. TFIID to the TATA box in mammalian and yeast cells (Davidson et al., 1983; Sawadogo and Roeder, 1985; Horikoshi et al., 1989), while others bind to upstream activator/enhancer sequences, e.g. GAL4. All these factors serve to recruit RNA polymerase by some active domain of their surfaces.

Yet not all eukaryotic promoters seem to function in a similar fashion. The mechanism by which DNA binding proteins communicate is not conserved between the 3 different classes of RNA (Schreck et al., 1989). Variation also exists across species. For example, yeast promoters have a slightly different set-up than their mammalian counterparts as far as initiation is concerned. In mammalian cells, initiation occurs 30 nt downstream of the TATA box element, the site recognized by TFIID. In yeast, the initiation site is not as close to the TATA element, nor as rigidly fixed, occurring at a distance between 60-120 nt downstream of the element. Yet the TFIID molecule itself cannot specify the distance to the initiation site, because the yeast TFIID binds to mammalian promoters and initiates transcription not 60-120 nt downstream as occurs in yeast but at the normal mammalian distance of 30 nt (Horikoshi et al., 1989). Rather, initiation sites in yeast are determined by conformity to consensus sequences and

not by a type of measuring process that fixes the initiation site a certain distance away from the TATA box (the strategy seen in mammalian cells). Yet there is still remarkable conservation of gene regulation mechanisms in cells as diverse as yeast and those of mammalian origin, since many regulatory proteins are functional even across kingdoms (Kakidani and Ptashne, 1988). Elimination of the TATA element in mammalian cells produces heterogenous mRNA initiation (Parker and Topol, 1984). Therefore, the TATA box is not absolutely necessary for transcription, although it is important in establishing polarity to the promoter (Ow et al., 1987). Another common promoter element seen in many housekeeping genes is the poly(dA-dT) tract, which, through possible exclusion of nucleosomes (Struhl, 1987) or recruitment of general transcription factors (Lue et al., 1989), is responsible for constitutive expression. Promoters of housekeeping genes often lack TATA box elements, and consequently have heterogenous initiation sites (Muller et al., 1988).

#### 1.6) *Regulatory schemes*

Increases in enzymatic activity can be mediated by many cellular events, including control of protein activity via phosphorylation; degradation of message and/or protein; translational control; nuclear transport; mRNA processing and transcriptional control (Darnell, 1988). The last, which is effected at the level of the promoter, is one of the more

important levels of control in both prokaryotes and eukaryotes.

A eukaryotic promoter element consists of a group of regulatory DNA sequences clustered around the transcription initiation site. These sequences are sequentially bound by transcription factors, which directly influence the initiation rate of RNA synthesis. Enhancers are segments of DNA located far upstream (or occasionally downstream) of promoters, to which regulatory proteins can bind. Their potency can be exponentially increased by concatenation (Axelrod *et al.*, 1990). When bound by activator proteins, they exert their function on promoters from large distances, regardless of orientation, by what is envisioned as a looping out of intervening DNA. In general, looping out of the DNA is aided by protein binding to *cis*-elements, since DNA bends upon protein binding (Adhya and Garges, 1990; Travers, 1990).

The presence of enhancer elements change the nature of the chromatin to "active chromatin", which is characterized by unique DNA topologies, lower concentrations of certain histones, and modifications at both the physical and organizational level of the DNA structure (Fedor and Kornberg, 1989). Activator proteins bind to these accessible enhancer elements, and their binding serves to recruit additional transcription factors, including RNA polymerase II, to the promoter site (Dyana, 1989) via specific activation domains. These activation domains, such as acid blobs or negative

noodles, form tighter more defined structures when the factor binds to DNA, and so are likely activated by conformational changes brought on by the binding to DNA itself (Sigler, 1988; Giniger and Ptashne, 1987). The domains are also surprisingly tolerant to mutational alterations, implying that, although form may be important, the general acidic characteristic of the activation domain is the critical component. Activators which contain acidic domains, such as GCN4 (Brandl and Struhl, 1989), have been speculated to interact with the basic domain of TFIID (Lewin, 1990; Stucka and Feldmann, 1990), RNA polymerase II's heptapeptide repeat (Brandl and Struhl, 1989) or indirectly with the above motifs via factors that act as communication relays called co-operators (Nishizawa et al., 1990).

Many activation proteins function only as dimers, and some do this via a leucine zipper domain (Schütte et al., 1989). Consequently, multifaceted avenues of transcriptional control can be achieved, depending on the combinations of regulatory protein interactions, their ability to form homocomplex or heterocomplex formations through association domains, and their actual DNA binding affinities to various promoter elements (Frankel and Kim, 1991; Dynan, 1989).

Sequences known to interfere with these processes are called repressors or silencers. Repressors function most effectively when positioned between an upstream activation site (UAS) and a TATA box (Brent, 1985). Silencers, on the

other hand, can work over distances of kilobases from an initiation site, apparently by altering the chromatin state so that transcription factors cannot access the DNA (Brent, 1985).

But how are the regulators themselves regulated? For GCN4, translational control is important for its production. Other activators have external modifiers of activity. For example, GAL4 activation is impeded by glucose induced binding of GAL80 to its activation domain. Another example is NF $\kappa$ -B, which becomes effective by being modified in such a way that an inhibitor of a translocation signal is removed, enabling it to enter the nucleus where it can carry out its role (Lenardo and Baltimore, 1989). Levels of phosphorylation also influence the potency of many transcription activators (Muller *et al.*, 1988).

### 1.7) *Gene regulation in yeast*

The UAS of yeast is required for transcription to occur. Deletion of this element has been shown to reduce transcription to a basal level. It lies approximately -1400 to -100 bases upstream of the transcription initiation site. The UAS, although quite similar to an enhancer, fails the definition as far as activation 3' of the promoter is concerned. Yeast promoters also contain negative regulation sites called operators, or URSs (upstream repressible sequences), which have been described for loci such as *MAT*

(mating type locus) and *PYK* (pyruvate kinase; for a review see Guarente, 1987). The repressor element of the *PYK* gene surprisingly functions in the presence of glucose. Glucose causes an induction of gene products involved in fermentation while it reduces the expression of genes from many other enzymatic systems, including those involved in gluconeogenesis, the tricarboxylic acid cycle, and oxidative phosphorylation (Nishizawa et al., 1989; Wright and Poyton, 1990).

Studies of the regulation of the *S. cerevisiae* genes encoding the enzymes of galactose metabolism have greatly enhanced our understanding of the mechanisms of gene regulation by transcriptional control. The promoter that controls both the *GAL1* and *GAL10* genes has been shown to contain GAL4 binding sites (Giniger et al., 1985), galactose-independent activation sites GAE<sub>1</sub> and GAE<sub>2</sub> (Finley and West, 1989), and operators responsible for glucose repression (West et al., 1987). Genes encoding peroxisomal proteins are similarly regulated by both repression and activation (Dommes et al., 1983). One can easily envision that environmental cues such as levels of glucose, oxygen or fatty acids act as signals that influence the activities of the  $\beta$ -oxidation enzymes of both *S. cerevisiae* and *C. tropicalis*, since these compounds determine the oxidative state of the cell or the energy sources available for consumption. These signal molecules may bind to and activate specific DNA binding

proteins directly. They may also influence activities of various secondary messengers like SNF1 or hexokinase II (protein kinases shown to be involved in mediating glucose repression; Cherry *et al.*, 1989; Wright and Poyton, 1990; Ma *et al.*, 1989; Entian, 1980) whereby they indirectly influence transcription factors and their activities.

### 1.8) *Techniques for transcription analysis*

For mammalian cells, tools used to analyze mechanisms of gene regulation have often involved *in vivo* systems where transfection is the means by which exogenous DNA is added to cells. This transfected DNA has been manipulated by molecular biological techniques to produce constructs that contain fusions of various promoters to sequences encoding enzymes with unique, measurable catalytic activities (e.g. chloramphenicol acetyl transferase (CAT),  $\beta$ -galactosidase (Heard *et al.*, 1987), or luciferase (Gould and Subramani, 1988)). Virtually identical systems exist for *S. cerevisiae*, where the constructs of interest are transformed into the yeast (Mannhaupt *et al.*, 1988).

Since *C. tropicalis* has no readily available transformation system, we had to develop either an *in vitro* transcription system or a heterologous *in vivo* system in order to study the processes involved in the induction of *C. tropicalis* genes encoding peroxisomal proteins (Rachubinski, 1990). *In vitro* systems are useful for providing information



about which factors are required for transcription, and how these factors associate (Manley, 1983). We attempted to set up such an *in vitro* system for *C. tropicalis*, but with little success. Subsequently, we proceeded to establish and characterize an *in vivo* system. Since *C. tropicalis* HDE could be expressed, induced and detected in *C. albicans* (with an antibody that did not cross-react extensively with the endogenous *C. albicans* HDE), we were hopeful that a similar *in vivo* heterologous system would be successful using *S. cerevisiae* as the expression host for the *HDE* gene. An analogous system had failed for investigators attempting to express an oleic acid- inducible P450 enzyme from *C. tropicalis* in *S. cerevisiae* (Sanglard and Loper, 1989). However, we could express and induce the *HDE* gene in this organism, and we thereby concluded that the *cis*-elements present in the *C. tropicalis* *HDE* promoter could functionally interact with the transcription factors present in *S. cerevisiae*. Further genetic studies using deletion constructs enabled us to locate and characterize a few upstream regulatory elements of the *HDE* gene.

## MATERIALS AND METHODS

### 2.1) List of Abbreviations

Ac	acetate
AOX	acyl-CoA oxidase
ATP	adenosine 5'-triphosphate
bp	base pair(s)
CATL	catalase
°C	degrees Celsius
cccDNA	covalently closed circular DNA
cdNA	complementary DNA
Ci	Curie(s)
cpm	counts per minute
d	day(s)
Da	Dalton(s)
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
ddATP	2',3'-dideoxyadenosine 5'-triphosphate
ddCTP	2',3'-dideoxycytosine 5'-triphosphate
ddGTP	2',3'-dideoxyguanosine 5'-triphosphate
ddNTP	2',3'-dideoxynucleoside 5'-triphosphate(s)
ddTTP	2',3'-dideoxythymidine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside 5'-triphosphate(s)
DTT	dithiothreitol
dTTP	deoxythymidine 5'-triphosphate
EDTA	ethylene diamine tetraacetic acid
EtBr	ethidium bromide
g	gram(s)
h	hour(s)
HDE	hydratase dehydrogenase epimerase
Hepes	N-2-hydroxyethylpiperazine-N'2-ethanesulfonic acid
kbp	kilobase pair(s)
kDa	kilodalton(s)
KGB	potassium glutamate buffer
L	litre(s)
M	molar
mA	milliampere(s)
mg	milligram(s)
μCi	microcurie(s)
μg	microgram(s)
μL	microlitre(s)
μm	micrometre(s)
μM	micromolar

min	minute(s)
mL	millilitre(s)
mM	millimolar
M <sub>r</sub>	relative molecular mass
mRNA	messenger RNA
ng	nanogram(s)
nm	nanometre(s)
NTP	nucleoside 5'-triphosphate
ORF	open reading frame
ocDNA	open circular DNA
PEG	polyethylene glycol
pmol	picomole(s)
PMSF	phenyl methyl sulfonyl fluoride
PNK	polynucleotide kinase
RCF	relative centrifugal force
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
s	second(s)
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TBE	Tris-borate-EDTA
TE	Tris-EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris(hydroxymethyl)aminomethane
tsp	transcription start point
Tween 40	polyoxyethylenesorbitan monopalmitate
UTR	untranslated region
UV	ultraviolet
V	volt(s)
W	watt(s)
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

## 2.2) Materials and Suppliers

### 2.2.1) Common reagents

acetic acid (glacial)	BDH Inc. (Toronto, Ontario)
acrylamide	Gibco/BRL Canada (Burlington, Ontario)
agarose	Gibco/BRL
ammonium acetate	Fisher Scientific (Unionville, Ontario)
ammonium persulfate	BDH
ampicillin	Sigma Chemical Company (St. Louis, Missouri, USA)
aqueous counting scintillant (ACS)	Amersham Canada Limited (Oakville, Ontario)
ATP	Pharmacia (Canada) Inc.

Bacto-agar	(Baie D'Urfé, Quebec) <b>Difco Laboratories</b> (Detroit, Michigan, USA)
Bacto-peptone	Difco
Bacto tryptone	Difco
Bacto yeast extract	Difco
Bacto yeast nitrogen base (w/o amino acids)	Difco
<i>Bam</i> HI linkers	<b>Institute of Molecular Biology and Biotechnology</b> (McMaster University, Hamilton, Ontario)
Bio-Rad protein assay dye reagent concentrate	<b>Bio-Rad Laboratories</b> (Canada) Ltd. (Mississauga, Ontario)
bis-acrylamide	Bio-Rad
boric acid	BDH
bovine serum albumin (BSA)- DNase-free	Pharmacia
bromophenol blue	BDH
1-butanol	Fisher
calcium chloride	BDH
chloroform	<b>Caledon Laboratories Ltd</b> (Georgetown, Ontario)
chromatography paper (Whatman 3MM)	<b>Whatman International Ltd</b> (Maidstone, England)
Coomassie brilliant blue	Bio-Rad
cesium chloride	Gibco/BRL
deoxyribonucleotides (dATP/dGTP/dCTP/dTTP)	Pharmacia
DTT	Sigma
EDTA	Fisher
EGTA	Fisher
ethanol	<b>Consolidated Alcohols Ltd</b> (Toronto, Ontario)
ethidium bromide	Sigma
filtration units	<b>Gelman Sciences</b> (Ann Arbor, Michigan, USA)
formaldehyde	BDH
formamide	BDH
glass beads	Sigma
glucose	BDH
glycerol	BDH
GTG-agarose	<b>FMC Bioproducts</b> (Rockland, Maine, USA)
histidine	Sigma
hemoglobin (bovine blood)	Sigma
IPTG (isopropyl $\beta$ -D- thiogalactoside)	Gibco/BRL
isoamyl alcohol	Fisher
isopropyl alcohol	Caledon

leucine	Sigma
lithium acetate	J.T. Baker Chemical Co. (Phillipsburg, New Jersey, USA)
lithium chloride	Baker
magnesium chloride	BDH
2-mercaptoethanol	BDH
molecular weight standards for	
(a) DNA:	
(i) 1 kbp DNA ladder	Gibco/BRL
(b) protein:	
(i) Dalton Mark VII-L molecular weight markers for SDS-PAGE	Sigma
(ii) $\beta$ -galactosidase ( <i>E. coli</i> )	Sigma
(iii) phosphorylase b (rabbit muscle)	Sigma
(iv) prestained SDS-PAGE molecular weight standards	Bio-Rad
nitrocellulose (pore size - 0.45 $\mu$ m)	Schleicher & Schuell, Inc. (Keene, New Hampshire, USA)
Nonidet P-40 (NP-40)	Sigma
oleic acid (clear)	Fisher
ovalbumin	Miles Laboratories Ltd. (Republic of South Africa)
petri dishes	Fisher
phenol	Fisher
photographic flash unit	Braun Canada Ltd. (Mississauga, Ontario)
PMSF	Sigma
polyethylene glycol 8000	BDH Inc.
polyvinylpyrrolidone	Sigma
potassium acetate	BDH
potassium chloride	BDH
potassium glutamate	Fluka Chemika (Buchs, Switzerland)
potassium phosphate (mono and di-basic)	Baker
random primers DNA labelling system	Gibco/BRL
RNA guard RNase inhibitor	Pharmacia
RNasin RNase inhibitor	Promega (Unionville, Ontario)
SDS	Biorad
Sephadex G-50 (medium)	Pharmacia

Sequenase/Version 2.0 DNA sequencing kit	United States Biochemical Corporation (Cleveland, Ohio, USA)
Spectra/Por molecular porous membrane tubing	Fisher
sodium acetate	BDH
sodium borate	Baker
sodium chloride	BDH
sodium citrate	Baker
sodium hydroxide	BDH
sodium phosphate (mono and dibasic)	BDH
sodium pyrophosphate	Fisher
sorbitol	BDH
spermidine	Sigma
spermine	Sigma
TEMED	Biorad
tetracycline	Sigma
Tris	<b>Boehringer Mannheim Canada</b> Ltée (Dorval, Quebec)
tRNA (calf liver)	Boehringer
Tween 40	Sigma
urea	Biorad
X-gal (5-bromo-4-chloro- 3-indolyl- $\beta$ -D-galactoside)	Gibco/BRL
xylene cyanol	Sigma
X-omat XAR-5, XK1 film	<b>Eastman Kodak</b> (Rochester, New York, USA)
zinc chloride	BDH

### 2.2.2) Radiochemicals

<sup>14</sup> C-methylated $\beta$ -galactosidase (2.6 $\mu$ Ci/mg protein)	Sigma
<sup>14</sup> C-methylated phosphorylase b (5.3 $\mu$ Ci/mg protein)	Sigma
<sup>14</sup> C-methylated proteins (molecular weight range 14,000-70,000 Da) (5-50 $\mu$ Ci/mg proteins)	Sigma
<sup>125</sup> I-protein A (>30 mCi/mg total protein A, 0.1 $\mu$ Ci/ $\mu$ L)	Amersham
[ $\alpha$ - <sup>32</sup> P]dATP(3,000 Ci/mmol, 10 $\mu$ Ci/ $\mu$ L)	Amersham; <b>ICN Biomedicals,</b> Inc. (Irvine, California, USA)
[ $\gamma$ - <sup>32</sup> P]dATP(3,000 Ci/mmol, 10 $\mu$ Ci/ $\mu$ L)	Amersham; <b>Dupont/NEN</b> Canada Inc. (Mississauga, Ontario)

### 2.2.3) Enzymes

Alkaline phosphatase (calf intestinal mucosa)	Pharmacia
--	-----------

AMV	Pharmacia
Bal 31 exonuclease ( <i>Alteromonas espejiana</i> )	Gibco/BRL
DNA polymerase I Klenow fragment ( <i>E. coli</i> )	New England Biolabs, Inc. (Beverly, Massachusetts, USA); Pharmacia; Gibco/BRL
Exonuclease III ( <i>E. coli</i> )	Gibco/BRL
Modified T7 DNA polymerase/Sequenase	USB
Nuclease S1 ( <i>Aspergillus oryzae</i> )	Pharmacia
Restriction endonucleases	Gibco/BRL; NEB; Pharmacia
Reverse transcriptase (avian myeloblastosis virus)	Pharmacia
RNase I "A" (bovine pancreas)	Pharmacia
T4 DNA ligase	Gibco/BRL; Pharmacia
T4 polynucleotide kinase	Pharmacia

## 2.3) Methods

### 2.3.1) Nucleic acids

#### 2.3.1.1) Plasmid Isolation

Plasmids were isolated from *Escherichia coli* DH5- $\alpha$  using a variation of the method of Birnboim and Doly (1979) as described by Maniatis et al. (1982), which is commonly known as the "mini-prep". 4 mL of LB-amp media, inoculated with an appropriate DH5- $\alpha$  transformant, was incubated overnight at 37°C. 1.5 mL of this overnight culture was pelleted in a microfuge tube. The pellet was resuspended in 100  $\mu$ L of a buffered glucose solution (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl (pH 8.0)) followed by addition of 200  $\mu$ L of a chromosomal DNA and protein denaturing solution (0.2 N NaOH, 1% SDS). 150  $\mu$ L of an acidic salt solution (3 M K<sup>+</sup>, 5 M Ac) was used to render insoluble and precipitate complexed chromosomal DNA, large RNA molecules and proteins. The supernatant containing native plasmid DNA and small RNA molecules was removed and treated with 500  $\mu$ L of phenol:chloroform:isoamyl alcohol (50:49:1) to remove residual proteins. The extracted aqueous phase was removed and nucleic acids were precipitated by addition of 1 mL of absolute ethanol. The resultant pellet was washed with 70% (v/v) ethanol to remove residual salts.

The DNA was dried in a vacuum chamber connected to a water aspirator, and then dissolved in 20  $\mu$ L of TE (pH 8.0) containing 20  $\mu$ g/mL RNase A to digest contaminating RNA.



Plasmids were analyzed by restriction endonuclease digestion followed by gel electrophoresis to confirm plasmid identity (see section 2.3.1.2). Plasmid DNA was stored at 4°C.

Large scale preparations of plasmid DNA employed the same solutions and techniques as above, except they involved larger volumes, and hence a larger centrifuge (Maniatis *et al.*, 1982).

Depending on the volumes manipulated, the GSA, JA-10, JA-20 and SS-34 rotors were used in their respective centrifuges and centrifuged at equivalent RCF by manipulating rpm and/or time.

