TARGETING OF *CANDIDA TROPICALIS* TRIFUNCTIONAL ENZYME TO PEROXISOMES IN YEAST: IDENTIFICATION OF A CARBOXY-TERMINAL TRIPEPTIDE PEROXISOMAL TARGETING SIGNAL

By

JOHN DAVID AITCHISON, B.Sc.

A Thesis Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

PEROXISOMAL TARGETING SIGNALS IN YEAST

DOCTOR OF PHILOSOPHY (1992)
(Biochemistry)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Targeting of *Candida tropicalis* trifunctional enzyme to peroxisomes in yeast: Identification of a carboxy-terminal tripeptide peroxisomal targeting signal

AUTHOR: John David Aitchison, B.Sc. (McMaster University)

SUPERVISOR: Dr. Richard A. Rachubinski, Associate Professor

NUMBER OF PAGES: xxvi, 283

ABSTRACT

The nature of the peroxisomal targeting signal in *Candida tropicalis* trifunctional enzyme, hydratase-dehydrogenase-epimerase (HDE) was investigated. The first part of this thesis describes the cloning and sequencing of the cDNA and gene encoding HDE. The second and third parts of the thesis describe the amino acid sequence (and general nature) of the signal that directs HDE to peroxisomes in yeast.

The cDNA and gene encoding *C. tropicalis* HDE was isolated from a λgt11 cDNA expression library and from a λEMBL3 genomic library, respectively. Primary sequence analysis of both the cDNA and the gene revealed a single open reading frame of 2718 nucleotides encoding a protein of 906 amino acids (calculated molecular mass = 99,481 Da).

The gene encoding HDE was expressed in the yeasts *Candida albicans* and *Saccharomyces cerevisiae*. The cellular location of HDE was determined by subcellular fractionation followed by western blot analysis of peroxisomal and cytosolic fractions using antiserum specific for HDE. HDE was found to be exclusively targeted to and imported into peroxisomes in both heterologous expression systems. Deletion and mutational analyses were used to determine the regions within HDE which are essential for its targeting to peroxisomes. Deletion of a carboxy-terminal tripeptide Ala-Lys-Ile completely abolished targeting of HDE to peroxisomes, whereas large internal deletions of HDE (amino acids 38-353 or 395-731) had no effect on HDE targeting to peroxisomes

in either yeast. Substitutions within the carboxy-terminal tripeptide (Ala→Gly and Lys→Gln) supported targeting of HDE to peroxisomes of *C. albicans* but not of *S. cerevisiae*.

Antiserum directed against the carboxy-terminal 11 amino acids of HDE, containing the Ala-Lys-Ile tripeptide (anti-AKI) was used to probe subcellular fractions from several yeasts and rat liver. The anti-AKI serum reacted exclusively with multiple peroxisomal proteins from the yeasts *C. tropicalis*, *C. albicans* and *Yarrowia lipolytica* suggesting a common motif for the targeting of some proteins to yeast peroxisomes, the prototype of which is the HDE carboxy-terminal tripeptide Ala-Lys-Ile.

The results presented herein have appeared, in part, in the following publications:

Publications

1. Nuttley, W.M., Aitchison, J.D. and Rachubinski, R.A. cDNA cloning and primary structure determination of peroxisomal hydratase-dehydrogenase-epimerase from the yeast *Candida tropicalis* pK233. Gene 69: 171-180 (1988).

2. Aitchison, J.D. and Rachubinski, R.A. *In vivo* import of *Candida tropicalis* hydratase-dehydrogenase-epimerase into peroxisomes of *Candida albicans*. Curr. Genet. 17: 481-486 (1990).

3. Aitchison, J.D., Sloots, J.A., Nuttley, W.M. and Rachubinski, R.A. Nucleotide sequence of the gene encoding peroxisomal trifunctional of *Candida tropicalis*. Gene 105: 135-136 (1991).

4. Sloots, J.A., Aitchison, J.D. and Rachubinski, R.A. Regulation of the gene encoding peroxisomal trifunctional enzyme of *Candida tropicalis*: identification of upstream responsive elements. Gene 105: 129-134 (1991).

5. Aitchison, J.D., Murray, W.W. and Rachubinski, R.A. The carboxy-terminal tripeptide Ala-Lys-Ile is essential for targeting of *Candida tropicalis* trifunctional enzyme to yeast peroxisomes. J. Biol. Chem. 266: 23197-23203 (1991).

6. Aitchison, J.D., Szilard, R.K., Nuttley, W.M. and Rachubinski, R.A. Antibodies directed against a yeast peroxisomal targeting signal specifically recognize peroxisomal proteins from various yeasts. Yeast (*in press*).

Abstracts and Presentations at Scientific Meetings:

1. Nuttley, W.M., Aitchison, J.D. and Rachubinski, R.A. cDNA cloning and primary-structure determination of peroxisomal hydratase-dehydrogenase-epimerase from the yeast *Candida tropicalis* pK233. Fourth Intl. Congr. Cell Biol. p. 341 (1988).

2. Aitchison, J.D. and Rachubinski, R.A. Expression of the gene encoding *Candida tropicalis* hydratase-dehydrogenase-epimerase in *Candida albicans*: Import into peroxisomes *in vivo* and induction by growth on oleic acid. Cold Spring Harbor Laboratory Meeting on Yeast Cell Biology, p. 164 (1989).

3. Aitchison, J.D. and Rachubinski, R.A. Import and assembly of *Candida tropicalis* hydratase-dehydrogenase-epimerase into peroxisomes of *Candida albicans*. J. Cell Biol. 109, Part 2, 56a (1989).

4. Rachubinski, R.A. and Aitchison, J.D. Import and assembly of *Candida tropicalis* hydratase-dehydrogenase-epimerase into peroxisomes of *Candida albicans*. J. Cell. Biochem. Supplement 14C, p. 24 (1990).

5. Aitchison, J.D. and Rachubinski, R.A. *Candida tropicalis* trifunctional enzyme is targeted to peroxisomes by its carboxy-terminal tripeptide. J. Cell Biol. 111: 195a (1990).

6. Aitchison, J.D., Murray. W.W. and Rachubinski, R.A. The carboxy-terminal tripeptide Ala-Lys-Ile is essential for targeting of *Candida tropicalis* trifunctional enzyme to yeast peroxisomes. Cold Spring Harbor Meeting on Yeast Cell Biology, p. 17 (1991).

7. Sloots, J.A., Aitchison, J.D. and Rachubinski, R.A. Glucose- and oleic acid-responsive elements in the gene encoding peroxisomal trifunctional enzyme of *Candida tropicalis*. Cold Spring Harbor Meeting on Yeast Cell Biology, p. 45 (1991).

8. Aitchison, J.D., Szilard, R.K. and Rachubinski, R.A. Antiserum to the carboxy-terminal peroxisomal targeting signal of *C. tropicalis* trifunctional enzyme recognizes multiple yeast peroxisomal proteins. J. Cell Biol. 115: p233a (1991).

9. Sloots, J.A., Aitchison, J.D. and Rachubinski, R.A. Glucose- and oleic acid-responsive elements in the gene encoding peroxisomal trifunctional enzyme of *Candida tropicalis*. J. Cell Biol. 115: p234a (1991).

10. Brade, A., Nuttley, W., Aitchison, J., Gaillardin, C. and Rachubinski, R.A. Peroxisomal assembly mutants in the yeast *Yarrowia lipolytica*. A.S.C.B. Molecular Mechanisms of Membrane Traffic. (In press).

ACKNOWLEDGEMENTS

I am grateful to my supervisor and friend Dr. Richard Rachubinski who has allowed me to develop my own ideas and led me in the right direction when I went astray. Rick has been a great source of inspiration throughout the course of this work. His endless enthusiasm has helped me to trudge forward through many frustrating days and weeks (months). I am indebted to everyone in the "Rach lab" past and present, Wayne Murray, Andrea Bodnar, Bill Nuttley, John Glover, Jim Sloots, Baowei Zhang, Tony Brade, Gary Eitzen, Rachel Szilard, Sandra Marcus, Jill Regeoczi and Shirley Jones for providing a truly unique and wonderful place to work. I can't imagine a more diverse group of people. Only science could have brought you all together. I am especially grateful to those of you who made it easy to forget about this place through sports, parties, and even daily conversations (and arguments) about matters having very little to do with science. You are what made this a terrific place to work. I will miss you all.

I would like to extend my appreciation to my supervisory committee members, Dr. Hara Ghosh, Dr. John Capone and Dr. Raymond Lee for their interest in and constructive criticism of this work.

I thank Bill Nuttley for a critical reading of this manuscript and Barbara Sweet for help with typing this thesis. I also thank Jim Fenner for lending me a laptop so that I never had to be too far away from the keyboard. -Thanks a lot-

I am grateful to the Natural Sciences and Engineering Research Council and The Ontario Graduate Studentship Program, for financial support during my graduate studies. I would also like to thank McMaster University and the Department of Biochemistry for financial support through scholarships and teaching assignments.

I would like to thank my family, especially my parents, Don and Sharron who have always been there when I needed them. Finally, I am truly indebted to my wife Leslie, whose love and encouragement have inspired me throughout the years. Your contributions were critical to this accomplishment.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A	Ampere
ADP	adenosine diphosphate
AKX	mutant form of HDE
AOX	fatty acyl-CoA oxidase
AQI	mutant form of HDE
ARS	autonomously replicating sequence
ATP	adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bp	base pair
BSA	bovine serum albumin
C.	*Candida*
CAT	catalase
CAT	gene encoding CAT
CATAKI	mutant form of *E. coli* CAT
CATAKI	gene encoding CATAKI
cDNA	DNA complementary to RNA
Ci	Curie(s)
cm	centimeter(s)

CoA	coenzyme A
cpm	counts per minute
D.	*Drosophila*
Da	dalton
ddNTP	2',3'-dideoxynucleoside 5'-triphosphate
DHFR	dihydrofolate reductase
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside 5'-triphosphate
ds	double stranded
DTT	dithiothreitol
E.	*Escherichia*
E. coli CAT	chloramphenicol acetyltransferase
E. coli CAT	gene encoding CAT
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
FAD	oxidized form of flavin adenine nucleotide
$FADH_2$	reduced form of flavin adenine nucleotide
Fig.	figure(s)
g	gram(s)
g_{av}	acceleration of gravity (average)

GKI	mutant form of HDE
g_{max}	acceleration of gravity (maximum)
H.	*Hansenula*
h	hour(s)
HD	gene encoding HD
HD	hydratase-dehydrogenase (bifunctional enzyme)
HDE	hydratase-dehydrogenase-epimerase (trifunctional enzyme)
HDE	gene encoding HDE
HEPES	n-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
hsp	heat shock protein
ICL	isocitrate lyase
IgG	immunoglobulin G
IgG(Fc)	constant region of immunoglobulin G
IPTG	isopropyl ß-D-thiogalactoside
k	kilo
KLH	keyhole limpet hemocyanin
l	litre
LB	Luria broth
μ	micro
m	milli
mm	millimeter(s)

M	molar
min	minute(s)
mol	mole(s)
M_r	relative molecular mass
mRNA	messenger ribonucleic acid
MS	malate synthase
MW	molecular weight
n	nano
NAD^+	nicotinamide adenine dinucleotide (oxidized form)
NADH	nicotinamide adenine dinucleotide (reduced form)
NBT	nitroblue tetrazolium
NP-40	Nonidet P-40
nt	nucleotide(s)
NTP	nucleoside 5'-triphosphate
OD_x	optical density measured at the wavelength of x nanometers
ORF	open reading frame
p	pico
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PEG	polyethylene glycol
pfu	plaque forming unit(s)

PMSF	phenylmethylsulfonyl fluoride
PNS	post-nuclear supernatant
POX-x	gene encoding PXP number x
PPO	2,5-diphenyloxazole
PTS	peroxisomal targeting signal
PXP-x	*C. tropicalis* peroxisomal protein number x
RNA	ribonucleic acid
rpm	revolutions per minute
S.	*Saccharomyces*
SDS	sodium dodecyl sulfate
SOD	superoxide dismutase
SRP	signal recognition particle
SSC	sodium chloride-sodium citrate buffer
TBE	Tris-borate-EDTA buffer
TBST	Tris-buffered saline plus Tween 20
TCA	trichloroacetic acid
TE	10 mM Tris-HCl (pH 7,5), 1 mM EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris(hydroxymethyl)aminoethane
tRNA	transfer ribonucleic acid
Tween 20	polyoxyethylenesorbitan monolaurate

V	volt(s)
X-gal	5-bromo-4-chloro-3-indolyl-ß-D-galactoside
XKI	mutant form of HDE
Y.	*Yarrowia*

1.0 INTRODUCTION

1.1 Overview

Eukaryotic cells have evolved a complex set of intracellular organelles, each of which is endowed with a specific complement of enzymes and performs unique metabolic functions. This compartmentalization of cellular functions provides a level of metabolic control not available to prokaryotic cells, but it also presents the cell with the problem of targeting proteins to their specific intracellular location(s). Proteins must be efficiently transported from their site of synthesis in the cytosol into their specific organelle(s). Such a process may require translocation across one or more hydrophobic membrane barriers or assymetric integration into specific membranes within the cell. The study of how this is achieved is currently under intense investigation by a number of different laboratories and has revealed a complex set of signals and cellular components that mediate these targeting and translocation events.

Proteins carry *cis*-acting amino acid sequences which serve to act as recognition sequences for protein sorting and the translocation machinery within the cell. Sequences that target proteins to the endoplasmic reticulum/secretory pathway, mitochondria or chloroplasts are often present as cleavable amino-terminal extensions. In contrast, most peroxisomal proteins are synthesized at their mature size and are

transported into the organelle without any post-translational modifications. Therefore, the information specific for targeting proteins to peroxisomes must lie within the mature molecule. This thesis is an analysis of the peroxisomal targeting signal (PTS) of a *Candida tropicalis* peroxisomal protein, hydratase-dehydrogenase-epimerase (HDE).

1.2 Background: What is a peroxisome?

The peroxisome was first discovered in thin sections of mouse kidney cells by Johannes A. Rhodin in 1954. The peroxisome appeared as an intracellular organelle, approximately 0.5 μm in diameter, bound by a single unit membrane and containing a granular matrix (cited in de Duve and Baudhuin, 1966). The name 'peroxisome' was given to the organelle by de Duve (1965) based on its characteristic production of H_2O_2 (by several different oxidases) and the presence of catalase to metabolize the peroxide (de Duve and Beaudin, 1966). Peroxisomes have since been shown to harbour a different complement of enzymes and perform a variety of essential metabolic tasks depending on cell type and environment, including the oxidation of glycolate, amino acids, alcohols, lactate, polyamines, urate and fatty acids to generate H_2O_2. In mammalian cells, peroxisomes play a role in bile acid, ether-linked glycerolipid and cholesterol biosyntheses, purine and amino acid metabolism and the pentose phosphate pathway (for review see Tolbert, 1981.)

Peroxisomes have been placed in a larger family of morphologically similar organelles called microbodies, which are classified based on their specific metabolic

roles in different organisms. Microbodies, which include peroxisomes, glyoxysomes and glycosomes are present in almost all eukaryotic cells (Lazarow and Fujiki, 1985; Hruban *et al.*, 1972). The different types of microbodies are classified based upon their specific metabolic functions. However, as more is learned about the biochemistry of each type of microbody, the distinctions between these organelles have become less clear.

1.2.1 Glyoxysomes

Glyoxysomes are found in plant cells and contain the enzymes of the glyoxylate cycle, which allows for the synthesis of carbohydrate from lipid-derived acetyl-CoA (Kornberg and Beevers, 1957; Cooper and Beevers, 1969a). Similar to peroxisomes, glyoxysomes also contain an H_2O_2-producing β-oxidation system, urate oxidase (the original oxidase found in rat liver peroxisomes; de Duve *et al.*, 1955) and catalase (Cooper and Beevers, 1969b). In addition, the glyoxylate cycle enzymes isocitrate lyase (ICL) and malate synthase (MS) are present in yeast peroxisomes. Although, in general, animal cells do not have the capability to synthesize carbohydrate from fatty acid, ICL and MS have also been found in peroxisomes of toad bladder (Goodman *et al.*, 1980).

1.2.2 Glycosomes

Glycosomes are found in trypanosomes and harbour glycolytic enzymes. Glycosomes have a diameter of 0.2-0.3 μm, are bound by a single unit membrane and

sometimes contain a crystalline core (Opperdoes *et al.*, 1977). In spite of these morphological similarities, until recently, glycosomes were thought to bear little biochemical relation to peroxisomes. Although an H_2O_2-producing oxidase has not been found in glycosomes, catalase has been localized to the microbodies of *Crithidia sp.* (Opperdoes *et al.*, 1977). In addition, like glyoxysomes and peroxisomes, glycosomes contain enzymes for the β-oxidation of fatty acids (Opperdoes, 1988).

Therefore, although not all microbodies conform to the strict definition of a peroxisome provided by de Duve (de Duve and Baudhuin, 1966), the morphologic similarities and the overlapping enzymatic complement of peroxisomes, glyoxysomes and glycosomes argue for a common evolutionary origin of microbodies and the subsequent specialization of the organelle in different organisms (see section 1.5). Thus, microbodies were all classified as peroxisomes by Osumi and Fujiki (1990).

The discussion that follows will focus on peroxisomes of mammalian cells and yeast; however, many aspects of peroxisome biogenesis and protein targeting may well apply to microbodies of other organisms (glyoxysomes and glycosomes).

1.3 Peroxisomal fatty acid β-oxidation

As mentioned above, all microbodies play a role in lipid metabolism by way of a fatty-acid β-oxidation system (Opperdoes, 1988; Kunau *et al.*, 1988), the properties and enzymes of which differ from their mitochondrial counterparts and even between organisms. The following discussion will be limited to the β-oxidation system and

enzymes from rat liver and the yeast *C. tropicalis*, because they are the best characterized.

1.3.1 Rat liver peroxisomal β-oxidation

As early as 1965, it was noted that clofibrate, a hypolipidemic drug (Oliver, 1963), caused the proliferation of rat hepatic peroxisomes (Hess *et. al.*, 1965). This fact, as well as the identification of a β-oxidation system in glyoxysomes (Cooper and Beevers, 1969b), led Lazarow and de Duve (1976) to investigate the role of a peroxisomal β-oxidation system in rat liver. It was originally believed that the oxidation of fatty acids in mammalian cells was carried out exclusively in mitochondria. However, the discovery of β-oxidation in peroxisomes of rat liver (Lazarow and de Duve, 1976; Lazarow, 1978) revealed an important contribution to long chain fatty acid metabolism by mammalian peroxisomes, and suggested that peroxisomes play a role in reducing serum lipid levels by induction of the peroxisomal β-oxidation system (Lazarow, 1978).

In mammalian cells, peroxisomes are used to break down long-chain fatty acids (C_{10}-C_{30}), whereas mitochondrial β-oxidation has a greater specificity for shorter chain fatty acids (up to C_{18}) (Lazarow, 1978). The shorter chain products ($<C_8$) of peroxisomal long chain fatty acid β-oxidation (and acetyl-CoA) are transported to mitochondria via carnitine esters for further oxidation (Lazarow, 1978; Hashimoto, 1982; Tolbert, 1981; Mannaerts and Debeer, 1982).

1.3.1.1 Fatty acyl-CoA oxidase

The peroxisomal β-oxidation pathway of rat liver peroxisomes is shown in Fig. 1.3.1.1 and was reviewed by Tolbert (1981). The fatty acid is transported into peroxisomes in a carnitine-independent manner, where peroxisomal acyl-CoA synthetase activates the fatty acid by the ATP-dependent addition of CoA. The activated acyl-CoA is dehydrogenated by the flavin-containing enzyme acyl-CoA oxidase (AOX) to produce *trans*-Δ^2-enoyl-CoA. The electrons are transferred to O_2 to produce H_2O_2. This is the most notable difference between β-oxidation in peroxisomes and mitochondria. In mammalian mitochondria this first step of the β-oxidation pathway is carried out by acyl-CoA dehydrogenase, and the electrons are transferred to a flavoprotein (electron-transfer flavoprotein) coupled to the electron transport chain (Stryer, 1988).

Purification of peroxisomal AOX revealed that the enzyme ($M_r \approx 139,000$) consists of 3 components: A, B and C ($M_r \approx 72,000$, 52,000 and 21,000, respectively). B and C are formed by the *in vivo* proteolytic cleavage of A, and the subunits associate to form the active enzyme in the following stoichiometries: A_2, B_2C_2, and ABC (Osumi and Hashimoto, 1980; Miura *et al.*, 1984).

The cDNA (Miyazawa *et al.*, 1987) and gene (Osumi *et al.*, 1987) encoding AOX have been cloned, and the cleavage site for A has been determined (Miyazawa *et al.*, 1987). The AOX mRNA is induced by treatment of rats with clofibrate and is alternatively spliced to yield two transcripts which potentially encode two different proteins differing in the amino acid sequence encoded by exon 3. However, the product

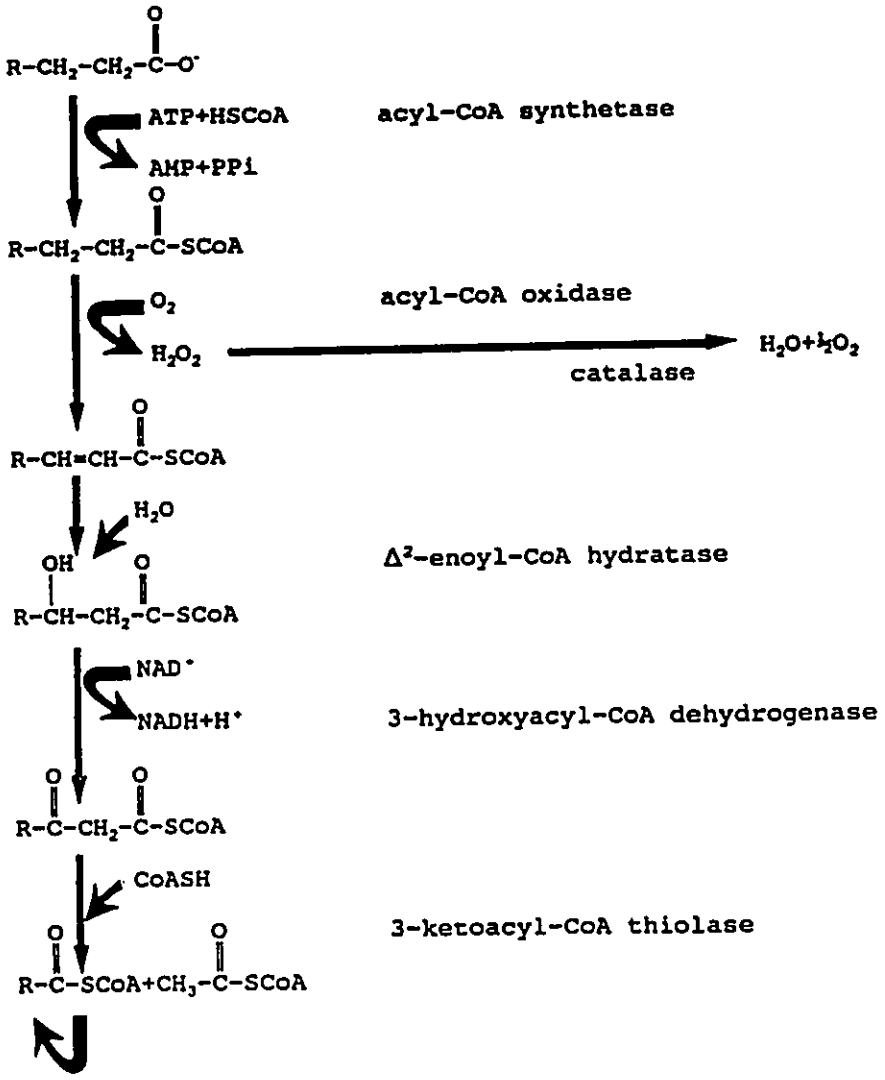


Figure 1.3.1.1 Peroxisomal β-oxidation. See text for details.

of one spliced mRNA predominates, and it is not known whether the second transcript is translated (Miyazawa *et al.*, 1987). Results of Schepers *et al.* (1990) indicate the presence of a second non-inducible peroxisomal AOX. The presence of a second gene, however, remains to be shown. Schepers *et al.* (1990) suggest that the second protein may be a result of the alternative splicing of the single transcript. This would suggest that upon induction and increased transcription of the AOX gene, there would be a corresponding preferential splicing of the transcript to yield the inducible AOX protein (Miura *et al.*, 1984; Schepers *et al.*, 1990).

1.3.1.2 Catalase

Catalase degrades the peroxide formed by AOX (and other peroxisomal oxidases) to form H_2O and O_2 (Lazarow and de Duve, 1976, Lazarow, 1978). Catalase is a heme-containing tetramer with a subunit M_r of ≈65,000 (Lazarow and Fujiki, 1985). Although catalase is only moderately induced upon clofibrate treatments compared to the overall increase in β-oxidation activity (approximately 2-fold *versus* up to 30-fold, Lock *et al.*, 1989), it is required for the degradation of the toxic H_2O_2 and continued activity of the β-oxidation pathway (Hashimoto and Hayashi, 1987). From the cDNA and deduced amino acid sequence, the subunit MW is determined to be 59,758 Da (Furuta *et al.*, 1987).

1.3.1.3 Hydratase-dehydrogenase

The next two reactions are the same as those found in mitochondria but are carried out by a single bifunctional enzyme in rat liver peroxisomes (hydratase-dehydrogenase, HD; Osumi and Hashimoto, 1979). The enoyl-CoA hydratase activity generates 3-hydroxyacyl-CoA, which is then dehydrogenated by the 3-hydroxyacyl-CoA dehydrogenase activity to give 3-ketoacyl-CoA. HD is a single polypeptide with a M_r ≈ 77,000 (Hashimoto, 1982). The cDNA has been cloned and sequenced (Osumi *et al.*, 1985). Recent evidence indicates that in addition to the above mentioned activities, HD also contains Δ^3,Δ^2-enoyl-CoA isomerase activity (Kilponen *et al.*, 1990; Palosaari and Hiltunen, 1990). In addition, it appears that there are two peroxisomal HD isozymes in rat liver, accounting for the apparent 3-hydroxylacyl-CoA epimerase activity seen in peroxisomal extracts. The second enzyme (termed the D-hydratase) shows activity towards D-3-hydroxylacyl-CoA and converts it to 2-trans-enoyl-CoA, which is then hydrated to L-3-hydroxylacyl-CoA and can be dehydrogenated by the first isozyme (L-hydratase) to 3-ketoacyl-CoA (Smeland *et al.*, 1989; Hiltunen *et al.*, 1989). Thus there is not a distinct epimerase enzyme, but the activity is attributed to the sequential action of two hydratases with opposite stereospecificity. The activities of HD are summarized in Fig. 1.3.1.3.1.

1.3.1.4 Thiolase

In the final step of peroxisomal β-oxidation, 3-ketoacyl-CoA is cleaved by

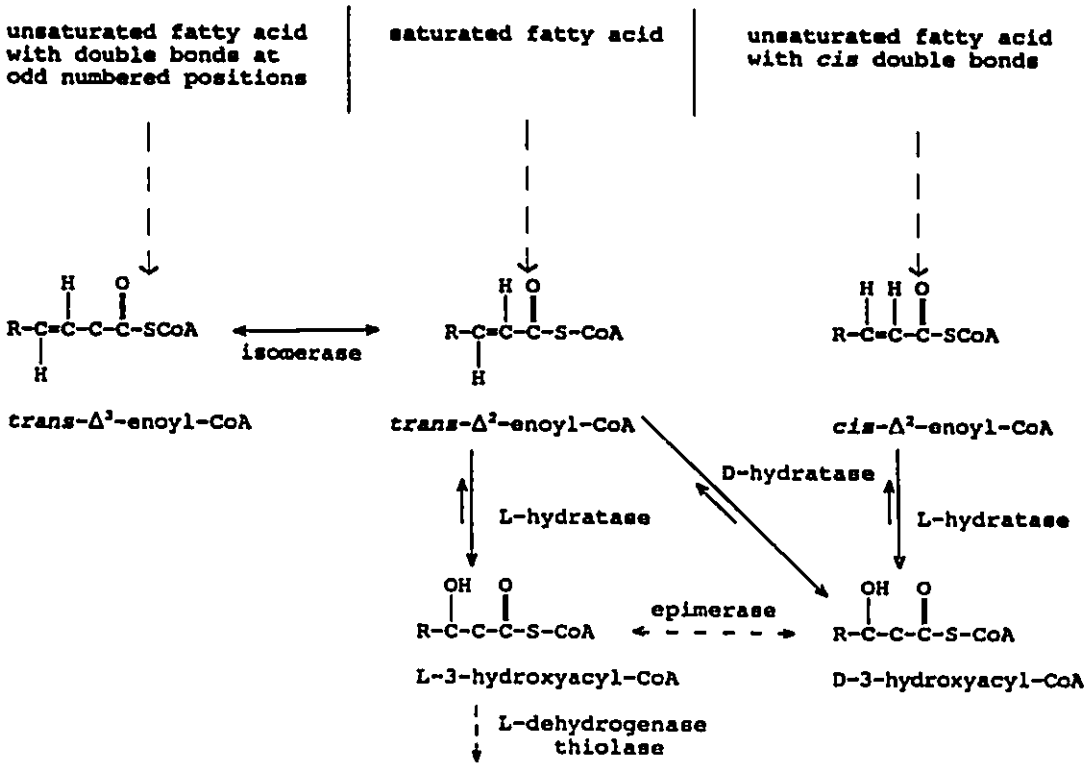


Figure 1.3.1.3.1 Enzymatic activities of HD. See text for details.

3-ketoacyl-CoA thiolase to shorten the activated fatty acid chain length by two C units. The fatty acid then re-enters the ß-oxidation pathway. Peroxisomal 3-ketoacyl-CoA thiolase is a dimer of identical 41,000 Da subunits (Miyazawa *et al.*, 1981; Fujiki *et al.*, 1985). Its expression is induced >10 fold by treatment of rats with clofibrate (Fujiki *et al.*, 1985; Bodnar and Rachubinski, 1990). Thiolase is unique among rat liver peroxisomal proteins in that it is synthesized with a ≈4000 Da amino-terminal extension which is absent from the mature peroxisomal form (Furuta *et al.*, 1982; Fujiki *et al.*, 1985). The role of this precursor in targeting the protein to peroxisomes is discussed in section 4. The thiolase cDNA was cloned by Hijikata *et al.* (1987). It encodes a 398 amino acid protein plus a 26 amino acid extension. The mature protein has a MW of 41,074 Da. Recently Bodnar and Rachubinski (1990) and Hijikata *et al.* (1990) independently cloned the cDNA and gene, respectively, encoding a second non-inducible thiolase. The second thiolase (thiolase 2) has a high identity (95.4%) to thiolase 1 and contains a precursor extension of 36 amino acids (Bodnar and Rachubinski, 1990). Thus rat liver peroxisomes contain a constitutively expressed (thiolase 2) and a clofibrate-inducible (thiolase 1) form of the enzyme.

1.3.1.5 Energy considerations

Although peroxisomal β-oxidation is not directly linked to an electron transport system as in mitochondria, the energy liberated in the second step (3-hydroxyacyl-CoA dehydrognease) is partly conserved in the form of glyceraldehyde-3-

phosphate (G3P) by coupling the reduction of dihydroxyacetone-phosphate (DHAP) to the oxidation of NADH in peroxisomes. G3P then diffuses to mitochondria, where it is converted to DHAP with the concomitant reduction of FAD to $FADH_2$ which can enter the mitochondrial electron transport chain (ETC) (see Fig. 1.3.2.5.1.). The energy of the first oxidation step by acyl-CoA oxidase is lost as heat. This, however, may also be useful to the organism (de Duve, 1983). In this regard, peroxisomes are induced in brown fat tissue of cold acclimated animals (Cannon *et al.*, 1982).

1.3.2 Yeast peroxisomal β-oxidation

In yeast, fatty acids are metabolized solely in peroxisomes (Kawamoto *et al.*, 1978; Tanaka *et al.*, 1982; de Duve, 1983). Therefore, when yeast cells are grown on medium containing fatty acids (or alkanes) as the sole carbon source, peroxisomes play a major role in the energy production and catabolic processes within the cell. Interestingly, yeast respond to growth on fatty acid-containing medium by the induction of peroxisomes and of the enzymes of β-oxidation (Osumi *et al.*, 1974; Osumi *et al.*, 1975; Fukui and Tanaka, 1979; Tanaka *et al.*, 1982). The induction of peroxisomal proteins occurs, at least partially, at the level of transcription (Kamiryo and Okazaki, 1984; Rachubinski *et al.*, 1985; Fujiki *et al.*, 1986). The ability to induce the proliferation of peroxisomes and peroxisomal proteins, along with the ease of genetic manipulation of most yeasts, has made them an attractive target for the study of peroxisome biogenesis and for the cloning of genes encoding peroxisomal proteins (see

below). *C. tropicalis* is the most extensively studied yeast in this regard and will be the topic of this discussion. It is interesting to note, however, that in other yeasts that utilize peroxisomal enzymes for the metabolism of methanol (*e.g. Hansenula polymorpha, Candida boidinii,* and *Pichia pastoris*) there is a similar induction of peroxisomes and peroxisomal enzymes with methanol as the carbon source.

The induction of peroxisomal proteins in *C. tropicalis* by growth on fatty acids was exploited by Kamiryo and Okazaki (1984) to identify a number of peroxisomal proteins. Eighteen peroxisomal proteins (PXPs -1 to -18) were identified on this basis. The corresponding genes have been named *POXs*. However, the nomenclature is not uniform, since many genes and their products are named based on their specific function (see below).

1.3.2.1 Fatty acyl-CoA oxidases

The biochemistry of the β-oxidation system in yeast is essentially the same as that of rat liver peroxisomes; however, the enzymes vary. AOX has 3 isoforms in *C. tropicalis* (PXP-2, -4, -5), each of which is encoded by a separate gene (Rachubinski *et al.*, 1985; Okazaki, *et al.*, 1986; 1987; Murray and Rachubinski, 1987a). A partial cDNA encoding the most abundant isoform, PXP-4 (Okazaki *et al.*, 1986), was isolated by differential hybridization and identified by hybridization selection-translation and immunoprecipitation (Rachubinski *et al.*, 1985). The gene was subsequently cloned by hybridization using the partial cDNA (Murray and Rachubinski, 1987a). Two highly

homologous genes were independently isolated encoding PXP-4 and -5 (Okazaki *et al.*, 1986). PXP-4 and -5 are 709 and 662 amino acids long, respectively (Okazaki *et al.*, 1986). It has been hypothesized that the two genes encode acyl-CoA oxidases with differing specificities, because PXP-4 is induced to a greater extent relative to PXP-5 by growth of cells on oleic acid compared to *n*-alkanes with shorter chain lengths (C_{10}-C_{12}; Okazaki *et al.*, 1986). Picataggio *et al.* (1991) confirmed, by gene disruption in *C. tropicalis*, that PXP-5 has a specificity for long chain fatty acids, whereas PXP-4 has a broad substrate spectrum.

As mentioned above, the gene cloned by Murray and Rachubinski (1987a) corresponds to *POX-4*. The two PXP-4s are identical over 94.2% of their 709 amino acid length but contain a stretch of non-identity between amino acids 359 and 394 in which no amino acids align correctly (Murray and Rachubinski, 1987a). Whether this is a result of allelic heterogeneity or of the presence of an alternate gene (and product) remains to be shown, although Southern blotting of *C. tropicalis* chromosomes has not identified separate genes for the two *POX-4* homologues (Okazaki *et al.*, 1987; Kamiryo *et al.*, 1991).

A third gene encoding AOX (*POX-2*) was recently identified by Okazaki *et al.* (1987). PXP-2 is 724 amino acids long and has an overall homology of 52% and 49% to PXP-4 and PXP-5, respectively (Okazaki *et al.*, 1987). Gel filtration experiments showed that AOX is present as homo-octamers of PXP-4 or 5 and also as a PXP-2/PXP-5 hetero-ocatamer (Lazarow and Fujiki, 1985; Okazaki *et al.*, 1986;

1987). The activity of the latter, remains to be shown (Okazaki *et al.*, 1987).

1.3.2.2 Catalase

As in rat liver, catalase (PXP-9; Kamiryo and Okazaki, 1984) is a tetrameric, heme-binding protein (Lazarow and Fujiki, 1985). Catalase mRNA is induced approximately 25-fold upon transfer of *C. tropicalis* from glucose- to oleic acid-containing medium (Yamada *et al.*, 1982). This induction was used to isolate a partial cDNA encoding peroxisomal catalase by differential hybridization (Rachubinski *et al.*, 1987), which was subsequently used to isolate a full-length cDNA clone (Murray and Rachubinski, 1987b) and gene (Murray and Rachubinski, 1989; see also Okada *et al.*, 1987). The protein is 484 amino acids long and has a MW of 54,767 Da (not including the initiator methionine).

1.3.2.3 Hydratase-dehydrogenase-epimerase

The hydratase and dehydrogenase activities of the peroxisomal β-oxidation system are carried out by a single enzyme. However, in contrast to rat liver HD, the *C. tropicalis* enzyme also contains an associated epimerase activity and is thus called the trifunctional enzyme hydratase-dehydrogenase-epimerase (HDE; Moreno de la Garza *et al.*, 1985). Whether this enzyme acts as a homodimer or as a monomer is currently unclear (Moreno de la Garza *et al.*, 1985; Ueda *et al.*, 1987). The cDNA and gene encoding HDE were cloned during the course of this study, and the results are presented

in detail in sections 3.1 - 3.3.

1.3.2.4 Thiolases

There are at least two thiolases in *C. tropicalis*. Thiolase 1 (PXP-12; Kamiryo *et al.*, 1991) is composed of 6 identical 41 kDa subunits and preferentially acts upon acetoacetyl-CoA (Kurihara *et al.*, 1989). Thiolase 3 (PXP-11; Kamiryo *et al.*, 1991) is a homodimer of 43 kDa subunits and exclusively cleaves 3-ketoacyl-CoAs with long-chain lengths (Kurihara *et al.*, 1989). A potential third thiolase (thiolase 2) remains uncharacterized because of its low acitivity, but it may be a derivative of thiolase 3 (Kurihara *et al.*, 1989). It is not known if *C. tropicalis* thiolase is synthesized as a larger precursor like its peroxisomal counterpart in rat liver.

1.3.2.5 Energy production from fatty acids in yeast

Since fatty acid β-oxidation occurs entirely in peroxisomes of yeast, it is interesting to ask how energy and substrates for other metabolic and biosynthetic processes are derived from fatty acids as a sole carbon source. In addition to shuttling G3P to mitochondria, as in mammalian cells, the presence of two enzymes of the glyoxylate cycle in peroxisomes allows for energy production and carbohydrate synthesis through the interaction of peroxisomal and mitochondrial enzymes. A scheme for the possible metabolic functions of peroxisomes and mitochondria in fatty acid-grown yeast is shown in Fig. 1.3.2.5.1 (Tanaka *et al.*, 1982). Fatty acids are activated in

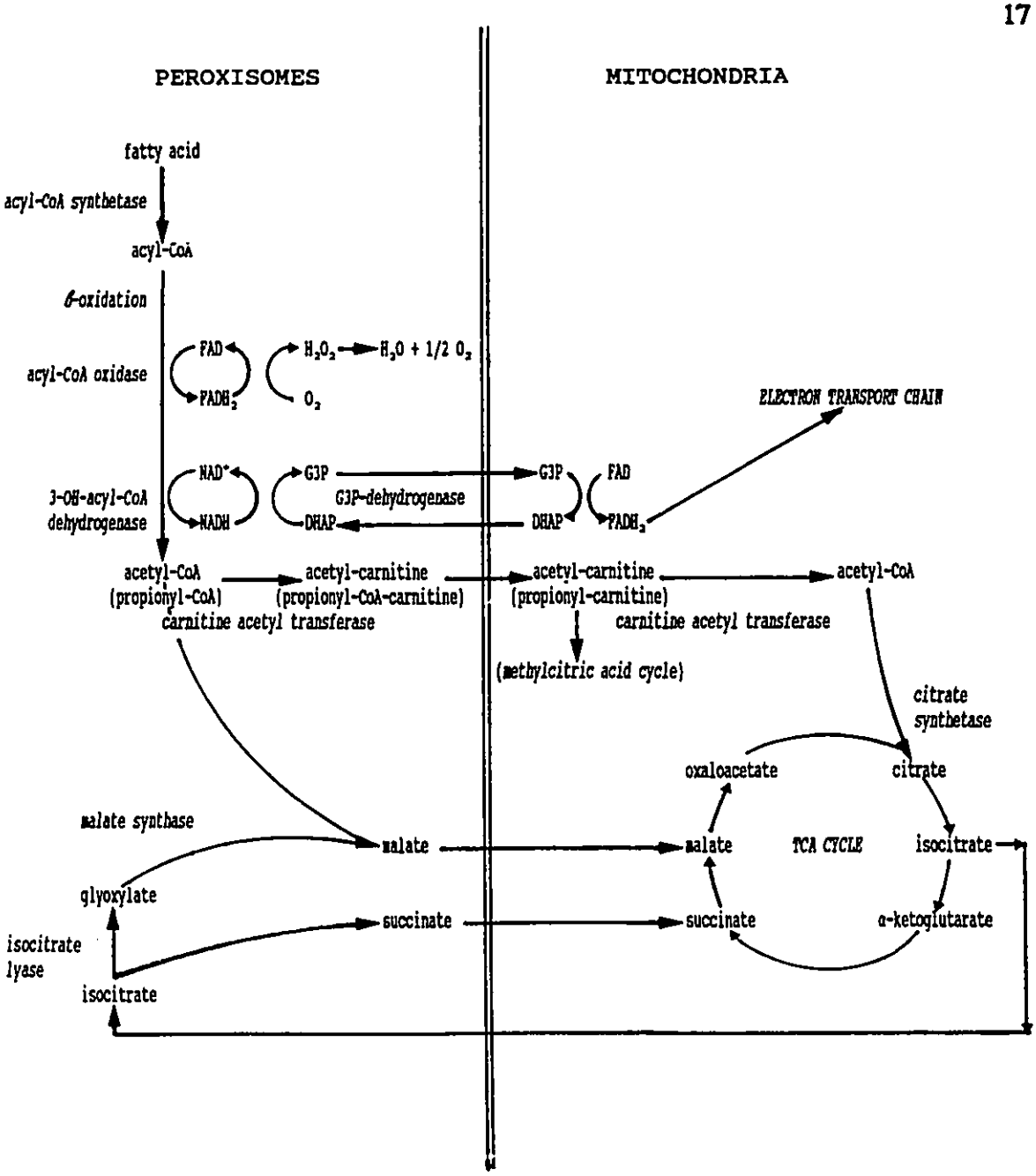


Figure 1.3.2.5.1 Metabolic interactions of yeast peroxisomes and mitochondria grown on fatty acids. See text for details.

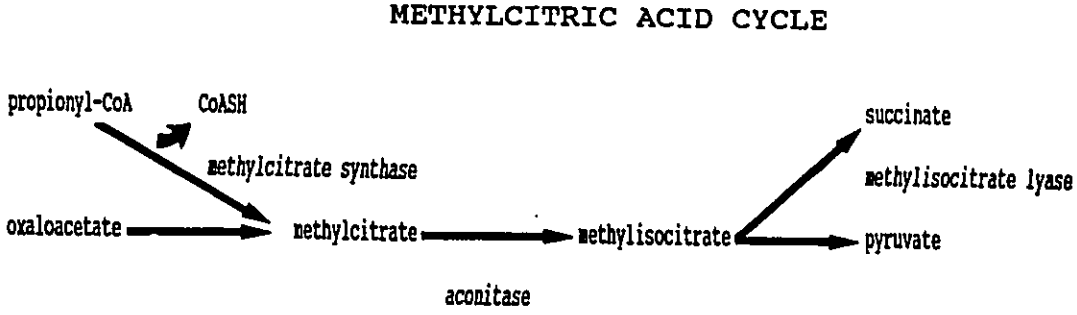


Figure 1.3.2.5.2 Methylcitric acid cycle. See text for details.

peroxisomes, microsomes, or mitochondria by acyl-CoA synthetase. In the last two organelles, these acyl-CoAs can be converted to lipid by glycerol-3-phosphate acyl transferase. In peroxisomes, the acyl-CoA is subjected to β-oxidation, and energy is obtained by coupling the conversion of 3-hydroxyacyl-CoA to trans-Δ^2-acyl-CoA with the reduction of NAD to NADH, followed by the reoxidation of NADH coupled to the reduction of DHAP to G3P by NAD-linked glycerol-3-phosphate dehydrogenase. G3P diffuses to mitochondria where its subsequent oxidaton to DHAP is coupled to the reduction of FAD to $FADH_2$ which can enter the electron transport chain for the production of ATP.

Acetyl-CoA and propionyl-CoA (generated from fatty acids of odd chain lengths) can be transported to mitochondria by an acetyl-carnitine shuttle after esterification with carnitine by carnitine acetyl transferase, which is located in both mitochondria and peroxisomes. Acetyl-CoA can thereby enter the tricarboxylic acid (TCA) cycle or be used in other biosynthetic reactions. The propionyl-CoA is metabolized through the mitochondrial methylcitric acid cycle (Fig. 1.3.2.5.2). The sequential action of methylcitrate synthase, aconitase and methylisocitrate lyase can use oxaloacetate and propionyl-CoA as substrates for the synthesis of succinate and pyruvate.

As an alternative to shuttling directly to mitochondria, acetyl-CoA in peroxisomes is condensed with glyoxylate by the action of malate synthase to form malate. Glyoxylate, in addition to succinate, is generated by the action of isocitrate lyase on isocitrate (from the TCA cycle). Malate and succinate can enter the

mitochondrial TCA cycle for the generation of ATP and the synthesis of carbohydrates, amino acids and other cellular constituents.

1.4 Biogenesis of peroxisomes and peroxisomal proteins

The relatively late discovery and characterization of peroxisomes have left the organelle comparatively unexplored. Consequently, much remains to be learned about its biogenesis. Peroxisomes do not contain DNA (Kamiryo *et al.*, 1982) or any protein synthetic machinery, and therefore peroxisomes rely on the transcription of nuclear genes and the import of proteins from the cytoplasm.

The classical model of peroxisome biogenesis suggested that peroxisomal proteins were synthesized on cytoplasmic polysomes bound to the endoplasmic reticulum (ER) and cotranslationally imported into the ER, where they accumulated in peripheral segments. Peroxisomes were thought to form by the pinching off of ER protrusions (Reddy and Svoboda, 1971; Lazarow, *et al.*, 1982). This hypothesis was based primarily on 2 lines of evidence. Firstly, morphological evidence from electron microsocopy of a number of rat, human and mouse hepatocytes suggested continuities between the ER and peroxisomes (Hruban *et al.*, 1963; Svoboda and Azarnoff, 1966; Essner *et al.*, 1967; Tsukasa *et al.*, 1968; Reddy and Svoboda, 1971; Novikoff *et al.*, 1973; Novikoff *et al.*, 1973; Novikoff and Novikoff, 1982). The second line of evidence was the original, but since disproven (see Table 1.4.1), evidence that catalase is synthesized on membrane-bound polysomes (Higashi and Peters, 1963). However,

other attempts to investigate the continuity between peroxisomes and ER both morphologically (Legg and Wood, 1970; Rigatuso *et al.*, 1970; Wedel and Berger, 1975; Gorgas, 1984; Yamamoto and Fahimi, 1987; Zaar *et al.*, 1987) and histochemically by the diffusion of the reaction products of catalase and glucose-6-phosphatase (Shio and Lazaro, 1981) failed to detect any continuity between the ER and peroxisomes.

In addition almost all peroxisomal proteins (see below) studied thus far have been shown to be synthesized on polysomes free in the cytoplasm (see Table 1.4.1). Many peroxisomal proteins have also been shown to be imported into peroxisomes after transit through the cytosol and do not pass through any other membrane system such as the ER or Golgi apparatus. This has been shown for both matrix and membrane proteins: rat liver HD, AOX, thiolase (Miura *et al.*, 1984) and catalase (Roa and Blobel, 1983), *C. boidinii* alcohol oxidase (Goodman *et al.*, 1984); cucumber malate synthase (Kruse and Kindl, 1983) and glycolate oxidase (Gerdes *et al.*, 1982); isocitrate lyase from castor bean (Roberts and Lord, 1981; Lord and Roberts, 1982) and *N. crassa* (Zimmerman and Neupert, 1980); and rat liver integral membrane proteins, PMPs 22 and 69 (Just and Hartl, 1987; summarized in Table 1.4.1).

Fujiki *et al.* (1982) found that while the peroxisomal and endoplasmic reticular membranes of rat liver contained similar phospholipid profiles, the polypeptide profiles were quite distinct. This also suggested that peroxisomes do not form by pinching off the ER.

Table 1.4.1 Site of synthesis of peroxisomal proteins

Protein	Source	Free/bound	Precursor/ mature	References
acyl-CoA oxidase	rat liver	free	mature	Rachubinski *et al.* 1984 Miura *et al.*, 1984
	C. tropicalis	-[1]	mature	Fujiki *et al.*, 1986 Kamiryo & Okazaki, 1982[2]
hydratase-dehydrogenase	rat liver	free	mature	Rachubinski *et al.*, 1984 Muira *et al.*, 1984
hydratase-dehydrogenase-epimerase	*C. tropicalis*	-	mature	Fujiki *et al.*, 1986 Nuttley *et al.*, 1988
thiolase	rat liver	free	precursor	Furuta *et al.*, 1982 Miura *et al.*, 1984 Fujiki *et al*, 1985
	S. cerevisiae	-	precursor?	R. Erdmann, pers. commun.
catalase	rat liver	free	mature	Lazarow & de Duve, 1973 Goldman & Blobel, 1978 Robbi & Lazarow, 1978 Robbi & Lazarow, 1982 Rachubinski *et al.*, 1984
	C. tropicalis	-	mature	Yamada *et al.*, 1982 Fujiki *et al.*, 1986
	H. polymorpha	-	mature	Roe and Blobel, 1983
	cucumber	-	precursor	Reizman *et al.*, 1980 Becker *et al.*, 1982
	pumpkin	-	precursor	Yamaguchi *et al.*, 1984
	sunflower	-	precursor	Eisig *et al.*, 1990
uricase	rat liver	free	mature	Goldman & Blobel, 1978 Rachubinski *et al.*, 1984

PMP-22	rat liver	free	mature	Fujiki *et al.*, 1984a Hashimoto *et al.*, 1986 Köster *et al.*, 1986 Just & Hartl, 1987 Suzuki *et al.*, 1987 Bodner & Rachubinski, 1990 Kamijo *et al.*, 1990
PMP-69	rat liver	free	mature	Hashimoto *et al.*, 1986 Köster *et al.*, 1986 Just & Hartl, 1987 Suzuki *et al.*, 1987 Bodner & Rachubinski, 1990 Kamijo *et al.*, 1990
PMP-50	rat liver	bound	mature	Bodner & Rachubinski, 1990
PMP-36	rat liver	free	mature	Köster *et al.*, 1986 Bodner & Rachubinski, 1990
PMP-35	CHO cells		mature	Tsukamoto *et al.*, 1991
alcohol oxidase	*C. boidinii*	(free)[3]	mature	Goodman *et al.*, 1984
	H. polymorpha	-	mature	Roa & Blobel, 1983 Roggenkamp *et al.*, 1984
uricase II	soybean	free	mature	Nguyen *et al.*, 198[illegible]
malate dehydrogenase	cucumber	-	precursor	Reizman *et al.*, 1980
	watermelon	free	precursor	Walk & Hock, 1978 Becker *et al.*, 1982 Gietl & Hock, 1982 Gietl & Hock, 1984 Gietl, 1990
malate synthase	cucumber	(free)	mature	Kruse *et al.*, 1981 Kruse & Kindl, 1983
	N. crassa	free	mature	Dessel *et al.*, 1982
	castor bean	-	mature	Lord & Roberts, 1982
isocitrate lyase	castor bean	free	mature	Roberts & Lord, 1981 Lord & Roberts, 1982
	cucumber	-	mature	quoted in Borst, 1986

	N. crassa	free	mature	Zimmerman & Neupert, 1980 Desel *et al.*, 1982
glycolate oxidase	cucumber	(free)	mature	Gerdes *et al.*, 1982
PXP-18	*C. tropicalis*	-	mature	Szabo *et al.*, 1989
PMP-20	*C. boidinii*	-	mature	Garrrd & Goodman, 1989
sterol carrier protein-2	rat liver	free	precursor	Fujiki *et al.*, 1989
carnitine acetyl-transferase	*C. tropicalis*	-	precursor	Ueda *et al.*, 1984
aldolase	*T. brucei*	free	mature	Clayton, 1987 Hartl *et al.*, 1987
glyceraldehyde-3-phosphate dehydrogenase	*T. brucei*	free	mature	Michaels *et al.*, 1986 Hartl *et al.*, 1987
glycerol-3-phosphate dehydrogenase	*T. brucei*	free	mature	Hartl *et al.*, 1987
triosphosphate isomerase	*T. brucei*	-	mature	Wierenga *et al.*, 1987

1. Not determined.

2. In the preliminary identification of *C. tropicalis* peroxisomal proteins (PXPs) Kamiryo and Okazaki (1984) noted that most *in vitro* synthesized PXPs had an identical M_r to their peroxisomal counterparts. The following could be readily identified: PXP-1, -2, -4, -5, -9, -11, -12, -18.

3. Interpreted as synthesized on free polysomes by incorporation into the cytosolic pool. The possibility exists that the protein is synthesized on bound polysomes and is directed to the cytosol rather than the lumen of the ER.

This battery of evidence has led to the hypothesis that peroxisomes arise by the budding or fragmentation of preexisting peroxisomes (Lazarow and Fujiki, 1985) and that proteins are post-translationally imported into the organelle without transit through the ER (Legg and Wood, 1970; Lazarow and Fujiki, 1985).

The "tails" and apparent reticular network by which peroxisomes were thought to associate with the ER have been ascribed to the presence of transient tubular connections between peroxisomes, forming a "peroxisomal reticulum" (Gorgas, 1984) in proliferating cells such as regenerating liver (Yamamoto and Fahimi, 1987). Such reticula were not observed in normal liver (Wedel and Berger, 1975).

Although tubular connections have not been observed in yeast, structures analogous to a peroxisomal reticulum are seen when peroxisomes are induced to proliferate. Clusters of irregularly shaped, sometimes interwoven, peroxisomes have been observed in *C. boidinii* (Veenhuis and Goodman, 1990), *S. cerevisiae* (Veenhuis *et al.*, 1987) and *C. tropicalis* (Osumi *et al.*, 1975). Similarly, small buds attached to larger peroxisomes have been detected in *H. polymorpha* (Veenhuis *et al.*, 1978).

Taken together, this evidence provides strong support for the suggestion that peroxisomes arise from preexisting peroxisomes and suggests that the observed clustering in yeast and the reticulum seen in regenerating (or clofibrate-treated) rat liver are the result of growth, budding and fragmentation of preexisting peroxisomes.

If it is the case that peroxisomes do not arise *de novo*, then all cells which give rise to peroxisome-containing cells must have at least one peroxisome. This has

been shown to be the case in yeast, grown under non-peroxisome-inducing conditions. Sections of *C. tropicalis* grown on malt-extract contained one or two peroxisomes near the cell periphery (Osumi *et al.*, 1975). Immunofluorescence (Thieringer *et al.*, 1991) and electron microscopy studies also detected peroxisomes in glucose- (Veenhuis *et al.*, 1987) and ethanol- (Distel *et al.*, 1987) grown *S. cerevisiae*. In the most detailed analysis thus far of peroxisome proliferation and assembly, Veenhuis and Goodman (1990) found that in *C. boidinii*, glucose-grown cells contained two to five small peroxisomes. Upon transfer of these cells to methanol-containing medium, one or two peroxisomes elongated, formed processes and pinched off. Active peroxisome proliferation yielded clusters of approximately 30 peroxisomes, which subsequently enlarged upon import of matrix proteins. The proliferation of peroxisomal membrane proteins prior to the synthesis and import of matrix proteins has also been observed in regenerating rat liver (Lueers *et al.*, 1990).

Peroxisome proliferation is therefore subject to complex control and it is not simply the import of matrix proteins, which causes peroxisome growth and subsequent fission, as has been suggested (Lazarow and Fujiki, 1985). This point is made evident by the overexpression of alcohol oxidase in *H. polymorpha* under nonmethylotrophic conditions. Such conditioning causes one or two peroxisomes to import and assemble the newly synthesized enzyme. Peroxisomes do not proliferate in response to the overexpression of alcohol oxidase, but rather, one or two peroxisomes become enlarged and irregularly shaped (Distel *et al.*, 1988)

1.5 Origin of peroxisomes

Peroxisomes are common to all eukaryotes except lower eukaryotes lacking mitochondria (Cavalier-Smith, 1987). This finding suggests that, although the current biochemical activities of peroxisomes vary between organisms, all peroxisomes arose from a single primitive organelle. de Duve (1970; 1982; 1983) has suggested that peroxisomes were a rudimentary respiratory organelle, which arose by endosymbiosis prior to the acquisition of mitochondria. It is hypothesized that peroxisomes lacked the ability to carry out oxidative phosphorylation but that they could oxidize all major foodstuffs. The several different oxidases currently found in peroxisomes are thought to be relics of the endosymbiont. The abduction of the aerobic prokaryotic precursor to mitochondria is believed to have permitted the specialization of peroxisomal biochemical functions (de Duve, 1983).

The conclusion that mitochondria and chloroplasts are endosymbionts is based primarily on the existence of DNA and protein synthetic machinery within the organelle, an alternative genetic code, and the existence of prokaryotic counterparts to mitochondria and chloroplasts. No such evidence has been found for peroxisomes. de Duve maintains the endosymbiotic theory for peroxisomes, and argues that only 10% of the mitochondria's DNA remains and therefore the complete loss of DNA by an earlier endosymbiont is not unreasonable. It is also possible that the genotoxic effects of H_2O_2 (Tomaszewski, 1986) generated within the peroxisomes may have helped to select for the transfer of DNA to the nucleus.

The persistence of peroxisomes after the appearance of mitochondria may be attributable to a requirement for compartmentalization of toxic H_2O_2. Hansen and Roggenkamp (1989) have shown that in the absence of catalase to breakdown H_2O_2, the yeast *H. polymorpha* grows very poorly, if at all, with methanol as the sole carbon source. Cytosolic catalase (from *S. cerevisiae*) is unable to complement the deficiency, suggesting either that H_2O_2 does not leak at an appreciable rate into the cytoplasm or that it must be degraded at its site of accumulation. It is generally believed that compartmentalization protects the rest of the cell from H_2O_2 (Chance *et al.*, 1979).

1.6 Zellweger's syndrome

Regardless of the origin of peroxisomes or the cause for their maintenance throughout evolution, the importance of the continued presence of peroxisomes is clearly demonstrated by the human autosomal recessive peroxisomal deficiency diseases where peroxisomes are either completely absent or defective in their assembly. Of the peroxisome deficiencies in humans, which include Zellweger's syndrome, neonatal adrenoleukodystrophy, infantile Refsum's disease and hyperpipecolic acidemia, Zellweger's syndrome demonstrates most dramatically the importance of functional peroxisomes in humans. Children afflicted with Zellweger's syndrome have severe neuronal, hepatic and renal defects and usually die within the first 4 months of life (reviewed in Moser, 1986, 1987, 1988; Moser *et al.*, 1991).

The connection of Zellweger's syndrome to peroxisomes was established by

the work of Goldfischer *et al.* (1973) who found a lack of morphologically recognizable peroxisomes in biopsies from patients with the disease. The biochemical manifestations of Zellweger's syndrome include increased levels of very long-chain fatty acids (Moser *et al.*, 1984), the accumulation of bile acid precursors (Hanson *et al.*, 1979), reduced plasmalogen levels (Heymans *et al.*, 1983; Datta *et al.*, 1984) and increased levels of pipecolic acid (Danks *et al.*, 1975). These various biochemical abnormalities can all be attributed to dysfunctional peroxisomes in Zellweger's patients. Enzymes required for the β-oxidation of very long-chain fatty acids (Singh *et al.*, 1984), in particular HD, AOX, and thiolase (Tager *et al.*, 1985; Suzuki *et al.*, 1986), are deficient. The activities of alkyl dihydroxyacetone phosphate synthase (Schrakamp *et al.*, 1985) and dihydroxyacetone phosphate synthase (Schutgens *et al.*, 1984), which catalyze the initial steps of peroxisomal plasmalogen biosynthesis (Hajra and Bishop, 1982), are also severely reduced. In addition, pipecolic acid degradation (Mihalik *et al.*, 1989) and bile acid synthesis (Pedersen and Gustafsson, 1980) by peroxisomes are diminished in cells from patients with Zellweger's syndrome (Moser, 1988).

Other peroxisomal proteins including catalase (Wanders *et al.*, 1984; Santos *et al.*, 1985), L-α-hydroxyacid oxidase and D-amino acid oxidase (Wanders *et al.*, 1984; Moser, 1988) seem to be present in (almost) normal amounts but are mislocalized to the cytosol.

The autosomal recessive mode of the inheritance of Zellweger's syndrome implies that the presence of a single genetic defect causes each episode of the disorder

and its pleiotrophic effects. Since it appears that several peroxisomal functions are aberrant, it has been suggested that the defect is either a result of the complete lack of peroxisomes, or alternatively, the lack of an ability to import matrix proteins. Many proteins unable to enter peroxisomes are degraded in the cytosol, while others persist (Santos *et al.*, 1985; 1988a; 1988b). These findings are supported by evidence from pulse-chase studies in fibroblasts from Zellweger's patients that indicate that AOX and thiolase are synthesized normally, but are not processed to their mature forms and are rapidly degraded.

Santos *et al.* (1988a; 1998b) demonstrated that peroxisomal membranes are indeed present in fibroblasts from Zellweger's patients. The 'peroxisomal ghosts', which contain peroxisomal 22, (Santos *et al.*, 1988a) 53 and 69 kDa (Santos *et al.*, 1988b) integral membrane proteins (IMPs), lack matrix proteins. Antiserum directed against the rat 22 kDa IMP cross-reacts with the human 22 kDa counterpart and was used in immunofluorescence studies to identify peroxisomal ghosts (Santos *et al.*, 1988a). Subsequently, two other peroxisomal membrane proteins were found to be associated with the peroxisomal ghosts (Santos *et al.*, 1988b). This evidence implicates a matrix protein import defect in Zellweger's cells and suggests that membrane proteins are integrated into the organelle in a manner different from that of matrix proteins. The absence of peroxisomal functions in Zellweger's patients is therefore attributed to the inability of soluble peroxisomal proteins to be imported into the organelle and has thus been a significant motivator for much of the research surrounding the mechanisms of

peroxisomal assembly. Most of these efforts have concentrated on the signals that direct peroxisomal proteins to their correct intracellular location. To put the current state of research on peroxisomal targeting in context, a brief overview of the signals that mediate transport to other organelles in cells will be presented.

1.7 Protein transport to organelles

One of the fundamental problems faced by a cell is how to direct proteins to specific intracellular locations where they are to perform their specific functions. The signal hypothesis provides an elegant explanation of how this might be achieved. The essential postulates of the signal hypothesis as it was initally proposed (Blobel and Sabatini, 1971; cited in Meyer, 1991) are that proteins which are translocated across (or integrated into) a particular cellular membrane carry organelle-specific signal sequences and that specific signal sequence recognition factors mediate the initial events of membrane translocation at the organelle (Blobel and Dobberstein, 1975a; Blobel, 1980). The initial hypothesis has been borne out by experimentation and holds true for all proteins which are specifically targeted to an intracellular organelle. Although there exist several exceptions (see Verner and Schatz, 1988; Meusch *et al.*, 1990; Meyer, 1991) and the precise intracellular pathways and mechanisms of translocation to, and across, membranes are yet to be elucidated, the general nature of signal-mediated intracellular protein trafficking is outlined below.

1.7.1 The secretory pathway

Proteins that are destined for secretion from the cell, (and for integration into the plasma membrane; see below) are initially directed to the ER membrane by a cleavable amino-terminal signal sequence. Soluble proteins are translocated across the ER membrane and, from here, begin their transit to the cell surface. Soluble proteins subsequently flow from the ER to the Golgi apparatus, through the Golgi stacks (from the *cis* to *trans* stacks) and from the *trans* Golgi network to the plasma membrane. Transport is mediated by lipoprotein transport vesicles which sequentially traverse the pathway by budding from the precursor compartment and fusing with the membrane of the destination compartment (reviewed in Rothman and Orci, 1992). During their transit, proteins are extensively post-translationally modified by proteolysis, glycosylation or fatty acylation. Such modifications are compartment-specific and are diagnositic of the position of a protein along the pathway. Proteins carry (and acquire) signals that allow for their retention in specific compartments along the pathway. It is generally accepted that in the absence of additional signals, proteins are directed to the plasma membrane for secretion by bulk flow (Pfeffer and Rothman, 1987; Verner and Schatz, 1988; Rothman and Orci, 1992).

Although the primary sequences that mediate the initial targeting of a protein to the ER vary considerably, signal sequences are usually 13 to 35 amino acids in length and have a tripartite structure. There is a basic amino-terminal region spanning the first 2 to 10 residues, followed by an α-helical stretch of at least 9 hydrophobic or neutral

residues, and a terminal helix-breaking amino acid (pro or gly). The signal peptide cleavage site is generally marked by the presence of small apolar amino acids approximately 3 and 5 amino acids downstream of the helix breaker (positons -1 and -3 relative to the cleavage site) (reviewed in Verner and Schatz, 1988; von Heijne, 1990).

Secretory proteins begin their synthesis on ribosomes free in the cytoplasm, but upon emergence of the signal sequence, synthesis is arrested by the binding of the signal recognition particle (SRP; Walter *et al.*, 1981; Walter and Blobel, 1981). In mammalian cells, SRP is an 11 S ribonucleoprotein comprised of six different polypeptides (9, 14, 19, 54, 68 and 72 kDa) and a 7 S (300 nucleotide) RNA (Walter and Blobel, 1980; 1982). The tertiary complex is then routed to the ER membrane where it is bound by the SRP receptor (or docking protein) (Walter *et al.*, 1979; Meyer and Dobberstein, 1980; Gilmore *et al.*, 1982; reviewed in Walter *et al.*, 1984). The SRP receptor is a dimer composed of SRα and SRβ subunits (reviewed in Andrews and Rachubinski, 1990). Binding to the SRP receptor causes the dissociation of the signal sequence and the ribosome-nascent chain complex from SRP in a cycle involving the binding of GTP to SRα and subsequent hydrolysis of the GTP (Connolly *et al.*, 1991). At this stage the signal sequence becomes associated with the signal sequence receptor α subunit (SSRα), a major, glycosylated 34 kDa, ER membrane protein (Wiedmann *et al.*, 1987; Thrift *et al.*, 1991; reviewed in Rapoport, 1990). SSRα and other proteins, including the β-subunit of the SSR, were identified by crosslinking to the signal sequence (Rapoport, 1990; Thrift *et al.*, 1991). The ribosome remains associated with

the ER membrane by, as yet, poorly characterized ribophorins (see Andrews and Rachubinski, 1990), and translation proceeds (Walter and Blobel 1981). Proteins are cotranslationally translocated across the ER membrane through a proteinaceous pore (Simon and Blobel, 1991), a component of which may be the signal sequence receptor (Rapoport, 1990; Thrift *et al.*, 1991). During protein translocation, the signal sequence is cleaved by a membrane-bound signal peptidase on the lumenal side of the ER membrane (Blobel and Dobberstein 1975b; reviewed in Verner and Schatz, 1988).

Membrane proteins that traverse the secretory pathway generally follow the same initial mechanisms for translocation across the ER as do soluble proteins. However, the cotranslational translocation of membrane proteins is halted by a hydrophobic 'stop transfer' sequence carboxy-terminal to the signal sequence (Blobel, 1980; Pugsley, 1989). Integration of the stop transfer sequence into the membrane is presumably mediated by a conformational alteration of the proteinaceous aqueous pore (Simon and Blobel, 1991).

After this initial translocation event, secondary signals within the proteins serve to remove proteins from vesicle-mediated bulk flow to the cell surface.

1.7.1.1 Retention in the endoplasmic reticulum

The sequence Lys-Asp-Glu-Leu-COOH (KDEL) serves to retain proteins in the lumen of the ER (reviewed in Pelham, 1990; 1991). Pelham and coworkers noted the presence of this tetrapeptide at the carboxy-terminus of three ER resident proteins:

glucose-regulated (grp)-78 (BiP), grp-94 and protein disulphide isomerase. Deletion or alteration of the tetrapeptide in BiP caused the secretion of the protein when expressed in COS cells. When the KDEL sequence was fused to the carboxy-terminus of a protein which is normally secreted (chicken lysozyme), the protein was efficiently retained within the ER (Munro and Pelham, 1987). These experiments demonstrated that the sequence KDEL is necessary and sufficient for the retention of proteins within the ER. Subsequently, several others have identified variations on the prototypical sequence which act as ER retention signals in mammalian cells and in other organisms (reviewed in Pelham, 1990; see also Robbi and Beaufay, 1991).

Proteins containing the ER-retention signal are delivered to the Golgi by bulk flow, as evidenced by Golgi-specific carbohydrate modifications (Pelham, 1988a; Dean and Pelham, 1990), but are prevented from continuing through the Golgi stacks by means of a receptor in the *cis*-Golgi (salvage compartment) (Lewis *et al.*, 1990) and perhaps even in later Golgi compartments (Rothman and Orci, 1992). Vesicles from the salvage compartment are recycled to the ER (retrograde transport), where the contents are released by fusion with the ER membrane (reviewed in Rothman and Orci, 1992).

Lewis *et al.* (1990) identified a receptor for the ER retention signal by the exploitation of yeast genetics. The yeast *Kluyveromyces lactis* recognizes the ER retention signals DDEL and HDEL, whereas *S. cerevisiae* recognizes only HDEL. Random mutagenesis of *S. cerevisiae* identified a gene (*ERD2*) which, when mutated (*erd2*), allowed for the secretion of normally ER resident proteins. Transformation of

the *K. lactis ERD2* gene into *S. cerevisiae* conferred the ability to retain DDEL-containing proteins, strongly suggesting that the *ERD2* gene product is responsible for specific recognition of proteins bearing the ER retention signal. The mammalian homologue of *ERD2* has also been isolated by PCR using degenerate oligonucleotide primers (Lewis and Pelham, 1990). The *ERD2* gene product is a 26 kDa IMP (Semenza *et al.*, 1990) (25 kDa in mammalian cells; Lewis and Pelham, 1990) which has been localized by immunofluorescence to the ER and Golgi apparatus (Lewis and Pelham, 1990).

In an independent effort, anti-idiotypic antibodies were used to identify a putative KDEL-binding protein in mammalian cells. This 72 kDa protein is specific to the Golgi is membrane-bound and has been suggested to be the salvage receptor (Vaux *et al.*, 1990). The ability of this protein to bind KDEL *in vivo* and act as a salvage receptor remains to be shown (Pelham, 1991). It has been suggested that perhaps both proteins operate in the salvage pathway; however their precise contributions to ER protein retention are unclear (Warren, 1990).

A similar, but distinct, ER retention signal has been proposed for ER membrane proteins. The carboxy-terminal signal Lys-Lys-X-X (KKXX) and Lys-X-Lys-X-X (KXKXX) is exposed at the cytoplasmic tail of several ER resident membrane proteins and is sufficient to retain other transmembrane proteins which normally exit from the ER (Nilsson *et al.*, 1989; Jackson *et al.*, 1990). The KDEL signal may also function in ER-resident membrane proteins. The *S. cerevisiae* ER membrane protein

sec 20p is a 50 kDa integral membrane protein which terminates with HDEL exposed to the lumenal side of the ER (Sweet and Pelham, 1991). It is likely that this signal operates as a retention signal by virtue of its lumenal orientation.

1.7.1.2 Retention in the Golgi apparatus

There are several proteins which must be retained in specific compartments of the Golgi apparatus where they are involved in Golgi-specific post-translational modifications such as carbohydrate trimming and additions, palmitoylation and tyrosine sulfations (reviewed in Pugsley, 1989). Two membrane proteins localized to the Golgi apparatus of mammalian cells (α-2,6 sialyltransferase, Munro, 1991, and coronavirus E1 protein, Swift and Machamer, 1991) are withdrawn from the bulk flow by way of a hydrophobic membrane-spanning segment. Deletion of, or substitutions within, these segments allow the proteins to be expressed at the cell surface. In addition, when these sequences were transferred to proteins which are normally destined to the plasma membrane, the chimeras were retained in the Golgi (Munro, 1991; Swift and Machamer, 1991). How these segments are recognized or how retention in the Golgi is achieved is unknown.

