BIOGENESIS OF PEROXISOMES IN THE YEAST *CANDIDA TROPICALIS* pK233: GENES, PROTEINS AND ORGANELLAR TARGETING

By

WILLIAM WAYNE MURRAY, B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

BIOGENESIS OF PEROXISOMES IN THE YEAST *CANDIDA TROPICALIS*

DOCTOR OF PHILOSOPHY (1992)
(Biochemistry)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Biogenesis of Peroxisomes in the Yeast Candida tropicalis pK233: Genes, Proteins and Organellar Targeting

AUTHOR: William Wayne Murray
B.Sc. (McMaster University)
M.Sc. (McMaster University)

SUPERVISOR: Dr. R.A. Rachubinski, Associate Professor

NUMBER OF PAGES: xxvi, 502

reading frame, without intervening sequences, of 2,127 nucleotides which encodes a polypeptide of 709 amino acids (M_r 79155). The deduced amino acid sequence of C. tropicalis AOx exhibits significant identity and homology with that of several other acyl-coenzyme A oxidases of C. tropicalis, as well as those of other yeasts (C. maltosa and S. cerevisiae) and rat. By comparison, sequencing of the C. tropicalis gene (or cDNA) encoding catalase (Cat) revealed a single open reading frame of 1,455 nucleotides, without intervening sequences, which encodes a polypeptide of 485 amino acids (M_r 54944). The deduced amino acid sequence of C. tropicalis Cat shows significant identity and homology with that of catalases from the yeast S. cerevisiae, mammals (rat, human and bovine), plants (maize, sweet potato and cottonseed) and Drosophila. Codon usage in the genes encoding C. tropicalis AOx and Cat is nonrandom, exhibiting a strong bias toward a group of preferred codons which is similar to that of several highly expressed genes of S. cerevisiae and E. coli, as well as other C. tropicalis genes encoding peroxisomal proteins.

Having achieved the characterization of the genes encoding C. tropicalis AOx and Cat, attempts were made to establish an in vitro system for the translocation of Cat into peroxisomes purified from C. tropicalis which could be used to determine the peroxisomal targeting signal (PTS) of Cat. For this purpose, N-terminal, C-terminal, N-/C-

ABSTRACT

Studies concerned with peroxisome biogenesis have established that peroxisomal proteins are encoded exclusively by nuclear genes, synthesized (with few exceptions) on free polyribosomes at their mature size and posttranslationally imported into preexisting organelles. From this, it was realized that the information responsible for the specificity of protein targeting to peroxisomes must reside within the mature amino acid sequence of the imported protein. Knowledge of the targeting sequences of peroxisomal proteins is, however, limited. Consequently, using the yeast C. tropicalis pK233 as a model system, this work endeavoured to clone and sequence a cDNA or gene encoding a peroxisomal protein so that the subsequent creation of deletion mutants of this cDNA/gene, together with the development of an in vitro and/or in vivo expression and translocation system for the corresponding polypeptides, would allow the amino acid sequence(s) comprising the peroxisomal addressing signal of this protein to be identified.

Related efforts resulted in the cloning of two genes (and one composite full-length cDNA) encoding different enzymes associated with the β-oxidation pathway of C. tropicalis peroxisomes. Sequencing of the gene encoding acyl-coenzyme A oxidase (AOx), revealed a single open

reading frame, without intervening sequences, of 2,127 nucleotides which encodes a polypeptide of 709 amino acids (M_r 79155). The deduced amino acid sequence of C. tropicalis AOx exhibits significant identity and homology with that of several other acyl-coenzyme A oxidases of C. tropicalis, as well as those of other yeasts (C. maltosa and S. cerevisiae) and rat. By comparison, sequencing of the C. tropicalis gene (or cDNA) encoding catalase (Cat) revealed a single open reading frame of 1,455 nucleotides, without intervening sequences, which encodes a polypeptide of 485 amino acids (M_r 54944). The deduced amino acid sequence of C. tropicalis Cat shows significant identity and homology with that of catalases from the yeast S. cerevisiae, mammals (rat, human and bovine), plants (maize, sweet potato and cottonseed) and Drosophila. Codon usage in the genes encoding C. tropicalis AOx and Cat is nonrandom, exhibiting a strong bias toward a group of preferred codons which is similar to that of several highly expressed genes of S. cerevisiae and E. coli, as well as other C. tropicalis genes encoding peroxisomal proteins.