Large scale preparations were purified by CsCl equilibrium density gradient centrifugation. The nucleic acid pellet was dissolved in a TE (pH 8.0)/CsCl solution at a density of approximately 1.6 g/mL (1 g CsCl/mL TE-DNA solution containing 800 µg/mL ethidium bromide (EtBr)) and centrifuged at 100,000 rpm in a Beckman TLV100 rotor for 4 h. This centrifugation step pellets RNA, buoys protein, and separates the less dense chromosomal and nicked plasmid DNA from the supercoiled plasmid DNA. The closed circular DNA band was removed from the polyallomer tube using a syringe and extracted with an equivalent volume of water-saturated 1-butanol until no colour could be seen in the organic phase, and then extracted an additional four times.

The plasmid DNA solution containing CsCl was desalted

by one of two methods: by dialysis overnight in 2 x 1 L exchanges of TE (pH 8.0), or by precipitation of DNA using 2 volumes of TE (pH 8.0) (to dilute the CsCl) and 6 volumes of absolute ethanol, followed by a 70% (v/v) ethanol wash, drying in a vacuum chamber and dissolving the DNA in TE (pH 8.0). The DNA was quantitated as described in section 2.3.1.11.

### *2.3.1.2) DNA Modification Reactions*

Standard procedures were used for restriction endonuclease digestion and plasmid construction (Maniatis et al., 1982). In general, restriction endonuclease digestions, along with many other enzymatic reactions, were performed in KGB (Hanish and McClelland, 1988). Otherwise, reactions were carried out according to the specifications of the manufacturers. Generally restriction endonuclease digestions were performed in 20  $\mu$ L final volumes with 500 ng or 1/10th the DNA obtained from a "mini-prep", and with a minimum 2-fold excess of required enzyme at the specified temperature for 1 h in an air incubator to minimize condensation of vapour to microfuge tube surfaces.

Concentrations of DNA in the digestions were limited to 500 ng/ $\mu$ L and volumes of enzyme (in 50% v/v glycerol) to  $\leq$  10% of the total volume in order to keep glycerol concentrations below 5%.

### 2.3.1.3) Gel Electrophoresis of DNA

DNA digests were analyzed by one of two methods, depending on the sizes of the fragments to be resolved. Small fragments of 25-1500 bp were fractionated on 5% polyacrylamide gels, whereas 0.7%-1.5% agarose gels were used to resolve larger fragments of 0.5-10 kbp (Maniatis *et al.*, 1982). 1% GTG-agarose gels were used to purify DNA fragments.

Agarose gels were the most common fractionation matrix used in the analysis of restriction endonuclease digestion products. 0.7%-1.0% agarose gels were made in 1X Tris-borate/EDTA (TBE) buffer (90 mM Tris-borate, 2 mM EDTA). Generally, 0.4 g of agarose was dissolved in 40 mL of 1X TBE by microwaving the solution to boiling and briskly stirring the heated solution at room temperature until the agarose cooled to 55°C. EtBr was added to 500 ng/mL and the agarose poured into a casting tray containing an 8, 12, or 16 well comb. The agarose solution was allowed to solidify at room temperature for 45 min.

To prepare the DNA for electrophoresis, 0.2 volumes of 6X DNA gel loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 40% sucrose) was added to the DNA solutions prior to gel loading. The gel was submerged in 1X TBE-EtBr buffer (TBE buffer containing 500 ng/mL EtBr) contained in a submarine electrophoresis chamber, and the DNA samples were loaded into the wells, alongside molecular size markers (approximately 300 ng diluted in 1X DNA sample buffer (10 mM

Tris-Ac (pH 7.6), 20% (v/v) glycerol, 0.025% bromophenol blue)). The gels were electrophoresed at 5 V/cm until the desired dye migration was achieved. DNA bands were visualized by UV irradiation of the gel, resulting in the fluorescence of the EtBr intercalated into the DNA.

Occasionally DNA was fractionated using 5% polyacrylamide gel electrophoresis. Glass plates, along with combs and spacers, were thoroughly cleaned and assembled using clamps ensuring that all plates and spaces were aligned properly. A 5% acrylamide solution containing 0.25% bis-acrylamide, 1X TBE, 0.067% ammonium persulfate, was degassed for 15 min, and TEMED then added to 0.03% (v/v) to initiate polymerization. The gel solution was poured between the plates, and the comb added to form wells. Throughout the polymerization, excess solution held at 4°C was added to the top of the gel to combat gel shrinkage.

For polyacryamide gels as well as agarose gels, 0.2 volumes of DNA sample buffer was added to the DNA solution, the sample loaded into the wells of the gel, and the samples subjected to electrophoresis using 1X TBE as the buffer.

To visualize the bands the gel was immersed for 15 min into 1X TBE-EtBr buffer and subsequently irradiated with UV light to illuminate the DNA bands.

#### **2.3.1.4) Production of DNA with blunt ends**

The following three sections describe the different

techniques used to make DNA ends blunt (Ausubel et al., 1988).

3'-DNA overhangs necessitated use of the 3'->5' exonuclease activity of T4 DNA polymerase in the presence of excess dNTP's. Under these conditions the enzyme removes single stranded nucleotides until double stranded DNA is encountered.

5'-DNA overhangs were made blunt by use of the 5'->3' polymerase capabilities of the Klenow fragment of *E. coli* DNA polymerase I, which lacks 5'->3' exonuclease activity but has the 3'->5' exonuclease activity.

S1 nuclease was used to remove single-stranded portions of DNA produced by Exo III digestions of DNA (see section 2.3.1.8.1).

#### **2.3.1.4.1) Using Klenow**

Klenow reactions were performed in KGB (0.5-2.0X) by adding dNTP's (to 20  $\mu$ M; from a 10X stock) to a solution containing DNA fragments possessing 5'-overhangs. DNA concentrations in the reactions were less than 200 ng/ $\mu$ L, and generally 1 U was used per  $\mu$ g of total DNA to be made blunt. The reaction was inactivated at 75°C for 10 min, followed by restriction endonuclease digestion(s) with a second enzyme if required.

#### **2.3.1.4.2) Using T4 DNA Polymerase**

DNA possessing 3'-overhangs was digested back to

duplex regions using T4 DNA Polymerase in the presence of high concentrations of dNTP's (100  $\mu$ M each dNTP). The reaction buffer used was KGB (0.5-2X) containing 0.1 U T4 DNA polymerase and up to 2  $\mu$ g of DNA.

#### 2.3.1.5) *Random Primer Labelling of DNA*

DNA fragments to be used as probes were radioactively labelled using the Random Primer Labelling Kit from Gibco/BRL. The protocol employed was essentially as suggested in the data sheets accompanying the labelling kit, except that the reaction volumes were reduced to 50%. Basically 12.5 ng of a gel purified fragment was diluted to 13  $\mu$ L in water and boiled for 2 min to denature the DNA strands. After cooling the DNA solution on ice, 7.5  $\mu$ L of the random primer buffer solution (670 mM Hepes, 17 mM MgCl<sub>2</sub>, 33 mM  $\beta$ -mercaptoethanol, 1.33 mg/mL BSA, 18 OD units/mL oligodeoxyribonucleotide primers (hexamer fraction) (pH 6.8)) was added along with 1  $\mu$ L of each nucleotide solution (0.5 mM), except for the nucleotide to be used as a label. 25  $\mu$ Ci of an [ $\alpha$ -<sup>32</sup>P]-labelled dNTP (3000 Ci/mmole, 10  $\mu$ Ci/ $\mu$ L) was added followed by 0.5  $\mu$ L of Klenow fragment (2.0 U). The reaction was incubated for 1 h at 25°C. The labelled probe was purified from unincorporated nucleotides by passing the total reaction over a G50-spin column (see section 2.3.1.6.2). 1  $\mu$ L of the eluate was added to 10 mL of aqueous counting scintillant and counted in a scintillation counter. Generally, specific activities of 2 x

$10^9$  cpm/ $\mu$ g were obtained.

### **2.3.1.6) Gel purification of DNA**

#### **2.3.1.6.1) From agarose gels**

DNA fragments requiring purification from contaminating fragments were fractionated on 1% GTG-agarose gels made using 0.5X TBE. The desired fragments were sectioned out of the gel using a scalpel, and the agarose slices containing DNA were placed in the wells of an IBI electroelutor apparatus. 0.5X TBE was added until the slice was almost submerged. 120  $\mu$ L of 7.5 M  $\text{NH}_4\text{OAc}$  was added to all six channel wells, regardless of the number of channels actually being used, thereby preventing short circuiting of the current via the empty channels. The DNA was electroeluted at 100 V for 1 h. 400  $\mu$ L was removed from each channel to recover the salt plug (which contained the DNA), and then 1 mL of absolute ethanol was added to precipitate the DNA. The DNA pellet was rinsed once with 70% (v/v) ethanol, dried, and dissolved in an appropriate volume of  $\text{H}_2\text{O}$ .

#### **2.3.1.6.2) Over G-50 Sephadex spin columns**

To purify labelled DNA (see section 2.3.1.5) from unincorporated nucleotides or to quickly change the buffer in which the DNA was dissolved, a gel matrix of G-50 was used to exclude the DNA (molecular cut off 10,000 Da) and to trap all smaller molecules, resulting in a DNA solution determined by

the buffer in which the gel beads were equilibrated (Maniatis et al., 1982).

To perform the above, a 1 mL syringe was plugged with siliconized glass wool, filled with TE (pH 8.0) equilibrated G-50 Sephadex suspension, and placed in a 15 mL polypropylene tube containing an uncapped-microfuge tube to collect the eluate. The column was centrifuged in a bench top centrifuge for 1 min to compress the gel in the column and refilled with more G-50 suspension, until the syringe contained a bed volume of 1 mL. The column was then rinsed with 100  $\mu$ L of TE (pH 8.0) buffer, re-centrifuged, and this process repeated. The DNA solution was diluted to 100  $\mu$ L and loaded onto the top of the G-50 column. The column was re-centrifuged, and the eluate containing the purified DNA was collected.

#### 2.3.1.7) *Ligation of DNA Fragments*

Ligation reactions were routinely performed using T4 DNA ligase because of its versatility in ligating DNA with both sticky and blunt ends. The DNA concentrations and reaction conditions depended largely on the nature of the fragments and the required end products. Generally, ligations were performed at DNA concentrations of approximately 5-25 ng/ $\mu$ L, as recommended by Dugacieck et al. (1975). The insert to vector ratio was 1:1, 2:1 and, occasionally, 5:1. Ligation volumes were generally small, approximately 10  $\mu$ L, to conserve DNA. ATP was added to a final concentration of 1 mM, along



with 1 U of T4 DNA ligase.

Ligation reactions were generally performed overnight at room temperature. However, if the ligation reaction contained fragments with complementary ends, a 2 h incubation at room temperature was sufficient to produce the desired end products.

### **2.3.1.8) Generation of HDE sequencing constructs**

#### **2.3.1.8.1) Exo III generation of constructs for use in sequencing HDE in the 3'→5' direction**

The plasmid 3T containing *HDE* cloned into the *Sal*I site of pGEM5Zf(+) was used as a starting material to complete the sequencing of *HDE* in both directions. The restriction endonuclease site used at the 3'-end of *HDE* to produce an Exo III susceptible end (5'-overhang) was *Nde*I. The restriction endonuclease site used at the 5'-end of *HDE* to produce an Exo III protected end (3'-overhang) was *Sac*I.

9 µg of plasmid 3T was cut with both *Nde*I and *Sac*I in 1X KGB (see section 2.3.1.2). The DNA was precipitated with ethanol and dissolved in 60 µL of 1X Exo III buffer (50 mM Tris-HCl (pH 7.5), 5 mM MgCl<sub>2</sub>, 5 mM DTT) containing 300 U Exo III. The incubation was at 37°C. 3 µL was removed every minute for 20 min and placed in 1X S1 buffer (50 mM NaAc (pH 4.5), 0.2 M NaCl, 1 mM ZnCl<sub>2</sub>) and frozen in a dry-ice ethanol bath. When all time points had been collected, the Exo III was inactivated by heating at 70°C for 10 min. 5 U of S1

nuclease was added to each microfuge tube and the microfuge tubes were incubated at room temperature for 30 min to cleave all single-stranded DNA. The S1 nuclease was inactivated by a pH shift to 8.0 upon addition of 5  $\mu$ L of 6X S1 stop buffer (0.5 M Tris-HCl (pH 8.0), 125 mM EDTA). 5  $\mu$ L of each time point was analyzed on an agarose gel to verify the extent and quality of the Exo III digestions (Henikoff, 1984).

The DNA from each time point was precipitated with ethanol, and each pellet was dissolved in 20  $\mu$ L of 1X KGB. 2 U of Klenow was added to each microfuge tube for 2 min at 37°C, followed by addition of 2  $\mu$ L of a 10X dNTP mix (250  $\mu$ M of each dNTP) and further incubation of the samples at 37°C for 10 min. Ligations of the digested plasmids were performed by adding 4  $\mu$ L of 10 mM ATP, 1 U of T4 DNA ligase, by diluting the reaction volume to 40  $\mu$ L, and by incubating the reactions at room temperature overnight.

50  $\mu$ L of DH5- $\alpha$  cells were transformed with plasmids produced from each ligation reaction (see section 2.3.5.2) and plated onto LB-amp plates. Plasmids were isolated from colonies using the 'mini-prep' method (see section 2.3.1.1) and analyzed by restriction endonuclease digestion to determine the size of the deletion in each transformant. Appropriate deletion constructs were then sequenced using the M13 reverse sequencing primer (see sections 2.3.1.9 and 2.3.1.10).

Because of the inability to find certain deletions

amongst the various clones, the following 'drop-out' constructs were generated to aid the 3'→5' sequencing strategy of the *HDE* gene. 3T was digested with *Bst*XI and *Hind*III to sequence upstream of nt 1333; *Pfl*MI and *Nde*I to sequence upstream of nt 2059; and *Eco*RI and *Nde*I to sequence upstream of nt 1759 (see left pointing arrows with boxed ends in Fig. 6). DNA ends were made blunt with T4 DNA polymerase and the linearized DNA ligated to produce the constructs of interest (see sections 2.3.1.2, 2.3.1.4, and 2.3.1.7, respectively).

#### **2.3.1.8.2) Completion of the 5'→3' sequencing of the *HDE* gene**

Dr. J.D. Aitchison had partially sequenced the *HDE* gene in the 5'→3' direction (see Fig. 6). Frozen Exo III deletion reactions (generated by Dr. Aitchison by the method given in section 2.3.1.8.1) were rethawed and used to transform *E. coli* DH5- $\alpha$  (see section 2.3.5.2). Plasmid DNA was isolated from the colonies (see section 2.3.1.1), and analyzed by restriction endonuclease digestion (see section 2.3.1.2). Two clones were found, one to complete the sequencing (see sections 2.3.1.9 and 2.3.1.10) downstream of nt 1056, and the other to complete the sequencing downstream of nt 3020 (see large arrows in Fig. 6).

*Pst*I was used to excise a *Pst*I fragment from 3T so as to produce upon ligation a construct (see sections 2.3.1.2 and section 2.3.1.7) that enabled us to sequence downstream of nt

-166 (see right-pointing arrows initiating from small boxes, Fig. 6).

### 2.3.1.9) DNA Sequencing Reactions

Double stranded sequencing was performed by the method of Sanger *et al.* (1977) as modified by Zhang *et al.* (1988). 3  $\mu\text{g}$  of 'mini-prep' DNA was denatured in 20  $\mu\text{L}$  of 0.2 M NaOH, 0.2 mM EDTA. Neutralization solution (2  $\mu\text{L}$  of 2 M  $\text{NH}_4\text{OAc}$  (pH 4.6)) was added, followed by 60  $\mu\text{L}$  of absolute ethanol to precipitate the denatured DNA. The DNA was pelleted in a microfuge for 10 min at 4°C, washed with 70% (v/v) ethanol, and dried. The DNA was dissolved in sequencing buffer (40 mM Tris-Cl (pH 7.5), 20 mM  $\text{MgCl}_2$ , 50 mM NaCl) which contained 1 pmol of sequencing primer in a total volume of 10  $\mu\text{L}$ . To anneal primers, the denatured DNA was heated to 65°C and slowly cooled to 35°C. The extension of the primers bound to the DNA was initiated by adding 1  $\mu\text{L}$  of 0.1 M DTT, 2  $\mu\text{L}$  of 5X diluted labelling mix (dGTP mix: 7.5  $\mu\text{M}$  of each of dGTP, dTTP, and dCTP), 2  $\mu\text{L}$  of 1:8 diluted Sequenase (2U) and 0.5  $\mu\text{L}$  of [ $\alpha$ - $^{32}\text{P}$ ]dATP (3000 Ci/mmol) to the reaction volume, mixing the contents and incubating the reaction at room temperature for 5 min. After 5 min, 3.5  $\mu\text{L}$  of the extension reaction was added to four microfuge tubes, each containing 2.5  $\mu\text{L}$  of one of the following termination mixes (ddG: 80  $\mu\text{M}$  dGTP, 80  $\mu\text{M}$  dATP, 80  $\mu\text{M}$  dCTP, 80  $\mu\text{M}$  dTTP, 50 mM NaCl, 8  $\mu\text{M}$  ddGTP; ddA: 80  $\mu\text{M}$  dGTP, 80  $\mu\text{M}$  dATP, 80  $\mu\text{M}$  dCTP, 80  $\mu\text{M}$  dTTP, 50 mM NaCl, 8  $\mu\text{M}$

ddATP; ddC: 80  $\mu$ M dGTP, 80  $\mu$ M dATP, 80  $\mu$ M dCTP, 80  $\mu$ M dTTP, 50 mM NaCl, 8  $\mu$ M ddCTP; ddT: 80  $\mu$ M dGTP, 80  $\mu$ M dATP, 80  $\mu$ M dCTP, 80  $\mu$ M dTTP, 50 mM NaCl, 8  $\mu$ M ddTTP) for 5 min. The labelling reactions were stopped by adding 4  $\mu$ L of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF) to each microfuge tube.

The DNA in the sequencing reactions was then fractionated on 6% denaturing polyacrylamide gels (1.5  $\mu$ L per lane; see section 2.3.1.10).

#### 2.3.1.10) Sequencing Gels

Products of sequencing reactions were fractionated on 30 cm x 40 cm x 0.4 mm 6% denaturing (7M urea) polyacrylamide gels at 60 W, after the samples had been heated at 70°C for 2 min (Maniatis et al., 1982). Samples were electrophoresed for various lengths of time, depending on the amount of sequence information desired.

To resolve up to 100 nucleotides, samples were electrophoresed until the bromophenol blue reached the bottom of the gel. To resolve between 100-250 nucleotides, the samples were electrophoresed until the xylene cyanol migrated to the bottom of the gel. New dye was added to the top of the gel, and the bromophenol blue allowed to migrate to the bottom of the gel. To resolve >250 bases, samples were electrophoresed until the xylene cyanol had migrated an equivalent of two gel lengths, plus the time needed for

bromophenol blue to migrate to the bottom of the gel.

When the samples had been electrophoresed for the desired length of time, the gel assembly was removed from the electrophoresis apparatus. The plates were then carefully separated and the gel overlaid with a piece of Whatman 3MM paper. The gel was dried using a gel drier for 1 h at 80°C. The dried gel was placed in contact with XAR-5 film for the required length of time necessary to visualize the bands.

#### 2.3.1.11) *Nucleic Acids Quantification*

Two methods were used to quantify nucleic acids, either that of UV absorbance at 260 nm or fluorometry. UV spectrophotometry was utilized to assay the concentration of CsCl purified plasmid from large scale plasmid preparations (see section 2.3.1.1). Extinction coefficients for nucleic acids at  $A_{260}$  are  $20 \text{ g}^{-1}\text{Lcm}^{-1}$  for DNA,  $25 \text{ g}^{-1}\text{Lcm}^{-1}$  for RNA and  $50 \text{ g}^{-1}\text{Lcm}^{-1}$  for nucleotides.  $50 \mu\text{L}$  of a dialysed plasmid solution diluted to 1 mL was read at 260 nm where the  $A_{260} = [\text{DNA stock}]$  in mg/mL. The absorbance was also read at 280 nm and the ratio of  $A_{260}/A_{280}$  was used as an indicator of DNA purity. A ratio greater than 1.8 means little contaminating protein (Maniatis *et al.*, 1982).