The Golgi retention signal may be different in yeast. The kex2 protease and dipeptidyl aminopeptidase A (DPAP A) seem to carry sequences on their cytoplasmic tails which retain them in the Golgi. Alterations of a specific Tyr residue from the carboxy-terminal cytoplasmic tail of kex2 protease (Fuller *et al.*, 1991) or of Phe

residues in the tail of DPAP A (Nothwehr *et al.*, 1991) cause the proteins to be mislocalized to the vacuole, where they are degraded. The precise amino acids involved in the retention of these proteins remain to be shown. Interestingly, these results also suggest that the default pathway for membrane proteins in yeast is the vacuole rather than the plasma membrane.

1.7.1.3 Targeting to lysosomes

Proteins destined for lysosomes are sorted from those destined for the plasma membrane in the *trans* Golgi network. Soluble lysosomal proteins are phosphorylated shortly after export from the ER, which marks them for transport to the lysosome. Phosphorylation takes place in two steps. *N*-acetylglucosamine-1-phosphate is added to the 6-hydroxy position of one or more mannose residues of *N*-linked oligosaccharide chains. Next, the *N*-acetylglucosamine residue is removed to generate mannose-6-phosphate (reviewed in Kornfeld and Mellman, 1989). Mannose-6-phosphate receptors (MPR) in the *trans* Golgi network bind specifically to the mannose-6-phosphate residues. Once bound, receptors migrate to a budding site in the Golgi membrane. The membrane pinches off, and the vesicle is transported to a prelysosomal compartment. The low pH of the prelysosomal compartment facilitates the release of lysosomal enzymes from the receptor. The receptor is subsequently recycled to the Golgi apparatus, while the lysosomal enzymes are transported to the lysosome. The precise mechanism of how this is achieved is unclear.

Two related, partially redundant integral membrane MPRs have been characterized. The smaller MPR is a homodimer of 28.5 kDa subunits, which binds ligands in a cation-dependent (CD) manner. The larger MPR is a 270 kDa monomer and binds ligands independently of cations (CI) (Kornberg and Mellman, 1989).

Although the MPRs are fairly well characterized, how the phosphotransferase recognizes lysosomal proteins to mark them for targeting to the lysosome is poorly understood. No amino acid sequence homology is evident among lysosomal proteins. The phosphotransferase may recognize simple primary or secondary structural domains on the lysosomal proteins (Kornberg and Mellman, 1989). Recently, Baranski *et al.* (1991) mapped the phosphotransferase recognition domain of the lysosomal hydrolase cathepsin D which, when transferred to a secreted protein (glycopepsinogen), targets the chimera to lysosomes. Mutagenesis studies and molecular modelling predictions revealed an extended β-loop containing a number of specific amino acids, as well as a specific Lys outside of the loop, required for recognition by the phosphotransferase. The specific amino acids involved in the protein-protein interaction are, however, not known (Baranski *et al.*, 1991). Nevertheless, it is evident that the recognition domain of soluble lysosomal proteins for the phosphotransferase is complex, involving multiple interacting sites.

There are several lysosomal membrane proteins which appear to be targeted to lysosomes by a mechanism independent of the MPRs. Like soluble lysosomal enzymes, they traverse the ER and Golgi stacks and appear to be sorted at the *trans*

Golgi network, but they do not contain the mannose-6-phosphate recognition marker. The specific signals mediating this targeting are not known (Kornfeld and Mellman, 1989).

1.7.1.4 Targeting to the yeast vacuole

The vacuole is the yeast equivalent of the lysosome of higher eukaryotes (Kornfeld and Mellman, 1989). As for protein targeting to lysosomes, proteins bound for vacuoles are sorted from the bulk flow at the *trans* Golgi. However, the signals which mediate this event seem to have diverged over the course of evolution. The most extensively studied vacuolar protein is carboxypeptidase Y (CPY), which serves as the prototype for soluble protein targeting to the vacuole (reviewed in Rothman *et al.*, 1989). CPY is synthesized as a prepropeptide (preproCPY) containing an ER signal sequence which, when cleaved, leaves an $\approx$10 residue amino-terminal sequence (proCPY) which, in turn, is cleaved upon activation of the enzyme at or near the vacuole (Stevens *et al.*, 1982; Rothman *et al.*, 1989). Deletion mutagenesis of preproCPY identified the propeptide (amino acids 21 to 34 of preproCPY) as containing essential vacuolar targeting information (Rothman *et al.*, 1989). Deletion and substitution mutagenesis subsequently revealed a tetrapeptide (Gln-Arg-Pro-Leu; QRPL) that is essential for vacuolar targeting (Valls *et al.*, 1990).

The ability of this tetrapeptide to sort non-vacuolar proteins to vacuoles has not been shown (Valls *et al.*, 1990). In addition, this sequence is not conserved as a

vacuolar targeting element in other propeptides, at least one of which (proteinase A; Klionsky *et al.*, 1988) has been shown to contain vacuolar targeting information (Valls *et al.*, 1990). More experiments are therefore required to define the signals that direct proteins to vacuoles in yeast.

1.7.2 Targeting to mitochondria

Although mitochondria contain their own DNA and protein synthetic machinery, over 90% of mitochondrial proteins are encoded by nuclear genes and are imported from the cytoplasm into the organelle. Most mitochondrial proteins are synthesized as larger precursors with amino-terminal presequences that contain the mitochondrial targeting information (see Hartl *et al.*, 1989; Pugsley, 1989, for exceptions). Presequences are generally 20-80 amino acids in length, contain a propensity of basic and hydroxylated amino acids, are devoid of acidic amino acids, and lack primary structural similarities. Although it is unclear what structural motif(s) mediate targeting to mitochondria, secondary structural predictions indicate that most presequences can adopt an amphipathic α-helix when exposed to a lipid membrane, with the hydrophobic face oriented toward the bilayer and the polar residues facing outward (von Heijne, 1986; reviewed in Verner and Schatz, 1988; Hartl *et al.*, 1989; Pugsley, 1989; Pfanner and Neupert, 1990). In addition, deletions or substitutions of positively charged amino acids within mitochondrial presequences severely hinder their ability to target proteins to mitochondria (Hartl *et al.*, 1989). It has been hypothesized that the

positive charge of the presequence responds to the electrochemical potential ($\Delta\Psi$) across the mitochondrial inner membrane, an essential component of mitochondrial protein import (Pfanner and Neupert, 1990).

Several components of the mitochondrial import machinery have been identified (reviewed in Neupert *et al.*, 1990; Pfanner and Neupert, 1990; Baker and Schatz, 1991; Pfanner *et al.*, 1991). The current understanding of how these factors interact to import mitochondrial proteins is presented below (reviewed in Pfanner and Neupert, 1990; Pfanner *et al.*, 1991).

Proteins are synthesized on free polysomes in the cytosol and are maintained in an unfolded, import competent state by the binding of cytosolic heat shock protein-70 (hsp70), presequence binding factor (Murikami and Mori, 1990), and probably other factors (Pfanner *et al.*, 1991). The role of hsps (also called chaperonins and polypeptide chain binding proteins) is emerging as a common theme for the import of proteins into intracellular organelles, as well as for the folding and assembly of proteins and multimeric complexes (for reviews see Meyer, 1988; Pelham, 1988b; Rothman, 1989; Baker and Schatz, 1991). *S. cerevisiae* mutants defective in cytosolic hsp70 are unable to import mitochondrial (and secretory; Chirico *et al.*, 1988) proteins (Deshaies *et al.*, 1988; Kang *et al.*, 1990). ATP hydrolysis is believed to facilitate the removal of hsps from the precursor (Neupert *et al.* 1990). ATP dependence has been demonstrated for import of folded precursors *in vitro* (Pfanner *et al.*, 1987; Verner and Schatz, 1987; Pfanner and Neupert, 1990), whereas proteins which are partially denatured with urea

prior to translocation are imported *in vitro* in an ATP-independent manner (Eilers *et al.*, 1988; Eilers and Schatz, 1988; reviewed in Pfanner and Neupert, 1990).

Precursor proteins bind to one of two outer membrane proteins [MOM19 (Söllner *et al.*, 1989) or MOM72=MAS70 (Riezman *et al.*, 1983; Söllner *et al.*, 1990)]. Proteins containing mitochondrial presequences are believed to interact preferentially at the surface of the outer mitochondrial membrane (OMM) via MOM19. In contrast, MOM72 recruits proteins with internal signal sequences (*e.g.* ADP/ATP carrier and possibly others; Söllner *et al.*, 1990). *In vivo* the two receptors can compensate for the loss of each other and thus are thought to have overlapping specificities (Pfanner and Neupert, 1990; Pfanner *et al.*, 1991). MOM19 (in a complex with general insertion protein GIP, a component of which is MOM38=ISP42; Vestweber *et al.*, 1989) and MOM72 are variously distributed over the OMM and are believed able to diffuse laterally to areas of contact between the inner mitochondrial membrane (IMM) and the OMM where translocation occurs. Protein bound to MOM72 is transferred to the MOM19-GIP complex at the contact site, which serves to insert proteins into the OMM (Pfanner and Neupert, 1990). The translocation apparatus appears to form an aqueous pore and is comprised of MOM19-GIP (MOM38), p32 (Pain *et al.*, 1990), MOM22, and probably several other components (Pfanner *et al.*,1991). Precursors are translocated through the putative channel in an unfolded conformation and are "pulled" across the membrane by the high affinity binding of hsps for unfolded proteins (Pfanner *et al.*, 1987; Neupert *et al.*, 1990; Pfanner and Neupert, 1990). Presequences are

cleaved by the matrix-located mitochondrial processing peptidase (MPP) in cooperation with the IMM-associated processing enhancing peptidase (PEP) (Hawlitshek *et al.*, 1988; Pfanner and Neupert, 1990). Proteins are then aided in their folding or are maintained in a translocation competent state for sorting within the mitochondrion by interaction with the matrix hsp60 (Pfanner and Neupert, 1990).

1.7.2.1 Intramitochondrial sorting

As mentioned above, the amino-terminal regions of precursor proteins encode mitochondrial targeting information. In addition, some proteins contain non-cleavable, internal mitochondrial targeting signals (*e.g.* ADP/ATP carrier). In general the amino-terminal presequence is sufficient to direct cytosolic proteins to the mitochondrial matrix (Hartl *et al.*, 1989). However, proteins directed to mitochondria must be sorted to one of at least 4 specific intramitochondrial locations: the mitochondrial matrix, IMM, OMM or the intermembrane space (IMS). Additional signals within mitochondrial proteins serve to label them for their specific destinations (reviewed in Hartl *et al.*, 1989).

Outer membrane proteins are made without cleavable presequences and are integrated directly from the cytosol into the membrane by GIP in a manner analogous to membrane proteins targeted to the secretory pathway. The 70 kDa protein from *Neurospora crassa* is the best characterized OMM protein with respect to its targeting. It contains an amino-terminal non-cleavable mitochondrial targeting signal, the first 12 amino acids of which are sufficient to direct cytosolic proteins to the mitochondrial

matrix. Carboxy-terminal to the mitochondrial targeting signal is a hydrophobic stretch of amino acids encoding a 'stop transfer' signal, which is responsible for terminating the complete transfer of the protein to the matrix and acts to anchor the protein in the OMM (for review and exceptions see Hartl *et al.*, 1989; see also Ngyuen *et al.*, 1988; Liu *et al.*, 1990).

Proteins targeted to the IMS are generally sorted by a so-called 'conservative sorting' mechanism. Proteins are directed to the matrix by the same pathway as matrix-localized proteins, but cleavage in the matrix generates an intermediate containing an additional amino-terminal extension (a hydrophobic stretch of 21 amino acids in the case of cytochrome b_2). It is thought that the second presequence contains sorting information for export from the matrix into the IMS in a second translocation event (for review and exceptions see Hartl *et al.*, 1989).

Proteins destined for the IMM also seem to be directed to the matrix prior to their integration into the membrane via the same mechanism as for other mitochondrial proteins. However, the difference between stop transfer sequences which are thought to play a role in directing proteins to the OMM and hydrophobic segments which allow for integration of proteins into the IMM is unclear (Hartl *et al.*, 1989). It is likely that the context of presentation (in particular, proximity to the amino terminus; Singer and Yaffe, 1990) of such hydrophobic segments is important to differential targeting (Ngyuen *et al.*, 1988; Liu *et al.*, 1990).

1.7.3 Targeting to chloroplasts

The limited chloroplast genome necessitates that most chloroplast enzymes be encoded by nuclear DNA and imported into the organelle from the cytoplasm in an ATP-dependent manner (de Boer and Weisbeek, 1991). Many aspects of chloroplast protein targeting (reviewed in Cashmore *et al.*, 1985; Schmidt and Mishkind, 1986; Keegstra, 1989; de Boer and Weisbeek, 1991) are similar to aspects of mitochondrial protein targeting, discussed above, although relatively little is known about the components which mediate specific targeting and sorting of chloroplast proteins.

Most proteins targeted to chloroplasts are synthesized as larger precursors containing amino-terminal extensions of variable length (28 to 83 amino acids; de Boer and Weisbeek, 1991; von Heijne and Nishikawa, 1991) which are subsequently cleaved by a stromal processing peptidase (Robinson and Ellis, 1984; Keegstra, 1989). Many such extensions have been shown to encode chloroplast targeting signals that are both necessary to target proteins to chloroplasts and sufficient to target foreign proteins to the organelle and are thus termed transit peptides (reviewed in van den Broeck *et al.*, 1985; Keegstra, 1989; de Boer and Weisbeek, 1991). Sequence analysis of several transit peptides has failed to reveal any obvious conserved primary or secondary structural motifs (De Boer and Weisbeek, 1991), although three domains within transit peptides have been proposed (Keegstra, 1989; von Heijne *et al.*, 1989). The amino-terminal domain ($\approx$ 10 amino acids) is generally uncharged and contains few Pro or Gly residues. The central domain is variable in length, rich in basic, Ser and Thr residues, and lacks

acidic residues. The carboxyl region (8-10 amino acids) often has the potential to form an amphiphilic β-strand and frequently contains Arg residues near positions -2 and -10 relative to the cleavage site (von Heijne *et al.*, 1989; de Boer and Weisbeek, 1991) which is often marked by the Val/Ile-X-Ala/Cys↓Ala consensus sequence (Gavel and von Heijne, 1990). The specific portion(s) of the transit peptide responsible for chloroplast protein targeting, or how these regions interact to form a functional transit peptide, are unknown. In addition, as the sequences of more transit peptides become available, these conserved blocks have been shown not to be present in all transit peptides (de Boer and Weisbeek, 1991). This may be because there are different transit peptide receptors specific for certain classes of transit peptides. There may also be a tertiary structural motif, not yet identified, that is common to all transit peptides. It has been suggested (von Heijne and Nishikawa, 1991) that the lack of structure in a transit peptide may be important for directing proteins to chloroplasts. However, this is difficult to imagine within the context of a specific chloroplast transit peptide receptor (see below).

Import into the chloroplast stroma is believed to occur through a proteinacous channel at sites of contact between the inner and outer chloroplast envelope membranes. ATP has been shown to be required for import of precursor proteins *in vitro*, but in contrast to precursor import into mitochondria, the existence of an electrochemical gradient is not a prerequisite for import (Keegstra, 1989; de Boer and Weisbeek, 1991). Hsps have also been shown to be involved in import, both on the cytoplasmic side and

on the stromal side of the putative channel (Marshall *et al.*, 1990; Waegemann *et al.*, 1990).

Two potential transit peptide receptors have been identified. A 66 kDa envelope protein (Keegstra, 1989; de Boer and Weisbeek, 1991) was identified by cross-linking to the precursor of the small subunit of ribulose bisphosphate carboxylase. Pain *et al.* (1988) also used the transit peptide of the small subunit of ribulose bisphosphate carboxylase to identify an outer envelope protein with anti-idiotypic antibodies. These anti-idiotypic antibodies bind to the surface of chloroplasts at contact sites and recognize a 30 kDa (36 kDa calculated MW with transit sequence) integral membrane protein of the envelope that functions as a transit peptide receptor (Schnell *et al.*, 1990).

1.7.3.1 Intrachloroplast sorting

Like proteins targeted to mitochondria, chloroplast proteins, once directed to the chloroplast by their transit peptides, must be sorted to their specific location within the organelle. There are six different intraorganellar locations available to precursor proteins: the stroma (matrix), the thylakoid lumen or membrane (harbours photosynthetic electron transport components), the inner or outer envelope membrane, or the intermembrane space. By analogy to mitochondrial targeting, it is expected that secondary signals within proteins allow for segregation of precursors.

Proteins destined for the thylakoid lumen and membrane, the inner membrane and the intermembrane space are suspected to be targeted to the stroma, where the

transit peptide is cleaved, prior to their transport to their respective destinations (Keegstra, 1989; de Boer and Weibeek, 1991). Thylakoid membrane proteins are thought to be inserted into the membrane by, as yet, unidentified membrane-spanning topogenic sequences within the mature protein. Proteins destined to the thylakoid lumen contain a second lumenal targeting signal adjacent to the cleaved stromal targeting signal. However, at least for plastocyanin, this signal probably interacts with lumenal sorting machinery prior to cleavage of the stromal signal (Bauerle and Keegstra, 1991). The thylakoid transfer signals are rich in hydrophobic residues and are believed to be structurally related to ER signal sequences (Bassham *et al.*, 1991). Upon transport into the thylakoid lumen, this secondary signal is also cleaved by a lumenal processing peptidase (Kirwin *et al.*, 1987; Keegstra, 1989). Transport into the thylakoid lumen *in vitro* is stimulated by the addition of ATP (Bauerle and Keegstra, 1991).

Very little is known about how proteins are directed to the chloroplast outer envelope, inner membrane or intermembrane space because of a lack of sequence information and of import studies into these compartments. Analogy to the mechanisms of mitochondrial protein sorting provides the framework for models around which these processes might be studied; however, stop transfer sequences do not appear to prevent transport of proteins into the stromal space (Lubben *et al.*, 1987). In this regard, however, these hydrophobic segments were inserted into the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase more than 100 amino acids from the amino terminus (Lubben *et al.*, 1987). One hypothesis suggests that these sequences must be

relatively close to the amino terminus to permit integration into the outer envelope membrane of chloroplasts (Singer and Yaffe, 1990). *In vitro* import of the 30 kDa protein is ATP-dependent, mediated by a presequence and does not require an electrochemical gradient (de Boer and Weisbeek, 1991). It is not known what information within proteins directs them to the envelope (de Boer and Weisbeek, 1991). A 6700 Da outer membrane protein has also been studied using *in vitro* import assays. It is integrated into the envelope, apparently without the aid of a proteinaceous receptor, and in an ATP-independent manner presumably by direct interaction with the membrane lipids (de Boer and Weisbeek, 1991).

1.7.4 Import into the nucleus

Macromolecules including protein and RNA are imported into and exported from the nucleus through the nuclear pore. The pore is a large proteinaceous complex that spans the double membrane of the nucleus. It is approximately 10^8 Da and 20-150 nm in diameter. The pore complex is composed of 8 globular subunits on both the inner and outer membrane giving the structure 8-fold rotational symmetry (reviewed in Newport and Forbes, 1987). Visualization of the structure in the plane of the membrane reveals an assymetric structure. Electron microscopy and 3-D reconstruction has shown that the two ring-like structures form a structure analagous to a 'lobster trap', with the narrow end oriented toward the nucleoplasm (Jarnik *et al.*, 1991). Filaments connect the outer and inner membrane complexes. Eight filaments also appear to extend toward

the centre of the pore and are visualized as spokes projecting inward to the centre of the pore. The resultant 9 nm (in *Xenopus* oocytes) aqueous channel often appears blocked by a 35 nm central plug which may be an integral component of the pore (Newport and Forbes, 1987). The complex allows the passive diffusion of proteins, dextrans or gold particles smaller than 40-60 kDa between the nucleus and cytoplasm. However, specific nuclear localization signals (NLSs) operate to direct proteins (both larger and smaller than 40-60 kDa) to the nucleus (reviewed in Roberts, 1989; Garcia-Bustos *et al.*, 1991; Silver, 1991).

The protypical NLS was identified in SV40 large T antigen. An internal, highly basic stretch of amino acids (Pro-Lys-Lys^{128}-Lys-Arg-Lys-Val) is necessary for nuclear localization (Kalderon *et al.*, 1984a) and is sufficient to direct non-nuclear proteins to the nucleus (Kalderon *et al.*, 1984b). Subsequently, several other NLSs have been identified (reviewed in Roberts, 1989; Garcia-Bustos *et al.*, 1991). Although all NLSs are rich in basic amino acids, there appears to be no primary structure conservation. Secondary structure predictions also fail to reveal a common structural motif. In contrast to localization signals for the ER, mitochondrion, and chloroplast, NLSs are not cleaved upon import of the protein into the organelle. This permits resorting of nuclear proteins after mitosis and the disassembly of the nuclear envelope. In addition, some nuclear proteins contain functionally redundant NLSs (Richardson, 1986; Roberts, 1989; Garcia-Bustos *et al.*, 1991; Silver, 1991). Additional signals provide an increased efficiency of import (Lanford *et al.*, 1986; Silver, 1991). Unlike

other organelles, proteins imported into the nucleus appear not to require unfolding for their import, because conjugation of NLSs to non-nuclear proteins (resulting in a branched structure; Lanford *et al.*, 1986) or to inert particles such as colloidal gold causes their accumulation in the nucleus (Silver, 1991). NLS-containing complexes as large as 25 nm can be accommodated by the gating response of the nuclear pore to NLSs (Silver, 1991).

Little is known about the mechanism of import. It is evident by electron microscopy that import into the organelle occurs at the nuclear pores. Import into the nucleus can be blocked by depletion of ATP, low temperature or by the binding of wheat germ agglutinin, which binds to *O*-linked *N*-acetylglucosamine residues common to several nuclear pore complex proteins (nucleoporins) (Silver, 1991). In such cases, import has been demonstrated to occur in two distinct temporal steps. First, rapid binding to the cytoplasmic side of the nuclear pore complex, followed by subsequent ATP-dependent import of the protein through the pore by a putative transporter complex (Newmeyer and Forbes, 1988; Richardson *et al.*, 1988). Accumulation of proteins at the nuclear pore has allowed visualization of proteins lining up apparently along fibres which extend from the nuclear pore into the cytoplasm (Richardson *et al.*, 1988; Silver, 1991).

Recognition and import of nuclear proteins is presumably mediated by NLS-binding proteins (NBP). *In vitro* binding studies have shown the existence of distinct sets of NBPs in the cytosol, at the nuclear pore complex, and in the nucleoplasm (Silver

et al., 1989; Adam and Gerace, 1991; reviewed in Silver, 1991). The identification of cytosolic NBPs suggest that proteins arrive at the pore by a specific mechanism, rather than by simple diffusion. NLS-receptors may therefore shuttle between the cytoplasm and the nuclear surface. Proteins may thus be transferred from the cytoplasm into the nucleoplasm by successive interaction with receptors during their transit into the nucleus. Protein import into nuclei is therefore envisioned to occur as follows: NLS-containing proteins are synthesized in the cytoplasm where they first interact with NBP and move to the nuclear pore complex. This interaction triggers the binding to additional NBPs (and nucleoporins), the hydrolysis of ATP and the gating of the nuclear pore. Proteins are translocated either after release from, or in association with, NBPs, which are subsequently recycled for interaction with another NLS-containing protein (Silver, 1991).

1.7.5 Targeting to peroxisomes

In general, peroxisomal proteins are synthesized at their mature size on free polysomes in the cytosol (see section 1.3; Table 1.3.1). Therefore proteins must be directed to peroxisomes post-translationally by virtue of signals within the amino acid sequences of the mature protein. At the beginning of this study there were no peroxisomal targeting signals identified. Recent evidence from S. Subramani's laboratory has pointed to a tripeptide (Ser-Lys-Leu) sequence at the carboxy-terminus of firefly (*Photinus pyralis*) luciferase which is necessary and sufficient for peroxisomal targeting in mammalian (monkey kidney CV-1) cells. Furthermore, mutational analysis

and comparison of this sequence to other peroxisomal proteins (Gould *et al.*, 1988) indicated the existence of a conserved consensus tripeptide peroxisomal targeting signal (PTS) having the generalized structure Ser/Ala/Cys - Lys/Arg/His - Leu (Gould *et al.*, 1989). However, many peroxisomal proteins do not contain this tripeptide at their carboxy-terminus. There is also no evidence that this tripeptide can act at internal locations. Also, some peroxisomal proteins, including rat peroxisomal thiolases, contain a cleavable amino-terminal presequence (see table 1.4.1), which has been hypothesized, by analogy to mitochondrial, chloroplast and ER targeting, to play a role in targeting to peroxisomes (see Borst, 1989).

2.0 MATERIALS AND METHODS

2.1 Materials

All materials were of the highest quality available and, where required, used according to the manufacturers' specifications.

2.1.1 Chemicals and reagents

agar	Difco Laboratories
agarose	Gibco/BRL Canada; Pharmacia (Canada) Inc.
albumin (bovine serum)	Sigma Chemical Company
ampicillin	Sigma Chemical Company
amino acid mix (minus methionine)	Promega Corp.
antipain	Sigma Chemical Company
β-hydroxybutyryl-Coenzyme A	Sigma Chemical Company
β-mercaptoethanol	BDH Chemicals
Bio-Rad protein assay dye reagent	Bio-Rad Laboratories (Canada) Ltd.
brij-35 (polyoxyethylene 23-lauryl ether)	Sigma Chemical Company
chymostatin	Sigma Chemical Company
Coomassie Brilliant Blue (R-250)	Gibco/BRL Canada
cytochrome-c (horse heart)	Sigma Chemical Company
dithiothreitol	Sigma Chemical Company
Freund's complete adjuvant	Sigma Chemical Company
Freund's incomplete adjuvant	Sigma Chemical Company
guanidine hydrochloride	Sigma Chemical Company
GTG agarose	FMC BioProducts
hemoglobin (bovine blood)	Sigma Chemical Company
hydrogen peroxide (30%)	Sigma Chemical Company

isopropyl β-D-thiogalactoside (IPTG)	Gibco/BRL Canada
leupeptin	Sigma Chemical Company
L-leucine	Sigma Chemical Company
L-methionine	Sigma Chemical Company
L-histidine	Sigma Chemical Company
malt extract	Difco Laboratories
nitrocellulose (pore size-0.45μm)	Schleicher and Schuell Inc.
Nonidet P-40 (NP-40)	Sigma Chemical Company
Nycodenz	Nycomed As Pharma
oleic acid	Fisher Scientific
oligo$(dT)_{12-18}$	Collaborative Research Inc.
oligo(dT) cellulose, Type 3	Collaborative Research Inc.
ovalbumin	Miles Laboratory (Pty) Ltd.
PANSORBIN *Staphylococcus aureus* cells (inactivated)	Calbiochem Corp.
pepstatin A	Sigma Chemical Company
peptone	Difco Laboratories
phosphorothioate dCTP	Amersham Canada Ltd.
Sephadex G-50 (medium)	Pharmacia (Canada) Inc.
sodium dithionite (sodium hydrosulfite)	Sigma Chemical Company
titanium oxysulfate hydrate	Aldrich Chemical Company Inc.
Triton X-100	Sigma Chemical Company
Tween 20 (polyoxyethylene sorbitan monolaurate)	Sigma Chemical Company
X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside)	Gibco/BRL Canada
X-ray film	Eastman Kodak Company
yeast extract	Difco Laboratories
yeast nitrogen base (without amino acids)	Difco Laboratories

2.1.2 Enzymes

β-glucuronidase (from Helix pomatia)	Sigma Chemical Company
chloramphenicol acetyl transferase	Pharmacia (Canada) Inc.
proteinase K	Sigma Chemical Company

(from *Tritirachium album*) ribonuclease A (RNase I "A"; bovine pancreas)	Pharmacia (Canada) Inc.
Zymolyase 100T	Seikagaku Kogyo; ICN Biomedicals Inc.

2.1.2.1 DNA modifying enzymes

DNA ligase (T4)	Gibco/BRL Canada; New England Biolabs; Promega Corp.
DNA polymerase I (*E. coli*)	New England Biolabs, Inc.; Pharmacia (Canada) Inc.; Gibco/BRL Canada
DNA polymerase I Klenow fragment (*E. coli*)	New England Biolabs, Inc. Pharmacia (Canada) Inc.; Gibco/BRL Canada
DNA polymerase, modified T7 (Sequenase)	United States Biochemical Corp.
DNase I (bovine pancreas)	Sigma Chemical Company
exonuclease III (*E. coli*)	Amersham Canada Ltd.; Gibco/BRL Canada
polynucleotide kinase (T4)	Pharmacia (Canada) Inc.
restriction endonucleases	Gibco/BRL Canada; New England Biolabs; Pharmacia (Canada) Inc.; Boehringer Mannheim
AMV reverse transcriptase	Pharmacia (Canada) Inc.; Gibco/BRL Canada
mung bean nuclease	Pharmacia (Canada) Inc.
S1 nuclease	Pharmacia (Canada) Inc.
*Eco*R1 methylase	Pharmacia (Canada) Inc.; Gibco/BRL Canada

2.1.3 Multi-component systems

enhanced chemiluminescence detection kit	Amersham Canada Ltd
gene clean kit	Bio 101
phosphorothioate mutagenesis kit	Amersham Canada Ltd.
rabbit reticulocyte lysate (nuclease treated)	Promega Corp.
random primers labeling system	Gibco/BRL Canada

2.1.4 Molecular weight standards

(i) 1 kb DNA ladder (75-12,216 bp) — Gibco/BRL Canada

(ii) Dalton Mark VII-L molecular weight markers for SDS-PAGE (bovine α-lactalbumin, 14,200; soybean trypsin inhibitor, 20,100; bovine trypsinogen, 24,000; bovine carbonic anhydrase, 29,000; rabbit glyceraldehyde dehydrogenase, 36,000; ovalbumin, 45,000; bovine albumin, 66,000 Da) — Sigma Chemical Company

(iii) phosphorylase b (rabbit muscle; 97,400 Da) — Sigma Chemical Company

(iv) β-galactosidase (*E. coli*; 116,000 Da) — Sigma Chemical Company

(v) prestained molecular weight markers for SDS-PAGE (egg lysozyme, 16,000; soybean trypsin inhibitor, 24,000; bovine carbonic anhydrase, 33,000; ovalbumin, 47,000; bovine albumin, 84,000; rabbit phosphorylase b, 110,000 Da) — Bio-Rad Laboratories (Canada) Ltd.

2.1.5 Radiochemicals

^{125}I-protein A (>30 mCi·mg^{-1} total protein A, 0.1 Ci·μl^{-1}) — Amersham Canada Ltd.

[γ-^{32}P]dATP (3,000 Ci·mmol^{-1}, 10 μCi·μl^{-1}) — Amersham; Dupont/NEN Canada Inc.

[α-^{32}P]dATP (3,000 Ci·mmol^{-1}, 10 μCi·μl^{-1}) — Amersham; ICN Biomedicals Inc.

L-[^{35}S]methionine (1150 Ci·mmol^{-1}, 10 μCi·μl^{-1}) — Dupont/NEN Canada Inc.

2.1.6 Plasmids and bacteriophage λ vectors

The bacteriophage vectors λgt11 (Young and Davis, 1983) and λEMBL3 (Frischauf *et al.*, 1983) were used to clone cDNAs and the *HDE* gene, respectively. The plasmids pGEM5Zf(+), pGEM7ZF(+) (Promega Corp.) and pUC118 (Vieira and Messing, 1987) were used for subcloning and were gifts from Dr. A.B. Futcher, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. The plasmids YCp50 (Kuo and Campbell, 1983), YEp13 (Broach, *et al.*, 1979) and pMK22 (Kurtz *et al.*, 1987) were used to express genes in yeast (see sections 3.4, 3.5, 3.6). YCp50 and YEp13 were gifts from Dr. A.B. Futcher. pMK22 was a gift from Dr. M.B. Kurtz, Squibb Laboratories, Rahway, NJ. Plasmid pCM4 is a recombinant of pBR327 and the *E.coli CAT* gene (Close and Rodriguez, 1982). pCM4 was purchased from Pharmacia (Canada) Inc.

2.1.7 Oligodeoxyribonucleotides

The M13 forward (5'-GTAAAACGACGACGGCCAGT-3') and reverse (5'-CAGGAAACAGCTATGAC-3') universal primers were purchased from United States Biochemical Corp. or Pharmacia (Canada) Inc. Custom oligodeoxyribonucleotides were synthesized using an Applied Biosystems 381A automated DNA synthesizer by the Central Facility of the Institute of Molecular Biology, McMaster University, Hamilton, ON. The specific uses of particular oligodeoxyribonucleotides are shown in Table 2.1.7.1.

Table 2.1.7.1 Oligodeoxyribonucleotides

Number	Sequence	Purpose
AB991	5'-CGGTGACAAAGGTAAGATCTAAT-3'	*In vitro* mutagenesis to generate *HDE* mutant *GKI*
AB993	5'-CAAAGCAAAGTAATGAAGTT-3'	*In vitro* mutagenesis to generate *HDE* mutant *AKX*
AB992	5'-CGGTGACAAAAAGATCTAAT-3'	*In vitro* mutagenesis to generate *HDE* mutant *XKI*
AB994	5'-TGACAAAGCACAAATCTAATGAA-3'	*In vitro* mutagenesis to generate *HDE* mutant *AQI*
AB864	5'-GTCGGTGACAAATAATGAAGTTGT-3'	*In vitro* mutagenesis to generate *HDE* mutant *ΔAKI*
AB865	5'-ATGGAAGGAAAGCGATG-3'	Sequencing of *HDE* mutations *GKI, AKX, XKI, AQI* and *ΔAKI*
AB1062	5'-GAAAGAAAGAAAGATCTATAACAATCA-3'	*In vitro* mutagenesis of *HDE* 5' region to generate *Bgl*II site
AB1069	5'-GAGTAGTATTTACCCAAACC-3'	Sequencing *HDE* 5' region carrying *Bgl*II site
AB995	5'-GGGCGGGGCGAAGATCTAATTTTTTTA-3'	*In vitro* mutagenesis to generate *E. coli CAT* mutant *CATAKI*
AB366	5'-CACCAAGCAATTGAAG-3'	Sequencing of HDE cDNA
AB367	5'-ATACCAAACTTGTCAA-3'	Sequencing of HDE cDNA

2.1.8 Antisera

Donkey anti-rabbit IgG linked to horseradish peroxidase was purchased from Amersham Canada Ltd. Rabbit anti-serum to *Escherichia coli* chloramphenicol acetyl transferase was a gift from Dr. J. Capone, Department of Biochemistry, McMaster University. Anti-SKL serum was a gift from Dr. S. Subramani, Department of Biology, University of California at San Diego, La Jolla, CA and was generated by immunizing rabbits with a chemically synthesized peptide (NH_2-CRYHLKPLQSKL-COOH) the last 9 amino acids of which correspond to the carboxy-terminal 9 amino acids AOX (Gould

et al., 1990a). Goat anti-mouse IgG conjugated to alkaline phosphatase was a gift from Dr. D.W. Andrews, Department of Biochemistry, McMaster University. Goat anti-guinea pig IgG (H and L) conjugated to alkaline phosphatase was purchased from Jackson Immunoresearch Laboratories, West Grove, PA.

Rabbit antisera to *C. tropicalis* HDE, fatty acyl-CoA oxidase, and catalase and rabbit, mouse and guinea pig antisera to the carboxy-terminal 12 amino acids of HDE conjugated to keyhole limpet hemocyanin (KLH) (anti-AKI sera) were prepared as outlined in sections 2.4.3 and 2.4.4.

2.1.9 Peptides

The synthetic peptides AKI (Acetyl-CAIKLVGDKAKI-COOH), ΔAKI (Acetyl-CAIKLVGDK-COOH) and UR (Acetyl-CSKRKEAE-amide) (>80% purity) were purchased from Multiple Peptide Systems, San Diego, CA. The AKI peptide, conjugated to keyhole limpet hemocyanin (KLH) through its amino-terminal cysteine by *m*-maleimidobenzoyl-*N*-hydroxysuccimide ester (Kitagawa and Aikawa, 1976), was also purchased from Multiple Peptide Systems.

2.1.10 Commonly used buffered solutions

The table below describes the composition of some commonly used buffers.

Table 2.1.10.1 Common buffers

5X KGB	1 M K-glutamate, 250 mM Tris-acetate (pH 7.6), 50 mM Mg-acetate, 250 $\mu g \cdot ml^{-1}$ bovine serum albumin, 2.5 mM 2-mercaptoethanol	Hanish and McClelland, 1988
20X SSC	3 M NaCl, 0.3 M trisodium citrate (pH 7.0)	Maniatis *et al.*, 1982
10X P	0.1 M Na_2HPO_4, 0.01 M $Na_2H_2P_2O_7$	Harley, 1987
Tris-saline	0.9% NaCl, 10 mM Tris-HCl (pH 7.4)	Burnette, 1981
Tris-saline -NP-40	Tris-saline plus 0.5% (w/v) NP-40	Burnette, 1981
TBST	20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% (w/v) Tween 20 (polyoxyethylene-sorbitan monolaurate)	Huynh *et al.*, 1985
Denhardt's Soln. (50X)	1% Ficoll, 1% polyvinylpyrrolidone, 1% bovine serum albumin.	Maniatis *et al.*, 1982
TE	10 mM Tris-HCl (pH 7.0-8.0), 1 mM EDTA (pH 8.0)	Maniatis *et al.*, 1982
TBE (10X)	0.89 M Tris-borate, 0.89 M boric acid, 0.02 M EDTA	Maniatis *et al.*, 1982

2.2 Culture conditions

2.2.1 Bacterial and bacteriophage strains and culture conditions

E. coli DH5α was purchased from Gibco/BRL and grown in Luria Broth [LB; 1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7.5 (plus 100 $\mu g \cdot ml^{-1}$ ampicillin or 12.5 $\mu g \cdot ml^{-1}$ tetracycline, as required); Maniatis *et al.*, 1982]. *E. coli* DH5αF' was purchased from Gibco/BRL and grown in TYPD [0.16% tryptone, 0.16% yeast extract, 0.5% NaCl, 0.25% potassium phosphate pH 7.5, 0.2% glucose (plus 150 $\mu g \cdot ml^{-1}$ ampicillin, as required)]. *E. coli* cells were transformed according to the procedures

outlined by the supplier. *E. coli* Y1090 and NM539 cells were grown to saturation in LB plus 0.2% maltose for infection with bacteriophage λ (Young and Davis, 1983). M13KO7 helper phage was purchased from Gibco/BRL. λgt11 and λEMBL3 bacteriophages were purchased from Promega Corp.

2.2.2 Yeast strains and culture conditions

The yeast strains used were *Candida tropicalis*, Berkhout strain pK233 (ATCC 20336; Tanabe *et al.*, 1966); *Candida albicans* strain SGY 243, which has a homozygous *ura3* deletion (Kelly *et al.*, 1987); and *Saccharomyces cerevisiae* strain DL-1 (*mat α, leu2, ura3, his3*). *C. tropicalis* was obtained from the American Type Culture Collection. *C. albicans* was a generous gift from Dr. M.B. Kurtz and *S. cerevisiae* was a gift from Dr. S. Subramani.

All yeasts were grown at 30°C in a rotary shaker. Media varied with the yeast strain and the purpose of growth. Table 2.2.2.1 describes the media used.

Table 2.2.2.1 Yeast growth media

Medium	Composition	Purpose	References
ME agar	3% malt extract, 1.5% agar	maintenance of C. tropicalis	Kamiryo et al., 1982
ME	3% malt extract	starter culture for C. tropicalis	Kamiryo et al., 1982
YPBO	0.3% yeast extract, 0.5% peptone, 0.5% K_2HPO_4, 0.5% KH_2PO_4, 1.0% Brij 35, 1.0% (w/v) oleic acid	oleic acid-containing medium for induction of peroxisomes and RNA encoding β-oxidation enzymes in C. tropicalis and C. albicans	Kamiryo et al., 1982 Fujiki et al., 1986 Aitchison and Rachubinski, 1990
YEPD	2% peptone, 1% yeast extract, 2% glucose	rich, glucose-containing medium for growth of C. tropicalis, C. albicans and S. cerevisiae, no plasmid selection	Sherman et al., 1986
YEPD agar	2% peptone, 1% yeast extract, 2% glucose, 2% agar	maintenance of untransformed C. tropicalis, C. albicans and S. cerevisiae	Sherman et al., 1986
minimal medium	0.7% yeast nitrogen base, 2% glucose	selective medium for growth of C. albicans transformants	Sherman et al., 1986 Kurtz et al., 1986
minimal medium agar	0.7% yeast nitrogen base, 2% glucose, 2% agar	selection and maintenance of C. albicans transformants	Sherman et al., 1986 Kurtz et al., 1986
C. alb rescue	0.7% yeast nitrogen base, 2% glucose, 1 M sorbitol, 2% agar	plating of C. albicans after protoplast transformation	Kurtz et al., 1986
minimal medium +his +leu	0.7% yeast nitrogen base, 2% glucose, 50 $\mu g \cdot ml^{-1}$ histidine, 50 $\mu g \cdot ml^{-1}$ leucine	selective medium for growth of S. cerevisiae transformants	Sherman et al., 1986
minimal medium +his +leu agar	0.7% yeast nitrogen base, 2% glucose, 50 $\mu g \cdot ml^{-1}$ histidine, 50 $\mu g \cdot ml^{-1}$ leucine, 2% agar	selection and maintenance of S. cerevisiae transformants	Sherman et al., 1986
SCIM	0.7% yeast nitrogen base 0.5% yeast extract, 0.5% peptone, 0.5% (w/v) Tween 40, 0.1% glucose, 1.0% (w/v) oleic acid, 50 $\mu g \cdot ml^{-1}$ histidine, 50 $\mu g \cdot ml^{-1}$ leucine	induction of peroxisomes in S. cerevisiae transformants	Erdmann et al., 1989
SCIM +ura	0.7% yeast nitrogen base 0.5% yeast extract, 0.5% peptone, 0.5% (w/v) Tween 40, 0.1% glucose, 1.0% (w/v) oleic acid, 50 $\mu g \cdot ml^{-1}$ histidine, 50 $\mu g \cdot ml^{-1}$ leucine, 50 $\mu g \cdot ml^{-1}$ uracil	induction of peroxisomes in non-transformed S. cerevisiae	Erdmann et al., 1989

2.3 Analyses of proteins

2.3.1 Protein determination

Total protein was determined by the method of Bradford (1976), as modified by Bio-Rad. Protein samples were diluted to 100 μl with 0.1 M sodium-phosphate buffer (7.0), and 1 ml of Bio-Rad protein assay reagent was added. The OD_{595} of each sample was determined. Ovalbumin (0, 2.5, 5, 10, 15, 20 μg) was used as a standard.

2.3.2 Separation of polypeptides by SDS-PAGE

Polypeptides were separated by SDS-PAGE essentially as described by Laemmli (1970), following modification by Fujiki *et al.* (1986). The separating gel consisted of 10% acrylamide (30:0.8 acrylamide:N,N'-methylene-bis-acryalmide), 0.37 M Tris-HCl (pH 8.8), 0.1% SDS, 0.5% (v/v) TEMED (N,N,N',N'-tetramethylethylene-diamine), 0.2% ammonium persulfate), and the stacking gel consisted of 3% acrylamide (30:0.8 acrylamide:N,N'-methylene-bis-acryalmide), 0.06 M Tris-HCl (pH 6.8), 0.1% SDS, 0.1% (v/v) TEMED, 0.1% ammonium persulfate). The gels were generally cast using the Hoefer vertical slab gel unit (model SE 400, Hoefer Scientific Instruments; separating gel dimensions: 12 cm x 14 cm x 0.75 mm). Proteins were denatured by boiling in SDS-PAGE sample buffer (0.0625 M Tris-HCl (pH 6.8), 2.0% SDS, 10% sucrose, 0.01 M DTT, 0.001% bromophenol blue) for 5 min before loading onto the gels. Gels were generally run at 50 V in SDS-PAGE running buffer (0.4 M glycine, 50 mM Tris-HCl (pH 8.8), 0.1% SDS) until the bromophenol blue reached the bottom of the gel (approximately 12 h).

2.3.3 Staining of SDS-polyacrylamide gels

SDS-polyacrylamide gels were stained for at least 1 h by immersion in 0.1% Coomassie Brilliant Blue (R-250) in 10% (v/v) acetic acid and 35% (v/v) methanol. Gels were destained by washing extensively in 10% (v/v) acetic acid and 35% (v/v) methanol. Gels were then processed further or dried on a Bio-Rad (model 483) gel dryer at 80°C.

2.3.4 Fluorography

SDS-polyacrylamide gels containing ^{35}S-labelled polypeptides were prepared for exposure to film by fluorography (Bonner and Laskey, 1974). Gels were stained and destained in the normal fashion, then dehydrated by washing in DMSO (dimethylsulfoxide) for 1 h (2 changes). The gels were then soaked in DMSO-PPO (DMSO containing 22% 2,5-diphenyloxazol) for 3 h with gentle agitation. Excess DMSO-PPO was removed, and the gels were soaked in water for 15 min. The gels were dried at 60°C (Bio-Rad model 483) and exposed to Kodak (Rochester, NY) X-AR5 film.

2.3.5 Electroelution of proteins

Polypeptides were separated by SDS-PAGE using a large Hoefer vertical slab gel unit (28 cm x 12 cm x 1.5 mm). Approximately 1 mg of total peroxisomal protein was loaded into the preparative well of the gel (11.5 cm wide). After electrophoresis the polypeptides were stained, and the bands of interest were excised with a razor blade. Each band was then placed into a serological pipette (plugged at the narrow end with

siliconized glass wool) and inserted into an isoelectric focussing (IEF) chamber (Bio-Rad, Tube-Cell model 155). Proteins were electroeluted into sealed dialysis tubing in SDS-PAGE running buffer for 2 days at 100 V (2.5-5 mA). The dialysis tubing was removed, and the protein was dialyzed against 50 mM ammonium bicarbonate for 2 days (2 changes) at room temperature. The protein in solution was frozen, lyophilized, and dissolved in H_2O.

2.4 Antibody production

2.4.1 Isolation of catalase, fatty acyl-CoA oxidase, and hydratase-dehydrogenase-epimerase

C. tropicalis peroxisomal fatty acyl-CoA oxidase (AOX), catalase (CAT) and hydratase-dehyhydrogenase-epimerase (HDE) were isolated by SDS-PAGE of purified peroxisomes followed by electroelution of the protein band of interest.

One mg of total *C. tropicalis* peroxisomal protein was run on a preparative SDS 7-15% polyacrylamide gel (separating gel dimensions: 14 cm x 28 cm x 1.5 mm), containing two wells (one of 1 cm width for the molecular weight standards, and one of 11.5 cm width for the sample). Electrophoresis was performed as described in section 2.3.2. After staining, the bands of interest were identified by comparison of the gel to Fig. 5 and 7 of Fujiki *et al.* (1986) and excised from the gel. The proteins were electroeluted into sealed dialysis tubing (section 2.3.5). The lyophilized samples of protein were each dissolved in 100 μl of H_20.

2.4.2 Popliteal lymph node injection in rabbits

Antiserum directed against *C. tropicalis* CAT, AOX or HDE was generated by direct injection of the antigen into the popliteal lymph nodes of rabbits. Anesthetized rabbits were injected with 0.2 ml of 2.5% Evan's blue dye into a fold of skin between the toes of the hind legs. After approximately 0.5 h, the lymph nodes of both hind legs were surgically exposed and identified by accumulation of the blue dye. Approximately 100 μg of purified protein was emulsified with an equal volume of Freund's complete adjuvant and injected directly into the popliteal lymph nodes. Rabbits were boosted 5 weeks post-operatively by subcutaneous injection at multiple sites with approximately 100 μg of antigen emulsified with an equal volume of Freund's incomplete adjuvant. Rabbits were bled 2 weeks after boosting, and the sera were collected by allowing the whole blood to coagulate at room temperature for 2 h, followed by a 2-4 h incubation at 4°C. Clots were then separated from sera by centrifugation at 2750 rpm in an IEC clinical centrifuge. Rabbits were subsequently boosted and bled to maintain a high titre of antibody.

2.4.3 Anti-AKI antibodies generated in guinea pigs

Antiserum directed against the carboxy-terminal 12 amino acids of HDE (containing the AKI tripeptide) was generated by subcutaneous injection of the antigen at multiple sites. 200 μg of the peptide (conjugated to keyhole lympet hemocyanin) was emulsified with an equal volume of Freund's complete adjuvant and injected. The animals were boosted 5 weeks after the initial injection with 200 μg of the peptide

emulsified with an equal volume of Freund's incomplete adjuvant. The guinea pigs were sacrificed 7 days later, and the sera were collected as in section 2.4.2.

2.4.4 Anti-AKI antibodies generated in rabbits

Antiserum directed against the carboxy-terminal 12 amino acids of HDE (containing the AKI tripeptide) was generated by subcutaneous injection of the antigen at multiple sites. 500 μg of the peptide (conjugated to KLH) was mixed with an equal volume of Freund's complete adjuvant and injected. The rabbits were boosted 6 weeks after the initial injection with 500 μg of the peptide emulsified with an equal volume of Freund's incomplete adjuvant. At week 8 after the initial injection, the rabbits were again boosted and then bled. After determination of the specificity of the antisera, the rabbits were sacrificed, and the sera were collected as in section 2.4.2.

Antisera were checked for their ability to bind antigen using the techniques of western blotting (Burnette, 1981; section 2.5) and immunoprecipitation (Fujiki *et al.*, 1984; section 2.21).

2.5 Western blotting

2.5.1 Protein A detection

Proteins separated by SDS-PAGE were transferred electrophoretically to nitrocellulose using a Bio-Rad transblot chamber filled with western transfer buffer (20 mM Tris, 150 mM glycine, 20% methanol), generally at 200 mA for 16 h. The nitrocellulose was blocked with 1% bovine hemoglobin (Sigma Chemical Co.) in Tris-

saline (Hb-TS) for 30 min at 37°C with gentle agitation. The blots were then sealed in a heat-sealable pouch (Kapak Corp.) with antiserum diluted in Hb-TS at the following concentrations:

Antiserum		Concentration^{-1}
source	target	
rabbit	CAT	500
rabbit	AOX	1000
rabbit	HDE	1000
rabbit	SKL	500
guinea pig	AKI	500
rabbit	AKI	500
rabbit	*E. coli* CAT	500

The blots were incubated with the diluted antiserum at room temperature with rocking for 90 min. The blots were then washed in Tris-saline for 10 min, twice in Tris-saline-NP-40 for 20 min each, and again in Tris-saline for 10 min, at room temperature with agitation. The blots were then sealed in a pouch containing Hb-TS plus 0.5 μCi [^{125}I]protein A · 10 ml^{-1} and incubated as above for 30 min. The blots were washed as above, air-dried briefly, and exposed to Kodak (Rochester, NY) X-AR5 film with one intensifying screen (DuPont) at -70°C.

2.5.2 Detection by enhanced chemiluminescence

Proteins were transferred to nitrocellulose as described in section 2.5.1, and the immunoreaction of proteins was detected by enhanced chemiluminescence essentially as suggested by the manufacturer (Amersham). The nitrocellulose was blocked with

milk-TBST (5% skim milk powder in TBST) for 1 h and then incubated with the primary antibody diluted in milk-TBST for 1 h. The blots were washed in TBST (once for 15 min, then twice for 5 min each). Blots were then incubated with donkey antirabbit IgG linked to horseradish peroxidase at a 1 in 25,000 dilution in milk-TBST for 0.5 h. The blots were washed in TBST (once for 15 min, then four times for 5 min each). The detection reagents were mixed (1:1) and poured onto the blots so as to cover each blot. After 1 min, excess reagent was drained and the blots were exposed to Kodak X-AR5 film for 30 s to 1 h.

2.5.3 Peptide competition assay

For competition assays, standard western blots were performed after pre-incubation of the antibodies with the indicated peptides at a concentration of 25 μM for 90 min at room temperature.

2.6 Preparation of yeast lysates

Yeast were grown to late-logarithmic phase (generally in 10 ml). The cells were harvested by centrifugation in an IEC clinical centrifuge and washed with water and then breakage buffer (50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.1 mM EDTA, 0.1 mM $ZnCl_2$, 1 mM PMSF). The pellet was resuspended in an equal volume of breakage buffer and transferred to a microfuge tube. Glass beads (Sigma, 425-600 μm diameter) were added to the level of the meniscus. The yeast cells were disrupted by vortexing for 1 min at 4°C. The cell lysate was centrifuged with a pulse of the microfuge to pellet cell

debris and glass beads. The layer of lipid was removed, and the supernatant containing protein was transferred to a fresh tube.

2.7 Peroxisome isolations

2.7.1 Peroxisome isolation from *Candida albicans* and *Candida tropicalis*

Peroxisomes were isolated from *C. albicans* and *C. tropicalis* as described (Kamiryo *et al.*, 1982; Aitchison and Rachubinski, 1990). Cells were grown to late log phase in either ME or YEPD and pelleted by centrifugation in an IEC clinical centrifuge at 2750 rpm for 5 min. The cells were washed with sterile H_2O, and approximately 0.5 ml of pelleted cells were resuspended in water and used to inoculate 1 l of YPBO. The cells were grown for 18-20 h and then collected by centrifugation (Sorvall GSA rotor, 5000 rpm (4000 x g_{max}) for 10 min). The cells were washed with water 3 times and converted to spheroplasts by treatment with 0.25 mg of Zymolyase 100T per g of cells per ml of Zymolyase digestion buffer (0.5 M KCl, 5 mM MOPS (pH 7.2), 10 mM Na_2SO_3) at 30°C with gentle rotation. Spheroplasts were collected by centrifugation at 4°C in a Sorvall HB-4 rotor at 3750 rpm (2300 x g_{max}) for 8 min. Collected spheroplasts were resuspended in 1-3 ml of F buffer (5% Ficoll 400, 0.6 M sorbitol, 2.5 mM MOPS (pH 7.2), 0.5 mM EDTA, 0.1% (v/v) ethanol) per g of original cells. The spheroplasts were mechanically homogenized in a grinding vessel (Potter-Elvejhem) with a tight fitting pestle by 20 strokes at 1000 rpm using a Cole-Palmer lab stirrer (model 4376-00). The homogenate was centrifuged for 10 min at 2500 rpm (1000 x g_{max}) at 4°C in the HB-4 rotor to remove cell debris and nuclei. The postnuclear supernatant was centrifuged at

11,250 rpm (20,000 x g_{max}) in the HB-4 rotor for 20 min at 4°C to yield a 20,000 x g_{max} pellet (20 kgP, enriched in peroxisomes). A portion of the 20 kgP was further fractionated to yield a purified peroxisomal fraction either on a sucrose step gradient (4.67 ml of 25%, 7 ml of 35%, 14 ml of 42%, 7 ml of 53% (w/w) sucrose in 2.5 mM MOPS (pH 7.2), 1 mM EDTA, 0.1% (v/v) ethanol) centrifuged in a Beckman VTi50 rotor at 32,000 rpm (100,000 x g_{max}) for 1 h at 4°C or on a Nycodenz step gradient (4.67 ml of 17%, 7 ml of 28%, 14 ml of 35%, 7 ml of 50% (w/v) Nycodenz in 5 mM Tris-HCl (pH 7.5), 3 mM KCl, 0.3 mM EDTA, 0.1% (v/v) ethanol) in a Beckman VTi50 at 35,000 rpm (130,000 x g_{max}) for 1 h at 4°C.

2.7.2 Peroxisome isolation from *Saccharomyces cerevisiae*

Peroxisomes were isolated from *S. cerevisiae* essentially as described for *C. albicans* and *C. tropicalis*, except for the following changes. Cells were grown overnight in YEPD or minimal medium +his +leu before transfer to 1 l of SCIM or of SCIM +ura for induction of peroxisomes. The spheroplasts were homogenized in SC disruption buffer (5 mM MES pH 6.0, 0.5 mM EDTA, 0.6 M sorbitol, 0.1% (v/v) ethanol; Lewin *et al.*, 1990) instead of F buffer. The 20 kgP was fractionated on a Nycodenz step gradient (1.1 ml of 17%, 2.5 ml of 25%, 0.7 ml of 35%, 0.5 ml of 50% (w/v) Nycodenz in SC disruption buffer) centrifuged in a Beckman VTi65 rotor at 35,000 rpm (135,000 x g_{max}) for 1 h at 4°C.

2.7.3 Peroxisome isolation from *Yarrowia lipolytica*

Peroxisomes were isolated from *Y. lipolytica* essentially as described for *C. albicans* and *C. tropicalis* except that the spheroplasts were homogenized in YL disruption buffer (5 mM MES (pH 5.5), 1 M sorbitol) and the 20 kgP was fractionated on a sucrose step gradient in YL disruption buffer.

2.7.4 Protease protection

Peroxisomes from *C. albicans* were recovered from the sucrose gradient and diluted with 42% (w/w) sucrose in 2.5 mM MOPS (pH 7.2), 1 mM EDTA, 0.1% (v/v) ethanol to a final protein concentration of 2.5 $mg \cdot ml^{-1}$. Peroxisomes were treated with trypsin or thermolysin for 30 min at 0°C in the presence or absence of 1% (w/v) Triton X-100 and 1% deoxycholate. Peroxisomes from *S. cerevisiae* were isolated on a Nycodenz gradient and used directly in the protease protection assay. Peroxisomes were treated with trypsin for 20 min at 0°C in the presence or absence of 1% (w/v) Triton X-100 and 1% deoxycholate. Samples were analyzed by SDS-PAGE and western blotting.

2.8 Marker enzyme analyses

2.8.1 Catalase assay

Catalase was used as a marker enzyme for peroxisomes. Catalase was assayed by the method of Leighton *et al.* (1968). 50 µl of sample was added to 50 µl of 2% (w/v) Triton X-100 and incubated on ice for at least 2 min. 1 ml of substrate (20

mM imidazole buffer (pH 7.0), 1 $mg \cdot ml^{-1}$ BSA, 0.01% (v/v) H_2O_2) was added and incubated for 5-15 min. at 0°C. The reaction was stopped by the addition of a saturated solution of $TiOSO_4$ (titanium (IV) oxysulfate in 2 N H_2SO_4), yielding a yellow complex (peroxy-titanium sulfate) indicative of residual H_20_2. The OD_{410} was measured.

2.8.2 Cytochrome *c* oxidase assay

Cytochrome *c* oxidase activity was monitored as the marker enzyme for mitochondria by the procedure of Cooperstein and Lazarow (1951). 2.5 ml of a solution of cytochrome *c* (0.38 $mg \cdot ml^{-1}$ cytochrome *c* dissolved in 0.3 M ammonium acetate, pH 7.4) and 250 μl of 10% (w/v) Triton X-100 were added to a 3 ml plastic spectrophotometer cuvette. The cytochrome *c* was reduced by the addition of a spatula tip of sodium hydrosulfite. This reduction was monitored as an increase in OD_{550}. The cuvette was mixed vigorously to oxidize slightly the cytochrome *c*. The sample was added, and the rate of cytochrome *c* oxidation was detected as a decrease in OD_{550}.

2.8.3 Beta-hydroxyacyl-Coenzyme A dehydrogenase assay

ß-hydroxyacyl-Coenzyme A dehydrogenase activity was assayed according to the procedure of Osumi and Hashimoto (1979) using ß-hydroxybutyryl-CoA as a substrate. 1-5 μl of sample was added to a 1 ml cuvette containing 0.1 M Tris-HCl (pH 10.2), 0.1 M KCl, 0.1% (w/v) Triton X-100, 1 mM NaN_3, 0.1 M NAD^+, 10 μM β-hybroxybutyryl-CoA. The production of NADH was monitored as the increase in OD_{340}

with time. The extinction coefficient E^{1M}_{1cm} = 6220 at 340 nm (Clarke and Switzer, 1977) was used to calculate the rate of NADH production.

2.9 Complementary DNA library construction and screening

2.9.1 Complementary DNA library construction

cDNAs were isolated from a λgt11 library containing inserts made from mRNAs of oleic acid-induced *C. tropicalis*. Cloning of cDNAs into the *Eco*RI site of the bacteriophage vector creates a gene fusion between the *E. coli lacZ* gene encoding β-galactosidase and the cDNA. This generates a fusion protein in the correct reading frame and in the correct orientation once in every six fusions. cDNAs of interest can thus be detected using antibody probes directed against a protein of interest and a secondary antibody conjugated to alkaline phosphatase (Young and Davis, 1983; Huynh *et al.*, 1985).

The λgt11 library was constructed by William Nuttley essentially following the method of Wickens *et al.* (1978). Poly(A)$^+$ RNA isolated from *C. tropicalis* grown in oleic acid-containing medium was reverse transcribed into single-stranded cDNA using AMV reverse transcriptase and oligo-$dT_{12\text{-}18}$ as a primer. After boiling for 2 min and quick-cooling on ice-water to denature the RNA-DNA duplex, the second DNA strand was synthesized using *E. coli* DNA polymerase I, which employs a 5' hairpin as a primer. The hairpin loop was digested with mung bean nuclease, and the DNA ends were made blunt with the large fragment of *E. coli* DNA polymerase (Klenow). The

cDNA was treated with *Eco*RI methylase to protect internal *Eco*RI restriction sites, and *Eco*RI linkers (5'-CCGGAATTCCGG-3') were ligated to the blunt ends. The ligation mix was treated with *Eco*RI to generate cohesive ends, and the cDNA was separated from liberated linkers by size fractionation using Sepharose CL-4B. The cDNA was then ligated into dephosphorylated λgt11 arms (Promega Biotech; Young and Davis, 1983). The ligation was packaged *in vitro* using the cell-free packaging extract Gigapack-Gold from Stratagene (Hohn and Murray, 1977).

2.9.2 Complementary DNA library screening

Packaged (and amplified) bacteriophage particles were mixed at an appropriate dilution with 100 μl of phage dilution buffer (10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$) and 200 μl of *E. coli* Y1090 cells previously grown to stationary phase in LB plus 0.2% maltose. The phage were incubated with cells for 30 min at room temperature, mixed with LB-soft top agar at 50°C, and poured onto LB agar plates pre-warmed to 42°C. The plates were incubated at 42°C for 3.5 h, overlaid with nitrocellulose filter disks presaturated with 10 mM IPTG, and incubated at 37°C for an additional 3.5 h.

The orientations of the filters were marked with a needle. Upon removal, the filters were treated with shaking at room temperature as follows:

1. 15 min wash in TBST to remove excess agar.
2. 15 min incubation with 7.5 ml TBST containing 1% BSA.

3. 30 min incubation with 7.5 ml TBST plus primary antibody (preadsorbed with *E. coli* extract for 30 min).
4. 10 min wash with 7.5 ml TBST (3 times).
5. 30 min incubation with secondary antibody diluted 1 : 7500.
6. 10 min wash with 7.5 ml TBST (3 times).

Positive plaques were detected by a colourimetric assay using alkaline phosphatase and the substrates nitroblue tetrazolium (NBT, 0.33 mg/ml) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP, 0.165 mg/ml) in alkaline buffer (100 mM Tris-HCl (pH 9.5), 100 mM NaCl, 5 mM $MgCl_2$).

2.9.3 Isolation of the HDE cDNA and AKI immunoreactive cDNAs

The HDE cDNA was isolated by William Nuttley. The AKI cDNAs were isolated with the assistance of Rachel Szilard. The HDE cDNA was isolated from the unamplified λgt11 library. Approximately 9,000 plaques were screened and 4 potentially positive clones were amplified. The AKI immunoreactive cDNAs were isolated from the amplified library after removal of the 4 HDE cDNAs. Approximately 60,000 plaques were screened and 22 positive clones were amplified. Rabbit anti-HDE serum was used at a dilution of 1:400. Guinea pig anti-AKI serum was used at a dilution of 1:500. Immunoreactive phage plaques were amplified and purified to homogeneity by successive isolations (corings) of the phage plaques, followed by replating.

2.10 Lambda EMBL3 genomic library construction and screening

The *C. tropicalis* genomic library was constructed in λEMBL3 by Richard Rachubinski by ligation of size-fractionated *C. tropicalis* DNA, which had been partially digested with *Sau*3AI, between the *Bam*HI sites of the λEMBL3 multiple cloning site (Murray and Rachubinski, 1987a). The ligation mix was packaged using Gigapack Gold (Stratagene).

Phage were incubated for 30 min at room temperature with an overnight culture of *E. coli* NM539 grown to stationary phase in LB plus 0.2% maltose, mixed with LB-soft top agar at 50°C and poured onto LB agar plates pre-warmed to 37°C. After incubation at 37°C for approximately 8 h, the plaques were overlaid with nitrocellulose filter disks for 1 min. The filters were placed in DNA denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 1 min and then in neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl (pH 8.0)) for 8 min. The filters were rinsed in 3X SSC and probed for the *HDE* gene by Southern hybridization (section 2.16.3) using the BA fragment of the HDE cDNA labelled with [α^{32}-P]-dATP (section 2.15.1).

Approximately 30,000 plaques were screened and 5 potentially positive clones were identified, two of which (3a and 5b), were amplified and purified to homogeneity by successive isolations (corings) of the phage plaques and replatings.

2.11 Lambda EMBL3 DNA isolation

Recombinant λEMBL3 DNA was isolated using modifications of the

procedures of Maniatis *et al.* (1982) and Davis *et al.* (1986). An exponentially growing culture of *E. coli* LE392 was obtained by inoculating 10 ml of LB medium plus 0.2% maltose and 10 mM $MgCl_2$ with 50 μl of an overnight culture, followed by incubation with shaking for 1 h at 37°C. The phage were amplified by addition of a phage plug (coring) to the cells, followed by incubation for 6 h with agitation. 0.5 ml of chloroform was added to the culture, and the culture was incubated with agitation for 15 min to complete cell lysis. The cell debris was pelleted by centrifugation in a Sorvall SS-34 at 8,000 rpm (7,700 x g_{max}) for 15 min at room temperature. The titre of the phage in the supernatant was determined by successive platings of 10-fold serial dilutions of supernatant. Phage particles were eluted from 2 plates containing approximately 10^5 plaques each. 3 ml of SM was added to the one plate, which was then agitated for 1 h at room temperature. The supernatant was removed and added to the second plate, which was treated in the same manner as the first. The eluted phage were further amplified by adding 100 μl of the phage-containing SM supernatant to an exponentially growing 8 ml culture of *E. coli* LE392 cells in LB plus 0.2% maltose. The culture was incubated as described above. After cell lysis was detected (approximately 7 h), chloroform was added, and the cell debris was removed by centrifugation as above. 1.5 ml of the supernatant (containing approximately 10^9 pfu $\cdot$ ml^{-1}) was added to a 500 ml culture of exponentially growing *E. coli* LE392. 500 ml of LB plus 0.2% maltose was inoculated with 1 ml of an overnight culture of *E. coli* LE392 followed by incubation at 37°C until evidence of cells could be detected (approximately 45 min to 1 h). The

culture was grown until complete lysis had occurred (7 to 8 h), and 50 ml of chloroform was added. The cell debris was removed by centrifugation in a Sorvall GSA rotor at 6,300 rpm (6,500 x g_{max}) for 15 min at room temperature. The phage were precipitated from the supernatant by the addition of 0.5 volume of 20% PEG 8000, 2 M NaCl, followed by incubation at 4°C overnight. The phage were pelleted by centrifugation in a Sorvall GSA rotor at 8,400 rpm (11,250 x g_{max}) for 30 min at 4°C. The phage were resuspended in a minimal volume of SM and purified by density gradient centrifugation through a 4.13 M CsCl gradient in a Beckman VTi50 rotor at 40,000 rpm (132,000 x g_{av}) for 20 h at 20°C. The phage particles were visible as an opaque band and were extracted with a 20 gauge needle and syringe. The phage were dialysed twice against 25 mM Tris-HCl (pH 8.0), 10 mM NaCl, 10 mM $MgCl_2$ for 1 h each and extracted (Maniatis *et al.*, 1982).

2.12 Transformation of *Escherichia coli*

Generally *E. coli* DH5α were transformed as outlined by Gibco/BRL for the selection, amplification and maintenance of recombinant plasmids. 50 μl of cells were thawed on ice from -70°C, and plasmid DNA, or 2-3 μl of a ligation mix, was added. The cells were incubated on ice for 30 min, heat-shocked at 37°C for 20 sec and returned to ice for 2 min. 1 ml of LB was added, and the cells were incubated with shaking at 37°C for 1 h. The cells were then plated at varying dilutions onto LB agar containing ampicillin (100 $\mu g \cdot ml^{-1}$) or tetracycline (12.5 $\mu g \cdot ml^{-1}$). When needed, 50

μl of 2% X-gal in dimethylformamide was applied to the plates to facilitate blue-white screening of recombinants.

2.13 Plasmid DNA isolation

Plasmid DNA was isolated from *E. coli* essentially as described by Birnboim and Doly (1979). Cells from a single colony were grown to saturation in LB containing either tetracycline at 12.5 $\mu g \cdot ml^{-1}$ or ampicillin at 100 $\mu g \cdot ml^{-1}$.

2.13.1 Small scale plasmid isolation

For a small scale plasmid isolation, cells from 1.5 ml of medium were pelleted by centrifugation in a microfuge for 1 min. The cells were resuspended in 100 μl of 50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA. Cellular DNA was denatured by the addition and gentle mixing of 200 μl of freshly prepared 1.0% SDS, 0.2 M NaOH. After a 5 min incubation on ice, 150 μl of potassium acetate (3 M with respect to K^+ and 5 M with respect to acetate) was added, followed by a 5 min incubation on ice. The resultant precipitated protein, chromosomal DNA and high molecular weight RNA were removed by centrifugation for 5 min at 4°C. The supernatant was first extracted with an equal volume of phenol:$CHCl_3$:isoamyl alcohol (51:50:1), and then with an equal volume of $CHCl_3$:isoamyl alcohol (50:1). The plasmid DNA was then pelleted from the aqueous phase by addition of 2 volumes of ethanol, followed by centrifugation in a microcentrifuge for at least 10 min. The supernatant was

removed, and the pellet was washed with 70% ethanol and re-centrifuged. The plasmid pellet was dried in a vacuum desiccator and dissolved in TE (pH 8.0) containing 20 $\mu g \cdot ml^{-1}$ RNAseA.

Recombinants were identified by restriction endonuclease mapping (section 2.14.1) and where necessary, by Southern blotting (section 2.16.2) and/or dideoxy-nucleotide sequencing (section 2.18).

2.13.2 Large scale plasmid isolation

For a large scale preparation of plasmid DNA (between 100 ml and 500 ml of culture) the procedure of Birnboim and Doly (1979; section 2.13.1) was scaled up and modified accordingly. The organic extraction steps were eliminated, and the plasmid DNA was precipitated from aqueous solution by the addition of 0.6 volume of 2-propanol, followed by centrifugation in a Sorvall SS-34 rotor at 10,000 rpm (12,000 x g_{max}) for 30 min at room temperature. The DNA pellet was washed with 70% ethanol, re-centrifuged and dried in a vacuum desiccator prior to dissolution in TE (pH 8.0). Plasmids were further purified by CsCl density gradient centrifugation in the presence of ethidium bromide as described by Radloff *et al.* (1967) and Maniatis *et al.* (1982). Centrifugation was performed either in a Beckman Type 65 rotor at 42,000 rpm (130,000 x g_{av}) for at least 48 h or in a Beckman TLV 100 rotor at 100,000 rpm (338,000 x g_{av}) for at least 3 h. The band corresponding to covalently closed circular plasmid was removed from the gradient using a 20 gauge needle and syringe, and the DNA was

extracted with water-saturated 1-butanol to remove the ethidium bromide. The CsCl was removed from the solution by dialysis against TE (pH 8.0) at 4°C (48 h, at least two changes of buffer).

2.14 Standard DNA manipulations

2.14.1 Restriction endonuclease digestions

DNA digestions with restriction endonucleases were generally carried out as described by Maniatis *et al.* (1982) with consideration of the manufacturers' suggestions. After digestion, 6X DNA sample dye I (0.25% bromophenol blue, 0.25% xylene cyanol, 40% sucrose) was added to samples. The samples were generally separated by agarose (0.8-1.0%) gel electrophoresis in 1X TBE.

2.14.2 Phenol-chloroform extractions

Samples were extracted with phenol-chloroform to remove contaminating protein from nucleic acid. Generally, an equal volume of phenol:chloroform:isoamyl alcohol (51:50:1) was added to a sample containing nucleic acid, followed by vortexing. The organic and aqueous phases were separated by centrifugation. The aqueous phase was removed to a fresh tube. An equal volume of chloroform:isoamyl alcohol (50:1) was added, and the sample was vortexed and centrifuged. The aqueous phase was removed to a fresh tube.