Having achieved the characterization of the genes encoding C. tropicalis AOx and Cat, attempts were made to establish an in vitro system for the translocation of Cat into peroxisomes purified from C. tropicalis which could be used to determine the peroxisomal targeting signal (PTS) of Cat. For this purpose, N-terminal, C-terminal, N-/C-

terminal and internal deletion mutants of Cat were constructed, expressed in cell-free systems and assayed for translocation into purified C. tropicalis peroxisomes. Unfortunately, all attempts at in vitro import proved to be inadequate for the translocation of Cat into purified C. tropicalis peroxisomes. Since a functional in vitro system could not be established, studies concerned with the in vivo expression and targeting of Cat from C. tropicalis to the peroxisomes of C. albicans and S. cerevisiae were initiated. Although, the in vivo expression of Cat sequences could not be attained in C. albicans, the expression of full-length and truncated Cat polypeptides was achieved in S. cerevisiae. Subsequent determination of the subcellular location of the full-length Cat polypeptide suggested that Cat from C. tropicalis was targeted to the peroxisomes of S. cerevisiae, reaffirming that heterologous signals for import into peroxisomes are recognized in S. cerevisiae. Interestingly, the deletion of amino acids 182 to 344 of Cat negated the targeting of Cat to S. cerevisiae peroxisomes. These preliminary in vivo studies suggest that at least part of the PTS of Cat from C. tropicalis resides at an internal location within its mature amino acid sequence.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my research supervisor, Dr. R.A. Rachubinski, for his guidance, encouragement and patience during the tenure of this work. Rick, the education I have received in your laboratory will provide an excellent foundation upon which to build future experience. Thank you.

I must also extend my sincere thanks to the members of my supervisory committee, Dr. J.P. Capone, Dr. P. Chang and Dr. K.B. Freeman, for the many helpful discussions and constructive criticisms which directed and shaped this thesis.

Many thanks are extended to my fellow lab workers, Mike Tyers ("Billy", "Mr. Mike"), Andrea Bodnar ("Chuck"), Jill Regoeczi ("Lots and lots of ..... garage."), John Aitchison ("Johnny A"), Bill Nuttley ("Nutter"), Jim Sloots ("Rookie"), John Glover ("Gloveman"), Baowei Zhang ("Zhanger") and Tony Brade for creating a stimulating environment promoting scientific thought, discussion and research. Equally important, were all the laughs enjoyed and memories retained along the way. I wish all of you the best of luck in your future endeavours.

The professional manner in which this thesis is typed is the result of a seemingly infinite number of hours

of work expended by Sue Mercer. Beyond this, Suzanne, you provided timely inspiration, encouragement and support, without which, this thesis would not have been completed. Thank you, for all of your efforts.

Technical assistance, which was greatly appreciated, is recognized as follows: John Aitchison, for production and characterization of the rabbit anti-C. tropicalis catalase antibody and help in attempting to develop the in vitro import system; Dr. D.W. Andrews, for help in establishing the in vitro wheat germ extract translation system; Andrea Bodnar, for assistance with peroxisome isolations and subsequent enzyme assays; John Glover, for assistance with the subcloning, expression and subcellular localization of Cat constructs in S. cerevisiae; Jill Regoeczi, for any type of help, at any and several times along the way; and Mike Tyers, for help with the double-stranded cDNA synthesis.