Fluorometry was generally used to assay the level of RNA or DNA in a solution because of its greater sensitivity compared to that of UV spectrophotometry and its ability to distinguish RNA, ocdDNA, cccDNA and linear (chromosomal) DNA by

using various pH's and reagents. For fluorometry 500 ng of linearized plasmid DNA (determined by UV spectroscopy) was used as a standard in 1 mL of pH 11.8 buffer (0.5  $\mu$ g/mL EtBr, 20 mM  $K_2HPO_4$ , 0.5 mM EDTA, adjusted to pH 11.8 with 10 N KOH). The fluorometer was zeroed with buffer alone and the scan set to 500 arbitrary units upon addition of the standard. The data presented on the following page in Table 1 was used to determine the concentration and identification of nucleic acid in sample microfuge tubes.

pH 11.8 buffer causes cccDNA to fluoresce at 72% of the level of linear DNA since it is contortionally restrained, whereas RNA gives no fluorescence at all because pH 11.8 prevents the secondary structure formations required for EtBr intercalation. Heat denaturation at pH 11.8, followed by quick cooling on ice, will reduce the signal of linear DNA to approximately 4% because of strand separation. It will not affect cccDNA, however, because the strands are intertwined and therefore at cooler temperatures have nucleation sites that can overcome the repulsion experienced by the strands at the high pH resulting in a zipper-like reannealing of the DNA strands (Morgan et al., 1979).

In pH 8.1 buffer (0.5  $\mu$ g/mL EtBr, 5 mM Tris-HCl (pH 8.1), 0.5 mM EDTA), RNA intercalates EtBr to 50% the level of linear DNA. Therefore to quantitate nucleic acid solutions, the signals for RNA were doubled and those for cccDNA were multiplied by a factor of 1.4. To determine if RNA samples

TABLE 1 Fluorescence of nucleic acids in the presence of ethidium bromide.  
(Reproduced from Morgan *et al.*, 1979)

Nucleic Acid	pH of solution	Approximate Relative Fluorescence	
		Before Heat	After Heat
Calf thymus DNA	8.1	100	50
rRNA ( <i>E. coli</i> )	8.1	50	50
Calf thymus DNA	11.8	71	0-3
100% cccDNA	11.8	51	51
Nicked or linear cccDNA	11.8	71	0-3
Topoisomerase-relaxed cccDNA	11.8	37	37
rRNA ( <i>E. coli</i> )	11.8	0	0



were contaminated with DNA, RNase A (1  $\mu$ g) was added to the samples, and any residual fluorescence was attributed to the presence of DNA.

#### 2.3.1.12) Colony Hybridization

Colony hybridization was performed according to the method of Grunstein and Hogness (1975). LB-amp plates containing between 100 and 1000 colonies were overlaid with nitrocellulose and marked for orientation with a sterile needle. 3.0 mL of a denaturing solution (0.5 M NaOH) was placed on a piece of Saran Wrap onto which the filters were floated with the colony side facing up. The nitrocellulose disks were removed from the solution and blotted on paper towels. The NaOH treatment was repeated for another 3 min, and the filters were re-blotted. The filters were subsequently placed on 3 mL of 1 M Tris-HCl (pH 7.6) for 3 min to neutralize the NaOH, blotted, and the process repeated. The final treatment involved rinsing the filters with 5 mL of 2X SSC (0.3 M NaCl, 0.03 M Na•citrate) for 5 min and removing residual colony debris. Filters were blotted and dried under a heat lamp for 10 min. The DNA was baked onto nitrocellulose at 80°C for 2 h *in vacuo*.

Prehybridization of the nitrocellulose was performed in 3 mL/disk of 1X prehybridization solution (1.25X SSC, 0.16X Denhardt's solution [0.0032% Ficoll, 0.0032% polyvinylpyrrolidone, 0.0032% BSA (Fraction V)], 0.00125% SDS,

20 mM sodium phosphate buffer (pH 7.1), 4  $\mu\text{g}/\text{mL}$  denatured salmon sperm DNA) at 65°C for 2 h. The prehybridization solution was removed and replaced with hybridization solution (same as prehybridization solution plus 30% formamide, and  $10^6$  cpm/mL of random-primer labelled probe; see section 2.3.1.5) and incubated overnight at 42°C. The disks were washed in 1X SSC at 50°C four times for 5 min each wash. The disks were placed on 1X SSC-moistened Whatman 3MM paper, wrapped in Saran Wrap and exposed to autoradiography with an intensifying screen. If background levels of radiation were too high the disks could be subsequently rewashed at increased stringency.

#### 2.3.1.13) RNA Isolation

RNA was isolated from *S. cerevisiae* according to the method of Schmitt et al. (1990). 50 mL of late log cultured cells were pelleted, washed with distilled water and resuspended in 400  $\mu\text{L}$  of AE buffer (50 mM NaAc (pH 5.3), 10 mM EDTA). 40  $\mu\text{L}$  of 10% SDS was added, the solution vortexed and 450  $\mu\text{L}$  of AE-saturated phenol was added. The solution was vortexed, placed at 65°C for 4 min and then placed at -70°C in a dry ice/ethanol bath until phenol crystals appeared. The sample was centrifuged at room temperature in a microfuge at maximum speed for 4 min, the aqueous phase removed, and the sample re-extracted with 450  $\mu\text{L}$  of phenol: $\text{CHCl}_3$ :isoamyl alcohol. 40  $\mu\text{L}$  of 3M NaAc (pH 5.3) was added to the aqueous phase, followed by 1 mL of ethanol, and the sample was

centrifuged for 10 min. The pellet was washed with 70% (v/v) ethanol and dried in a vacuum desiccator. The pellets were dissolved in a minimal volume of water, approximately 20  $\mu$ L. The concentration of RNA was determined by fluorometry using pH 8.1 buffer (see section 2.3.1.11).

#### **2.3.1.14) Primer Extension**

Primer extension was performed as outlined in Ausubel *et al.* (1988). 100 ng of primer AB 1069 was end-labelled using T4 polynucleotide kinase (PNK) with 20  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] ATP at 37°C for 45 min. The reaction was performed in 1X PNK buffer (50 mM Tris-HCl (pH 7.6), 10 mM MgCl<sub>2</sub>, 5 mM DTT, 0.1 mM spermidine-HCl, 0.1 mM EDTA), containing 4 U of T4 PNK. The labelled oligonucleotide was purified from unincorporated label by three standard ethanol precipitations and counted in a scintillation counter. 500,000 cpm of oligonucleotide were added to 50  $\mu$ g of total RNA and precipitated with ethanol. The pellet was dissolved in 30  $\mu$ L of hybridization buffer (1 M NaCl, 167 mM Hepes (pH 7.5), 330  $\mu$ M EDTA), and hybridized overnight at 30°C. The RNA was precipitated, washed with 75% (v/v) ethanol/25% (v/v) 0.1 M NaAc (pH 5.2), and allowed to air dry. The pellet was dissolved in 25  $\mu$ L of buffer containing 1.25  $\mu$ L of RNasin, 0.56 mM dNTP's, 2.5  $\mu$ L 10X reverse transcriptase buffer (0.5 M Tris-HCl (pH 8.0), 50 mM MgCl<sub>2</sub>, 50 mM DTT, 0.5 M KCl). 40 U of AMV reverse transcriptase was added to the reaction, which was then

incubated at 42°C for 90 min. The reaction was terminated by the addition of 1  $\mu$ L 0.5 M EDTA (pH 8.0), and the RNA digested with 1  $\mu$ L of 1 mg/mL RNase A for 30 min at 37°C. The reaction volume was extracted with phenol/chloroform, precipitated, and dissolved in 3  $\mu$ L of TE (pH 8.0) and 4  $\mu$ L of sequencing buffer (see section 2.3.1.9).

The products were fractionated on a sequencing gel (see section 2.3.1.10), along with a sequencing ladder generated using the same primer annealed to plasmid 3T and extended using T7 DNA polymerase.

### 2.3.2) Proteins

#### **2.3.2.1) Yeast Lysate Preparation**

50 mL cultures of *S. cerevisiae* DL-1 transformants or *C. tropicalis* were cultured to late exponential phase ( $A_{600}=0.4$ ) in various media (see section 2.3.4). The cells were pelleted, washed in sterile H<sub>2</sub>O, and resuspended in an equal volume of breakage buffer (50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.1 mM EDTA, 0.1 mM ZnCl<sub>2</sub>, 15 mM PMSF). Glass beads were added to just below the level of the meniscus, and the microfuge tubes placed on ice. The microfuge tubes were vortexed for 30 s, followed by cooling on ice. This procedure was repeated twice. The crude extracts were pelleted for 10 s, and the supernatant containing the proteins was removed and collected.

Total protein content was determined by the method of

Bradford (1976), using ovalbumin as the standard. Buffers, which were identical to those containing protein samples, were used as blanks.

### **2.3.2.2) Sodium Dodecyl Sulfate Polyacrylamide Gel**

#### ***Electrophoresis (SDS-PAGE)***

Proteins in crude extracts, placed in 1X sample buffer, (50 mM Tris-HCl (pH 6.8), 0.1 mM DTT, 2% SDS, 0.1% bromophenol blue, 10% glycerol) were fractionated by SDS-polyacrylamide gel electrophoresis on 10% polyacrylamide gels (Fujiki et al., 1984). A 10% polyacrylamide gel solution (10% acrylamide, 2.6% bis-acrylamide, 367 mM Tris-HCl (pH 8.8), 1% SDS, 0.2% ammonium persulfate) was degassed, TEMED added to 0.05% (v/v), and poured between glass plates separated by 1.5 mm spacers assembled in a gel apparatus. The solution was poured to 0.75 the height of the glass plates and overlaid with 2-propanol to prevent interference with the polymerization process by oxygen.

A gel stacking solution (10% acrylamide, 2.6% bis-acrylamide, 1% SDS, 60 mM Tris-HCl (pH 6.8), 0.5% ammonium persulfate, 0.5% (v/v) TEMED) was placed over the polymerized resolving gel, and a comb was inserted into the stacking solution. The stacking gel was allowed to polymerize for 30 min before the gel was loaded.

The gel was loaded with proteins dissolved in 1X sample buffer and subjected to electrophoresis in 1X running

buffer (380 mM glycine, 50 mM Tris-HCl (pH 8.8), 0.1% SDS) until the bromophenol blue migrated to the bottom of the gel.

### 2.3.2.3) Immunoblotting

Protein gels destined for immunoblotting were electrophoresed (see section 2.3.2.2) until the dye front reached the bottom of the gel. The bottom right corner of the gel was removed, the gel inverted on a piece of Whatman 3MM paper saturated with transfer buffer (20 mM Tris-HCl, 150 mM glycine, 20% (v/v) methanol) and then placed atop a scotch pad located on a Western blot sandwich folder. Nitrocellulose was placed over the gel and trimmed to size with a razor blade. Another piece of Whatman 3MM paper and a scotch pad, both moistened in transfer buffer, were overlaid on the nitrocellulose, and all air bubbles were squeezed out. The whole assembly was sandwiched in the Western blot transfer apparatus.

The transfer was performed at 200 mA overnight in transfer buffer. The nitrocellulose containing the electrophoretically transferred proteins was then blocked in a 1% hemoglobin solution in 1X Tris-saline (0.9% NaCl, 10 mM Tris-HCl (pH 7.5)) for 30 min at 37°C. Primary polyclonal antibodies raised against HDE were added at 1:1000 dilution to blocking solution and incubated at room temperature for 1 h on a rocker platform to ensure adequate circulation of antibody solution. The antibody-containing solution was removed, and

2 x 20 min in 1X Tris-saline containing 0.05% (w/v) NP-40, and finally a repeat of the first wash. <sup>125</sup>I-labelled Protein A was then added to fresh 1% blocking solution (0.1  $\mu$ Ci/mL solution), the blots rocked for 30 min at room temperature. The same series of washes were used to remove the uncomplexed protein A. The blots were then air dried and exposed to preflashed XAR-5 film with an intensifying screen (Laskey and Mills, 1977; Burnette, 1981).

### 2.3.3) Vectors and Constructs

#### 2.3.3.1) *Subcloning of HDE into YCp50 for Pilot Experiments*

A 4.2 kbp *SalI* fragment containing *HDE* (Aitchison and Rachubinski, 1990) was gel purified, electroeluted and ligated, using an insert to vector ratio of 2:1 and a DNA concentration of 10 ng/ $\mu$ L (Dugacieck *et al.*, 1975), into similarly treated *SalI*-cut YCp50 (Rose *et al.*, 1987) to produce YCp50H. This ligation would insertionally inactivate the tetracycline resistance gene. Transformed *E. coli* DH5- $\alpha$  cells were therefore replica plated onto both tetracycline and ampicillin plates. Tet<sup>Amp</sup> transformants were identified, grown, and analyzed by restriction endonuclease digestions of isolated plasmid DNA to identify recombinants and orientation of *HDE* in recombinants (see section 2.3.1.2). Of the first 60 transformants tested, only one recombinant was found, named YCp50H2, with the upstream region of *HDE* adjacent to the *URA3* gene (Fig. 3). Later, upon further screening of 80

transformants, a recombinant YCp50H1 was found with the *HDE* gene situated in the opposite orientation.

#### **2.3.3.2) Subcloning of *HDE* into YCp50 for *Bal31* Deletions**

A 5.4 kbp *ScaI-SacI* fragment containing *HDE* was excised from 3T, made blunt with T4 DNA Polymerase and gel purified on 1% GTG agarose. YCp50 was linearized by digesting with *SalI* and made blunt using Klenow fragment (see section 2.3.1.4.1). The *ScaI-SacI* fragment containing *HDE* was blunt end-ligated into the YCp50 vector using a j:i ratio of 1, insert to vector ratio of 2:1 and DNA concentration of 25 ng/ $\mu$ l. This produced plasmids YCp505H or YCp503H (see Fig. 8).

Competent DH5- $\alpha$  cells were transformed with DNA from the ligation reaction and screened for recombinants containing *HDE* by colony hybridization (see section 2.3.1.12). Recombinants were confirmed and insert orientation determined by restriction endonuclease digestions with *EcoRI* (see section 2.3.1.2).

#### **2.3.3.3) *Bal 31* Deletion Reactions**

*Bal 31* digestions were performed on plasmid YCp505H to generate 5'-deletions of the upstream regions of *HDE* according to Ausubel et al. (1987). YCp505H was digested with *XbaI* to produce the substrate ends required for *Bal 31* digestion. The digestion products were concentrated to 1  $\mu$ g/ $\mu$ L by a standard



ethanol precipitation. 16  $\mu\text{g}$  of linearized plasmid was digested in 1X KGB with 2 U of *Bal* 31 in 80  $\mu\text{L}$  at 37°C. At 1 min, 3 min, 5 min, 8.5 min, 13 min, 18 min, 24 min and 30 min, 10  $\mu\text{L}$  of the digestion was removed and quickly frozen at -70°C in a dry ice/ethanol bath. All samples were then heat inactivated at 70°C for 10 min. 1  $\mu\text{L}$  from each microfuge tube was removed and digested by addition of 9  $\mu\text{L}$  of a *Bgl*III cocktail (1X KGB, 2 U *Bgl*III) for 1 h at 37°C.

The products were analyzed on a 1% agarose gel to assess the extent and rate of digestion. 4  $\mu\text{L}$  of each time point was taken, pooled, and made blunt with T4 DNA polymerase (see section 2.3.1.4.2). The DNA was precipitated by a standard ethanol precipitation and dissolved in 10  $\mu\text{L}$  of water.

*Bam*HI linkers (8mer; GGGATCCC) were ligated overnight to the DNA under two conditions (50 mM Tris-Cl (pH 7.6), 10 mM  $\text{MgCl}_2$ , 1 mM ATP, 5 or 15% PEG 8000, 2.8  $\mu\text{g}$  DNA, 100 ng *Bam*HI linkers) in a total volume of 15  $\mu\text{L}$  (Pfeiffer and Zimmerman, 1983). The buffer and excess linkers were removed the following day by diluting each ligation to 50  $\mu\text{L}$  and passing the contents over a G50-column (see section 2.3.1.6.2). The ligation was then digested with 100 U of *Bam*HI in 75  $\mu\text{L}$  to ensure that all the *Bam*HI linkers were digested.

The DNA molecules containing sticky *Bam*HI ends, were purified again over a G50-column in order to rid the reaction of *Bam*HI fragments and were subsequently ligated in a 50  $\mu\text{L}$

volume at 10 ng/ $\mu$ L to recircularize the plasmids.

Fluorometry was used to follow the recovery of DNA at each step to ensure ligation reactions were performed at suitable DNA concentrations. DH5- $\alpha$  cells were transformed with 3  $\mu$ L of each ligation (*i.e.* 5% and 15% PEG 8000 linker-ligation) and colonies analyzed by restriction endonuclease digestions of isolated mini-plasmid DNA.

Appropriately sized deletions were sequenced using a primer 5'-ACTATCGACTACGCG-3' purchased from BRL, which would bind just upstream of the *Bam*HI site of the plasmid pBR322, a plasmid from which YCp50 is derived.

#### 2.3.3.4) *Production of Positive and Negative Control Constructs*

Plasmid YCp505Hwt (wild type) (5WT) was generated by cutting out the pGEM5Zf(+) sequences lying upstream of the *HDE* gene to produce a product with almost the identical flanking sequences found in all the deletion constructs (*i.e.* a clone that lacks a *Bam*HI linker at the *Bam*HI site and contains 9 extra nucleotides from the multiple cloning site upstream of the *Sal*I site of 3T). This was accomplished by digesting YCp505H with *Not*I and *Bam*HI in 1X KGB (see section 2.3.1.2) followed by Klenow treatment to make DNA ends blunt (see section 2.3.1.4.1), and a subsequent ligation to recircularize the plasmid (see section 2.3.1.7).

A negative control, YCp505H<sup>-</sup>, was derived by digesting

the plasmid YCp505H with *Sal*I, which excised the *Sal*I-fragment containing the *HDE* gene from the plasmid.

#### **2.3.4) Strains and Media**

The yeasts *Candida tropicalis* Berkhout strain pK233 (ATCC 20336) (Tanabe et al., 1966) and *S. cerevisiae* strain DL-1 (*MAT* $\alpha$ , *leu2-3, 2-112*; *his3-11, 3-15*; *ura3-251, 3-372, 3-328*) (Van Loon et al., 1983) were cultured in minimal medium (YNE; 0.67% yeast nitrogen base, 0.05% yeast extract, 0.5% (w/v) Tween 40) (Erdmann et al., 1989), containing one of the following as a carbon source: 5% glucose; 0.2% glucose; 0.2% glucose, 0.2% (w/v) oleic acid [glucose-oleic acid]; 2% (v/v) glycerol; 2% (v/v) ethanol; 2% (v/v) glycerol, 2% (v/v) ethanol [glycerol-ethanol]; 2% (v/v) glycerol, 2% (v/v) ethanol, 0.2% (w/v) oleic acid; or 2% (w/v) oleic acid. Leucine (30  $\mu$ g/mL) and histidine (20  $\mu$ g/mL) were added where necessary to meet auxotrophic requirements.

#### **2.3.5) Transformation Procedures**

##### **2.3.5.1) *S. cerevisiae* DL-1**

10 mL of a logarithmically growing culture ( $OD_{600}=0.5$ ) of *S. cerevisiae* strain DL-1 cultured in YEPD media was pelleted and washed twice with distilled water. The yeast pellet was resuspended in 100  $\mu$ L of a lithium salt solution (10 mM Tris-HCl, 1 mM EDTA, 0.1 M LiAc (pH 7.5)), agitated and incubated at 30°C for 1 h. The yeast cells were then

incubated with 50  $\mu\text{g}$  of salmon sperm DNA, along with 1  $\mu\text{g}$  of plasmid DNA, mixed and incubated a further 30 min at 30°C without agitation. 700  $\mu\text{L}$  of a PEG solution (40% PEG 8000, 0.1 M LiAc, 1X TE (pH 8.0)) was added to the mixture and incubated as previously. The cells were centrifuged for 2 s and the supernatant removed. The cell pellet was washed twice with TE (pH 8.0), resuspended in 100  $\mu\text{L}$  TE (pH 8.0), plated on selective media and incubated for 2-3 days until colonies appeared (Ito et al., 1983). Ura<sup>+</sup> prototrophs were selected on YNB-plates supplemented with leucine and histidine.

#### **2.3.5.2) *Escherichia coli* DH5- $\alpha$**

Competent DH5- $\alpha$  cells were purchased from Gibco/BRL (Burlington, ON) and transformed according to the product data sheets provided. In brief, DH5- $\alpha$  cells were thawed on ice (50  $\mu\text{L}$  per transformation), 2  $\mu\text{L}$  of a DNA solution was added, and the cells were incubated on ice for 30 min. The cells were then heat shocked at 37°C for 20 s, diluted to 1 mL with LB and incubated at 37°C for 1 h. The cells were subsequently concentrated by centrifugation, plated on LB-amp plates and incubated overnight at 37°C to allow for colony formation.

#### **2.3.6) Densitometric scanning**

Autoradiographic films were scanned on a Hoefer densitometer by maximizing the amount of digitized signal that could be analyzed per scan (within the limits of the

equipment). Each scan was performed in one sitting, in triplicate, and the average signal was used for comparative purposes. Bands were scanned only when the intensity of each signal was within the linear range of the film. Signals too weak or intense on one exposure were quantitated from longer or shorter exposures, and mathematically manipulated to correspond to data from other films.