2.14.3 Precipitation of DNA

Nucleic acid was generally precipitated from solution as follows: 0.1 volume of 3 M sodium-acetate (or an equal volume of 4 M ammonium-acetate) and 2.5 volumes of ethanol (or 0.6 volume of 2-propanol) were added to a solution containing nucleic acid. The samples were mixed, generally chilled for at least 10 min and then centrifuged at 12,000 x g_{max} for 5-30 min. The supernatant was removed, and the pellet was washed with 70% ethanol and recentrifuged. The pellets were dried *in vacuo* and dissolved in an appropriate volume of either TE (pH 8.0) or water.

2.14.4 Isolation of DNA fragments

DNA fragments were separated by agarose gel electrophoresis using 0.8-1% GTG (SeaKem) agarose. The bands of interest were excised with a razor blade. The DNA was purified from the agarose gel slice by unidirectional electroelution (International Biotechnologies unidirectional electroeluter, model UEA). Gel slices were placed in the chamber of the electroeluter. 80 μl of 7.5 M ammonium acetate, 0.01% bromophenol blue, was added to the V-slot of the electroeluter. DNA was eluted from the gel at 100 V for 1 h in 0.5X TBE and trapped within the high salt in the V-slot. The DNA/salt solution (350 μl) was removed from the V-slot, and the DNA was pelleted by the addition of 1 ml of ethanol followed by centrifugation.

2.14.5 Polishing 5' overhangs

DNA fragments with 5' overhangs were made blunt by the addition of 1 U of the large fragment of *E. coli* DNA polymerase (Klenow fragment) to 14 μl of DNA (at approximately 100 ng per μl) in 1X KGB. The reaction mixture was incubated for 2 min at 37°C. Four μl of dNTPs (0.25 mM each), 2 μl H_2O and 2 μl 5X KGB were added, and the reaction was incubated for an additional 10 min.

2.14.6 Ligations

DNA fragments were ligated according to the conditions outlined in the International Biotechnologies catalogue, Vol. II. Generally, ligations of DNA fragments with complementary overhangs were done at a 2:1 insert:vector molar ratio at 20 ng $\cdot$ μl^{-1} for >2 h at room temperature. Blunt-end ligations were done at a 2 to 4:1 insert:vector ratio at 50 ng $\cdot$ μl^{-1} overnight at room temperature (Dugaiczyk *et al.*, 1975).

2.14.7 Phosphorylation of oligodeoxyribonucleotides

Oligodeoxyribonucleotides were phosphorylated at their 5' end by incubating 200 pmols of the oligodeoxyribonucleotide with 5 U of T4 polynucleotide kinase in 50 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT, 3.3 μM ATP (final volume, 30 μl) for 45 min at 37°C. The reaction was stopped by heating at 65°C for 10 min.

2.15 Radiolabelling of DNA

2.15.1 Random primer labelling

Large fragments of DNA (greater than 500 bp) were labelled by the random primer method of Feinberg and Vogelstein (1983, 1984) as outlined by the kit manufacturer (BRL). 25 ng of linear DNA was denatured by boiling, then quick-cooled on ice. 15 μl of random primers buffer mixture (0.67 M HEPES, 0.17 M Tris-HCl, 17 mM $MgCl_2$, 33 mM ß-mercaptoethanol, 1.33 $mg \cdot ml^{-1}$ BSA, 18 OD_{260} $units \cdot ml^{-1}$ oligodeoxyribonucleotide primers (hexamer fraction) (pH 6.8)), was added to the denatured DNA in 23 μl of H_2O. 2 μl each of 0.5 M dCTP, 0.5 M dGTP and 0.5 M dTTP (in 3 mM Tris-HCl (pH 7.0), 0.2 mM EDTA), 5 μl (50 μCi) of [α-^{32}P]dATP, (3000 $Ci \cdot mmol^{-1}$, 10 $mCi \cdot ml^{-1}$) and 3 units (3 $units \cdot \mu l^{-1}$) of Klenow fragment were added, and the reaction was incubated at room temperature for at least 1 h. The labelled DNA was separated from unincorporated nucleotides by Sephadex G-50 spin-column chromatography (Maniatis *et al.*, 1982). Incorporated ^{32}P was quantitated by scintillation counting (Beckman LS3801, window setting = 0 to 1000). Typically, the specific activity of the labelled DNA was 5-10 x 10^6 $cpm \cdot \mu g^{-1}$.

2.15.2 End labelling of oligodeoxyribonucleotides

Oligodeoxyribonucleotides were radiolabelled at their 5' end by incubation of 6 pmols of the nonphosphorylated oligodeoxyribonucleotide with 10 μCi of [γ-^{32}P]ATP (3,000 $Ci \cdot mmol^{-1}$) and 3-5 units of T4 polynucleotide kinase in 50 mM Tris-HCl (pH

8.0), 10 mM $MgCl_2$ and 5 mM DTT for 40 min at 37°C in a 30 μl volume. The labelled DNA was separated from unincorporated nucleotides by spotting an aliquot of the reaction mixture on DEAE (diethylaminoethyl) cellulose paper (Whatman DE-81) and removing the free nucleotide by washings in 1X P (3 washes of 5 min each at room temperature). The paper was allowed to air-dry, and the incorporated ^{32}P was quantitated by scintillation counting as above (section 2.15.1). Typically, the specific activity of the labelled DNA was 2.5-5 x 10^7 cpm $\cdot$ μg^{-1}.

2.16 Identification of recombinants

2.16.1 Colony hybridization

Colony hybridization was performed by the technique of Grunstein and Hogness (1975) as modified by Meselson (1980). Colonies were either grown directly on, or transferred to, nitrocellulose by overlaying bacterial colonies with nitrocellulose disks. After removal of the disks, the bacteria were lysed and the DNA was denatured by overlaying the disks onto 3 ml of 0.5 M NaOH. The blots were neutralized by overlaying the disks onto 3 ml of 1 M Tris-HCl (pH 7.6) and then onto 3 ml of 0.5 M Tris-HCl (pH 7.5), 1.5 M NaCl. Each step was performed twice. The nitrocellulose disks were then washed of bacterial debris by rinsing in 2X SSC. The disks were then probed for recombinant insert sequences by hybridization (section 2.16.3). All selected recombinants were confirmed by Southern blotting (Southern, 1975) using purified plasmid DNA digested with diagnostic restriction enzymes.

2.16.2 Southern blotting

DNA was transferred to nitrocellulose from agarose gels essentially as described by Southern (1975) and modified by Ausubel *et al.* (1989). The gels were immmersed in 0.4 M NaOH, 0.8 M NaCl for 1 h (3 changes) to denature complementary DNA strands. The gels were neutralized by immersion in 1.5 M Tris-HCl (pH 7.5), 1.5 M NaCl for 1 h (3 changes) and DNA was transferred to nitrocellulose in 6X SSC either by capillary action overnight or with the aid of a positive pressure transfer chamber (Stratagene) (50 mm•Hg, 1 h). The DNA was fixed to the nitrocellulose either by baking *in vacuo* (Napco vacuum oven model 5831, Fisher Scientific) for 2 h at 80°C or by exposure to UV light (120,000 $\mu J \cdot cm^{-2}$ λ = 254 nm; UV Stratalinker 1800, Stratagene).

2.16.3 Hybridization of DNA

DNA immobilized onto nitrocellulose by colony hybridization (section 2.16.1) or by Southern blotting (section 2.16.2) was probed for DNA sequences complementary to radiolabelled DNA by the methods of Murray and Rachubinski (1987a) and Rachubinski *et al.* (1985) in the case of DNA fragments and by the method of Harley (1987) in the case of end-labelled oligodeoxyribonucleotides.

Blots probed with DNA fragments labelled by the random primer method (section 2.15.1) were prehybridized in 1.25X SSC, 0.16X Denhardt's solution, 4 $\mu g \cdot ml^{-1}$ sheared salmon sperm DNA, 0.001% SDS, 0.02 M sodium-phosphate (pH 7.0) for 3 h

at 65°C. After prehybridization, the blots were incubated for at least 16 h in 1.25X SSC, 0.16X Denhardt's solution, 4 $\mu g \cdot ml^{-1}$ sheared salmom sperm DNA, 0.001% SDS, 0.02 M sodium-phosphate (pH 7.0), 30% deionized formamide, plus 0.2-1 x 10^6 cpm of ^{32}P-labelled DNA (denatured by incubation in boiling water for 5 min followed by quick-cooling on ice) probe per ml at 50°C. The blots were then washed in 1X SSC at 55°C for 1 h (4 changes) and exposed to film.

Blots probed with end-labelled oligodeoxyribonucleotides (section 2.15.2) were prehybridized in 5X SSC, 1X P, 5X Denhardt's solution for at least 0.5 h at room temperature and then hybridized in the same solution plus the labelled oligodeoxyribonucleotide (2-4 x 10^5 $cpm \cdot ml^{-1}$) overnight at room temperature. The blots were then washed in 2X SSC at 35°C for 1 h (4 changes) before exposure to film.

2.17 Exonuclease III deletions

Deletion mutants of the genomic clone for HDE were constructed in the vector pGEM5Zf(+). The *HDE* gene was digested with exonuclease III in both directions (Henikoff, 1984). Transformations were in *E. coli* DH5α. Recombinant pGEM5Zf(+) harbouring the *HDE* gene was digested with *Sph*I and *Sac*II in the multiple cloning site of the vector. 10 μg of linearized plasmid (approximately 2 pmol of susceptible ends) was extracted with phenol:$CHCl_3$:isoamyl alcohol (51:50:1) (section 2.14.2), precipitated with ethanol (section 2.14.3) and resuspended in 100 μl of EXO III buffer (50 mM Tris Cl (pH 7.5), 5 mM $MgCl_2$, 5 mM DTT, 50 $\mu g \cdot ml^{-1}$ BSA, 20 mM KCl). 650 U (2.5 μl)

of exonuclease III was added, followed by incubation at 37°C. 3 μl aliquots were removed at 30 sec time intervals and added to tubes containing 5 μl of 5X S1 buffer (250 mM sodium-acetate (pH 4.5), 1 M NaCl, 5 mM $ZnCl_2$) and 13 μl H_2O. The tubes were then placed in a dry ice-ethanol bath. In total, 30 samples were removed. Exonuclease III was inactivated by heating the samples at 70°C for 10 min. The samples were cooled to room temperature, and 4 U (4 μl) of S1 nuclease was added. The S1 nuclease digestion was allowed to proceed for 30 min at room temperature. The digestion was stopped by a pH shift from the addition of 5 μl of 6X S1 stop buffer (0.5 M Tris-HCl (pH 8.0), 0.125 M EDTA) to each sample. The samples were then heated at 70°C for 10 min, and the DNA was precipitated (section 2.14.3). The pellet was dissolved in 10 μl of 1X KGB. 1 μl was removed for analysis by agarose gel electrophoresis. The remainder of the sample was treated with Klenow fragment to generate blunt ends (section 2.14.5). Plasmids were recircularized by ligation overnight at room temperature, and ligation mixes were used to transform *E. coli* DH5α (section 2.12).

2.18 DNA sequencing

DNA sequencing was done by the method of Sanger *et al.* (1977) using chemically modified T7 DNA polymerase (Tabor and Richardson, 1987; United States Biochemicals) or recombinant T7 DNA polymerase (Pharmacia). Plasmid DNA samples, purified by the method of Birnboim and Doly (1979), were prepared for double-stranded dideoxynucleotide sequencing by the method of Zhang *et al.* (1988). Approximately 3

μg of plasmid DNA was denatured by the addition of NaOH to 0.4 M. After a 5 min incubation at room temperature, the solution was neutralized by the addition of ammonium acetate and EDTA to a final concentration of 0.2 M and 0.2 mM, respectively. The DNA was precipitated by the addition of 2.5 volumes of ice-cold ethanol (section 2.14.3). The dried DNA pellets were dissolved in H_2O. Sequencing was then carried out as described by the maunfacturer (Sequenase™ Manual) using [α-^{32}P]dATP (3000 Ci$\cdot$mmol^{-1}) as the radiolabelled nucleotide, except that twice as much primer (1 pmol) was used in the annealing reaction. The DNA fragments were separated by electrophoresis in 6% acrylamide (40:1 acrylamide:bis-acrylamide), 8 M urea, in 1X TBE at 60 watts. The gels were dried and exposed to Kodak X-Omat K-1 film (Rochester, NY).

2.19 Isolation of RNA

2.19.1 Total RNA isolation from *Candida tropicalis*

Total RNA was isolated (Maccecchini *et al.*, 1979; Murray and Rachubinski, 1987b) from *C. tropicalis* grown in medium containing either oleic acid or glucose. Cells were pelleted at 5000 rpm (4000 x g_{max}) in a Sorvall GSA rotor, washed with sterile water, and frozen in liquid N_2. The frozen cells were pulverized with a pestle in a mortar immersed in liquid N_2. The crushed cells were added to 150 mM NaCl, 50 mM Tris-Cl (pH 8.0), 5 mM EDTA, 5% SDS. The mixture was extracted 3 times for 20 min each with 5 mM EDTA-saturated phenol:$CHCl_3$: isoamyl alcohol (51:50:1). Nucleic

acid was precipitated by the addition of 0.1 volume 3 M sodium-acetate and 2.5 volumes ethanol at -20°C and pelleted by centrifugation at 10,000 rpm (12,000 x g_{max}) in a Sorvall SS-34 for 30 min at 4°C. Precipitates were resuspended in 3 M LiCl and kept at 0°C for 24 h to selectively dissolve DNA. RNA was collected by centrifugation at 10,000 rpm in a Sorvall SS-34 (12,000 x g_{max}) for 30 min at 4°C. The RNA pellet was dried *in vacuo* and dissolved in H_2O. Poly (A)$^+$RNA was isolated by oligo (dT)-cellulose chromatography (Aviv and Leder, 1972).

2.19.2 Oligo (dT)-cellulose chromatography

Poly (A)+ RNA was isolated from total RNA by oligo (dT)-cellulose chromatography essentially as described by Maniatis *et al.* (1982). The column (1 ml bed volume) was prepared for RNA binding by successive passages of 30 volumes of binding buffer (10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1 mM EDTA, 0.5% SDS), of elution buffer (10 mM Tris-HCl (pH 7.5), 0.5% SDS) and again of binding buffer. RNA (1-10 mg $\cdot$ ml^{-1} in 0.5 M NaCl, 1 mM EDTA, 0.5% SDS) was heated at 65°C for 10 min and quick-cooled on ice before application to the column. The RNA was passed through the column three times. Poly (A)$^+$ RNA was eluted from the column by the addition of 3 ml of elution buffer. Eluted RNA was precipitated by the addition of potassium-acetate (pH 5.5) to 0.2 M and of 2.5 volumes of ethanol, followed by incubation overnight at -20°C. The RNA was pelleted by centrifugation at 10,000 rpm in a Sorvall SS-34 rotor (12,000 x g_{max}) for 30 min at 4°C, washed with 70% ethanol at

4°C, dried *in vacuo*, and dissolved in sterile H_2O.

2.20 Analysis of RNA

2.20.1 Analysis of RNA slot blot

Poly $(A)^+$ RNA was denatured by boiling and added to a solution so as to yield an appropriate final concentration in 50 $\mu g \cdot ml^{-1}$ calf liver tRNA, 3X SSC, 5 mM methyl mercury (Andrews *et al.*, 1982). Samples containing either 4 ng or 1 ng of poly $(A)^+$ RNA in 100 μl volumes were blotted to nitrocellulose using a slot-blot apparatus (Minifold II Slot-Blot System, model SRC072/0, Schleicher and Schuell Inc.). The RNA was bound to the nitrocellulose by baking *in vacuo* (Napco vacuum oven model 5831, Fisher Scientific) for 2 h at 80°C. The blot was prehybridized overnight in 3X SSC, 10X Denhardt's, 0.1% SDS, 50 $\mu g \cdot ml^{-1}$ salmon sperm DNA, 10 $mg \cdot ml^{-1}$ poly (A) at 65°C overnight. Hybridization was with a ^{32}P-radiolabelled (Feinberg and Vogelstein, 1983) recombinant pUC118 plasmid containing a partial cDNA encoding HDE (BA) (section 2.15.1) at a concentration of 1-3 x 10^6 $cpm \cdot ml^{-1}$ in 3X SSC, 10X Denhardt's, 0.1% SDS, 50 $\mu g \cdot ml^{-1}$ salmon sperm DNA, 10 $mg \cdot ml^{-1}$ poly (A), 10% dextran sulfate for 15 h at 65°C. The blots were then washed briefly in 1X SSC, 0.1% SDS at room temperature, followed by successive 20 min washes at 65°C in 1X SSC, 0.1% SDS; 0.3X SSC, 0.1% SDS; 0.1X SSC, 0.1% SDS (Ryall *et al.*, 1984). Signals were normalized by hybridization of radiolabelled oligo-$(dT)_{18}$ (Harley, 1987). For oligo-$(dT)_{18}$ hybridization, the blot was prehybridized with 5X SSC, 1X P, 5X Denhardt's

solution for 30 min at room temperature. 80 pmols (8×10^5 cpm) end-labelled (section 2.15.2) oligo-$(dT)_{18}$ were hybridized at room temperature for 4 h as in section 2.16.3. The blots were washed in 2X SSC at 35°C for 1 h (4 changes). Slot blots were scanned by densitometry using a GS-300 densitometer (Hoefer Scientific, San Francisco, CA).

2.20.2 *In vitro* translation

RNA was translated *in vitro* in a micrococcal nuclease treated rabbit reticulocyte lysate system (Pelham and Jackson, 1976) purchased from Promega Corp. Translations were performed as outlined by the manufacturer using L-^{35}S-methionine (1000-1500 Ci $\cdot$ mmol^{-1}) as the radiolabelled amino acid. RNA was heated to 65°C for 5 min, then quick-cooled on ice (Mortensen *et al.*, 1984) prior to addition to the translation lysate. The translation mix was incubated for 60 min at 30°C. Translation products were analyzed by trichloroacetic acid (TCA) precipitation (Mans and Novelli, 1961; section 2.20.3), immunoprecipitation (Fujiki *et al.*, 1984; section 2.2.1), SDS-PAGE (Laemmli, 1970; section 2.3.2) and fluorography (Laskey and Mills, 1975; section 2.3.4).

2.20.3 Trichloroacetic acid precipitation of translation products

Five μl of the *in vitro* translation mix was diluted with 19 volumes of NP-40 mix plus methionine (10 mM Tris-HCl (pH 8.5), 150 mM NaCl, 200 mM methionine, 0.02% NaN_3, 1% (w/v) NP-40). Twenty μl of the diluted translation mix was spotted

in duplicate onto 1 cm x 1 cm Whatman 3 mm chromatography paper and allowed to air dry. The protein was precipitated onto the filter paper by placing the paper in 10% TCA containing 20 mM methionine and incubating on ice for 20 min. The filters were then washed sequentially in 5% TCA containing 20 mM methionine at 98-100°C for 20 min, in 4 changes of 5% TCA at room temperature, in 2 changes of ethanol at room temperature and in boiling ether. The filters were allowed to air-dry and then placed into 10 ml of aqueous scintilation fluid (ACS), and the preciptated ^{35}S-methionine was quantitated by scintillation counting using a Beckman LS 3801 scintillation counter with a window setting of 0 to 1000.

2.21 Immunoprecipitation of translation products

Translation products were immunoprecipitated by the procedure of Fujiki *et al.* (1984). Translation mixes were diluted 10-fold and adjusted to 1% (w/v) NP-40, 10 mM Tris-HCl (7.4), 150 mM NaCl, 10 mM methionine, 25 $\mu g \cdot ml^{-1}$ leupeptin, 25 $\mu g \cdot ml^{-1}$ antipain, 12.5 $\mu g \cdot ml^{-1}$ chymostatin and 12.5 $\mu g \cdot ml^{-1}$ pepstatin. The diluted mixes were centrifuged at 41,000 rpm (150,000 x g_{max}) in a Beckman Type 65 rotor for 1 h at 4°C to remove ribosomal complexes. SDS and EDTA were added to the post-ribosomal supernatants to final concentrations of 0.1% and 2 mM, respectively. Antiserum was incubated with post-ribosomal supernatants for 90 min at room temperature, and then overnight at 4°C. The antigen-antibody complexes were adsorbed to *Staphylococcus aureus* cells prewashed with binding buffer (10 mM Tris-HCl (pH

7.4), 1% (w/v) NP-40, 0.1% SDS, 150 mM NaCl, 10 mM methionine, 2 mM EDTA, 25 $\mu g \cdot ml^{-1}$ leupeptin, 25 $\mu g \cdot ml^{-1}$ antipain, 12.5 $\mu g \cdot ml^{-1}$ chymostatin, 12.5 $\mu g \cdot ml^{-1}$ pepstatin) for 90 min at room temperature. The complexes were collected by centrifugation (1 min, microfuge), and washed 3 times with binding buffer, twice with 10 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.05% NP-40, 10 mM methionine, 25 $\mu g \cdot ml^{-1}$ leupeptin, 25 $\mu g \cdot ml^{-1}$ antipain, 12.5 $\mu g \cdot ml^{-1}$ chymostatin and 12.5 $\mu g \cdot ml^{-1}$ pepstatin, once with 10 mM Tris-HCl (pH 7.4), 150 mM NaCl, 10 mM methionine, 25 $\mu g \cdot ml^{-1}$ leupeptin, 25 $\mu g \cdot ml^{-1}$ antipain, 12.5 $\mu g \cdot ml^{-1}$ chymostatin and 12.5 $\mu g \cdot ml^{-1}$ pepstatin, and then once with 10 mM Tris-HCl (pH 7.4), 10 mM methionine, 25 $\mu g \cdot ml^{-1}$ leupeptin, 25 $\mu g \cdot ml^{-1}$ antipain, 12.5 $\mu g \cdot ml^{-1}$ chymostatin and 12.5 $\mu g \cdot ml^{-1}$ pepstatin. The washed antigen-antibody complexes were dissociated by boiling in SDS-PAGE sample buffer and analyzed by SDS-PAGE (section 2.3.3) and fluorography (section 2.3.4).

2.22 Hybridization-selection translation

Hybridization-selection translation was done by the procedure of Parnes *et al.* (1981) as modified by Rachubinski *et al.* (1985). Five μg of recombinant pUC118 containing a partial HDE cDNA insert (BA) in 20 μl of H_2O was denatured by heating in boiling water for 10 min, followed by quick-cooling on ice. 20 μl of 1 M NaOH was added, and the solution was incubated at room temerature for 20 min. The solution was then neutralized by the addition of 20 μl of 2 M NaCl, 0.3 M sodium-citrate, 0.5 M

Tris-HCl (pH 8.0), 1 M HCl. The solution was spotted onto nitrocellulose disks in 5 μl aliquots. The disks were air-dried for 1 h, washed in 50 ml of 6X SSC, and baked *in vacuo* (Napco vacuum oven model 5831, Fisher Scientific) for 2 h at 80°C. 30 μg of poly $(A)^+$ RNA isolated from oleic acid-grown *C. tropicalis* cells was heated to 70°C for 10 min in 0.5 ml of hybridization solution (65% (v/v) deionized formamide, 20 mM 1,4-piperazinediethanesulfonic acid (Pipes) (pH 6.4), 0.2% SDS, 0.4 M NaCl, 0.1 $mg \cdot ml^{-1}$ yeast tRNA) and hybridized to immobilized DNA overnight at 50°C in a siliconized glass vial. The filters were transferred to sterile 50 ml plastic tubes and washed at 65°C nine times with 50 ml volumes of 0.01 M Tris-HCl (pH7.6), 0.15 M NaCl, 1 mM EDTA, 0.5% SDS and twice with 50 ml of 0.01 M Tris-HCl (pH 7.6), 0.15 M NaCl, 1 mM EDTA. The hybridized RNA was eluted in a siliconized microfuge tube by the addition of 300 μl H_2O and 30 μg of tRNA, followed by heating in a boiling water bath for 1 min. The samples were quick-frozen in a dry-ice ethanol bath and thawed slowly on ice (5 h). The RNA was extracted with phenol:chloroform:isoamyl alcohol (51:50:1) and precipitated by the addition of 0.1 volume of sodium-acetate and 2.5 volumes of ethanol (section 2.14.3). The precipitates were dried and dissolved in 5 μl of water for translation *in vitro* (section 2.20.2). The translation products were immunoprecipitated as described in section 2.21. For the competition asssay, 1 μg of purified HDE was added to the immunoprecipitation reaction (Fujiki *et al.*, 1984; Rachubinski *et al.*, 1985). Protein products were analyzed by SDS-PAGE (section 2.3.2) followed by fluorography (section 2.3.4).

2.23 Yeast transformations

C. albicans spheroplasts were transformed as previously described (Kurtz *et al.*, 1986). *S. cerevisiae* were transformed either by the lithium acetate method (Ito et al., 1983) or by electroporation (Delorme, 1989).

2.23.1 Transformation of *Candida albicans*

C. albicans was transformed by the protoplast yeast transformation procedure (Beggs, 1978; Kurtz *et al.*, 1986). 0.5 ml of an overnight culture of *C. albicans* SGY 243 grown in YEPD was subcultured into 50 ml of fresh YEPD and grown to a concentration of 1-2 x 10^7 cells $\cdot$ ml^{-1}. Cells were pelleted in an IEC clinical centrifuge for 5 min at 2750 rpm at 4°C. The supernatant was discarded, and the cells were resupended in 5 ml of SED (1 M sorbitol, 25 mM EDTA (pH 8.0), 50 mM DTT). The cells were incubated for 10 min at 30°C, pelleted at 2750 rpm as above and resuspended in 5 ml of SCE (1 M sorbitol, 0.1 M sodium-citrate (pH 5.8), 10 mM EDTA (pH 8.0)). 50 μl of β-glucoronidase (110,600 units $\cdot$ ml^{-1}) was added, and the cells were incubated at 30°C for 30 min with gentle agitation. After 80-90% of the cells were converted to spheroplasts (20-30 min), the spheroplasts were collected by centrifugation at 1000 rpm for 5 min in an IEC clinical centrifuge at 4°C. The pellet was washed in 5 ml of CaS (1 M sorbitol, 10 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$) at 4°C. The pellet was resuspended in 0.5 M CaS and divided into 2-5 aliquots. 2-5 μg of plasmid DNA (recombinant pMK22) was added to the spheroplasts, followed by incubation without

agitation at room temperature for 15 min. 1 ml of PEG solution (20% PEG 4000, 10 mM $CaCl_2$, 10 mM Tris-HCl (pH 7.5)) was added, and the cells were pelleted in an IEC clinical centrifuge at 2750 rpm at room temperature. The cells were resuspended gently in 100 μl of SOS (1 M sorbitol, 33% YEPD, 6.5 mM $CaCl_2$) and plated directly onto *C. alb.* rescue plates. Plates were incubated for 2-4 days at 30°C. Colonies were transferred to minimal medium plates for further growth and detection of integrants. This procedure routinely yielded transformation frequencies of approximately 10^3 per μg of plasmid DNA.

2.23.2 Transformation of *Saccharomyces cerevisiae*

S. cerevisiae were transformed either by treatment with lithium-acetate or by electroporation.

2.23.2.1 Lithium-acetate transformation of *Saccharomyces cerevisiae*

An overnight culture of *S. cerevisiae* (DL-1) grown in YEPD was diluted at least 100-fold and then grown to a concentration of 1-2 x 10^7 cells per ml. The cells were then collected by centrifugation in an IEC clinical centrifuge for 5 min at 2750 rpm. The cells were washed with sterile H_2O, resuspended in 0.01 volume (approximately 100 μl) of LiAc (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.1 M lithium-acetate) and incubated for 1 h at 30°C with agitation. 50 μg of sonicated, denatured calf thymus DNA and 1 μg of plasmid DNA were added to the cells followed

by incubation at 30°C without agitation for 30 min. 700 μl of PEG-LiAc (40% Peg 8000, 0.1 M Li- acetate, 1X TE (pH 7.5)) was added, and the cells were incubated for an additional 30 min at 30°C. The cells were pelleted by centrifugation in a microfuge. The pelleted cells were washed twice with TE (pH 7.5) and resuspended in 100 μl of TE (pH 7.5). The cells were plated directly onto minimal medium with the appropriate supplements. The plates were incubated for 2-4 days at 30°C.

2.23.2.2 Electroporation of *Saccharomyces cerevisiae*

S. cerevisiae were electroporated essentially as described by Delorme (1989) using a BRL Cell-Porator series 1600 and voltage booster series 1602. Cells were grown overnight in YEPD and subcultured into 50 ml of fresh YEPD. Cells were then grown for at least 2 doublings to a concentration of 5 x 10^7 cells·ml^{-1} and pelleted by centrifugation at 2750 rpm in an IEC clinical centrifuge. Cells were washed with sterile water, resuspended in 100 μl of 1 M sorbitol, and divided into 20 μl aliquots. 1 μg (1 μl) of plasmid DNA was added to each aliquot, and the cells were suspended between the electrodes of the electroporation cell (1.5 mm distance between electrodes). The electroporation chamber was filled with ice-water, and the cells were electroporated at 250 V (low Ω). The capacitor on the Cell-Porator was set at 330 μF, and the voltage booster was set at a resistance of 4 kΩ.

2.24 Isolation of DNA from *Candida albicans*

Total DNA was isolated from *C. albicans* essentially as described by Holm *et al.* (1986). Cells were grown to late-logarithmic phase in 100 ml of YEPD and collected by centrifugation (Sorvall GSA rotor, 5000 rpm (4,000 x g_{max}), 10 min). Cells were washed in 10 ml SCE (1 M sorbitol, 0.1M sodium-citrate (pH 5.8), 10 mM EDTA (pH 8.0)) and recentrifuged in an IEC clinical centrifuge at 2750 rpm for 5 min. The cells were resuspended in 10 ml SCE, and 67 μl of 14.1 M β-mercaptoethanol was added. The cells remained at room temperature for 10 min. 2 mg of Zymolyase 100T was then added, and the cells were converted to spheroplasts by incubation at 37°C with gentle shaking for 40 min. The spherolasts were collected by centrifugation (1200 rpm, IEC clinical centrifuge, 10 min, room temperature). The pellet was washed in 10 ml SCE and then resuspended in 10 ml of GuHCl buffer (4.5 M guanidine-HCl, 0.1 M EDTA (pH 8.0), 0.15 M NaCl). Sarkosyl (N-laurylsarcosine sodium salt) was added to a final concentration of 0.5%, and the mixture was incubated at 65°C for 10 min. 10.5 ml of ethanol at -20°C was added. The DNA was pelleted by accelerating the rotor (Sorvall SS34) to 10,000 rpm, and holding the rotor speed for 30 sec. The pellet was dissolved in 10X TE (pH 8.0) (facilitated by incubation at 65°C) and then extracted twice with phenol:chloroform:isoamyl alcohol (51:50:1). The nucleic acid was precipitated from the aqueous phase by the addition of 0.6 volume of 2-propanol. The pellet was dissolved in 2 ml TE (pH 8.0), extracted 5 times with phenol:chloroform:isoamyl alcohol (51:50:1) and twice with chloroform:isoamyl alcohol (50:1). The DNA was precipitated

with sodium-acetate and ethanol (section 2.14.3).

2.25 Single-stranded DNA isolation

Single-stranded DNA was isolated from recombinant pGEM vectors by the method of Vieira and Messing (1987). DH5αF' cells harbouring recombinant pGEM phagemids were grown overnight in LB containing 150 $\mu g \cdot ml^{-1}$ ampicillin, then diluted 100-fold into 50 ml of TYPD (1.6% tryptone, 1.6% yeast extract, 0.5% NaCl, 0.25% K_2HPO_4, 0.2% glucose) plus 150 $\mu g \cdot ml^{-1}$ ampicillin and grown with agitation to an OD_{550} of 0.05. Kanamycin was added to a concentration of 75 $\mu g \cdot ml^{-1}$, and M13K07 (Bio-Rad) helper bacteriophage were added to a concentration of 1 x 10^9 pfu $\cdot ml^{-1}$. The cells were grown with agitation for an additional 18 h. The cells were separated from the liberated bacteriophage by centrifugation at 10,000 rpm (12,000 x g_{max}) in a Sorvall SS34 rotor for 10 min. The supernatant was recentrifuged at 11,000 rpm (14,000 x g_{max}) for 15 min. RNAseA was added to the supernatant to a concentration of 3 $\mu g \cdot ml^{-1}$, and the supernatant was incubated at room temperature for 30 min. 0.25 volume of 20% PEG 8000, 2.5 M NaCl was added, and the DNA solution was incubated at 4°C for at least 4 h with rotation to precipitate the phage. The phage were pelleted by centrifugation at 12,000 rpm (17,000 x g_{max}) in a Sorvall SS34 rotor for 20 min at 4°C. The pellet was dried and resuspended in TE (pH 8.0). The nucleic acid was then separated from the phage protein by extraction with an equal volume of phenol (2 times), phenol:chloroform:isoamyl alcohol (51:50:1) (4 times) and chloroform:isoamyl alcohol

(50:1) (2 times). The single-stranded DNA was then precipitated by the addition of 0.1 volume of 7.0 M ammonium-acetate and 2.5 volumes of ethanol (section 2.14.3).

2.26 *In vitro* mutagenesis

Oligodeoxyribonucleotide-directed mutagenesis was performed on single-stranded DNA recombinant pGEM phagemids (Promega Biotech) by the phosphorothioate method (Taylor *et al.*, 1985) as described by the supplier (Amersham), except that *in vitro* synthesized double-stranded DNA was used directly for transformation of DH5α *E. coli* cells.

2.27 Immunofluorescence

Log-phase cell cultures were fixed by the addition of formaldehyde to 3.7% and incubated for 1 h at room temperature, then overnight at 4°C. The cells were washed in 1.2 M sorbitol, 100 mM KPi (pH 7.5), treated with Zymolyase 100T for 30 min at 37°C and processed essentially as described by Pringle *et al.* (1991). Primary antibodies were used at either a 50^{-1} (anti-AKI, anti-SKL) or 100^{-1} (anti-AOX) dilution of the serum. Secondary antibodies were used as a mixture of rhodamine-conjugated goat anti-guinea pig IgG and fluorescein-conjugated goat anti-rabbit IgG. The labelled samples were mounted in 75% glycerol, 2.5% 1,4-diazabicyclo[2,2,2]octane, viewed with oil immersion optics on a Zeiss MC63A equipped for fluorescence and photographed on Kodak TMAX-400 film.

2.28 Computer analysis

Computer analysis was done using an IBM personal computer with the MICROGENIE (Queen and Korn, 1984; Beckman, Palo Alto, CA) or PC/GENE (Intelligenetics, Inc., Mountain View, CA) software.

The hydropathy profile (Fig. 3.2.6.1.1) was determined using the Peptide Structure program of the Computer Genetics Group software package of the University of Wisconsin (Kyte and Doolittle, 1982; Devereux *et al.*, 1984).

3.0 RESULTS AND DISCUSSION

3.1 Overview

The complexities of the biogenesis of peroxisomes are only now beginning to be unravelled. Unlike proteins directed to other organelles, most peroxisomal proteins are synthesized without precursor extensions on polysomes free in the cytoplasm and are post-translationally imported into peroxisomes. This thesis is aimed at determining the nature of signals that direct proteins specifically to peroxisomes. I undertook to answer this question using an *in vivo* expression system in the yeasts *Candida albicans* and *Saccharomyces cerevisiae*. This strategy was chosen for several reasons. *In vivo* expression of a gene and monitoring the ability of the protein product to be imported into peroxisomes *in vivo* circumvents many problems associated with current peroxisomal *in vitro* import systems (see section 3.8). Traditionally, *C. tropicalis* has been used as a model organism for studying peroxisomes in yeast (Lazarow and Fujiki, 1985), because *C. tropicalis* peroxisomes are inducible and can be easily isolated from cells grown on alkanes or fatty acids. Although peroxisomes have been studied extensively in *C. tropicalis*, *in vivo* targeting studies have been hindered by the lack of a good transformation system. *C. albicans* is closely related to *C. tropicalis*, and the advent of an efficient transformation system (Kurtz *et al.*, 1986; 1987) suggested that *C. albicans*

might provide a means to study peroxisomal targeting *in vivo*. Peroxisomes can be induced and isolated from *C. albicans* (Aitchison and Rachubinski, 1990) which, at the beginning of this study, was not possible with *S. cerevisiae*. The ability to isolate peroxisomes from *C. albicans* (and subsequently from *S. cerevisiae*) provided the means to study the import of truncated versions of a peroxisomal protein in an effort to establish the amino acid sequences required for that protein's import into peroxisomes.

C. tropicalis peroxisomal trifunctional enzyme, hydratase-dehydrogenase-epimerase (HDE) was chosen for these studies. The full-length cDNA encoding HDE was cloned from a λgt11 expression library using rabbit anti-HDE serum. The *HDE* gene was subsequently cloned. The cloned gene contains the entire HDE coding sequence plus 5' and 3' non-coding portions sufficient for inducible expression in both *C. albicans* and *S. cerevisiae*. Mutant and wild-type forms of the gene were expressed in the two yeasts to identify the amino acids responsible for targeting HDE to peroxisomes.

3.2 Cloning of the HDE cDNA

3.2.1 Generation of antisera

The first step in the cloning of the *HDE* gene was the generation of antiserum directed against HDE. Peroxisomes were isolated from *C. tropicalis* grown on oleic acid-containing medium, and the polypeptides were separated by SDS-PAGE. A typical polypeptide profile of peroxisomes purified from *C. tropicalis* (and *C. albicans*) is shown

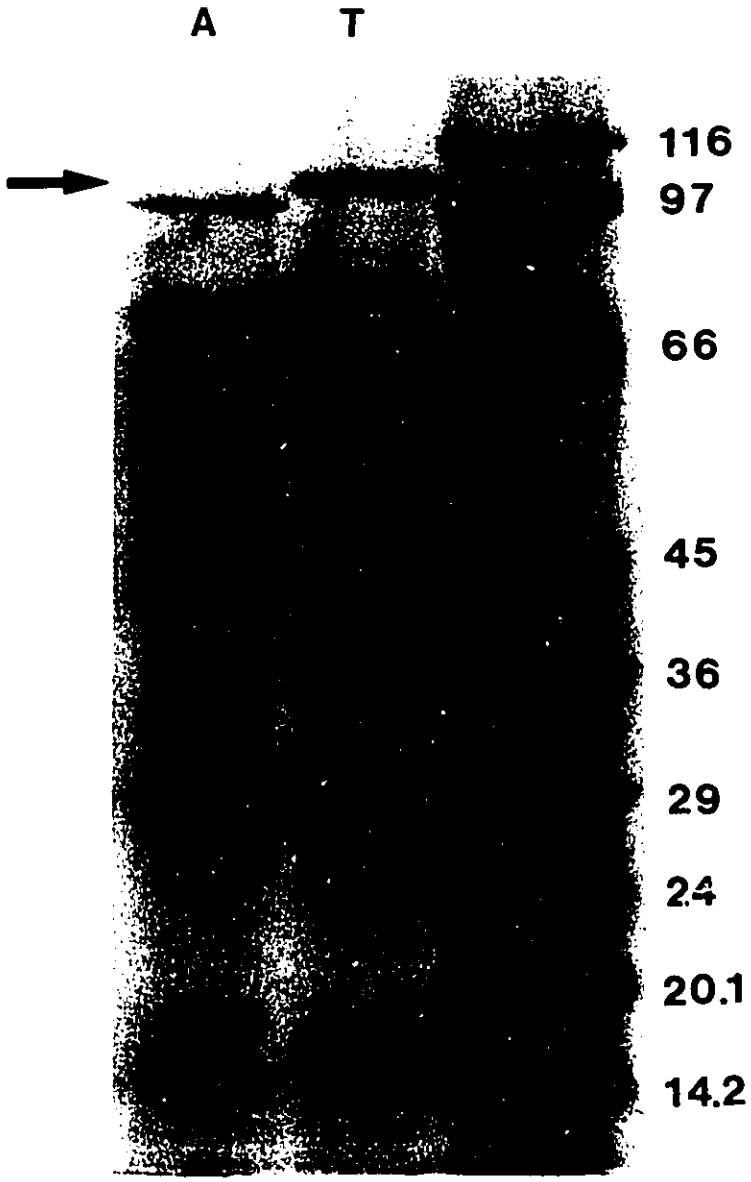


Figure 3.2.1.1 Coomassie-stained 7-15% polyacrylamide gel of peroxisomes purified from *C. tropicalis* (lane T) and *C. albicans* (A). Note the position of HDE (denoted by the arrow). 20 μg of protein was loaded in each lane. The molecular weight standards are in the right lane.

in Fig. 3.2.1.1. The band corresponding to HDE was excised with a razor blade, electroeluted, lyophilized and injected into rabbits for the generation of antiserum. At the same time, anti-AOX and anti-catalase sera were also generated. The specificities of the antisera were confirmed by western blotting and immunoprecipitation of proteins from *in vitro* translated poly $(A)^+$ RNA isolated from oleic acid-grown cells.

3.2.2 Screening of the cDNA expression library[1]

A cDNA expression library was constructed in λgt11 from poly $(A)^+$ RNA isolated from oleic acid-grown *C. tropicalis*. The double-stranded cDNA was ligated into the λgt11 arms, packaged *in vitro* (Hohn and Murray, 1977) and used to infect *E. coli* Y1090 cells (Davis *et al.*, 1986). Approximately 9000 plaques from the unamplified library were screened directly by the double-antibody technique, firstly with anti-HDE at a 400^{-1} dilution and secondly with anti-rabbit IgG (Fc) alkaline phosphatase conjugate at a 7500^{-1} dilution. Four immunoreactive phage plaques were detected by colourimetric reaction with the substrates nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP). The four plaques were amplified and purified to homogeneity by subsequent platings and screenings. Recombinant bacteriophage DNA was isolated by plate lysis (Maniatis *et al.*, 1982) and analyzed by restriction endonuclease digestion. The 4 recombinants yielded insert sizes of 1.1, 1.6, 2.7 and 2.8 kbp, when digested with *Eco*RI. One fragment of 1.1 kbp was present in each clone suggesting that the 4 clones

[1]The *C. tropicalis* cDNA library was constructed and screened by William Nuttley.

were related, and that the 1.1 kbp piece encoded the 3' portion of the gene. The largest clone, HDE BN4 (2.8 kbp) was selected for further analysis. It was hoped that this clone encoded the full-length HDE, since the protein is approximately 100 kDa, which should be encoded by ≈2700 nucleotides. The *Eco*RI fragments of clone HDE BN4 were subcloned into pUC118 to generate plasmids pBA, pSA, pBB and pSB (see Fig. 3.2.2.1).

3.2.3 Identification of the HDE cDNA by hybridization-selection translation

To confirm the identity of the HDE cDNA, plasmid pBA was used to select the mRNA encoding HDE from total poly $(A)^+$ RNA (Fig. 3.2.3.1). Five μg of pBA was denatured, bound to nitrocellulose and hybridized with 30 μg of poly $(A)^+$ RNA isolated from *C. tropicalis* grown on oleic acid-containing medium. After washing, the specifically hybridized RNA was eluted and translated in an *in vitro* rabbit reticulocyte lysate system. Fig. 3.2.3.1 shows the results of this analysis. Lane 11 shows 20 μg of purified peroxisomes separated by SDS-PAGE and stained with Coomassie blue. The arrow indicates HDE. Lanes 1 and 2 are the total ^{35}S-labelled *in vitro* translation products from poly $(A)^+$ RNA isolated from oleic acid- (lane 1) and glucose- (lane 2) grown *C. tropicalis*. Immunoprecipitation using anti-HDE serum from RNA translation products of oleic acid-grown cells specifically immunoprecipitated a polypeptide of ≈ 100 kDa, comigrating with HDE (lane 3). Comparatively little HDE was immunoprecipitated from translation products of RNA isolated from glucose-grown cells (lane

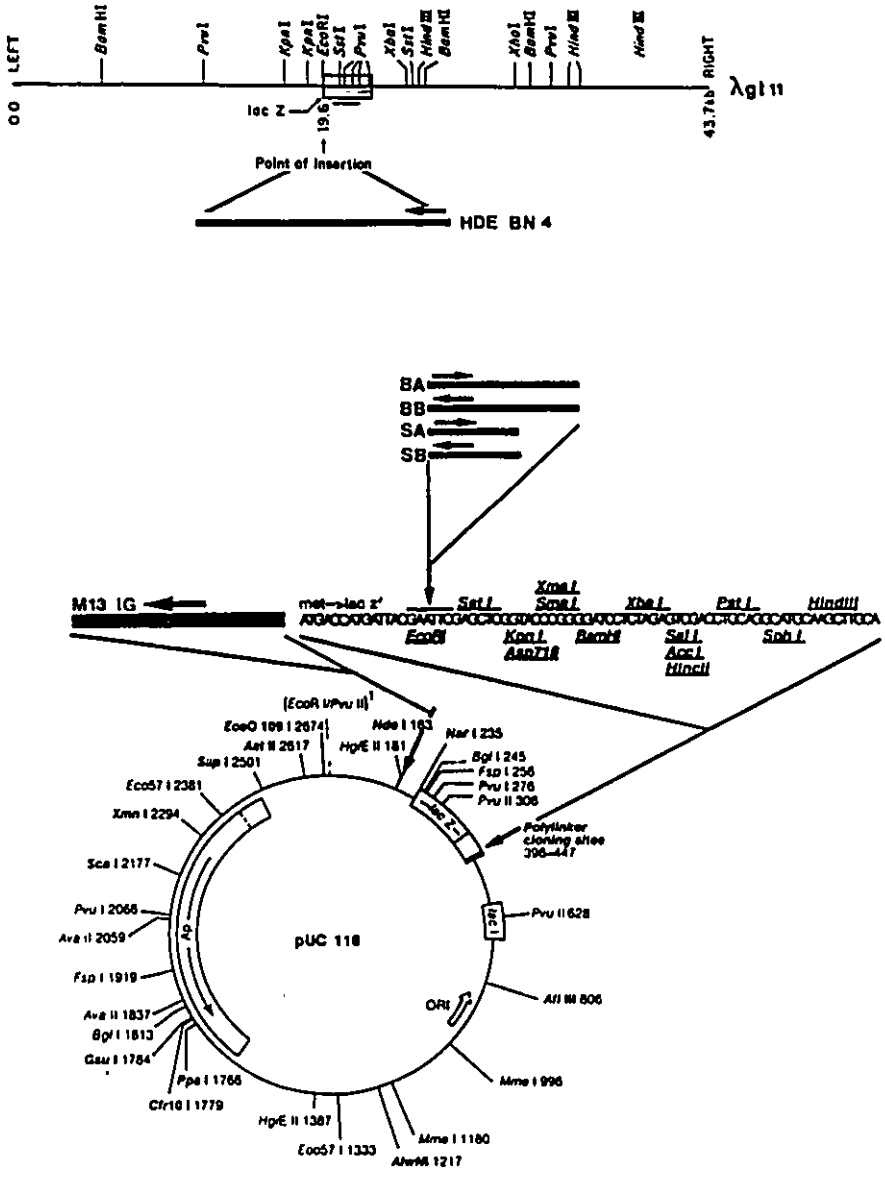


Figure 3.2.2.1 Subcloning of HDE BN4. DNA from the anti-HDE immunoreactive λgt11 bacteriophage plaque was isolated and digested with *Eco*RI. Two fragments of 1.7 and 1.1 kb were isolated from the restriction digestion and subcloned into pUC118 in both orientations to give plasmids pBA, pBB, pSA, pSB as shown.

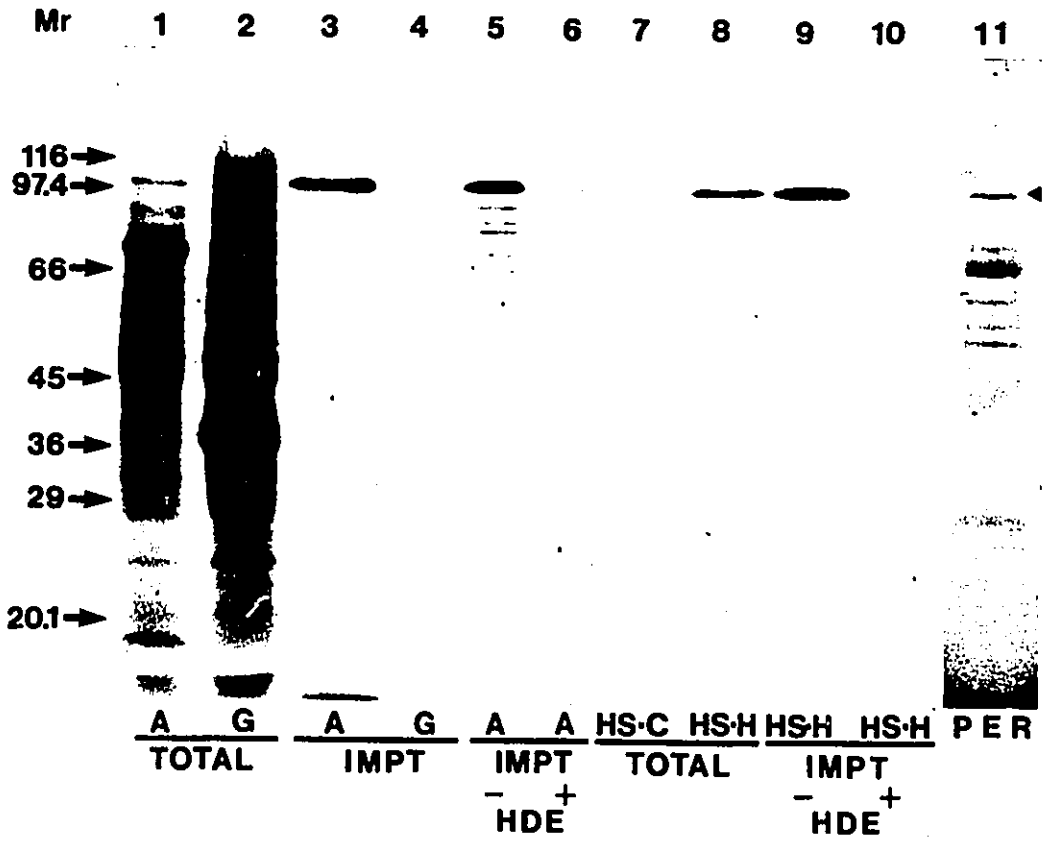


Figure 3.2.3.1 Hybridization-selection translation identification of HDE. Five μg of CsCl-purified pBA was bound to nitrocellulose and hybridized with 30 μg of poly $(A)^+$ RNA from oleic acid-grown *C. tropicalis*. The specifically hybridized mRNA was eluted and translated *in vitro* in the presence of L-^{35}S-Met. The resulting proteins were analyzed by SDS-PAGE and fluorography. The fluorogram is as follows: lanes 1 and 2, total translation products of RNA from cells grown in oleic acid (A) and glucose (G); lanes 3 and 4, translation products immunoprecipitated with anti-HDE serum from an equal number of TCA-precipitable counts of translation mixtures A and G; lanes 5 and 6, translation products from A immunoprecipitated with anti-HDE serum in the absence (lane 5) or presence (lane 6) of 1 μg of purified, unlabelled HDE; lanes 7 and 8, total translation products of mRNA hybridization-selected with pUC118 (HS•C; lane 7) or pBA (HS•H; lane 8); lanes 9 and 10, anti-HDE total translation products from HS•H immunoprecipitated with anti-HDE serum in the absence (lane 9) or presence (lane 10) of 1 μg of purified, unlabelled HDE. 20 μg of purified peroxisomal protein stained with Coomassie blue is shown in lane 11. The arrow at the right shows the position of HDE.

4). This difference was calculated by densitometric analysis to be 9.3-fold, and reflects the induction of HDE-specific RNA by growth of cells in oleic acid-containing medium. The immunoprecipitation was specifically competed by the addition of unlabelled HDE purified from isolated peroxisomes. The addition of 1 μg of unlabelled HDE to the immunoprecipitation reaction caused an 85% reduction in the amount of *in vitro* translated HDE immunoprecipitated, as determined by densitometric scanning (compare lanes 5 and 6). *In vitro* translation of the mRNA selected by hybridization to plasmid pBA yielded a predominant polypeptide comigrating with HDE (lane 8). There was no *in vitro* translation product generated with control pUC118 plasmid (lane 7). The identity of the hybrid-selected translation product was determined to be HDE, because it was immunoprecipitated with anti-HDE serum (lane 9), and the immunoprecipitation competed out by 82% when 1 μg of unlabelled purified HDE was added to the reaction (compare lanes 9 and 10).

3.2.4 Slot blot analysis of HDE-encoding RNA

In order to quantitate the increase in HDE-encoding RNA in oleic acid-grown cells as compared to glucose-grown cells, an RNA slot blot was done (Fig. 3.2.4.1). One and 4 ng of poly $(A)^+$ RNA from oleic acid- and glucose-grown cells were bound to nitrocellulose using a slot blot apparatus and hybridized to plasmid pBA labelled with ^{32}P by the random primer method. Densitometric analysis of the signals obtained from the autoradiogram showed an 8.1-fold increase in HDE-encoding mRNA levels in oleic

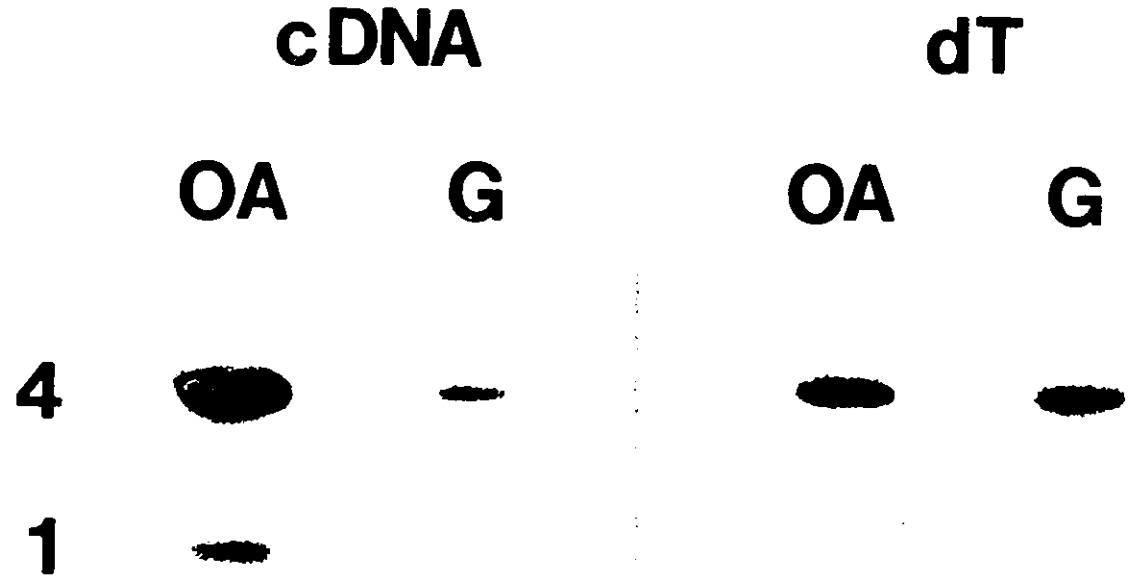


Figure 3.2.4.1 Slot blot analysis of HDE mRNA. One (lower row) and 4 (upper row) ng of poly $(A)^+$ RNA from oleic acid- (OA) or glucose-grown (G) cells was bound to nitrocellulose and hybridized with ^{32}P-labelled HDE cDNA (pBA) or oligo-dT.

acid-grown cells compared to glucose-grown cells. The signals were normalized by hybridization of ^{32}P-labelled oligo-dT_{18} to quantitate the amount of poly $(A)^+$ RNA bound to the nitrocellulose. The 8.1-fold induction in HDE-encoding mRNA levels is similar to the 9.3-fold increase found by immunoprecipitation of HDE from *in vitro* translated RNA from oleic acid-grown cells *versus* glucose-grown cells (Fig. 3.2.3.1).

The induction of mRNAs encoding HDE, catalase and the AOXs has been noted previously (Fujiki *et al.*, 1986). This induction has been used to clone the cDNAs encoding various forms of AOX and catalase (Rachubinski *et al.*, 1985; Okazaki *et al.*, 1986; Murray and Rachubinski, 1987a; Rachubinski *et al.*, 1987) by differential hybridization. Several other mRNAs have been identified which are induced upon growth of *C. tropicalis* on alkane- (Fujiki *et al.*, 1986) or oleic acid- (Kamiryo and Okazaki, 1984) containing medium compared to glucose-containing medium. Many of these mRNAs have been identified as encoding peroxisomal proteins (Kamiryo and Okazaki, 1984; Fujiki *et al.*, 1986). These results suggest that the increased levels of many peroxisomal proteins found in oleic acid- (or alkane-) grown cells can be attributed, at least in part, to increased RNA levels, and that the expression of genes encoding peroxisomal proteins is coordinately regulated. Possible mechanisms of this control of this increase in RNA levels are discussed in section 3.3.4.

3.2.5 Nucleotide sequence of the HDE cDNA[2]

The entire nucleotide sequence of the cDNA encoding HDE was determined by dideoxynucleotide sequencing. The sequencing strategy is shown in Fig. 3.2.5.1. Overlapping nested deletions were created with exonuclease III in pUC118 bearing the *Eco*RI cDNA fragments (pBA, pBB, pSA). Sequencing was performed using the M13 universal primer. Where independent sequence information could not be obtained using this primer, primers (AB366 and AB367) were synthesized and used to prime the sequencing reaction.

The junctions of the 5' and 3' *Eco*RI fragments were sequenced independently from a *Kpn*I fragment from the genomic clone which spans the internal *Eco*RI site (section 3.3; Fig. 3.3.3.1). The nucleotide sequence of this fragment ruled out the possibility of the loss of a small cDNA fragment within this region during the subcloning of the cDNA subfragments into pUC118. The complete nucleotide and derived amino acid sequences are shown in Fig. 3.2.5.2. There is a single open reading frame (ORF) of 2718 nucleotides encoding a protein of 906 amino acids (including the initiator methionine), using the proposed initiation codon at position +1 shown in Fig. 3.3.5.2. It was reasoned that this is the correct initiation codon, because the calculated MW of the protein using this codon is 99,481 Da, which is in close agreement with the M_r of HDE calculated by SDS-PAGE ($\approx$100,000; Fig. 3.2.3.1; Moreno de la Garza *et al.*, 1985; Fujiki *et al.*, 1986; Ueda *et al.*, 1987). Use of the next in-frame initiation codon

---

[2]HDE cDNA was sequenced with William Nuttley.

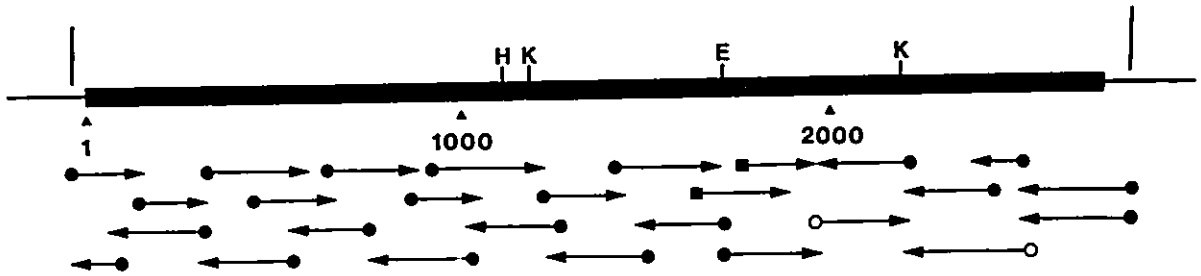


Fig. 3.2.5.1 Partial restriction map and sequencing strategy for the cDNA encoding HDE. The cDNA ORF is represented by the thick line. Vertical lines show the locations of *Eco*RI linkers flanking the cDNA insert. The restriction sites shown are: E, *Eco*RI; H, *Hind*III; K, *Kpn*I. Exonuclease III deletions were constructed in pBA, pBB, and pSA. The direction and extent of sequencing is indicated by arrows. Closed circles represent the starting points of dideoxy chain termination sequencing with the M13 forward primer. Open circles represent the starting points of sequencing with primers AB366 and AB367. The closed squares represent the starting points of sequencing with the M13 forward primer of a portion of a *Kpn*I-*Kpn*I fragment from pGEM/3T (harbouring the genomic cloning encoding HDE). The numbering is as presented in Fig. 3.2.5.2

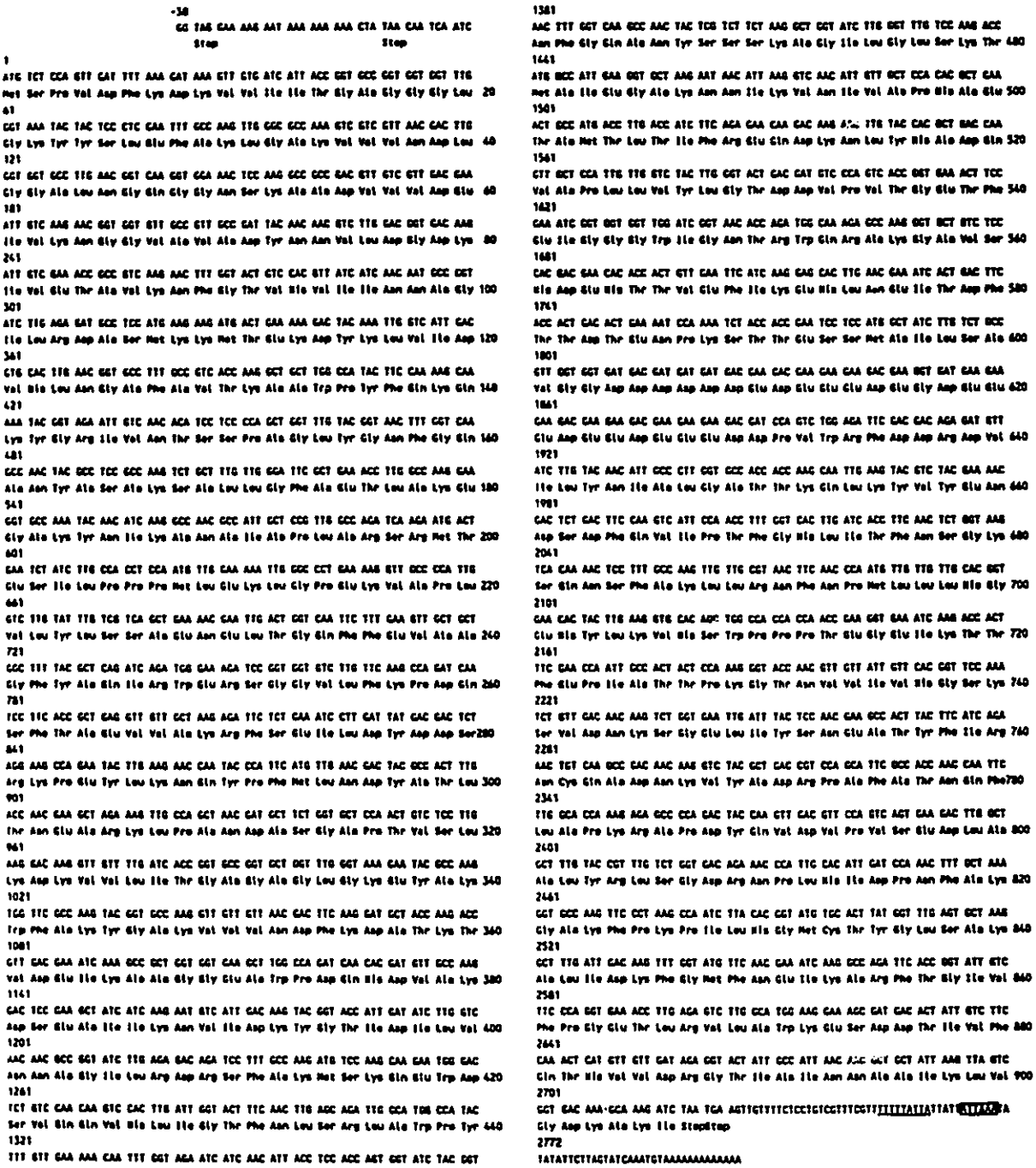


Figure 3.2.5.2 Nucleotide sequence and deduced amino acid sequence of the cDNA for HDE. Nucleotides are numbered at the left of the columns beginning with the first residue of the initiation codon. Nucleotides 5' to this position are designated by negative numbers. Amino acids are numbered at the right of the columns. The in-frame termination codons are indicated by 'Stop'. A putative polyadenylation consensus sequence is boxed. A putative consensus sequence for transcription termination is underlined.

downstream (nucleotides 319-321) would result in a protein of calculated MW of 91,958 Da. There are no ATG codons 5' to the proposed initiation codon, while there are two in-frame stop codons (TAG; nucleotides -36 to -34 and TAA; nucleotides -12 to -10). Upstream in-frame stop codons are a common feature of *C. tropicalis* genes encoding peroxisomal proteins. Stop codons are present upstream of the ORFs encoding catalase (CAT; Murray and Rachubinski, 1987b; Okada *et al.*, 1987), the three forms of fatty acyl-CoA oxidase (PXP-2 (Okazaki *et al.*, 1987), PXP-4(a) (Okazaki *et al.*, 1986; Murray and Rachubinski, 1987a) and PXP-5 (Okazaki *et al.*, 1986)), PXP-18 (Szabo, *et al.*, 1989) and isocitrate lyase (ICL; Atomi *et al.*, 1990).

In addition, the proposed initiation codon for the *HDE* gene lies in a consensus sequence (5'- A/Y A A/U A AUG U C U -3') for translational start as identified in the yeast *S. cerevisiae* (Cigan and Donahue, 1987), which is distinct from its mammalian counterpart (5'-GCCGCCA/GCCAUGG-3'; Kozak, 1986; 1989). The consensus sequence for yeast translation initiation is absent from the next in-frame initiation codon (nucleotides 319-321). The consensus sequence for yeast is found around the initiation codon in genes encoding other peroxisomal proteins of *C. tropicalis*:

	-4	-3	-2	-1	+1	+4	+5	+6
Consensus[3]	(A/Y)	A	(A/T)	A	ATG	T	C	T
HDE[4]	A	A	T	C	ATG	T	C	T
POX-2[5]	C	A	C	G	ATG	G	C	T
POX-4(2)[6]	C	A	T	A	ATG	A	C	T
POX-5[7]	T	A	T	C	ATG	C	C	T
POX-9[8]	C	A	A	T	ATG	G	C	T
POX-18[9]	A	A	A	C	ATG	T	C	A
ICL[10]	A	A	C	C	ATG	G	C	T

The underlined nucleotide in the consensus sequence (A, position -3) is proposed to be important for initiation by yeast ribosomes (Cigan and Donahue, 1987). Approximately 75% of yeast genes contain an A at position -3 relative to the initiation of translation. In addition, mutational studies indicate that nucleotide changes in this position decrease the efficiency of translation initiation (reviewed in Cigan and Donahue, 1987). The conservation of a C at position +5 and a T at position +6 suggests that these nucleotides may be important as well. Approximately 50% of yeast genes contain

---

[3]Cigan and Donahue (1987)

[4]Nuttley *et al.*, (1988); Aitchison *et al.*, (1991b)

[5]Okazaki *et al.*, (1987)

[6]Okazaki *et al.*, (1986); Murray and Rachubinski, (1987a)

[7]Okazaki *et al.*, (1986)

[8]Murray and Rachubinski (1987b); Okada *et al.*, (1987)

[9]Szabo *et al.*, (1989)

[10]Atomi *et al.*, 1990.

a C at position +5 relative to the initation site (Cigan and Donahue, 1987).

The 3' noncoding region of the HDE cDNA is 71 nucleotides long including a stretch of 13 A residues at the 3' terminus, which probably represents the poly A tail of the mRNA. There is a perfect match to the eukaryotic polyadenylation consensus sequence (5'-ATTAAA-3'; Birnstiel *et al.*, 1985) 23 nucleotides upstream of the poly A tail (nucleotides 2764-2969; boxed in Fig. 3.3.5.2). There is also a variant (5'-TTTTTATTA-3'; nucleotide 2750-2758) of the consensus signal for transcription termination proposed to operate in *S. cerevisiae* (5'-TTTTATA-3'; Henikoff and Cohen, 1984). The presence of this sequence 3' to the coding sequence of genes encoding other *C. tropicalis* peroxisomal proteins (catalase, Murray and Rachubinski, 1987b; *POX-2*, Okazaki *et al.*, 1987; *POX-4(a)*, Okazaki *et al.*, 1986, Murray and Rachubinski, 1987a; *POX-5*, Okazaki *et al.*, 1986; *POX-18*, Szabo, *et al.*, 1989; *ICL*, Atomi *et al.*, 1990), as well as the conservation of the polyadenylation consensus signal and the consensus sequence surrounding the initiation codon, suggest that both yeasts recognize similar genetic signals for the control of gene expression. This similarity has proven useful for the heterologous expression of *C. tropicalis* genes in other yeasts (see sections 3.4, 3.5, 3.6).

3.2.6 Codon frequencies in HDE cDNA

The derived amino acid sequence of HDE is shown below the cDNA nucleotide sequence in Fig. 3.2.5.2. As shown in Table 3.2.6.1, there is a definite

codon usage bias exhibited by this gene. This is a common feature of many highly expressed genes in yeast and bacteria and reflects the use of more prevalent tRNA isoacceptor species (reviewed in Sharp *et al.*, 1986). Table 3.2.6.1 compares the codon bias of *HDE* and other genes encoding peroxisomal proteins of *C. tropicalis*. It is evident that the genes encoding other peroxisomal proteins of *C. tropicalis* share the same 23 biased codons. Eighty-six percent of the codons encoded by the HDE cDNA represent the 23 preferentially used codons. This is comparable to the bias exhibited by other genes encoding peroxisomal proteins of *C. tropicalis*: *HDE* gene, 85.8%; *POX-4a*, 82.7%; *POX-4*, 86.6%; *POX-5*, 84.4%; *POX-2*, 85.2%; CAT cDNA, 85.2%; *CAT* gene, 85.2%; *ICL*, 88.9%; *POX-18*, 85%. Although the protein products of these genes are not present constitutively in large amounts, they are all co-induced upon transfer of the yeast to an oleic acid-containing medium (Okazaki *et al.*, 1986; 1987; Murray and Rachubinski, 1987a; 1987b; Nuttley *et al.*, 1988; Szabo *et al.*, 1989; Atomi *et al.*, 1990). The codon bias may permit the genes to be expressed at a high level in response to the medium change. Also, recent evidence suggests that, in addition to other factors, the use of highly expressed codons may also help the mRNAs to escape rapid degradation (Parker *et al.*, 1991). The preferred codons in *C. tropicalis* genes encoding peroxisomal proteins are identical to the 23 biased codons identified for the highly expressed genes of *S. cerevisiae* (Bennetzen *et al.*, 1982; Sharp *et al.*, 1986). This finding suggests a similar population of major tRNA isoacceptor species in the two yeasts and also points to a similarity in the control of gene expression in the two yeasts. In addition, the codon

Table 3.2.6.1
Codon usage in genes encoding peroxisomal proteins of *Candida tropicalis*

Amino acid	Codon[1]	Genes[2]									
		HDE cDNA	*HDE* gene	*POX-4a*	*POX-4*	*POX-5*	*POX-2*	CAT cDNA	*CAT* gene	*ICL*	*POX-18*
Ala	GCT	37	37	28	32	23	27	27	27	40	4
	GCC	47	47	23	28	27	20	13	13	23	10
	GCA	5	5	0	0	1	5	1	1	2	0
	GCG	0	0	0	0	1	0	0	0	0	0
Arg	CGT	3	3	0	0	5	5	0	0	1	0
	CGC	0	0	0	0	0	2	0	0	0	0
	CGA	0	0	0	0	0	0	0	0	0	0
	CGG	0	0	0	0	0	0	0	0	0	0
	AGA	23	23	36	34	29	36	20	20	22	1
	AGG	1	1	2	0	0	1	0	0	0	0
Asn	AAT	4	4	5	3	3	4	4	4	0	0
	AAC	48	48	27	28	26	27	25	25	18	9
Asp	GAT	22	23	12	11	5	2	5	5	7	3
	GAC	44	43	33	39	36	45	23	23	27	7
Cys	TGT	1	1	8	8	8	9	2	2	3	0
	TGC	1	1	0	0	1	0	1	1	0	0
Gln	CAA	23	23	29	28	18	23	20	20	22	4
	CAG	1	1	2	2	5	9	1	1	2	0
Glu	GAA	56	56	31	34	34	32	29	29	37	5
	GAG	2	2	4	3	9	11	0	0	2	1
Gly	GGT	68	68	43	43	46	46	22	24	33	7
	GGC	3	3	5	5	1	1	5	3	2	1
	GGA	2	2	0	0	2	1	2	2	1	1
	GGG	0	0	0	0	0	1	0	0	0	0
His	CAT	1	1	0	1	1	2	3	3	0	0
	CAC	16	16	10	9	12	13	19	19	15	0
Ile	ATT	25	25	20	19	17	26	5	6	10	1
	ATC	31	31	21	25	19	23	11	11	26	4
	ATA	0	0	2	1	1	0	0	0	0	0

[1]The 23 preferentially used codons are underlined.