Grateful acknowledgements should be expressed to the Ontario Ministry of Colleges and Universities for financial support in the form of Ontario Graduate Scholarships, as well as to McMaster University for its financial support through scholarships and teaching assistantships.

Finally, I thank my family for always being there when I needed them. Your contribution to this thesis, through love, support and encouragement, has been the most important, bar none.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A	ampere(s)
A	adenosine
ADH-I	alcohol dehydrogenase I from S. cerevisiae
ADH-II	alcohol dehydrogenase II from S. cerevisiae
AGT1	L-alanine:glyoxylate aminotransferase 1
ala(A)	alanine
Amp^r	resistance to ampicillin
Amp^s	sensitivity to ampicillin
AMV	Avian myeloblastosis virus
AOx	acyl-coenzyme A oxidase (M_r 79006) from C. tropicalis
AOX1	acyl-coenzyme A oxidase from C. maltosa
AOx TypeI	acyl-coenzyme A oxidase from rat
AP	ammonium persulfate
arg(R)	arginine
asn(N)	asparagine
asp(D)	aspartic acid (aspartate)
ATP	adenosine 5'-triphosphate
Bluo-gal	halogenated indolyl-β-D-galactoside
bp	base pair(s)
Brij 35	polyoxyethylene 23-lauryl ether
C	cytidine
Cat	catalase from C. tropicalis

CAT	chloramphenicol acetyltransferase from E. coli
CAT-1	catalase isozyme-1 from maize (Zea mays L.)
CBI	codon bias index
cDNA	DNA complementary to RNA
CHO cells	Chinese hamster ovary cells
Ci	curie(s)
CoA	coenzyme A
cpm	count(s) per minute
CPR	NADPH-cytochrome P-450 oxidoreductase from C. tropicalis
CTA1	S. cerevisiae gene encoding catalase A
CTP	cytidine 5'-triphosphate
CTT1	S. cerevisiae gene encoding catalase T
cys(C)	cysteine
d	day(s)
Da	dalton(s)
DAB	diaminobenzidine
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
ddNTP	2',3'-dideoxynucleoside 5'-triphosphate
°C	degrees Celsius
dGTP	deoxyguanosine 5'-triphosphate
DHAP	dihydroxyacetone phosphate
DHCA	3α,7α-dihydroxy-5β-cholestanoic acid
DHFR	dihydrofolate reductase
14DM	cytochrome P450 lanosterol 14α-demethylase from

	C. tropicalis
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside 5'-triphosphate
DOC	deoxycholic acid
DTT	dithiothreitol
dTTP	deoxythymidine 5'-triphosphate
E-I	enolase I from S. cerevisiae
E-II	enolase II from S. cerevisiae
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
ER	endoplasmic reticulum
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
g	gram(s)
G	guanosine
G3PDH-I	glyceraldehyde-3-phosphate dehydrogenase I from S. cerevisiae
G3PDH-II	glyceraldehyde-3-phosphate dehydrogenase II from S. cerevisiae
GAPDH	glyceraldehyde-phosphate dehydrogenase
g_{av}	acceleration of gravity (average)
gln(Q)	glutamine

glu(E)	glutamic acid (glutamate)
gly(G)	glycine
g_{max}	acceleration of gravity (maximum)
GTP	guanosine 5'-triphosphate
h	hour(s)
HDE	hydratase-dehydrogenase-epimerase from C. tropicalis
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
his(H)	histidine
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A
HSP	heat shock protein
HXK1	hexokinase PI from S. cerevisiae
HXK2	hexokinase PII from S. cerevisiae
ile(I)	isoleucine
IMP	integral membrane protein
IPTG	isopropyl β-D-thiogalactoside
J	joule(s)
k	kilo
l (or L)	litre(s)
lacI	lactose repressor from E. coli
leu(L)	leucine
lpp	outer membrane lipoprotein from E. coli
lys(K)	lysine
m	metre(s)
μ	micro