## RESULTS AND DISCUSSION

### 3.1) *Expression of genes encoding HDE, acyl-CoA oxidase and catalase in Candida tropicalis*

Before proteins could be isolated from either *C. tropicalis* or *S. cerevisiae*, growth curves were plotted of these yeasts cultured in media containing oleic acid. Earlier experiments had indicated that, of all the carbon sources tested, oleic acid produced the lowest cell densities of *S. cerevisiae* at stationary phase. Consequently cells were harvested during late logarithmic phase, rather than mid-log phase or early log phase, since the late phase would provide sufficient cell mass from 50 mL cultures for isolation of proteins. Harvesting cells during a logarithmic growth phase would guarantee cells of a consistent metabolic state, regardless of the batch of cells grown or small fluctuations in cell density. All other media used in these experiments produced equivalent or increased *C. tropicalis* cell densities. Increased cell densities were also observed at late logarithmic phase for cells of *S. cerevisiae* cultured in media containing a carbon source other than oleic acid.

Proteinaceous lysates were prepared from *C. tropicalis* cells cultured in media containing different carbon sources. This was achieved by disrupting cells with glass beads in an

environment inhibiting protein degradation. Lysates, containing extracted proteins, were then analyzed by immunoblotting to assess the effects of various carbon sources on expression of two genes encoding the  $\beta$ -oxidation enzymes HDE and acyl-CoA oxidase (AOX), and the gene encoding the enzyme catalase (CATL). Although Tween 40 was needed only for the solubilization of oleic acid, it was nevertheless consistently added to all media as an invariant component. For the purposes of the discussion, catalase will be discussed as a ' $\beta$ -oxidation' enzyme because of its close association with the  $\beta$ -oxidation process, although in the strict sense it is not part of the pathway.

The addition of glucose to both high and low concentrations in YNE media greatly repressed *HDE* expression (35-fold and 29-fold, respectively) and repressed, though to a lesser extent, *CATL* (7.6-fold and 4.8-fold, respectively) and *AOX* (10-fold and 7.8-fold, respectively) expression (Fig. 2; Table 2; compare lanes - to lanes  $G^5$  and  $G^{0.2}$ , respectively). These results suggest that glucose repression and derepression play an important role in the regulation of expression of *HDE*, *CATL* and *AOX* in *C. tropicalis*, since absence of glucose produced high levels of expression, and presence of glucose, even at low concentrations, repressed expression, although not to the extent observed at high concentrations of glucose. This derepression, although slight, is consistent with the hypothesis that as one

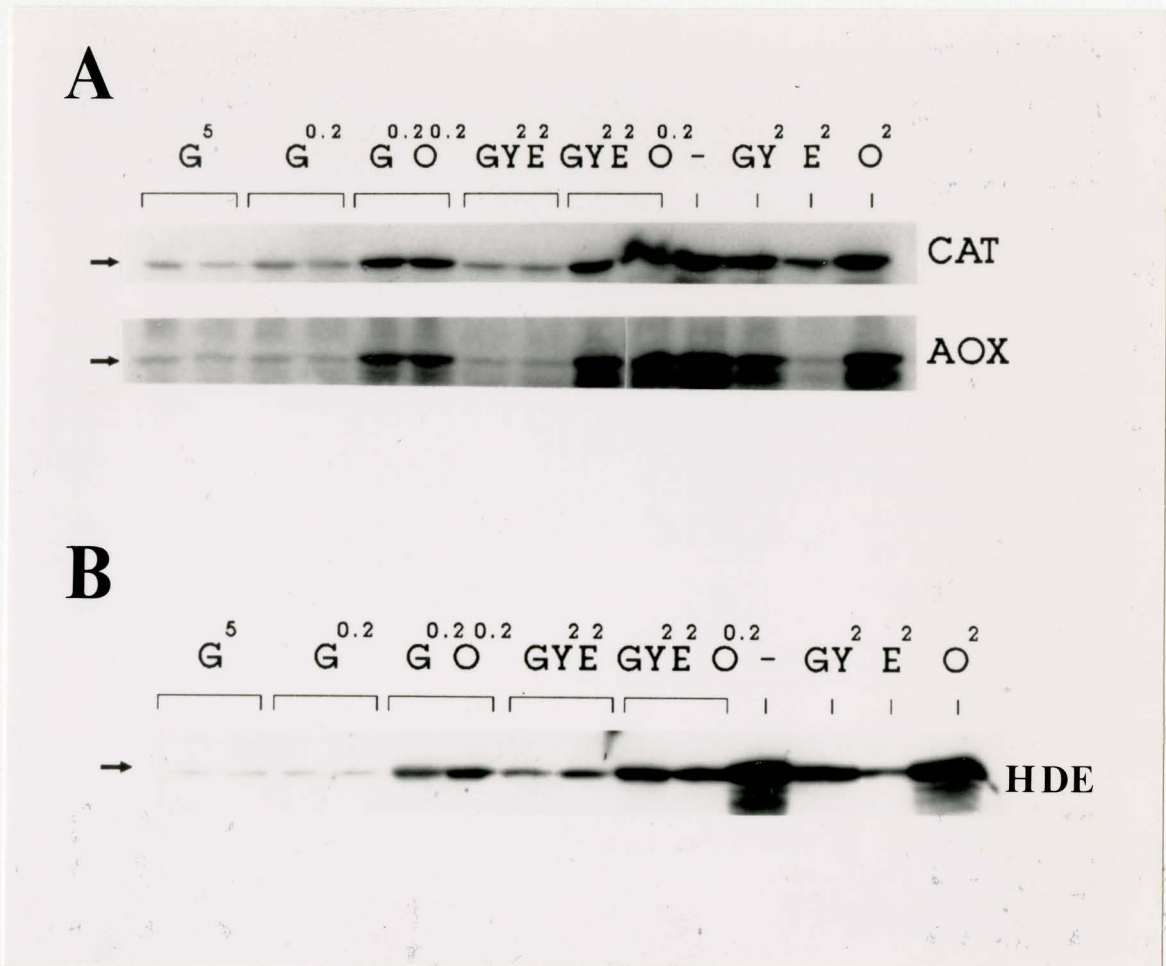


Figure 2. Effects of various carbon sources on expression of *HDE*, *AOX* and *CAT* in *Candida tropicalis* Berkhout strain pK233 (ATCC 20336). Panel A. Western blot analysis of total protein from *C. tropicalis* using antisera to *CATL* and *AOX*. Panel B. Western blot analysis of total protein from *C. tropicalis* using antiserum to *HDE*. Proteins were isolated from cells cultured in minimal medium (YNE; 0.67% yeast nitrogen base, 0.05% yeast extract, 0.5% (w/v) Tween 40) (Erdmann et al., 1989), containing one of the following as a carbon source: 5% glucose; 0.2% glucose; 0.2% glucose, 0.2% (w/v) oleic acid [glucose-oleic acid]; 2% (v/v) glycerol; 2% (v/v) ethanol; 2% (v/v) glycerol, 2% (v/v) ethanol [glycerol-ethanol]; 2% (v/v) glycerol, 2% (v/v) ethanol, 0.2% (w/v) oleic acid [glycerol-ethanol-oleic acid]; or 2% (w/v) oleic acid. Quantitative analyses were the average of at least two independent observations, except for *C. tropicalis* grown in YNE alone or YNE plus one of 2% (v/v) glycerol, 2% (v/v) ethanol, or 2% (w/v) oleic acid. G=glucose; GY=glycerol; E=ethanol; O=oleic acid; [-]=YNE. Superscripts refer to percentages of a particular carbon source. Equal amounts of protein were run in each lane.



CARBON SOURCE PRESENT	HDE SIGNAL DETECTED		CATL SIGNAL DETECTED		AOX SIGNAL DETECTED	
	SIGNAL	AVERAGE	SIGNAL	AVERAGE	SIGNAL	AVERAGE
G <sup>5</sup>	1.0	1.0	1.1	1.0	1.1	1.0
G <sup>5</sup>	1.0	± 0	0.9	± 0.1	0.9	± 0.1
G <sup>0.2</sup>	1.1	1.3	1.8	1.6	1.5	1.3
G <sup>0.2</sup>	1.4	± 0.2	1.3	± 0.4	1.1	± 0.3
G <sup>0.2</sup> O <sup>0.2</sup>	7.1	7.7	3.8	4.0	5.6	5.9
G <sup>0.2</sup> O <sup>0.2</sup>	8.3	± 0.8	4.2	± 0.3	6.2	± 0.4
GY <sup>2</sup> E <sup>2</sup>	3.5	4.6	1.7	1.8	1.4	1.5
GY <sup>2</sup> E <sup>2</sup>	5.7	± 1.6	1.8	± 0.1	1.6	± 0.1
GY <sup>2</sup> E <sup>2</sup> O <sup>0.2</sup>	16	15	4.6	4.6	7.1	7.1
GY <sup>2</sup> E <sup>2</sup> O <sup>0.2</sup>	14	± 1.4	-		-	
- (YNE)	35		7.6		10.2	
GY <sup>2</sup>	24		7.9		7.1	
E <sup>2</sup>	6.0		3.6		3.2	
O <sup>2</sup>	46		11		13.3	

TABLE 2 Relative densitometric signals obtained from bands present in Fig. 2. Each band was scanned in triplicate and the average value recorded in the table under the heading "SIGNAL". Averages and standard deviations obtained from independent samples, loaded adjacently on gels, are tabulated under the headings "AVERAGE".

energy source is being exhausted, the cells, through derepression, are diversifying their metabolic repertoire so that alternative energy sources can be utilized.

Similar observations have been made for many yeast genes, especially those encoding proteins involved in gluconeogenesis or the metabolism of other nonfermentive carbon energy sources. One such gene is the gene encoding alcohol dehydrogenase (*ADH2*). The product of this gene is involved in the metabolism of ethanol, and its response to the presence or absence of glucose has been examined by Cherry *et al.* (1989). They discovered that ADR1, a yeast transcription factor known to activate the *ADH2* gene, is phosphorylated rapidly upon addition of glucose to the media, thereby inactivating ADR1's ability to stimulate transcription of *ADH2*. The enzyme responsible for phosphorylating ADR1 is a cyclic-AMP dependent-protein kinase (cAPK), and could possibly be the protein hexokinase II, which functions not only in the phosphorylation of glucose to glucose-6-phosphate, but contains a glucose and cAMP-responsive kinase activity (Herrero *et al.*, 1989).

Although some discrepancies exist in the literature, the current hypothesis is that cAMP in fact mediates catabolite repression in yeast. When no glucose is present in the media, *i.e.* cells are cultured in nonfermentable carbon sources, then genes mediating fermentive catabolic processes are derepressed. Addition of a fermentable carbon source,

such as glucose, causes a transient increase in cAMP levels. This process is mediated itself by a glucose-repressible protein, so that when cells are cultured in glucose, this intermediate protein in the pathway is not present, and therefore no cAMP spike is observed in such cells when additional glucose is presented to them (Myboni et al., 1990)

Perhaps identical systems regulate the glucose responses observed with the  $\beta$ -oxidation genes. Analysis of the upstream region of *HDE* revealed a potential binding site for ADR1 (Fig. 7). In fact, it has been demonstrated that the mRNA levels from the *S. cerevisiae HDE* gene are affected in cells containing a defective *ADR1* gene (Simon et al., 1991). They demonstrated that *ADR1* influences genes encoding peroxisomal proteins most notably in cells cultured in sub-optimal growth conditions, i.e. limiting nutrient supplies. Since the metabolism of ethanol also requires a functional glyoxylate system, ADR1 therefore might conceivably play a general role in the regulation of all genes linked to the glyoxylate pathway.

The addition of 0.2% oleic acid to low concentrations of glucose increased expression of *HDE* 6.4-fold, of *CATL* 2.5-fold and of *AOX* 4.5-fold (Fig 2., compare lanes G<sup>0.2</sup>O<sup>0.2</sup> to lanes G<sup>0.2</sup>), demonstrating an induction of the expression of these genes by oleic acid. High levels of expression were seen for all three genes when cells were cultured in YNE media alone (containing only Tween 40) (Fig. 2, lanes -), indicating that

Tween 40 can be used as a sole carbon and energy source in *C. tropicalis* and is metabolized by the  $\beta$ -oxidation pathway. This is not surprising, since Tween 40 is largely (90%) composed of an ester of palmitate.

The addition of 2% oleic acid to YNE increased *HDE* (1.3-fold), *CATL* (1.4-fold), and *AOX* (1.3-fold) expression (Fig. 2, compare lanes O<sup>2</sup> to lanes -). Due to the observed errors of earlier comparisons, one can only speculate that this effect is real, and if so, that oleic acid and Tween 40 act in similar yet distinct additive fashions.

Expression of these enzymes in cells cultured in oleic acid was not repressed by glucose to the same extent as cells cultured in YNE media (Fig. 2, compare ratio of signal intensity of lanes O<sup>2</sup> versus G<sup>0.2</sup>O<sup>0.2</sup> to lanes - versus G<sup>0.2</sup>). Because cells cultured in YNE media responded in a more pronounced fashion to the presence of glucose, it seems that different regulatory mechanisms operate in effecting the activation/repression of these two carbon sources. At first glance this comparison may not seem valid, since cells in one instance were cultured in media containing 2% oleic acid, instead of 0.2% oleic acid. Nevertheless, these solutions were probably equivalent, since oil droplets could still be observed during late logarithmic growth in both media, implying that solubilized oleic acid concentrations were probably comparable.

*HDE* (3.8-fold), *CATL* (1.1-fold) and *AOX* (1.2-fold)

expression was increased when cells were cultured in glycerol-ethanol as opposed to growth in low concentrations of glucose (Fig. 2, compare lanes GY<sup>2</sup>E<sup>2</sup> to lanes G<sup>0.2</sup>). These results demonstrate that expression of  $\beta$ -oxidation enzymes is the consequence of many environmental cues, not just the presence of glucose or oleic acid. The presence of nonfermentable carbon sources also influence the expression levels of these genes, adding weight to the observation that general catabolite repression/derepression is an important regulator of  $\beta$ -oxidation activity.

The addition of 0.2% oleic acid to nonfermentable glycerol-ethanol further increased expression of *HDE* 3.3-fold, of *CATL* 2.6-fold and of *AOX* 4.7-fold (Fig. 2, compare lanes GY<sup>2</sup>E<sup>2</sup>O<sup>0.2</sup> to lanes GY<sup>2</sup>E<sup>2</sup>), reiterating oleic acid's role as an inducer of  $\beta$ -oxidation gene expression. The inductions seen here are remarkably similar to those observed upon addition of oleic acid to low concentrations of glucose, with the exception of the enzyme *HDE*. These similar activations, independent of the additional carbon sources present, indicate that activation and repression/derepression are probably mediated by two somewhat independent mechanisms, at least under derepressive conditions.

Ethanol, a nonfermentable carbon source, increased expression relative to glucose but was also observed to repress expression of  $\beta$ -oxidation genes in other circumstances. Ethanol reduced *HDE* (5.2-fold), *CATL* (4.4-

fold) and AOX (4.7-fold) expression when added to glycerol-containing media (Fig. 2, compare lanes GY<sup>2</sup>E<sup>2</sup> to lanes GY<sup>2</sup>). Yet, even though both glycerol and ethanol independently produced increases in gene expression with respect to low concentrations of glucose (Fig 1., compare lanes GY<sup>2</sup> and E<sup>2</sup> to lanes G<sup>0.2</sup>), the addition of ethanol to glycerol containing media reduced levels of expression to that observed for cells cultured in media containing ethanol only (compare lanes GY<sup>2</sup>E<sup>2</sup> to lanes GY<sup>2</sup> and E<sup>2</sup>). If both glycerol and ethanol act by positive mechanisms, one would predict that expression levels in media containing both carbon sources would at least equal or exceed expression produced by the carbon source giving the highest levels of expression, in this case, glycerol. Since ethanol represses expression when added to media containing glycerol, it cannot be induction by glycerol in the absence of glucose, but rather differential levels of derepression resulting in increased expression of these genes in glycerol alone. Presence of ethanol in the media derepresses these genes to a lesser extent.

Repressive regulation therefore plays an important role in the control of  $\beta$ -oxidation activities. Glucose, ethanol and glycerol all repressed expression of these genes (Fig. 2, compare lanes -, to each of the lanes G<sup>5</sup>, G<sup>0.2</sup>, GY<sup>2</sup> and E<sup>2</sup>), in a manner consistent with the following scenario.

It appears that the cell has an order of carbon source preference with respect to Tween 40 when cultured in media

containing alternative carbon sources. Using expression levels of these three genes as references, all carbon sources, other than Tween 40 and oleic acid, repressed levels of expression. This repression indicates that the cells, at the time of harvesting, are not yet committed to metabolizing Tween 40. The extent of gene repression reflects the prioritized position of the alternative carbon source with that of Tween 40. Since glucose repressed expression the most, its metabolism can be considered the most preferred by the cell when presented with the alternative of Tween 40 (Fig. 2, compare lanes - to lanes G<sup>5</sup> and G<sup>0.2</sup>, respectively). As the carbon source is shifted to ethanol and then glycerol, derepression increases steadily (Fig 1, compare lanes - to lanes E<sup>2</sup> and GY<sup>2</sup>).

This hypothesis of carbon source priority is demonstrated rather convincingly when cells are cultured in media containing both glycerol and ethanol. In such media, expression levels most resemble those observed for cells cultured in ethanol. Expression levels are not an average of the expression seen for ethanol and glycerol independently. Because there is expression of these  $\beta$ -oxidation genes in all the media tested, regardless of the carbon source, the  $\beta$ -oxidation pathway is probably always active, but controlled in its ability to efficiently catabolize Tween 40 because of limited enzyme concentrations. As the alternative carbon source is changed from glucose to ethanol to glycerol, this

catabolic rate is derepressed, and a larger proportion of the energy produced by the cells is attributable to the metabolism of Tween 40. Therefore, it appears that when *C. tropicalis* is cultured on carbon sources not metabolized by the  $\beta$ -oxidation pathway, a general negative control mechanism involving repression/derepression effects the expression of *HDE*, *AOX* and *CATL*, and positive activation is restricted to molecules that can be metabolized by the  $\beta$ -oxidation pathway.

The coordinate expression of *HDE*, *AOX* and *CATL* by different carbon sources implies that these genes share common regulatory elements. In order to investigate these observations, the upstream portion of one of the genes, namely *HDE*, was dissected and analyzed for functional sub-sequences. The sub-sequences obtained from upstream regions that mediated various aspects of gene regulation were then compared with the upstream regions of other genes encoding enzymes induced by oleic acid.

### 3.2) Expression of *HDE* in *S. cerevisiae*.

In order to localize the elements responsible for mediating the changes in *HDE* expression observed in *C. tropicalis* for the various carbon sources tested, we produced upstream deletion constructs. Because *C. tropicalis* is diploid, and no convenient transformation system exists for this yeast, we had to look for another organism to express *HDE*. Since it had already been demonstrated that a



heterologous *in vivo* system could be used to express and induce *HDE* in the yeast *Candida albicans* (Aitchison and Rachubinski, 1990), a trial experiment was set up to demonstrate whether or not *HDE* could be expressed similarly in the more genetically manipulable yeast *S. cerevisiae*. *HDE* was subcloned into the *SalI* site of YCp50 (Rose *et al.*, 1987) and one orientation was obtained, construct YCp50H2 (see section 2.3.3.1 and Fig. 3). This construct, along with the plasmid YCp50 alone, was then transformed into *S. cerevisiae*.