[2]HDE cDNA, Nuttley *et al.*, 1988; *HDE* gene, Aitchison *et al.*, 1991b; *POX-4a*, Murray and Rachubinski, 1987a; *POX-4* and *POX-5*, Okazaki *et al.*, 1986; *POX-2*, Okazaki *et al.*, 1987; CAT cDNA, Murray and Rachubinski, 1987b; *CAT* gene, Okada *et al*, 1987; *ICL*, isocitrate lyase, Atomi *et al.*, 1990; *POX-18*, Szabo *et al.*, 1989.

Amino acid	Codon	Genes									
		HDE cDNA	*HDE* gene	*POX-4a*	*POX-4*	*POX-5*	*POX-2*	CAT cDNA	*CAT* gene	*ICL*	*POX-18*
Leu	CTT	2	2	1	1	3	2	0	0	0	0
	CTC	1	1	4	5	5	3	2	2	0	0
	CTA	0	0	1	0	0	0	0	0	0	0
	CTG	0	0	3	0	1	0	0	1	0	0
	TTA	2	2	2	0	1	2	0	0	0	2
	TTG	61	61	53	56	48	64	28	27	37	12
Lys	AAA	16	16	4	5	6	5	4	5	8	2
	AAG	55	55	38	40	37	46	24	23	38	20
Met	ATG	13	13	19	16	17	13	9	9	12	4
Phe	TTT	15	15	13	12	9	7	10	10	6	2
	TTC	27	27	16	16	24	18	22	22	14	3
Pro	CCT	3	3	7	5	3	2	8	8	1	0
	CCC	0	0	1	0	0	0	0	0	1	0
	CCA	33	33	21	18	23	27	24	23	17	4
	CCG	1	1	3	0	0	1	0	1	0	0
Ser	TCT	16	16	16	13	4	12	9	6	9	0
	TCC	20	20	26	28	25	24	13	15	19	1
	TCA	3	3	4	3	1	1	1	2	0	1
	TCG	2	2	2	1	3	2	0	0	0	0
	AGT	3	3	3	3	4	1	1	1	1	0
	AGC	3	3	3	3	3	0	1	0	0	1
Thr	ACT	22	22	15	15	8	21	11	10	9	0
	ACC	31	31	23	26	32	20	22	24	23	3
	ACA	1	1	5	3	0	2	1	1	1	0
	ACG	0	0	2	0	0	1	0	0	0	0
Tyr	TAT	3	3	2	2	3	8	4	3	1	0
	TAC	30	30	24	26	22	22	14	15	21	0
Trp	TGG	11	11	10	8	9	10	8	8	12	1
Val	GTT	33	33	19	23	16	23	11	10	13	3
	GTC	32	32	26	26	25	13	19	19	13	10
	GTA	0	0	1	1	0	0	0	0	0	0
	GTG	3	3	3	1	1	3	1	1	1	0
Sum aa[3]		906	906	709	709	662	724	485	485	550	127
Mf codons[4]		779	778	586	614	559	617	413	413	489	108
% Mf codons[5]		86.0	85.9	82.6	86.6	84.4	85.2	85.2	85.2	88.9	85.0

[3]Sum of amino acid residues.

[4]Number of amino acids encoded by the 23 most frequent codons.

[5]Percentage of the number of amino acids encoded by the most frequent codons.

biases in the above sequences are similar, but not identical, to the codon biases of highly expressed genes in *E. coli*. In particular, the codons preferred for Leu, Cys, Gln, Arg, Lys and Pro in *E. coli* are different from those of the genes discussed above (Sharp *et al.*, 1986).

3.2.7 Analysis of the HDE amino acid sequence of HDE

3.2.7.1 Hydropathy profile

The hydropathy profile for HDE was analyzed using the method of Kyte and Doolittle (1982) and is shown in Fig. 3.2.7.1.1. The protein shows an overall hydrophilic nature. One area of extreme hydrophilicity is found between amino acids 604 and 630. In this stretch of 27 amino acids, there are 26 acidic residues separated by a single Gly residue at amino acid position 617. Use of the FAST SCAN program of PCGENE to scan the Swiss-Prot data base for proteins showing similarity to HDE revealed a number of proteins with similar acidic stretches. The relationships of these proteins to HDE and the function of the acidic region of HDE are unknown (but see below, section 3.2.8.2).

Overall the HDE protein contains 124 (13.7%) acidic (Asp and Glu), 115 (12.7%) basic (Arg, Lys, His), 86 (9.5%) aromatic (Phe, Trp, Tyr) and 289 (31.9%) hydrophobic (aromatic, Leu, Ile, Met, Val) residues. The calculated pI of the protein is 5.10.

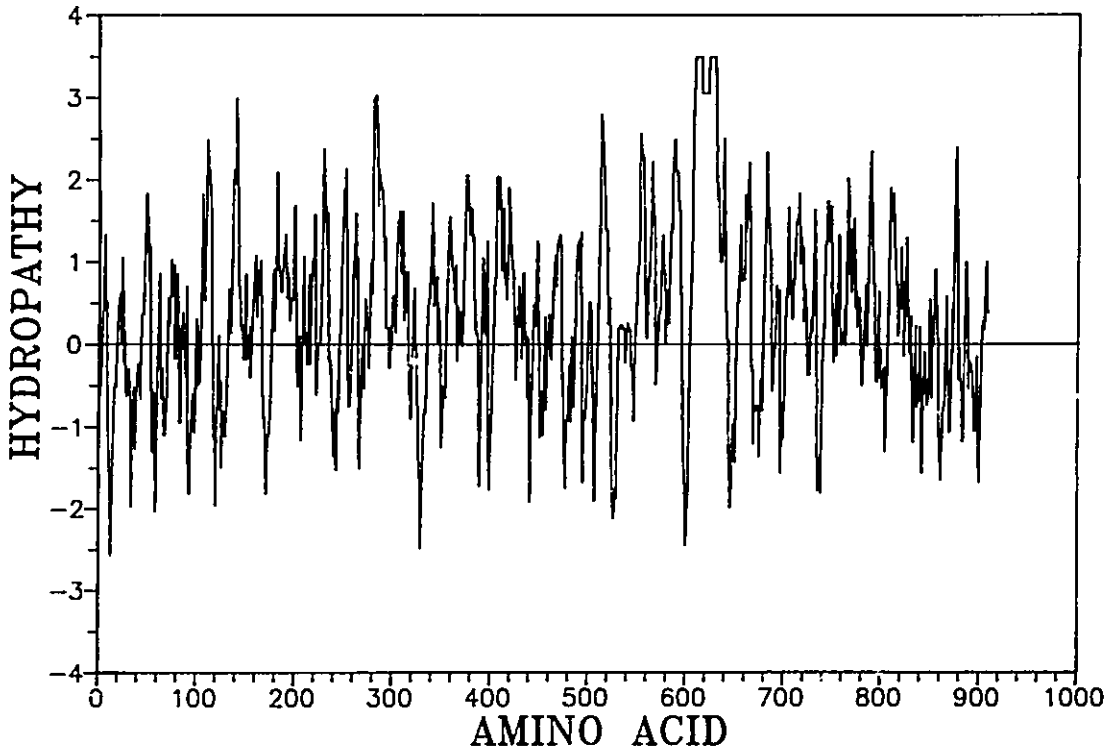


Figure 3.2.7.1.1 Hydropathy profile of HDE. Hydrophilicity is designated by positive values. Amino acids are numbered as in Fig. 3.2.5.2. Analysis was done by the method of Kyte and Doolittle (1982) using a window of 7 amino acids.

3.2.7.2 Repeat region

The HDE cDNA (and gene; see section 3.3) contains a tandem repeat region which appears to have arisen during the evolution of the *HDE* gene (Fig. 3.2.7.2.1). Nucleotides 16 to 856 show a 60.0% identity with a second area of the gene located between nucleotides 958 to 1767. Alignment of these sequences by the MICROGENIE ALIGN program yielded 506 matches, 301 mismatches and 37 unmatched nucleotides. There is also a strong amino acid identity encoded by these repeat regions (Fig.3.2.7.2.2). Alignment of the amino acids encoded by the nucleotide repeat regions (amino acids 6 to 265 and 319 to 588) showed 41.7% amino acid identity. The similarity between these regions becomes 56.5% when conservative substitutions are permitted. Shorter stretches of amino acids, in particular, between amino acids 7 to 40 and 320 to 353 and between amino acids 153 to 172 and 456 to 475 show even greater conservation, 78.7 and 70.0%, respectively.

Analysis of the HDE amino acid sequence using the sites and signatures program of PCGENE detected two insect-type dehydrogenase/ribitol dehydrogenase family signatures at amino acids 163 to 173 (YASAKSALLGF) and 467 to 477 (YSSSKAGILGL). These sites are located in each of the duplicated regions. This suggests that one, or both, of these regions of the protein contain the dehydrogenase activity of the trifunctional enzyme. In agreement with this suggestion, Baker (1990) found that each of the duplicated regions shows similarity to a family of oxidoreductases involved in metabolizing sugars, alcohol and toluene and in synthesizing antibiotics and

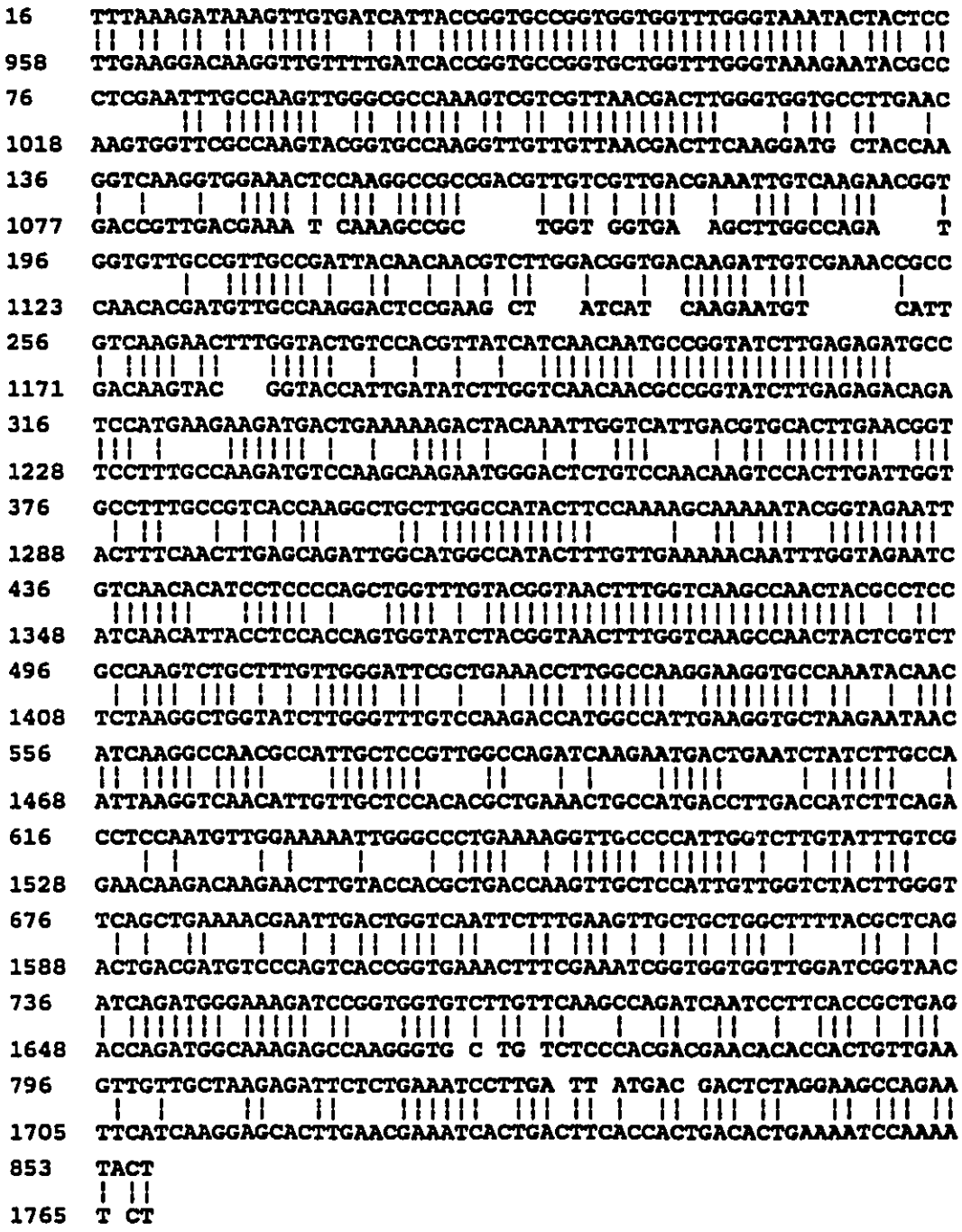


Figure 3.2.7.2.1 Nucleotide homology between the repeated regions of the HDE cDNA. Numbering is as in Fig. 3.3.5.2. Vertical lines indicate nucleotide identity. 60% of the nucleotides are matched (506 matched, 301 mismatched and 37 unmatched nucleotides over the 844 nucleotide length).

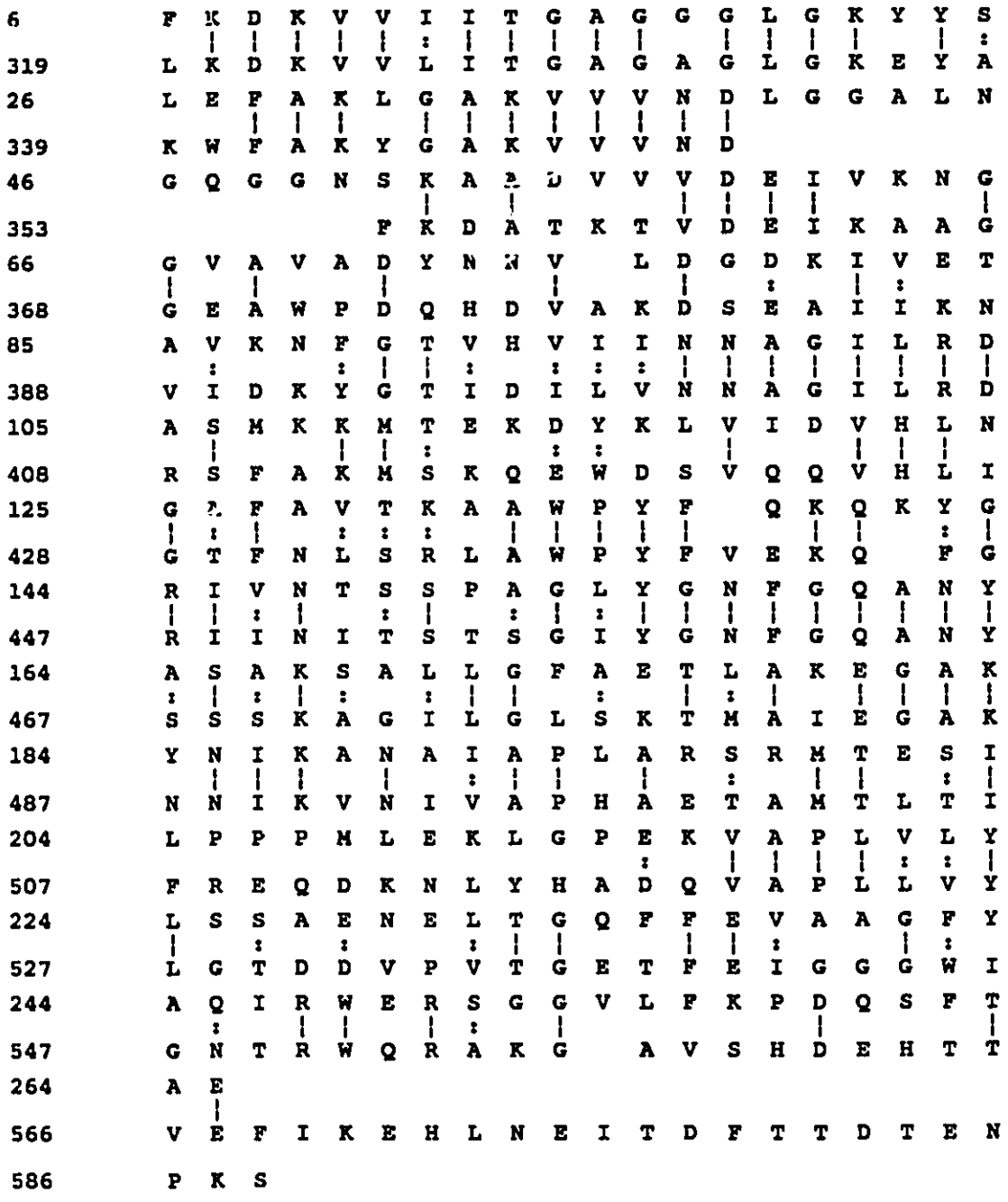


Figure 3.2.7.2.2 Amino acid homology between the duplicated regions of HDE. Numbering is as in Fig. 3.2.5.2. Vertical lines represent amino acid identity. Dots represent conservative amino acid substitutions. 41.7% of the amino acids are matched (118 matched, 129 mismatched and 36 unmatched amino acids over the 283 amino acid length).

steroids. In particular, Baker (1990) found the duplicated regions show a very high similarity to amino acids 4 to 224 of *Bacillus megaterium* glucose dehydrogenase. Acetacetyl-CoA reductase also shows amino acid similarity to the repeat regions of HDE (FAST SCAN program of PCGENE; SWISS-PROT 20, Feb, 1992). The similarities are shown in Appendix 1.

Together the tandemly duplicated domains of HDE, plus a short 90 amino acid linker, comprise 560 of the 906 amino acids, or 65%, of the protein. The remaining 35% of the protein is unlikely to encode separate hydratase and epimerase domains of the enzyme (Baker, 1990). Search of the SWISS-PROT (version 20, Feb. '92) database using the FAST SCAN program of PCGENE failed to detect any proteins with significant similarity to the carboxy-terminal region of the protein outside the repeated domains. It is possible that one of the duplicated domains has acquired, through mutation, either the hydratase or the epimerase activity following the gene duplication event. Alternatively, the repeated regions may contain dehydrogenase (or hydratase) activities with opposite stereospecificities, which may account for the observed 3-hydroxyacyl-CoA epimerse activity of HDE. As discussed in section 1.2.1, two rat liver hydratases with different specificities are responsible for the observed 3-hydroxyacyl-CoA epimerase activity of rat liver peroxisomes (see below, section 3.2.8.1).

3.2.8 Comparison of *C. tropicalis* HDE to its homologues

3.2.8.1 Mammalian bifunctional enzyme

Comparison of HDE to its mammalian peroxisomal counterpart shows that the enzymes are quite distinct. In mammalian peroxisomes, the second and third steps (2-enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase) of the β-oxidation of fatty acids are carried out by a single bifunctional enzyme (section 1.2.1), which lacks the epimerase activity of *C. tropicalis* HDE (Moreno de la Garza, 1985). Recently, it has been shown that rat hydratase-dehydrogenase (HD) actually possesses a third activity not presented in HDE, Δ^3,Δ^2-enoyl-CoA isomerase (Palosarri and Hilutnen, 1990). In addition, the 3-hydroxyacyl-CoA epimerase activity observed in rat liver peroxisomes is due to two 2-enoyl-CoA hydratases with opposite stereospecificities (Hiltunen *et al.*, 1989; Smeland *et al.*, 1989; see section 1.2.1).

As well as possessing different enzymatic activities, *C. tropicalis* HDE shows no primary amino acid sequence similarity to HD. These facts, combined with the inclusion of HDE in the oxidoreductase superfamily, indicate that HDE and HD probably arose independently (Baker, 1990).

3.2.8.2 *S. cerevisiae* and *Neurospora crassa* trifunctional enzymes

Although HDE bears little similarity to its mammalian counterpart HD, comparison of *C. tropicalis* HDE to *S. cerevisiae* and *Neurospora crassa* trifunctional

Figure 3.2.8.2.1 Amino acid alignment of *C. tropicalis*, *S. cerevisiae* and *N. crassa* trifunctional enzymes. Identical amino acids are in upper case. Analysis was done using the MICROGENIE align program.

```
C. tropicalis   1  M     s  p  v  d  F  K  D  k  V  V  i  I  T  G  A  G  G  G
S. cerevisiae   1  M  p  g  n  L  s  F  K  D  r  V  V  V  I  T  G  A  G  G  G
N. crassa       1  M  a  e  q  L  r  F  d  g  q  V  V  V  v  T  G  A  G  G  G

C. tropicalis  20  L  G  K  y  Y  s  L  e  F  A  k  l  G  A  K  V  V  V  N  D
S. cerevisiae  21  L  G  K  v  Y  a  L  a  y  A  S  R  G  A  K  V  V  V  N  D
N. crassa      21  L  G  K  a  Y  c  L  f  F  g  S  R  G  A  s  V  V  V  N  D

C. tropicalis  40  L  G  G  a  L  n  G  q  G  g  N  S     K  A  A  D  V  V  V
S. cerevisiae  41  L  G  G  t  L  g  G  s  G  h  N  S     K  A  A  D  l  V  V
N. crassa      41  L  G  a  s  f  k  G  e  G     N  S  t  K  A  A  D  V  V  V

C. tropicalis  59  D  E  I  v  K  n  G  G  v  A  V  A  d  Y  n  n  V     l  d
S. cerevisiae  60  D  E  I  K  K  A  G  G  i  A  V  A  N  Y  D  S  V  n  E  N
N. crassa      60  n  E  I  K  a  A  G  G  k  A  V  A  N  Y  D  S  V     E  N

C. tropicalis  78  G  D  K  I  v  E  T  A  v  K  n  F  G  t  V  h  V  i  I  N
S. cerevisiae  80  G  e  K  I  I  E  T  A  I  K  E  F  G  R  V  D  V  L  I  N
N. crassa      79  G  D  K  I  I  E  T  A  I  K  E  F  G  R  i  D  i  L  I  N

C. tropicalis  98  N  A  G  I  L  R  D  a  S  m  K  K  M  T  E  k  D  y  k  L
S. cerevisiae 100  N  A  G  I  L  R  D  v  S  F  a  K  M  T  E  r  e  f  a  s
N. crassa      99  N  A  G  I  L  R  D  i  S  F  K  n  M  k  d  e  D  w  d  L

C. tropicalis 118  V  i  D  V  H  L  n  G  a  f  a  v  t  k  A  A  W  P  Y  F
S. cerevisiae 120  V  v  D  V  H  L  t  G  g  Y  K  l  s  R  A  A  W  P  Y  m
N. crassa      119  i  f  k  V  H  v  k  G  a  Y  K  t  a  R  A  A  W  P  Y  F

C. tropicalis 138  q  K  Q  K  y  G  R  I  v  N  T  s  S  P  A  G  L  y  G  N
S. cerevisiae 140  R  s  Q  K  F  G  R  I  I  N  T  A  S  P  A  G  L  F  G  N
N. crassa      139  R  K  Q  K  F  G  R  v  I  N  T  A  S  a  A  G  L  F  G  N

C. tropicalis 158  F  G  Q  A  N  Y  a  s  A  K  s  a  L  l  G  F  A  E  T  L
S. cerevisiae 160  F  G  Q  A  N  Y  S  A  A  K  m  G  L  V  G  l  A  E  T  L
N. crassa      159  F  G  Q  A  N  Y  S  A  A  K  l  G  m  V  G  F  t  E  T  L

C. tropicalis 178  A  K  E  G  A  K  Y  N  I  k  a  N  a  I  A  P  L  A  R  S
S. cerevisiae 180  A  K  E  G  A  K  Y  N  I  n  v  N  s  I  A  P  L  A  R  S
N. crassa      179  A  K  E  G  l  K  Y  N  I  i  s  N  v  I  A  P  i  A  a  S

C. tropicalis 198  R  N  T  E  s  i  L  P  P  p  m  L  e  k  L  G  P  E  K  V
S. cerevisiae 200  R  N  T  E  n  V  L  P  P  h  i  L  k  q  L  G  P  E  K  i
N. crassa      199  R  N  T  E  t  V  m  P  P  d  l  L  a  l  m  k  P  E  w  V

C. tropicalis 218  a  P  L  V  L  Y  L     s  s  a  e  n  e  l  T  G  q  f  F
S. cerevisiae 220  V  P  L  V  L  Y  L     t  h  e  s  T  k  v  s  n  S  I  F
N. crassa      219  V  P  L  V  a  v  L  v  h  k  n  n  T  s  e  T  G  S  I  F

C. tropicalis 237  E  V  A  A  G  F  y  A  Q  i  R  W  E  R  S  g  G  v  L  F
S. cerevisiae 239  E  l  A  A  G  F  f  g  Q  L  R  W  E  R  S  S  G  q  i  F
N. crassa      239  E  V  g  g  G  h  v  A  k  L  R  W  E  R  S  S  G  l  L  l

C. tropicalis 257  K  P  D     q  S  f  T  a  E  v  v  a  k  r  f  s  E  I  l
S. cerevisiae 259  n  P  D  p  k  t  Y  T  p  E  a  i  l  n  k  w  K  E  I  t
N. crassa      259  K  a  D     e  S  Y  T        p  g  a  i  i  K  k  w  d

C. tropicalis 276  D  Y  d  D  s  r  K  P  e  y  l  K  n  Q  Y  P  f  m  L  N
S. cerevisiae 279  D  Y  r  D        K  P     f  n  K  t  Q  h  P  y  q  L  s
N. crassa      275  q  v  t  D  f  s  n  P        q  Y  P  t  g  p  N

C. tropicalis 296  D  Y  a  t  L  t  n  E  A  r  K  L  P  a  N  d  a  s  g  a
S. cerevisiae 296  D  Y  n  d  L  i  t  k  A  k  K  L  P  P  N  e  q  g  a  v
N. crassa      290  D  f  l  a  L  l  e  E  s  l  K  L  g  P  N        d  p  g

C. tropicalis 316  p  t  V  S  L  K  d  K  V  V  L  i  T  G  A  G  A  G  L  G
S. cerevisiae 316  k  i  k  S  L  c  n  K  V  V  v  V  T  G  A  G  g  G  L  G
N. crassa      308  e  k  V  d  f  K  g  r  V  a  L  V  T  G  g  G  A  G  i  G

C. tropicalis 336  K  e  Y  A  k  W  F  A  k  Y  G  A  K  V  V  V  N  D  f  K
S. cerevisiae 336  K  s  h  A  i  W  F  A  R  Y  G  A  K  V  V  V  N  D  i  K
N. crassa      328  r  a  Y  c  l  a  F  A  R  a  G  A  s  V  V  V  N  D  l  v

C. tropicalis 356  D  a  t  k  t  V  d  E  I     K  a  a  G     G  e  A  W  P
S. cerevisiae 356  D  P  f  s  V  V  e  E  I  n  K  l  y  G  e  G  t  A  I  P
N. crassa      348  n  P  d  d  V  V  n  E  I     K  k  m  G     G  k  A     v
```

C. tropicalis	374	D	q	H	D	V	A	k	D	s	e	A	I	I	K	n	v	I	D	K	y
S. cerevisiae	376	D	s	H	D	V	v	t	e	s	p	l	I	I	q	t	A	I	s	K	F
N. crassa	365	g	s	k	f	s	A	e	D	g	d	A	v	v	K	s	A	I	D	s	F
C. tropicalis	394	G	t	I	D	I	L	V	N	N	A	G	I	L	R	D	r	S	F	e	K
S. cerevisiae	396	q	R	V	D	I	L	V	N	N	A	G	I	L	R	D	K	S	F	l	K
N. crassa	385	G	R	V	D	I	v	V	N	N	A	G	I	L	R	D	K	a	F	h	n
C. tropicalis	414	N	s	k	q	E	W	D	s	V	q	q	V	H	L	I	G	T	F	n	L
S. cerevisiae	416	N	k	D	e	E	W	f	a	V	l	k	V	H	L	f	s	T	F	s	L
N. crassa	405	N	d	D	s	l	W	D	p	V	m	n	V	H	a	r	G	T	y	k	v
C. tropicalis	434	S	r	l	A	W	P	Y	F	v	e	K	Q		f	G	R	I	I	N	i
S. cerevisiae	436	S	K	A	v	W	P	i	F		t	K	Q	K	s	G	f	I	I	N	T
N. crassa	425	t	K	A	A	W	P	Y	F		l	K	Q	K	y	G	R	v	i	N	T
C. tropicalis	453	T	S	T	S	G	I	Y	G	N	F	G	Q	A	N	Y	S	s	s	K	A
S. cerevisiae	455	T	S	T	S	G	I	Y	G	N	F	G	Q	A	N	Y	a	A	A	K	A
N. crassa	444	T	S	T	S	G	I	Y	G	N	F	G	Q	A	N	Y	S	A	A	K	c
C. tropicalis	473	g	I	L	G	l	S	K	T	m	A	i	E	G	A	K	n	N	I	k	V
S. cerevisiae	475	A	I	L	G	F	S	K	T	I	A	L	E	G	A	K	r	g	I	i	V
N. crassa	464	A	I	L	G	F	S	r	a	I	A	L	E	G	A	K	y	N	I	y	V
C. tropicalis	493	N	i	v	A	P	H	A	E	T	A	M	T	l	T	I	F	r	E		q
S. cerevisiae	495	N	v	I	A	P	H	A	E	T	A	M	T	K	T	I	F	s	E	k	E
N. crassa	484	N	t	I	A	P	n	A	g	T	A	M	T	K	T	I	l	p	E		E
C. tropicalis	512	d	k	N	l	y	h	A	D	Q	V	A	P	L	l	V	y	L			
S. cerevisiae	515	L	s	N	h	F	d	A	s	Q	V	s	P	L	v	V	l	L	s	s	e
N. crassa	503	L	v	q	s	F	k	p	D	y	V	A	P	L		V		L			s
C. tropicalis	529				g	t	d	d	v	p	V	T	G	e	t	F	E	i	G	G	G
S. cerevisiae	535	e	l	q	k	y	s	g	r	r	V	i	G	q	L	F	E	V	G	G	G
N. crassa	519	l	c	s	d	k	v	p	k	k	p	T	G	g	L	y	E	V	G	s	G
C. tropicalis	546	W	i	G	n	T	R	W	Q	R	s	k	G		s	V	S	h	D	E	h
S. cerevisiae	555	W	C	G	Q	T	R	W	Q	R	S	s	G		y	V	S	i	k	E	t
N. crassa	539	W	C	G	Q	T	R	W	Q	R	S	g	G	h	g	f	p	v	D	v	p
C. tropicalis	565	t	T	v	E	f	I	K	E	N	l	N	e	I	T	D	F	t	t	d	T
S. cerevisiae	574	i	e	P	E	e	I	K	E	n	W	N	h	I	T	D	F	s	r	n	T
N. crassa	559	l	T	P	E	e	v	v	k	H	W	N	d	I	v	t	F	d	s	r	s
C. tropicalis	585	e	N	P			k	S		T	t	E	S	S	M	A	i	L	s	a	v
S. cerevisiae	594	i	N	P			s	S		T	e	E	S	S	M	A	t	L	q	s	v
N. crassa	579	d	h	P	e	k	s	S	d	s	i	E	k	i	M	A			n	m	e
C. tropicalis	602	g	g	d	d	d	d	d	d	e	d	e	e	e	d	e	g	d	e	e	e
S. cerevisiae	611																			q	K
N. crassa	597	n	r	v	g	e	g	k	s	g	a	a	e	n	e	h	l	a	a	i	K
C. tropicalis	622	d	e	e	d	e	E	e	D	D	p	v	w	r	F	d	d	R	D	V	I
S. cerevisiae	613	a	h	s	s	k	E	l	D	D	g	l	f	k	y	T	t	k	D	c	I
N. crassa	617	k	f	t	g	v	E	g	k	g	t	e	y	t	F	T	e	R	D	V	c
C. tropicalis	642	L	Y	N	i	a	L	G	A	T	t	K	q	L	K	Y	v	Y	E	N	D
S. cerevisiae	633	L	Y	N	L	G	L	G	c	T	s	K	e	L	K	Y	t	Y	E	N	D
N. crassa	637	L	.	N	L	G	i	G	A	k	r	t	d	i	K	Y	i	f	E	g	n
C. tropicalis	662	s	D	F	Q	V	i	P	T	F	g	h	l	I	t	F	n	s	g	k	s
S. cerevisiae	653	p	D	F	Q	V	l	P	T	F		a	V	I	P	F	m	q	a	t	a
N. crassa	657	e	D	F	e	V	v	P	T	F		g	V	I	P	p	f	n	t	e	m
C. tropicalis	682	q	n	S	F	a	k	L	l	r	N	F	N	P	M	i	L	L	H	G	E
S. cerevisiae	672	t	l	a	m	D	n	L	V	d	N	F	N	y	a	M	L	L	H	G	E
N. crassa	676	p	f	S	F	D	d	i	V	p	N	F	s	P	M	M	L	L	H	G	E
C. tropicalis	702	h	Y	L	K	V	h	s	w	P	p	P	T	e	G	e	i	K	T	t	f
S. cerevisiae	692	Q	Y	f	K	l	c	t	p	t	m	P	s	n	G	t	L	K	T	l	a
N. crassa	696	Q	Y	L	e	V	r	k	y	P	i	P	T	s	G	r	L	v	s	k	g
C. tropicalis	722	e	P	i	a	t	t	p	K		G	t	n	v	v	i	V	h	G	s	k
S. cerevisiae	712	K	P	L	q	V	l	D	K	n	G	k	A	A	l	V	V	g	G	f	e
N. crassa	716	K	l	L	e	V	v	D	K		G	s	A	A	i	V	k	q	G	i	t

C. tropicalis	741	s	v	D	n	K	s	G	E	L	I	y	s	N	E	a	T	y	F	I	R
S. cerevisiae	732	T	y	D	i	K	t	k	k	L	I	a	Y	N	E	g	s	f	F	I	R
N. crassa	735	T	f	n	a	e	t	G	E	e	l	f	Y	N	E	m	T	v	F	l	R
C. tropicalis	761			n	c	q	a	d	n	K	v	y	A	D			R	p	a	f	A
S. cerevisiae	752		g	a	h	v	p	p	e	K	e	v	r	D	g	k	R	a	k	F	A
N. crassa	755	g	c	g	g	f	g	g	q	K	k	p	A	D			R			g	A
C. tropicalis	777	t	n	q	F	l	a	P		k	R	a	P	D	y	q	V	d	V	p	v
S. cerevisiae	771	v	q	n	F	e	v	P	h	g	k	v	P	D	f	e	a	E	i	s	T
N. crassa	771	s	t	a	a	n	k	P	p	a	R	s	P	D	a	v	V	E	V	q	T
C. tropicalis	796	s	E	D	l	A	A	L	Y	R	L	S	G	D	r	N	P	L	H	I	D
S. cerevisiae	791	n	k	D	Q	A	A	L	Y	R	L	S	G	D	f	N	P	L	H	I	D
N. crassa	791	t	E	e	Q	A	A	i	Y	R	L	S	G	D	y	N	P	L	H	v	D
C. tropicalis	816	P	n	F	A	K	g	a	K	F	P	k	P	I	L	H	G	m	C	T	y
S. cerevisiae	811	P	t	l	A	K	a	v	K	F	P	t	P	I	L	H	G	L	C	T	l
N. crassa	811	P	a	F	A	K	v	g	g	F	k	v	P	I	L	H	G	L	C	s	f
C. tropicalis	836	G	l	S	A	K	A	L	i	d	K	f	G	m	F	n	E	I	K	a	R
S. cerevisiae	831	G	I	S	A	K	A	L	f	E	h	Y	G	p	y	e	E	l	K	V	R
N. crassa	831	G	I	a	g	K	A	v	y	E	K	Y	G	k	F	k	n	I	K	V	R
C. tropicalis	856	F	T	G	i	V	F	P	G	e	T	L	r	V	l	A	W	K	e	s	d
S. cerevisiae	851	F	T	n	v	V	F	P	G	d	T	L	k	V	k	A	W	K		q	G
N. crassa	851	F	a	G	t	V	n	P	G	q	T	L	v	t	e	m	W	K		e	G
C. tropicalis	876	d	t	i	V	F	Q	T	h	V	v	d	R	G	t	I	A	I	n	n	A
S. cerevisiae	870	s	v	V	V	F	Q	T	i	d	t	t	R	n	v	I	v	l	d	n	A
N. crassa	870	n	k	V	V	F	Q	T	k	V	k	e	t	G	k	l	A	I	s	g	A
C. tropicalis	896	A	i	K	L	v	g	d	K	a	K	i									
S. cerevisiae	890	A	v	K	L	s	q	a	K	s	K	l									
N. crassa	890	A	a	e	L	a															

enzymes reveals striking similarity[11]. An alignment of the three amino acid sequences is shown in Fig. 3.2.8.2.1. Alignment shows a similarity in the sequences, especially at the amino termini. There is 34.7% triple identity amongst the sequences. Alignments of *C. tropicalis* HDE with *S. cerevisiae* and *N. crassa* trifunctional enzymes showed 53.0 and 46.9% amino acid identity, respectively, over the entire lengths of the proteins (Appendix 2). Both the *S. cerevisiae* and *N. crassa* trifunctional enzymes contain the duplicated regions with homology to the oxidoreductase superfamily (Baker, 1990). Each region contains the insect-type alcohol dehydrogenase/ribitol dehydrogenase signature as detected using PCGENE (amino acids 165-175 and 469-479 in *S. cerevisiae* trifunctional enzyme, and amino acids 164-174 and 458-468 in *N. crassa* trifunctional enzyme). The duplicated regions show strong identity amongst the three enzymes: amino acids 6 to 224 and 397 to 467 in *C. tropicalis* HDE show 52.9% and 50.6% triple identity, respectively. This suggests that the repeat is not just a relic of a non-functional gene duplication which occurred in the evolution of the enzyme, but that each domain of the repeat serves a functional, indispensible role in the activity of the enzyme. The acidic amino acid stretch in *C. tropicalis* HDE (amino acids 604-630) is not present in either the *S. cerevisiae* or *N. crassa* counterparts. This suggests that the acidic nature of this stretch is not essential for the activity of *C. tropicalis* HDE. In addition, there is little similarity in this region between the proteins of either *S. cerevisiae* or *N. crassa*,

---

[11]The amino acid sequences for *S. cerevisiae* and *N. crassa* trifunctional enzymes were kindly provided by W.-H. Kunau (personal communication)

suggesting that this stretch of amino acids may act as a spacer between domains containing different activities of the trifunctional enzyme.

3.3 Cloning and sequencing the *HDE* gene

3.3.1 Restriction analysis of the *HDE* gene 5' and 3' flanking sequences

Prior to cloning the *HDE* gene, it was important to determine the restriction sites flanking the *HDE* gene in the *C. tropicalis* genome. Total *C. tropicalis* DNA was digested with various restriction enzymes and analyzed by Southern blotting using either the 5' (BA) or 3' (SA) cDNA fragments as probes (Fig. 3.3.1.1). Analysis of the approximate sizes of each restriction fragment was used to generate the partial restriction map shown in Fig. 3.3.1.2. The position of the *Sal*I sites (approximately -700 and +3600, relative to the initiation codon) allowed the use of *Sal*I for the digestion of λEMBL3/HDE clones 3a and 5b (see section 3.3.2), because these sites do not disrupt the gene or the immediate flanking regions.

3.3.2 Cloning the *HDE* gene

A *C. tropicalis* λEMBL3 genomic DNA library was screened using the ^{32}P-labelled HDE cDNA *Eco*RI fragment from pBA. Approximately 30,000 plaques were screened, and 5 potential positives were identified. Clones 3a and 5b were amplified and purified to homogeneity. The DNA from clones 3a and 5b was isolated and digested with *Sal*I. The agarose gel and Southern blot analysis of these digests are shown in Fig.

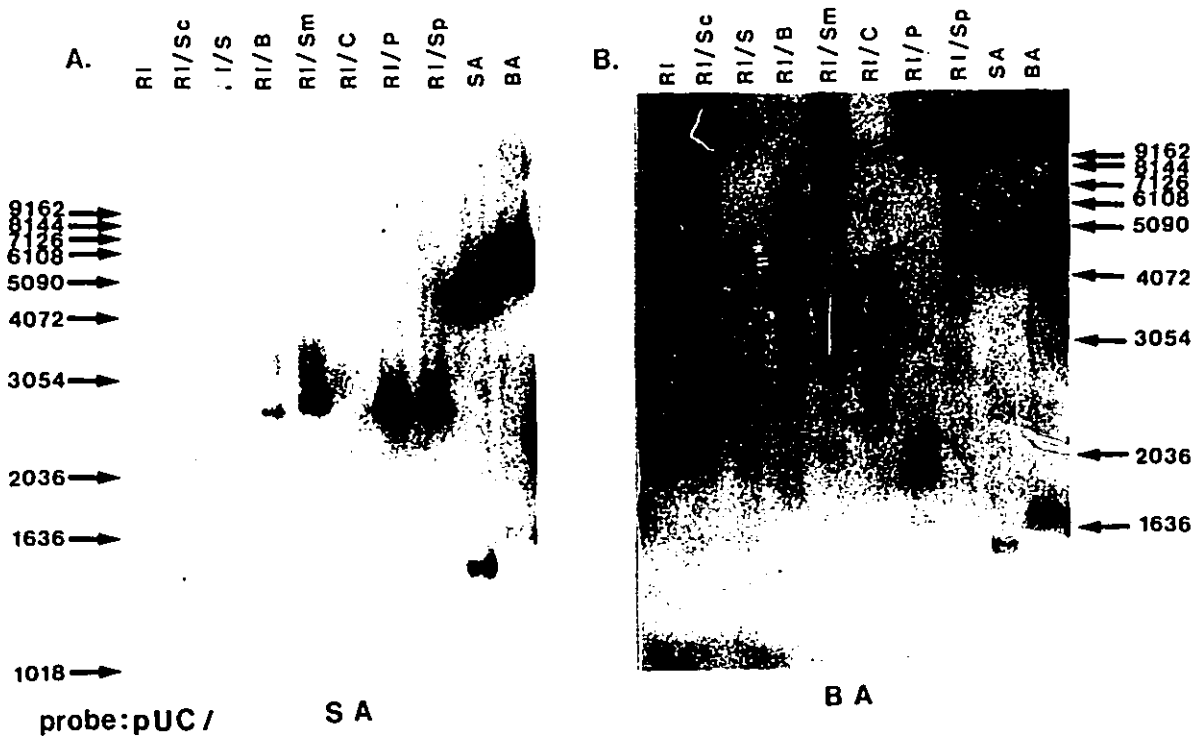


Figure 3.3.1.1 Southern analysis of sequences flanking the HDE gene. Eight μg of *C. tropicalis* genomic DNA was digested with 20 U of each restriction enzyme for 5 h. 4 μg of DNA from each digestion was separated on a 0.8% agarose gel by electrophoresis, blotted to nitrocellulose, and probed with plasmid pUC118 containing the 3' end of the HDE cDNA (pSA; Panel A) or the 5' end of the HDE cDNA (pBA; Panel B), radiolabelled with [α-^{32}P]-dATP. Lanes SA and BA each contain 25 ng of control plasmids pSA and pBA linearized with *Eco*RI. Total genomic DNA was digested with *Eco*RI alone (lanes RI) or in conjunction with *Sac*I (lanes RI/Sc), *Sal*I (lanes RI/S), *Bam*HI (lanes RI/B), *Sma*I (lanes RI/Sm), *Cla*I (lanes RI/C), *Pst*I (lanes RI/P) or *Sph*I (lanes RI/Sp). The migrations of the 1 kb ladder fragments are shown at the side of each panel.

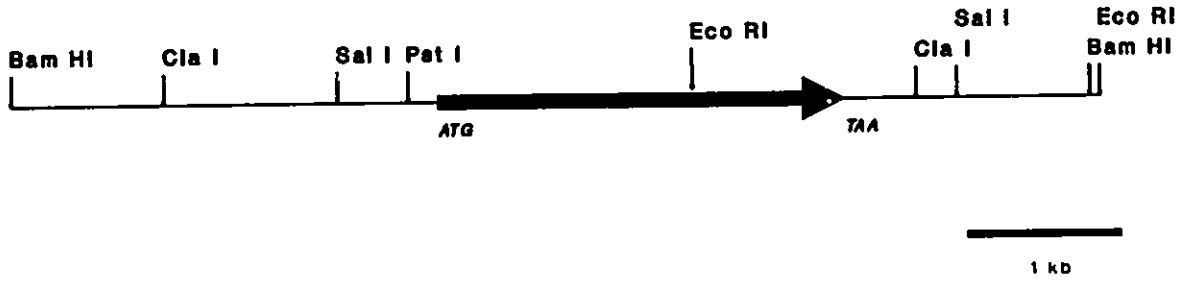


Figure 3.3.1.2 Partial restriction map of DNA flanking the *HDE* gene. Analysis of the approximate sizes of fragments generated by restriction digestion of total genomic *C. tropicalis* DNA (Fig. 3.3.1.1) was used to generate the restriction map shown. The *HDE* gene ORF is indicated by the arrow. Locations of restriction sites are approximate.

3.3.2.1. The fragments which were determined to carry the *HDE* gene are identified by the arrows 3T and 5BT. Further separation of these fragments showed that clone λEMBL3/HDE 3a contained two closely migrating fragments (at ≈ 4 kbp). The upper fragment (3T; ≈ 4.2 kbp) was determined to contain the *HDE* gene. There were 3 closely migrating fragments of ≈ 4 kbp in clone λEMBL3/HDE 5b. Further analysis demonstrated that the intermediate fragment (5BT; ≈ 4.0 kbp) contained the *HDE* gene. The fragments 3T and 5BT were ligated into pGEM5Zf(+) (Fig. 3.3.2.2).

Positive subclones were confirmed by restriction analysis and Southern blotting. Recombinants of pGEM5ZF(+) containing either fragment 3T or 5BT were identified and designated pGEM/3T and pGEM/5BT, respectively. Restriction digests were performed, and the patterns generated were compared to the known cDNA sequence to determine the orientations of the inserts. pGEM/5BT recombinants were found in both orientations, whereas pGEM/3T recombinants were found in only one orientation.

Clone λEMBL3/HDE 3a contained a larger fragment of the *C. tropicalis* genome (4.2 kbp) than λEMBL3/HDE 5b. This suggests that the *Sal*I sites flanking the *HDE* gene was not present in fragment 5BT. pGEM/3T was therefore used for all subsequent analyses.

3.3.3 Sequencing the *HDE* gene[12]

Exonuclease III digestion from both ends of the *HDE* gene in pGEM/3T

---

[12]Sequencing of the HDE gene was done in collaboration with James Sloots

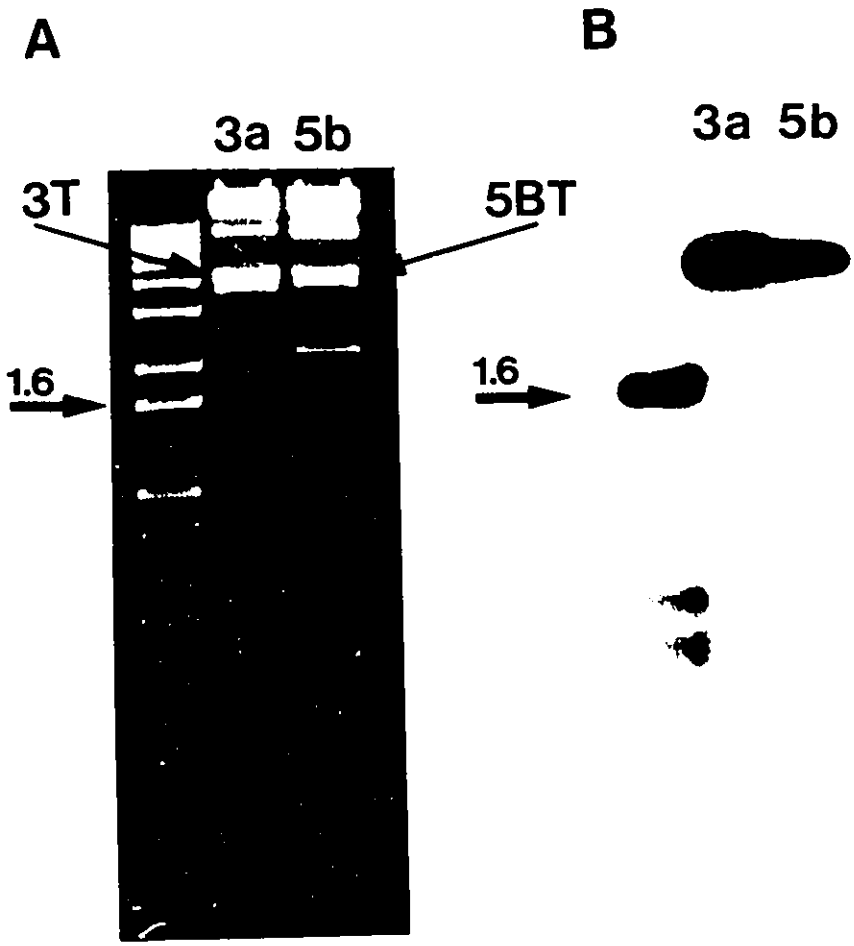


Figure 3.3.2.1 Southern blot analysis of λEMBL3 HDE clones 3a and 5b. One μg of DNA of each clone was digested with *Sal*I. The digests were separated by 0.8% agarose gel electrophoresis and stained with ethidium bromide (Panel A). The DNA was transferred to nitrocellulose and probed with the 5' fragment of the HDE gene in pUC118 (pBA) radiolabelled with α-^{32}P-dATP (Panel B). The bands containing HDE were identified by autoradiography and aligned with the ethidium bromide stained gel. Fragments 5BT and 3T were determined to carry the *HDE* gene (indicated by arrows in panel A) and were excised from the gel and subcloned into pGEM5Zf(+). The left lane of each panel contains the 1 kb ladder. The 1.6 kb piece (1636 nucleotides) is indicated in each case.

yielded 300 to 400 nucleotide deletion intervals across the 4.2 kbp *HDE* gene. Deletions from the 5' end of the gene were created using *Sph*I (3'- overhang; protected from exonuclease III digestion) and *Sac*II (5'- overhang; susceptible to exonuclease III digestion). Deletions from the 3' end of the gene were created using *Sac*I (3'- overhang; protected from exonuclease III digestion) and *Nde*I (5'- overhang; susceptible to exonuclease III digestion). Sequencing was carried out using the M13 forward and reverse primers. The summary of the sequencing strategy is shown in Fig. 3.3.3.1. In addition to exonuclease III constructs, 4 additional "drop-out" constructs were generated and sequenced to cover areas of the gene that could not be sequenced from the available exonuclease III constructs. The drop-out constructs were generated as follows: 1. *Bst*XI/*Hind*III: pGEM/3T was digested with *Hind*III at position 1169 in the *HDE* gene and with *Bst*XI in the multiple cloning site 3' to the gene. The ends were made blunt with T4 DNA polymerase and the plasmid was recircularized by ligation. 2. *Pfl*MI/*Nde*I: pGEM/3T was digested with *Pfl*MI at positions 2062 and 2394 in the *HDE* gene and with *Nde*I in the multiple cloning site 3' to the gene. The plasmid was recircularized as above. 3. *Eco*RI/*Nde*I: pGEM/3T was digested with *Eco*RI at position 1762 in the *HDE* gene and with *Nde*I in the multiple cloning site 3' to the gene. The plasmid was recircularized as for drop-outs 1 and 2. Drop-outs 1, 2 and 3 were sequenced in the 3' to 5' direction (with respect to the direction of transcription of the gene) using the M13 reverse primer. Drop-out 4 (*Pst*I/*Pst*I) was constructed by digestion of pGEM/3T with *Pst*I. Recircularization of the plasmid resulted in the drop-out of nucleotides 3' from the

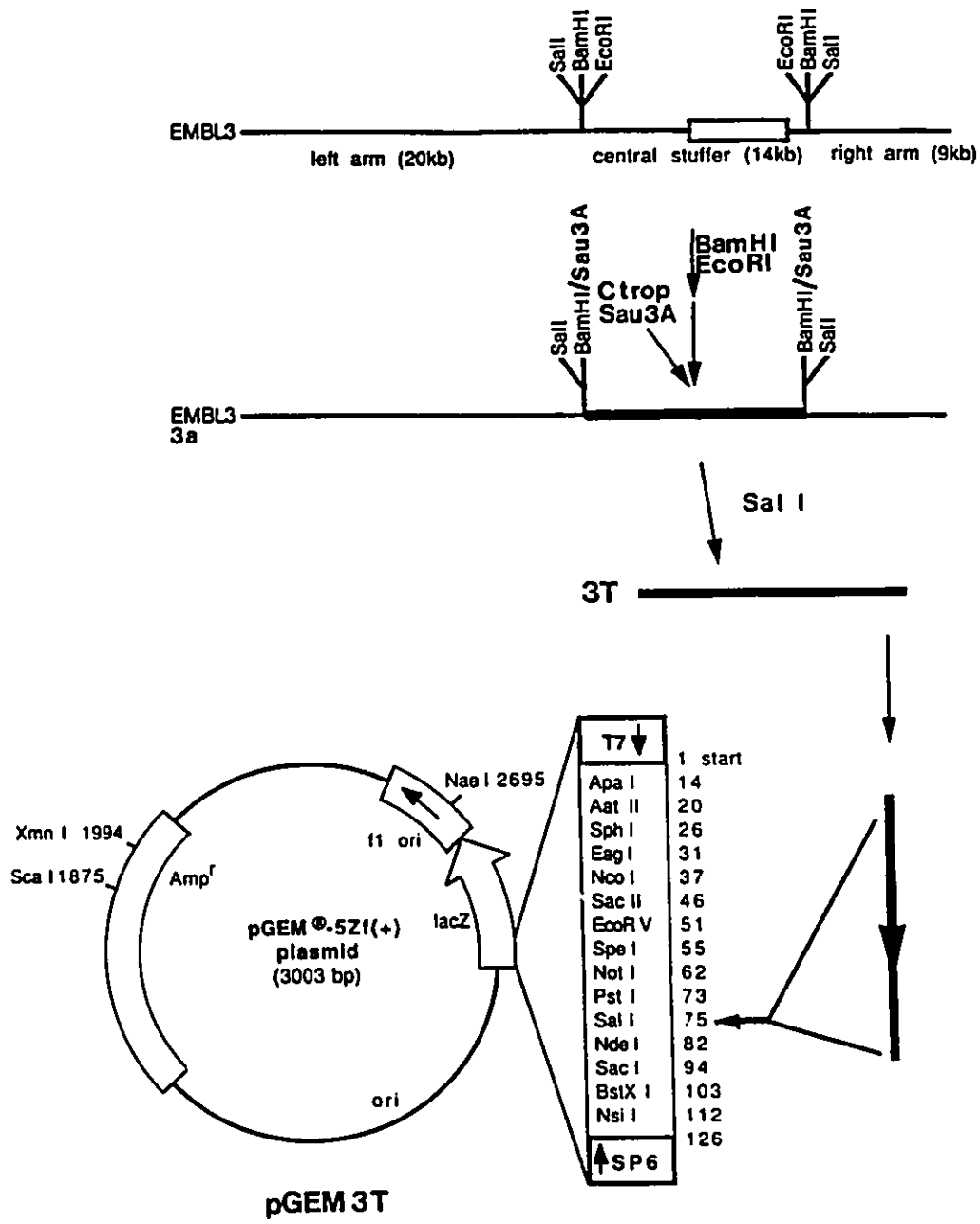


Figure 3.3.2.2 Cloning strategy of the *HDE* gene. λEMBL3 DNA was digested with *Bam*HI and *Eco*RI. *Sau*3AI-digested and size-fractionated (≈20 kbp) *C. tropicalis* DNA was ligated into λEMBL3 arms and plaques were screened with ^{32}P-labelled HDE cDNA to select clone λEMBL 3a. Digestion of λEMBL 3a with *Sal*I yielded several fragments, including 3T as shown. Fragment 3T (≈4.2 kbp) was detected by Southern hybridization and ligated into pGEM5Zf(+) to generate pGEM/3T. The orientation of the *HDE* gene in pGEM/3T is indicated by the arrow. Adapted from Davis *et al.*, 1986 and the Promega Biotech 1990 catalogue.

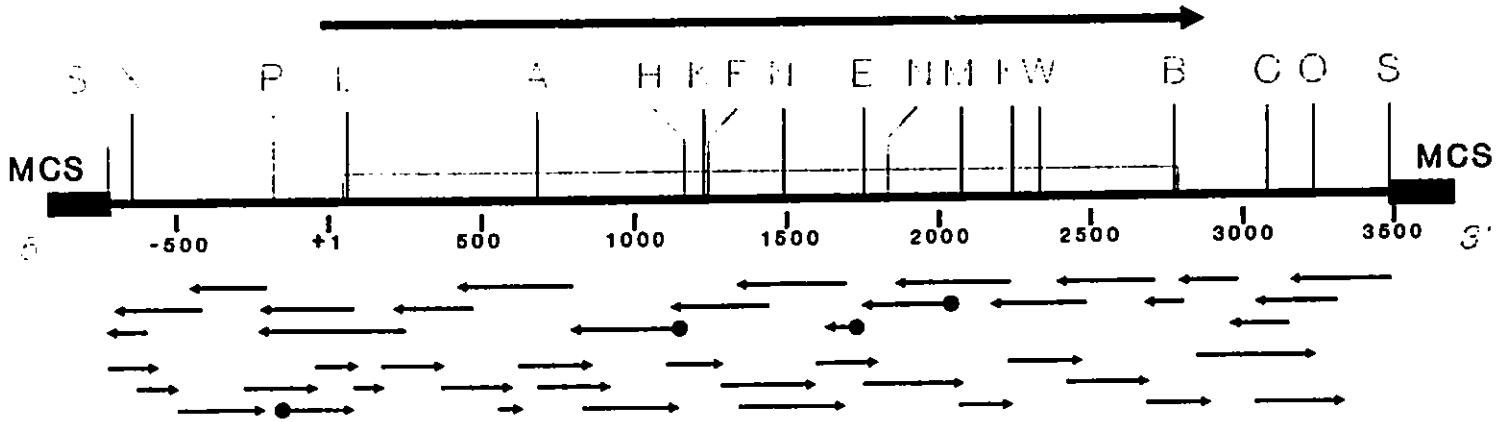


Figure 3.3.3.1 Summary of the sequencing of the HDE gene. The solid line represents the *C. tropicalis* genomic DNA fragment containing the *HDE* gene. The larger boxes at the 5' and 3' end of the geneomic DNA fragment represent the multiple cloning site of pGEM5Zf(+). Numbers below the genomic fragment correspond to the numbering of the *HDE* gene in Fig. 3.3.4.1. The open box shows the postition of the ORF. The extent and direction of the HDE transcript is indicated by the arrow above the genomic DNA fragment. The *HDE* gene was sequenced from exonuclease III deletions generated in pGEM/3T from both ends of the gene. Deletions from the 5' end of the gene were created using *Sph*I (3'-overhang) and *Sac*II (5'-overhang). Deletions from the 3' end of the gene were created using *Sac*I (3'-overhang) and *Nde*I (5'-overhang). The direction and extent of sequence determination is indicated by the arrows below the genomic DNA fragment. Sequences indicated by left pointing arrows were generated using the M13 reverse primer and sequences indicated by the right pointing arrows were generated using the M13 forward primer. The sequences generated by 'drop-out' restriction digests are indicated by a closed circle at the starting points of sequencing. The 5' MCS contains recognition sequences for *Nco*I, *Sac*II, *Eco*RV, *Not*I, *Pst*I and *Sal*I. The 3' MCS contains recognition sequences for *Sal*I, *Nde*I, *Sac*I, *Bst*XI and *Nde*I. The following restriction sites are indicated above the genomic DNA fragment: S, *Sal*I; X, *Xba*I; P, *Pst*I; L, *Aha*I; A, *Apa*I; H, *Hin*dIII; K, *Kpn*I; F, *Eco*RV; N, *Nco*I; E, *Eco*RI; M, *Pfl*MI; W, *Atw*NI; B, *Bgl*II; C, *Cla*I; O, *Xho*I. See text (section 3.3.3) for details.

*Pst*I site in the multiple cloning site (5' to the gene) to the *Pst*I site at position (-171) in the 5' non-transcribed region of the *HDE* gene. The region 3' of nucleotide -171 was sequenced with the M13 forward primer.

3.3.4 Analysis of the *HDE* gene

The nucleotide sequence of the *HDE* gene derived from clone pGEM/3T is shown in Fig. 3.3.4.1. The numbering begins at the major transcription start point as identifed by primer extension analysis[13]. The coding sequence of the *HDE* gene, beginning at nucleotide 61, is identical to the cDNA except for the presence of a single silent nucleotide (C to T) transition at position 2049 in the gene (position 1988 in the cDNA) and is indicated by t in Fig. 3.3.4.1. The sequence varies from the derived sequence of the cDNA upstream of nucleotide 42. This may be due to allelic differences, or to an artefact generated in the cloning of the cDNA. There are 725 nucleotides upstream from the transcription start site and 700 nucleotides downstream from the TAA stop codon. The TATA box is located between nucleotides -105 and -100.

Analysis of the regulation of the *HDE* gene by heterologous expression in *S. cerevisiae* has identified regions responsible for the induction of *HDE* expression in oleic acid-containing medium and repression in glucose-containing medium (Sloots *et al.*, 1991). The presence of glucose-repressive element(s) between nucleotides -466 and

---

[13]The major transcription start site was mapped by James Sloots.

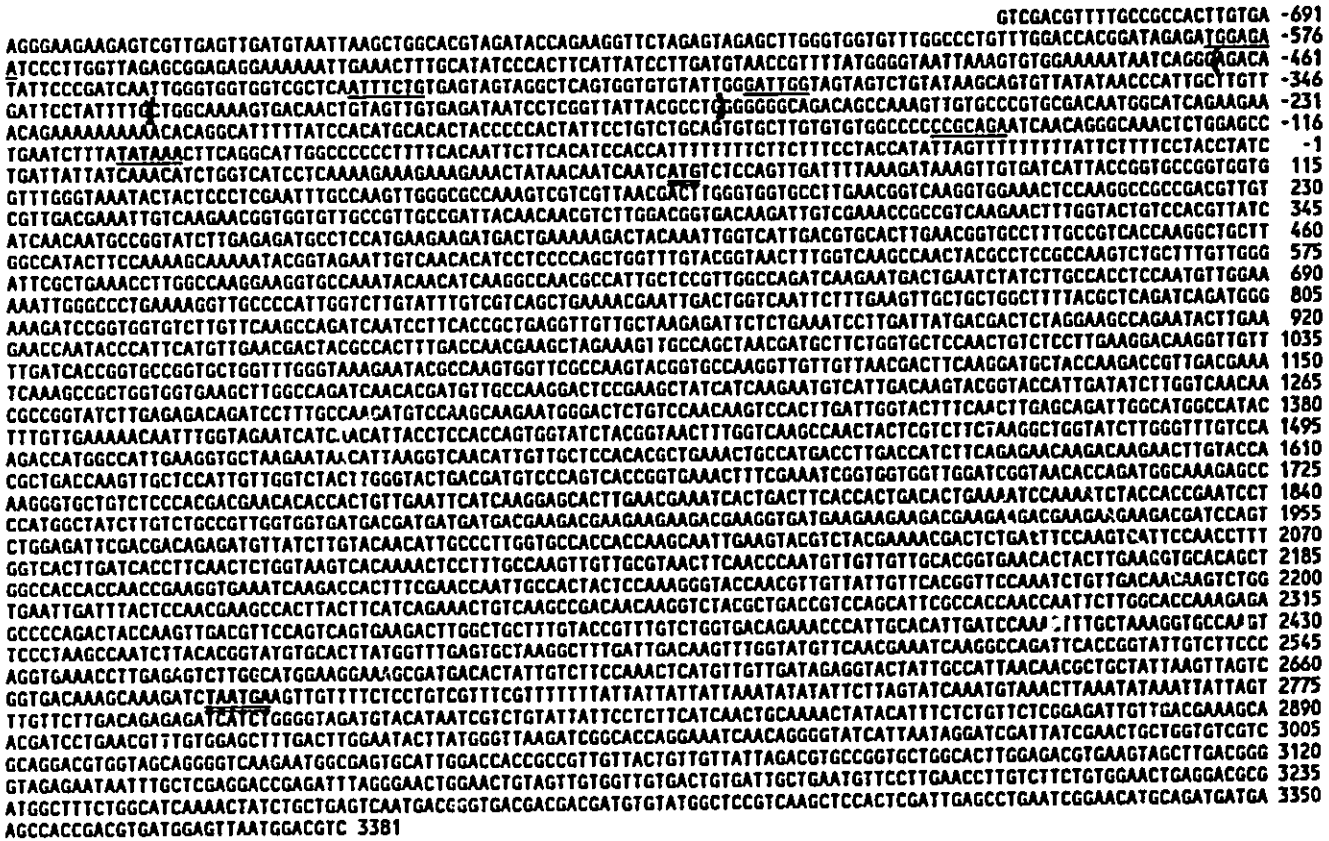


Figure 3.3.4.1 Sequence of the *HDE* gene. Numbering begins at the transcription start point (nucleotide +1). The TATA box appears in boldface and is underlined (nucleotides -105 to -100). The start and stop codons are in boldface and doubly underlined. The glucose- and oleic acid-responsive regions are bracketed (nucleotides -466 to -333 and -333 to -281). Consensus sequences as shown in Table 3.3.4.2 are highlighted in yellow. Potential yeast activator and repressor binding sites, as identified in *S. cerevisiae*, are underlined: ADR1-binding site (nucleotides -581 to -575), glucose-responsive site of *SUC2* (nucleotides -429 to -423), HAP2-HAP3-responsive site (nucleotides -393 to -388) and upstream repression sequence of *CAR1* (nucleotides -146 to -140). The sequence is deposited in the EMBL and GenBank data bases (accession No. X57854).

-333 was demonstrated by deletion mutagenesis (Sloots *et al.*, 1991). Within this region lies a sequence identified as a glucose-responsive site of *SUC2* from *S. cerevisiae* (Sarokin and Carlson, 1986). Comparison of the glucose-responsive region of the *HDE* gene with the upstream regions of genes encoding other oleic acid-inducible genes of *C. tropicalis* allowed for the identification of 2 putative glucose-responsive elements in this region of the *HDE* gene. These potential sequences and the derived consensus sequences are shown in Table 3.3.4.2. It is noteworthy that the conserved sequence between nucleotides -428 and -418 overlaps with the sequence identified by homology to the *SUC2* glucose-responsive element. A similar comparison of the oleic acid-responsive region of the *HDE* gene with the upstream regions of genes encoding other oleic acid-responsive genes of *C. tropicalis* allowed for the identification of a putative oleic acid-responsive element in this region of the *HDE* gene (Sloots *et al.*, 1991). This sequence and the derived consensus sequence are also shown in Table 3.3.4.2. In addition to these potential regulatory signals, the *HDE* upstream region also contains sequences homologous to other potential yeast activator and repressor binding sites identified in *S. cerevisiae*. There is a HAP2-HAP3-responsive site (nucleotides -393 to -388; Forsburg and Guarente, 1988) and an upstream repression sequence of *CAR1* (nucleotides -146 to -140; Luche *et al.*, 1990; see Fig. 3.3.4.1). Interestingly, there is also an ADR1 binding site (nucleotides -581 to -575). The *ADR1* gene has been shown to be a positive regulator of the transcription of genes encoding peroxisomal proteins in *S. cerevisiae* (Simon *et al.*, 1991).

Table 3.3.4.2

Sequences of the responsive regions of the *HDE* gene conserved in other *Candida tropicalis* genes encoding oleic acid-inducible proteins[1]

Genes[2]	Glucose[3] (distance)[5]		Oleate[4] (distance)[5]
	A	B	
	-468 -459	-428 -418	-295 -287
HDE	GGGAGACATA	TTTCTGTGAGT	CGGTTATTA
	-482 -473	-330 -320	-359 -351
POX-4[6]	GAGAGAGAGA	**A**TTGTCTGAGT	CGGTTATTC
	-396 -387	-194 -184	-223 -215
POX-18	GGGAGAGAGA	TTTTTGTATGT	CGGTAGTTA
	-580 -571	-437 -427	-253 -261
P450alk	GGTAGACTTA	TTCATGTGAAA	TGGTTGTTG
	-422 -431	-600 -590	-695 -687
CAT	GGGAAACAGA	TTTGTGTGAGG	TGGTTATT**.**
		T	C
Consensus	GGGAGAGATA	TTTNTGTGAGA	TGGTTATTA
	C G	G	C G G

[1]The sequences of the glucose-responsive (at -466 to -333) and the oleic acid-responsive (at -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*.

[2]*HDE*, trifunctional enzyme; *POX-4*, fatty acyl-CoA oxidase (Okazaki *et al.*, 1986); *POX-18*, peroxisomal 18-kDa protein (Szabo *et al.*, 1989); *P450alk*, alkane-inducible cytochrome P450 (Sanglard and Loper, 1989); *CAT*, catalase (Murray and Rachubinski, 1989b).

[3]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*.

[4]Sequence in the oleic acid-responsive region of *HDE* conserved in other oleic acid-responsive genes of *C. tropicalis*.

[5]Numbering is as shown in Fig. 3.3.4.1.

[6]Nucleotides in bold type deviate from the derived consensus sequences.

3.4 Expression of the *HDE* gene and subcellular localization of HDE *in vivo*

3.4.1 Transformation of *C. albicans*

The expression of the *HDE* gene and the import of HDE into peroxisomes were studied in the yeast *C. albicans* (strain SGY 243; Kurtz *et al.*, 1987). *C. albicans* SGY 243 contains a *ura3* homozygous deletion (Kelly *et al.*, 1987) and can be complemented by the introduction of the *C. albicans/E. coli* shuttle vector pMK22 containing the wild-type *URA3* gene (Kurtz *et al.*, 1987). pMK22 also contains a *C. albicans* autonomous replication sequence (*ARS*), which allows for replication in *C. albicans* and the gene encoding tetracycline resistance for selection of transformants in *E. coli* (Kurtz *et al.*, 1987). The genomic DNA *Sal*I fragment containing the *HDE* gene plus its flanking regions was subcloned into pMK22 to generate pMK22/HDE50. The subcloning strategy is outlined in Fig. 3.4.1.1. Briefly, the *Sal*I fragment was isolated from pGEM/3T, made blunt with the Klenow fragment of DNA polymerase I, and subcloned into the unique *Sca*I site of pMK22 to generate pMK22/HDE50. Recombinants were selected by tetracycline resistance and colony hybridization. Southern blotting and restriction analysis were used to confirm recombinants and to determine the orientation of the *HDE* gene.