m	milli
M	molar
MCS	multiple cloning site
MES	2-(N-morpholino)ethanesulfonic acid
met(M)	methionine
min	minute(s)
mol	mole(s)
MOPS	3-(N-morpholino)propanesulfonic acid
M_r	relative molecular mass
mRNA	messenger ribonucleic acid
n	nano
NAD	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate, reduced
NLS	nuclear localization signal
NMR	nuclear magnetic resonance
NP-40	Nonidet P-40
NTP	nucleoside 5'-triphosphate
Nycodenz	5-(N-2,3-dihydroxypropylacetamido)-2,4,6-triiodo-N,N'-bis(2,3-dihydroxypropyl)isopthal-amide
OD_x	optical density at a wavelength of x nm
ompA	outer membrane protein A from E. coli
ORF	open reading frame(s)
P_{450}Alk	cytochrome P450 from C. tropicalis
PEG	polyethylene glycol
PGK	phosphoglycerate kinase from S. cerevisiae

PH1	hyperoxaluria type 1
phe(F)	phenylalanine
pI	isoelectric point
PK	pyruvate kinase from S. cerevisiae
pmol	picomole(s)
PMP	peroxisomal membrane protein
PMSF	phenylmethylsulfonyl fluoride
PNS	post-nuclear supernatant
POX1	acyl-coenzyme A oxidase from S. cerevisiae
POX2	C. tropicalis gene encoding PXP-2
POX4	C. tropicalis gene encoding PXP-4
POX5	C. tropicalis gene encoding PXP-5
POX9	C. tropicalis gene encoding PXP-9
POX18	C. tropicalis gene encoding PXP-18
pro(P)	proline
psi	pound(s) per square inch
PTS	peroxisomal targeting signal
PXP-2	acyl-coenzyme A oxidase (M_r 81711) from C. tropicalis
PXP-4	acyl-coenzyme A oxidase II (M_r 78554) from C. tropicalis
PXP-5	acyl-coenzyme A oxidase I (M_r 74048) from C. tropicalis
PXP-9	catalase from C. tropicalis
PXP-18	nonspecific lipid-transfer protein from C. tropicalis

RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolution(s) per minute
s	second(s)
SCP-2	sterol carrier protein 2
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
ser(S)	serine
T	thymidine
TCA	trichloroacetic acid
TEMED	N,N,N',N'-tetramethylethylenediamine
Tet^r	resistance to tetracycline
Tet^s	sensitivity to tetracycline
THCA	3α,7α,12α-trihydroxy-5β-cholestanoic acid
thr(T)	threonine
TLCK	Nα-p-tosyl-L-lysine chloromethyl ketone
T_m	thermal denaturation temperature
Tris	Tris(hydroxymethyl)aminomethane
trp(W)	tryptophan
tsp	transcription start point(s)
tufA	elongation factor TU-A from E. coli
tufB	elongation factor TU-B from E. coli
Tween 40	polyoxyethylenesorbitan monopalmitate
Tween 80	polyoxyethylenesorbitan monooleate
tyr(Y)	tyrosine

U	uridine
UTP	uridine 5'-triphosphate
V	volt(s)
val(V)	valine
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

INTRODUCTION

1.1 MICROBODIES (PEROXISOMES) AND RELATED PARTICLES

1.1.1 Historical Perspective

In 1954, at the Karolinska Institute laboratory of Fritjof Sjostrand, Johannes A.G. Rhodin, a Swedish graduate student, introduced the term "microbody" into the field of electron microscopy (Rhodin, 1954). "Microbody" was used to define a previously unidentified, nondescript, cytoplasmic body present in the convoluted tubule cells of the mouse kidney, characterized as being approximately 0.5 μm in diameter, surrounded by a single membrane and filled with a fine granular matrix. Two years later, "granules" resembling kidney microbodies, but containing a dense core with a regular crystalloid structure, were described in the parenchymal cells of rat liver by Rouiller and Bernhard (1956) and Gänsler and Rouiller (1956), who hypothesized, incorrectly, that these hepatic microbodies might be the precursors of mitochondria. Concomitantly, biochemical studies initiated by Novikoff et al. (1953), and confirmed by de Duve et al. (1953, 1955), revealed that rat liver uricase (urate oxidase) exhibits sedimentation properties similar to those of acid phosphatase (ie. lysosomes), but intermediate between those of mitochondria and microsomes. By contrast, however, urate oxidase does not display the