Cells containing plasmid with or without the *HDE* insert were cultured in media containing glucose or oleic acid. Cell lysates were fractionated by electrophoresis, and immunoblotted (see Fig. 4). Cells transformed with the control vector YCp50 contained no immunogenic protein of the expected molecular weight. Cells transformed with YCp50H2 produced detectable amounts of *HDE*. Not only was *HDE* expressed in the cells, but it also was expressed in a regulated fashion, for when cells were cultured in media containing oleic acid as opposed to glucose, a 2.7-fold increase in the level of *HDE* protein was observed. Therefore *HDE* was both expressed and induced in *S. cerevisiae*, and the antibodies directed against *HDE* could be used to assay levels of *C. tropicalis* *HDE* because they did not cross-react with endogenous *HDE*. This inability to detect the *HDE* homologue of *S. cerevisiae* was rather unexpected, since these cells are known to contain an endogenous trifunctional enzyme (Einerhand

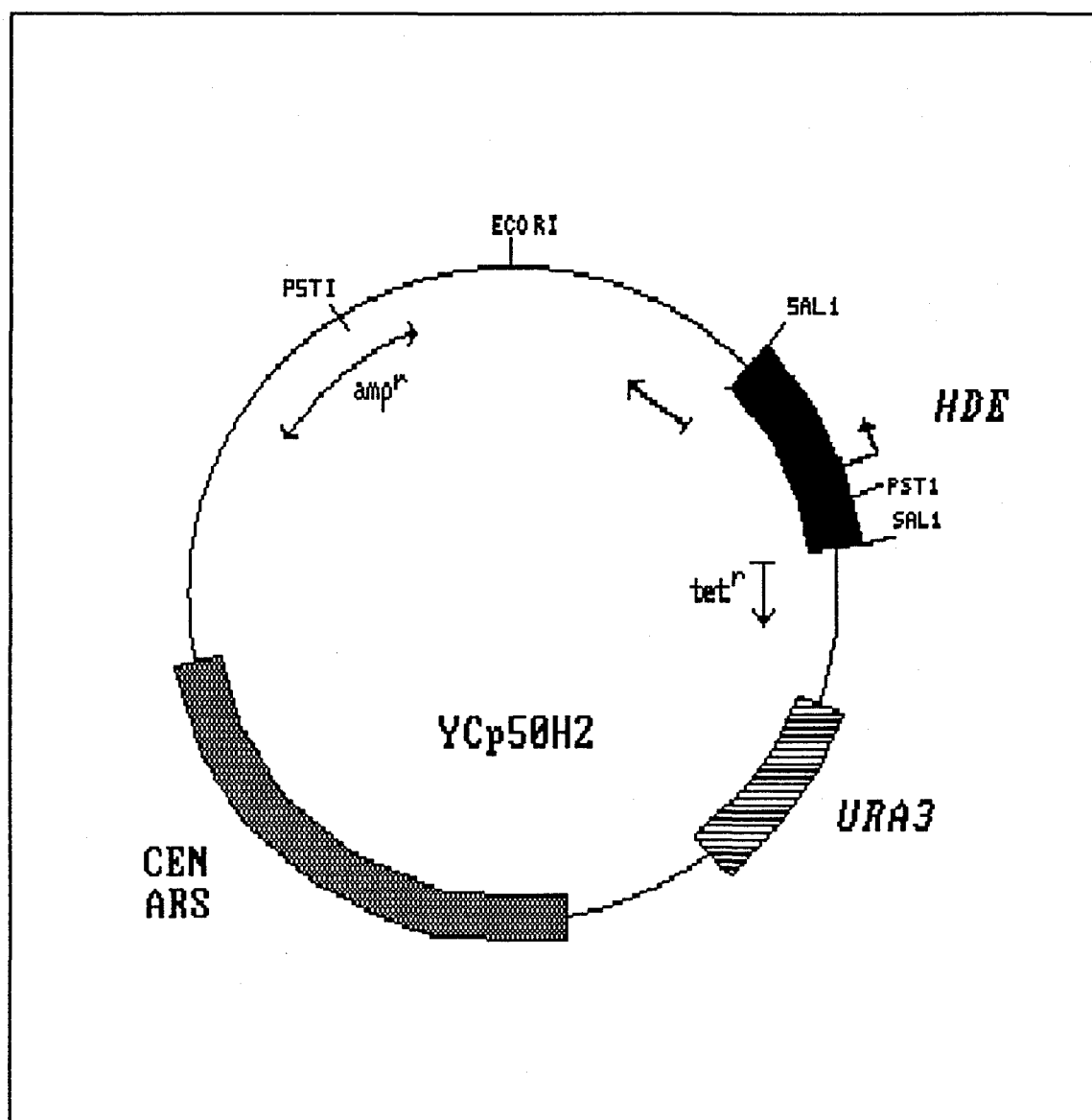


Figure 3. Subcloning of the *HDE* gene into the plasmid YCp50. Plasmid 3T was cut with *Sal*I to excise the *HDE* gene (see Fig. 6). *HDE* was inserted into the *Sal*I site of YCp50, insertionally inactivating the tetracycline resistance gene. Selection for DH5- $\alpha$  cells which were resistant to ampicillin and sensitive to tetracycline resulted in a transformant harbouring the recombinant plasmid YCp50H2.

*et al.*, 1991).

One antigen at a much lower molecular weight was recognized by the HDE antiserum (see small left pointing arrow, Fig. 4). This antigen is regulated in a fashion opposite to that of HDE. The concentration of this cross-reacting species was increased in cells cultured in media containing glucose and non-detectable in cells cultured in media containing oleic acid (compare lanes G and O).

HDE contains a duplicated 280-residue domain of unknown function (Nuttley *et al.*, 1988). These two domains and the linker sequence joining them make up 65% of the total residues of the protein. Upon further study of this region Baker (1990) discovered that these two domains bear striking similarity to oxidoreductases that are involved in the synthesis of antibiotics and the metabolism of toluene, alcohol and sugars. One of the enzymes examined, namely glucose dehydrogenase, has 53 (24%) positions along the length of this conserved domain where trios of amino acids are identical to those observed in the duplicated region of the HDE enzyme (Baker, 1990). The author therefore speculates that it is this region of HDE that contains the dehydrogenase activity of the enzyme. It is interesting to also speculate that this cross-reacting protein contains a dehydrogenase epitope similar to that of HDE and those found for the above family of oxidoreductases, and it is this region that the HDE antiserum chiefly recognizes.

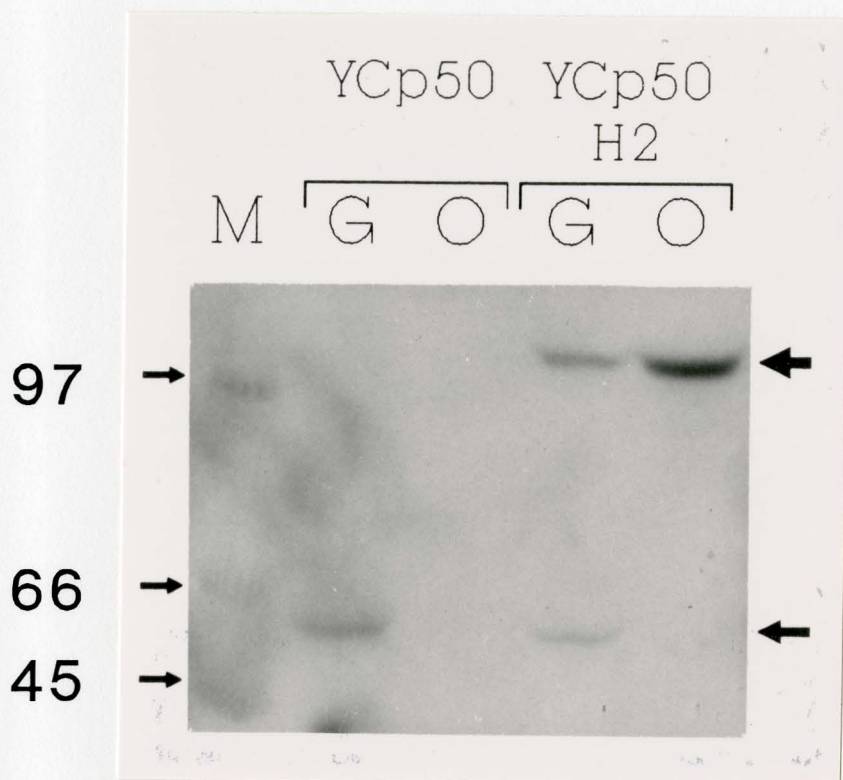


Figure 4. Expression and induction of *HDE* in *S. cerevisiae*. *S. cerevisiae* was transformed with YCp50H2 (Fig. 3) or YCp50 by the method of Ito *et al.* (1983). Transformants containing YCp50 or YCp50H2 were cultured in media containing either 2% glucose (G) or 0.1% oleic acid (O). Lane (M) contained  $^{14}\text{C}$ -methylated proteins (molecular weight range 14,000-70,000 Da; phosphorylase b, 97,400 Da) with the location of each species indicated by an arrow preceded by the size of the band in kDa. The cross-reacting species is flagged by the small arrow on the right. Position of signal attributed to presence of HDE is indicated by the large arrow on the right hand side of the figure. Lysates of transformants were fractionated in the four lanes to the right of the molecular weight markers. Lysates of cells containing plasmid YCp50 were fractionated in the first two lanes adjacent to the marker lane. Lysates of cells containing plasmid YCp50H2 was fractionated in the last two lanes. Equal amounts of protein were fractionated in each lane.

Western blotting results, in conjunction with the observation that this species is absent when cells are cultured in oleic acid, and present when cells are cultured in glucose, suggests that this cross-reacting species functions in fermentive metabolism as a dehydrogenase.

### 3.3) Primer extension of *HDE* mRNA

Total RNA was isolated from cells (see section 2.3.1.13) and used to extend the primer AB1069 to map the 5' mRNA start site of the *HDE* gene. Many start sites were found, (perhaps AMV reverse transcriptase pause sites), but only one major start site was mapped at 60 nt upstream of the A of the ATG translational start (see large arrow in Fig. 5). This start site is 107 nt downstream of the first T of the TATA box (see Fig. 7), a position within the usual distance of 60 to 120 nt from the TATA box (Guarente, 1987). This initiation site matches neither of two common consensus sequences that are found at more than 50% of all start sites in yeast, namely, RRYRR (where Y = pyrimidine and R = purine) or TC(G/A)A; Guarente, 1987.

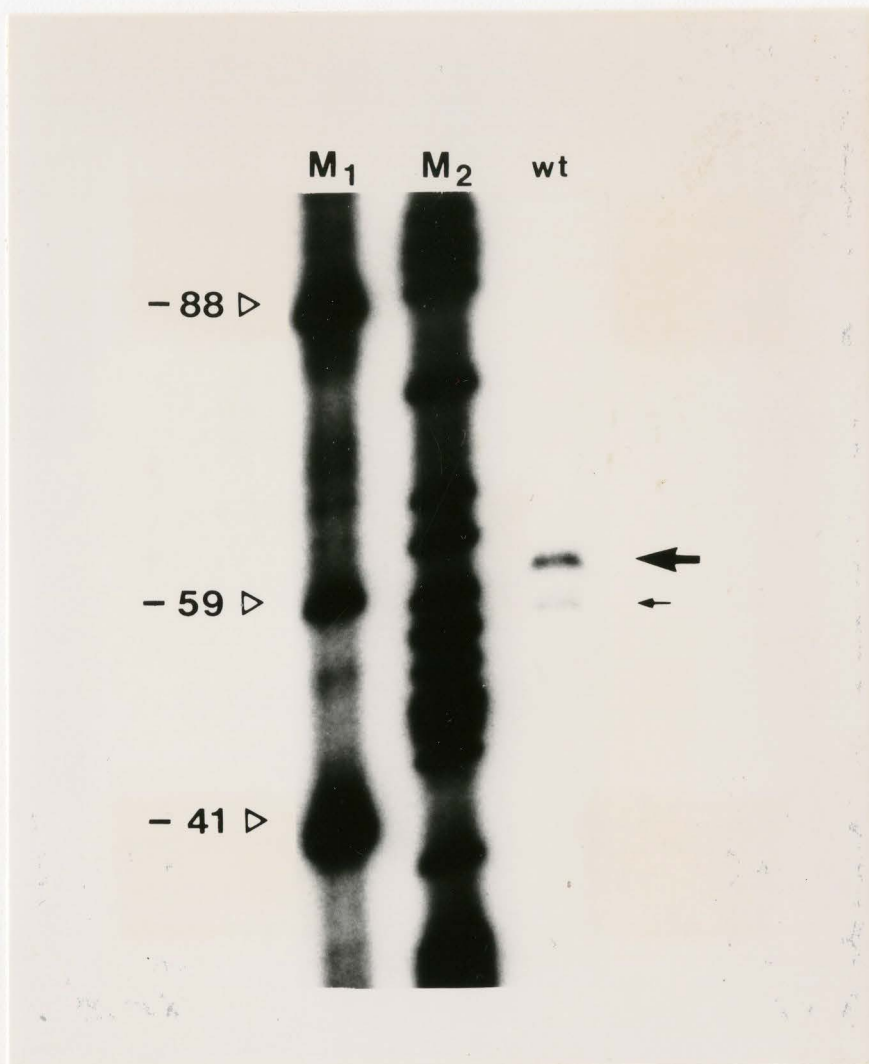


Figure 5. Mapping of the RNA transcription start point of *HDE*. Total RNA isolated from *S. cerevisiae* was hybridized with primer AB1069 and extended as described in section 2.3.1.14. The products of this extension were fractionated in the lane marked (wt). A sequencing ladder was generated using the same primer bound to plasmid 3T and extended with T7 DNA polymerase. The lanes marked M<sub>1</sub> and M<sub>2</sub> correspond to the C and T terminal nucleotides of the sequencing ladder respectively. Numbering is given as negatives, indicating nt positions upstream of the A of the start codon ATG. The large arrow indicates the major transcription start point, while the smaller arrow beneath it indicates the minor transcription start point (see Fig. 7 for positions in the *HDE* sequence *i.e.* +1 major transcription start point and +4 minor transcription start point).

### 3.4) Sequencing of the *HDE* gene

In order to characterize and localize the regulatory regions of the *HDE* gene, it was necessary to sequence the available upstream region of a 4.2 kbp genomic clone containing *HDE*. Partial sequencing in the 5'→3' direction of the gene (direction of transcription) had been completed earlier by Dr. Aitchison, using Exo III-generated deletions (Fig. 6). In order to complete the sequencing of this direction, specific Exo III digestion reactions, which had been prepared and stored at -20°C by Dr. Aitchison, were rethawed and used to transform DH5- $\alpha$  cells. Upon analysis of approximately 80 transformants, two useful clones were found carrying constructs that could be used to fill in the gaps of the sequence information (see large arrows, Fig. 6).

The sequences downstream of nucleotide -166 were obtained by generating a clone which had the *Pst*I-fragment removed from the *HDE* gene, thereby placing nucleotide -166 downstream of the 3' end of the M13 forward primer-binding site (see box initiated arrow below *Pst*I recognition site in Fig 6). This clone provided the necessary sequence information to complete the sequencing of the entire *HDE* gene, the sequence of which is presented in Fig. 7. The only area not sequenced in this direction was the last 100 nt, approximately 500 nt downstream of the mRNA termination site.

The opposite strand was sequenced using a series of exonuclease III-generated deletion constructs of plasmid 3T

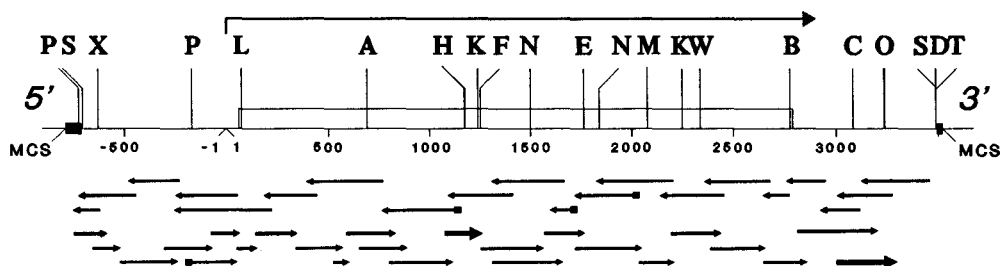


Figure 6. Sequencing strategy used to sequence the *HDE* gene. This map illustrates the position of the *HDE* gene as it is found in plasmid 3T. The line, divided by the numerical values and flanked by the MCS sequences, corresponds to the *HDE* gene cloned into pGEM5ZF(+) from a  $\lambda$ gt11 isolate. The numbering of the nucleotides is as in Fig. 7. The open box indicates the ORF of the gene. The direction and length of the transcript is illustrated by the arrow above the ORF. The orientation of the gene is indicated by both the direction of the arrow over the ORF and the italicized 5' and 3' symbols to the left and the right of the gene respectively. The bold uppercase letters interspersed above the representation of the genomic sequence of *HDE* mark the locations of specific restriction endonuclease recognition sites as follows:

A = <i>Apa</i> I	B = <i>Bgl</i> II	C = <i>Cla</i> I
D = <i>Nde</i> I	E = <i>Eco</i> RI	F = <i>Eco</i> RV
H = <i>Hind</i> III	K = <i>Kpn</i> I	L = <i>Aha</i> I
M = <i>Pfl</i> MI	N = <i>Nco</i> I	O = <i>Xho</i> I
P = <i>Pst</i> I	S = <i>Sal</i> I	T = <i>Sac</i> I
W = <i>Alw</i> NI	X = <i>Xba</i> I	

The M13 forward primer was used to generate sequences mapped by the right pointing arrows, and the reverse primer used to generate sequences mapped by the left pointing arrows. The two large right pointing arrows identify the sequences used to complete the sequencing in the 5'→3' direction, which were obtained from deletion constructs generated by Dr. J.D. Aitchison. The arrows initiating from small boxes identify the sequences obtained using restriction endonuclease-generated 'drop out' constructs.



Figure 7. Nucleotide sequence of the gene encoding peroxisomal trifunctional enzyme. Sequences were determined by dideoxysequencing methods using deletion constructs (see sections 2.3.1.8 and 2.3.1.9). The nt +1 and arrow in bold indicate the location of the major transcription start point, and the other arrow a minor transcription start point. The anchored arrows indicate the endpoints of the named deletion constructs. The sequence in bold type and underlined is similar to the glucose-responsive element of *SUC2*. The single underlined sequence corresponds to a consensus nt sequence implicated in mediating transcriptional responses to the presence of oleic acid. The sequence in bold maps a consensus to the TATA box. The bold typed, doubly underlined codons indicate the start codon at the beginning of the ORF, and the two consecutive stop codons at the end of the ORF. The bold typed, lowercase **t** at position 2049 is the only change (C→T) observed in the ORF between the cDNA sequence and the genomic sequence.