The recombinant plasmid pMK22/HDE50 was used to transform *C. albicans* SGY 243 to uracil prototrophy. Transformants were picked and transferred to fresh selective plates. Although pMK22/HDE50 contains sequences that permit the plasmid

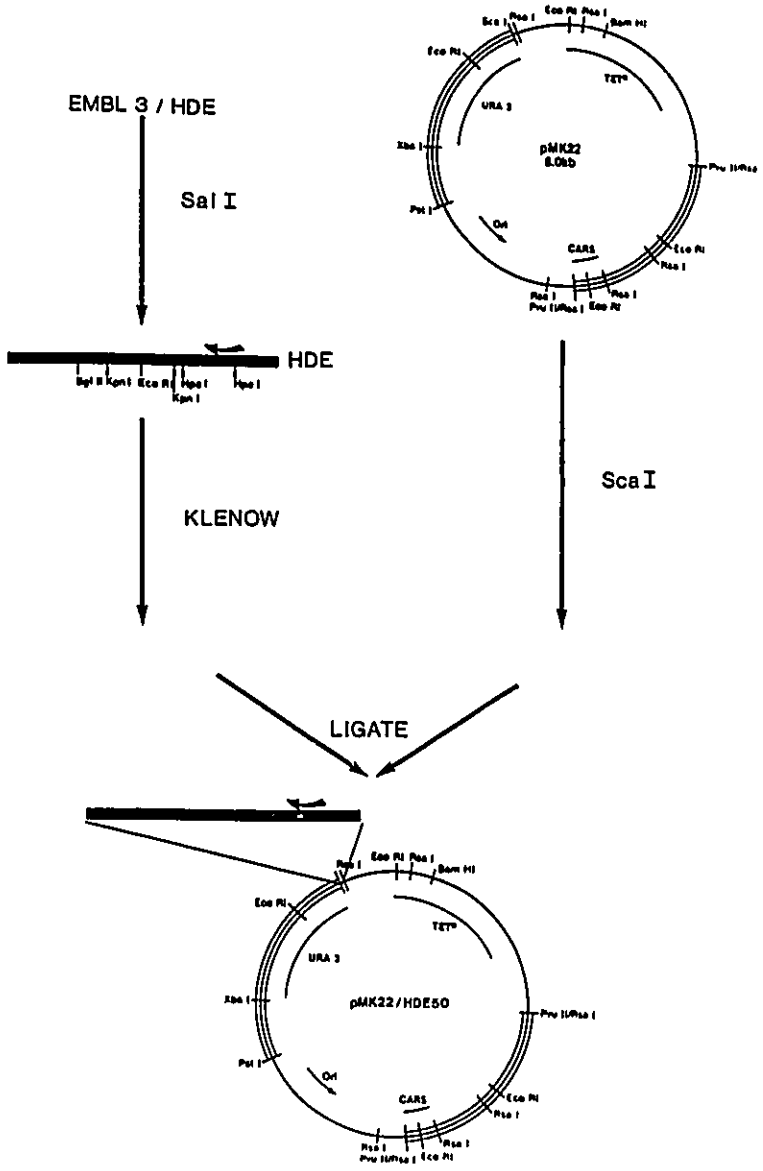


Figure 3.4.1.1 Subcloning of the gene encoding HDE into pMK22. A 4.2 kbp fragment containing the *HDE* gene was isolated from the recombinant EMBL3/HDE phage (λEMBL 3a). The fragment was ligated into pMK22 to generate pMK22/HDE50. pMK22/HDE50 was then used to transform *C. albicans* SGY 243 (*ura3*) to uracil prototrophy. The arrow indicates the direction of transcription of the HDE gene. The partial restriction map of pMK22 was reproduced from Kurtz *et al.* (1987).

to replicate autonomously, the plasmid was unstable in *C. albicans*, even when the cells were grown under selective pressure. To circumvent this problem, clones which had spontaneously integrated the plasmid into the genome were selected. Integrants were chosen by three criteria: rapid growth on selective medium, lack of plasmid loss when grown on rich medium for more than 10 generations, and a positive signal on Southern blots probed with both pBR322 and HDE cDNA.

Selection of integrants allowed for the maintenance of the clones on selective medium, as well as for the coinduction of peroxisomes and *C. tropicalis* HDE in the rich, non-selective medium YPBO without loss of the plasmid. Early attempts to induce peroxisomes in selective medium plus oleic acid were unsuccessful. It is thought that the plasmid replicates autonomously for a few generations and then integrates in tandem multimers into the genome by homologous recombination.

A Southern blot of total genomic DNA isolated from both untransformed SGY 243 and a ura$^+$ transformant, clone L-1, probed with both the HDE cDNA fragment BA and pBR322 is shown in Fig. 3.4.1.2. Using this analysis, and comparison of clone L-1 to other integrated transformants (not shown), clone L-1 was determined to contain integrated pMK22/HDE50 DNA at a copy number of 5-8 (lane 7). Control DNA from untransformed SGY 243 did not hybridize to either pBR322 or the HDE cDNA fragment BA (Panels B and C, lanes 6, respectively), indicating a lack of homology between the *HDE* gene and *C. albicans* genomic DNA sequences. This result would suggest that the integration event occurred at the *ura3* locus, which contains a 0.6 kbp

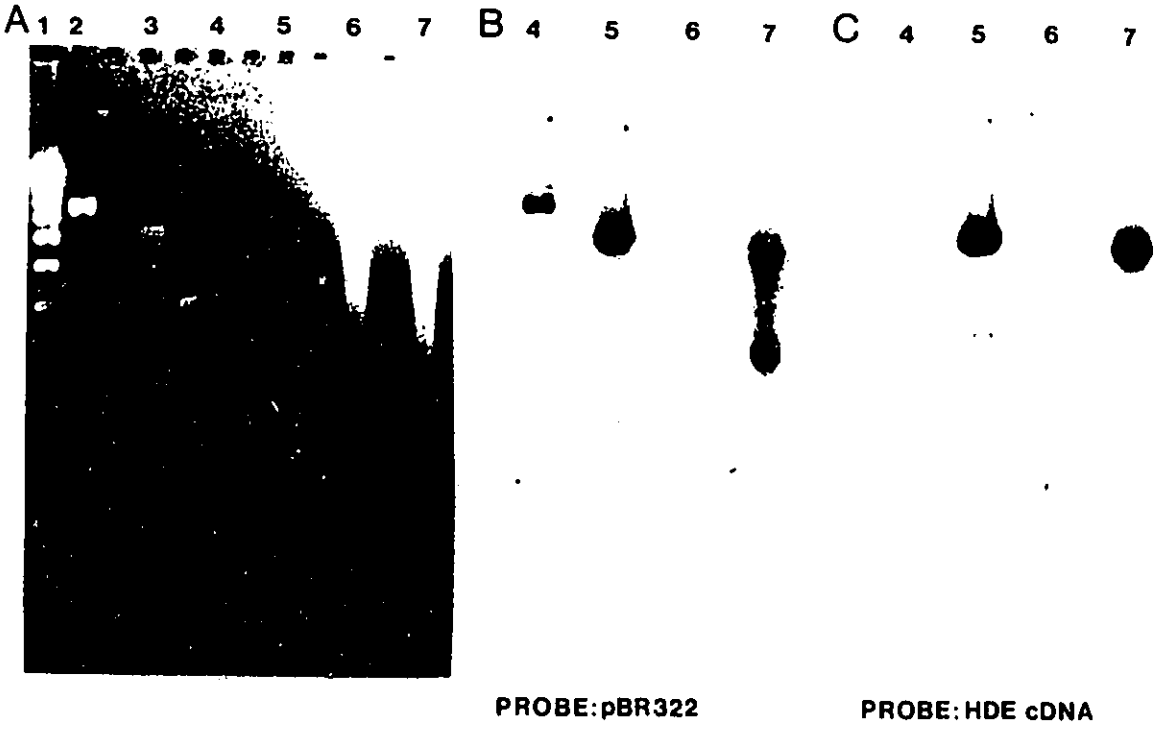


Figure 3.4.1.2 Southern blot analysis of *C. albicans* SGY 243 and L-1. Genomic DNA was digested with *Pst*I. pMK22 and pMK22/HDE50 were also digested with *Pst*I. DNA was separated on a 0.8% agarose gel, stained with ethidium bromide, and then transferred to nitrocellulose. The blot was first probed with radiolabeled pBR322. The blot was boiled in 0.1 X SSC, 0.1% SDS to remove pBR322 and then probed with radiolabeled HDE cDNA (fragment BA; nucleotides -38 to 1703). Panel A, ethidium bromide stained agarose gel. One kb ladder (lane 1), 80 ng pMK22 (lane 2), 80 ng pMK22/HDE50 (lane 3), 20 ng of pMK22 (lane 4), 20 ng of pMK22/HDE50 (lane 5), 5 μg *C. albicans* SGY 243 genomic DNA (lane 6), 5 μg L-1 genomic DNA (lane 7). Panel B, Southern blot analysis of Panel A probed with radiolabeled pBR322. Panel C, Southern blot analysis of Panel A probed with radiolabeled HDE cDNA fragment BA.

region homologous to the *URA3* containing fragment in pMK22 (Kelly et al. 1987) or at a homologous *ARS* sequence. Alternatively, limited homology between the endogenous *HDE* gene and the *HDE* gene of *C. tropicalis*, which was not detected under the stringency of the wash conditions used in this blot, may be responsible for the integration event. This is a reasonable scenario considering that there are regions of high conservation in the *HDE* genes from *C. tropicalis*, *S. cerevisiae* and *N. crassa* (section 3.2.7.2). In addition, restriction analysis and Southern blotting of genomic DNA from different ura$^+$ transformants indicated that the location of integration varies between clones (not shown).

3.4.2 Expression of *HDE* and induction of *HDE*

Yeast lysates were prepared from SGY 243, L-1 and 21-1 (a pMK22/HDE50 integrant with an estimated copy number of 1) grown in either oleic acid- (YPBO) or glucose- (YEPD) containing medium. Growth of *C. tropicalis* on oleic acid has been shown to induce *HDE* gene expression 8- to 9-fold over growth on glucose-containing medium (Nuttley et al. 1988; section 3.2.4). To determine if *HDE* expression was induced in *C. albicans*, polypeptides from lysates and purified peroxisomes were separated by SDS-PAGE and transferred to nitrocellulose for western blot analysis (Fig. 3.4.2.1). HDE (100 kDa; Moreno de la Garza *et al.* 1985; Nuttley *et al.* 1988; section 3.2.3) was detectable in lysates from clones L-1 (panel A, lane L, o) and 21-1 (panel A, lane 21-1, o) grown on oleic acid. Control cells did not contain an immunoreactive

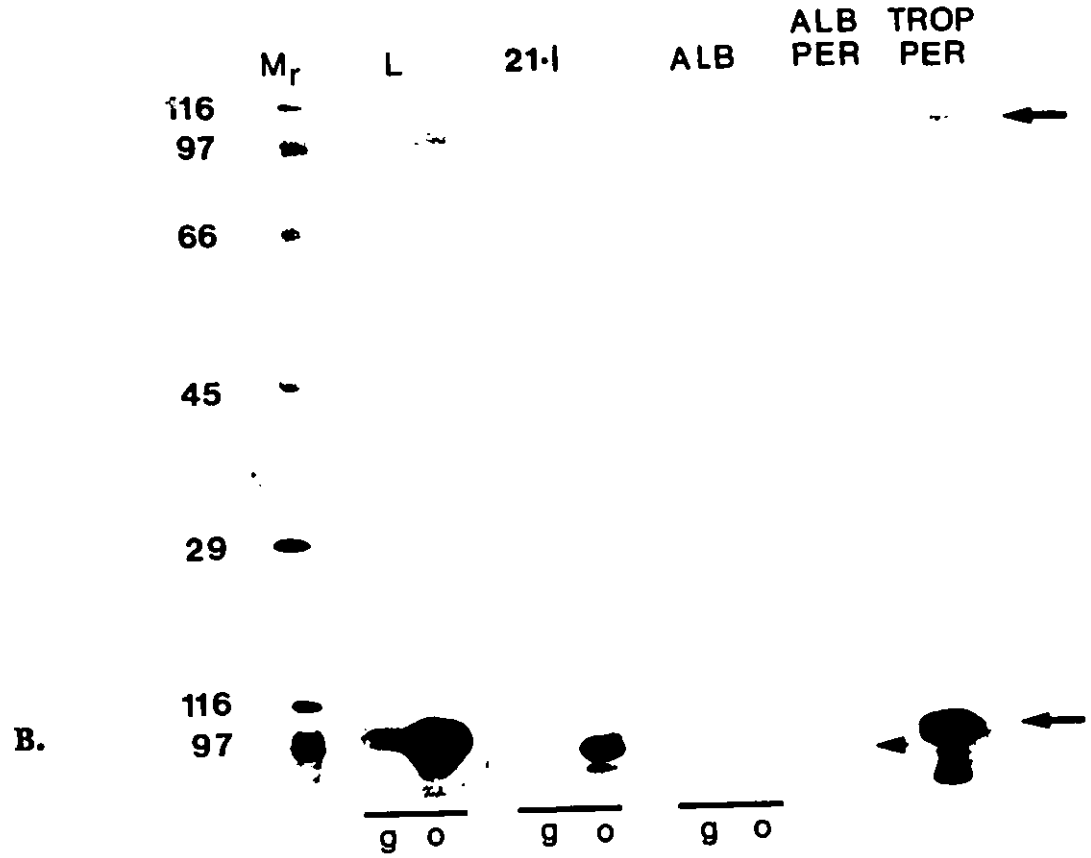


Figure 3.4.2.1 Expression of *HDE* in *C. albicans* and induction of HDE in oleic acid-containing medium. Panel A. Western blot probed with anti-HDE antibody. Polypeptides of either yeast lysates (grown on oleic acid or glucose medium) or purified peroxisomes were separated by SDS-PAGE, transferred directly to nitrocellulose and probed with antibody to HDE. The arrow indicates HDE (100 kDa). Lanes: M_r, ^{14}C M_r standards; L, 21-l, ALB, 150 μg of yeast lysate from clones L-1, 21-l and untransformed *C. albicans* SGY 243, respectively, grown on either glucose (g) or oleic acid (o) medium; ALB PER, 20 μg of purified peroxisomes from *C. albicans* SGY243 grown on oleic acid medium; TROP PER, 20 μg of purified peroxisomes isolated from *C. tropicalis* grown on oleic acid medium. The film was exposed for 16 h. Panel B. The blot shown in panel A was re-exposed for 72 h. The arrowhead shows a weak cross-reaction of anti-HDE with a *C. albicans* peroxisomal polypeptide (97-98 kDa), presumably endogenous trifunctional enzyme.

polypeptide of this M_r (lane ALB, o). However, upon overexposure of the blot (panel B), a weak reaction at 97-98 kDa was detected in lysates from both control ALB, o; (lanes L, o; 21-1, o; arrowhead) and transformed cells. This polypeptide is present in peroxisomes purified from control cells (lane ALB PER) and presumably corresponds to endogenous trifunctional enzyme (hydratase-dehydrogenase-epimerase). Densitometric analyses indicated that HDE was induced in *C. albicans* approximately 3.6-fold (compare lanes L, g and L, o), suggesting that similar mechanisms of gene expression operate in *C. albicans* and *C. tropicalis*.

3.4.3 HDE is targeted to peroxisomes of *C. albicans*

In order to determine if HDE was targeted to peroxisomes of *C. albicans*, peroxisomes from SGY 243, *C. tropicalis* and L-1 grown on oleic acid-containing medium were isolated on a sucrose density gradient. Peroxisomal proteins were separated by SDS-PAGE. Fig. 3.4.3.1 shows a partial protein profile of peroxisomes purified from SGY 243 (lane A), L-1 (lane L), and *C. tropicalis* (lane T). Comparison of lanes A and L reveals similar protein profiles in peroxisomes from control and transformed cells except for an additional band at 100 kDa present in the peroxisomal fraction isolated from L-1. This polypeptide comigrated with HDE in the *C. tropicalis* peroxisomal fraction (lane T, arrow). Western blot analysis of an identical gel probed with anti-HDE antibody (Fig. 3.4.3.1, panel B) indicated that the 100 kDa species present in L-1 peroxisomes was indeed HDE. At this exposure, no reaction was detected

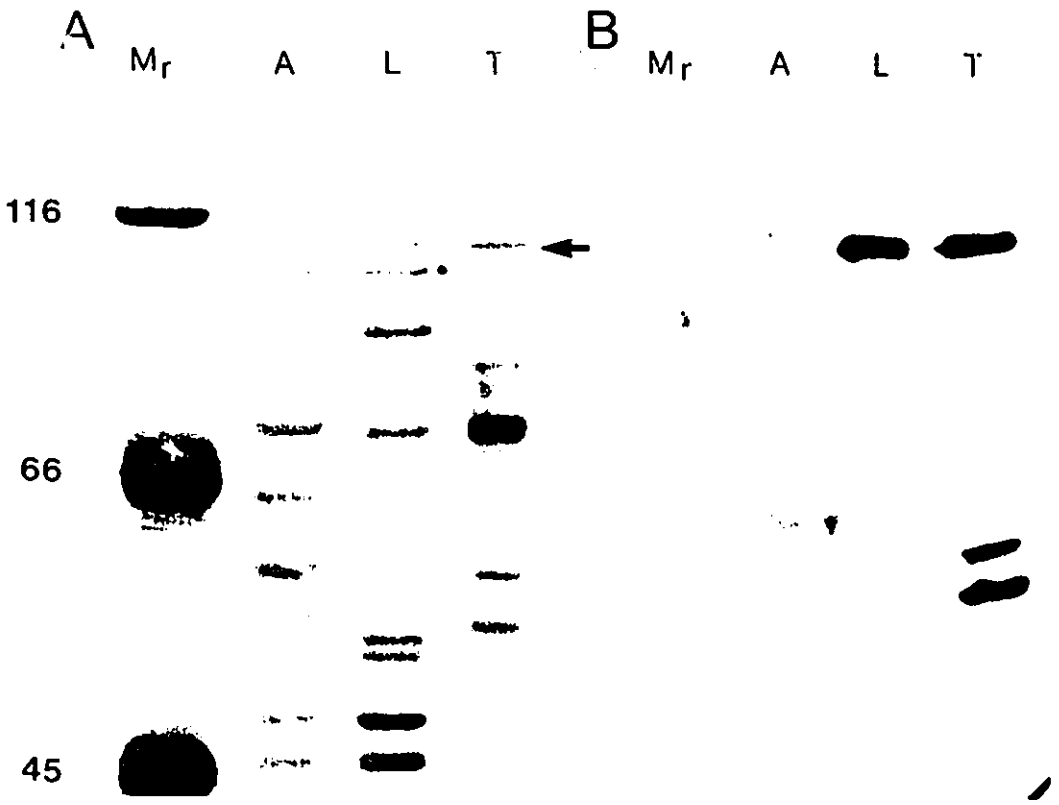


Figure 3.4.3.1 Targeting of HDE to peroxisomes of *C. albicans*. Peroxisomes were isolated from oleic acid-grown cells. Polypeptides were separated by SDS-PAGE (7% acrylamide) and either stained with Coomassie blue (Panel A) or transferred to nitrocellulose and probed with anti-HDE antibody (Panel B). HDE is indicated by the arrow. Lanes: M_r, molecular weight markers (kDa) as in Fig. 3.4.2.1.; A, peroxisomes isolated from *C. albicans*; L, peroxisomes isolated from *C. albicans* transformed with pMK22/HDE50 (clone L1); T, peroxisomes isolated from *C. tropicalis*. Arrow indicates HDE. Closed circle indicates probable endogenous trifunctional enzyme.

in peroxisomes isolated from control cells (lane A). The two bands at M_r 55-60 are breakdown products of HDE seen upon freezing and thawing of *C. tropicalis* peroxisomes (cf. Fig. 5, Fujiki et al. 1986).

Fig. 3.4.3.2 shows the subcellular fractionation profile from L-1. To our knowledge peroxisomes had not been isolated from *C. albicans* prior to this study. A good separation of peroxisomes and mitochondria was achieved, as determined by separation of catalase and cytochrome *c* oxidase activities, respectively. The densities of *C. albicans* peroxisomes (1.21 $g \cdot cm^{-3}$) and mitochondria (1.17 $g \cdot cm^{-3}$) peak fractions were similar to those found for *C. tropicalis* peroxisomes (1.23 $g \cdot cm^{-3}$) and mitochondria (1.17 $g \cdot cm^{-3}$; Nuttley *et al.*, 1990). Mitochondrial contamination of peroxisomes was calculated by the amount of cytochrome *c* oxidase activity in the peroxisomal fraction and was determined to be 5.8%. HDE (as detected by immunoblotting and ß-hydroxyacyl coenzyme A dehydrogenase activity) localized to peroxisomes of *C. albicans*. There was no significant anti-HDE antibody reaction or dehydrogenase activity in the mitochondrial fractions. Taken together, these results demonstrate that HDE is targeted to peroxisomes of *C. albicans* (clone L-1).

3.4.4 Dehydrogenase activity of HDE

The specific dehydrogenase activity of HDE was determined in both yeast lysates and purified peroxisomes. It was reasoned that if HDE is active in L-1, then the specific dehydrogenase activity should be greater in both lysates and purified peroxisomes

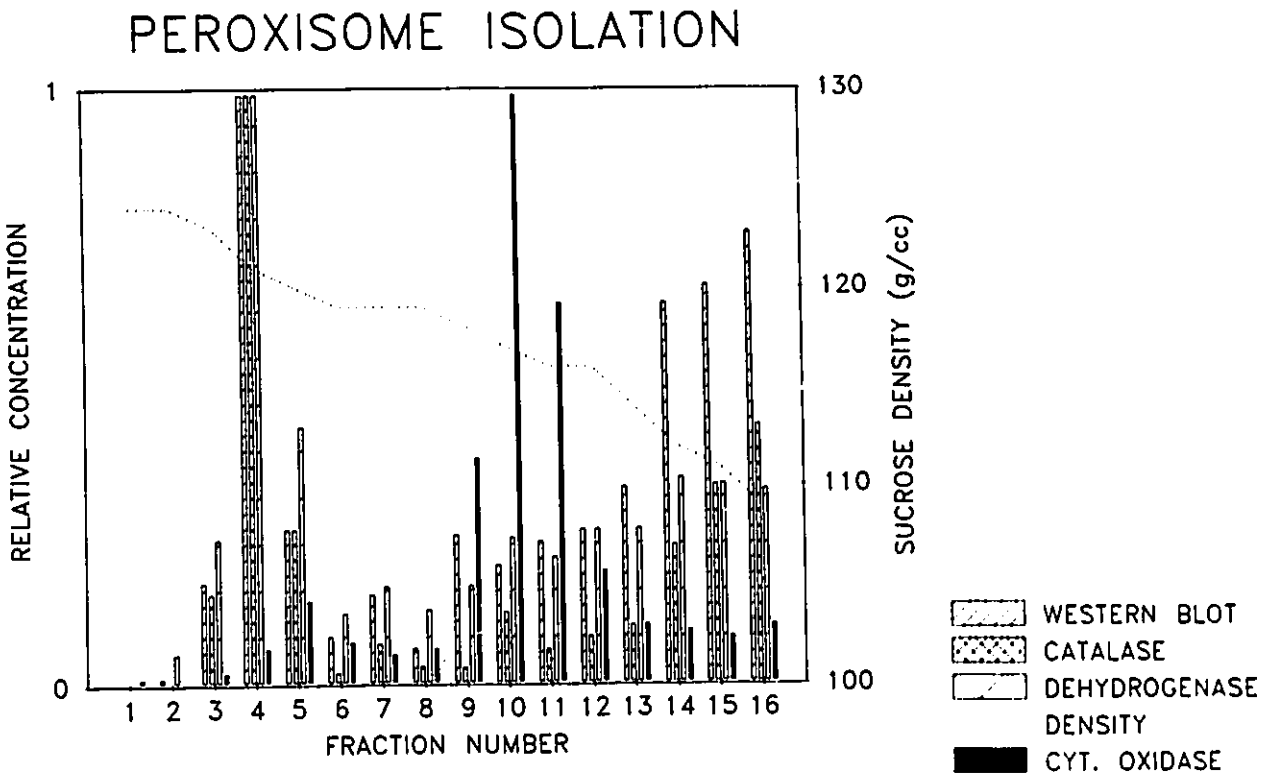


Figure 3.4.3.2 HDE fractionates with peroxisomes from *C. albicans*. Peroxisomes were purified from L-1 on a sucrose step gradient. 2 ml fractions were collected beginning from the bottom of the gradient and analyzed for catalase activity (peroxisomes), cytochrome *c* oxidase activity (mitochondria), ß-hydroxyacyl-CoA dehydrogenase activity (HDE), and immunoreactivity (western blot) with anti-HDE antibody. The relative activities of each fraction were determined and the peak fractions were normalized to 1. The density ($g \cdot cm^{-3} \times 10^2$) of each fraction is given on the right axis.

prepared from L-1 than from control cells (SGY 243) due to the presence of HDE in addition to endogenous trifunctional enzyme. The hydratase and epimerase activities were not assayed. Lysates showed a 4.2 ± 1.3-fold increase in dehydrogenase specific activity, while purified peroxisomes showed an increase of 2.2 ± 0.3-fold (Fig. 3.4.4.1). The average activities in peroxisomes isolated from SGY 243 and L-1 were 0.5 μmols • s^{-1} • mg $protein^{-1}$ and 1.1 μmols • s^{-1} • mg $protein^{-1}$, respectively. The two fold increase in the specific activity of peroxisomal dehydrogenase activity in L-1 compared to wild-type cells is in agreement with what is seen on the Coomassie- stained gel of peroxisomes purified from L-1 (Fig. 3.4.3.1, lane L), since the presumptive trifunctional enzyme of *C. albicans* (closed circle) and HDE (arrow) are in approximately equimolar amounts. The differential increase in specific dehydrogenase activity in lysates compared to purified peroxisomes suggests that more HDE is expressed in L-1 than is targeted into peroxisomes, and that the targeting of HDE to peroxisomes is not a prerequisite for dehydrogenase activity of HDE. It is not expected that this difference in specific activity of HDE between lysate and purified peroxisomes is due to leakage of HDE from peroxisomes, since this would require differential leakage between HDE and the endogenous dehydrogenase activity.

The ability to target HDE to peroxisomes in an *in vivo* heterologous expression system set the stage for the expression of mutant forms of *HDE* to determine which portion(s) of the gene encode the peroxisomal targeting signal (PTS) of HDE. In addition, as more is learned about the nature of the PTS(s), it will become increasingly

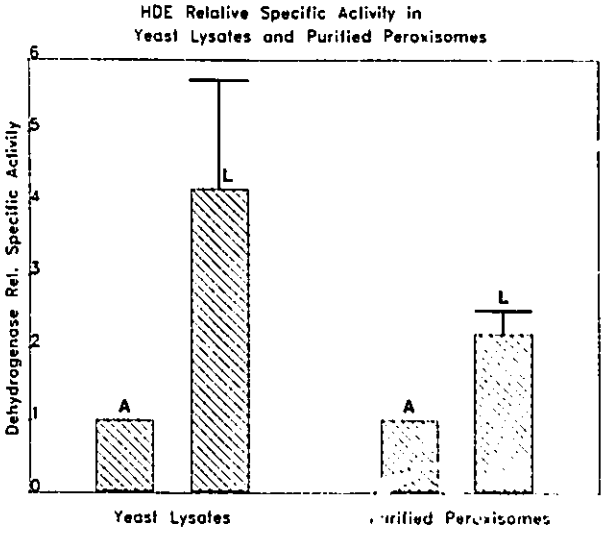


Figure 3.4.4.1 Imported HDE is enzymatically active. The specific activity of ß-hydroxyacyl-CoA dehydrogenase was determined in both *C. albicans* SGY 243 (A) and clone L-1 (L) in yeast lysates and purified peroxisomes. Equal amounts of protein were assayed for activity.

important to quantitate the efficiency of different PTSs in different contexts. To date, the only quantitative data regarding peroxisomal protein import efficiency come from *in vitro* import studies (Small et al. 1988; Miyazawa et al. 1989; see section 3.8). The ability to isolate peroxisomes from *C. albicans* that contain heterologous peroxisomal proteins provides the tools for the assessment of import efficiencies *in vivo*.

Although the absolute numbers vary according to gene copy number, *HDE* expression is induced when transformants are grown in oleic acid-containing medium. This suggests that, in addition to using *in vivo* heterologous expression of *HDE* as a model for studying PTSs, a similar system could be used to study the expression of peroxisomal genes which respond to different growth conditions. This has recently been accomplished for *HDE* by expression in *S. cerevisiae* (Sloots *et al.*, 1991).

3.5 Subcellular location of mutant forms of *C. tropicalis* HDE expressed in *C. albicans*

When HDE is made in *C. albicans*, it is targeted to peroxisomes (section 3.4; Aitchison and Rachubinski, 1990). This heterologous expression system was used to delineate the PTS within HDE. *HDE* deletion and substitution mutants were expressed in *C. albicans*, and the subcellular targeting of HDE was assayed initially by immunodetection within the 20,000 x g pellet (20 kgP; mitochondria and peroxisomes) or supernatant fraction (20 kgS; mainly cytosol) from a yeast homogenate, using anti-HDE serum. In all cases, equivalent cellular fractions of 20 kgP and 20 kgS were loaded

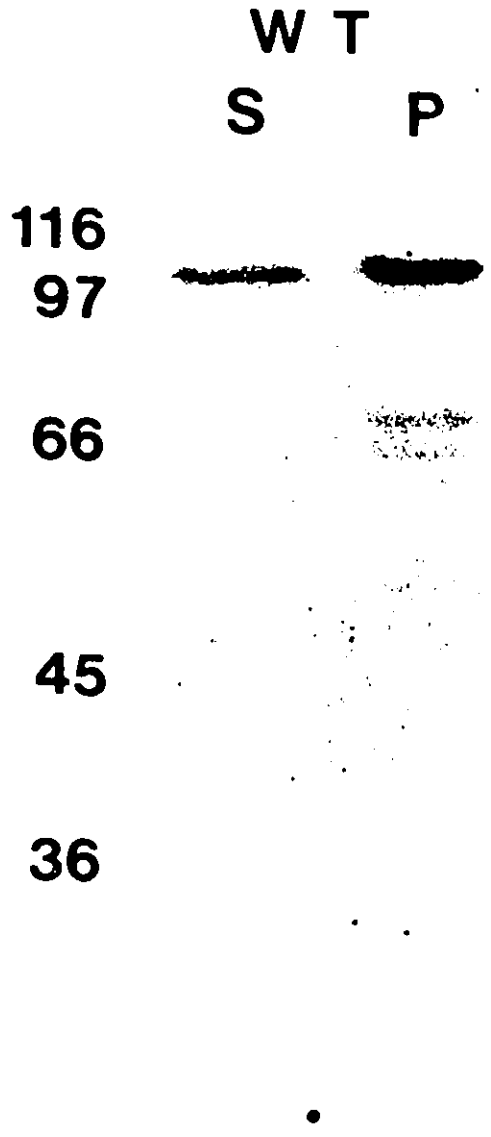


Figure 3.5.1 Western blot analysis of wild-type HDE in organellar and cytosolic fractions. *HDE* was expressed in *C. albicans* as described in section 3.4. The 20 kgP (P) and 20 kgS (S) fractions were prepared and resolved by SDS-PAGE. HDE was identified by western blotting using antiserum specific for HDE. 100 μg of protein was loaded in lane P. The equivalent cellular fraction of 20 kgS was loaded in lanes S.

onto an SDS-polyacrylamide gel for western blot analysis. This allows for direct comparison of the proportion of pelletable and non-pelletable protein. Analysis of wild-type HDE by this assay is shown in Fig. 3.5.1. The detection of HDE in the 20 kgS is due, at least in part, to organelle breakage during isolation as evidenced by catalase activity leaked from peroxisomes into the supernatant (Alexson *et al.*, 1985). Further fractionation of the 20 kgP by density gradient centrifugation to separate peroxisomes and mitochondria, followed by detection within gradient fractions using SDS-PAGE and western blotting, allowed for identification of the specific organellar location of the proteins. Absence of HDE from the 20 kgP coupled with its presence in the 20 kgS indicates the mutant protein remains cytosolic and thus lacks sequences required for sorting to peroxisomes.

3.5.1 Large internal deletions

Two large internal deletions of the *HDE* gene were expressed in *C. albicans* SGY 243. The organellar or cytosolic location of the HDE mutants was assayed by SDS-PAGE and western blotting. pMK22/H1/DO was created by digestion of pMK22/HDE50 with *Hpa*I, which cuts twice in the *HDE* gene (nucleotides 109 and 1051 in the coding sequence). Religation of the plasmid created an in-frame deletion of amino acids 38 to 353. pMK22/K1/DO was created in a similar manner using *Kpn*I digestion (nucleotides 1180 and 2188 in the coding sequence) of pMK22/HDE5[illegible] followed by religation. This created an in-frame deletion of amino acids 395 to 731.

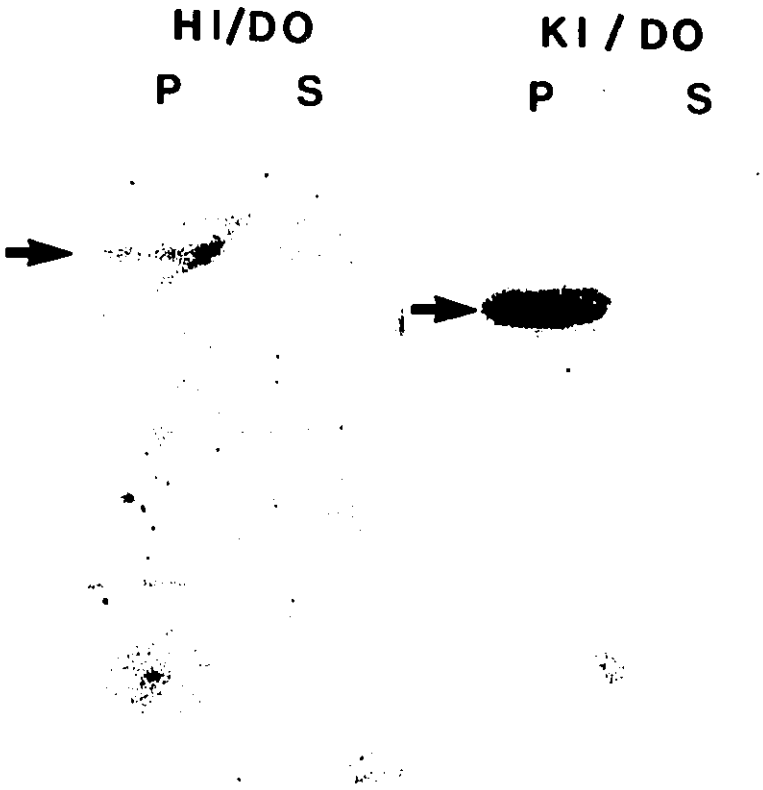


Figure 3.5.1.1 Western blot analysis of HDE deletion mutants H1/DO and K1/DO in organellar and cytosolic fractions. H1/DO and K1/DO were expressed in *C. albicans*. The 20 kgP (P) and 20 kgS (S) fractions were prepared and resolved by SDS-PAGE. HDE was identified by western blotting using antiserum specific for HDE. 100 μg of protein was loaded in lanes P. The equivalent cellular fraction of 20 kgS was loaded in lanes S. H1/DO, *Hpa*I deletion mutant (calculated molecular mass, 65604 kDa); K1/DO, *Kpn*I deletion mutant (calculated molecular mass, 61837 kDa).

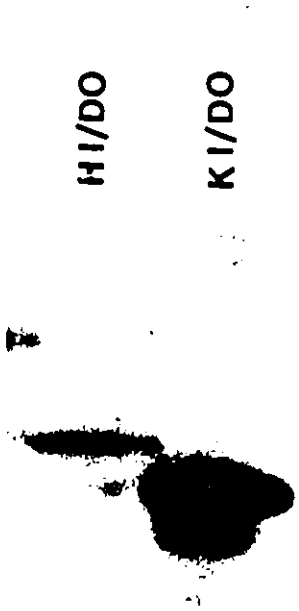


Figure 3.5.1.2 Western blot analysis of purified peroxisomal fractions expressing H1/DO and K1/DO. 20 μg of purified peroxisomes from *C. albicans* expressing H1/DO and K1/DO was loaded in each lane.

C. albicans cells were transformed with pMK22/H1/DO and pMK22/K1/DO. Clones with integrated copies of the plasmid were chosen for further analysis. The cells were grown in oleic acid-containing medium to coinduce peroxisomes and the mutant forms of the *HDE* gene. Western blots of the 20 kgP and 20 kgS fractions from these clones are shown in Fig.3.5.1.1. Both of these mutant forms of HDE with internal deletions remained organellar (lanes P) and were not mislocalized to the cytosol (lanes S). The absence of H1/DO and K1/DO peptides in the 20 kgS is thought to be the result of degradation of the polypeptides in the cytosol; however, this has not been investigated further. To determine if these deletions were correctly targeted to peroxisomes, the 20 kgP fractions were further separated by sucrose and Nycodenz density gradient centrifugation. The peroxisomal proteins were separated by SDS-PAGE and probed by western blotting with anti-HDE serum (Fig. 3.5.1.2). This analysis indicated that the H1/DO and K1/DO polypeptides were correctly targeted to highly purified peroxisomes (section 3.4.3), suggesting that these deletions did not disrupt the peroxisomal targeting signal of the protein.

3.5.2 Deletion of the carboxy-terminal tripeptide

As discussed in section 1.7.5 firefly luciferase is targeted to peroxisomes by its carboxy-terminal tripeptide Ser-Lys-Leu (Gould *et al.*, 1989). To investigate a similar role for the carboxy-terminus of HDE, the codons encoding the three terminal amino acids of HDE (Ala-Lys-Ile) were deleted by site-directed mutagenesis of pGEM/3T. The deletion mutants were first screened for the absence of a *Bgl*II restriction site that is

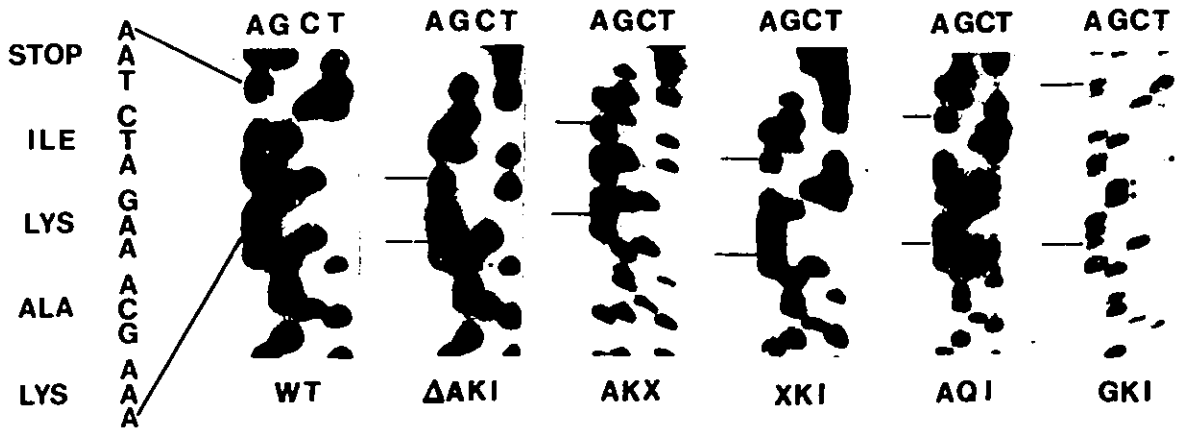


Figure 3.5.2.1 Dideoxynucleotide sequence analysis of mutant forms of *HDE*. Mutants in the carboxyl-terminal 3 codons of the *HDE* gene were generated as described in section 3.5.2.1. The wild-type (WT) sequence is shown at the left side (5' to 3'; bottom to top). Horizontal lines on the sequence of mutants indicate the boundaries of the sequence shown at left. Dots on the sequence of AQI and GKI indicate the altered nucleotides.

located within the nucleotide sequence encoding Ala-Lys-Ile-COOH (see Fig. 3.2.5.2), and then sequenced (Fig. 3.5.2.1). The altered genes were subcloned into pMK22 in the same manner and orientation as the wild-type gene and transformed into *C. albicans* SGY 243. Integrants were chosen.

Expression of the gene was first analyzed by western blotting of total lysates from oleic acid and glucose-grown cells (Fig. 3.5.2.2). ΔAKI-1A was the first clone to be assayed for expression. The gene was expressed at sufficiently high levels for detection and was therefore chosen for use in all subsequent assays.

The ΔAKI mutant form of HDE was assayed for its ability to be targeted to peroxisomes. Fig. 3.5.2.3 shows that deletion of the carboxy-terminal three amino acids (Ala-Lys-Ile) of HDE completely abolished its immunodetection within the organellar fraction (20 kgP) and caused mislocalization of HDE to the cytosolic fraction (20 kgS) (Fig. 3.5.2.3, ΔAKI).

The terminal Ile alone was deleted from the carboxy terminus to give the mutant AKX (Fg. 3.5.2.1). AKX-8 was selected for the targeting studies (Fig. 3.5.2.2). Subcellular fractionation of AKX-8, like ΔAKI, showed the protein to be mislocalized to the cytosolic (20 kgS) fraction, (Fig. 3.5.2.3, AKX). The localization of these mutants to the cytosolic fraction is due to mistargeting and not to peroxisome breakage during the isolation, since there were no differences in catalase leakage into the 20 kgSs for these different preparations as compared to clones harbouring the wild-type gene. Taken together this evidence suggests that the intact carboxy-terminal tripeptide Ala-Lys-

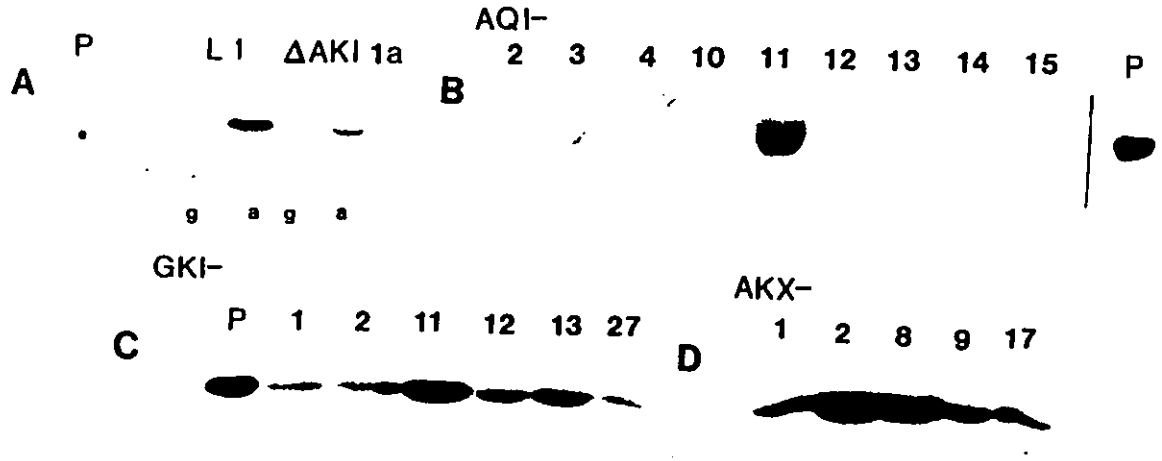


Figure 3.5.2.2 Expression of carboxy-terminal mutants forms of HDE in *C. albicans*. Expression of HDE was monitored by western blotting of yeast lysates (200 μg protein) using anti-HDE serum. The amount of HDE present on the blot was estimated by comparison to the anti-HDE immunoreaction observed from 20 μg of *C. tropicalis* purified peroxisomes (lanes P). Panel A: Lysates from L-1 or cells transformed with pMK22/ΔAKI (clone ΔAKI 1a) grown on either glucose- (g) or oleic acid- (a) containing medium, Panel B: Lysates from cells transformed with pMK22/AQI grown on oleic acid-containing medium. Panels C and D: Lysates from cells transformed with pMK22/GKI and pMK22/AKX respectively, grown on oleic acid-containing medium.

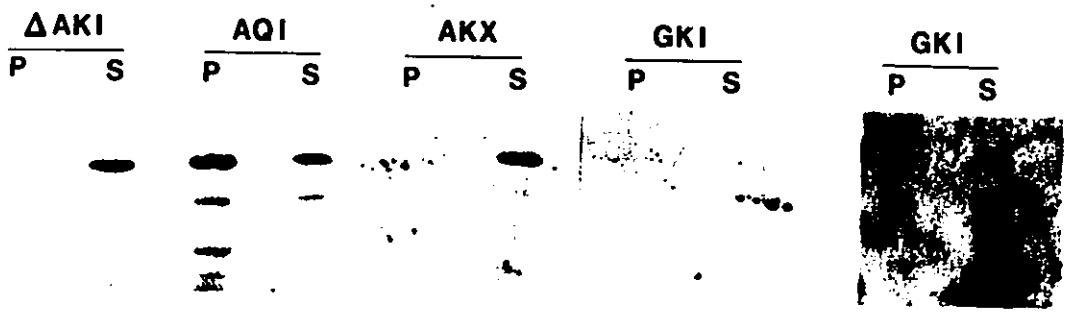


Figure 3.5.2.3 Western blot analysis of carboxy-terminal mutants of HDE expressed in *C. albicans*. The 20 kgP (P) and 20 kgS (S) fractions were prepared and resolved by SDS-PAGE. HDE was identified by western blotting using antiserum specific for HDE. 100 μg of protein was loaded in lanes P. The equivalent cellular fraction of 20 kgS was loaded in lanes S. Lanes ΔAKI, AQI, AKX, GKI correspond to carboxy-terminal tripeptide mutants ΔAKI, AQI, AKX, GKI, respectively, as described in the text. The blot at the extreme right was exposed for 5 times longer than those to the left.

Ile is required for targeting HDE to peroxisomes.

As discussed in section 1.7.5, Gould *et al.* (1989) have proposed, through analysis of other peroxisomal proteins and mutagenesis of luciferase, a consensus tripeptide PTS having the structure: Ser/Ala/Cys - Lys/Arg/His - Leu - CO_2H. The Ala-Lys-Ile tripeptide at the carboxy-terminus of HDE, although similar, does not fit into the consensus PTS as described by Gould. In fact, mutagenesis studies indicated that Ile could not substitute for Leu at the carboxy terminus of luciferase in the context of a Ser-Lys-Ile tripeptide (Gould *et al.*, 1989). This evidence suggests that HDE contains a similar, but distinct, form of the tripeptide PTS.

3.5.3 Substitutions within the tripeptide PTS

Analysis of the proposed consensus PTS suggests that a small uncharged amino acid is required at position -3 (Ser, Ala, or Cys) and that a basic amino acid is required at position -2 (Lys, His, or Arg) (Gould *et al.*, 1989). To further test this hypothesis, we created mutations within the HDE carboxy-terminal tripeptide which changed Ala→Gly at position -3 and Lys→Gln at position -2 (GKI and AQI, respectively). The mutant HDE proteins GKI and AQI were generated in the same manner as the ΔAKI and AKX mutant forms of HDE in pHDE/3T. The sequences of these mutant forms of *HDE* are shown in Fig. 3.5.2.1. The genes were subcloned into pMK22 in the same manner and orientation as the wild-type gene to generate pMK22/GKI and pMK22/AQI. *C. albicans* cells were transformed with the plasmids, integrants were selected, and the

mutants were expressed in oleic acid-containing medium (Fig. 3.5.2.2). The subcellular location of HDE in clones GKI-11 and AQI-11 was determined. HDE mutants GKI and AQI were both detected within the 20 kgP fraction, suggesting that these substitutions could still function to target HDE to peroxisomes (Fig. 3.5.2.3, panels GKI and AQI, respectively). GKI was expressed at low levels, requiring a longer exposure of the blot (Fig. 3.5.2.3, extreme right). The copy number of the integrated gene may reflect the differing levels of expression in each case (section 3.4); however, the copy number has not been determined. The specific organellar fraction to which these proteins were targeted was identified as peroxisomal by sucrose and Nycodenz density gradient centrifugation (section 3.4.3) followed by SDS-PAGE and western blot analysis (Fig. 3.5.3.1). Panel B shows a Coomassie-stained SDS-polyacrylamide gel of peak peroxisomal fractions. The arrowhead indicates the position of HDE. It is evident that peroxisomes isolated from cells transformed with wild-type HDE (WT) and mutants AQI and GKI contained an additional polypeptide band at molecular mass 100 kDa corresponding to HDE (lanes AQI and GKI). The identity of this polypeptide band as HDE was confirmed immunologically by western blot analysis (Panel A). The faster migrating immunoreactive bands are due to breakdown of HDE (section 3.4.3).

The Coomassie-stained gel also showed the relative levels of wild-type HDE and mutants. There was no overrepresentation of these polypeptides in the peroxisomal fractions. Fractionations were repeated on Nycodenz gradients (not shown) so as to band peroxisomes at a density different from that on sucrose gradients (1.19 $g \cdot cm^{-3}$ *versus*

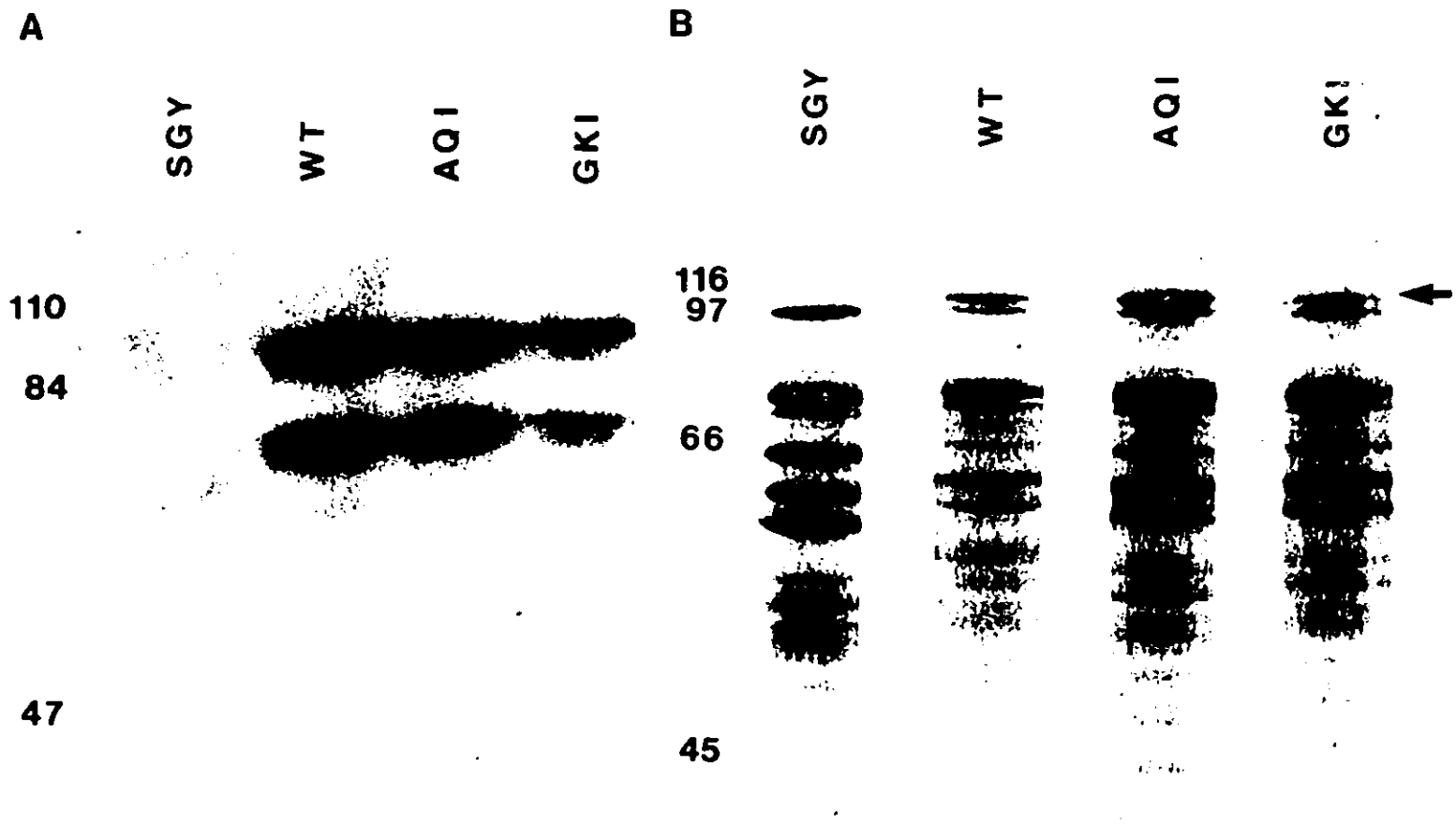


Figure 3.5.3.1 Mutants AQI and GKI are peroxisomal. Western blot analysis (panel A) and Coomassie-stained SDS-PAGE analysis (panel B) of purified peroxisomal fractions (20 μg of protein) containing wild-type *HDE* and *HDE* mutants expressed in *C. albicans*. SGY, non-transformed *C. albicans* SGY 243; WT, wild-type *HDE* expressed in SGY 243; AQI, GKI mutant forms of *HDE* expressed in SGY 243. Arrow denotes the migration of HDE.

1.23 $g \cdot cm^{-3}$, respectively) due to differential uptake of Nycodenz and sucrose by peroxisomes (Nuttley *et al.*, 1990). Wild-type HDE and mutants localized to the 20 kgP were again found exclusively in peroxisomal fractions indicating the peroxisomal targeting of wild-type HDE, GKI, AQI, H1/DO and K1/DO. Wild-type HDE and mutants which were detectable in peroxisomal fractions could not be detected in mitochondrial fractions. A summary of the subcellular locations of HDE mutants expressed in *C. albicans* is presented in Fig. 3.5.3.2.

3.5.4 Protease protection

The specific intraperoxisomal location of wild-type HDE and a selected mutant, AQI, was confirmed by protease protection experiments. Peroxisomes isolated from *C. albicans* harbouring wild-type HDE and mutant AQI were treated with 2, 10 and 20 μg of trypsin and 20 μg of thermolysin per 100 μg of purified peroxisomal protein in the presence or absence of 1% (w/v) Triton X-100 and 1% deoxycholate. Fig. 3.5.4.1 shows the Coomassie-stained gels of the purified peroxisomes and the western blots detecting HDE. Wild-type HDE (panels A and B), AQI (panels C and D), and endogenous peroxisomal proteins were protected from external protease in the absence of detergent but were proteolytically digested in the presence of detergent. Therefore wild-type HDE and mutant AQI are localized to the peroxisomal matrix.

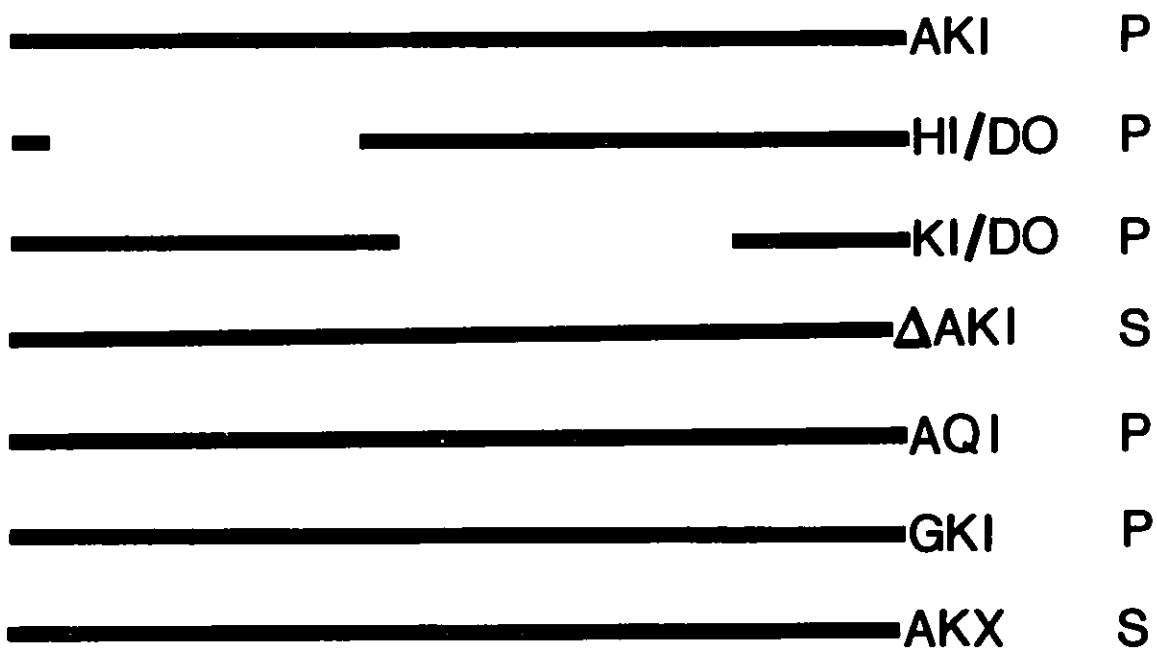


Figure 3.5.3.2 Summary of HDE mutants and their subcellular locations in *C. albicans*. P, peroxisomal; S, 20 kgS fraction (cytosol). AKI, wild-type HDE, 906 amino acids; H1/DO, deletion of amino acids 38-353; K1/DO, deletion of amino acids 395-731; ΔAKI, deletion of carboxy terminal amino acids Ala-Lys-Ile; AQI, replacement of Lys at 905 (-2) with Gln; GKI, replacement of Ala at 904 (-3) with Gly; AKX, deletion of Ile at 906 (-1).

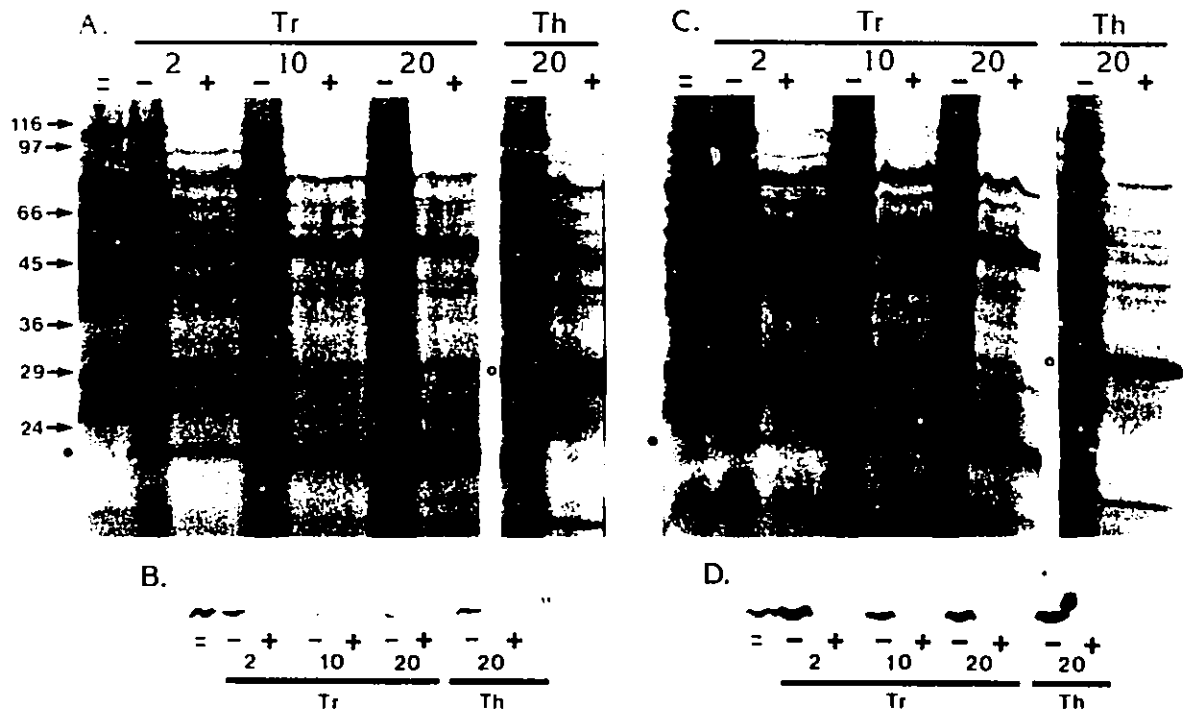


Figure 3.5.4.1 Protease protection of wild-type HDE and mutant AQI in peroxisomes of *C. albicans*. Peroxisomes were treated with 2, 10 and 20 μg of trypsin (lanes Tr) or 20 μg of thermolysin (lanes Th) per 100 μg of peroxisomal protein in the absence (lanes -) or presence (lanes +) of 1% (w/v) Triton X-100 and 1% deoxycholate at 0°C for 30 min. Lanes =: mock treated peroxisomes. Panel A: Coomassie-stained SDS-polyacrylamide gel of peroxisomes isolated from *C. albicans* transformed with wild-type *HDE* and treated as above. 100 μg of starting peroxisomal protein was loaded in each lane. The downward pointing arrow denotes the position of HDE. Panel B: Western blot analysis of the SDS-polyacrylamide gel in panel A treated with anti-HDE serum. Panel C: Coomassie-stained SDS-polyacrylamide gel of peroxisomes isolated from *C. albicans* transformed with mutant *HDE* (AQI) and treated as above. 100 μg of starting peroxisomal protein was loaded in each lane. The downward pointing arrow denotes the position of AQI. Panel D: Western blot analysis of the SDS-polyacrylamide gel in Panel C treated with anti-HDE serum. Numbers at left of figure represent molecular weight markers (in kDa). The closed circle denotes the migration of trypsin. The open circle denotes the migration of thermolysin.

3.5.5 Summary

Previous experiments addressing the targeting of peroxisomal fatty acyl-CoA oxidase of *C. tropicalis* demonstrated the presence of a bipartite peroxisomal targeting signal (Small *et al.*, 1988), distinct from the C-terminal tripeptide PTS of mammalian and insect peroxisomal proteins. Neither of the topogenic sequences identified by Small *et al.* contains a Ser-Lys-Leu PTS or a variant thereof as described by Gould *et al.* (1989) (Borst, 1989). In addition none of the cloned genes encoding peroxisomal proteins in *C. tropicalis* (HDE, Nuttley *et al.*, 1988; catalase, Murray and Rachubinski, 1987b; Okada *et al.*, 1987; the three fatty acyl-CoA oxidases: *POX-2*, Okazaki *et al.*, 1987; *POX-4(a)*, Okazaki *et al.*, 1986; Murray and Rachubinski, 1987a; and *POX-5*, Okazaki *et al.*, 1986; *POX-18*, Szabo *et al.*, 1989 and isocitrate lyase, Atomi *et al.*, 1990) code for a consensus carboxy-terminal PTS. This suggested that *C. tropicalis* may possess a unique mechanism for targeting proteins to peroxisomes and/or that different receptors exist which recognize different PTSs. The results presented above show that at least one protein of *C. tropicalis* is targeted to peroxisomes by a carboxy-terminal PTS in a manner similar to that found for mammalian and insect peroxisomal proteins.

3.6 Expression of *HDE* in *S. cerevisiae*

To determine if the PTS in HDE is operable in *S. cerevisiae*, the gene encoding HDE was inserted into the plasmid YCp50 for expression in *S. cerevisiae*.

3.6.1 Expression in *S. cerevisiae* of heterologous genes encoding peroxisomal proteins

Several others have attempted to study peroxisomal biogenesis through the expression of genes encoding peroxisomal proteins in yeasts. Unfortunately, efforts to localize PTSs using this approach have met with limited success. Distel *et al.* (1987) expressed *H. polymorpha* alcohol oxidase in *S. cerevisiae* to determine if this heterologous expression system would be useful for the identification of the alcohol oxidase PTS. *S. cerevisiae* lacks endogenous alcohol oxidase, therefore the protein is not essential and could be detected by anti-alcohol oxidase antibodies without the interference of endogenous proteins. The gene was placed behind a constitutive *S. cerevisiae* promoter. Alcohol oxidase was synthesized in the heterologous expression system and detected by western blotting of yeast lysates. When the gene was placed on a high copy 2 μm circle-based plasmid, the level of alcohol oxidase was 5 to 7 times higher than when the gene was expressed from a single copy CEN-based plasmid. Analysis by electron microscopy of the ability of the cells to import alcohol oxidase into peroxisomes demonstrated that when expressed at normal levels (from the CEN-based plasmid), alcohol oxidase was imported into peroxisomes. This suggested that the targeting signal within alcohol oxidase could be recognized by the import machinery of *S. cerevisiae* peroxisomes. However, the protein was not assembled into its octameric functional form. High level expression of alcohol oxidase from the 2 μm-based plasmid resulted in the formation of inclusion bodies within the cytoplasm and no import of

alcohol oxidase.

Similar results were obtained by expression of DHFR-catalase A fusion proteins in *S. cerevisiae* (Binder *et al.*, 1990). Small fusions of the amino-terminal 57 amino acids of catalase A to DHFR or of the carboxy-terminal 15 amino acids of catalase A to DHFR were not directed to peroxisomes as detected by subcellular fractionation. This indicated that the targeting information within catalase A was neither amino- nor carboxy-terminal. When larger fusions were tested, large protein aggregates were formed. The hybrids, found as aggregates, were originally thought to be peroxisomal because they copurified with peroxisomes by subcellular fractionation; however, immunofluorescence and immunoelectron microscopy demonstrated the presence of inclusion bodies in the cytosol, not surrounded by membranes. The formation of inclusion bodies was attributed to oversynthesis of the fusion proteins by virtue of their expression from the multi-copy plasmid YEp352. Approximately 1-2% of the total protein content in transformants was attributed to highly expressed fusion proteins. Expression of *C. tropicalis* catalase deletion constructs in *S. cerevisiae* from the multi-copy plasmid YEp13 also caused the formation of aggregated proteins. In this study, *HDE* was expressed in *S. cerevisiae* on a single copy plasmid (YCp50) behind its own inducible promoter. This was used to circumvent problems associated with the potential saturation of the peroxisomal import apparatus. The maintenance of the *HDE* promoter permits the coinduction of *HDE* expression and the proliferation of peroxisomes in *S. cerevisiae* (Veenhuis *et al.*, 1987; Sloots *et al.*, 1991). It was hypothesized that this

approach would prevent accumulation of HDE in the cytosol prior to the induction of peroxisomes (and import machinery), which may contribute to protein inclusion body formation.

3.6.1.1 Subcloning of *HDE* and mutants into YCp50

Wild-type and mutant forms of *HDE* were sucloned into the unique *SalI* site of YCp50 to generate the recombinant plasmids YCp-HDE, YCp-H1/DO, YCp-K1/DO, YCp-AKX, YCp-GKI and YCp-AQI. The recombinants were selected by tetracycline sensitivity, ampicillin resistance and restriction mapping. The recombinant plasmid YCp-HDE containing the wild-type *HDE* gene is shown in Fig. 3.6.1.1. All recombinants selected for further analysis contained the *HDE* gene in the same orientation. The single-copy plasmid YCp50 was expressly chosen for expression of the *HDE* gene behind its own promoter so as to coinduce peroxisomes and the expression of *HDE* and to provide normal levels of the single copy *HDE* gene (Sloots *et al.*, 1991). In addition, the orientation of the *HDE* gene in all recombinants was the same as the *URA3* gene. Insertion of the gene in the opposite orientation resulted in an increase in the level of *HDE* expression. The use of clones expressing HDE at normal levels (single copy, single promoter) reduces the probability of overexpression of HDE and the formation of inclusion bodies.

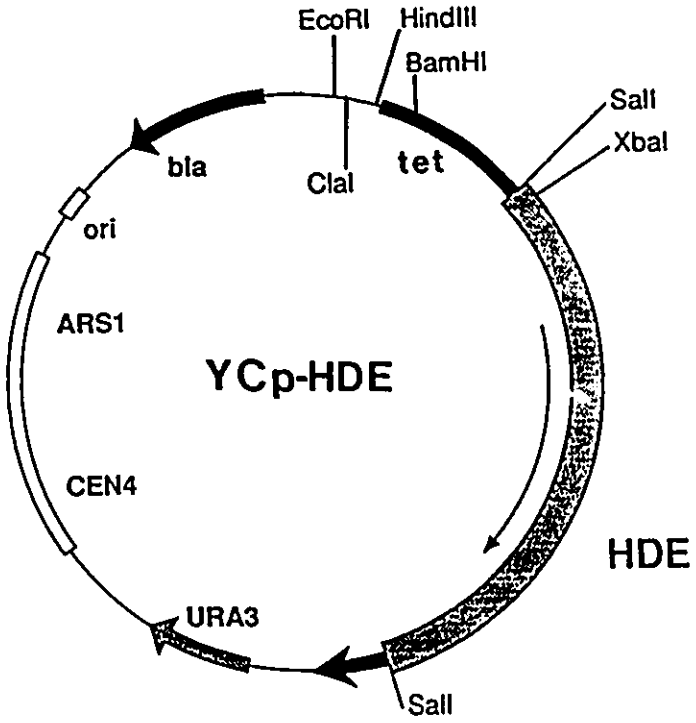


Figure 3.6.1.1.1 Recombinant plasmid YCp-HDE. The plasmid pGEM/3T and YCp50 were digested with *SalI*. The 4.2 kb fragment bearing the *HDE* gene was ligated into YCp50 disrupting the tetracycline resistance marker. The orientation of the *HDE* gene is indicated by the arrow. Mutant forms of *HDE* were ligated into YCp50 in the same manner.

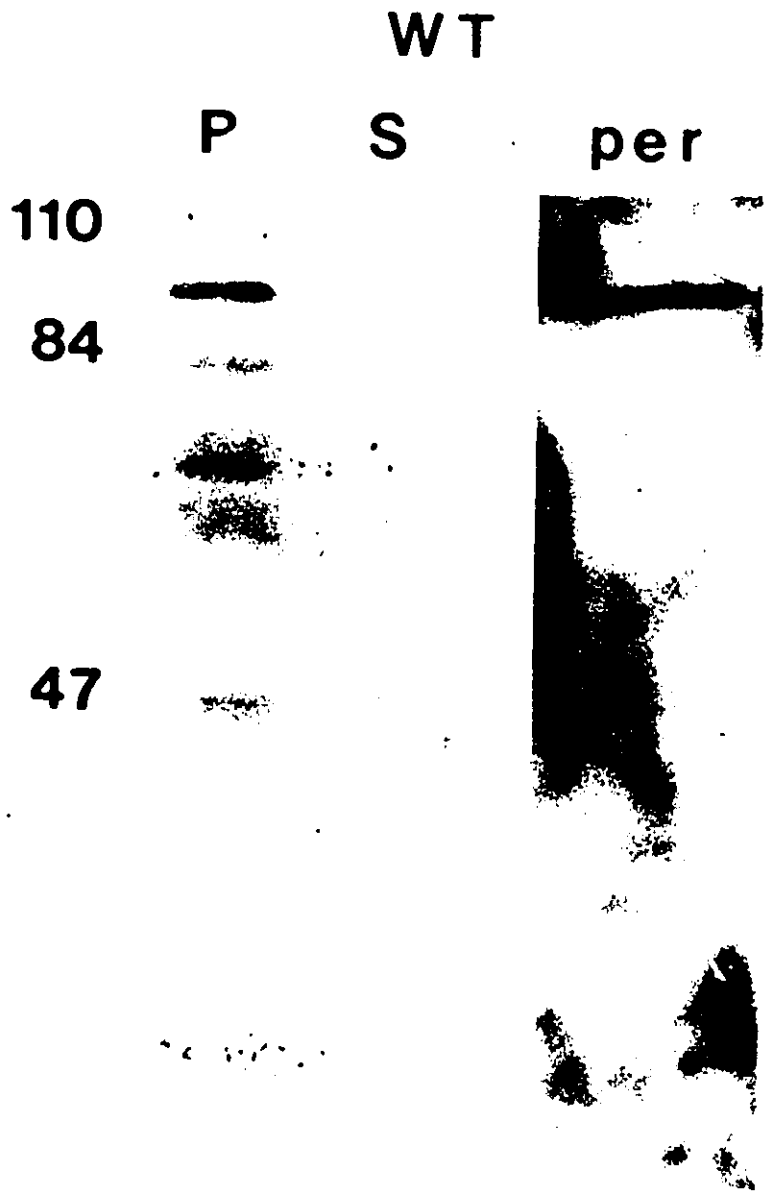


Figure 3.6.2.1 Localization of wild-type HDE to peroxisomes of *S. cerevisiae*. The 20 kgP (P), 20 kgS (S) and peroxisomal (per) fractions were prepared and resolved by SDS-PAGE. HDE was identified by western blotting using anti-HDE serum. 100 μg of the 20 kgP fraction (lane P) and an equivalent cellular fraction of the 20 kgS (lane S) were loaded. 40 μg of peroxisomal protein isolated by Nycodenz density gradient centrifugation was loaded in lane P. Numbers at the left of the figure represent the migrations of the molecular weight markers (kDa).

3.6.2 Subcellular location of wild-type HDE synthesized in *S. cerevisiae*

The recombinant plasmids were used to transform *S. cerevisiae* DL-1 to uracil prototrophy. HDE was detected by western blots. Cells transformed with control plasmids YEp13 or YCp50 did not yield an immunoreactive polypeptide within the same molecular weight region as HDE (Fig. 3.6.4.1, YEP; Sloots *et al.*, 1991). Fractionation of cells transformed with wild-type *HDE* demonstrated that HDE was pelletable (Fig. 3.6.2.1, lane P) and colocalized with peroxisomes isolated on Nycodenz gradients (lane per). HDE was not detectable in the mitochondrial fraction. Therefore the PTS of HDE is functional in *S. cerevisiae*.