solubility and latency properties of the lysosomal hydrolases (de Duve et al., 1955). These differences led to the postulation that urate oxidase is associated with an alternative group of granules, "resembling large microsomes rather than small mitochondria, and differing from lysosomes by a greater uniformity in sedimentation properties" (de Duve et al., 1955). Within five years, these hypothetical granules were sufficiently characterized to ascertain that catalase and D-amino acid oxidase were also present within microbodies (de Duve et al., 1960); within eight years, upon the discovery that the injection of rats with the mild detergent Triton WR-1339 causes a selective and drastic decrease in the density of lysosomes (in which the detergent accumulates), the isolation of microbodies, independent of any contaminating "particles" (lysosomes), was achieved via density equilibration in a sucrose gradient (Wattiaux et al., 1963). Further investigations into other biological systems established complementary results to those of de Duve et al. (1960), not only in rat kidney, but more significantly, in the ciliate, Tetrahymena pyriformis (Baudhuin et al., 1965). In addition, two other enzymes, L-α-hydroxyacid oxidase and, in kidney, L-a ino acid oxidase, were found to be associated with microbodies (Baudhuin et al., 1965). Collectively, these findings strengthened the hypotheses that microbodies are important sites of hydrogen peroxide metabolism, potentially in a wide variety of

organisms, and that the association of oxidases with catalase is biologically meaningful (de Duve, 1965a). Consequently, de Duve (1965b) proposed the name 'peroxisome' to describe those organelles (previously called microbodies) containing at least one oxidase to form hydrogen peroxide and catalase to decompose it (ie. the name refers to hydrogen peroxide metabolism, not to the presence of peroxidases). The subsequent development of a special zonal rotor by Beaufay (1966) enabled the large-scale preparation of highly-purified rat liver peroxisomes (Leighton et al., 1968) and opened the door to the more than twenty years of research devoted to the characterization of this organelle that has followed. As a consequence, the properties, functions, species distribution and differences, biogenesis, genetic regulation, pathological conditions and associated defects and evolutionary history of peroxisomes have become increasingly understood. Certain of these disciplines (ie. those pertaining directly to this thesis) will be reviewed in depth herein. For a comprehensive review of those areas outside the scope of this thesis, the following recent reviews should be consulted: Tolbert, 1981; Kindl and Lazarow, 1982; de Duve, 1983; Osumi and Hashimoto, 1984; Masters and Crane, 1984, 1985; Lazarow and Fujiki, 1985; Borst, 1986, 1989; Fahimi and Sies, 1987; Lazarow, 1987a, 1987b; Rodricks and Turnball, 1987; Moser, 1988; Mannaerts and Van Veldhoven, 1990. Those reviews suited to gaining a

more extensive historical perspective of the discovery and early research devoted to microbodies/peroxisomes include: de Duve, 1965a, 1989; de Duve and Baudhuin, 1966; Hogg, 1969; Tolbert, 1971; Masters and Holmes, 1977; Beevers, 1979; Fukui and Tanaka, 1979.