↗5WT  
 GTCGA CGTTTTGCCG CCACCTGTGA AGGGAAGAAG AGTCGTTGAG TTGATGTAAT TAAGCTGGCA CGTAGATACC AGAAGGTTCT AGAGTAGAGC  
 -710 -700 -690 -680 -670 -660 -650 -640 -630  
 TTGGGTGGTG TTTGGCCCTG TTTGGACCAC GGATAGAGAT GGAGAATCCC TTGGTTAGAG CGGAGAGGAA AAAATTGAAA CTTTGCATAT CCCACTTCAT  
 -620 -610 -600 -590 -580 -570 -560 -550 -540 -530  
 TATCCTTGAT GTAACCGTTT TATGGGGTAA TAAAGTGTG GAAAAATAAT CAGGGAGACA TATTCCCGAT CAATTGGGTG GTGGTCGCTC AATTTCTGTG  
 -520 -510 -500 -490 -480 -470 -460 -450 -440 -430  
 AGTAGTAGGC TCAGTGGTGT GTATTGGGAT TGGTAGTAGT CTGTATAAGC AGTGTATAT AACCCATTGC TTGTGTATTC CTATTTTGTCT GGCAAAAGTG  
 -420 -410 -400 -390 -380 -370 -360 -350 -340 -330  
 ACAACTGTAG TTGTGAGATA ATCCTCGGTT ATTACGCCTG GGGGGGCGA CAGCCAAAGT TGTGCCCGTG CGACAATGGC ATCAGAAGAA ACAGAAAAAA  
 -320 -310 -300 -290 -280 -270 -260 -250 -240 -230  
 AAAACACAGG CATTITATC CACATGCACA CTACCCCCAC TATTCTGTG TGCAGTGTGC TTGTGTGTGG CCCCCGCGAG AATCAACAGG GCAAACCTCTG  
 -220 -210 -200 -190 -180 -170 -160 -150 -140 -130  
 GAGCCTGAAT CTTTATATAA ACTTCAGGCA TTGGCCCCC TTTTACAAT TCTTCACATC CACCATTTTT TTTCTTCTTT CCTACCATAT TAGTTTTTTT  
 -120 -110 -100 -90 -80 -70 -60 -50 -40 -30  
 TTATTCTTTT CCTACCTATC TGATTATTAT CAAACATCTG GTCATCCTCA AAAGAAAGAA AGAAACTATA ACAATCAATC  
 -20 -10 +1 11 21 31 41 51

61

ATGTCCTCAGTTGATTTTAAAGATAAAGTTGTGATCATTACCGGTGCCGGTGGTGGTTGGGTAATACTACTCCCTCGAATTTGCCAAGTTGGGCGCCA 160  
 AAGTCGTCGTTAACGACTTGGGTGGTGCCTTGAACGGTCAAGGTGGAACTCCAAGGCCGCGCAGCTTGTCTGTGACGAAATTTGCAAGAACGGTGGTGT 260  
 TGCCGTTGCCGATTACAACAACGCTTGGACGGTGAACAAGATTGTGAAACCCGCCCTCAAGAACTTTGGTACTGTCCACGTTATCATCAACAATGCCGGT 360  
 ATCTTGAGAGATGCCCTCCATGAAGAAGATGACTGAAAAAGACTACAATTTGGTCACTTGAACGGTGCCTTTGCCGTCACCAAGGCTGCTT 460  
 GGCCATACTTCCAAAAGCAAAATACGGTAGAATTGTCAACACATCCTCCCGAGCTGGTTTGTACGGTAACTTTGGTCAAGCCAACCTACGCCCTCCGCCAA 560  
 GACTGCTTTGTTGGGATTCGCTGAAACCTTGGCCAAGGAAGTGGCAAAATACAACATCAAGGCCAACGCCATTGCTCCGTTGGCCAGATCAAGAATGACT 660  
 GAATCTATCTTGGCCACTCCAATGTTGAAAAAATGGGCCCTGAAAAGGTTGCCCATTTGGTCTTGTATTTGTCGTGAGCTGAAAACGAATGACTGGTC 760  
 AATCTTTGAAGTTGCTGCTGGCTTTTACGCTCAGATCAGATGGGAAAGATCCGGTGGTGTCTTGTTCAGCCAGATCAATCCTTACCAGCTGAGGTTGT 860  
 TGCTAAGAGATTCTCTGAAATCCTTGATTATGACGACTCTAGGAAGCCAGAATACTTGAAGAACAATACCCATTTCATGTTGAACGACTACGCCACTTTG 960  
 ACCAACGAAGCTAGAAAGTTGCCAGCTAACGATGCTTCTGGTCTCCAACCTGTCTCCTTGAAGGACAAGGTTGTTTGTATCACCAGTGGCCGGTGGTGGT 1060  
 TGGTAAAGAATAACGCAAGTGGTTCGCCAAGTACGGTGGCAAGGTTGTTGTTAACGACTTCAAGGATGCTACCAAGACCCTTACGAAATCAAAGCCGC 1160  
 TGTTGGTGAAGCTTGGCCAGATCAACACGATGTTGCCAAGGACTCCGAAGCTATCATCAAGAATGTCATTGACAAGTACGGTACCATTGATATCTTGGTC 1260  
 AACACGCGCGTATCTTGAGAGACAGATCCTTGGCAAGATGTCCAAGCAAGAATGGGACTGTGTCACCAAGTCCACTTGATTGGTACTTTCAACTTGA 1360  
 GCAGATTGGCATGGCCATACTTTGTTGAAAAACAATTTGGTGAATCATCAACATACCTCCACCAGTGGTATCTACGGTAACTTTGGTCAAGCCAACTA 1460  
 CTCGTCTTCAAGGCTGGTATCTTGGGTTTGTCCAAGACCATGGCCATTGAAGGTGCTAAGAATAACATTAAGGTCAACATTTGTTGCTCCACACGCTGAA 1560  
 ACTGCCATGACCTTACCCTCTTCAAGAGAACGAAGAAGTGTACCACGCTGACCAAGTGTCTCCATTGTTGGTCTACTTGGGACTGACGATGTCC 1660  
 CAGTCACCGGTGAAACTTTCGAAATCGGTGGTGGTGGATCGGTAAACCCAGATGGCAAAGGCCAAGGTTGCTGTCTCCACGACGACACACCCTGT 1760  
 TGAATTCATCAAGGAGCACTTGAACGAAATCACTGACTTCAACACTGACTGAAAAATCCAAAATCTACCACCGAATCCTCCATGGCTATCTTGTCTGCC 1860  
 GTTGGTGGTATGACGATGATGATGACGAAAGCAAGCAAGGATGATGAAGAAGAAGCAAGAAGCAAGAAGCAAGAAGCAAGATCCAGTCTGGA 1960  
 GATTGACGACGAGATGTTATCTTGTACAACATTTGCCCTTGGTGGCCACCACCAAGCAATTTGAAGTACGCTACGAAAAACGACTCTGAETTCAGGTCAT 2060  
 TCCAACCTTTGGTCACTTGTACCTTCAACTCTGGTAAAGTCAAAAACCTCTTGGCAAGTGTGCGTAACTTCAACCCAATGTTGTTGTTGCACGGT 2160  
 GAACACTACTTGAAGGTGCACAGCTGGCCACCACCAACCGAAGGTGAAATCAAGACCACTTTCGAACCAATTTGCCACTACTCCAAGGGTACCAACGTTG 2260  
 TTATTGTTACGGTTCCAAATCTGTTGACAACAAGTCTGGTGAATGATTACTCCAACGAAGCCACTTACTTTCATCAGAACTGTCAAGCCGACAAACA 2360  
 GGCTACGCTGACCGTCCAGCATTGCCACCAACCAATTTCTGGCACCAGAGAGAGCCAGACTACCAAGTTGACGTTCCAGTCAAGTGAAGACTTGGCT 2460  
 GCTTTGTACGGTTTGTCTGGTGGACAGAAACCCATTGCACATTGATCCAACTTTGCTAAAGGTGCCAAGTCCCTAAGCCAATCTTACACGGTATGTGCA 2560  
 CTTATGGTTTGGTGTAAAGCTTTGATTGACAAGTTGGTATGTTCAACGAAATCAAGGCCAGATTACCCGGTATTTGCTTCCAGGTGAAACCTTGG 2660  
 AGCTTGGCATGGAAGGAAAGCGATGACACTATTGCTTCCAACTCATGTTGTTGATAGAGTACTATGCCATTAACAACGCTGCTATTAAGTTAGT 2760  
 GGTGACAAGCAAGATCTAATGAAGTGTGTTTCTCCTGTCGTTTCGTTTTTTATTATTATTATAAATATATATCTTAGTATCAAAATGTAACCTTAA 2860  
 ATATAAATATATAGTTTGTCTTGGACAGAGATCATCTGGGGTAGATGTACATAATCGTCTGATTATTCCTTTCATCAACTGCAAACTATACATTT 2960  
 CTCTGTTCTCGGAGATTTGACGAAAGCAACGATCCTGAACGTTTGTGGAGCTTTGACTTGGAACTTATGGGTAAAGTACGGCACCAGGAAATCAAC 3060  
 AGGGGTATCATTAAAGGATCGATTGCAACTGCTGGTGTGCTGCGAGGAGCTGGTAGCAGGGGTCAAGAATGGCAGTGCATTGGACCACCGCCGTTG 3160  
 TTAAGTGTGTTTATAGACGTCGGTGTGGCACTTGGAGACTGAAAGTACGTTGACGGGGTAGAGAATAATTTGCTCGAGGACCAGATTTAGGGAAC 3260  
 GGAAGTGTAGTTGTGGTGTGACTGTGATTGCTGAATGTTCTTGAACCTTGTCTTCTGTGGAAGTGAAGCAGCGATGGCTTCTGGCATCAAACTATC 3360  
 TGCTGAGTCAATGACGGGTGACGACGACGATGTTATGGCTCGTCAAGCTCCACTCGATTGAGCCTGAATCGGAACATGCAGATGATGAAGCCACCGAC 3460  
 GTGATGGAGTTAATGGACGTC 3481

cut with the enzymes *NdeI* and *SacI* (see section 2.3.1.8.1). In order to fill in some of the remaining gaps, deletion 'drop outs' were generated by cutting the plasmid 3T with the restriction endonucleases described in section 2.3.1.8.1 and fusing the resultant plasmids (see box-initiated right pointing arrows in Fig. 6). These constructs enabled completion of the sequencing of the *HDE* gene in the 3'→5' direction.

### 3.5) Expression of *HDE* deletion constructs in *S. cerevisiae*

In order to localize, and eventually isolate the upstream elements involved in the regulation of the *HDE* gene, it was necessary to introduce the gene into the genetically manipulable yeast *S. cerevisiae*. The plasmid chosen as a vector for the *HDE* gene was the *CEN*-containing plasmid YCp50. *HDE* was subcloned into YCp50 producing plasmid YCp505H (see Fig. 8 below and section 2.3.3.4) and then transformed into *S. cerevisiae* (see section 2.3.5.1).

The YCp50 vector assures plasmid stability and an average gene dosage of one copy per cell (Struhl, 1983). This allows scientists to mimic the natural *in vivo* scenario existing within all cells, namely, one copy of each gene per haploid genome. Other multicopy plasmids (such as *ARS* based-plasmids) can potentially distort transcriptional studies because of uncontrollable and unpredictable effects due to copy number, plasmid loss, and abnormally high expression

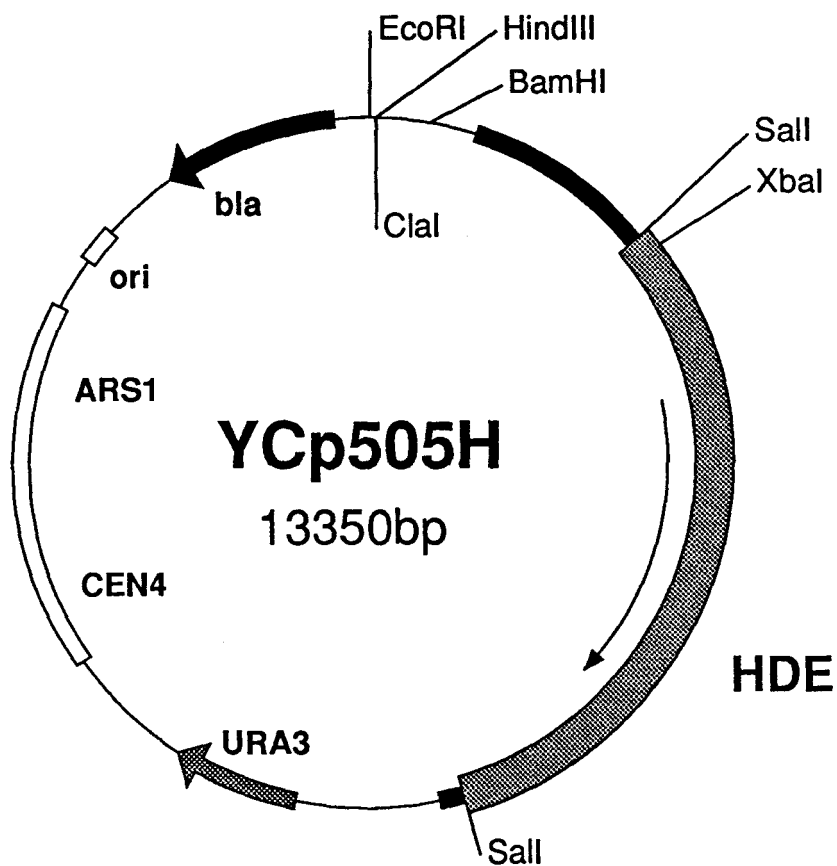


Figure 8. Subcloning of the *HDE* gene for expression in *S. cerevisiae*. A 4.2 kbp *Sal*I fragment containing the *HDE* gene (Aitchison *et al.*, 1990) was ligated into *Sal*I-digested pGEM5zf(+) to produce recombinant plasmid 3T. YCp505H was constructed by ligating the *Sca*I-*Sac*I fragment containing the *HDE* gene from 3T into the *Sal*I site of YCp50. 1.8 kbp of pGEM5zf(+) (closed box) flanked the 5' end of the *HDE* gene to act as a buffer for subsequent *Bal* 31 digestion. The thin black arrow indicates the direction of transcription of the *HDE* gene.

levels of genes. A consequence of increased gene copy number is that these genes squelch important transcription factors required by many other genes. As a result, not only are the number of transcripts per template reduced, but gross influences on transcriptional rates of other similarly regulated genes also occurs. This influences many other metabolic pathways. Excessive protein concentrations, another consequence of increased gene copy number, can cause the formation of inactive, non-functional protein aggregates (Binder *et al.*, 1991). All these influences and variables make interpretation of data rather difficult.

YCp50 was tested for rate of plasmid loss from cells in non-selective media, and demonstrated a rate of 3.4% per generation. Interestingly, Rose *et al.* reported that larger *CEN* plasmid derivatives are more mitotically stable than smaller ones (Rose *et al.*, 1987). In fact, we observed just that. The plasmid YCp50H2 had a plasmid loss rate of only 1.3% (see Table 3).

Transformed *S. cerevisiae* could have been cultured in non-selective rich media and maintained the plasmid, but nevertheless, selective media were used. The reason for this is that rich media contain not only nutrients, but also additional carbon sources, *e.g.* yeast extract contains glucose at 0.5%, carbon sources which could influence gene regulation and thereby complicate the results and interpretation of data. Unfortunately, the selective media

PLASMID	YEPD	MEAN	YNB	MEAN	%LOSS/12 GENERATIONS	%LOSS/GENERATION
	COLONIES		COLONIES			
YCp50	181	189 ± 8	125	126 ± 9	33.5%	3.4%
	197		117			
	189		135			
YCp50H2	285	295 ± 13	247	253 ± 5	14.5%	1.3%
	291		254			
	310		257			

**TABLE 3** PLASMID LOSS OF YCp50 BASED PLASMIDS

Cells containing plasmids were cultured for 12 generations in YEPD media. Appropriate dilutions were made, and equivalent aliquots of culture were plated on selective (YNB) or non-selective (YEPD) plates. The number of colonies present on the plates after 3 days was recorded in the column describing the media.

used in these experiments were more challenging as far as culturing yeast cells was concerned, because, doubling times were lengthened and cell yields reduced.

The only ingredient in the YNE medium which could compromise its selective nature was the 0.05% yeast extract. Erdmann *et al.* (1989) had demonstrated that this amount of extract was required in order for *S. cerevisiae* to grow in oleic acid. Yeast extract alone, even in the presence of 0.5% (w/v) Tween 40, did not promote cell growth (Erdmann *et al.* 1989), unlike that seen above for *C. tropicalis* (see Fig 1, lane -, for results of *C. tropicalis* enzyme expressions in presence of Tween 40 alone).

Deletions (Fig. 9 and Fig. 7) in the upstream region of *HDE* were constructed to delineate the 5' elements mediating the responses detailed above. Upstream of all the deletion endpoints were DNA sequences of bacterial origin. Such sequences would not be expected to contain regulatory sequences that would influence eukaryotic transcription. The possibility that cryptic eukaryotic DNA elements might exist in these regions would be remote. Therefore the effects observed with various deletion constructs were attributed to the deletions themselves, and not to positional effects of transcription elements in the *HDE* promoter with cryptic elements located in these flanking bacterial upstream sequences.



Figure 9. Generation of 5' deletion constructs. Bal 31 exonuclease was used to produce deletions in the upstream region of the *HDE* gene. Bal 31 deletion end points were determined by dideoxy sequencing (Sanger et al., 1977). Numbering is as in Fig. 7.



*HDE* expression of various deletion constructs in *S. cerevisiae* was assayed by immunoblotting cellular lysates with anti-HDE serum. This serum contained an antibody specific for the heterologously synthesized HDE. The assumption made in these studies is that protein levels are indicative of increases in transcriptional activity. This assumption is valid if no excessively variable post-transcriptional or translational controls exist to further regulate the protein concentrations of the  $\beta$ -oxidation enzymes. This is a rather plausible assumption, especially if one considers that the purpose of increasing transcriptional rates is to increase the levels of translatable template and therefore levels of functional enzyme. Therefore, assaying protein concentrations should provide reasonable data about rates of transcription in various media, and, in conjunction with deletion constructs of upstream regulatory regions, information about the location of the *cis*-transcriptional elements that mediate and elicit the observed regulatory responses.

The removal of sequences from nt -715 (5WT) to nt -553 (5R) had little effect on the levels of *HDE* expression in either low concentrations of glucose or glucose-oleic acid (Fig. 10, compare lanes 5WT and 5R). These sequences contained elements such that their collective removal served to increase *HDE* expression 1.4-fold and 1.7-fold for cells cultured in low concentrations of glucose or glucose-oleic acid respectively. An ADR1 binding site (-581;TGGAGAA;-575)

was found in this region, which matched the core region shown to be protected from DNase I by ADR1. This factor has been demonstrated to regulate the expression of genes encoding peroxisomal proteins (Simon et al., 1991). However, very slight effects on expression of genes encoding peroxisomal proteins were seen in *adr1* cells cultured in media containing oleic acid, and no effects were seen on basal expression when cells were cultured in glucose. Expression of genes encoding peroxisomal proteins were reduced only in *adr1* mutants cultured in media containing ethanol (*i.e.* a non-fermentable carbon source). Therefore, it is not entirely unexpected that these sequences between nt -715 and -553 (even though they contain an ADR1 binding site) have no effect on the observed regulation of *HDE*, because the media used in these experiments contained glucose, which is known to repress ADR1 function. An interesting experiment, which would conclusively indicate the functional presence of an ADR1 binding site in this region, would be to culture the yeast cells containing these deletion constructs in media containing ethanol only. Differences in expression observed between construct 5WT and 5R would be likely attributable to this potential binding site.

Deletion of the next 53 nt to -500 (5E) combined to give a 3-fold and 2-fold increase in *HDE* expression when cells were cultured in low concentrations of glucose and glucose-oleic acid respectively, indicating the presence of additional

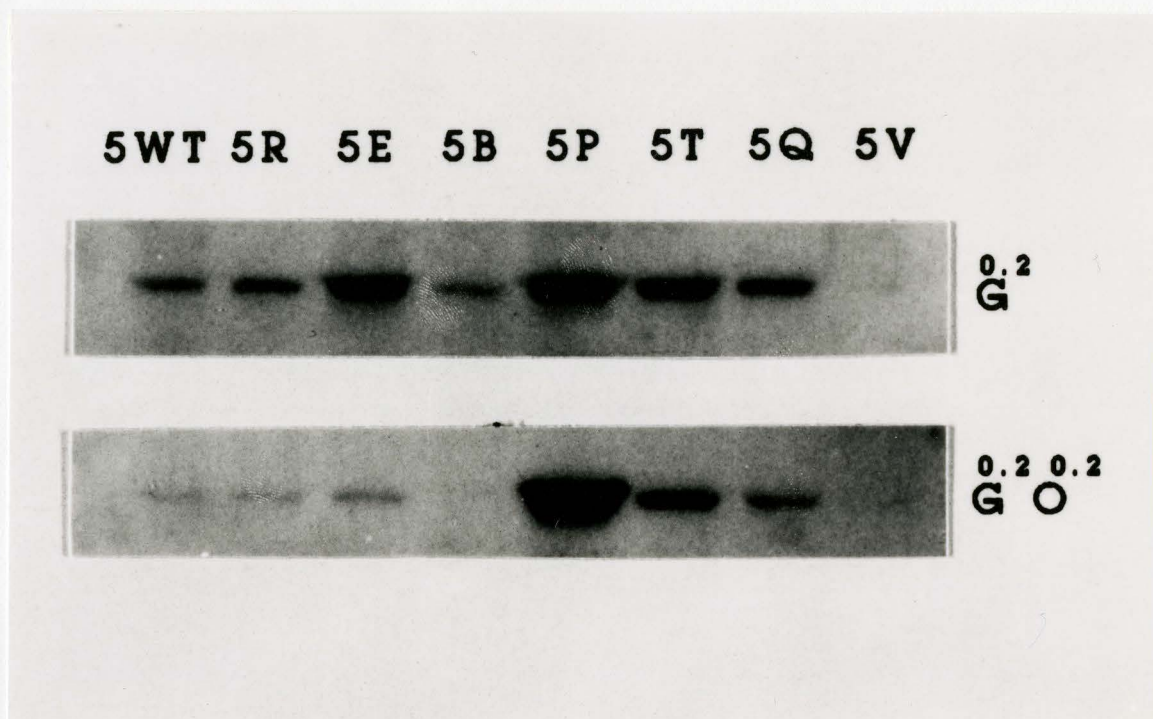


Figure 10. Expression of deletion constructs in *S. cerevisiae*. *S. cerevisiae* DL-1 (Van Loon *et al.*, 1983) was transformed by the method of Ito *et al.* (1983). Transformants harboring various deletion constructs were grown in YNE supplemented with leucine (30  $\mu\text{g}/\text{mL}$ ) and histidine (20  $\mu\text{g}/\text{mL}$ ) containing either low glucose (upper panel,  $G^{0.2}$ ) or glucose-oleic acid (lower panel,  $G^{0.2}O^{0.2}$ ; see Fig. 2 for details of media composition and abbreviations), lysed and analyzed for HDE expression by immunoblotting. Lanes of each panel contain equal amounts of protein.