3.6.3 Protease protection

The intraperoxisomal location of HDE in *S. cerevisiae* was confirmed by protease protection experiments (Fig. 3.6.3.1). Trypsin was added to purified peroxisomes at a concentration of 5 μg per 100 μg of purified peroxisomal protein in the presence or absence of 1% (w/v) Triton X-100 and 1% deoxycholate. Peroxisomal proteins were degraded in the presence of protease and detergent as detected by the Coomassie-stained gel (panel A, lane +) but were protected from protease in the absence of detergent (lane -). HDE was monitored in this assay by western blotting. HDE was protected from external protease in the absence of detergent (panel B, lane -) but not in the presence of detergent (lane +), in the same fashion as endogenous peroxisomal proteins. These results demonstrate the matrix location of heterologously expressed HDE

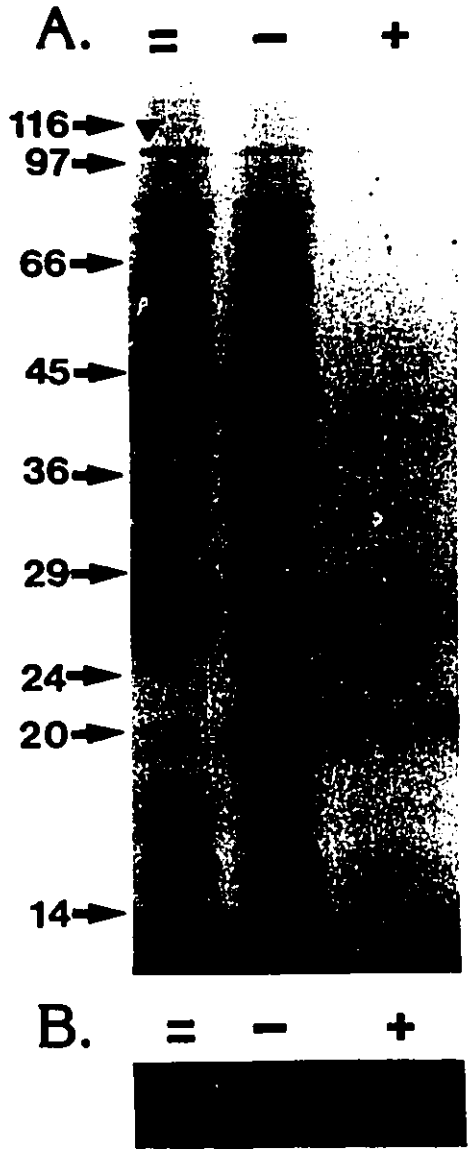


Figure 3.6.3.1 Wild-type HDE is imported into peroxisomes in *S. cerevisiae* as assayed by protease protection. Peroxisomes were treated with 5 μg of trypsin per 100 μg of peroxisomal protein in the absence (lane -) and in the presence (lane +) of 1% (w/v) Triton X-100 and 1% deoxycholate at 0°C for 20 min. Lanes =, mock treated peroxisomes. Panel A: Coomassie-stained SDS-polyacrylamide gel of peroxisomes isolated from *S. cerevisiae* transformed with YCp-HDE. Panel B: Western blot analysis of panel A using anti-HDE serum. Numbers at the left of the figure represent molecular weight standards (kDa).

in peroxisomes of *S. cerevisiae*.

3.6.4 Subcellular location of mutant forms of HDE synthesized in *S. cerevisiae*

The subcellular location of mutant forms of HDE was also determined (Fig. 3.6.4.1). Deletion of the carboxy-terminal tripeptide Ala-Lys-Ile of HDE caused mislocalization of the mutant to the cytosol in *S. cerevisiae* as determined by western blot analysis of 20 kgS and 20 kgP fractions (Fig. 3.6.4.1, ΔAKI). Similarly, deletion of only the last amino acid (AKX) also caused mistargeting of HDE (lanes AKX). These results are the same as those found for *C. albicans*. An additional mutant in which the amino acid (Ala) at position -3 was deleted (XKI), which changes the three carboxy-terminal amino acids to Lys-Lys-Ile, also abolished targeting of HDE (lanes XKI), demonstrating the importance of the intact tripeptide. In addition, K1/DO and H1/DO were localized to the peroxisomal fraction isolated from *S. cerevisiae*. Curiously, the mutants AQI and GKI, which were peroxisomal in *C. albicans*, were also not peroxisomal when expressed in *S. cerevisiae* (lanes AQI and GKI). The integrity of the peroxisomes during the isolation in these experiments was established by catalase activity in the 20 kgP (greater than 50%), demonstrating that the localization of these mutants to the cytoslic fraction was due to mislocalization and not to complete peroxisomal breakage during the isolations. Although the carboxy-terminal substitutions AQI and GKI were not imported into peroxisomes in *S. cerevisiae*, the results from HDE targeting in *S. cerevisiae* and *C. albicans* indicate that the tripeptide Ala-Lys-Ile is an essential feature

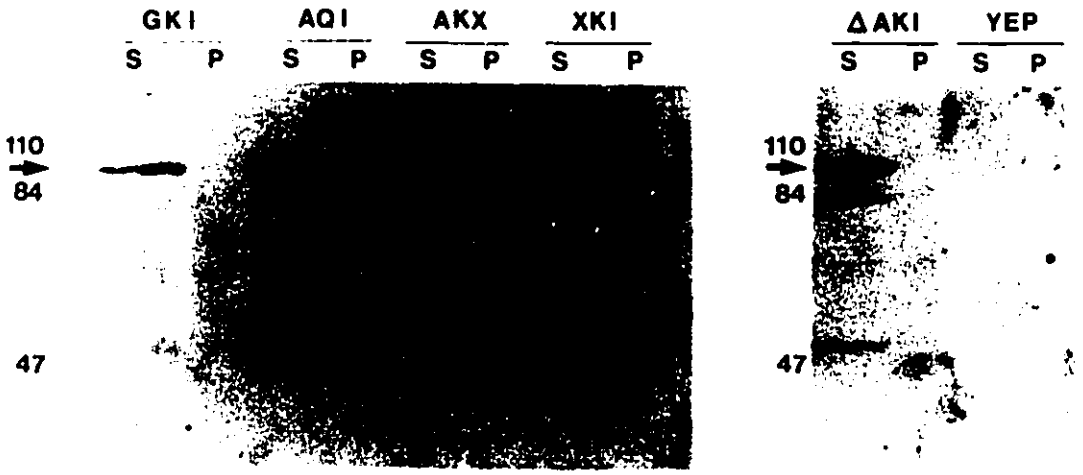


Figure 3.6.4.1 Carboxyl-terminal mutant forms of HDE are not targeted to peroxisomes in *S. cerevisiae*. The 20 kgP (P) and 20 kgS (S) fractions of cells transformed with carboxyl-terminal mutant forms of *HDE* were prepared and analysed by western blotting using anti-HDE serum. 100 μg of protein was loaded in lanes P, and an equivalent cellular fraction of the 20 kgS fraction was loaded in lanes S. The carboxyl-terminal mutants are indicated above their respective lanes. YEP corresponds to *S. cerevisiae* transformed with control plasmid YEp13, not containing the HDE gene. The molecular weight standards (kDa) are indicated at the left of each panel.

of the PTS.

3.6.5 CATAKI

The critical test for the action of Ala-Lys-Ile as a PTS is to determine if the tripeptide can target passenger proteins to peroxisomes. *E. coli* CAT fused to the Ser-Lys-Leu tripeptide (CAT-SKL) is imported into peroxisomes when expressed in CV-1 cells (Gould *et al.*, 1989). Ala-Lys-Ile was fused to *E. coli* CAT and expressed in *S. cerevisiae* to assay the sufficiency of the Ala-Lys-Ile tripeptide PTS. The gene encoding *E. coli* CAT (flanked by two *Bam*HI sites) was cloned into the *Bam*HI site in the multiple cloning site of plasmid pGEM7Zf(+) to generate pGEM/CAT. Single-stranded DNA was isolated, and the codons encoding the Ala-Lys-Ile tripeptide of HDE were added to the *E. coli CAT* gene immediately upstream of the stop codon by oligonucleotide-directed site-specific mutagenesis (to generate pGEM/CATAKI). The HDE promoter in pGEM/3T was similarily mutated by site-directed mutagenesis for cloning of the *E. coli CAT* and *CATAKI* genes behind this promoter to allow for oleic acid induction of *CATAKI* and the wild-type *E. coli* CAT. A *Bgl*II site was inserted into the 5' portion of the *HDE* gene by insertion of nucleotides GATCT between nucleotides 45 and 46 (Fig. 3.3.4.1). Digestion of the plasmid pGEM/3T with *Bgl*II resulted in the removal of the coding sequence of the *HDE* gene maintaining intact 5' and 3' noncoding regions of the gene to give plasmid pHDE. *E. coli CAT* and *CATAKI* were removed from pGEM/CAT and pGEM/CATAKI by *Bam*HI digestion and ligated into the complementary ends

generated by digestion of pHDE with *Bgl*II. The resulting hybrid genes were subloned into YCp50 in the same manner as for the *HDE* gene (section 3.6.1.2) to produce YCp/CAT and YCp/CATAKI. *S. cerevisiae* DL-1 cells were transformed with plasmids YCp/CAT and YCp/CATAKI. The subcellular locations of *E. coli* CAT and CATAKI were determined. Both the wild-type *E. coli* CAT and the CATAKI fusion protein were localized to the cytosolic fractions (Fig. 3.6.5.1). It therefore must be stated that the presence of Ala-Lys-Ile at the carboxy-terminus is not sufficient to direct *E. coli* CAT to peroxisomes in *S. cerevisiae*.

However, it is premature to disqualify the Ala-Lys-Ile tripeptide as a *bona fide* PTS based on this observation alone. The fusion of Ser-Lys-Leu to the carboxy-terminus of mouse dihydrofolate reductase (DHFR-SKL) is also not sufficient to direct the passenger protein to peroxisomes in *S. cerevisiae* (S. Subramani, personal communication). This is in spite of recent evidence indicating that Ser-Lys-Leu-COOH operates as a PTS in *S. cerevisiae*. *S. cerevisiae* trifunctional enzyme (Fig. 3.2.8.2.1) and a number of other *S. cerevisiae* peroxisomal proteins terminate with Ser-Lys-Leu (section 3.8). In addition, firefly luciferase is directed to peroxisomes in *S. cerevisiae* (Gould *et al.*, 1990b). The inabilty of the DHFR-SKL and CATAKI fusion proteins to be directed to *S. cerevisiae* peroxisomes suggests that the context of PTS presentation is important. Additional evidence for the importance of the context of presentation of the PTS has been shown for the the Ser-Lys-Leu PTS, since deletion of the amino-terminal 58 amino acids of luciferase, which maintains an intact Ser-Lys-Leu tripeptide at the

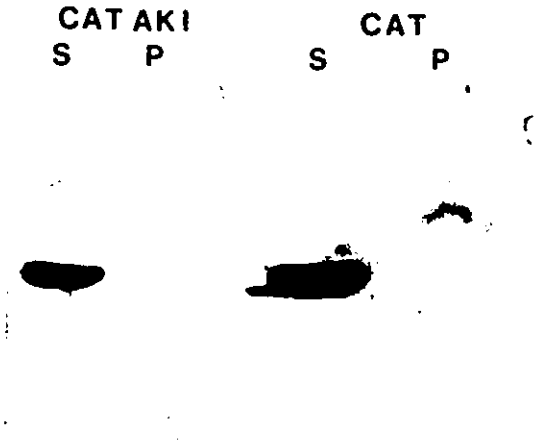


Figure 3.6.5.1 *E. coli* CAT and CATAKI are not directed to peroxisomes in *S. cerevisiae*. The 20 kgP (P) and 20 kgS (S) fractions of cells transformed with YCp/CAT and YCp/CATAKI were prepared and analysed by western blotting using anti-CAT serum (lanes CAT and CATAKI, respectively). 100 μg of protein was loaded in lanes P, and an equivalent cellular fraction of the 20 kgS fraction was loaded in lanes S.

carboxy terminus of the protein, renders luciferase cytosolic in CV-1 cells (Gould *et al.*, 1989). Moreover, some fusions of *E. coli* CAT to human catalase fragments containing the catalase PTS remain cytosolic (Gould *et al.*, 1988). It is also possible that CATAKI oligomerization in the cytosol prevents its import into peroxisomes of *S. cerevisiae*. Chloramphenicol acetyltransferase enzyme assays on yeast lysates prepared from cells harbouring YCp-CAT and YCp-CATAKI indicate that both cytosolic forms are enzymatically active. Thus the fusion of Ala-Lys-Ile to *E. coli* CAT does not significantly alter its conformation or its ability to form an enzymatically active tetramer. Results consistent with the suggestion that oligomerization prevents import into peroxisomes have been obtained with yeast apoiso-1-cytochrome *c* (containing a mitochondrial targeting signal) fused to *E. coli* CAT. This fusion was not imported into mitochondria when expressed in *S. cerevisiae* until a deletion was made in the *E. coli* CAT protein to disrupt oligomerization (Nye and Scarpula, 1990). To investigate the possibility that oligomerization might impede import, the nucleotides encoding the oligomerization domain of *E. coli* CAT were deleted from plasmids pGEM/CAT and pGEM/CATAKI. The deletions were subIconed into pHDE and then into YCp50 as above. However, attempts to express these constructs were unsuccessful. It is likely that deletion of these domains resulted in an unstable protein that was degraded shortly after expression. This has not been investigated further.

The sufficiency of Ala-Lys-Ile-COOH to act as a PTS may be better assayed by the use of an alternative (monomeric) reporter protein. In addition it may be

necessary to insert a spacer region between the PTS and the reporter protein to separate the tripeptide from the rest of the (folded) protein.

3.6.6 Summary

The carboxy-terminal tripeptide Ala-Lys-Ile is required for correct targeting of HDE to peroxisomes in both *C. albicans* and *S. cerevisiae*. To investigate further the essential properties of the HDE tripeptide sequence, substitution mutants were constructed to change the Ala→Gly at position -3 and the Lys→Gln at position -2. These substitutions were chosen to investigate the role of a small amino acid at position -3 and to determine if the basic amino acid at position -2 could be substituted with an uncharged amino acid containing an amide. The HDE mutants with these changes (GKI and AQI) are both efficiently targeted to peroxisomes in *C. albicans*, but not in *S. cerevisiae* (sections 3.5 and 3.6, respectively). The fact that GKI and AQI are correctly targeted to peroxisomes in *C. albicans*, but AKX is not, allows for the formal possibility that all that is required for targeting HDE to peroxisomes in *C. albicans* is the terminal Ile. This is unlikely, given the importance of a tripeptide PTS in peroxisomal targeting in mammalian cells (Gould *et al.*, 1989) and the requirement of the complete, intact tripeptide in targeting HDE to peroxisomes of *S. cerevisiae* (XKI is not targeted to peroxisomes). Taken together, these data suggest that *C. tropicalis*, *C. albicans* and *S. cerevisiae* can recognize a common peroxisomal targeting motif (all import wild-type HDE); however, evolutionary divergence (Barns *et al.*, 1991) does not permit *C.*

albicans and *S. cerevisiae* to recognize the same variations in this motif.

Luciferase has been shown to be imported into peroxisomes of *S. cerevisiae* (Gould *et al.*, 1990b), indicating that, in addition to Ala-Lys-Ile, this yeast recognizes PTSs consisting of the tripeptide Ser-Lys-Leu, although the absolute identity of the luciferase PTS functioning in *S. cerevisiae* has not been determined. An antiserum directed against a peptide containing Ser-Lys-Leu at its carboxy terminus reacts specifically with peroxisomal proteins of *S. cerevisiae* and of mammalian cells but not with peroxisomal proteins from *C. tropicalis* (see section 3.7)

Gould *et al.* (1989) have also shown that substitution of Ile for Leu in the Ser-Lys-Leu PTS could not support targeting of luciferase to peroxisomes of CV-1 cells. Whether this represents differences in the yeast and mammalian systems or whether the Ala at position -3 plays a compensatory role for the Ile at position -1 remains to be shown (but see section 3.7). In this regard, investigations involving import of HDE into peroxisomes of CV-1 cells should prove interesting. It is noteworthy that not all combinations of the PTS as identified by Gould *et al.* (1989) are imported with equal efficiency in CV-1 cells (Swinkels and Subramani, 1990), suggesting that each amino acid in the PTS cannot be considered independently and that its context within the surrounding amino acids is important.

The HDE mutant AQI demonstrates the ability of an uncharged, amide-containing amino acid to substitute for Lys within the PTS in *C. albicans*. It has been shown that fusion of the last 27 amino acids of human catalase, including Ala-Asn-Leu

at the carboxy-terminus, to *E. coli* CAT targets the fusion protein to peroxisomes in CV-1 cells (Gould *et al.*, 1988). Perhaps the Asn at position -2 in human catalase acts in the same capacity as Gln at position -2 in the AQI mutant.

The results presented in 3.5 and 3.6 are the first detailed analysis of a yeast PTS. They demonstrate the existence of an additional PTS tripeptide which is operable in yeast and suggest that there may be other peroxisomal proteins that do not contain the canonical PTS (and variations) as described by Gould *et al.* (1989) but which are still targeted to peroxisomes by way of an as yet unrecognized carboxy-terminal tripeptide.

It is unknown whether a tripeptide PTS can act at an internal location within a protein. The tripeptide Ala-Lys-Ile is present at internal locations in other peroxisomal proteins for which a targeting signal has not yet been identified, including *H. polymorpha* dihydroxyacetone synthase, which does not contain any previously identified PTS variant (Janowicz *et al.*, 1985; Gould *et al.*, 1989) and *Drosophila melanogaster* catalase (Orr *et al.*, 1990). The Ala-Lys-Ile tripeptide (or variants described herein) is not present in either of the topogenic PTSs described for *C. tropicalis* fatty acyl-CoA oxidase (Small *et al.*, 1988; Borst, 1989). A search of the current protein data bases (SWISS PROT; PseqIp; Protein; New; CAN/SND Molecular Biology Data Base System, NRCC) using the PSQ package (version 4.2/4.7; see also section 3.7) revealed only one protein terminating with Ala-Lys-Ile which is not targeted to either mitochondria or the secretory pathway: *Drosophila virilis* Cu-Zn superoxide dismutase (SOD) (Kwiatowski and Ayala, 1989). Whether this protein is peroxisomal is unknown; however, recent evidence

suggests that a human version of SOD is peroxisomal, and the carboxy terminus contains the tripeptide Ser-Arg-Leu which may be important for its targeting to peroxisomes (Keller *et al.*, 1991b). The ability of this tripeptide to act as a PTS in mammalian cells remains to be investigated. The presence of this tripeptide PTS in other peroxisomal proteins is examined in sections 3.7 and 3.8.

The ability of GKI and AQI mutants to be imported into peroxiomes in *C. albicans* does not demonstrate their existence in nature. Although at a qualitative level it appears that these mutants are imported into peroxisomes with reasonable efficiency, subtle differences in import efficiencies operating in nature may have profound effects on the maintenance of PTSs on an evolutionary time scale.

3.7 Anti-AKI antibodies[14]

C. tropicalis HDE contains a previously unidentified carboxy-terminal tripeptide PTS (sections 3.5 and 3.6; Aitchison *et al.*, 1991a). To investigate the possibility that this tripeptide may act as a general PTS for other proteins, anti-AKI antibodies were raised against the carboxy-terminal 12 amino acids of trifunctional enzyme and were used to search for this tripeptide in proteins of the yeasts *C. tropicalis*, *C. albicans*, *Y. lipolytica* and *S. cerevisiae* and of rat liver. This approach has been used previously with anti-SKL antibodies to investigate the presence of antigenically similar

[14]The experiments discussed in this section were carried out with Rachel Szilard and William Nuttley.

epitopes in peroxisomal, glyoxysomal and glycosomal proteins from mammalian cells, methylotrophic yeast and plant cells (Gould *et al.*, 1990a; Keller *et al.*, 1991a). In these studies anti-SKL antibodies identified a number of immunoreactive proteins in these microbodies but did not detect proteins in peroxisomes from *C. tropicalis* (Keller *et al.*, 1991a). This evidence, as well as the identification of the novel Ala-Lys-Ile-containing PTS in trifunctional enzyme (section 3.6), suggested that the peroxisomal proteins of *C. tropicalis* might contain an antigenically distinct form of the PTS.

Anti-AKI antibodies were generated in guinea pigs by the injection of the synthetic peptide AKI (Acetyl-CAIKLVGDKAKI-COOH) coupled through its amino terminus to keyhole limpet hemocyanin. The reaction of anti-AKI antibodies was detected by both western blotting and immunofluorescence. The reactions were compared to that obtained using anti-SKL antibodies. Anti-SKL antibodies were a gift from Dr. S. Subramani and were generated in rabbits against the peptide NH_2-CRYHLKPLQSKL-COOH (the carboxy-terminal 9 amino acids of rat acyl-CoA oxidase) linked to keyhole limpet hemocyanin (Gould *et al.*, 1990a).

3.7.1 Peroxisome isolations

Peroxisomal fractions were isolated from *C. tropicalis*, *C. albicans*, *Y. lipolytica* and *S. cerevisiae*. The profiles of the fractionations from these yeasts are shown in Fig. 3.7.1.1. In each case the purity of the peroxisomal fraction was greater than 95% as judged by marker enzyme analyses. The peak peroxisomal and

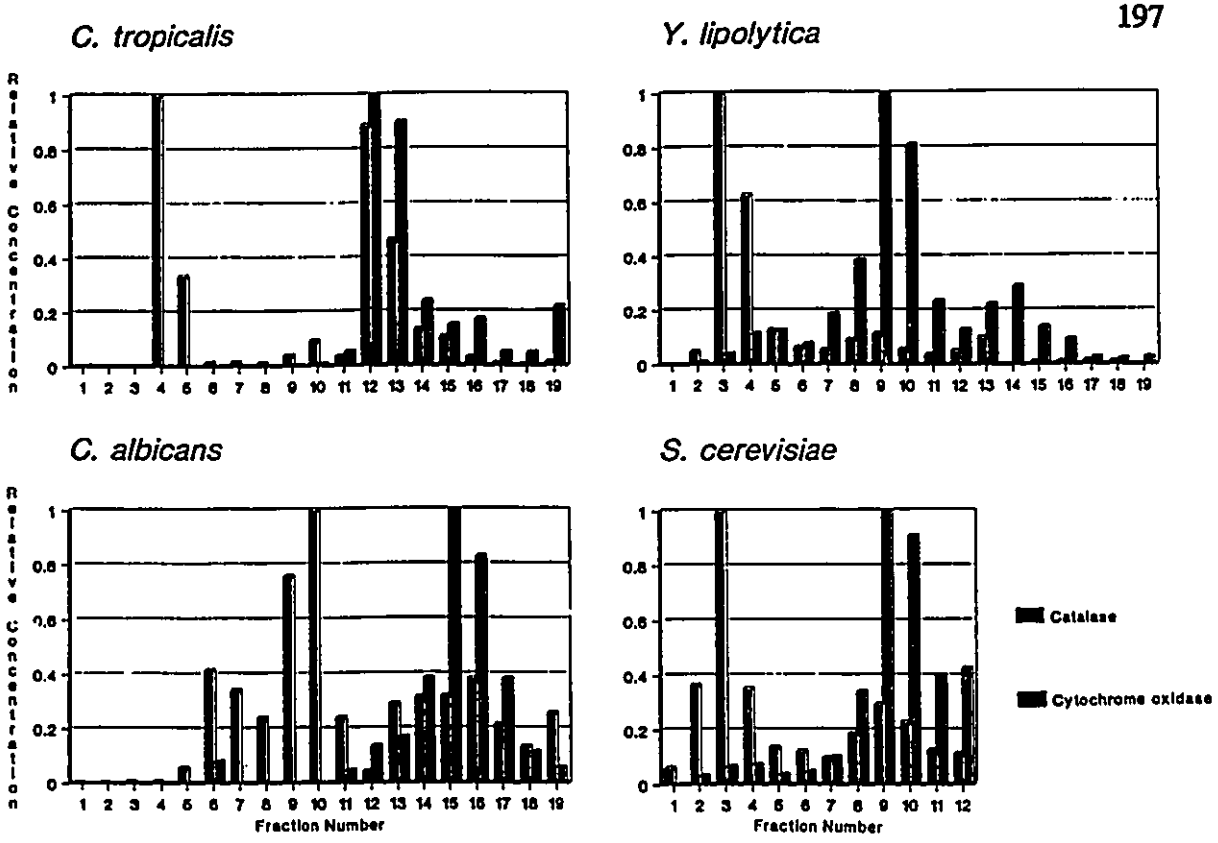


Figure 3.7.1.1 Subcellular fractionation profiles of *C. tropicalis*, *C. albicans*, *Y. lipolytica* and *S. cerevisiae*. Mitochondria and peroxisomes from *C. tropicalis*, *C. albicans*, *Y. lipolytica* and *S. cerevisiae* were purified from a 20,000 x g pellet by density gradient centrifugation. Fractions were collected from the bottom of the gradient (fraction 1) and analysed for catalase activity (peroxisomes) and cytochrome *c* oxidase activity (mitochondria). The activities of the peak fractions were normalized to one.

mitochondrial fractions were probed by western blotting with anti-AKI and anti-SKL antibodies. Rat liver peroxisomes were isolated previously by Dr. Andrea Bodnar.

3.7.2 Characterization of anti-AKI and anti-SKL reactive proteins in *C. tropicalis* and rat liver

Anti-AKI antibodies were used to probe purified peroxisomal proteins of *C. tropicalis* and rat liver by western blotting (Fig. 3.7.2.1). The Coomassie-stained polypeptide profiles of the isolated peroxisomes are shown in panel A. Anti-AKI antibodies reacted with at least 8 polypeptides of *C. tropicalis* peroxisomes in addition to HDE (100 kDa). These polypeptides have relative molecular masses of ≈69,000, 54,000, 38,000, 35,000, 33,000, 29,000, 27,000 and 18,000 (panel B, *C. trop.*, lane PER). The polypeptide at approximately 28 kDa could be resolved into two polypeptides (see Fig. 3.7.3.1. lane -). There was no reaction of the antibodies with mitochondrial proteins of *C. tropicalis* (panel B, *C. trop.*, lane MIT), or with mitochondrial (panel B, rat, lane MIT) or peroxisomal (panel B, rat, lane PER) proteins of rat liver. There was no specific cytosolic protein that reacted with anti-AKI antibodies. Immunoreactive proteins in the cytosolic fraction comigrated with immunoreactive peroxisomal proteins and were attributed to leakage from peroxisomes broken during homogenization (sections 3.5 and 3.6). There was no immunoreaction with preimmune serum.

Peroxisomal and mitochondrial fractions of *C. tropicalis* and rat liver were

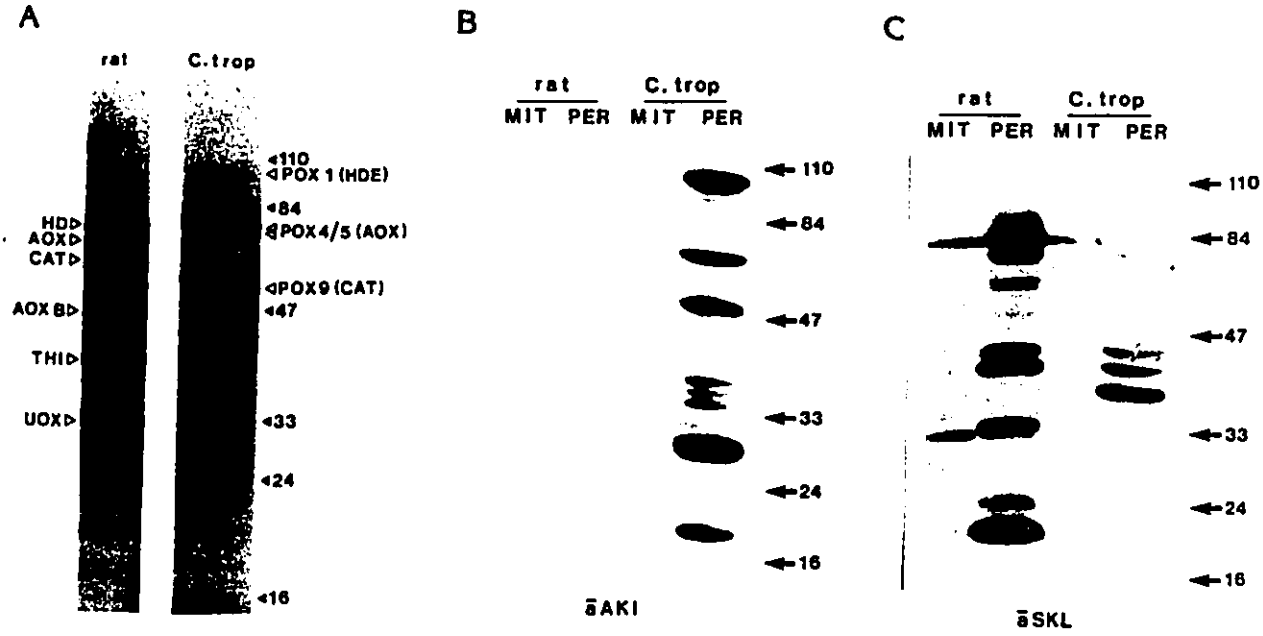


Figure 3.7.2.1 Western blot analysis of anti-AKI and anti-SKL reactive protiens in *C. tropicalis* and rat liver. Panel A: 100 μg of purified peroxisomal protein from rat liver or *C. tropicalis* were separated by SDS-PAGE and stained with Coomassie blue. The bands corresponding to major peroxisomal proteins are indicated: HD, bifunctional protein (hydratase-dehydrogenase); AOX, fatty acyl-CoA oxidase; CAT, catalase; AOX B, fatty acyl-CoA oxidase subunit B; THI, thiolase; UOX, urate oxidase; POX1 (HDE), trifunctionai enzyme (hydratase-dehydrogenase-epimerase); POX 4/5 (AOX), PXP-4 and PXP-5 subunits of fatty acyl-CoA oxidase; POX 9 (CAT), catalase. Panel B: 100 μg of mitochondrial protein (*lanes MIT*) or 100 μg of peroxisomal protein (*lanes PER*) from rat liver and *C. tropicalis* were probed with anti-AKI antibodies by western blotting. Panel C: A blot identical to that shown in *panel B* was probed with anti-SKL antibodies. The numbers at the right of each panel indicate the migrations of molecular weight standards (kDa).

also probed with antibodies directed against the carboxy-terminal PTS (Ser-Lys-Leu) of rat acyl-CoA oxidase (anti-SKL antibodies; Fig. 3.7.2.1, panel C). Consistent with the observations of Gould *et al.* (1990a), a number of rat liver peroxisomal proteins reacted with the anti-SKL antibodies (rat, lane PER). The immunoreactive proteins in the rat mitochondrial fraction have also been previously observed. The polypeptides of relative molecular masses ≈80,000 and 26,000 in the mitochondrial fraction were attributed to contamination with peroxisomes, but the presence of additional immunoreactive polypeptides in the mitochondrial fraction have not been explained (Gould *et al.*, 1990a; but see below; section. 3.7.7). Four peroxisomal proteins (M_r ≈73,000, 43,000, 40,000 and 37,000) from *C. tropicalis* reacted with the anti-SKL antibodies (panel C, *C. trop*, lane PER). The anti-AKI reactive *C. tropicalis* peroxisomal proteins showed different electrophoretic mobilities than those peroxisomal polypeptides that reacted with the anti-SKL antibodies (compare panels B and C). The polypeptides which appear to give a relatively weak reaction on the western blots may be due to a low abundance of these proteins, or to the reaction of the antibodies with an epitope similar to, but not identical to, the carboxy-terminal Ala-Lys-Ile or Ser-Lys-Leu.

3.7.3 Specificity of the anti-AKI antibodies

A competition experiment was performed to determine the specificity of anti-AKI antibodies for the carboxy-terminal Ala-Lys-Ile PTS. Isolated *C. tropicalis* peroxisomes were probed by western blotting with anti-AKI antibodies preadsorbed with

the AKI peptide, ΔAKI peptide (identical to the AKI peptide, but lacking the carboxy-terminal three amino acids), or an unrelated peptide (UR) (Fig. 3.7.3.1). Preadsorption of the antibodies with the AKI peptide greatly diminished binding of the antibodies to the blotted polypeptides (lane AKI) compared to a mock preincubation of the antibodies (lane -). There was no change in the profile of anti-AKI reactive polypeptides when the antibodies were preadsorbed to an unrelated peptide (lane UR) or to the ΔAKI peptide (lane ΔAKI). These results demonstrated that the recognition of the peroxisomal proteins with anti-AKI antibodies was dependent on the presence of the epitope(s) generated by the carboxy-terminal Ala-Lys-Ile PTS. The positions of *C. tropicalis* catalase and fatty acyl-CoA oxidase were also determined on this blot by probing with antibodies specific for these proteins. These proteins did not correspond to anti-AKI reactive proteins, demonstrating that the anti-AKI antibodies do not recognize either catalase or AOX, neither of which contains a consensus carboxy-terminal PTS (Murray and Rachubinski, 1987a; 1987b).

The specificity of antibodies for the AKI motif was also demonstrated by the absence of any reaction with peroxisomal proteins from rat liver, many of which terminate in Ser-Lys-Leu. In fact, the antibody did not detect *any* antigenically similar epitopes in either peroxisomal or mitochondrial proteins of rat liver. The antibodies did not recognize the carboxy-terminal tripeptide Ala-Lys-Leu of peroxisomal sterol carrier protein 2 (Morris *et al.*, 1988) or the carboxy-terminal Ala-Asn-Leu of peroxisomal catalase (Furuta *et al.*, 1986). Internal Ala-Lys-Ile tripeptides present in rat

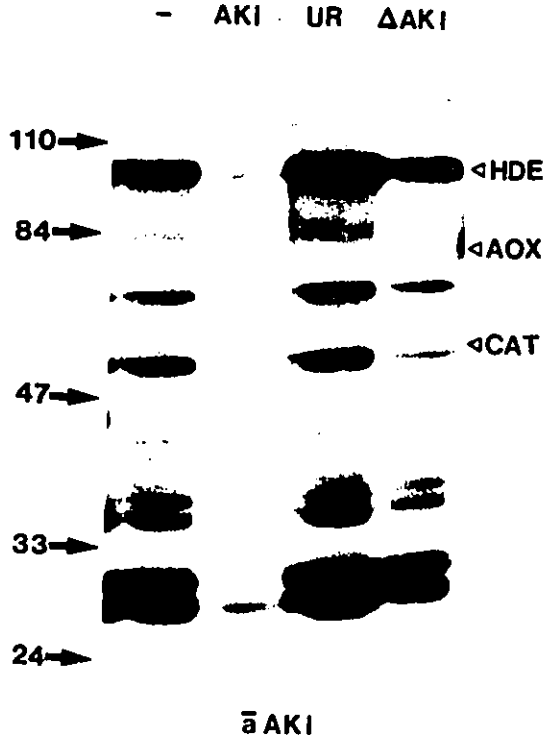


Figure 3.7.3.1 AKI peptide competes specifically for recognition by anti-AKI antibodies. 100 μg of *C. tropicalis* peroxisomal protein was probed with anti-AKI antibodies either preadsorbed in the absence of an added peptide (lane -), or preadsorbed with the AKI peptide (lane AKI), an unrelated peptide (lane UR), or the ΔAKI peptide (lane ΔAKI). See section 3.7.3 for details. The numbers at the left indicate the migrations of molecular mass standards (kDa). HDE, trifunctional enzyme; AOX, fatty acyl-CoA oxidase; CAT, catalase.

mitochondrial carbamoyl phosphate synthase (Nyunoya *et al.*, 1985) and fumarate hydratase (Suzuki *et al.*, 1989) were also not detected (Fig. 1, *panel B, rat, MIT*). In addition, the anti-AKI antibodies did not react with peroxisomal proteins of *H. polymorpha* on western blots (data not shown) demonstrating that the internal Ala-Lys-Ile of dihydroxyacetone synthase (Janowicz *et al.*, 1985) is not recognized by the antibodies. The absence of reaction with internal Ala-Lys-Ile tripeptides indicates that the carboxy-terminal location of the PTS is important for recognition by the anti-AKI antibodies.

3.7.4 Double labelling immunofluorescence studies using anti-AKI antibodies

To characterize further the subcellular localization of the anti-AKI and anti-SKL reactive proteins, *C. tropicalis* cells were examined for the pattern generated with anti-AKI, anti-SKL and anti-AOX antibodies using double labelling indirect immunofluorescence.

3.7.4.1 Anti-AKI and anti-AOX antibodies

Cells were probed first with guinea pig anti-AKI antibodies and rabbit anti-AOX antibodies (Fig. 3.7.4.2.1, column A). The peroxisomal matrix localization of AOX (Fujiki *et al.*, 1986) allows its immunoreaction to serve as a control for the detection of peroxisomes by this procedure. The fluorescence patterns generated by the anti-AOX and anti-AKI antibodies are superimposable and show the punctate pattern characteristic of peroxisomes (Gould *et al.*, 1987). The proteins recognized by these

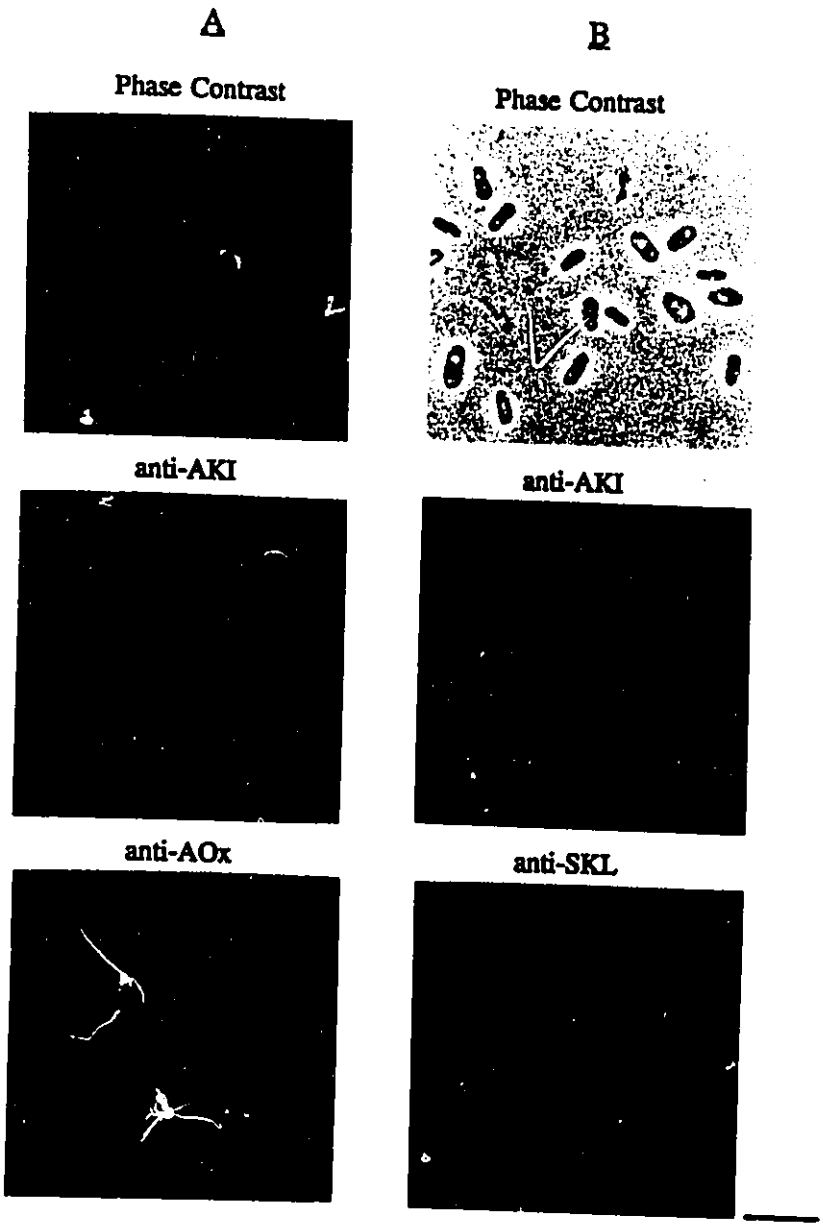


Figure 3.7.4.2.1 Double labelling indirect immunofluorescence analysis of *C. tropicalis*. Column A: Upper panel: phase contrast. Middle panel: Immunofluorecence with anti-AKI antibodies. Lower panel: Immunofluorescence with anti-AOX antibodies. Column B: Upper panel: phase contrast. Middle panel: Immunofluorecence with anti-AKI antibodies. Lower panel: Immunofluorescence with anti-SKL antibodies. Bar represents 10μm.

antibodies were localized exclusively to peroxisomes, as there was an absence of fluorescence over non-peroxisomal structures within the cell, in particular the nucleus, mitochondria, cell membrane and cytoplasm. No fluorescence was detected with preimmune serum.

3.7.4.2 Anti-AKI and anti-SKL antibodies

Immunofluorescence analysis of *C. tropicalis* with anti-AKI antibodies and anti-SKL antibodies (Fig. 3.7.4.2.1, column B) yielded identical punctate patterns and a lack of fluorescence over the nucleus, mitochondria, cell membrane and cytoplasm similar to double labelling experiments with anti-AKI and anti-AOX antibodies. Both anti-AKI and anti-SKL antibodies are therefore specific for peroxisomal proteins of *C. tropicalis*, again suggesting that peroxisomal proteins of *C. tropicalis* terminate with both PTSs.

The combination of western blotting of subcellular fractions and double labelling immunofluorescence of whole cells demonstrates the exclusive localization of anti-AKI reactive and anti-SKL reactive proteins to peroxisomes in *C. tropicalis*. The presence of an antigenically similar epitope at the carboxy-termini of many *C. tropicalis* peroxisomal proteins detected in this study suggests that these proteins may all be recognized by the same putative PTS receptor for import into peroxisomes. However, whether the anti-SKL and anti-AKI reactive proteins would be recognized by a single putative receptor is unknown (but see section 3.8).

3.7.5 Membrane proteins

The polypeptides identified in these studies do not correspond to any of the major peroxisomal membrane proteins of *C. tropicalis* (Nuttley *et al.*, 1990). This result suggests that integral membrane proteins are targeted to peroxisomes by a different signal. To date, no peroxisomal membrane proteins have been shown to contain a carboxy-terminal consensus PTS (Kamijo *et al.*, 1990; McCammon *et al.*, 1990b; Höhfeld *et al.*, 1991; Tsukamoto *et al.*, 1991).

3.7.6 Characterization of anti-AKI reactive proteins in *C. albicans*, *Y. lipolytica* and *S. cerevisiae*

3.7.6.1 *C. albicans*

The Coomassie-stained peroxisomal polypeptide profile of *C. albicans* is shown in Fig. 3.7.6.1.1 (panel A). Three peroxisomal polypeptides (M_r ≈95,000, 30,000, 27,000; the lower band can be resolved into two polypeptides) reacted strongly with anti-AKI antibodies (panel B, lane PER) on western blots. There are two additional polypeptides which gave weaker signals (M_r ≈46,000 and 33,000). Proteins of the mitochondrial fraction did not react with anti-AKI antibodies (panel B, lane MIT). An identical blot was probed with anti-SKL antibodies and two strongly immunoreactive polypeptides (M_r ≈52,000 and 38,000) and a weakly reactive polypeptide (M_r ≈45,000) were detected in the peroxisomal fraction (panel C, lane PER). Like *C. tropicalis*, anti-SKL and anti-AKI antibodies reacted with a different subset of peroxisomal proteins.

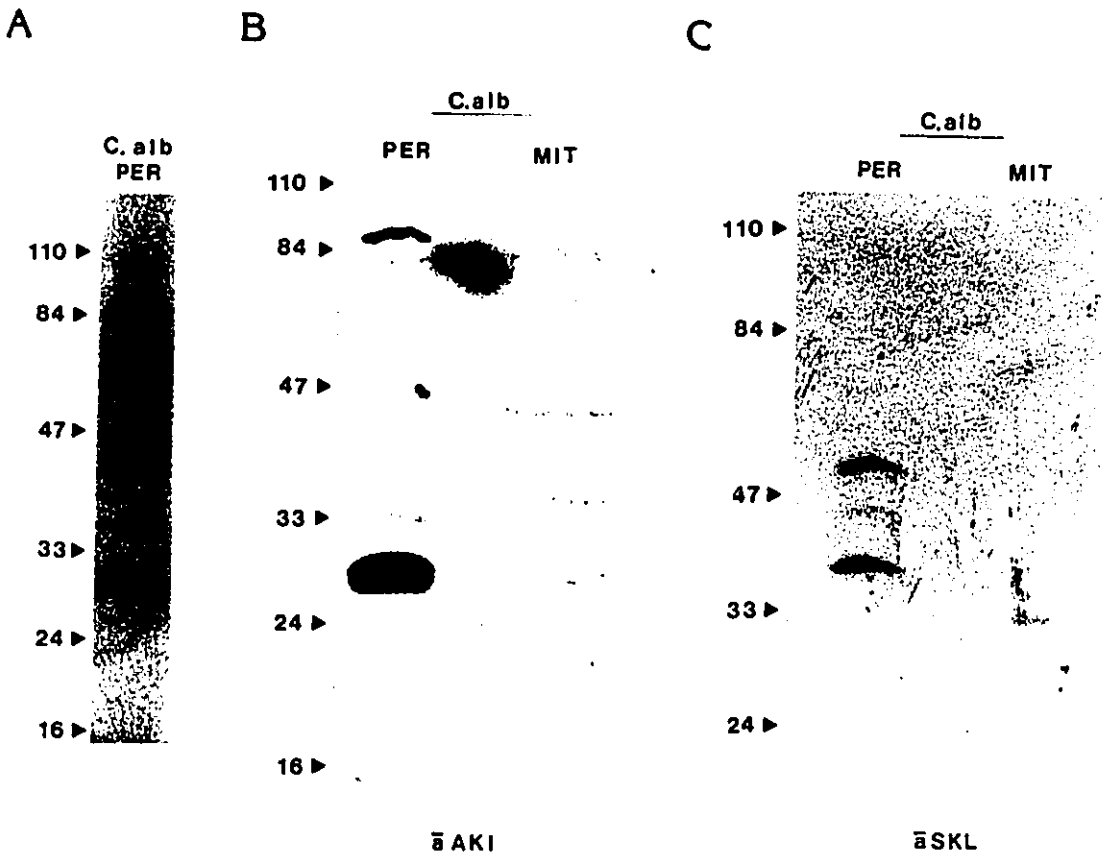


Figure 3.7.6.1.1 Western blot analysis of anti-AKI and anti-SKL reactive proteins in *C. albicans*. Panel A: 50 μg of purified peroxisomal protein from *C. albicans* were separated by SDS-PAGE and stained with Coomassie blue. Panel B: 50 μg of peroxisomal protein (lane PER) or 50 μg of mitochondrial protein (lane MIT) from *C. albicans* were probed with anti-AKI antibodies by western blotting. Panel C: A blot identical to that shown in *panel B* was probed with anti-SKL antibodies. The numbers at the left of each panel indicate the migrations of molecular weight standards.

However, the weakly reactive polypeptides at M_r ≈45,000 to 46,000, which were apparent in both anti-SKL and anti-AKI probed western blots may, or may not, be the same polypeptide (see below; section 3.7.6.2). All anti-AKI or anti-SKL reactive proteins in the cytosolic fraction were attributable to leakage from peroxisomes damaged during homogenization. There was no reaction with preimmune serum.

3.7.6.2 *Y. lipolytica*

The reaction of anti-AKI and anti-SKL antibodies with *Y. lipolytica* peroxisomal proteins was also examined (Fig. 3.7.6.2.1). The polypeptide profile of the isolated peroxisomes generated by SDS-PAGE is shown in Fig. 3.7.6.2.1 (panel A). When the peroxisomal and mitochondrial fractions were probed with anti-SKL antibodies, there were three immunoreactive polypeptides that were exclusively peroxisomal (M_r ≈66,000, 62,000 and 33,000; panel B, lane PER). The immunoreactive polypeptides in the mitochondrial fraction (panel B, lane MIT) were due to peroxisomal contamination of the mitochondrial fraction. When an identical blot was probed with anti-AKI antibodies, two peroxisomal polypeptides reacted (panel C, $\bar{a}$AKI, lane PER), which comigrated with two of the anti-SKL reactive polypeptides (66 and 33 kDa). To determine if the anti-SKL and anti-AKI immunoreactive polypeptides could be distinguished, the blot was reprobed with anti-SKL antibodies (panel C, $\bar{a}$AKI + $\bar{a}$SKL). Although the signal became more intense at 66 kDa, a separate anti-SKL reactive polypeptide at either 66 kDa or 33 kDa could not be identified. This result suggests that

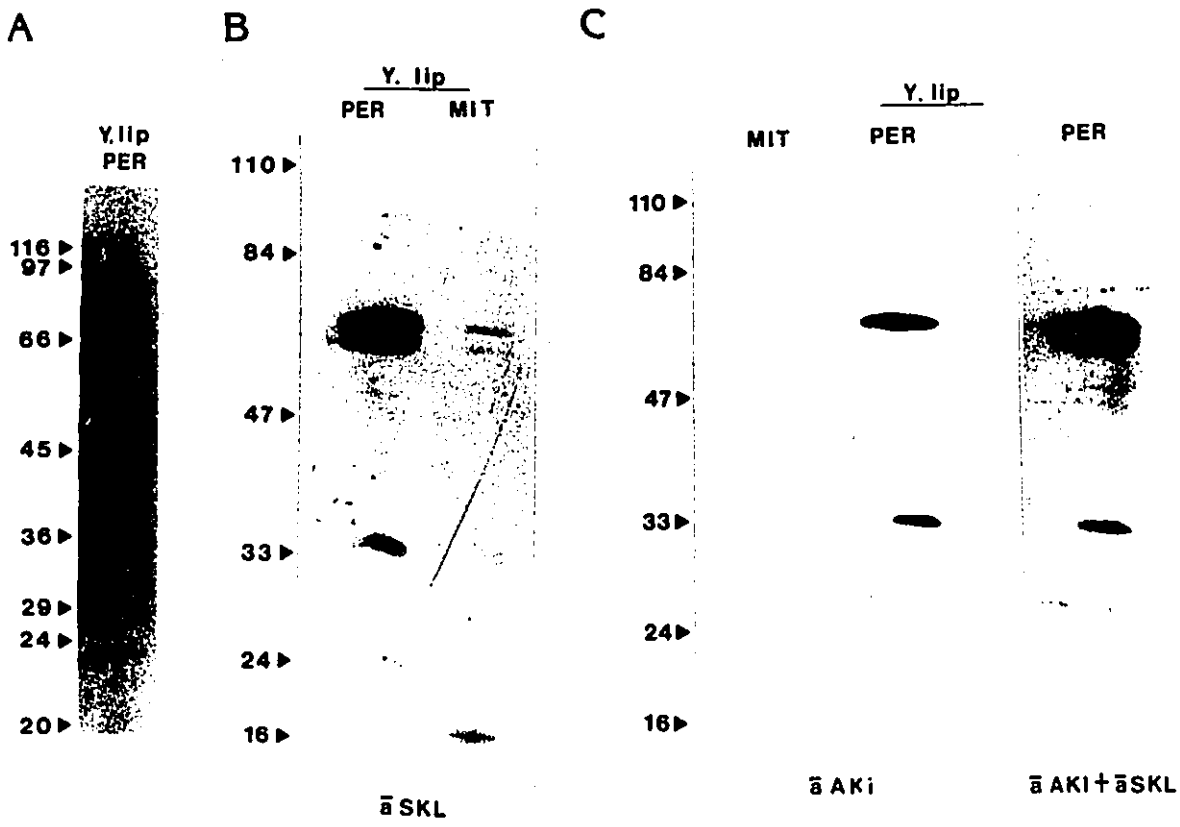


Figure 3.7.6.2.1 Western blot analysis of anti-SKL and anti-AKI reactive proteins in *Y. lipolytica*. Panel A: 75 μg of purified peroxisomal protein from *Y. lipolytica* were separated by SDS-PAGE and stained with Coomassie blue. Panel B: 75 μg of peroxisomal protein (lane PER) or 75 μg of mitochondrial protein (lane MIT) from *Y. lipolytica* were probed with anti-SKL antibodies by western blotting. Panel C: A blot identical to that shown in panel B was probed with anti-AKI antibodies and then probed with anti-SKL antibodies. The numbers at the left of each panel indicate the migrations of molecular weight standards.

the 66 and 33 kDa polypeptides may react with both the anti-AKI and anti-SKL antibodies. This would be a unique situation, because reaction of a protein with both anti-AKI and anti-SKL antibodies has not been observed in any of the other yeasts or in rat peroxisomal proteins, and would probably necessitate a carboxy terminus different from either Ala-Lys-Ile or Ser-Lys-Leu which reacts with both antibodies. Alternatively, there may be two sets of polypeptides at 66 and 33 kDa, each of which contains at least two polypeptides: one that reacts with the anti-SKL antibodies, and a second that reacts with anti-AKI antibodies. There was no reaction of cytosolic proteins with either the anti-SKL or anti-AKI antibodies.

3.7.6.3 Immunofluoresence analysis of anti-AKI reactive proteins in *C. albicans* and *Y. lipolytica*

The localization of the anti-AKI reactive proteins in *C. albicans* and *Y. lipolytica* was also investigated by indirect immunofluorescence. Both yeasts displayed a punctate pattern of fluorescence characteristic of peroxisomes (Fig. 3.7.6.3.1). There was a distinct lack of fluorescence over non-peroxisomal structures, demonstrating that the anti-AKI antibodies recognize only proteins localized to peroxisomes in these yeasts. There was no fluorescence detected with preimmune serum.

The anti-AKI antibodies reacted with several peroxisomal proteins from *C. albicans* and *Y. lipolytica*, suggesting that peroxisomal proteins from these yeasts may be targeted to peroxisomes by a sequence similar to the PTS identified in *C. tropicalis*

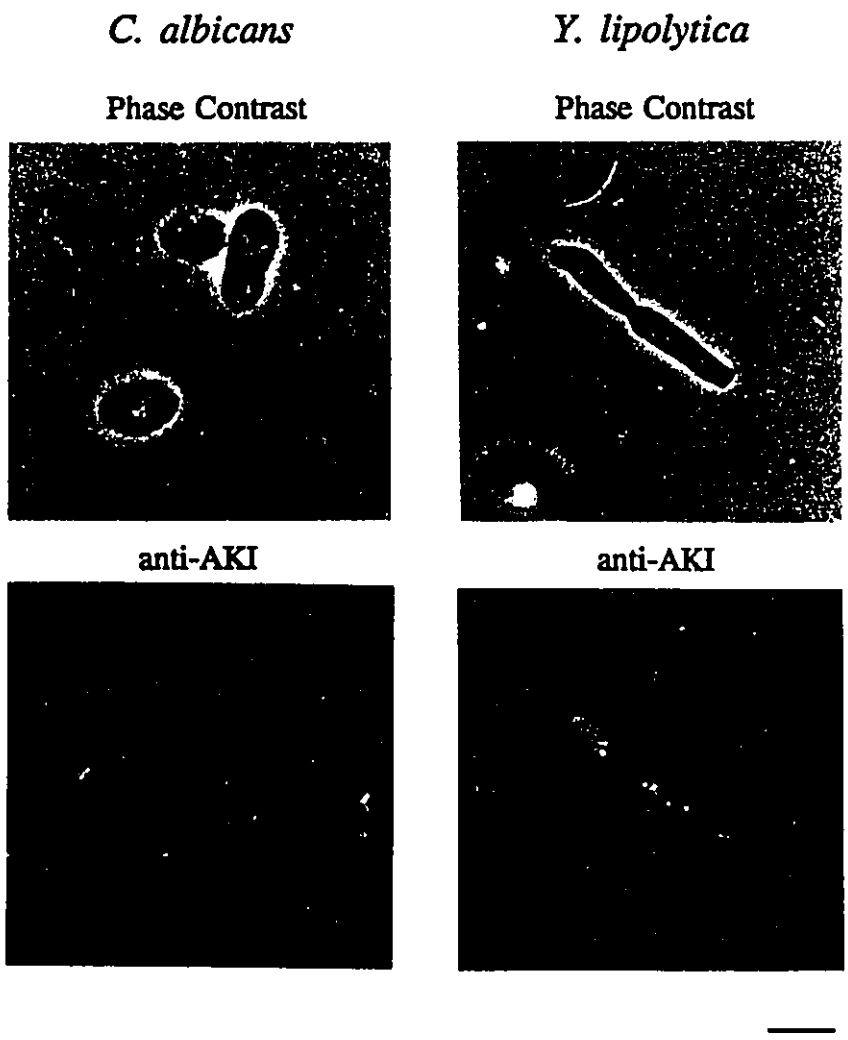


Figure 3.7.6.3.1 Indirect immunofluorescence analysis of *C. albicans* and *Y. lipolytica*. *C. albicans* and *Y. lipolytica* were visualized with phase contrast and indirect immunofluorescence microscopy using anti-AKI antibodies. Bar represents 1 μm.

trifunctional enzyme. Anti-SKL antibodies also detected a number of peroxisomal proteins from these yeasts. Reaction of a different subset of peroxisomal proteins from these yeasts with anti-SKL antibodies indicates that there is a redundancy in the signals that can be recognized as PTSs in yeast. In fact, two proteins of *Y. lipolytica* peroxisomes may react with both antibodies, suggesting a different subclass of PTS which is neither Ala-Lys-Ile-COOH or Ser-Lys-Leu-COOH. This is not surprising since there appears to be a similar redundancy in the PTSs in mammalian cells (Gould *et al.*, 1989).

3.7.6.4 Anti-SKL and anti-AKI reactive proteins in *S. cerevisiae*

Peroxisomes from *S. cerevisiae* contained four polypeptides which reacted strongly with anti-SKL antibodies (M_r ≈95,000, 84,000, 62,000 and 32,000) and one polypeptide which reacted weakly (M_r ≈54,000; Fig. 3.7.6.4.1). The polypeptide at M_r 95,000 probably corresponds to *S. cerevisiae* trifunctional enzyme which terminates with Ser-Lys-Leu (MW 98,703 Da, section 3.2.8.2). Only one weakly reactive polypeptide (M_r ≈42,000) was observed when peroxisomes of *S. cerevisiae* were probed with anti-AKI antibodies (Fig. 3.7.6.4.1, panel B). There was no reaction of *S. cerevisiae* mitochondrial proteins with either anti-AKI or anti-SKL antibodies.

In contrast to peroxisomal proteins of *C. tropicalis*, *C. albicans* and *Y. lipolytica*, peroxisomal proteins from *S. cerevisiae* and rat liver were not detected with anti-AKI antibodies (with the exception of one weakly reactive protein in *S. cerevisiae*

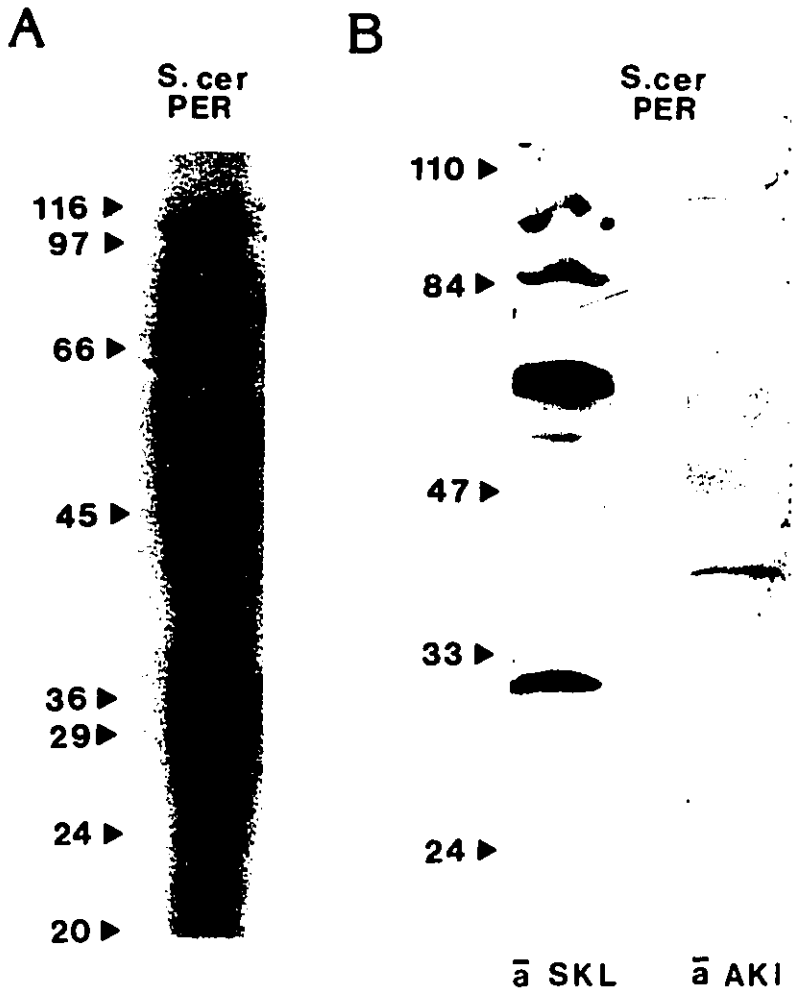


Figure 3.7.6.4.1 Western blot analysis of anti-SKL and anti-AKI reactive proteins in *S. cerevisiae*. Panel A: 100 μg of purified peroxisomal protein from *S. cerevisiae* was separated by SDS-PAGE and stained with Coomassie blue. Panel B: 100 μg of peroxisomal protein was probed with anti-SKL or anti-AKI antibodies by western blotting. The numbers at the left of each panel indicate the migrations of molecular mass standards (kDa).

peroxisomes) but were detected with anti-SKL antibodies. This result suggests that peroxisomal proteins from *S. cerevisiae* and rat liver may not contain an Ala-Lys-Ile PTS. However, *S. cerevisiae* does recognize Ala-Lys-Ile-COOH as a PTS, because *C. tropicalis* trifunctional enzyme is imported into *S. cerevisiae* peroxisomes, and deletion of this PTS renders the protein cytosolic (section 3.6; Aitchison *et al.*, 1991a). A listing of the M_r of the anti-AKI and anti-SKL reactive proteins is given in Table 3.7.6.4.2.

3.7.7 Other proteins terminating in SKL or AKI

Anti-SKL antibodies react with a number of mitochondrial and endoplasmic reticular proteins from rat liver (Gould *et al.*, 1990; see also Fig. 1, *panel C, lane MIT*). Although a carboxy-terminal Ser-Lys-Leu is sufficient for directing cytosolic proteins to peroxisomes in mammalian cells (Gould *et al.*, 1989), a scan of the SWISSPROT data base (version 18, August, 1991) using the PC/GENE sequence analysis program revealed a number of non-peroxisomal proteins which terminate in Ser-Lys-Leu: two mitochondrial proteins, ATP synthase protein 8 (encoded in the mitochondrial genome) from *Candida glabrata* (Clark-Walker *et al.*, 1985) and *S. cerevisiae* (Simon and Faye, 1984; molecular mass ≈5000 Da, unresolved by the gel system used in this study) and ribosomal protein MRP7 precursor (Fearon and Mason, 1988); one chloroplast protein, which is encoded by the chloroplast genome, ATP synthase ϵ chain from liverwort (Ohyama *et al.*, 1986); and two secreted proteins, α-amylase precursors from human (Nishide *et al.*, 1986; Horii *et al.*, 1987), mouse (Hagenbuchle *et al.*, 1980), pig

Approximate M_r of anti-AKI and anti-SKL immunoreactive peroxisomal proteins

	anti-AKI[1]	anti-SKL[2]
Candida tropicalis	100 (HDE[3])	73
	69	43
	54	40
	38	37
	35	
	33	
	29	
	27	
	18	
Yarrowia lipolytica	66	66
	33	62
		33
Candida albicans	94	52
	33 (faint)	45 (faint)
	30	38
	27	
Saccharomyces cerevisiae	40 (faint)	95 (TFE[4])
		84
		62 (doublet)
		54 (faint)
		32
Rat Liver, clofibrate treated		81 (HD[5])
		72 (AOX[6])
		62
		43
		40
		32
		24
		21

1. M_r of anti-AKI reactive proteins.

2. M_r of anti-SKL reactive proteins.

3. HDE; MW 99,481 Da (section 3.2).

4. Trifunctional enzyme (HDE); MW 98,703 Da (section 3.2.8.2).

5. HD; MW 78,526 (Osumi *et al.*, 1985).

6. AOX; MW 74,526 (Miyazaa *et al.*, 1987)

(Pasero *et al.*, 1986), and rat (McDonald *et al.*, 1980) and serine carboxypeptidase I from barley (Sorensen *et al.*, 1986; Doan and Fincher, 1988). The ATP synthases are encoded by the mitochodrial and chloroplastic genomes, and are therefore not expected to come in contact with the putative peroxisomal import machinery. On the other hand, MRP7 and proteins which travel through the secretory pathway are nuclear encoded and synthesized on cytoplasmic ribosomes.

This evidence suggests that a signal peptide at the amino-terminus of a protein may be dominant over a carboxy-terminal PTS, either by being recognized by the import machinery of other organelles prior to recognition by the peroxisomal import machinery or by altering the conformation of the protein sufficiently to mask a PTS. In this regard, it has been shown that the fusion of a mitochondrial targeting signal to the amino-terminus of luciferase directs the fusion protein to mitochondria in yeast (Aflalo, 1990). In addition, it has been shown that certain deletions in the amino terminus of luciferase, which keep the PTS intact, affect the ability of the protein to be imported into peroxisomes (Gould *et al.*, 1987). These observations provide a plausible explanation for the presence of the Ser-Lys-Leu tripeptide at the carboxy terminus of at least one cytoplasmic protein (malonyl-CoA decarboxylase from western graylag goose; Jang *et al.*, 1989). A similar search of the data base revealed only four eukaryotic proteins, in addition to *C. tropicalis* trifunctional enzyme, known to terminate with Ala-Lys-Ile: glutamic acid-rich protein from *Plasmodium falciparum* (Triglia *et al.*, 1988) and rat OX40 antigen precursor (Mallett *et al.*, 1990) are expressed at the cell surface and

contain amino terminal signal sequences; RNA splicing protein MRS1 from *S. cerevisiae* (Kreike *et al.*, 1987) is mitochondrial, and Cu-Zn superoxide dismutase from *Drosophila virilis* (Kwiatowski and Ayala, 1989), the human homolog of which has been shown to be peroxisomal (Keller *et al.*, 1991b). The reason MRS1 was not detected by anti-AKI antibodies is unknown. The quantities of MRS1 may be below the levels of detection of our western blot assay, or modifications (either post-transcriptional or post-translational) may remove the predicted Ala-Lys-Ile carboxy terminus.

3.7.8 Utility of AKI antibodies

The use of anti-AKI antibodies and anti-SKL antibodies allows for identification of different carboxy-terminal PTSs. Anti-AKI and anti-SKL antibodies react with different subsets of proteins. Anti-AKI antibodies do not recognize any peroxisomal proteins from rat liver (at least two of which terminate with SKL; rat HD (Osumi *et al.*, 1985) and rat AOX (Miyazawa *et al.*, 1987). In addition, anti-AKI antibodies do not react with rat peroxisomal sterol carrier protein 2, which terminates with Ala-Lys-Leu (Morris *et al.*, 1988; see Fig. 3.7.2.1, panel B, lane rat, PER). Anti-SKL antibodies recognize carboxy-terminal Ser-Lys-Leu tripeptides (Gould *et al.*, 1990a). It has previously been shown that anti-SKL antibodies react with rat sterol carrier protein-2 (Gould *et al.*, 1990a), which terminates with Ala-Lys-Leu (Morris *et al.*, 1988). However, the antibody reaction with most rat liver peroxisomal proteins can not be competed with peptides containing Ala-Lys-Leu or Ser-Arg-Leu (Keller *et al.*,

1991a), which have been shown to act as PTSs (Gould *et al.*, 1989). These antibodies also do not react with trifunctional enzyme (Fig. 3.7.2.1, panel 2, lane PER), which contains the carboxy-terminal Ala-Lys-Ile tripeptide (Nuttley *et al.*, 1988).

The anti-AKI antibodies described herein will be a valuable tool for further investigations into yeast peroxisomal biogenesis. The approach described above can be used to extend our current knowledge of peroxisomal targeting signals in other organisms. The anti-AKI antibodies also provide a means to clone previously unidentified peroxisomal proteins from yeast expression libraries, which will lead to further insights into the structural limitations for yeast PTSs. In this regard, 22 cDNA clones have been isolated from an HDE-depleted oleic acid-grown *C. tropicalis* cDNA library. The identity of these clones is currently under investigation. The anti-AKI antibodies can also be used to generate anti-idiotypic antibodies. These can be used, in conjunction with ligand binding experiments, to identify AKI-binding proteins in *C. tropicalis*. In addition, the anti-AKI antibodies are currently being used to screen for peroxisomal assembly mutants in *Y. lipolytica*.

4.0 SUMMARY AND PERSPECTIVES

4.1 PTS-I

At least some peroxisomal proteins are directed to the organelle by a carboxy-terminal tripeptide PTS. The prototype PTS was identified by Dr. S. Gould in the laboratory of Dr. S. Subramani as the carboxy-terminal tripeptide Ser-Lys-Leu (Gould *et al.*, 1987, 1988, 1989). Mutational analysis permitted the elucidation of a consensus tripeptide, subsequently designated PTS-I, which was capable of sorting luciferase to peroxisomes in CV-1 cells: Ser/Ala/Cys-Lys/Arg/His-Leu (the leucine being carboxy-terminal and invariant) (Gould *et al.*, 1989). The addition of a single Ser downstream of this PTS rendered the resultant protein cytosolic (Gould *et al.*, 1989) suggesting that the location of the tripeptide at the carboxy terminus is essential for its action. Independent of these investigations, rat AOX was also shown to be targeted to peroxisomes by its carboxy-terminal 5 amino acids (Leu-Gln-Ser-Lys-Leu) *in vitro* (Miyazawa *et al.*, 1989).

The carboxy-terminal 14 amino acids of pig *D*-amino acid oxidase, 15 amino acids of rat AOX, 15 amino acids of rat HD (Gould *et al.*, 1988), and 12 amino acids of *C. boidinii* PMP20 (Gould *et al.*, 1990b) were also shown to be capable of directing *E. coli* CAT fusion proteins to peroxisomes in CV-1 cells. In retrospect, this is not

surprising because these peptides all contain variations of PTS-I (Gould *et al.*, 1989; see Table 4.2.1). These experiments indicated that the targeting mode of luciferase is common to other peroxisomal proteins. In addition, they immediately suggested that peroxisomal targeting was conserved between insects (firefly), yeast (*C. boidinii*) and mammals (rat, pig, monkey). Recently, it has been shown that luciferase is targeted to peroxisomes when expressed in yeast (*S. cerevisiae*), plant (*Nicotiana tabacum*) (Gould *et al.*, 1990) and embryonic frog (*Xenopus*) neurons (Holt *et al.*, 1990). The *E. coli* CAT-SKL fusion protein is also peroxisomal (glycosomal) when expressed in *T. brucei* (Fung and Clayton, 1991).

The question of cross-species PTS-I conservation has also been addressed using antibodies directed against a peptide corresponding to the carboxy-terminal 12 amino acids of rat AOX (and HD), which specifically react with polypeptides terminating with Ser-Lys-Leu (Gould *et al.*, 1990a). These antibodies specifically reacted with multiple peroxisomal proteins from mammalian cells (Gould *et al.*, 1990a), glyoxysomal proteins from *Pichia pastoris*, germinating castor beans and *Neurospora crassa* as well as with glycosomal proteins from *Trypanosoma brucei* (Keller *et al.*, 1991a). This evidence indicates that peroxisomes from different organisms harbour PTS-I-containing proteins and that peroxisomes, glycosomes and glyoxysomes all share a common evolutionary origin (see section 1.5).

4.2 A more general PTS-I

The experiments discussed above failed to yield any immunoreactive polypeptides in the yeast *C. tropicalis*. In addition, although several *C. tropicalis* genes encoding peroxisomal proteins have been cloned and sequenced (see section 1.3.2, and Table 4.2.1), none of them encodes a canonical PTS-I. Experiments addressing the nature of the peroxisomal targeting signal of *C. tropicalis* AOX (PXP-4) *in vitro* pointed to an amino-terminal and an internal region of the protein (approximately 100 amino acids each) that were capable of directing DHFR fusion proteins to peroxisomes (Small *et al.*, 1988). Neither of these regions contains an SKL-like motif (Borst, 1989). These findings suggested that *C. tropicalis* might possess a unique means of targeting proteins to peroxisomes. The identification of the Ala-Lys-Ile tripeptide PTS of HDE indicates that PTS-I is indeed conserved in *C. tropicalis* and that substitutions in the signal vary amongst organisms. In addition, the identification of several peroxisomal proteins of *C. tropicalis* and of other yeasts that contain epitopes immunologically related to the HDE-PTS (section 3.7) suggests that this form of the PTS may be common to several different proteins.

The experiments described in section 3.5 suggest that Gly (position -3) and Gln (position -2) can be added to the consensus PTS-I sequence. In addition, recent experiments suggest that Met may substitute for Leu in the luciferase targeting signal (Swinkels and Subramnai, 1990; unpublished observations cited in Subramani, 1991, 1992). Therefore it appears that at least one mechanism of peroxisomal targeting is

evolutionarily conserved and that the PTS-I consensus sequence needs to be expanded.

Although there are obvious species variations (section 3.7), and it is likely that not all of amino acid combinations work equally efficiently as PTSs (Swinkels and Subramani, 1990), the PTS-I consensus sequence can be presented as follows:

Pos	-3	-2	-1	-COOH
	Ser	Lys	Leu	
	Ala	Arg	Ile	
	Cys	His	Met	
	Gly	Gln		

Analysis of this consensus sequence suggests that the requirements for the PTS may be a small, uncharged amino acid in position -3, a basic or hydrophilic amide-containing amino acid at position -2, and a hydrophobic amino acid at position -1. It is likely, using this more general PTS-I consensus sequence, that several peroxisomal proteins that do not conform to the original consensus sequence for PTS-I (Gould *et al.*, 1989) are targeted to peroxisomes by a PTS-I motif.

A current list of the carboxy-termini of known peroxisomal proteins is shown in Table 4.2.1. It is evident from this list that several proteins fit the proposed expanded PTS-I requirements outlined above. In particular, human (and rat) catalase terminates with Ala-Asn-Leu and fusion of this sequence plus 24 additional amino acids immediately upstream of this tripeptide to *E. coli* CAT is sufficient to direct the resultant protein to peroxisomes in CV-1 cells. Similar variations have been described for the carboxy-terminal ER retention signals (see section 1.7.1.2).

Table 4.2.1 Carboxy termini of peroxisomal proteins

Proteins which fit consensus PTS-1:*

Protein	Carboxy terminus	Reference
rat AOX	SKL[§]	Miyazawa *et al.*, 1987; Gould *et al.*, 1988, 1989; Miyazawa *et al.*, 1989
rat HD	SKL[§]	Osumi *et al.*, 1985; Gould *et al.*, 1988, 1989
pig, mouse, human D-amino acid oxidase	SHL[§]	Ronchi *et al.*, 1982; Gould *et al.*, 1988, 1989; Tada *et al.*, 1990; Momoi *et al.*, 1988; Gould *et al.*, 1988, 1989
pig, mouse, baboon, rat uricase	SRL	Wallrath *et al.*, 1990; Alvarez *et al.*, 1989
Drosophila melanogaster uricase	SHL	Wallrath *et al.*, 1990
Photinus pyralis luciferase	SKL	de Wet *et al.*, 1987; Keller *et al.*, 1987; Gould *et al.*, 1987, 1989
P. plagiophthalamus luciferase	SKL	Wood *et al.*, 1989
cucumber malate synthase	SKL[§]	Smith and Lever, 1986; Gould *et al.*, 1988
Brassica napus malate synthase	SRL	Comai *et al.*, 1989
spinach glycolate oxidase	ARL	Volokita and Sommerville, 1987
S. cerevisiae trifunctional enzyme	SKL	see section 3.2.8.2
S. cerevisiae citrate synthase	SKL	Lewin *et al.*, 1990
S. cerevisiae DAL-7 protein	SKL	Yoo and Cooper, 1989
C. boidinii PMP-20 (A and B)	AKL[§]	Garrard and Goodman, 1989; Gould *et al.*, 1990a
T. brucei glucose-6-phosphate isomerase	SHL	Marchand *et al.*, 1989
T. brucei glyceraldehyde-3-phosphate dehydrogenase	AKL	Michels *et al.*, 1986
T. cruzi glyceraldehyde-3-phosphate dehydrogenase	ARL	Kendall *et al.*, 1990
rat sterol-carrier protein 2	AKL	Keller *et al.*, 1989; Billheimer *et al.*, 1990 (synthesized as a precursor)

Proteins which fit the more general consensus PTS-1[●]

C. tropicalis HDE	AKI[§]	this study; Nuttley *et al.*, 1988; Aitchison *et al.*, 1991
L. cruciata luciferase	AKM	Masuda *et al.*, 1989
cotton and *Ricinus communis* isocitrate lyase	ARM	Turley *et al.*, 1990 Beeching and Northcote, 1987
B. napus isocitrate lyase	SRM	Comai *et al.*, 1989
H. polymorpha alcohol oxidase	ARF	Ledeboer *et al.*, 1985
C. tropicalis isocitrate lyase	AKV	Atomi *et al.*, 1990
maize catalase 3	ANM	Redinbaugh *et al.*, 1988
D. melanogaster catalase	SKF	Orr *et al.*, 1990
human, rat catalase	ANL[□]	Quan *et al.*, 1986; Furuta *et al.*, 1986; Gould *et al.*, 1988

Proteins which do not fit the consensus sequence:

matrix proteins:

H. polymorpha amine oxidase	CGK	Bruinenberg *et al.*, 1989
C. tropicalis PXP-4(a)	LSK	Okazaki *et al.*, 1986; Murray and Rachubinski, 1987a
C. tropicalis PXP-5	LSS	Okazaki *et al.*, 1986
C. tropicalis PXP-2	LSK	Okazaki *et al.*, 1987
S. cerevisiae AOX (POX-1)	INK	Dmochowska *et al.*, 1990
C. maltosa AOX	LSK	Hill *et al.*, 1988
H. polymorpha dihydroxyacetone synthase	GKA	Janowicz *et al.*, 1985
maize catalase 1	PSM	Redinbaugh *et al.*, 1988
maize catalase 2	PEA	Bethards *et al.*, 1987
C. tropicalis catalase	PRK	Okada *et al.*, 1987; Murray and Rachubinski, 1987
bovine catalase	PKN	Schroeder *et al.*, 1982
sweet potato catalase	PTM	Sakojo *et al.*, 1987
castor bean, cotton (1 and 2) catalase	PSI	Gonzalez, 1991
N. crassa trifunctional enzyme	ELA	see section 3.2.8.2
C. tropicalis PXP-18	PKL	Szabo *et al.*, 1989
human superoxide dismutase	IAQ	Levanon *et al.*, 1985; Keller *et al.*, 1991b

integral membrane proteins:

S. cerevisiae PAS3	FKP	Hohfield *et al.*, 1991
rat PMP 70	FGS	Kamijo *et al.*, 1990
C. boidinii PMP 47	AKE†	McCammon *et al.*, 1990a
CHO PAF1	NAL	Tsukamoto *et al.*, 1990

proteins synthesized as precursors:

rat thiolase 1 and 2	PGN†•	Furuta *et al.*, 1982; Bodnar and Rachubinski, 1991; Swinkels *et al.*, 1991
human thiolase	PGN	Bout *et al.*, 1988
watermelon malate dehydrogenase	IRS	Gietl, 1990
rat sterol-carrier protein 2	AKL	Billheimer *et al.*, 1990

---

* - consensus signal: S/A/C-K/R/H/-L (Gould *et al.*, 1989)
@ - consensus signal defined in section 3.8.1
δ - shown to act in targeting to peroxisomes
† - shown not to be involved in targeting to peroxisomes
• - targeted to peroxisomes by amino-terminal presequence (PTS-II)
Σ - located within the PTS but not precisely defined

4.3 PTS-II

Rat liver thiolase is unique among peroxisomal proteins in that it is synthesized as a precursor protein bearing an amino-terminal extension, which is cleaved upon import of the protein into peroxisomes (section 1.2.1.4). Thiolase does not contain a PTS-I (Bodnar and Rachubinski, 1990), and, therefore, by analogy to signal sequences for the ER, mitochondrion and chloroplast, it has been hypothesized that this extension may contain the protein's targeting information (Borst, 1989). Swinkels *et al.* (1991) investigated the role of the amino-terminal extension in targeting thiolase to peroxisomes. Recently, it was determined that the peroxisomal targeting information in thiolase resides in the amino-terminal extension. Deletion of the amino-terminal extension from peroxisomal thiolase rendered the protein cytosolic, and fusion of amino acids -26 to -16 within the prepiece to the amino terminus *E. coli* CAT directed the fusion protein to peroxisomes. This signal has therefore been designated PTS-II. The precise characteristics of PTS-II have not been fully defined, and the presence of this signal within the other precursor proteins (listed in Table 4.2.1) remains to be shown. Comparison of the amino-terminal signal of thiolase with other peroxisomal preseqences for which sequence information is available, including human and yeast thiolases (*S. cerevisiae*, *C. tropicalis* and *Y. lipolytica*) and watermelon malate dehydrogenase, revealed the presence of two conserved dipeptides which may form part of PTS-II. There is an Arg-Leu/Ile and a His/Gln-Leu dipeptide separated by 5 amino acids present in each of the amino terminal regions. The precise role, if any, of these dipeptides in

peroxisomal targeting remains to be shown.

4.4 Other PTSs

Several peroxisomal proteins do not contain PTS-I or an amino-terminal extension in the prototype of PTS-II and are therefore believed to be targeted to peroxisomes by a distinct PTS. As discussed in section 4.1, analysis of the PTS of *C. tropicalis* AOX (PXP-4) by *in vitro* methods suggests that this protein contains two redundant, non-overlapping topogenic sequences which sort it to peroxisomes, one of which is internal (amino acids 309-462), the other of which is located at the amino terminus (amino acids 1-112) of the PXP-4 molecule (Small *et al.*, 1988). Neither sequence contains a variation of the conserved tripeptide PTS (Borst, 1989). It will, of course, be interesting to determine if these sequences operate as PTSs *in vivo*, if they are present in other peroxisomal proteins, and if present, if they operate in other organisms.

Peroxisomal integral membrane proteins do not react with anti-PTS-I antibodies (section 3.7; Gould *et al.*, 1990a; Keller *et al*, 1991a) and those whose sequences are known (CHO and human PAF-I, Tsukamoto *et al.*, 1990, Shimozawa *et al.*, 1992; *S. cerevisiae* PAS-3, Höhfield *et al.*, 1991; rat PMP 70, Kamijo *et al.*, 1990; *C. boidinii* PMP 47, McCammon *et al.*, 1990; see Table 4.2.1) do not contain a PTS-I. In addition, they appear to lack amino-terminal presequences. It is likely that membrane proteins are targeted to peroxisomes by a signal different from either PTS-I or PTS-II.

As discussed in section 1.7, many integral membrane proteins are targeted to (or retained in) other organelles by signals distinct from their soluble counterparts.

It has been suggested that glycosomal phosphoglycerate kinase (PGK) is directed to glycosomes by its carboxy terminus. PGK is uniquely suited for studying glycosomal targeting sequences, because the glycosomal form differs from the cytosolic form by the presence of a carboxy-terminal extension (Borst, 1989). Initial cloning and sequencing of the genes encoding PGK from *T. brucei* indicated that the cytosolic and glycosomal forms of the protein are 95% identical in amino acid sequence, but differ primarily by a higher propensity of positively charged amino acids and a 20 amino acid carboxy-terminal extension present in the glycosomal counterpart (Osinga *et al.*, 1985). Similar analysis of genes encoding the glycosomal and cytosolic forms of PGK from *Crithidia fasiculata*, a kinetoplastid related to *T. brucei*, demonstrated that the two forms of PGK in this organism differ by only 2 conserative amino acid substitutions near the amino terminus and a 38 amino acid carboxy-terminal extension present in the glycosomal form (Swinkels *et al.*, 1988). These results indicate that the positive charge is probably not important to the glycosomal targeting of *T. brucei* PGK as suggested by Wieringa *et al.* (1987) and that the carboxy-terminal extension encodes the glycosomal sorting signal (Swinkels *et al.*, 1988). There are no striking structural similarities between the two carboxy-terminal extensions. They are both rich in small hydrophobic and hydroxylated amino acids and are devoid of charged residues (Swinkels *et al.*, 1988). Interestingly, this is also a feature of the PTS-II of rat liver thiolase (Swinkels *et al.*,

1991). The roles of the carboxy-terminal extensions in targeting foreign proteins to peroxisomes (or glycosomes) have not been investigated directly.

By analogy to protein import into other organelles, protein import into peroxisomes is probably mediated by a series of proteinaceous receptors, perhaps present within the cytoplasm as well as at the surface of the peroxisome. It is also likely that different classes of PTSs are recognized, at least initially, by different PTS receptors. Certain classes of Zellweger's syndrome import thiolase, but not PTS-I containing proteins (Balfe *et al.*, 1990; Walton *et al.*, 1992). Similarly, a yeast mutant has been isolated which is specifically incapable of translocating thiolase into peroxisomes (Marzioch *et al.*, 1991). The question of whether proteins interact with a common translocator protein (such as mitochondrial GIP; Pfaller *et al.*, 1988) after interaction with PTS-specific receptors remains open to speculation.

4.5 Mechanism of Import

Most information regarding peroxisomal biogenesis has focussed on the sequences required for protein sorting to peroxisomes. This information has been obtained mainly using *in vivo* analyses and therefore has yielded little information on the mechanism of protein import into peroxisomes. Studies of the mechanism of peroxisomal protein import *in vitro* have been hampered by a number of problems unique to peroxisomes (see Murray. W.W., PhD thesis, McMaster University, 1992). *In vitro*

import assays have relied on the conversion of proteins to a particle-associated and protease resistant form (which can be abolished by the addition of detergents) when added to isolated peroxisomes (Fujiki and Lazarow, 1985). As discussed in section 1.4, most peroxisomal proteins are imported into peroxisomes without any detectable modifications. Indeed, this is true for the proteins studied thus far using *in vitro* import assays. Thus import cannot be detected by a simple M_r shift by SDS-PAGE, as is possible for *in vitro* import into microsomes, mitochondria or chloroplasts. Consequently, the assay is susceptible to artefactual protein aggregation, which may be interpreted as true protein import.

The limited reports of *in vitro* import into peroxisomes suggest that translocation into the organelle occurs in two steps. There is a temperature-independent binding of proteins to the surface of peroxisomes, followed by ATP- and temperature-dependent translocation (Imanaka *et al.*, 1987; Small and Lazarow, 1987; Small *et al.*, 1988; Thieringer *et al.*, 1991; but see Small *et al.*, 1987). Temperature dependence of protein import has also been confirmed by microinjection experiments (Walton *et al.*, 1992; see below).

One candidate for an ATPase involved in protein translocation is rat liver PMP 70. This integral membrane protein contains six proposed transmembrane segments, an ATP-binding domain exposed to the cytosol, is homologous to several members of the P-glycoprotein family of proteins involved in active transport across membranes in both prokaryotes and eukaryotes, but shows little similarity to proton or

ion pump ATPases (Kamijo *et al.*, 1990; see below). Mild protease treatment of peroxisomes (and PMP 70) destroys the ability of peroxisomes to uptake proteins *in vitro*. It should be stressed, however, that the function of PMP 70 is unknown and the implication that this protein is involved in protein transport across the peroxisomal membrane is entirely conjecture.

It is not known if an electrochemical gradient is required for the import of proteins into peroxisomes. ^{31}P-NMR studies (Nicolay *et al.*, 1987) and fluorescence quenching experiments (Douma *et al.*, 1987) indicate that there is a 1.1 - 1.3 pH unit proton gradient across the peroxisomal membranes of *H. polymorpha, C. utilis* and *Trichosporon cutaneum X4*). An ATPase has been detected associated with the peroxisomal membrane of *H. polymorpha* by cytochemical analysis (Douma *et al.*, 1987). It is proposed that this protein is a proton translocating ATPase based on biochemical (Douma *et al.*, 1987) and immunological (Douma *et al.*, 1990) similarities to the mitochondrial ATPase. The putative peroxisomal ATPase reacts with antibodies against the F_1 and β subunits of mitochondrial ATPase from *S. cerevisiae* (Douma *et al.*, 1990). In addition, a similar membrane ATPase has been reported to copurify with rat liver peroxisomes (del Valle *et al.*, 1988). Both of these ATPases are sensitive to oligomycin like their mitochondrial counterparts (Douma *et al.*, 1987; del Valle *et al.*, 1989).

It is unlikely that such a pH gradient exists in the *in vitro* import assays reported. The presence of a peroxisomal reticulum necessitates that upon isolation of spherical organelles (used for *in vitro* import assays) peroxisomes must break and reseal,

thus dissipating the proposed gradient. Isolated peroxisomes are notoriously fragile and permeable to small solutes such as sucrose, phosphorylated sugars and nucleotide cofactors (reviewed in Borst, 1989). This permeability would presumably also dissipate the pH gradient. *In situ* analyses have shown that protein import into peroxisomes of *C. boidinii* is disrupted by the addition of the proton ionophore carbonyl cyanide-*m*-chlorophenylhydrazone (CCCP) (Bellion and Goodman, 1987). However, this disruption may be a secondary effect resulting from the depletion of cellular ATP rather from than dissipation of the peroxisomal proton gradient (Borst, 1989).

It is unknown if proteins are unfolded prior to or during import into peroxisomes. Microinjection of peroxisomal proteins into mammalian cells, has provided some insight into this process. Human serum albumin (HSA) coupled to SKL-containing peptides is efficiently imported into peroxisomes, indicating that branched proteins can be imported into the organelle (Walton *et al.*, 1992). In addition, microinjection of luciferase bound to the luciferin analogue dehydroluciferin, which presumably prevents complete unfolding of luciferase, is imported into the organelle (Walton *et al.*, 1990; Paul Walton, personal communication). Furthermore, import into peroxisomes does not disrupt SKL-HSA-biotin conjugates complexed with streptavidin-Texas Red. Microinjection of the complex results in Texas Red fluorescence within peroxisomes (Paul Walton, personal communication). These preliminary experiments suggest that extensive unfolding of peroxisomal proteins is not required for protein import into peroxisomes, at least upon microinjection. The upper size limit of the putative

translocation channel is, however, unknown.

4.6 Peroxisomal Assembly Mutants

Recent attempts to identify components of the peroxisomal import machinery in the absence of a reliable *in vitro* import assay have focussed on the use of peroxisomal assembly mutants in CHO cells and yeast. In yeast these mutants are detectable by their inability to use oleic acid as a sole carbon source (see section 1.3.2) combined with the cytosolic location (detected by presence in the 20 kgS fraction) of normally peroxisomal matrix proteins (Erdmann *et al.*, 1989). Peroxisomal assembly mutants, in this regard, mimic the human genetic disorder, Zellweger's syndrome (section 1.6). Complementation of assembly mutants has revealed two complementing genes in *S. cerevisiae* (*PAS-1* and *-3*) and one in CHO cells (*PAF-1*). *PAS-1* encodes a 117 kDa soluble protein that contains two highly conserved putative ATP-binding regions and shows structural similarity to proteins involved in vesicle-mediated protein transport through the secretory pathway and in control of the cell cycle. However the role of this protein in protein transport into peroxisomes is unknown (Erdmann *et al.*, 1991). *PAS-3* encodes a 50 kDa peroxisomal integral membrane protein thought to be anchored to the membrane by a single membrane spanning region near the amino terminus thereby exposing the bulk of the protein to the cytosol (Höhfield *et al.*, 1991). *PAF-1* is a peroxisomal integral membrane protein, containing two potential membrane spanning domains. Limited protease treatment of peroxisomes suggests that the majority of the

protein is exposed to the cytosolic face of peroxisomes (Tsukamoto *et al.*, 1991). Interestingly, a nonsense mutation in *PAF-1* has been shown to be responsible for one episode of Zellweger syndrome (Shimozawa *et al.*, 1992). The functions of these proteins are unknown. Since the clones were identified by genetic complementation of peroxisomal assembly mutants, many roles relating to the biogenesis of peroxisomes for these proteins can be envisaged.

4.7 Summary

The results described herein demonstrate that HDE is transported to peroxisomes in yeast by a carboxy-terminal PTS containing the amino acids Ala-Lys-Ile-COOH. Antibodies directed against a peptide containing this PTS react with several proteins from different yeasts but not with proteins from rat liver. These results suggest that, Ala-Lys-Ile-COOH is a common PTS in yeast. The stage is now set for experiments designed to identify and isolate the putative PTS-I receptor from yeast through a combination of biochemical and genetic means.

5.0 BIBLIOGRAPHY

Adam, S.A. and Gerace L. (1991) Cytosolic proteins that specifically bind nuclear location-signals are receptors for nuclear import. Cell 66, 837-847.

Aflalo, C. (1990) Targeting of cloned firefly luciferase to yeast mitochondria. Biochemistry 29, 4758-4766.

Aitchison, J.D. and Rachubinski, R.A. (1990) *In vivo* import of *Candida tropicalis* hydratase-dehydrogenase-epimerase into peroxisomes of *Candida albicans*. Curr. Genet. 17, 481-486.

Aitchison, J.D., Murray, W.W. and Rachubinski, R.A. (1991a) The carboxy-terminal tripeptide Ala-Lys-Ile is essential for targeting *Candida tropicalis* trifunctional enzyme to yeast peroxisomes. J. Biol. Chem. 266, 23197-23203.

Aitchison, J.D., Sloots, J.A., Nuttley, W.M. and Rachubinski, R.A. (1991b) Sequence of the gene encoding *Candida tropicalis* peroxisomal trifunctional enzyme. Gene 105, 135-136.

Alexson, S.E.H., Fujiki, Y., Shio, H. and Lazarow, P.B. (1985) Partial disassembly of peroxisomes. J. Cell. Biol. 101, 294-304.

Alvares, K., Nemali, M.R., Reddy, P.G., Wang, X., Rao, M.S. and Reddy, J.K. (1989) The nucleotide sequence of a full length cDNA clone encoding rat liver urate oxidase. Biochem. Biophys. Res. Commun. 158, 991-995.

Andrews, D.W. and Rachubinski, R.A. (1990) Secretion and organelle biogenesis in: Tip growth in plant and fungal cells. Heath, I.B. ed., Academic Press, pp. 317-343.

Andrews, G.K., Dziadek, M. and Tamaoki, T. (1982) Expression and methylation of the mouse α-fetoprotein gene in embryonic, adult, and neoplastic tissues. J. Biol. Chem. 257, 5148-5153.

Angermuller, S., Bruder, G., Fahimi, H.D., Volke, A. and Wesch, H. (1987) Localization of xanthine oxidase in crystalline cores of peroxisomes - A cytochemical and biochemical study. Eur. J. Cell Biol. 45, 137-144.

Atomi, H., Ueda, M., Hikida, M., Hishida, T,, Teranishi, Y. and Tanaka, A. (1990) Peroxisomal isocitrate lyase of the *n*-alkane-assimilating yeast *Candida tropicalis*: gene analysis and characterization. J. Biochem. 107, 262-266.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1989) Current Protocols in Molecular Biology Vol. 1, Greene Publishing Associates and Wiley - Interscience, New York.

Aviv, H. and Leder, P. (1977) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA 69, 1408-1412.

Baker, K.P. and Schatz, G. (1991) Mitochondrial proteins essential for viability mediate protein import into yeast mitochondria. Nature 349, 205-208.

Baker, M.E. (1990) A common ancestor for *Candida tropicalis* and dehydrogenases that synthesize antibiotics and steroids. FASEB J. 4, 3028-3032.

Balfe, A., Hoefler, G., Chen, W.W. and Watkins, P.A. (1990) Aberrant subcellular localization of peroxisomal 3-ketoacyl-CoA thiolase in the Zellweger syndrome and rhizomelic chondrodysplasia punctata. Pediatr. Res. 27, 304-310.

Baranski, T.J., Koelsch, G., Hartsuck, J.A. and Kornfeld, S. (1991) Mapping and molecular modeling of a recognition domain for lysosomal enzyme targeting. J. Biol. Chem. 266, 23365-23372.

Barns, S.M., Lane, D.J., Sogin, M.L., Bibeau, C. and Weisburg, W.G. (1991) Evolutionary relationships among pathogenic *Candida* species and relatives. J. Bacteriol. 173, 2250-2255.

Bassham, D.C., Bartling, D., Mould, R.M., Dunbar, B., Weisbeek, P., Hermann, R.G. and Robinson, C. (1991) Transport of proteins into chloroplasts. Delineation of envelope "transit" and thylakoid "transfer" signals within the presequences of three imported thylakoid lumen proteins. J. Biol. Chem. 266, 23606-23610.

Bauerle, C. and Keegstra, K. (1991) Full-length plastocyanin precursor is translocated across isolated thylakoid membranes. J. Biol. Chem. 266, 5876-5883.

Becker, W.M., Riezman, H., Weir, E.M., Titus, D.E. and Leaver, C.J. (1982) *In vitro* synthesis and compartmentalization of glyoxysomal enzymes from cucumber. Ann. N.Y. Acad. Sci. 386, 329-348.

Beeching, J.R. and Northcote, D.H. (1987) Nucleic acid (cDNA) and amino acid sequences of isocitrate lyase from caster bean. Plant Mol. Biol. 8, 471-475.

Beevers, H. (1982) Glyoxysomes in higher plants. Ann. N.Y. Acad. Sci. 386, 243-253.

Beggs, J.D. (1978) Transformation of yeast by a replicating hybrid plasmid. Nature 275, 104-109.

Bellion, E. and Goodman, J.M. (1987) Proton ionophores prevent assembly of a peroxisomal protein. Cell 48, 165-173.

Bennetzen, J.L. and Hall, B.D. (1982) Codon selection in yeast. J. Biol. Chem. 257, 3026-3031.

Bethards, L.A., Scandalios, J.G. and Skadsen, R.W. (1987) Isolation and characterization of a cDNA clone for the cat2 gene in maize and its homology with other catalase. Proc. Natl. Acad. Sci. USA 84, 6830-6834.

Billheimer, J.T., Strehl, L.L., Davis, G.L., Straus, J.F. III, and Davis, L.G. (1990) Characterization of a cDNA encoding rat sterol carrier protein-2. DNA Cell Biol. 9, 159.

Binder, M., Schanz, M. and Hartig, A. (1991) Vector-mediated overexpression of catalase A in the yeast *Saccharomyces cerevisiae* induces inclusion body formation. Eur. J. Cell Biol. 54, 305-312.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7, 1513-1523.

Birnstiel, M.L., Busslinger, M. and Struhl, K (1985) Transcription termination and 3' processing: The end is in site! Cell 41, 349-359.

Blobel, G. (1980) Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA 77, 1496-1500.

Blobel, G. and Sabatini, D.D. (1971) In: Biomembranes Vol. II, Manson, L.A. ed., Plenum Publishing, pp. 193-195.

Blobel, G. and Dobberstein, B. (1975a) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J. Cell Biol. 67, 835-851.

Blobel, G. and Dobberstein, B. (1975b) Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. J. Cell Biol. 67, 852-862.

Bodnar, A.G., and Rachubinski, R.A. (1990) Cloning and sequence determination of cDNA encoding a second rat liver peroxisomal 3-ketoacyl-CoA thiolase. Gene 91, 193-199.

Bodnar, A.G. and Rachubinski, R.A. (1991) Characterization of the integral membrane polypeptides of rat liver peroxisomes isolated from untreated and clofibrate-treated rats. Biochem. Cell Biol. 69, 499-508.

Bonner, W.M. and Laskey, R.A. (1974) A film detection method for tritium-labelled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. 46, 83-88.

Borst, P. (1986) How proteins get into microbodies (peroxisomes, glyoxysomes, glycosomes). Biochim. Biophys. Acta 866, 179-203.

Borst, P. (1989) Peroxisome biogenesis revisited. Biochim. Biophys. Acta 1008, 1-13.

Bout, A., Teunissen, Y., Hashimoto, T., Benne, R. and Tager, J.M. (1988) Nucleotide sequence of human peroxisomal 3-oxoacyl-CoA thiolase. Nucl. Acids Res. 16, 10369.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-284.

Broach, J.R., Strathern, J.N. and Hicks, J.B. (1979) Transformation in yeast: development of a hybrid cloning vector and isolation of the *CAN1* gene. Gene 8, 121-133.

Bruinenberg, P.G., Evers, M., Waterham, H.R., Kuipers, J., Arnberg, A.C. and Geert, A.B. (1989) Cloning and sequencing of the peroxisomal amino oxidase gene from *Hansenula polymorpha*. Biochim. Biophys. Acta 1008, 157-167.

Burnette, W.M. (1981) "Western blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated Protein A. Anal. Biochem. 112, 195-203.

Cannon, B., Alexson, S. and Nedergard, J. (1982) Peroxisomal ß-oxidation in brown fat. Ann. N.Y. Acad. Sci. 386, 40-57.

Cashmore, A., Szabo, L., Timko, M., Kausch, A., Van den Broeck, G., Schreier, P., Bohnert, H., Herrera-Estrella, L., Van Montagu, M. and Schell, J. (1985) Import of polypeptides into chloroplasts. Biotechnology 3, 803-808.

Cavalier-Smith, T. (1987) Molecular evolution: Eukaryotes with no mitochondria. Nature 326, 332-333.

Chance, B., Sies, H. and Boveris, A. (1979) Hydroperoxide metabolism in mammalian organs. Physiol. Rev. 59, 527-605.

Chirico, W.J., Waters, M.G. and Blobel, G. (1988) 70 K heat shock related proteins stimulate protein translocation into microsomes. Nature 332, 805-810.

Cigan, M. and Donahue, T.F. (1987) Sequence and structural features associated with translational initiator regions in yeast - a review, Gene 59, 1-18.

Clarke, J.M. and Switzer, R.L. (1977) In: Experimental Biochemistry, 2nd Ed., W.H. Freeman and Company, pp. 256.

Clark-Walker, G.D., McArthur, C.R. and Sriprakash, K.S. (1985) Location of transcriptional control signals and transfer RNA sequences in *Torulopsis glabrata* mitochondria DNA. EMBO J. 4, 465-473.

Clayton, C.E. (1987) Import of fructose bisphosphate aldolase into the glycosomes of *Trypanosoma brucei*. J. Cell Biol. 105, 2649-2654.

Close, T.J. and Rodriguez, R.L. (1982) Construction and characterization of the chloramphenicol-resistance gene cartridge: A new approach to transcriptional mapping of extrachromosomal elements. Gene 20, 305-316.

Comai, L., Baden, C.S. and Harada, J.J. (1989) Deduced sequence of a malate synthase polypeptide encoded by a subclass of the gene family. J. Biol. Chem. 264, 2778-2782.

Connolly, T., Rapiejko, P.J. and Gilmore, R. (1991) Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. Science 252, 1171-1173.

Cooper, T.G. and Beevers, H. (1969a) Mitochondria and glyoxysomes from castor bean endosperm. J. Biol. Chem. 244, 3507-3513.

Cooper, T.G. and Beevers, H. (1969b) ß-oxidation in glyoxysomes from castor bean endosperm. J. Biol. Chem. 244, 3514-3520.

Cooperstein, S.J. and Lazarow, A. (1951) A microspectrophotometric method for the determination of cytochrome oxidase. J. Biol. Chem. 189, 665-670.

Danks, D.M., Tippet, P., Adams, C. and Campbell, P. (1975) Cerebro-hepato-renal syndrome of Zellweger. A report of eight cases with comments upon the incidence, the liver lesion, and a fault in pipecolic acid metabolism. J. Pediatr. 86, 382-387.

Datta, N.S., Wilson, G.N. and Hajra, A.K. (1984) Deficiency of enzymes catalyzing the biosynthesis of glycerol-ether lipids in Zellweger syndrome. N. Engl. J. Med. 311, 1080-1083.

Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) In: Basic Methods in Molecular Biology, Elsevier Science Publishing Co., New York, N.Y.

Dean, N. and Pelham, H.R.B. (1990) Recycling of proteins from the Golgi compartment to the ER in yeast. J. Cell Biol. 111, 369-377.

de Boer, A.D. and Weisbeek, P.J. (1991) Chloroplast protein topogenesis: import, sorting and assembly. Biochim. Biophys. Acta. 1071, 221-253.

de Duve, C. (1965) Functions of microbodies (peroxisomes). J. Cell Biol. 27, 25a.

de Duve, C. (1970) Evolution of the peroxisome. Ann. N.Y. Acad. Sci. 168, 369-381.

de Duve, C. (1982) Peroxisomes and related particles in historical perspective. Ann. N.Y. Acad. Sci. 386, 1-4.

de Duve, C. (1983) Microbodies in the living cell. Sci. Amer. 248 (5), 74-84.

de Duve, C. and Baudhuin, P. (1966) Peroxisomes (microbodies and related particles). Physiol. Rev. 46, 323-357.

de Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R. and Appelmans, F. (1955) Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat liver tissues. Biochem. J. 60, 604-617.

Delorme, E. (1989) Transformation of *Saccharomyces cerevisiae* by electroporation. Appl. Environ. Microbiol. 55 2242-2246.

del Valle, R., Soto, U., Necochea, C. and Leighton, F. (1988) Detection of an ATPase activity in rat liver peroxisomes. Biochem. Biophys. Res. Commun. 156, 1353-1359.

Desel, H., Zimmerman, R., Janes, M., Miller, F. and Neupert, W. (1982) Biosynthesis of glyoxysomal enzymes in *Neurosporra crassa*. Ann. N.Y. Acad. Sci. 386, 377-388.

Deshaies, R.J., Koch, B.D., Werner-Washburne, M., Craig, E.A. and Schekman, R. (1988) A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. Nature 332, 800-805.

Devereux, J., Haeberli, P. and Smithies, P. (1984) A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. 12, 387-395.

de Wet, J.R., Wood, K.W., DeLuca, M., Helsinki, D,R. and Subramani, S. (1987) Firefly luciferase gene-structure and expression in mammalian cells. Mol. Cell. Biol. 7, 725-737.

Distel, B., Veenhuis, M. and Tabak, H.F. (1987) Import of alcohol oxidase into peroxisomes of *Saccharomyces cerevisiae*. EMBO J. 6, 3111-3116.

Distel, B., Van der Ley, I., Veenhuis, M. and Tabak, H.F. (1988) Alcohol oxidase expressed under nonmethylotrophic conditions is imported, assembled, and enzymatically active in peroxisomes of *Hansenula polymorpha*. J. Cell Biol. 107 (5), 1669-1675.

Dmochowska, A., Dignard, D., Maleszka, R. and Thomas, D.Y. (1990) Structure and transcriptional control of the *Saccharomyces cerevisiae POX 1* gene encoding acyl-Coenzyme A oxidase. Gene 88, 247-252.

Doan, N.P. and Fincher, G.B. (1988) The A- and B-chains of carboxypeptidase I from germinated barley originate from a single precursor polypeptide. J. Biol. Chem. 263, 11106-11110.

Douma, A.C., Veenhuis, M., Sulter, G.J. and Harder, W. (1987) A proton-translocating adenosine triphosphatase is associated with the peroxisomal membrane of yeasts. Arch. Microbiol. 147, 42-47.

Douma, A.C., Veenhuis, M., Waterham, H.R. and Harder, W. (1990) Immunological demonstration of the peroxisomal ATPase of yeasts. Yeast 6, 45-51.

Dugaiczyk, A., Boyer, H.W. and Goodman, H.M. (1975) Ligation of *Eco*RI endonuclease-generated DNA fragments into linear and circular structures. J. Mol. Biol. 96, 171-184.

Eilers, M. and Schatz, G. (1988) Protein unfolding and the energetics of protein translocation across biological membranes. Cell 52, 481-483.

Eilers, M., Hwang, S. and Schatz, G. (1988) Unfolding and refolding of a purified precursor protein during import into isolated mitochondria. EMBO J. 7, 1139-1145.

Eising, R., Trelease, R.N. and Hi, W. (1990) Biogenesis of catalase in glyoxysomes and leaf-type peroxisomes of sunflower catyledons. Arch. Biochem. 278, 258-264.

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

Erdmann, R., Wiebel, F.F., Flessau, A., Rytka, J., Beyer, A., Fröhlich, K.-U. and Kunau, W.-H. (1991) *PAS 1*, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. Cell 64, 499-510.

Essner, E. (1969) Localization of peroxidase activity in microbodies of fetal mouse liver. J. Histochem. Cytochem. 17, 454-466.

Fearon, K. and Mason, T.L. (1988) Structure and regulation of a nuclear gene in *Saccharomyces cerevisiae* that specifies MRP7, a protein of the large subunit of the mitochondrial ribosome. Mol. Cell. Biol. 8, 3636-3646.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13.

Feinberg, A.P. and Vogelstein, B. (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137, 266-267.

Forsburg, S.L. and Guarante, L. (1988) Mutational analysis of upstream activation sequence 2 of the *cyc 1* gene of *Saccharomyces cerevisiae*: a *HAP2-HAP3*-responsive site. Mol. Cell. Biol. 8, 647-654.

Frischauf, A-M., Lebrach, M., Poustka, A. and Murray, N. (1983) Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. 170, 827-842.

Fujiki, Y. and Lazarow, P.B. (1985) Post-translational import of fatty acyl-CoA oxidase and catalase into peroxisomes of rat liver *in vitro*. J. Biol. Chem. 260, 5603-5603.

Fujiki, Y., Fowler, S., Shio, H., Hubbard, A.L. and Lazarow, P.B. (1982) Polypeptide and phospholipid composition of the membranes of rat liver peroxisomes: Comparison with endoplasmic reticulum and mitochondrial membranes. J. Cell Biol. 93, 103-110.

Fujiki, Y., Rachubinski, R.A. and Lazarow, P.B. (1984) Synthesis of a major integral membrane polypeptide of rat liver peroxisomes on free polysomes. Proc. Natl. Acad. Sci. USA 81, 7127-7131.

Fujiki, Y., Rachubinski, R.A., Mortensen, R.M. and Lazarow, P.B. (1985) Synthesis of 3-ketoacyl-CoA thiolase of rat liver peroxisomes on free polyribosomes as a larger precursor. Biochem. J. 226, 697-704.

Fujiki, Y., Rachubinski, R.A., Zentella-Dehesa, A. and Lazarow, P.B. (1986) Induction, identification, and cell-free translation of mRNAs coding for peroxisomal proteins in *Candida tropicalis*. J. Biol. Chem. 261, 15787-15793.

Fujiki, Y., Tsuneoka, M. and Tashiro, Y. (1989) Biosynthesis of nonspecific lipid transfer protein (sterol carrier protein 2) on free polyribosomes as a larger precursor in rat liver. J. Biochem. 106, 1126-1131.

Fukui, S. and Tanaka, A. (1979) Yeast peroxisomes. Trends Biochem. Sci. 4, 246-249.

Fuller, R.S., Wilcox, C.A. and Redding, K. (1991) Disruption of the Golgi-retention signal in Kex2 protease: default targeting to the vacuole. Yeast Cell Biol., Cold Spring Harbor, 136.

Fung, K. and Clayton, C. (1991) Recognition of a peroxisomal tripeptide entry signal by the glycosomes of *Trypanosoma brucei*. Mol. Biochem. Parasitol. 45, 261-264.

Furuta, S., Miyazawa, S. and Hashimoto, T. (1982) Biosynthesis of enzymes of peroxisomal ß-oxidation. J. Biochem. 92, 319-326.

Furuta, S., Hayashi, H., Hijikata, M., Miyazawa, S., Osumi, T. and Hashimoto, T. (1986) Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liver catalase. Proc. Natl. Acad. Sci. USA 83, 313-317.

Garcia-Bostos, J., Heitman, J. and Hall, M.N. (1991) Nuclear protein localization. Biochim. Biophys. Acta. 1071, 83-101.

Garrard, L.J. and Goodman, J.M. (1989) Two genes encode the major membrane-associated protein of methanol-induced peroxisome from *Candida boidinii*. J. Biol. Chem. 264, 13929-13937.

Gavel, Y. and von Heijne, G. (1990) A conserved cleavage motif in chloroplast transit peptides. FEBS Lett. 261, 455-458.

Gerdes, H.-H., Behrends, W. and Kindl, H. (1982) Biosynthesis of a microbody matrix enzyme in greening catyledons. Glycollate oxidase synthesized *in vivo* and *in vitro*. Planta 156, 572-578.

Gietl, C. (1990) Glyoxysomal malate dehydrogenase from watermelon is synthesized with an amino-terminal transit peptide. Proc. Natl. Acad. Sci. USA. 87, 5773-5777.

Gietl, C. and Hock, B. (1982) Organelle-bound malate dehydrogenase isozymes are synthesized as higher molecular weight precursors. Plant Physiol. 70, 483-487.

Gietl, C. and Hock, B. (1984) Import of *in vitro*-synthesized glyoxysomal malate hydrogenase into isolated watermelon glyoxysomes. Planta 162, 261-267.

Gilmore, R., Walter, P. and Blobel, G. (1982) Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. J. Cell Biol. 95, 470-477.

Goldfischer, S., Moore, C.L., Johnson, A.B., Spiro, A.J., Valsamis, M.P., Wisniewski, H.K., Ritch, R.H., Norton, W.T., Rapin, I. and Gartner, L.M. (1973) Peroxisomal and mitochondrial defects in the cerebro-hepato-renal syndrome. Science 182, 62-64.

Goldman, B.M. and Blobel, G. (1978) Biogenesis of peroxisomes: Intracellular site of synthesis of catalase and uricase. Proc. Natl. Acad. Sci. USA 75, 5066-5070.

Gonzalez, E. (1991) The C-terminal domain of plant catalases. Eur. J. Biochem. 199, 211-215.

Goodman, D.B.P., Davis, W.L. and Jones, R.G. (1980) Glyoxylate cycle in toad urinary bladder: Possible stimulation by aldosterone. Proc. Natl. Acad. Sci. USA 77, 1521-1525.

Goodman, J.M., Scott, C.W., Donahue, P.N. and Atherton, J.P. (1984) Alcohol oxidase assembles post-translationally into the peroxisome of *Candida boidinii*. J. Biol. Chem. 259, 8485-8493.

Gorgas, K. (1984) Peroxisomes in sebaceous glands. V. Complex peroxisomes in the mouse preputial gland: Serial secretory and three-dimensional reconstruction studies. Anat. Embrol. 169, 261-270.

Gould, S.J., Keller, G.-A. and Subramani, S. (1987) Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. J. Cell Biol. 105, 2923-2931.

Gould, S.J., Keller, G.-A. and Subramani, S. (1988) Identification of a peroxisomal targeting signal located at the carboxy terminus of four peroxisomal proteins. J. Cell Biol. 107, 897-905.

Gould, S.J., Keller, G.-A., Hosken, M., Wilkinson, J. and Subramani, S. (1989) A conserved tripeptide sorts proteins to peroxisomes. J. Cell Biol. 108, 1657-1664.

Gould, S.J., Krisans, S., Keller, G.-A. and Subramani, S. (1990a) Antibodies directed against the peroxisomal targeting signal of firefly luciferase recognize multiple mammalian peroxisomal proteins. J. Cell Biol. 110, 27-34.

Gould, S.J., Keller, G.-A., Schneider, M., Howell, S.H., Garrard, L.J., Goodman, J.M., Distel, B., Tabak, H. and Subramani, S. (1990b) Peroxisomal protein import is conserved between yeast, plants, insects and mammals. EMBO J. 9, 85-90.

Grunstein, M. and Hogness, D.S. (1975) Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA 72, 3961-3965.

Hagenbüchle, O., Bovey, R., and Young, R.A. (1980) Tissue specific expression of mouse α-amylase genes: nucleotide sequence of isozyme mRNAs from pancreas and salivary gland. Cell 21, 179-187.

Hajra, A.K. and Bishop, J.E. (1982) Glycerolipid biosynthesis in peroxisomes via the acyl-dihydroxyacetone phosphate pathway. Ann. N.Y. Acad. Sci. 386, 170-182.

Hanahan, D. and Meselson, M. (1980) Plasmid screening at high colony density. Gene 10, 63-67.

Hanish, J. and McClelland, M. (1988) Activity of DNA modification and restriction enzymes in KGB, a potassium glutamate buffer. Gene Anal. Techn. 5, 105-107.

Hansen, H. and Roggenkamp, R. (1989) Functional complementation of catalase-defective peroxisomes in a methylotrophic yeast by import of the catalase A from *Saccharomyces cerevisiae*. Eur. J. Biochem. 184, 173-179.

Hanson, R.F., Szczepanik-Van Leeuwen, P., Williams, G.C., Grabowski, G. and Sharp, H.L. (1979) Defects of bile acid synthesis in Zellweger's syndrome. Science 203, 1107-1108.

Harley, C.B. (1987) Hybridization of oligo(dT) to RNA on nitrocellulose. Gene Anal. Techn. 4, 17-22.

Hartig, A., Ogris, M., Cohen, G. and Binder, M. (1990) Fate of highly expressed proteins destined to peroxisomes in *Saccharomyces cerevisiae*. Curr. Genet. 18, 23-27.

Hartl, F.-U., Pfanner, N., Nicholson, D.W. and Neupert, W. (1989) Mitochondrial protein import. Biochim. Biophys. Acta 988, 1-45.

Hashimoto, T. (1982) Individual peroxisomal ß-oxidation enzymes. Ann. N.Y. Acad. Sci. 386, 5-12.

Hashimoto, F. and Hayashi, H. (1987) Significance of catalase in peroxisomal fatty acyl-CoA ß-oxidation. Biochim. Biophys. Acta 921, 142-150.

Hawlitschek, G., Schneider, H., Schmidt, B., Trapschug, M., Hartl, F.-U. and Neupert, W. (1988) Mitochondrial protein import: Identification of processing peptidase and of PEP, a processing enhancing protein. Cell 53, 795-806.

Henikoff, S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28, 351-359.

Henikoff, S. and Cohen, E.H. (1984) Sequence responsible for transcription termination on a gene segment in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 4, 1515-1520.

Hess, R., Stäubli, W. and Riess, W. (1965) Nature of the hepatomegalic effect produced by ethyl-chlorophenoxy-isobutyrate in the rat. Nature 208, 856-858.

Heymans, H.S.A., Schutgens, R.B.H., Tan, R., van den Bosch, H. and Borst, P. (1983) Severe plasmalogen deficiency in tissues of infants without peroxisomes (Zellweger syndrome). Nature 306, 69-70.

Higashi, T. and Peters, T. Jr. (1963) Studies on rat liver catalase II. Incorporation of C14-leucine into catalase of liver cell fractions *in vivo*. J. Biol. Chem. 238, 3952-3954.

Hijikata, M., Ishii, N., Kagamiyama, H., Osumi, T. and Hashimoto, T. (1987) Structural analysis of cDNA for rat peroxisomal 3-ketoacyl-CoA thiolase. J. Biol. Chem. 262, 8151-8158.

Hijikata, M., Wen, J.-K., Osumi, T. and Hashimoto, T. (1990) Rat peroxisomal 3-ketoacyl-CoA thiolase gene. J. Biol. Chem. 265, 4600-4606.

Hill, D.E., Boulay, R. and Rogers, D. (1988) Complete nucleotide sequence of the peroxisomal acyl-CoA oxidase from the alkane-utilizing yeast *Candida maltosa*. Nucl. Acids Res. 16, 365-366.

Hiltunen, J.K., Palosaari, P.M. and Kunau, W.-H. (1989) Epimerization of 3-hydroxyacyl-CoA esters in rat liver: involvement of two 2-enoyl-CoA hydratases. J. Biol. Chem. 264, 13536-13540.

Höhfeld, J., Veenhuis, M. and Kunau, W. (1991) PAS 3, a *Saccharomyces cerevisiae* gene encoding a peroxisomal integral membrane protein essential for peroxisome biogenesis. J. Cell Biol. 114, 1167-1178.

Hohn, B. and Murray, K. (1977) Packaging recombinant DNA molecules into bacteriophage particles *in vitro*. Proc. Natl. Acad. Sci. USA 74, 3259-3263.

Holm, C., Meeks-Wagner, D.W., Fangman, W.L. and Botstein, D. (1986) A rapid, efficient method for isolating DNA from yeast. Gene 42, 169-173.

Holt, C.E., Garlick, N. and Cornel, E. (1990) Lipofection of cDNAs in the embryonic central nervous system. Neuron. 4, 203-214.

Horii, A., Emi, M., Tomita, M., Nishide, T., Ogawa, M., Mori, T. and Matsubara, K. (1987) Primary structure of human pancreatic α-amylase gene and its comparison with human salivary α-amylase gene. Gene 60, 57-64.

Hruban, Z., Swift, H. and Wissler, R.W. (1963) Alterations in the fine structure of hepatocyes produced by ß-3-thienylalanine. J. Ultrastruct. Res. 8, 236-250.

Hruban, Z., Vigil, E.L., Stesers, A. and Hopkins, E. (1972) Microbodies constituent organelles of animal cells. Lab. Invest. 27, 184-191.

Huynh, T.V., Young, R.A. and Davis, R.W. (1985) DNA Cloning: A Practical Approach, Vol. 1 (D.M. Glover, ed.) IRL Press, Oxford, pp. 49-78.

Imanaka, T., Small, G.M. and Lazarow, P.B. (1987) Translocation of acyl-CoA oxidase into peroxisomes requires ATP hydrolysis but not a membrane potential. J. Cell Biol. 105, 2915-2922.

Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. 153, 163-168.

Jackson, M.R., Nilsson, T. and Peterson, P.A. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. EMBO J. 9, 3153-3162.

Jang, S.-H., Cheesbrough, T.M. and Kolattukudy, P.E. (1989) Molecular cloning, nucleotide sequence, and tissue distribution of malonyl-CoA decarboxylase. J. Biol. Chem. 264, 3500-3505.

Janowicz, Z.A., Eckart, M.R., Drewke, G., Roggenkamp, R.P. and Hollenberg, C.P. (1985) Cloning and characterization of the DAS gene encoding the major methanol assimilatory enzyme for the methylotrophic yeast *Hansenula polymorpha*. Nucl. Acids. Res. 13, 3043-3062.

Jarnik, M., Aebi, U. and Ris, H. (1991) Towards a complete 3-D model of the nuclear pore complex (NPC). J. Cell Biol. 115, 458a.

Just, W.W. and Hartl, F.-U. (1987) Peroxisomes in Biology and Medicine (H.D. Fahimi and H. Sies, eds.) Springer-Verlag, Berlin, Germany, pp. 402-416.

Kalderon, D., Richardson, W.D., Markham, A.F. and Smith, A.E. (1984a) Sequence requirements for nuclear location of simian virus 40 large-T antigen. Nature 311, 33-38.

Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984b) A short amino acid sequence able to specify nuclear location. Cell 39, 499-509.

Kamijo, K., Taketani, S., Yokota, S., Osumi, T. and Hashimoto, T. (1990) The 70-kDa peroxisomal membrane protein is a member of the Mdr (P-glycoprotein)-related ATP-binding protein superfamily. J. Biol. Chem. 265, 4534-4540.

Kamiryo, T. and Okazaki, K. (1984) High-level expression and molecular cloning of genes encoding *Candida tropicalis* peroxisomal proteins. Mol. Cell. Biol. 4, 2136-2141.

Kamiryo, T., Abe, M., Okazaki, K., Kato, S. and Shimamoto, N. (1982) Absence of DNA in peroxisomes of *Candida tropicalis*. J. Bacteriol. 152, 269-274.

Kamiryo, T., Mito, N., Niki, T. and Suzuki, T. (1991) Assignment of most genes encoding major peroxisomal polypeptides to chromosomal band V of the asporogenic yeast *Candida tropicalis*. Yeast 7, 503-511.

Kawamoto, S., Nozaki, C., Tanaka, A. and Fukui, S. (1978) Fatty acid ß-oxidation systems in microbodies of *n*-alkane-grown *Candida tropicalis*. Eur. J. Biochem. 83, 609-613.

Keegstra, K. (1989) Transport and routing of proteins into chloroplasts. Cell 56, 247-253.

Keller, G.-A., Gould, S., Deluca, M. and Subramani, S. (1987) Firefly luciferase is targeted to peroxisomes in mammalian cells. Proc. Natl. Acad. Sci. USA 84, 3264-3268.

Keller, G.-A., Scallen, T.J., Clarke, D., Maher, P.A., Krisans, S.K. and Singer, S.J. (1989) Subcellular localization of sterol carrier protein-2 in rat hepatocytes: Its primary localization to peroxisomes. J. Cell Biol. 108, 1353-1361.

Keller, G.A., Grisansm, S., Gould, S J., Sommer, J.M., Wang, C.C., Schliebs, W., Kunau, W., Brody, S. and Subramani, S. (1991a) Evolutionary conservation of a microbody targeting signal that targets proteins to peroxisomes, glyoxysomes, and glycosomes. J. Cell. Biol. 114, 893-904.

Keller, G.-A., Warner, T.G., Steimer, K.S. and Hallewell, R.A. (1991b) Cu, Zn superoxide dismutase is a peroxisomal enzyme in human fibroblasts and hepatoma cells. Proc. Natl. Acad. Sci. USA 88, 7381-7385.

Kelly, R., Miller, S.M., Kurtz, M.B. and Kirsch, D.R. (1987) Directed mutagenesis in *Candida albicans*: one-step gene disruption to isolate *ura3* mutants. Mol. Cell. Biol. 7, 190-208.

Kendall, G., Wilderspin, A.F., Ashall, F., Miles, M.A. and Kelly, J.M. (1990) *Trypanosoma cruzi* glycosomal glyceraldehyde-3-phospnate dehydrogenase does not conform to the 'hotspot' topogenic signal model. EMBO J. 9, 2751-2758.

Kilponen, J.M., Palosaari, P.M. and Hiltunen, J.K (1990) Occurrence of a long-chain Δ^3,Δ^2-enoyl-CoA isomerase rat liver. Biochem. J. 269, 223-226.

Kirwin, P.M., Elderfield, P.D. and Robinson, C. (1987) Transport of proteins into chloroplasts. Partial purification of a thylakoidal processing peptidase involved in plastocyanin biogenesis. J. Biol. Chem. 262, 16386-16390.

Kitagawa, T. and Aikawa, T. (1976) Enzyme coupled immunoassay of insulin using a novel coupling reagent. J. Biochem. 79, 233-236.

Klionsky, D.J., Banta, L.M. and Emr. S.D. (1990) Intracellular sorting and processing of a yeast vacuolar hydrolase: the proteinase A propeptide contains vecuolar targeting information. Mol. Cell. Biol. 8, 2105-2116.

Kornberg, H.L. and Beevers, H. (1957) A mechanism of conversion of fat to carbohydrate in castor beans. Nature 180, 35-36.

Kornfeld, S. and Mellman, I. (1989) The biogenesis of lysosomes. Ann. Rev. Cell Biol. 5, 483-525.

Köster, A., Heisig, M., Heinrich, P.C. and Just, W.W. (1986) *In vitro* synthesis of peroxisomal membrane polypeptides. Biochem. Biophys. Res. Commun. 137, 626-632.

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283-292.

Kozak, M. (1989) The scanning model for translation: an update. J. Cell Biol. 108, 229-241.

Kreike, J., Schulze, M., Ahne, F., and Franz-Lang, B. (1987) A yeast nuclear gene, MRS 1 involved in mitochondrial RNA splicing: nucleotide sequence and mutational analysis of two overlapping open reading frames on opposite strands. EMBO J. 6, 2123-2129.

Kruse, C., Frevert, J. and Kindl, H. (1981) Selective uptake by glyoxysomes of *in vitro* translated malate synthase. FEBS Lett. 129, 36-38.

Kruse, C. and Kindl, H. (1983) Oligomerization of malate synthase during glyoxysome biosynthesis. Arch. Biochem. Biophys. 223, 629-638.

Kunau, W.-H., Bühne, S., de la Garza, M., Kionka, C., Mateblowski, M., Schultz-Borchard, U. and Thieringer, R. (1988) Comparative enzymology of β-oxidation. Biochem. Soc. Trans. 16, 418-420.

Kuo, C.-L. and Campbell, J.L. (1983) Cloning of *Saccharomyces cerevisiae* DNA replication genes: isolation of the *CDC8* gene and two genes that compensate for the *cdc8-1* mutation. Mol. Cell. Biol. 3, 1730-1737.

Kurihara, T., Ueda, M. and Tanaka, A. (1989) Peroxisomal acetoacetyl-CoA thiolase and 3-ketoacyl-CoA thiolase from *n*-alkane-utilizing yeast, *Candida tropicalis*: purification and characterization. J. Biochem. 106, 474-478.

Kurtz, M.B., Cortelyou, M.W. and Kirsch, D.R. (1986) Integrative transformation of *Candida albicans*, using a cloned *Candida ADE2* gene. Mol. Cell. Biol. 6, 142-149.

Kurtz, M.B., Cortelyou, M.W., Miller, S.M. and Kirsch, D.R. (1987) Development of autonomously replicating plasmids for *Candida albicans*. Mol. Cell. Biol. 7, 209-217.

Kwiatowski, J. and Ayala, F.J. (1989) *Drosophila virilus* Cu-Zn superoxide dismutase gene sequence. Nucl. Acids. Res. 17, 2133.

Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lanford, R.E., Kanda, P. and Kennedy, R.C. (1986) Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. Cell 46, 575-582.

Lazarow, P.B. (1978) Rat liver peroxisome catalyze the ß-oxidation of fatty acids. J. Biol. Chem. 253, 1522-1528.

Lazarow, P.B. (1987) What is a peroxisome? Toxicol. Industr. Health 3, 1-5.

Lazarow, P.B. and Fujiki, Y. (1985) Biogenesis of peroxisomes. Ann. Rev. Cell Biol. 1, 489-530.

Lazarow, P.B. and de Duve, C. (1973) Synthesis and turnover of rat liver peroxisomes. J. Cell Biol. 59, 507-524.

Lazarow, P.B. and de Duve, C. (1976) A fatty acyl-CoA oxidizing system in rat liver peroxisomes: enhancement by clofibrate, a hypolipidemic drug. Proc. Natl. Acad. Sci. USA 73, 2043-2046.

Lazarow, P.B., Robbi, M., Fujiki, Y. and Wong, L. (1982) Biogenesis of peroxisomal proteins *in vivo* and *in vitro*. Ann. N.Y. Acad. Sci. 386, 285-300.

Ledeboer, A.M., Edens, L., Maat, J., Visser, C., Bos, J.W. and Verrips, C.T. (1985) Molecular cloning and characterization of a gene coding for methanol oxidase in *Hansenula polymorpha*. Nucl. Acids Res. 13, 3063-3082.

Legg, P.G. and Wood, R.L. (1970) New observations on microbodies. J. Cell Biol. 45, 118-129.

Leighton, F., Poole, B., Beaufay H., Baudhuin, P., Coffey, J.W., Fowler, S. and de Duve, C. (1968) The large-scale separation of peroxisomes, mitochondria, and lysosomes from the livers of rats injected with Triton WR-1339. J. Cell Biol. 37, 482-512.

Levanon, D., Lieman-Hurwitz, J., Dafni, M., Widderson, M., Sherman, L., Bernstein, Y., Laver-Rudich, Z., Danciger, E., Stein, O. and Groner, Y. (1985) Architecture and anatomy of the chromosomal locus in human chromosome 21 encoding the Cu/Zn superoxide dismutase. EMBO J. 4, 77-84.

Lewin, A.S., Hines, V. and Small, G. (1990) Citrate synthase encoded by the *CIT2* gene of *Saccharomyces cerevisiae* is peroxisomal. Mol. Cell. Biol. 10, 1399-1405.

Lewis, M.J. and Pelham, H.R.B. (1990) A human homologue of the yeast HDEL receptor. Nature 348, 162-163.

Lewis, M.J., Sweet, D.J. and Pelham, H.R.B. (1990) The *ERD2* gene determines the specificity of the luminal ER protein retention system. Cell 61, 1359-1363.

Liu, X., Freeman, K.B. and Shore, G.C. (1990) An amino-terminal signal sequence abrogates the intrinsic membrane-targeting information of mitochondrial uncoupling protein. J. Biol. Chem. 265, 9-12.

Lock, E.A., Mitchell, A.M. and Elcombe, C.R. (1989) Biochemical mechanisms of induction of hepatic peroxisome proliferation. Ann. Rev. Pharmacol. Toxicol. 29, 145-163.

Lord, J.M. and Roberts, L.M. (1982) Glyoxysome biogenesis via the endoplasmic reticulum in castor bean endosperm. Ann. N.Y. Acad Sci. 386, 362-374.

Lubben, T.H., Bansberg, J. and Keegstra, L. (1987) Stop-transfer regions do not halt translocation of proteins into chloroplasts. Science 238, 1112-1113.

Luche, R.M., Sumrada, R. and Cooper, T.G. (1990) A *cis*-acting element present in multiple genes serves as a repressor protein binding site for the yeast *CAR1* gene. Mol. Cell. Biol. 10, 3889-3895.

Lueers, G., Beier, K., Hashimoto, T., Fahimi, H.D. and Voelkl, A. (1990) Biogenesis of peroxisomes: sequential biosynthesis of the membrane and matrix proteins in the course of hepatic regeneration. Eur. J. Cell Biol. 52, 175-184.

Maccecchini, M.-L., Rudin, Y., Blobel, G. and Schatz, G. (1979) Import of proteins into mitochondria: Precursor forms of the extramitochondrially made F_1-ATPase subunits in yeast. Proc. Natl. Acad. Sci. USA 76, 343-347.

Mallett, S., Fossum, S., and Barclay, A.N. (1990) Characterization of the MRC OX40 antigen of activated CD4 positive T lymphocytes--a molecule related to nerve growth factor receptor. EMBO J 9:1063-1068

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Laboratory, Cold Spring Harbor, NY.

Mannaerts, G.P. and Debeer, L.J. (1982) Mitochondrial and peroxisomal ß-oxidation of fatty acids in rat liver. Ann. N.Y. Acad. Sci. 386, 30-39.

Mans, R.J. and Novelli, G.D. (1961) Measurement of the incorporation of radioactive amino acids into protein by a filter-paper disk method. Arch. Biochem. Biophys. 94, 48-53.

Marchand, M., Kooystra, I., Wierenga, R.K., Lambier, A., Van Beeumen, J., Opperdoes, F.R. and Michels, P.A.M. (1989) Glucose phosphate isomerase from *Trypanosoma brucei*. Eur. J. Biochem. 184, 455-464.

Marshall, J.S., De Rocher, A.E., Keegstra, K. and Vierling, E. (1990) Identification of heat shock protein hsp70 homologues in chloroplasts. Proc. Natl. Acad. Sci. USA 87, 374-378.

Marzioch, M., Erdmann, R., Leismann, D. and Kunau, W.-H. (1991) Two yeast genes required for peroxisomal import of 3-oxoacyl CoA thiolase. Yeast Cell Biology, Cold Spring Harbor, p. 21.

Masters, C. and Holmes, R. (1977) Peroxisomes: new aspects of cell physiology and biochemistry. Physiol. Rev. 57, 816-882.

Masuda, T., Tatsumi, H. and Nakano, E. (1989) Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly *Luciola cruciata*. Gene 77, 265-270.

Matsushita, Y., Kitakawa, M. and Isono, K. (1989) Cloning and analysis of the nuclear genes for two mitochondrial ribosomal proteins in yeast. Mol. Gen. Genet. 219, 119-124.

McCammon, M.T., Dowds, C.A., Orth, K., Moomaw, C.R., Slaughter, C.A. and Goodman, J.M. (1990) Sorting of a peroxisomal membrane protein PMP 47 from *Candida boidinii* into peroxisomal membranes of *Saccharomyces cerevisiae*. J. Biol. Chem. 265, 20098-20105.

McDonald, R.J., Drerar, M.M., Swain, W.F., Pictet, R.T., Thomas, G., and Rutter, W.J. (1980) Structure of a family of rat amylase genes. Nature 287, 117-122.

McGuinness, M.C., Moser, A.B., Moser, H.W. and Watkins, P.A. (1990) Peroxisomal disorders: complementation analysis using beta-oxidation of very long-chain fatty acids. Biochem. Biophys. Res. Commun. 172, 364-369.

Meusch, A., Hartmann, E., Rohde, K., Rubatelli, A., Sitia, R. and Rapoport, T.A. (1990) A novel pathway for secretory proteins? Trends Biochem. Sci. 15, 86-88.

Meyer, D.I. (1988) Preprotein conformation: the year's major theme in translocation studies. Trends Biochem. Sci.13, 471-474.

Meyer, D.I. (1991) Protein translocation into the endoplasmic reticulum: a light at the end of the tunnel. Trends Cell Biol. 1, 154-159.

Meyer, D.I. and Dobberstein B. (1980) A membrane component essential for vectoriai translocation of nascent proteins across the endoplasmic reticulum: Requirements for its extraction and reassociation with the membrane. J. Cell Biol. 87, 498-502.

Michels, P.A.M., Poliszczak, A., Osinga, K.A., Misset, O., Van Beeumen, J., Wierenga, R.J., Borst, P. and Opperdoes, F.R. (1986) Two tandemly linked identical genes code for the glycosomal glyceraldehyde-phosphate dehydrogenase in *Trypanosoma brucei*. EMBO J. 5, 1049-1056.

Mihalik, S.J., Moser, H.W., Watkins, P.A., Danks, D.M., Poulos, A. and Rhead, W.H. (1989) Peroxisomal L-pipecolic acid oxidation is deficient in liver from Zellweger syndrome patients. Pediatr. Res. 25, 548-552.

Miura, S., Mori, M., Takiguchi, M. Tatibana, M., Furuta, S., Miyazawa, S. and Hashimoto, T. (1984) Biosynthesis and intracellular transport of enzymes of peroxisomal β-oxidation. J. Biol. Chem. 259, 6397-6402.

Miyazawa, S., Furuta, S, Osumi, T., Hashimoto, T. and Vi, N. (1981) Properties of peroxisomal 3-ketoacyl-CoA thiolase from rat liver. J. Biochem. 90, 511-519.

Miyazawa, S., Hayashi, H., Hijikata, M., Ishii, N., Furuta, S., Kagamiyama, H., Osumi, T. and Hashimoto, T. (1987) Complete nucleotide sequence of cDNA and predicted amino acid sequence of rat acyl-CoA oxidase. J. Biol. Chem. 262, 8131-8137.

Miyazawa, S., Osumi, T., Hashimoto, T., Ohno, K., Miura, S. and Fujiki, Y. (1989) Peroxisome targeting signal of rat liver acyl-coenzyme A oxidase resides at the carboxy terminus. Mol. Cell. Biol. 9, 83-91.

Momoi, K., Fukui, K., Watanabe, F. and Miyake, Y. (1988) Molecular cloning and sequence analysis of cDNA encoding human kidney D-amino acid oxidase. FEBS Lett. 238, 180-184.

Moreno de la Garza, M., Schultz-Borchard, U., Crabb, J.W. and Kunau, W.-H. (1985) Peroxisomal ß-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, 285-291.

Morris, H.R., Larsen, B.S. and Billheimer, J.T. (1988) A mass spectrometric study of the structure of sterol carrier protein SCP2 from rat liver. Biochem. Biophys. Res. Commun, 154: 476-482.

Mortensen, R.M., Rachubinski, R.A., Fujiki, Y. and Lazarow, P.B. (1984) Heating RNA before cell free translation is essential for the efficient and reproducible synthesis of several peroxisomal proteins. Biochem. J. 223, 547-550.

Moser, A.E., Singh, I., Brown, F.R., Solish, G.I., Keeley, R.I., Benke, P.J. and Moser, H.W. (1984) The cerebrohepatorenal (Zellweger) syndrome. Increased levels and impaired degradation of very-long-chain fatty acids and their use in prenatal diagnosis. N. Engl. J. Med. 310, 1141-1146.

Moser, H.W. (1986) Peroxisomal disorders. J. Pediatr. 108, 89-91.

Moser, H.W. (1987) New approaches in peroxisomal disorders. Dev. Neurosci. 9, 1-18.

Moser, H.W. (1988) The peroxisome: Nervous system role of a previously underrated organelle. Neurology 38, 1617-1627.

Moser, H.W., Bergin, A. and Cornblath, D. (1991) Peroxisomal disorders. Biochem. Cell Biol. 69, 463-474.

Munro, S. (1991) Sequences within and adjacent to the transmembrane segment of α-2,6-sialyltransferase specify Golgi retention. EMBO J. 10, 3577-3588.

Munro, S. and Pelham, H.R.B. (1987) A C-terminal signal prevents secretion of luminal ER proteins. Cell 48, 899-907.

Murakami, K. and Mori, M. (1990) Purified presequence binding factor (PbF) forms an import-competent complex with a purified mitochondrial precursor protein. EMBO J. 9, 3201-3208.

Murray, W.W., and Rachubinski, R.A. (1987a) The primary structure of a peroxisomal fatty acyl-CoA oxidase from the yeast *Candida tropicalis* pK233. Gene 51, 119-128.

Murray, W.W. and Rachubinski, R.A. (1987b) The nucleotide sequence of complementary DNA and the deduced amino acid sequence of peroxisomal catalase of the yeast *Candida tropicalis* pK233. Gene 61, 401-413.

Murray, W.W. and Rachubinski, R.A (1989) Nucleotide sequence of peroxisomal catalase from the yeast *Candida tropicais* pK223: Identification of an upstream *Bam*HI site polymorphism. Nucl. Acids Res. 17, 3600.

Neupert, W., Hartl, F.-U., Craig, E. and Pfanner, M. (1990) How do polypeptides cross the mitochondrial membranes? Cell 63, 447-450.

Newmeyer, D.D. and Forbes, D.J. (1988) Nuclear import can be separated into distinct steps *in vitro*: nuclear pore binding and translocation. Cell 52, 641-653.

Newport, J.W. and Forbes, D.J. (1987) The nucleus: structure, function, and dynamics. Ann. Rev. Biochem. 56, 535-565.

Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H. and Verma, D.P.S. (1985) Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. Proc. Natl. Acad. Sci. USA 82, 5040-5044.

Nguyen, M., Bell, A.W. and Shore, G.C. (1988) Protein sorting between mitochondrial membranes specified by position of the stop-transfer domain. J. Cell Biol. 106, 1499-1505.

Nicolay, K., Veenhuis, M., Douma, A.C. and Harder, W. (1987) A ^{31}P NMR study of the internal pH of yeast peroxisomes. Arch. Microbiol. 147, 37-41.

Nilsson, T., Jackson, M.R. and Peterson, P.A. (1989) Short cytoplasmic sequences serve as retention signals for transmembrane proteins on the endoplasmic reticulum. Cell 58, 707-718.

Nishide, T., Nakamura, Y., Emi, M., Yamamoto, T., Ogawa, M., Mori, T. and Matsubara, K. (1986) Primary structure of human salivary alpha-amylase gene. Gene 41, 299-304.

Nothwehr, S.F., Roberts, C.J. and Stevens, T.H. (1991) Evidence that membrane proteins traversing the yeast secretory pathway are delivered to the vacuole by default. Yeast Cell Biol., Cold Spring Harbor, 135.

Novikoff, A.B. and Novikoff, P.M. (1982) Microperoxisomes and peroxisomes in relation to lipid metabolism. Ann. N.Y. Acad. Sci. 386, 138-152.

Novikoff, P.M., Novikoff, A.B., Quintana, N. and Davis, C. (1973) Studies on microperoxisomes. III Observations on human and rat hepatocytes. J. Histochem. Cytochem. 21, 540-558.

Nuttley, W.M., Aitchison, J.D. and Rachubinski, R.A. (1988) cDNA cloning and primary structure determination of the peroxisomal trifunctional enzyme hydratase-dehydrogenase-epimerase from the yeast *Candida tropicalis* pK233. Gene 69, 171-180.

Nuttley, W.M., Bodnar, A.G., Mangroo, D. and Rachubinski, R.A. (1990) Isolation and characterization of membranes from oleic acid-induced peroxisomes of *Candida tropicalis*. J. Cell Sci. 95, 463-470.

Nye, S.H. and Scarpulla, R.C. (1990) *In vivo* expression and mitochondrial targeting of yeast apoiso-1-cytochrome c fusion proteins. Mol. Cell. Biol. 10, 5753-5762.

Nyunoya, H., Broglie, K.E., Widgren, E.E. and Lusty, C.J. (1985) Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat. J. Biol. Chem. 260, 9346-9356.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchant polymorpha* chlorplast DNA. Nature 322, 572-574.

Okada, H., Ueda, M., Sugaya, T., Atomi, H., Mozaffar, S., Hishida, T., Teranishi, Y., Okazaki, K., Takechi, T., Kamiryo, T. and Tanaka, A. (1987) Catalase gene of the yeast *Candida tropicalis*. Eur. J. Biochem. 170, 105-110.

Okazaki, T., Takechi, T., Kambara, N., Fukui, S., Kubota, I. and Kamiryo, T. (1986) 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, 1232-1236.

Okazaki, K., Tan, H., Fukui, S., Kubota, I. and Kamiryo, T. (1987) Peroxisomal acyl-coenzyme A oxidase multigene family of the yeast *Candida tropicalis*; nucleotide sequence of a third gene and its protein product. Gene 58, 37-44.

Opperdoes, F.R. (1988) Glycosomes may provide clues to the import of peroxisomal proteins. Trends Biochem. Sci. 13, 255-260.