DELETION CONSTRUCT	5WT	5R	5E	5B	5P	5T	5Q	5V
MEDIA	YNE PLUS 0.2% GLUCOSE							
	RELATIVE DENSITOMETRIC SIGNAL							
BLOT # 1	1.0	1.0	2.7	0.8	6.8	3.0	2.0	0.2
BLOT # 2	1.0	2.4	3.3	0	6.5	1.8	1.7	0.5
AVERAGE	1.0	1.7	3.0	0.4	6.7	2.4	1.9	0.4
STANDARD DEVIATION	± 0	± 1.0	± 0.4	± 0.6	± 0.2	± 0.8	± 0.2	± 0.2
MEDIA	YNE PLUS 0.2% GLUCOSE, 0.2% OLEIC ACID							
	RELATIVE DENSITOMETRIC SIGNAL							
BLOT # 3	1.0	1.3	2.1	-	34	7.0	1.5	-
BLOT # 4	ND	1.4	2.0	-	20	4.0	2.7	-
AVERAGE	1.0	1.4	2.0	0	27	6.5	2.1	0
STANDARD DEVIATION		± 0.1	± 0.1		± 10	± 2.1	± 0.8	

TABLE 4: Relative densitometric signals obtained from autoradiographic films. Blots numbered 1 and 3 are Western blots which gave rise to the photoreproductions seen in Fig. 10. Blots numbered 2 and 4 are independent Western blots; photoreproductions not shown. ND = not determined. - = not detectable. Each band was scanned in triplicate and the average value recorded in the table under the corresponding deletion construct heading. Averages and standard deviations obtained from independent experiments are tabulated to the left of the labels "AVERAGE" and "STANDARD DEVIATION"

repressors between -715 and -500 (Fig. 10, compare lanes 5E and 5R to lanes 5WT).

Repressor binding sites have been found in many yeast genes, and serve to moderate gene expression in a negative fashion (Hodge et al., 1990; Nishizawa et al., 1989; Kronstad et al., 1987; Luche et al., 1990; Brindle et al., 1990; West, Jr., et al., 1987). Studies with the *GAL1-GAL10* divergent promoter have indicated how they function, namely, by interfering with activator binding and/or function (Finley and West, Jr., 1989). Another sequence with repressive properties is located in the region between nt -466 and -333. Removal of these 133 nt caused a 6.6-fold increase and a 36-fold increase in *HDE* expression relative to 5WT in glucose and glucose-oleic acid, respectively (Fig. 8, compare lanes 5P to lanes 5WT). Two sequences within this region are conserved in a number of oleic acid-responsive genes of *C. tropicalis* (see Table 5 and Fig. 7). One of these sequences is similar to the consensus sequence  $\left[ \begin{array}{c} \text{ATTTC} \\ \text{TT} \end{array} \text{CG} \right]$  for regulation of *SUC2* by glucose in *S. cerevisiae* (Sarokin and Carlson, 1986). However, this sequence is more involved in derepressive expression, as a positive element for the *SUC2* gene. This element, when fused to a reporter gene, mediated a two to three-fold decrease in expression when glucose was added to the media. This glucose response is most likely mediated by the lack of an activator's presence\potency. The same sequence was observed flanking a *MIG1* binding site in the upstream region of the *SUC2* gene

(Nehlin and Ronne, 1990). *MIG1* has been identified as a glucose repressor, and therefore it possibly competes with an activator for binding to this site. Yet no consensus sequence for the *MIG1* site was observed in the upstream region of *HDE*. However, another highly conserved sequence found in this upstream region of *HDE* and other oleic acid-responsive genes matched a small portion of a sequence observed to have glucose-repressive properties in the upstream region of *SUC2* (Sarokin and Carlson, 1984). This purine-rich sequence, located from -142 to -133 in *SUC2*, is GGAAGAAAGA; a good match to the sequence seen in the oleic acid-responsive genes namely -468  $\left[ \begin{array}{c} \text{GGGAGAGATA} \\ \text{C} \quad \text{A} \end{array} \right]$  -459. One can go on to speculate that these common sequences may function in a similar fashion in all the genes that carry them.

Sequences between -500 (5E) and -466 (5B) regulate positively *HDE* expression, because when these sequences are removed, levels of *HDE* expression decreased to approximately 40% of 5WT in cells cultured in low concentrations of glucose, and to below limits of detection for cells cultured in glucose-oleic acid (Fig. 10, compare lanes 5B to lanes 5E and 5WT). Again a consensus sequence common to the upstream regions of the genes *AOX*, *CATL*, *HDE* was located in the region flanked by the deletion endpoints of -500 (5E) and -466 (5B). The consensus sequence  $\left[ \begin{array}{c} \text{TGTGCAAAAAATAAAC} \\ \text{C} \quad \text{ATCG} \end{array} \right]$  (-484 to -470) can be found with a few variations at least twice in each upstream region (see Table 5). The removal of this region

caused noticeable effects in cells grown in media containing glucose-oleic acid and in media containing just low concentrations of glucose. Therefore this repeated element is a strong candidate for a derepressive element responding to low concentrations of fermentable carbon sources and possibly the presence of an alternative carbon source, because when this region is removed, the response to the presence of oleic acid is reduced.

*HDE* expression in 5P (-333) and 5T (-281) is 6.7-fold and 2.4-fold that of 5WT, respectively, in low concentrations of glucose, and 27-fold and 6.5-fold that of 5WT, respectively, in glucose-oleic acid (Fig. 10, compare lanes 5P, 5T, and 5WT). This indicates that this region contains an oleic acid-responsive element, because the ratio of 5P/5T (relative to 5WT) is 1.5-fold higher for cells cultured in glucose-oleic acid versus low concentrations of glucose. Therefore, sequences unique to 5P, *i.e.* -333 to -281, positively mediate responses to oleic acid. A nonanucleotide sequence [consensus =  $\begin{array}{c} \text{TGGTT} \\ \text{C} \end{array} \text{ATT} \begin{array}{c} \text{A} \\ \text{G} \end{array} \text{TTA}$ ] between nt -333 and -281 of *HDE* is conserved in oleic acid-responsive genes of *C. tropicalis* (see Table 5 and Fig. 11). This sequence may be a binding site for an oleic acid-responsive transcriptional activator. This is also supported by reports by Einerhand *et al.* (1991). They describe a sequence they call the " $\beta$ -oxidation box", which they observed in three *S. cerevisiae* genes. This  $\beta$ -oxidation box closely resembles (*i.e.* 89% identity) the sequence

TABLE 5

Sequences of the responsive regions of the *HDE* gene conserved in other *C. tropicalis* genes encoding oleic acid-inducible proteins<sup>a</sup>

GENES <sup>b</sup>	Glucose <sup>c</sup> (distance) <sup>e</sup>			Oleate <sup>d</sup> (distance) <sup>e</sup>
	A	B	C	A
<i>HDE</i>	GGGAGACATA -468 -459	TTTCTGTGAGT -428 -418	TGTGGAAAAATAATC -484 -470 TGTGGAGAAAAATGC -211 -197	CGGTTATTA -295 -287
<i>POX4</i>	GAGAGAGAGA -482 -473	ATTGTCTGAGT -330 -320	TGTGCAAAAATTCAC -252 -266 TGTGCAAAAAGTCAC -158 -172	CGGTTATTC -359 -351
<i>POX18</i>	GGGAGAGAGA -396 -387	TTTTTGTATGT -194 -184	ND	CGGTAGTTA -223 -215
<i>P450alk</i>	GGTAGACTTA -580 -571	TTCATGTGAAA -437 -427	ND	TGGTTGTTG -253 -261
<i>CATL</i>	GGGAAACAGA -422 -431	TTTGTGTGAGG -600 -590	CATGGAAAAAAAACC -188 -202 TGGGGAGAAATAAGC -134 -148	TGGTTATTA -695 -687
CONSENSUS	GGGAGA <b>G</b> A <b>T</b> C <b>G</b> A	TTTNTGTGAGA T G	TGTGGAAAAAAAAN C TTC	T <b>C</b> GGTT <b>A</b> TTA C <b>G</b> T <b>T</b> A <b>G</b>

<sup>a</sup>The sequences of the glucose-responsive (-486 to -333) and the oleic acid-responsive (-333 to -281) regions of the *HDE* gene were compared to the sequences of the upstream regions of other oleic acid-responsive genes of *C. tropicalis*.

<sup>b</sup>*HDE*, trifunctional enzyme; *POX4*, fatty acyl-CoA oxidase (Okazaki et al., 1986); *POX18*, peroxisomal 18 kiloDalton protein (Szabo et al., 1989); *P450alk*, alkane-inducible cytochrome P450 (Sanglard and Loper, 1989); *CATL*, catalase (Murray and Rachubinski, 1989).

<sup>c</sup>Sequences in the glucose-responsive region of *HDE* conserved in other oleic acid-responsive genes of *C. tropicalis*. Column A, conserved sequence A; column B, conserved sequence B similar to the *SUC2* glucose-responsive sequence of *S. cerevisiae*; column C, conserved upstream sequence found in duplicate in *CATL*, *AOX*, and *HDE*.

<sup>d</sup>Sequence in the oleic acid-responsive region of *HDE* conserved in other oleic acid-responsive genes of *C. tropicalis*.

<sup>e</sup>Distance in nt upstream of the A (+1) of the start codon, except for *HDE* where distance is given in nt upstream of the *tsp* (which lies 60nt upstream of A of start codon).

Nucleotides highlighted in bold deviate from the derived consensus sequences.



depicted above. This group discovered that when cells were grown in oleic acid, expression of *S. cerevisiae* thiolase was reduced in cells transformed with a promoter containing the conserved GG of the above consensus changed to TT, implicating this sequence's importance in oleic acid regulation.

Construct 5P consistently gave the highest levels of *HDE* expression, indicating that repressive control, by upstream sequences, mediates regulation of the gene. This type of control mediates reduced levels of expression in situations where some substrates or metabolites are more preferred when present.

Activation of construct 5P was only observed when oleic acid was added to media. Mechanisms that possibly explain the large increase in activation for construct 5P in the presence of oleic acid are two-fold: 1) that a repressive site upstream of -333 is removed in construct 5P, giving increased levels of expression in low concentrations of glucose and, 2) the additional presence of an oleic acid-responsive factor in cells cultured in oleic acid-containing media now can bind more effectively to its recognition site, possibly the " $\beta$ -oxidation-like box", because of the lack of interference from proteins binding to upstream sequences. The presence of glucose, even though at low concentrations in this study, still negatively interferes with the levels of expression.

Removal of nt downstream of -281 (5T) to give 5Q and 5V progressively reduced *HDE* expression to the limits of

detection in both glucose and glucose-oleic acid (Fig. 10), suggesting that the minimal length for a functional *HDE* promoter lies between nt -281 and -145. These deletions removed a consensus sequence located between nt -211 to -197 similar to the sequence between nt -484 to -470.

The difference between constructs 5T and 5Q is the absence of two Gs in 5Q. This small difference does not affect the expression of *HDE* for cells cultured in low concentrations of glucose but has a 3.3-fold effect for cells cultured in glucose-oleic acid. These two G residues, lacking in construct 5Q, are the last two nucleotides of a consensus sequence TCTGCCCCCCC found in *HDE*, *AOX* and *CATL*. Since the response is present only in cells cultured in media containing glucose-oleic acid, it can be speculated this site is also important in positively mediating responses to oleic acid.

Since construct 5V produced low levels of expression, it can again be argued that the regions upstream of nt -715 do not influence the expression of *HDE* in the positive fashion which perhaps could account for some of the previous results. An activating element located in this region would become more potent as its distance from the transcription start site was reduced by each deletion, and therefore increases in expression would be expected. Because expression in the construct is so low, it can be safely assumed that the elements which influenced *HDE* expression were to a large extent removed by the deletions themselves and the nucleic

acid sequences now juxtaposed to the transcription start site indeed have no, or at least very little, effect on the observed expression of the *HDE* gene.

To extend the above results obtained from cells cultured in glucose and glucose-oleic acid media, 5WT (full-length *HDE*) along with 5P and 5T constructs (deletion constructs with end-points that delineate the oleic acid-responsive region) were chosen to investigate the effects of various carbon sources. Since only 5P showed detectable levels of *HDE* expression in high concentrations of glucose (Fig. 11, panel G<sup>5</sup>), sequences upstream of -333 contain elements involved in glucose repression, in agreement with results presented above indicating the presence of a negative regulatory region between nt -466 and -333. As there is no detectable *HDE* expression seen with 5T, the positive elements responsible for the expression seen with 5P are located between nt -333 and -281 (Fig. 11, panel G<sup>5</sup>). Low concentrations of glucose increased *HDE* expression in both 5P (3.9-fold) and 5T (Fig. 11, compare G<sup>0.2</sup> to G<sup>5</sup>), indicating that derepression must be involved, and that at least some of the elements responsible for mediating this effect are located downstream of nt -281. Addition of oleic acid to low concentrations of glucose increased expression of 5WT to detectable levels, 5P 14-fold and 5T 1.3-fold, (Fig. 11, compare G<sup>0.2</sup>O<sup>0.2</sup> to G<sup>0.2</sup>), providing further evidence of an oleic acid-responsive element positioned between nt -333 and -281.

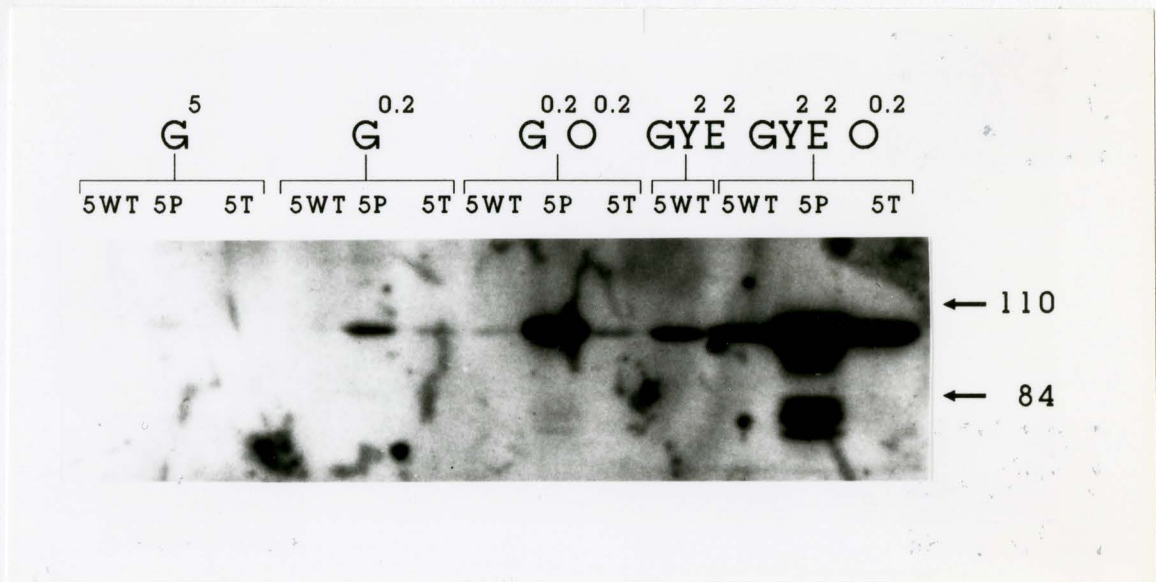


Figure 11. Effects of various carbon sources on *HDE* expression by constructs 5WT, 5P, and 5T. Transformants harbouring constructs 5WT, 5P and 5T were grown in various carbon sources. Yeast lysates were prepared and analyzed for *HDE* expression by immunoblotting. Carbon sources and methods are as in Fig. 2. Equal amounts of protein were run in each lane. The numbers at right refer to molecular weight standards and their  $M_r$  values are 110,000 and 84,000.

	CARBON SOURCE(S) PRESENT IN YNE MEDIA				
	G <sup>5</sup>	G <sup>0.2</sup>	G <sup>0.2</sup> O <sup>0.2</sup>	GY <sup>2</sup> E <sup>2</sup> O <sup>0.2</sup>	GY <sup>2</sup> E <sup>2</sup>
DELETION CONSTRUCT	RELATIVE DENSITOMETRIC SIGNAL				
5WT	-	-	1.2	8.2	11
5P	1	3.9	55	190	ND
5T	-	0.9	1.1	49	ND

Table 6 Relative densitometric signals obtained from the audioradiograph of the western blot used to produce photoreproduction depicted in Fig. 11. Each band was scanned in triplicate; the average value tabled. - = not detected; ND = not determined.

Substitution of nonfermentable glycerol-ethanol for low concentrations of glucose in media containing oleic acid resulted in increased expression for 5WT (9.3-fold), 5P (3.4-fold) and 5T (45-fold) (Fig. 11, compare  $GY^2E^{2O^{0.2}}$  and  $G^{0.2O^{0.2}}$ ), indicating that sequences responsive to glycerol-ethanol, or more generally, the lack of a fermentable carbon source, are downstream of -281 (5T). The effect of glycerol-ethanol on *HDE* expression is quite pronounced in *S. cerevisiae*; around 3 times more than that observed for *HDE* in *C. tropicalis*. From these results we can see that catabolite repression/derepression, of at least *HDE*, is a much more potent regulator of gene expression in *S. cerevisiae* than in *C. tropicalis*.

These results, and others not presented in this dissertation, certainly bore out the complexity of the regulatory mechanisms operating on expression of these genes. Yet it is no wonder that regulation of all genes, by both *cis*- and *trans*-elements, is an exceptionally complex organizational network of many factors, binding sites and interferences. The robustness of cells in their responses to environmental stimuli necessitates an extremely integrated, co-ordinated, and therefore complex system of regulatory control. In this way, the many challenging situations possible in the environmental setting, can be managed by the cell's ability to produce an internal homeostatic state. This can be achieved only by regulating and controlling the many diverse gene

products available to the cell, so that with these controls, the cell can create an internal environment able to cope with the many insults with which it is presented.

### **(3.6) Future Studies**

In order to construct a picture of the overall mechanisms of regulatory control that exist in the cell, it would be best to assay the levels of activity of the enzymes of interest, since the overall purpose of regulation is to control net enzymatic activity. Post-transcriptional/translational modifications, such as phosphorylation, modify activities of enzymes. These are not discernable when assaying for protein concentrations only. Comparisons could be made between changes in enzymatic activities and enzyme concentrations, and differences ascribed to post-translational control mechanisms such as phosphorylation, inhibition control or other stochastic effects (e.g. inaccessibilities of substrates to enzymes).

Also, a host of mutants exist with deficiencies in regulatory pathways such as catabolite repression/derepression (*SNF1*, *SSN6*; Wright and Poyton, 1990), signal transduction (*RAS2*, *SRA1-13*; Bissinger et al., 1989), and activator proteins (*ADR1*, Simon et al., 1991). If genes encoding peroxisomal proteins were expressed in such mutants, an indication of which pathways affect expression could be ascertained, and the regions responsible for mediating these

effects could be tested in both the mutants and the wild type organisms.

A powerful assay, used by many scientists, is the fusion of assayable coding regions (such as  $\beta$ -galactosidase) to the upstream elements of genes of interest. Mutagenic studies of cells carrying two similar constructs, each consisting of an identical upstream region fused to a different assayable protein, can quickly yield mutants with deficiencies in activation or repression. These mutants can then be rescued by transformation with a genomic library, and the genes mediating these regulatory anomalies isolated and characterized. Cell extracts prepared from mutant and wild type cells, cultured in various media, and in conjunction with mobility shift assays, can then reveal how and where proteins associate to the upstream region.

One approach, which would clarify the results and interpretation of these experiments, would be to grow the yeasts on media containing a detergent other than Tween 40, which does not act as an activator of the enzymes of the  $\beta$ -oxidation pathway.



## REFERENCES

- Adhya, S. and Garges, S.: Positive Control. *J. Biol. Chem.* 265 (1990) 10797-10800.
- Aitchison, J.D. and Rachubinski, R.A.: In vivo import of *Candida tropicalis* hydratase-dehydrogenase-epimerase into peroxisomes of *Candida albicans*. *Curr. Genet.* 17 (1990) 481-486.
- Aitchison, J.D., Sloots, J.A., Nuttley, W.M. and Rachubinski, R.A.: Sequence of the gene encoding *Candida tropicalis* peroxisomal trifunctional enzyme. *Gene* 105 (1991a) 135-136.
- Aitchison, J.D., Murray, W.W. and Rachubinski, R.A.: The carboxyl-terminal tripeptide Ala-Lys-Ile is essential for targeting *Candida tropicalis* trifunctional enzyme to yeast peroxisomes. *J. Biol. Chem.* 266 (1991b) 23197-23203.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.D., Smith, J.A. and Struhl, K. (Eds.), *Current Protocols in Molecular Biology*. John Wiley & Sons, New York, NY. 1987.
- Axelrod, N.J., Carmichael, G.G. and Farabough, P.J.: Enhancer and promoter elements from Simian Virus 40 and Polyomavirus can substitute for an upstream activation sequence in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 10 (1990) 947-957.
- Baker, M.E.: A common ancestor for *Candida tropicalis* and dehydrogenases that synthesize antibiotics and steroids. *FASEB J.* 4 (1990) 3028-3032.
- Binder, M., Schanz, M. and Hartig, A.: Vector-mediated overexpression of catalase A in the yeast *Saccharomyces cerevisiae* induces inclusion body formation. *Eur. J. Cell Biol.* 54 (1991) 305-312.
- Birnboim, H.C. and Doly, J.: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* 7 (1979) 1513-1523
- Bissinger, P.H., Wieser, R., Hamilton, B. and Ruis, H.: Control of *Saccharomyces cerevisiae* catalase T gene

- (CTT1) expression by nutrient supply via the RAS cyclic AMP pathway. *Mol. Cell. Biol.* 9 (1989) 1309-1315.
- Borst, P.: Peroxisome biogenesis revisited. *Biochim. Biophys. Acta* 1088 (1989) 1-13.
- Bradford, M.M.: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* 72 (1976) 248-254.
- Brandl, C.J. and Struhl, K.: Yeast GCN4 transcriptional activator protein interacts with RNA polymerase II *in vitro*. *Proc. Natl. Acad. Sci. USA* 86 (1989) 2652-2656.
- Bremer, J. and Osmundsen, H.: Fatty acid oxidation and its regulation. In: Numa, S., (Ed.), *Fatty acid metabolism and its regulation*. Elsevier, New York, NY, 1984, pp.113-154.
- Brent, R.: Repression of transcription in yeast. *Cell* 42 (1985) 2-4.
- Brindle, P.K., Holland, J.P., Willett, C.E., Innis, M.A. and Holland, M.J.: Multiple factors bind the upstream activation sites of the yeast enolase genes *ENO1* and *ENO2*: ABF1 protein, like repressor activator protein RAP1, binds *cis*-acting sequences which modulate repression or activation of transcription. *Mol. Cell. Biol.* 10 (1990) 4872-4885.
- Burnette, W.N.: Western blotting: electrophoretic transfer of proteins from SDS-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* 112 (1981) 195-203.
- Cherry, J.R., Johnson, T.R., Dollard, C., Shuster, J.R. and Denis, C.L.: Cyclic AMP-dependent protein kinase phosphorylates and inactivates the yeast transcriptional activator ADR1. *Cell* 56 (1989) 409-419.
- Darnell, Jr., J.E.: Variety in the level of gene control in eukaryotic cells. *Nature* 297 (1982) 365-371.
- Davidson, B.L., Egly, J.M., Mulvihill, E.R. and Chambon, P.: Formation of stable pre-initiation complexes between eukaryotic class B transcription factors and promoter sequences. *Nature* 301 (1983) 680-686.

- de Duve, C. and Baudhuin, P.: Peroxisomes (microbodies and related particles). *Physiol. Rev.* 46 (1966) 323-357.
- del Rio, L.A., Sandalio, L.M. and Palma, J.M.: A new cellular function for peroxisomes related to oxygen free radicals? *Experientia* 46 (1990) 989-992.
- Dommes, P., Dommes, V. and Kunau, W.-H.:  $\beta$ -oxidation in *Candida tropicalis*. *J. Biol. Chem.* 258 (1983) 10846-10852.
- Douma, A.C., Veenhuis, M., Waterham, H.R. and Harder, W.: Immunocytochemical demonstration of the peroxisomal ATPase of yeast. *Yeast* 6 (1990) 45-51.
- Dugacieck, A., Boyer, H.W. and Goodman, H.M.: Ligation of *EcoRI* endonuclease-generated DNA fragments into linear and circular structures. *J. Mol. Biol.* 96 (1975) 171-184.
- Dynan, W.S.: Modularity in promoters and enhancers. *Cell* 58 (1989) 1-4.
- Einerhand, A.W.C., Voorn-Brouwer, T.M., Erdmann, R., Kunau, W.-H. and Tabak, H.F.: Regulation of transcription of the gene coding for peroxisomal 3-oxoacyl-CoA thiolase of *Saccharomyces cerevisiae*. *Eur. J. Biochem.* 200 (1991) 113-122
- Entian, K.-D.: Genetic and biochemical evidence for hexokinase PII as a key enzyme involved in carbon catabolite repression in yeast. *Mol. Gen. Genet.* 178 (1980) 633-637.
- Erdmann, R., Veenhuis, M., Mertens, D. and Kunau, W.-H.: Isolation of peroxisome-deficient mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 86 (1989) 5419-5423.
- Fahimi, H.D. and Sies, H., (Eds.) In: Peroxisomes in biology and medicine. Springer-Verlag, Berlin, 1987.
- Fedor, M.J. and Kornberg, R.D.: Upstream activation sequence dependent alteration of chromatin structure and transcription activation on the yeast *GAL1-GAL10* genes. *Mol. Cell. Biol.* 9 (1989) 1721-1732.
- Finley, Jr., R.L. and West, Jr., R.W.: Differential repression of GAL4 and adjacent transcription activators by operators in the yeast *GAL* upstream activation sequence. *Mol. Cell. Biol.* 9 (1989) 4282-4290.

- Frankel, A.D. and Kim, P.S.: Modular structure of transcription factors: Implications for gene regulation. *Cell* 65 (1991) 717-719.
- Fujiki, Y., Rachubinski, R.A. and Lazarow, P.B.: Synthesis of a major integral membrane polypeptide of rat liver peroxisomes on free polysomes. *Proc. Natl. Acad. Sci. USA* 81 (1984) 7127-7131.
- Giniger, E. and Ptashne, M.: Transcription in yeast activated by a putative amphipathic  $\alpha$ -helix linked to a DNA binding unit. *Nature* 330 (1987) 670-672.
- Giniger, E., Varnum, S.M. and Ptashne, M.: Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* 40 (1985) 767-774.
- Gould, S.J., Keller, G.-A., Scheider, M., Distel, B., Tabak, H. and Subramani, S.: Peroxisomal protein import is conserved between yeast, plants, insects and mammals. *EMBO J.* 9 (1990) 85-90.
- Gould, S.J. and Subramani, S.: Firefly luciferase as a tool in molecular and cell biology. *Anal. Biochem.* 175 (1988) 5-13.
- Grunstein, M. and Hogness, D.S.: Colony hybridization: a method for the isolation of cloned DNA's that contain a specific gene. *Proc. Natl. Acad. Sci. USA* 72 (1975) 3961-3965.
- Guarente, L., Lalonde, B., Gifford, P. and Alani, E.: Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. *Cell* 36 (1984) 503-511.
- Guarente, L.: Regulatory proteins in yeast. *Ann. Rev. Genet.* 21 (1987) 425-452.
- Hanish, J. and McClelland, M.: Activity of DNA modification and restriction enzymes in KGB, a potassium glutamate buffer. *Gene Analyt. Techn.* 5 (1988) 105-107.
- Hashimoto, T.: Individual peroxisomal  $\beta$ -oxidation enzymes. In: Kindl, H. and Lazarow, P.B. (Eds.), *Peroxisomes and glyoxysomes*. *Ann. N.Y. Acad. Sci.* 386 (1982) 5-12.
- Heard, J.M., Herbomel, P., Ott, M.O., Mottura-Rollier, A., Weiss, M. and Yaniv, M.: Determinants of rat albumin promoter tissue specificity analyzed by an improved transient expression system. *Mol. Cell. Biol.* 7 (1987)

2425-2434.

- Helmann, J.D. and Chamberlin, M.J.: Structure and function of bacterial sigma factors. *Ann. Rev. Biochem.* 57 (1988) 839-872.
- Henikoff, S.: Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28 (1984) 351-359.
- Herrero, P., Fernández, R. and Moreno, F.: The hexokinase isoenzyme PII of *Saccharomyces cerevisiae* is a protein kinase. *J. Gen. Microbiol.* 135 (1989) 1209-1216.
- Hodge, M.R., Singh, K. and Cumsy, M.G.: Upstream activation and repression elements control transcription of the yeast *COX5b* gene. *Mol. Cell. Biol.* 10 (1990) 5510-5520.
- Horikoshi, M., Wang, C.K., Fugii, H., Cromlish, J.A., Weil, P.A. and Roeder, R.G.: Purification of a yeast TATA box-binding protein that exhibits human transcription factor IID activity. *Proc. Natl. Acad. Sci. USA* 86 (1989) 4843-4847.
- Ito, H., Fukuda, Y., Murata, K. and Kimura, A.: Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* 153 (1983) 163-168.
- Kakidani, H. and Ptashne, M.: GAL4 activates gene expression in mammalian cells. *Cell* 52 (1988) 161-167.
- Kawamoto, S., Nozaki, C., Tanaka, A. and Fukui, S.: Fatty acid  $\beta$ -oxidation in microbodies of *n*-alkane-grown *Candida tropicalis*. *Eur. J. Biochem.* 83 (1978) 609-613.
- Kindl, H.: Lipid degradation in higher plants. In: Numa, S., (Ed.) *Fatty acid metabolism and its regulation*. Elsevier, New York, NY 1984, pp.181-204.
- Kispal, G. and Srere, P.A.: Studies on yeast peroxisomal citrate synthase. *Arch. Biochem. Biophys.* 286 (1991) 132-137.
- Kronstad, J.W., Holly, J.A. and MacKay, V.L.: A yeast operator overlaps an upstream activation site. *Cell* 50 (1987) 369-377.
- Laskey, R.A. and Mills, A.D.: Enhanced autoradiographic detection of  $^{32}\text{P}$  and  $^{125}\text{I}$  using intensifying screens

- and hypersensitized film. FEBS Lett. 82 (1977) 314-316.
- Lazarow, P.B. and Fujiki, Y.: Biogenesis of peroxisomes. Ann. Rev. Cell Biol. 1 (1985) 489-530.
- Lenardo, M.J. and Baltimore, D.: NF $\kappa$ -B: A pleiotropic mediator of inducible and tissue-specific gene control. Cell 58 (1989) 227-229.
- Lewin, B.: Commitment and activation at Pol II promoters: A tail of protein-protein interactions. Cell 61 (1990) 1161-1164.
- Lillie, J.W. and Green, M.R.: Activator's target in sight. Nature 341 (1989) 279-280.
- Luche, R.M., Sumrada, R. and Cooper, T.G.: A *cis*-acting element present in multiple genes serves as a repressor protein binding site for the yeast *CAR1* gene. Mol. Cell. Biol. 10 (1990) 3884-3895
- Lue, N.F., Buchman, A.R. and Kornberg, R.D.: Activation of yeast RNA polymerase II transcription by a thymidine-rich upstream element *in vitro*. Proc. Natl. Acad. Sci. USA 86 (1989) 486-490.
- Ma, H., Bloom, L.M., Walsh, C.T. and Botstein, D.: The residual enzymatic phosphorylation activity of hexokinase II mutants is correlated with glucose repression in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 9 (1989) 5643-5649.
- Maniatis, T., Fritsch, E.F. and Sambrook, J.: Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, (1982).
- Manley, J.M.: Analysis of the expression of genes encoding animal mRNA by *in vitro* techniques. Prog. Nucl. Acids Res. Mol. Biol. 30 (1983) 195-244.
- Mannhaupt, G., Pilz, U. and Feldmann, H.: A series of shuttle vectors using chloramphenicol acetyltransferase as a reporter enzyme in yeast. Gene 67 (1988) 287-294.
- Moreno de la Garza, M., Schultz-Borchard, U., Crabb, J.W. and Kunau, W.-H.: Peroxisomal  $\beta$ -oxidation system of *Candida tropicalis*: Purification of a multifunctional protein possessing enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase and 3-hydroxyacyl-CoA epimerase activities. Eur. J. Biochem. 148 (1985) 285-291.

- Morgan, A.R., Lee, J.S., Pulleyblank, D.E., Murray, N.L. and Evans, D.H.: Ethidium fluorescence assays. Part I. Physicochemical studies. Nucl. Acids Res. 7 (1979) 547-569
- Muller, M.M., Gerster, T. and Schaffner, W.: Enhancer sequences and the regulation of gene transcription. Eur. J. Biochem. 176 (1988) 485-495.
- Murray, W.W. and Rachubinski, R.A.: Nucleotide sequence of peroxisomal catalase from the yeast *Candida tropicalis* pK233: identification of an upstream *Bam*HI site polymorphism. Nucl. Acids Res. 17 (1989) 3600.
- Myboni, K., Van Aelst, L., Argüelles, J.C., Jans, A.W.H. and Thevelein, J.M.: Glucose-induced hyperaccumulation of cyclic AMP and defective glucose repression in yeast strains with reduced activity of cyclic AMP-dependent protein kinase. Mol. Cell. Biol. 10 (1990) 4518-4523.
- Nehlin, J.O. and Ronne, H.: Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumour finger proteins. EMBO J. 9 (1990) 2891-2891.
- Newport, J.W. and Forbes, D.J.: The nucleus: Structure, function and dynamics. Ann. Rev. Biochem. 56 (1987) 535-565.
- Nishizawa, M., Araki, R. and Teranishi, Y.: Identification of an upstream activating sequence and a upstream repressible sequence of the pyruvate kinase gene of the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. 9 (1989) 442-451.
- Nishizawa, M., Suzuki, Y., Nogi, Y., Matsumoto, K. and Fukasawa, T.: Yeast GAL11 protein mediates the transcriptional activation signal of two different transacting factors, GAL4 and general regulatory factor I/repressor/activator site binding protein 1/translation upstream factor. Proc. Natl. Acad. Sci. USA 87 (1990) 5373-5377.
- Nuttley, W.M., Aitchison, J.D. and Rachubinski, R.A.: cDNA cloning and primary structure determination of the peroxisomal trifunctional enzyme hydratase-dehydrogenase-epimerase from the yeast *Candida tropicalis* pK233. Gene 69 (1988) 171-180.
- Okazaki, K., Takechi, T., Kambara, N., Fukui, S., Kubota, I. and Kamiryo, T.: Two acyl-coenzyme A oxidases in peroxisomes of the yeast *Candida tropicalis*: primary

- structures deduced from genomic DNA sequence. Proc. Natl. Acad. Sci. USA 83 (1986) 1232-1236.
- Ow, D.W., Jacobs, J.D. and Howell, S.H.: Functional regions of the cauliflower mosaic virus 35 S RNA promoter determined by use of the firefly luciferase gene as a reporter of promoter activity. Proc. Natl. Acad. Sci. USA 84 (1987) 4870-4874.
- Parker, C.S. and Topol, J.: A Drosophila RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. Cell 36 (1984) 357-369.
- Pfeiffer, B.H. and Zimmerman, S.B.: Polymer-stimulated ligation: enhanced blunt or cohesive-end ligation of DNA or deoxyribonucleotides by T4 DNA ligase in polymer solutions. Nucl. Acids Res. 11 (1983) 7853-7871.
- Rachubinski, R.A.: Genetic methods for and gene structure in other *Candida* species. In: Kirsch, D.R., Kelly, R. and Kurtz, M.B. (Eds.), The Genetics of *Candida*. CRC Press, Boca Raton, FL, 1990, pp. 177-186.
- Rose, M.D., Novick, P., Thomas, J.H., Botstein, D. and Fink, G.R.: A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. Gene 60 (1987) 237-243.
- Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463-5467.
- Sanglard, D. and Loper, J.C.: Characterization of the alkane-inducible cytochrome P450 (P450alk) gene from the yeast *Candida tropicalis*: identification of a new P450 family. Gene 76 (1989) 121-136.
- Sarokin, L. and Carlson, M.: Upstream region required for regulated expression of the glucose-repressible *SUC2* gene of *Saccharomyces cerevisiae*. Mol. Cell. Biol. 4 (1984) 2750-2757.
- Sarokin, L. and Carlson, M.: Short repeated elements in the upstream regulatory region of the *SUC2* gene of *Saccharomyces cerevisiae*. Mol. Cell. Biol. 6 (1986) 2324-2333.
- Sawadogo, M. and Roeder, R.G.: Factors involved in specific transcription by human RNA polymerase II: Analysis by a rapid and quantitative *in vitro* assay. Proc. Natl.



- Acad. Sci. USA 82 (1985) 4394-4398.
- Schibler, U. and Sierra, F.: Alternative promoters in developmental gene expression. *Ann. Rev. Genet.* 21 (1987) 237-257.
- Schmitt, M.E., Brown, T.A. and Trumppower, B.L.: A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucl. Acids Res.* 18 (1990) 3091.
- Schreck, R., Carey, M.F. and Grummt, I.: Transcriptional enhancement is brought about by different molecular mechanisms for class I and II RNA polymerase genes. *EMBO J.* 8 (1989) 3011-3017.
- Sigler, P.B.: Acid blobs and negative noodles. *Nature* 333 (1988) 210-212.
- Simon, M., Gerhard, A., Rapatz, W., Spevak, W. and Ruis, H.: The *Saccharomyces cerevisiae* *ADR1* gene is a positive regulator of transcription of genes encoding peroxisomal proteins. *Mol. Cell. Biol.* 11 (1991) 699-704.
- Skoneczny, M., Chelstowska, A. and Rytka, J.: Study of the coinduction by fatty acids of catalase A and acyl-CoA oxidase in standard and mutant *Saccharomyces cerevisiae* strains. *Eur. J. Biochem.* 174 (1988) 297-302.
- Sloots, J.A., Aitchison, J.D. and Rachubinski, R.A.: Glucose-responsive and oleic acid-responsive elements in the gene encoding the peroxisomal trifunctional enzyme of *Candida tropicalis*. *Gene* 105 (1991) 129-134.
- Struhl, K.: The new yeast genetics. *Nature* 303 (1983) 391-396.
- Struhl, K.: Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. USA* 82 (1985) 8419-8423.
- Struhl, K.: Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* 49 (1987) 295-297.
- Stucka, R. and Feldmann, H.: An element of symmetry in yeast TATA-box binding protein transcription factor IID. *FEBS Lett.* 261 (1990) 223-225

- Szabo, L.J., Small, G.M. and Lazarow, P.B.: The nucleotide sequence of *POX18*, a gene encoding a small oleate-inducible peroxisomal protein from *Candida tropicalis*. *Gene* 75 (1989) 119-126.
- Tanabe, I., Okada, J. and Ono, H.: Isolation and determination of yeasts utilizing kerosene as a sole source of carbon. *Agr. Biol. Chem.* 30 (1966) 1175-1182.
- Tanaka, A., Osumi, M. and Fukui, S.: Peroxisomes of alkane-grown yeast: fundamental and practical aspects. *Ann. N.Y. Acad. Sci.* 386 (1982) 183-199.
- Thompson, J., Solomon, R., Pellegrino, M., Sakai, K., Lewin, M., Feild, M., Castrovinci, M., Sacramore, L. and Gillespie, D.: A noise free molecular hybridization procedure for measuring RNA in cell lysates. *Anal. Biochem.* 181 (1989) 371-378.
- Tolbert, N.E.: Metabolic pathways in peroxisomes and glyoxysomes. *Ann. Rev. Biochem.* 50 (1981) 133-57.
- Travers, A.A.: Why bend DNA? *Cell* 60 (1990) 177-80.
- van der Klei, I.J., Harder, W. and Veenhuis, M.: Biosynthesis and assembly of alcohol oxidase, a peroxisomal matrix protein in methylotrophic yeasts: a review. *Yeast* 7 (1991) 195-209.
- Van Loon, A.P., Van Eijk, E. and Grivell, L.A.: Biosynthesis of the ubiquinol-cytochrome *c* reductase complex in yeast. Discoordinate synthesis of the 11-kd subunit in response to increased gene copy number. *EMBO J.* 2 (1983) 1765-1770.
- Veenhuis, M., Douma, A.C., Harder, W. and Osumi, M.: Degradation and turnover of peroxisomes in the yeast *Hansenula polymorpha* induced by selective inactivation of peroxisomal enzymes. *Arch. Microbiol.* 134 (1983) 193-203.
- Veenhuis, M. and Goodman, J.M.: Peroxisomal assembly: Membrane proliferation precedes the induction of the abundant matrix proteins in the methylotrophic yeast *Candida boidinii*. *J. Cell Sci.* 96 (1990) 583-590.
- West, Jr., R.W., Chen, S., Putz, H., Butler, G. and Banerjee, M.: *GAL1-GAL10* divergent promoter region of *Saccharomyces cerevisiae* contains negative control elements in addition to functionally separate and possibly overlapping upstream activation sequences.

Genes Dev. 1 (1987) 1118-1131.

Wright, R.M. and Poyton, R.O.: Release of two *Saccharomyces cerevisiae* cytochrome genes, COX6 and CYC1, from glucose repression requires the SNF1 and SSN6 gene products. Mol. Cell. Biol. 10 (1990) 1297-1300.

Zhang, H., Scholl, R., Browse, J. and Somerville, C.: Double stranded sequencing as a choice for DNA sequencing. Nucl. Acids Res. 16 (1988) 1220.