Opperdoes, F.R., Borst, P., Bakker, S. and Keene, W. (1977) Localization of glycerol-3-phosphate oxidase in the mitochondrion and particulate NAD^+-linked glycerol-3-phosphate dehydrogenase in the microbodies of the bloodstream form of *Trypanosoma brucei*. Eur. J. Biochem. 76, 29-39.

Orr, E.C., Bewley, G.C. and Orr, W.C. (1990) cDNA and deduced amino acid sequence of *Drosophila* catalase. Nucl. Acids Res. 18, 3663.

Osinga, K.A., Swinkels, B.W., Gibson, W.C., Borst, P., Veeneman, G.H., Van Boom, J.H., Michels, P.A.M. and Opperdoes, F.R. (1985) Topogenesis of microbody enzymes: a sequence comparison of the genes for the glycosomal (microbody) and cytosolic phosphoglycerate kinases of *Trypanosoma brucei*. EMBO J. 4, 3811-3817.

Osumi, M., Miwa, N., Teranishi, Y., Tanaka, A and Fukui, S. (1974) Ultrastructure of *Candida* yeasts grown on *n*-alkanes. Arch. Microbiol. 99, 181-201.

Osumi, T. and Hashimoto, T. (1979) Peroxisomal ß-oxidation system of rat liver. Copurification of enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase. Biochem. Biophys. Res. Commun. 89, 580-584.

Osumi, T. and Hashimoto, T. (1980) Purification and properties of acyl-CoA oxidase from rat liver. J. Biochem. 87, 1735-1746.

Osumi, T., Fukuzumi, F., Teranishi, Y., Tanaka, A. and Fukui, S. (1975) Development of microbodies in *Candida tropicalis* during incubation in a n-alkane medium. Arch. Microbiol. 103, 1-11.

Osumi, T., Ishii, N., Hijikata, M., Kamijo, K., Ozasa, H., Furata, S., Miyazawa, S., Kondo, K., Inoue, K., Kagamiyama, H. and Hashimoto, T. (1985). Molecular cloning and nucleotide sequence of the cDNA for rat peroxisomal enoyl-CoA: hydratase-3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme. J. Biol. Chem. 260, 8905-8910.

Osumi, T., Ishii, N., Miyazawa, S. and Hashimoto, T. (1987) Isolation and structural characterization of the rat acyl-CoA oxidase gene. J. Biol. Chem. 262, 8138-8143.

Pain, D., Kanwar, Y.S. and Blobel, G. (1988) Identification of a receptor for protein import into chloroplasts and its localization to envelope contact zones. Nature 331, 232-237.

Pain, D., Murikami, H. and Blobel, G. (1990) Identification of a receptor for protein import into mitochondria. Nature 347, 444-449.

Palosaari, P.M. and Hiltunen, J.K. (1990) Peroxisomal bifunctional protein from rat liver is a trifunctional enzyme possessing 2-enoyl-CoA-hydratase, 2-hydroxyacyl-CoA dehydrogenase and Δ^3, Δ^2-enoyl-CoA isomerase activities. J. Biol. Chem 265, 2446-2449.

Parker, R., Caponigro, G., Heaton, B., Muhlrad, P. and Decker, C. (1991) mRNA turnover in yeast. J. Cell Biol. 115, 247a.

Parnes, J.R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E. and Seidman, J.G. (1981) Mouse β_2-microglobulin cDNA clones: A screening procedure for cDNA clones corresponding to rare mRNAs. Proc. Natl. Acad. Sci. USA. 78, 2253-2257.

Pasero, L., Mazzei-Pierron, Y., Abadie, B., Chicheportiche, Y. and Marchis-Mouren, G. (1986) Complete amino acid sequence and location of the five disulfide bridges in porcine pancreatic alpha-amylase. Biochim. Biophys. Acta 869, 147-157.

Pedersen, J.I. and Gustafsson, J. (1980) Conversion of 3α, 7α,12α-trihydroxy-5β-cholestanoic acid into cholic acid by rat liver peroxisomes. FEBS Lett. 121, 345-348.

Pelham, H.R.B. (1988a) Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. EMBO J. 7, 913-918.

Pelham, H. (1988b) Heat shock proteins. Coming in from the cold. Nature 332, 776-777.

Pelham, H.R.B. (1991) Recycling of proteins between the endoplasmic reticulum and Golgi complex. Curr. Opin. Cell Biol. 3, 585-591.

Pelham, H.R.B. and Jackson, R.J. (1976) An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. 67, 247-256.

Pfaller, R. and Neupert, W. (1987) High-affinity binding sites involved in the import of porin into mitochondria. EMBO J. 6, 2635-2642.

Pfaller, R., Steger, H.F., Rassow, J., Pfanner, N. and Neupert, W. (1988) Import pathways of precursor proteins into mitochondria: multiple receptor sites are followed by a common membrane insertion site. J. Cell Biol. 107, 2483-2490.

Pfanner, N. and Neupert, W. (1990) The mitochondrial protein import apparatus. Ann. Rev. Biochem. 59, 331-353.

Pfanner, N., Hartl, F.-U., Guiard, B. and Neupert, W. (1987) Mitochondrial precursor proteins are imported through a hydrophilic membrane environment. Eur. J. Biochem. 169, 289-293.

Pfanner, N., Tropschug, M. and Neupert, W. (1987) Mitochondrial protein import: nucleoside triphosphates are involved in conferring import competence to precursors. Cell 49, 815-823.

Pfanner, N., Söllner, T. and Neupert, W. (1991) Mitochondrial import receptors for precursor proteins. Trends Biochem. Sci 16, 63-67.

Pfeffer, S.R. and Rothman, J.E (1987) Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. Ann. Rev. Biochem. 56, 829-852.

Picataggio, S., Deanda, K. and Millenz, J. (1991) Determination of *Candida tropicalis* acyl-Coenzyme A oxidase isozyme function by sequential gene disruption. Mol. Cell. Biol. 11, 4333-4339.

Pringle, J.R., Adams, A.E.M., Drubin, D.G., and Haarer, B.K. (1991) Immunuofluorescence methods for yeast. Methods Enzymol. **194**, 565-602.

Pugsly, A.P. (1989) Protein Targeting, Academic press, Inc. San Diego, California.

Quan, F., Korneluk, R.G., Tropak, M.B. and Gravel, R.A. (1986) Isolation and characterization of the human catalase gene. Nucl. Acids Red. 14, 5321-5335.

Queen, C. and Korn, L.J. (1984) A comprehensive sequence analysis program for the IBM personal computer. Nucl. Acids Res. 12, 581-599.

Rachubinski, R.A., Fujiki, Y., Mortensen, R.M. and Lazarow, P.B. (1984) Acyl-CoA oxidase and hydratase-dehydrogenase, two enzymes of the peroxisomal β-oxidation system, are synthesized on free polysomes of clofibrate-treated rat liver. J. Cell Biol. 99, 2241-2246.

Rachubinski, R.A., Fujiki, Y. and Lazarow, P.B. (1985) Cloning of cDNA coding for peroxisomal acyl-CoA oxidase from the yeast *Candida tropicalis* pK233. Proc. Natl. Acad. Sci. USA 82, 3973-3977.

Rachubinski, R.A., Fujuki, Y. and Lazarow, P.B. (1987) Isolation of cDNA clones coding for peroxisomal proteins of *Candida tropicalis*: identification and sequence of a clone for catalase. Biochim. Biophys. Acta 909, 35-43.

Radloff, R., Bauer, W. and Vinograd, J. (1967) A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: The closed circular DNA in HeLa cells. Proc. Natl. Acad. Sci. USA 57, 1514-1521.

Rapoport, T. (1990) Protein translocation across the ER membrane. Trends Biochem. Sci. 15, 355-358.

Reddy, J. and Svoboda, D. (1971) Microbodies in experimentally altered cells VIII. Continuities between microbodies and their possible biologic significance. Lab. Invest. 24, 74-81.

Redinbaugh, M.G., Wadsworth, G.J., and Scandalois, J.G. (1988) Characterization of catalase transcripts and their differential expression in maize. Biochem. Biophys. Acta 951, 104-116.

Richardson, W.D., Roberts, B.L. and Smith, A.E. (1986) Nuclear location signals in polyoma virus large T. Cell 44, 77-85.

Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A. and Dingwall, C. (1988) Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translation through nuclear pores. Cell 52, 655-664.

Riezman, H., Weir, E.M., Leaver, C.J., Titus, D.E. and Becker, W.M. (1980) Regulation of glyoxysomal enzymes during germination of cucumber. 3. *In vitro* translation and characterization of four glyoxysomal enzymes. Plant Physiol. 65, 40-46.

Riezman, H., Hase, T., van Loon, A.P., Grivell, L.A., Suda, K. and Schatz, G. (1983) Import of proteins into mitochondria: a 70 kilodalton outer membrane protein with a large carboxy-terminal deletion is still transported to the outer membrane. EMBO J. 2, 2161-2168.

Rigatuso, J.L., Legg, P.G. and Wood, R.L. (1970) Microbody formation in regenerating rat liver. J. Histochem. Cytochem. 18, 893-900.

Roa, M. and Blobel, G. (1983) Biosynthesis of peroxisomal enzymes in the methylotrophic yeast *Hansenula polymorpha*. Proc. Natl. Acad. Sci. USA 80, 6832-6876.

Robbi, M. and Lazarow, P.B. (1978) Synthesis of catalase in two cell-free protein-synthesizing systems and in rat liver. Proc. Natl. Acad. Sci. USA 75, 4344-4348.

Robbi, M. and Lazarow, P.B. (1982) Peptide mapping of peroxisomal catalase and its precursor. J. Biol. Chem. 257, 964-970.

Robbi, M. and Beaufay, H. (1991) The COOH terminus of several liver carboxyl esterases targets these enzymes to the lumen of the endoplasmic reticulum. J. Biol. Chem. 266, 20498-20503.

Roberts, B. (1989) Nuclear location signal-mediated protein transport. Biochim. Biophys. Acta 1008, 253-280.

Roberts, L.M. and Lord, J.M. (1981) Synthesis and post-translational segregation of glyoxysomal isocitrate lyase from castor bean endosperm. Eur. J. Biochem. 119, 43-49.

Robinson, C. and Ellis, R.J. (1984) Transport of proteins into chloroplasts. Partial purification of a chloroplast protease involved in the processing of imported precursor polypeptides. Eur. J. Biochem.142, 337-342.

Roggenkamp, R., Janowicz, Z., Stanikowski, B. and Hollenberg, C.P. (1984) Biosynthesis and regulation of the peroxisomal methanol oxidase from the methylotrophic yeast *Hansenula polymorpha*. Mol. Gen. Genet. 194, 489-493.

Roggenkamp, R., Didion, T. and Kowallik, K.V. (1989) Formation of irregular grant peroxisomes by overproduction of the crystalloid core protein methanol oxidase in the methylotrphic yeast *Hansenula polymorpha*. Mol. Cell. Biol. 9, 988-994.

Ronchi, S., Minchiotti, L., Galliano, M., Curti, B., Swenson, R.P., Williams, C.H. Jr. and Massey, V. (1982) The primary structure of D-amino acid oxidase from pig kidney. J. Biol. Chem. 257, 8824-8834.

Rothman, J.E. (1989) Polypeptide chain binding proteins: catalysts of protein folding and related processes in cells. Cell 59, 591-601.

Rothman, J.E. and Orci, L. (1992) Molecular dissection of the secretory pathway. Nature 355, 409-415.

Rothman, J.H., Yamashiro, C.T., Kane, P.M. and Stevens, T.H (1989) Protein targeting to the yeast vacuole. Trends Biochem. Sci. 14, 347-350.

Ryall, J., Rachubinski, R.A., Nguyen, M., Rozen, R., Broglie, K. and Shore, G.C. 1984) Regulation and exp ..tion of carbamylphosphate synthetase I mRNA in developing rat liver and Morris hepatoma 5123D. J. Biol. Chem. 259, 9172-9176.

Sakajo, S., Nakamura, K. and Asahi,T. (1987) Molecular cloning and nucleotide sequence of full-length cDNA for sweet potato catalase mRNA. Eur. J. Biochem. 165, 437-442.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Sanglord, D. and Loper, J.C. (1989) Characterization of the alkane-inducible cytochrome P450 ($P450_{alk}$) gene from the yeast *Candida tropicalis*: identification of a new *P450* family. Gene 76, 121-136.

Santos, M.J., Ojeda, J.M., Garrido, J. and Leighton, F. (1985) Peroxisomal organization in normal and cerebrohepatorenal (Zellweger) syndrome fibroblasts. Proc. Natl. Acad. Sci. USA 82, 6556-6560.

Santos, M.J., Imanaka, T., Shio, H., Small, G.M. and Lazarow, P.B. (1988a) Peroxisomal membrane ghosts in Zellweger syndrome-aberrant organelle assembly. Science 239, 1536-1538.

Santos, M.J., Imanaka, T., Shio, H. and Lazarow, P.B. (1988b) Peroxisomal integral membrane proteins in control and Zellweger fibroblasts. J. Biol. Chem. 263, 10502-10509.

Sarokin, L. and Carlson, M. (1986) Short repeated elements in the upstream regulatory region of *SUC2* gene of *Saccharomyces cerevisiae*. Mol. Cell Biol. 6, 2324-2333.

Schepers, L., Van Veldhoven, P.P., Casteels, M., Eyssen, H.J. and Mannaerts, G.P. (1990) Presence of three acyl-CoA oxidases in rat liver peroxisomes. J. Biol. Chem. 265, 5242-5246.

Schmidt, G.W. and Mishkind, M.L. (1986) The transport of proteins into chloroplasts. Ann. Rev. Biochem. 55, 879-912.

Schnell, D.J., Blobel, G. and Pain, D. (1990) The chloroplast import receptor is an integral membrane protein of chloroplast envelope contact sites. J. Cell Biol. 111, 1825-1838.

Schrakamp, G., Roosenboom, C.F.P., Schotgens, R.B.H., Wanders, R.J.A,, Heymans, H.S.A., Tager, J.M. and Van Den Bosch, H. (1985) Alkyl dihydroxyacetone phosphate synthase in human fibroblasts and its deficiency in Zellweger syndrome. J. Lipid Res. 26, 867-873.

Schram, A.W., Strijland, A., Hashimoto, T., Wanders, R.J.A., Schutgens, R.B.H., van den Bosch, H. and Tager, J.M. (1986) Biosynthesis and maturation of peroxisomal β-oxidation enzymes in fibroblasts in relation to the Zellweger syndrome and infantile Refsum disease. Proc. Natl. Acad. Sci. USA 83, 6156-6158.

Schroeder, W.A., Shelton, J.R., Shelton, J.B., Robertson, B., Apell, G., Fang, R.S. and Bonaventura, J. (1982) The complete amino acid sequence of bovine liver catalase and the partial sequence of bovine erythrocyte catalase. Arch. Biochem. Biophys. 214, 397-421.

Schutgens, R.B.H., Romeyn, G.J., Wanders, R.J.A., Van Den Bosch, H., Schrakamp, G. and Heymans, H.S.A. (1984) Deficiency of acyl-CoA: dihydroxyacetone phosphate acyl transferase in patients with Zellweger (cerebro-hepato-renal) syndrome. Biochem. Biophys. Res. Commun. 12, 179-184.

Semenza, J.C., Hardwick, K.G., Dean, N. and Pelham, H.R.B. (1990) *ERD2*, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. Cell 61, 1349-1357.

Sharp, P.M., Tuohy, T.M.F. and Mosurski, K.R. (1986) Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucl. Acids Res. 14, 5125-5143.

Sherman, F., Fink, G.R. and Hicks, J.B. (1986) In: Laboratory Course Manual for Methods in Yeast Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Shimozawa, N., Tsukamoto, T., Suzuki, Y., Orii, T., Shirayoshi, Y., Mori, T. and Fujiki, Y. (1992) A human gene responsible for Zellweger syndrome that affects peroxisome assembly. Science 255, 1132-1134.

Shio, H. and Lazarow, P.B. (1981) Relationships between peroxisomes and endoplasmic reticulum investigated by combined catalase and glucose-6-phosphatase cytochemistry. J. Histochem. Cytochem. 29, 1263-1272.

Silver, P.A. (1991) How proteins enter the nucleus. Cell 64, 489-497.

Silver, P., Sadler, I. and Osborne, M.A. (1989) Yeast proteins that recognize nuclear localization sequences. J. Cell Biol. 109, 983-989.

Simon, M. and Faye, G. (1984) Organization and processing of the mitochondrial oxi3/oli2 multigenic transcript in yeast. Mol. Gen. Genet. 196, 266-274.

Simon, S.M. and Blobel, G. (1991) A protein-conducting channel in the endoplasmic reticulum. Cell 65, 371-380.

Simon, M., Adam, G., Rapatz, W., Spevak, W. and Ruis, H. (1991) The *Saccharomyces cerevisiae APR1* gene is a positive regulator of transcription of genes encoding peroxisomal proteins. Mol. Cell. Biol. 11, 691-704.

Singer, S.J. and Yaffe, M.P. (1990) Embedded or not? hydrophobic sequences and membranes. Trends Biochem. Sci. 15, 369-373.

Singh, I., Moser, A.E., Goldfischer, S. and Moser, H.W. (1984) Lignoceric acid is oxidized in the peroxisome: Implications for the Zellweger cerebro-hepato-renal syndrome and adrenoleukodystrophy. Proc. Natl. Acad. Sci. USA 81, 4203-4207.

Sloots, J.A., Aitchison, J.D. and Rachubinski, R.A. (1991) Glucose-responsive and oleic acid-responsive elements in the gene encoding the peroxisomal trifunctional enzyme of *Candida tropicalis*. Gene 105, 129-134.

Small, G.M. and Lazarow, P.B. (1987) Import of the carboxy-terminal portion of acyl-CoA oxidase into peroxisomes of *Candida tropicalis*. J. Cell Biol. 105, 247-250.

Small, G.M., Imanaka, T., Shio, H. and Lazarow, P.B. (1987) Efficient association of *in vitro* translation products with purified, stable *Candida tropicalis* peroxisomes. Mol. Cell. Biol. 7, 1848-1855.

Small, G.M., Szabo, L.J. and Lazarow, P.B. (1988) Acyl-CoA oxidase contains two targeting sequences each of which can mediate protein import into peroxisomes. EMBO J. 7, 1167-1173.

Smeland, T.E., Chin-hung, C., Cuebas, D. and Schulz, H. (1989) The 3-hydroxyacyl-CoA epimerase activity of rat liver peroxisomes is due to the combined actions of two enoyl-CoA hydratases: a revision of the epimerase-dependent pathway of unsaturated fatty acid oxidation. Biochem. Biophys. Res. Commun. 160, 988-992.

Smith, S.M. and Leaver, C.J. (1986) Glyoxysomal malate synthase of cucumber. Molecular cloning of a cDNA and regulation of enzyme synthesis during germination. Plant Physiol. 81, 762-767.

Söllner, T., Griffiths, G., Pfaller, R., Pfanner, N. and Neupert, W. (1989) MOM19, an import receptor for mitochondrial precursor proteins. Cell 59, 1061-1070.

Söllner, T., Pfaller, R., Griffiths, G., Pfanner, N. and Neupert, W. (1990) A mitochondrial import receptor for the ADP-ATP carrier. Cell 62, 107-115.

Sorensen, S.B., Breddam, K. and Svendsen, I. (1986) Primary structure of carboxypeptidase I from malted barley. Carlsberg Res. Commun. 51, 475-485.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Stevens, T.H., Esmon, B. and Schekman, R. (1982) Early stages in the yeast secretory pathway are required for transport of carboxypeptidase Y to the vacuole. Cell 30, 439-448.

Stryer, L. (1988) In: Biochemistry, 3rd Ed. W.H. Freeman and Co. NY, p. 475.

Subramani, S. (1991) Peroxisomal targeting signals - the end and the beginning. Curr. Sci. 61, 28-32.

Subramani, S. (1992) Targeting of proteins into the peroxisomal matrix. J. Mem. Biol. 125, 99-106.

Suzuki, T., Sato, M., Yoshida, T. and Tuboi, S. (1989) Rat liver mitochondrial and cytosolic fumerases with identical amino acid sequences are encoded from a single gene. J. Biol. Chem. 264, 2581-2586.

Suzuki, Y., Orii, T., Mori, M., Tatibana, M. and Hashimoto, T. (1986) Deficient activities and proteins of peroxisomal ß-oxidation enzymes in infants with Zellweger syndrome. Clin. Chim. Acta 156, 191-196.

Suzuki, Y., Orii, T., Takiguchi, M., Mori, M., Hijikata, M. and Hashimoto, T. (1987) Biosynthesis of membrane polypeptides of rat liver peroxisomes. J. Biochem. 101, 491-496.

Svoboda, D.J. and Azarnoff, D.L. (1966) Response of hepatic microbodies to a hypolipidemic agent, ethylchlorophenoxyisobutyrate (CPIB). J. Cell Biol. 30, 443-450.

Sweet, D.J. and Pelham, H.R.B. (1991) The role of the *S. cerevisiae* SEC20 gene in secretion. Yeast Cell Biology, Cold Spring Harbor, p. 53.

Swift, A.M. and Machamer, C.T. (1991) A Golgi retention signal in a membrane-spanning domain of coronavirus E1 protein. J. Cell Biol. 115, 19-30.

Swinkels, B.W., Evers, R. and Borst, P. (1988) The topogenic signal of the glycosomal (microbody) phosphoglycerate kinase of *Crithidia fasciculata* resides in a carboxy-terminal extension. EMBO J. 7, 1159-1165.

Swinkels, B.W. and Subramani, S. (1990) Targeting efficiencies of all possible permutations of the consensus tripeptide peroxisomal targeting signal. J. Cell Biol. 111, 195a.

Swinkels, B.W., Gould, S.J., Bodnar, A.G., Rachubinski, R.A. and Subramani, S. (1991) A novel cleavable peroxisomal targeting signal at the amino terminus of the rat 3-ketoacyl-CoA thiolase. EMBO J. 10, 3255-3262.

Szabo, L.J., Small, G.M. and Lazarow, P.B. (1989) The nucleotide sequence of *POX 18*, a gene encoding a small oleate-inducible peroxisomal protein from *Candida tropicalis*. Gene 75, 119-126.

Tabor, S. and Richardson, C.C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA 84, 4767-4771.

Tada, M., Fukui, K., Momoi, K. and Miyake, Y. (1990) Cloning and expression of a cDNA encoding mouse kidney D-amino acid oxidase. Gene 90, 293-297.

Tager, J.M., Ven Der Beek, W.A.T.H., Wanders, R.J.A., Hashimoto, T., Heymans, H.S.A., Van Den Bosch H., Schutgens, R.B.H. and Schram, A.W. (1985) Peroxisomal ß-oxidation enzyme proteins in the Zellweger syndrome. Biochem. Biophys. Res. Commun. 126, 1269-1275,

Tanabe, I., Okada, J. and Ono, H. (1966) Isolation and determination of yeasts utilizing kerosene as a sole source of carbon. Agr. Biol. Chem. 30, 1175-1182.

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

Taylor, J.W., Ott, J., and Eckstein, F. (1985) The rapid generation of oligonucleotide-directed mutations at a high frequency using phosphorthioate-modified DNA. Nucleic Acids Res. 13, 8765-8785

Thieringer, R., Shio, H., Han, Y., Cohen, G. and Lazarow, P.B. (1991) Peroxisomes in *Saccharomyces cerevisiae*: Immunofluorescence analysis and import of catalase A into isolated proxisomes. Mol. Cell. Biol. 11, 510-522.

Thrift, R.N., Andrews, D.W., Walter, P. and Johnson, A.E. (1991) A nascent membrane protein is located adjacent to ER membrane proteins throughout its integration and translation. J. Cell Biol. 112, 809-821.

Tolbert, N.E. (1981) Metabolic pathways in peroxisomes and glyoxysomes. Ann. Rev. Biochem. 50, 133-157.

Tomaszewski, K.E., Agarwal, D.K. and Melnick, R.L. (1986) *In vitro* steady-state levels of hydrogen peroxide after exposure of male F344 rats and female B6C3F$_1$ mice to hepatic peroxisome proliferators. Carcinogenesis 7, 1871-1876.

Triglia, T., Stahl, H.D., Crewther, P.E., Silva, A., Anders, R.F. and Kemp, D.J. (1988) Structure of a plasmodium falciparum gene that encodes a glutamic acid-rich protein (GARP). Mol. Biochem. Parasitol. 31, 199-201.

Tsukada, H., Mochizuki, Y. and Konishi, T. (1968) Morphogenesis and development of hepatocytes of rats during pre- and postnatal growth. J. Cell Biol. 37, 231-243.

Tsukamoto, T., Miura, S. and Fujiki, Y. (1991) Restoration by a 35K membrane protein of peroxisome assembly in a peroxisome-deficient mammalian cell mutant. Nature 350, 77-81.

Turley, R.B., Choe, S.M. and Trelease, R.N. (1990) Characterization of a cDNA clone encoding the complete amino acid sequence of cotton isocitrate lyase. Biochim. Biophys. Acta 1049, 223-226.

Ueda, M., Tanaka, A., Horikawa, S., Numa, S. and Fukui, S. (1984) Synthesis *in vitro* of precursor-type carnitine acetyltransferase with messenger RNA from *Candida tropicalis*. Eur. J. Biochem. 138, 451-457.

Ueda, M., Morikawa, T., Okada, H. and Tanaka, A. (1987) Relationship between enoyl-CoA hydratase and a peroxisomal bifunctional enzyme, enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase, from an *n*-alkane-utilizing yeast, *Candida tropicalis*. Agric. Biol. Chem. 51, 2197-2205.

Valls, L.A., Winther, J.R. and Stevens, T.H. (1990) Yeast carboxypeptidase Y vacuolar targeting signal is defined by four propeptide amino acids. J. Cell Biol. 111, 361-368.

van den Broeck, G., Timko, M.P., Kausch, A.P., Cashmore, A.R., van Montagu, M. and Herrera-Estrella, L. (1985) Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,5-bisphosphate carboxylase. Nature 313, 358-363.

van Loon, A.P., van Eijk, E. and Grivell, L.A. (1983) 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, 1765-1770.

Vaux, G., Tooze, J. and Fuller, S. (1990) Identification by anti-idiotype antibodies of an intracellular membrane protein that recognizes a mammalian endoplasmic reticulum retention signal. Nature 345, 495-502.

Veenhuis, M. and Goodman, J.M. (1990) Peroxisomal assembly: Membrane proliferation preceeds the induction of the abundant matrix proteins in the methylotrophic yeast *Candida boidinii*. J. Cell Sci. 96, 583-590.

Veenhuis, M., van Dijken, J.P., Pilon, S.A.F. and Harder, W. (1978) Development of crystalline peroxisomes in methanol-grown cells of the yeast *Hansenula polymorpha* and its relation to environmental conditions. Arch. Microbiol. 117, 153-163.

Veenhuis, M., Mateblowski, M., Kunau, W.-H. and Harder, W. (1987) Proliferation of microbodies in *Saccharomyces cerevisiae*. Yeast 3, 77-84.

Verner, K. and Schatz, G. (1987) Import of incompletely folded precursor protein into isolated mitochondria requires an energized inner membrane, but no added ATP. EMBO J. 6, 2449-2456.

Verner, K. and Schatz, G. (1988) Protein translocation across membranes. Science 241, 1307-1313.

Vestweber, D., Brunner, J., Baker, A. and Schatz, G. (1989) A 42 K outer membrane protein is a component of the yeast mitochondrial protein import site. Nature 341, 205-209.

Vieira, J. and Messing, J. (1987) Production of single-stranded plasmid DNA. Meth. Enzymol. 153 3-11.

Volokita, M. and Sommerville, C.R. (1987) The primary structure of spinach glycolate oxidase deduced from the DNA sequence of a cDNA clone. J. Biol. Chem. 262, 15825-15828.

von Heijne, G. (1986) Mitochondrial targeting sequences may form amphiphilic helices. EMBO J. 5, 1335-1342,

von Heijne, G. (1988) Transcending the impenetrable: how proteins come to terms with membranes. Biochim. Biophys. Acta 947, 307-333.

von Heijne, G. (1990) Protein targeting signals. Curr. Opin. Cell Biol. 2, 604-608.

von Heijne, G. and Nishikawa, K. (1991) Chloroplast transit peptides. The perfect random coil? FEBS Lett. 278, 1-3.

von Heijne, G., Steppuhn, J. and Herrmann, R.G. (1989) Domain structure of mitochondrial and chloroplast targeting peptides. Eur. J. Biochem. 180, 535-545.

Waegemann, K., Paulsen, H. and Soll, J. (1990) Translocation of proteins into isolated chloroplasts requires cytosolic factors to obtain competence. FEBS Lett. 261, 89-92.

Walk, R.-A. and Hock, B. (1978) Cell-free synthesis of glyoxysomal malate dehydrogenase. Biochem. Biophys. Res. Commun. 81, 636-643.

Wallrath, L.L., Burnett, J.B. and Friedman, T.B. (1990) Molecular characterization of the *Drosophila melanogaster* urate oxidase gene, an ecdysone-repressible gene expressed only in maloighian tubes. Mol. Cell. Biol. 10, 5114-5127.

Walter, P. and Blobel, G. (1980) Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. Proc. Natl. Acad. Sci. USA 77, 7112-7116.

Walter, P. and Blobel, G. (1981) Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. J. Cell Biol. 91, 557-561.

Walter, P. and Blobel, G. (1982) Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature 299, 691-698.

Walter, P., Jackson, R.C., Marcus, M.M., Lingappa, V.R. and Blobel, G. (1979) Tryptic dissection and reconstitution of translocation activity for nascent presecretory proteins across microsomal membranes. Proc. Natl. Acad. Sci. USA 76, 1795-1799.

Walter, P., Ibrahimi, I. and Blobel, G. (1981) Translocation of proteins across the endoplasmic reticulum I. Signal recognition particle (SRP) binds to *in vitro*-assembled polysomes synthesizing secretory protein. J. Cell Biol. 91, 545-550.

Walter, T., Gilmore, R. and Blobel, G. (1984) Protein translocation across the endoplasmic reticulum. Cell 38, 5-8.

Walton, P.A., Gould, S.J., Feramisco, J.R. and Subramani, S. (1992) Transport of microinjected proteins into peroxisomes of mammalian cells: Inability of Zellweger cll lines to import proteins with the SKL tripeptide targeting signal. Mol. Cell. Biol. 12, 531-541.

Wanders, R.J.A., Kos, M., Roest, B., Merjer, A.J., Schrakamp, G., Heymans, H.S.A., Tegelaers, W.H.H., Van Den Bosch, H., Schutgens, R.B.H. and Tager, J.M. (1984) Activity of peroxisomal enzymes and intracellular distribution of catalase in Zellweger syndrome. Biochem. Biophys. Res. Commun. 123, 1054-1061.

Warren, G. (1990) Salvage receptors: Two of a kind? Cell 62, 1-2.

Wedel, F.P. and Berger, E.R. (1975) On the quantitative stereo-morphology of microbodies in rat hepatocytes. J. Ultrastruct. Res. 51, 153-165.

Wickens, M.P., Buell, G.N. and Schimke, R.T. (1978) Synthesis of double-stranded DNA complementary to lysozyme, ovomucoid, and ovalbumin mRNAs. J. Biol. Chem. 253, 2483-2495.

Wiedmann, M., Kurzchalia, T.V., Hartmann, E. and Rapoport, T.A. (1987) A signal sequence receptor in the endoplasmic reticulum membrane. Nature 328, 830-833.

Wierenga, R.K., Swinkels, B., Michels, P.A.M., Osinga, K., Misset, O., Van Beeuman, J., Gibson, W.C., Postma, J.P.M., Borst, P., Opperdoes, F.R. and Hol, W.G.J. (1987) Common elements on the surface of glycolytic enzymes from *Trypanosoma brucei* may serve as topogenic signals for import into glycosomes. EMBO J. 6, 215-221.

Wood, K.V., Lam, Y.A., Seliger, H.H. and McElroy, W.D. (1989) Complementary DNA coding click beetle luciferases can elicit broluminscence in different colors. Science 244, 700-702.

Yamada., T., Tanaka, A., Horikawa, S., Numa, S. and Fukui, S. (1982) Cell-free translation and regulation of *Candida tropicalis* catalase messenger RNA. Eur. J. Biochem. 129, 251-255.

Yamaguchi, J., Nishimura, M. and Akazawa, T. (1984) Maturation of catalase precursor proceeds to a different extent in glyoxysomes and leaf peroxisomes of pumpkin cotyledons. Proc. Natl. Acad. Sci USA 81, 4809-4813.

Yamamoto, K. and Fahimi, H.D. (1987) Three-dimensional reconstitution of a peroxisomal reticulum in regenerating rat liver: evidence of interconnections between heterogenous segments. J. Cell Biol. 105, 713-722.

Yoo, H.S. and Cooper, T.G. (1989) The *DAL7* promoter consists of multiple elements that cooperatively mediate regulation of the gene's expression. Mol. Cell. Biol. 9, 3231-3243.

Young, R.A. and Davis, R.W. (1983) Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. USA 80, 1194-1198.

Zaar, K., Völkl, A. and Fahimi, H.D. (1987) Association of isolated bovine kidney cortex peroxisomes with endoplasmic reticulum. Biochim. Biophys. Acta 897, 135-142.

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

Zimmermann, R. and Neupert, W. (1980) Biogenesis of glyoxysomes, synthesis and intracellular transfer of isocitrate lyase. Eur. J. Biochem. 112, 225-233.

Zoeller, R.A. and Raetz, C.R.H. (1986) Isolation of animal cell mutants deficient in plasmalogen biosynthesis and peroxisome assembly. Proc. Natl. Acad. Sci. USA 83, 5170-5174.

Appendix 1

Alignment of HDE repeat regions with similar sequences identified by FAST SCAN.

The FAST SCAN program of PC/GENE was used to scan the SWISS-PROT 20 protein data base (version 6.0, February 1992) for sequences similar to HDE. The three proteins with the most significant similarities were aligned with the repeat regions of HDE using the MULTIPLE SEQUENCE ALIGNMENT program. The sequences are as follows: HDEREP1, Amino acids 6 to 265 of HDE; HDEREP2, amino acids 319 to 588 of HDE; PHBB$ZOORA, acetoacetyl-CoA reductase from *Zooglea ramigera* (241 amino acids); DHG$BACSU, glucose dehydrogenase from *Bacillus subtilis* (260 amino acids); DHGB$BACME, glucose dehydrogenase b from *Bacillus megaterium* (262 amino acids). The similarity of the sequences shown below supports the inclusion of HDE into the superfamily of oxidoreductases (Baker, 1990) as discussed in section 3.2.7.2.

'*' indicates that the position in the alignment is perfectly conserved.

'.' indicates that the position in the alignment is well conserved.

Consensus length: 302

Identity: 35 (11.6%)

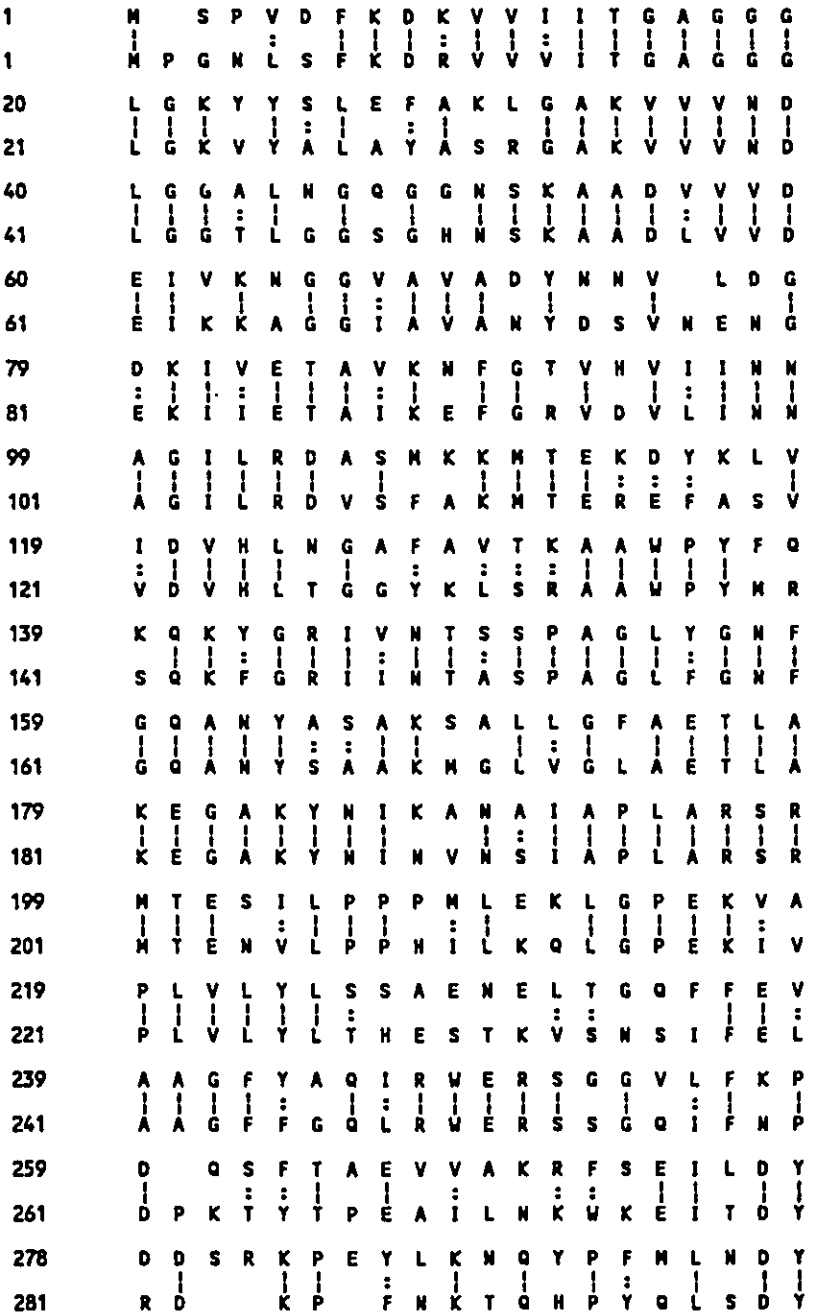
Similarity: 72 (23.8%)

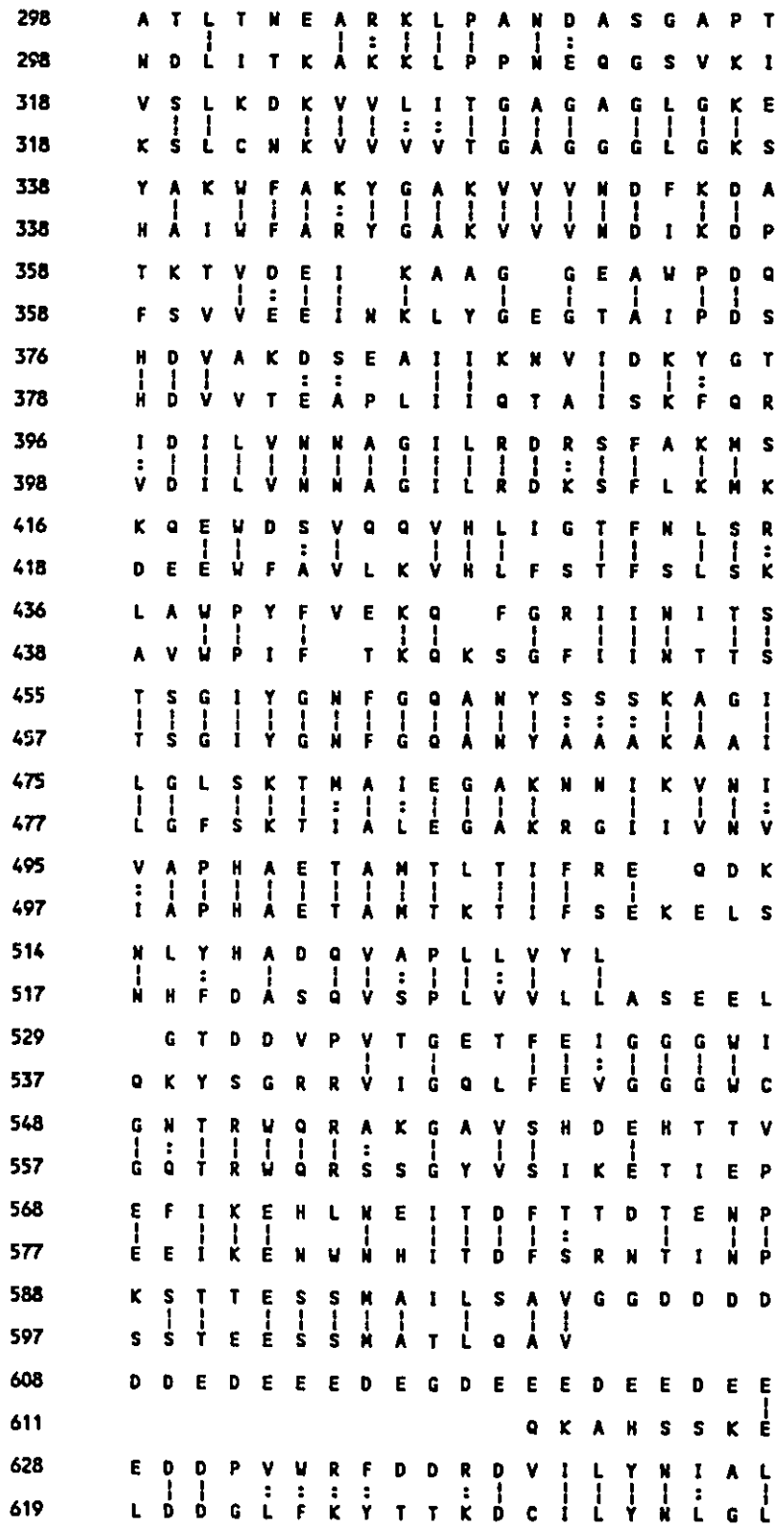
Alignment

```
HDEREP1      FKDKV---VVIITGAGGGLGKYYSLEFAKLGAKVVVNDLGGALNGQGGNS     47
HDEREP2      LKDKV---VLI-TGAGAGLGKEYAKWFAKYGAKVVVNDF-----------     35
PHBB$ZOORA   MSR-----VALVTGGSRGIGAAISIALKAAGYKV-----AASYAGNDDAA     40
DHG$BACSU    MYPDLKGKVVAITGAASGLGKAMAIRFGKEQAKVVIN-----YYSNKQDP     45
DHGB$BACME   MYKDLEGKVVVITGSSTGLGKSMAIRFATEKAKVVVN-----YRSKEDEA     45
             .       *   **.. *.*   . .     **

HDEREP1      KAADVVVDEIVKNGGVAVA-DYN-NVL-DGDKIVETAVKNFGTVHVIINN     94
HDEREP2      KDATKTVDEIKAAGGEAWP-DQH-DVAKDSEAIIKNVIDKYGTIDILVNN     83
PHBB$ZOORA   KP-------FKAETGIAVY-KWDVSSYEACVEGIAKVEADLGPIDVLVNN     82
DHG$BACSU    N--EV-KEEVIKAGGEAVVVQGDVTKEEDVKNIVQTAIKEFGTLDIMINN     92
DHGB$BACME   N--SVLEEEIKKVGGEAIAVKGDVTVESDVINLVQSAIKEFGKLDVMINN     93
             .            .* *   . . .   .    . ..  . * .....**

HDEREP1      AGILRDASMKKMTEKDYKLVIDVHLNGAFAVTKAAWPYFVKQKY-GRIVN    143
HDEREP2      AGILRDRSFAKMSKQEWDSVQQVHLIGTFNLSRLAWPYFVEKQF-GRIIN    132
PHBB$ZOORA   AGITKDAMFHKMTPDQWNAVINTNLTGLFNMTHPVWSGMRDRSF-GRIVN    131
DHG$BACSU    AGLENPVPSHEMPLKDWDKVIGTNLTGAFLGSREAIKYFVENDIKGNVIN    142
DHGB$BACME   AGMENPVSSHEMSLSDWNKVIDTNLTGAFLGSREAIKYFVENDIKGTVIN    143
             **. .     .*.  ... *  ..* * *  .. .   . ..   * ..*

HDEREP1      TSSPAGLYGNFGQ-ANYASAKSALLGFAETLAKEGAKYNIKANAIAP---    189
HDEREP2      ITSTSGIYGNFGQ-ANYSSSKAGILGLSKTMAIEGAKNNIKVNIVAP---    178
PHBB$ZOORA   ISSINGQKGQMGQ-ANYSAAKAGDLGFTKALAQEGAAKGITVNAICPGYI    180
DHG$BACSU    MSSVHA--FPWPLFVHYAASKGGIKLMTETLALEYAPKGIRVNNIGPGAI    190
DHGB$BACME   MSSVHEWKIPWPLFVHYAASKGGMKLMTETLALEYAPKGIRVNNIGPGAI    193
              .*  .         ..*...*..   .....* * *  .* .* . *

HDEREP1      --------LARSRMTESILPPPMLEKLG-PEKVAPLVLYLSSAENE-LTG    229
HDEREP2      --------HAETAMTLTIFREQDKNLYH-ADQVAPLLVYLGTDDVP-VTG    218
PHBB$ZOORA   GTEM-VRAIPEKVLNERIIPQIPVGRLGEPDEIARIVVFLASDEAGFITG    229
DHG$BACSU    NTPINAEKFADPKQKADVESMIPMGYIGEPEEIAAVAAWLASKEASYVTG    240
DHGB$BACME   NTPINAEKFADPEQRADVESMIPMGYIGEPEEIAAVA-WLASSEASYVTG    242
                      .       . .    .    ....* .  .*.. .   .**

HDEREP1      QFFEVAAGFYAQIRWERSGGVLFKPDQS--------------FTAE----    261
HDEREP2      ETFEIGGGWIGNTRWQRAKGAVSHDEHTTVEFIKEHLNEITDFTTDTENP    268
PHBB$ZOORA   STISANGGQF----------------------FV----------------    241
DHG$BACSU    ITLFADGG-----------------------------MTQYPSFQAG----    258
DHGB$BACME   ITLFADGG-----------------------------MTQYPSFQAG----    260
               .  ..*

HDEREP1      --    261
HDEREP2      KS    270
PHBB$ZOORA   --    241
DHG$BACSU    RG    260
DHGB$BACME   RG    262
```

Appendix 2

Amino acid alignment of HDE and *S. cerevisiae* trifunctional enzyme.

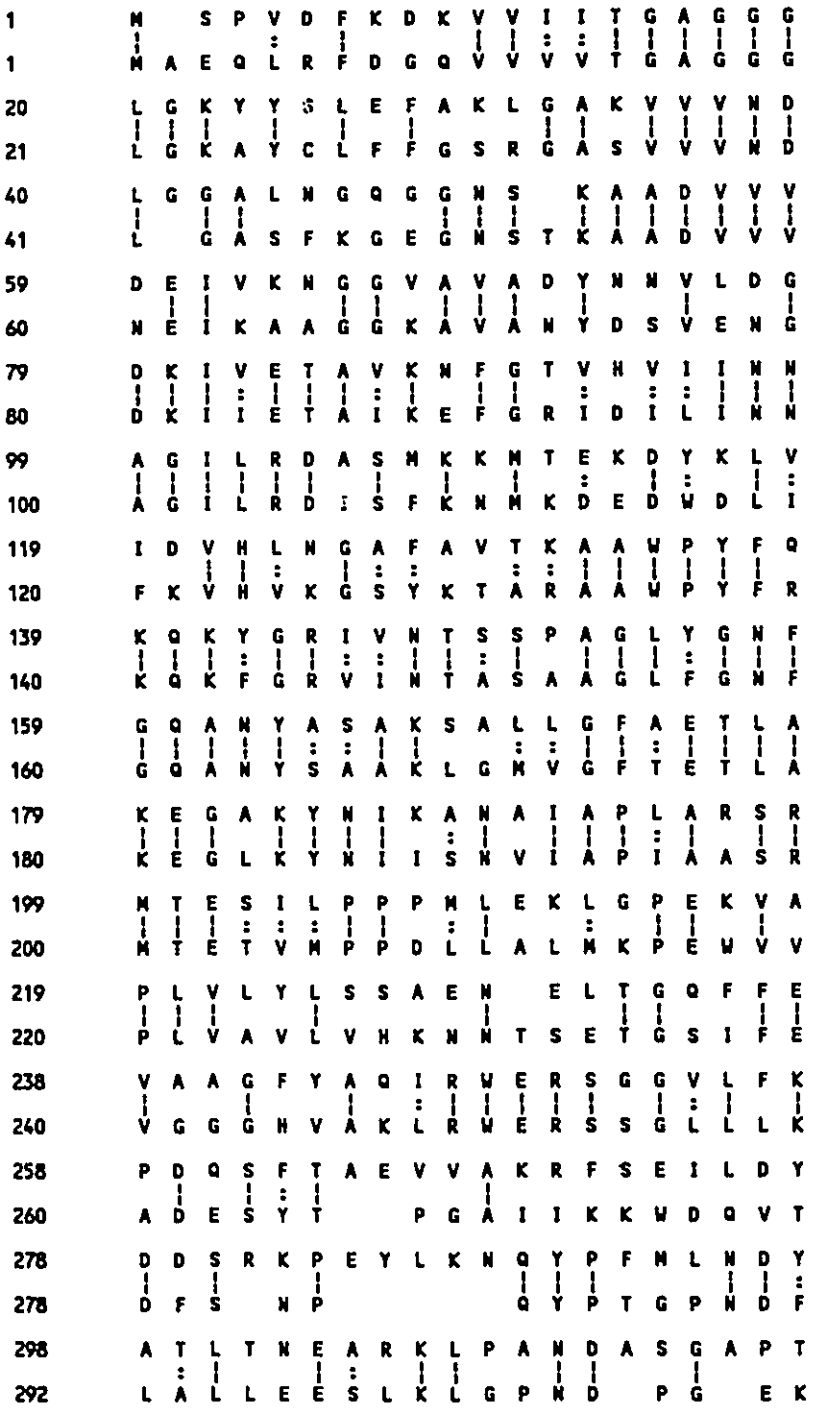


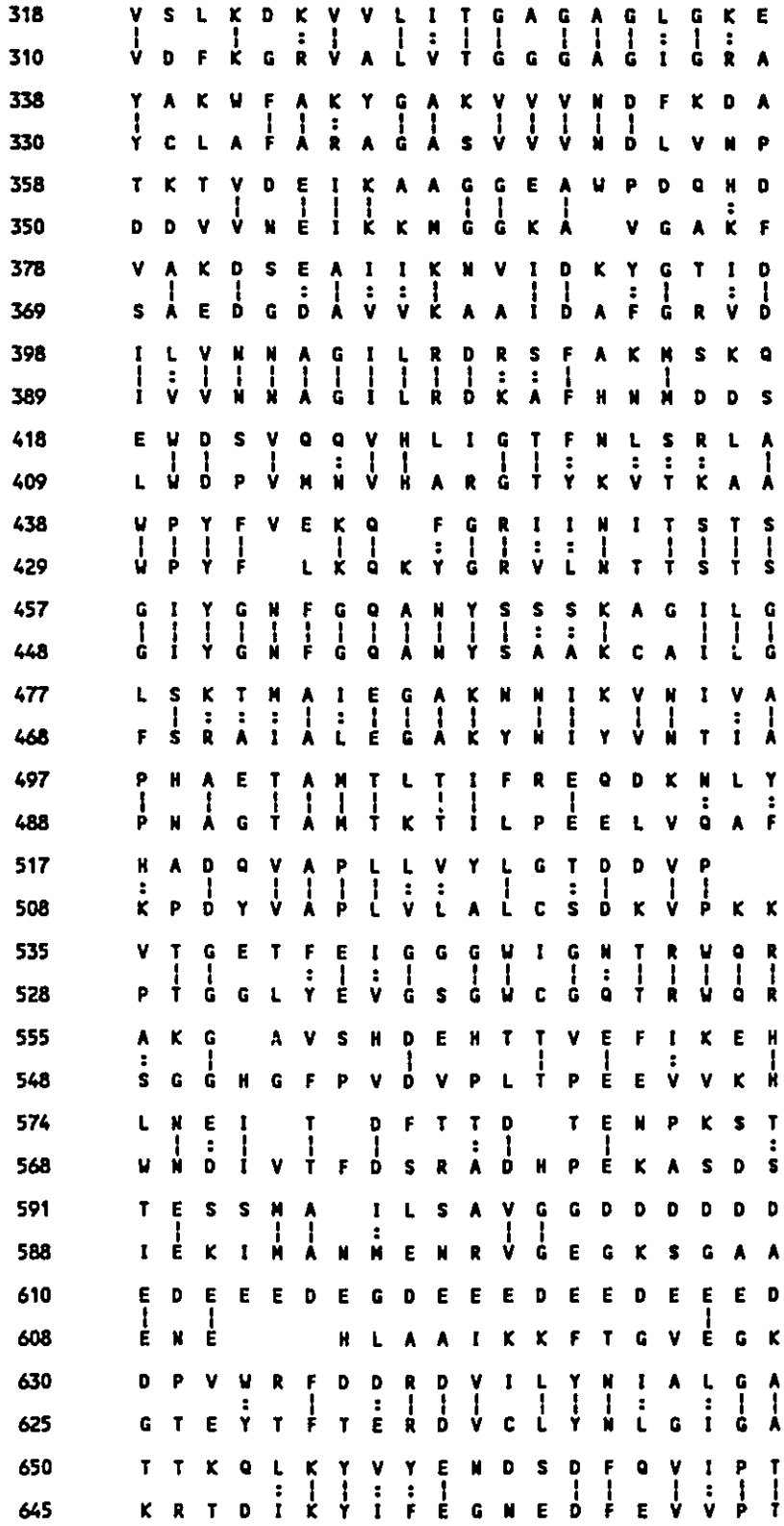


```
648  G A T T K Q L K Y V Y E N D S D F Q V I
     | | | : | | | | | | | | | | | | | | :
639  G C T S K E L K Y T Y E N D P D F Q V L

668  P T F G H L I T F N S G K S Q N S F A K
     | : | |   : | |   | |   |   |   : |   |
659  P T F   A V I P F M Q A T A T L A M D N

688  L L R N F N P M L L L H G E H Y L K V H
     | : |   | | | |   |   : | | | | | | : |
678  L V D N F N Y A M L L H G E Q Y F K L C

708  S W P P P T E G E I K T T F E P I A T T
     : |       | : |   | | | | | | | : |
698  T P T M P S N G T L K T L A K P L Q V L

728  P K   G T N V V I V H G S K S V D N K S
     |   |   | |   |   : : | | | | | : | | :
718  D K N G K A A L V V G G F E T Y D I K T

747  G E L I Y S N E A T Y F I R   N C Q A D
     |   | | | |   | | | | : : | | | |
738  K K L I A Y N E G S F F I R G A H V P P

766  N K V Y A D   R P A F A T N Q F L A P
     |   |       | | | | | | : : | | | | |
758  E K E V R D G K R A K F A V Q N F E V P

784    K R A P D Y Q V D V P V S E D L A A L
       | : | | | : |   : : |   | | | | | |
778  H G K V P D F E A E I S T N K D Q A A L

803  Y R L S G D R N P L H I D P N F A K G A
     | | | | | | |   | | | | | | | | | | |
798  Y R L S G D F N P L H I D P T L A K A V

823  K F P K P I L H G M C T Y G L S A K A L
     | | | | | | | | | : | | |   | : | | | |
818  K F P T P I L H G L C T L G I S A K A L

843  I D K F G M F N E I K A R F T G I V F P
       : : | |   | | | : | | | | | | : | | |
838  F E H Y G P Y E E L K V R F T N V V F P

863  G E T L R V L A W K E S D D T I V F Q T
     | : | | : |   | | | |   |   |   : | | | |
858  G D T L K V K A W K   Q G S V V V F Q T

883  H V V D R G T I A I N N A A I K L V G D
       |   |   | | : |   | | | | : | | |
877  I D T T R N V I V L D N A A V K L S Q A

903  K A K I
     | : | :
897  K S K L
```

Matches = 490 Mismatches = 392 Unmatched = 42
Length = 924 Matches/length = 53.0 percent

Amino acid alignment of HDE and *N. crassa* trifunctional enzyme.

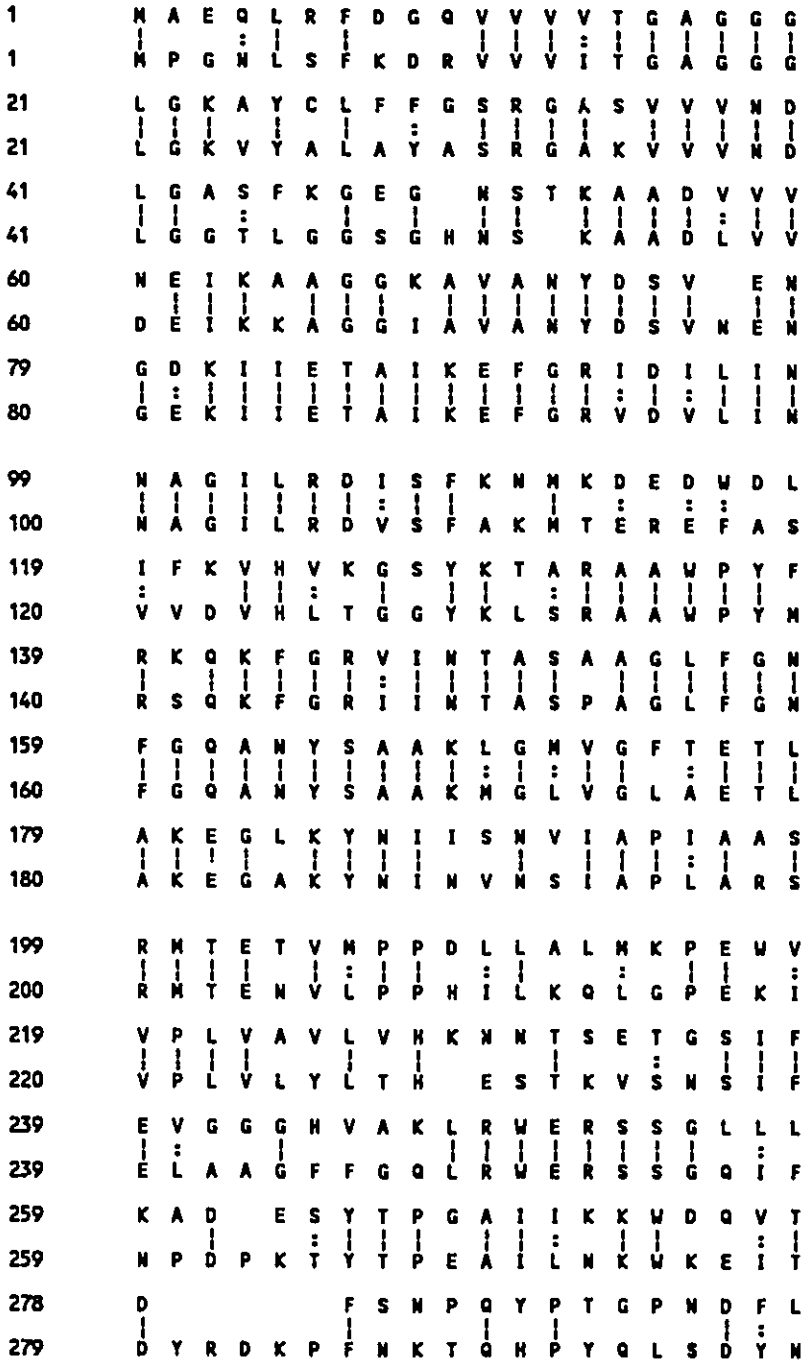


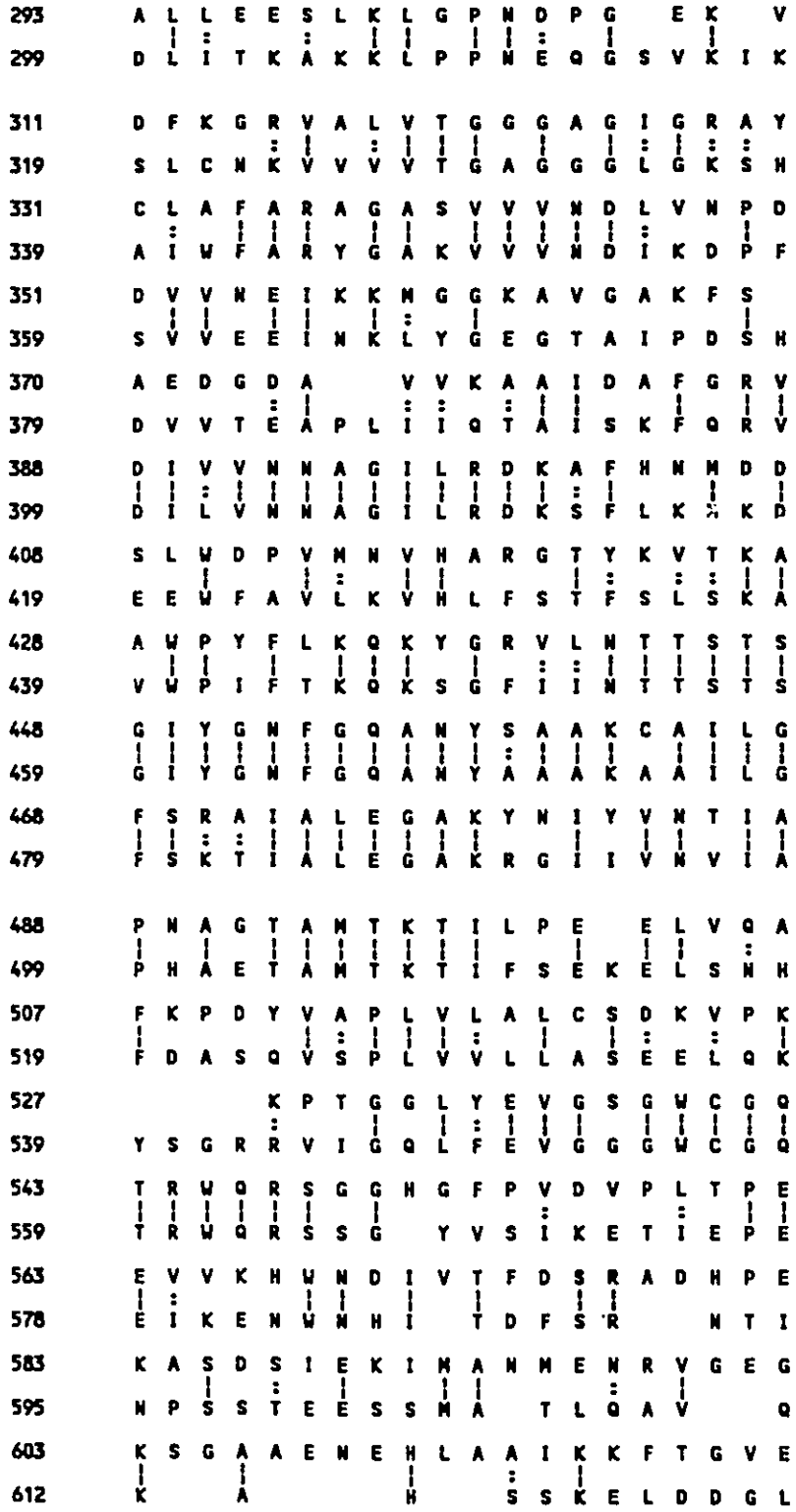

318 V S L K D K V V L I T G A G A G L G K E
310 V D F K G R V A L V T G G G A G I G R A
338 Y A K W F A K Y G A K V V V N D F K D A
330 Y C L A F A R A G A S V V V N D L V N P
358 T K T V D E I K A A G G E A W P D Q H D
350 D D V V N E I K K N G G K A V G A K F
378 V A K D S E A I I K N V I D K Y G T I D
369 S A E D G D A V V K A A I D A F G R V D
398 I L V N N A G I L R D R S F A K N S K Q
389 I V V N N A G I L R D K A F H N M D D S
418 E W D S V Q Q V H L I G T F N L S R L A
409 L W D P V M N V H A R G T Y K V T K A A
438 W P Y F V E K Q F G R I I N I T S T S
429 W P Y F L K Q K Y G R V L N T T S T S
457 G I Y G N F G Q A N Y S S S K A G I L G
448 G I Y G N F G Q A N Y S A A K C A I L G
477 L S K T M A I E G A K N N I K V N I V A
468 F S R A I A L E G A K Y N I Y V N T I A
497 P H A E T A M T L T I F R E Q D K N L Y
488 P N A G T A M T K T I L P E E L V Q A F
517 H A D Q V A P L L V Y L G T D D V P
508 K P D Y V A P L V L A L C S D K V P K K
535 V T G E T F E I G G G W I G N T R W Q R
528 P T G G L Y E V G S G W C G Q T R W Q R
555 A K G A V S H D E H T T V E F I K E H
548 S G G H G F P V D V P L T P E E V V K H
574 L N E I T D F T T D T E N P K S T
568 W N D I V T F D S R A D H P E K A S D S
591 T E S S M A I L S A V G G D D D D D D D
588 I E K I M A N M E N R V G E G K S G A A
610 E D E E E D E G D E E E D E E E E D
608 E N E H L A A I K K F T G V E G K
630 D P V W R F D D R D V I L Y N I A L G A
625 G T E Y T F T E R D V C L Y N L G I G A
650 T T K Q L K Y V Y E N D S D F Q V I P T
645 K R T D I K Y I F E G N E D F E V V P T

```
670       F  G  H  L  I  T  F  N  S  G  K  S  Q  N  S  F  A  K  L  L
                    :              |  |              |  |           :  :
665       F  G  V  I  P  P  F  N     T  E  M  P  F  S  F  D  D  I  V

690       R  N  F  N  P  M  L  L  L  N  G  E  H  Y  L  K  V  N  S  W
             |  |     |  |  :  |  |  |  |  |  |  |  |  |  |  :  :  :
684       P  N  F  S  P  M  M  L  L  N  G  E  Q  Y  L  E  V  R  K  Y

710       P  P  P  T  E  G  E  I  K  T  T  F  E  P  I  A  T  T  P  K
             |  |  |     |     :        :              :              |
704       P  I  P  T  S  G  R  L  V  S  K  G  K  L  L  E  V  V  D  K

730       G  T  N  V  V  I  V     H  G  S  K  S  V  D  N  K  S  G
             |     |  |  |  |        |  |        :        |  :  |
724       G     S  A  A  I  V  K  Q  G  I  T  T  F  N  A  E  T  G  E

748       E  L  I  Y  S  N  E  A  T  Y  F  I  R  N  C           Q  A  D
             |  |  |  |  |  |  |  |  |  |  |  :  |  |              |  |  |
743       E  L  F  Y     N  E  M  T  V  F  L  R  G  C  G  G  F  G  G

766       N  K  V  Y  A  D  R  P  A  F  A  T  N  Q  F  L  A  P  K  R
          :  |              |  |  |     |  :  |              |  |  |
762       Q  K  K  P  A  D  R  G  A     S  T  A  A  N  K  P  P  A  R

786       A  P  D  Y  Q  V  D  V  P  V  S  E  D  L  A  A  L  Y  R  L
          :  |  |           |  :  |        :  |  :     |  |  :  |  |  |
781       S  P  D  A  V  V  E  V  Q  T  T  E  E  Q  A  A  I  Y  R  L

806       S  G  D  R  N  P  L  H  I  D  P  N  F  A  K  G  A  K  F  P
          |  |  |     |  |  |  |  :  |  |     |  |  |  |              |
801       S  G  D  Y  N  P  L  H  V  D  P  A  F  A  K  V  G  G  F  K

826       K  P  I  L  H  G  M  C  T  Y  G  L  S  A  K  A  L  I  D  K
             |  |  |  |  |  :  |  :  :  |  :  :     |  |  :     :  |
821       V  P  I  L  H  G  L  C  S  F  G  I  A  G  K  A  V  Y  E  K

846       F  G  N  F  N  E  I  K  A  R  F  T  G  I  V  F  P  G  E  T
          :  |     |        |  |        |  |  :  |     |  |  |  |  |
841       Y  G  K  F  K  N  I  K  V  R  F  A  G  T  V  N  P  G  Q  T

866       L  R  V  L  A  W  K  E  S  D  D  T  I  V  F  Q  T  H  V  V
          |              |  |  |              :  |  |  |  |  :  |
861       L  V  T  E  M  W  K  E     G  N  K  V  V  F  Q  T  K  V  K

886       D  R  G  T  I  A  I  N  N  A  A  I  K  L  V  G  D  K  A  K
          :     |     :  |  |              |  |        |
880       E  T  G  K  L  A  I  S  G  A  A  A  E  L  A

906       I
```

Matches = 432 Mismatches = 447 Unmatched = 42
Length = 921 Matches/length = 46.9 percent

Amino acid alignment of *N. crassa* and *S. cerevisiae* trifuntional enzymes.





```
623       G  K  G  T  E  Y  T  F  T  E  R  D  V  C  L  Y  N  L  G  I
                                  |  |  :  |  |  |  |  |  |  |  :
624       F  K           Y  T     T     K  D  C  I  L  Y  N  L  G  L

643       G  A  K  R  T  D  I  K  Y  I  F  E  G  N  E  D  F  E  V  V
          |              :  :  |  |     :  |                    |  :
639       G  C  T  S  K  E  L  K  Y  T  Y  E  N  D  P  D  F  Q  V  L

663       P  T  F  G  V  I  P  P  F  N  T  E  M  P  F  S  F  D  D  I
          |  |  |     |  |  |           :  :                 |     :
659       P  T  F  A  V  I  P  F  N  Q  A  T  A  T  L  A  M  D  N  L

683       V  P  N  F  S  P  M  M  L  L  H  G  E  Q  Y  L  E  V  R  K
          |     :  |                 |  |  |  |  |  |  |           :
679       V  D  N  F  N  Y  A  M  L  L  H  G  E  Q  Y  F  K  L  C  T

703       Y  P  I  P  T  S  G  R  L  V  S  K  G  K  L  L  E  V  V  D
                :  |  :     |     |        :           |     |  |  |
699       P  T  M  P  S  N  G  T  L  K  T  L  A  K  P  L  Q  V  L  D

723       K     G  S  A  A  I  V  K  Q  G  I  T  T  F  N  A  E  T  G
          |     |  |  |  |  :  |                 |  :           |
719       K  N  G  K  A  A  L  V  V  G  G  F  E  T  Y  D  I  K  T  K

742       E  E  L  F  Y  N  E  M  T  V  F  L  R  G  C  G  G  F  G  G
                :     |  |  |     :     |  :  |  |
739       K  L  I  A  Y  N  E  G  S  F  F  I  R  G     A  M  V  P  P

762       Q  K  K  P  A  D     R  G  A  S  T  A     A  N        K  P
             |              :  |              |     |  |           |
758       E  K  E  V  R  D  G  K  R  A  K  F  A  V  Q  N  F  E  V  P

778       P  A  R  S  P  D  A  V  V  E  V  Q  T  T  E  E  Q  A  A  I
                :     |  |              |  :     |        :  |  |  :
778       H  G  K  V  P  D  F  E  A  E  I  S  T  N  K  D  Q  A  A  L

798       Y  R  L  S  G  D  Y  M  P  L  H  V  D  P  A  F  A  K  V  G
          |  |  |  |  |  |  :  |  |  |  :  |  |  |  :     |  |
798       Y  R  L  S  G  D  F  M  P  L  H  I  D  P  T  L  A  K  A  V

818       G  F  K  V  P  I  L  H  G  L  C  S  F  G  I  A  G  K  A  V
             |              |  |  |  |  |  :     |  |  :     |  |  :
818       K  F  P  T  P  I  L  H  G  L  C  T  L  G  I  S  A  K  A  L

838       Y  E  K  Y  G  K  F  K  M  I  K  V  R  F  A  G  T  V  M  P
          :  |  :  |  |     :              :  |  |  |  :           |
838       F  E  H  Y  G  P  Y  E  E  L  K  V  R  F  T  N  V  V  F  P

858       G  Q  T  L  V  T  E  M  W  K  E  G  N  K  V  V  F  Q  T  K
          |  |  |  |                 |  |  |  |  |  |  |  |  |  |
858       G  D  T  L  K  V  K  A  W  K  Q  G  S  V  V  V  F  Q  T  I

878       V  K  E  T  G  K  L  A  I  S  G  A  A  A  E  L  A
                            :     :        |  |           |  :
878       D  T  T  R  N  V  I  V  L  D  M  A  A  V  K  L  S  Q  A  K


898       S  K  L
```

Matches = 461 Mismatches = 410 Unmatched = 52
Length = 923 Matches/length = 49.9 percent