1.1.2 Nomenclature

Within the relevant literature, three different names, microbody, glyoxysome and microperoxisome, have been used to describe the subcellular organelle most often referred to as the peroxisome. As a general, morphological term, "microbody" is used to describe a species-specific organelle which may be native to prokaryotic and/or eukaryotic organisms, is bounded by a single unit membrane, possesses a homogeneous matrix and exhibits a diameter of between 0.2 and 1.0 μm (Tolbert, 1981; Lazarow and Fujiki, 1985). Functionally, "microbodies" are multi-purpose organelles containing a β-oxidation system and a diversity of other (often inducible) organism- and tissue-specific enzyme systems, which may share a common evolutionary origin that is distinct from that of mitochondrial systems (Borst, 1986, 1989). Although the term "microbody" has previously been used to describe similar structures in several cell types (Hruban and Rechcigl, 1969), cytochemical and cell fractionation studies have established that "microbodies" are usually peroxisomes. Documented exceptions to this

generality include: (i) the hydrogenosomes of Trichomonads (reviewed by Müller, 1975, 1988), (ii) the glycosomes of Trypanosomatids (reviewed by Opperdoes, 1982, 1987a, 1987b, 1988) and (iii) the catalase-less microbodies of Neurospora crassa (Kunau et al., 1987).

Hydrogenosomes were originally described in trichomonad flagellates (eg. Trichomonas vaginalis, a parasite of humans) and have since been detected in a number of free-living anaerobic ciliates and in ciliates and a fungus of the rumen (Müller, 1988; Finlay and Fenchel, 1989). Within this organelle, which is known to be delimited by two membranes (Honinberg et al., 1984), the principle biochemical reaction is the oxidative decarboxylation of pyruvate, with the accompanying formation of acetyl-CoA. The process is mediated by pyruvate:ferredoxin oxidoreductase and is coupled, via hydrogenase, to the anaerobic formation of H_2. In subsequent reactions, substrate-level phosphorylation of ADP serves as the means to conserve the energy of the thioester bond of acetyl-CoA (Müller, 1975, 1988). This pathway and its associated enzymes, together with recent inferences concerning protein targeting to hydrogenosomes (section 1.4.3.2; Keller et al., 1991), suggest that hydrogenosomes are markedly distinct from peroxisomes [and other microbodies (ie. glyoxysomes and glycosomes)].

Glycosomes, on the other hand, were originally

described in trypanosomatid hemoflagellates (eg. Trypanosoma brucei, another human parasite) and have since been found in all representatives of this family, including Trypanosoma cruzi, Crithidia spp. and Leishmania mexicana (Opperdoes, 1982). Characteristically, glycosomes resemble most other microbodies in the following respects: (i) they possess a diameter of between 0.2 and 0.8 μm, (ii) they are bounded by a single bilayer membrane, (iii) they sometimes contain a crystalloid core and (iv) they equilibrate at 1.23 g/cc in sucrose gradients (Opperdoes, 1987a, 1987b). Contrary to other microbodies, however, glycosomes are responsible biochemically for the conversion of glucose and glycerol to 3-phosphoglycerate and accordingly, contain the initial enzymes of glycolysis (hexokinase, glucosephosphate isomerase, phosphofructokinase, fructosebisphosphate aldolase, triosephosphate isomerase, glyceraldehyde-phosphate dehydrogenase and phosphoglycerate kinase) and glycerol catabolism (glycerol kinase and glycerol-3-phosphate dehydrogenase); in general, they do not contain H_2O_2-producing oxidases or catalase (Opperdoes, 1987a). In addition to the enzymes of glucose and glycerol metabolism, select enzymes involved in pyrimidine biosynthesis (orotate phosphoribosyltransferase and orotidine carboxylase), purine salvage (hypoxanthine-guanine phosphoribosyltransferase), ether-lipid biosynthesis [dihydroxyacetonephosphate (DHAP) acyltransferase, acyl-CoA reductase and alkyl/acyl DHAP

reductase], carbon dioxide fixation (phosphoenolpyruvate carboxykinase) and the β-oxidation of fatty acids have also been localized to glycosomes (Opperdoes, 1987a, 1987b). Thus, based upon its biochemistry, the glycosome, like the hydrogenosome, is an innately different microbody from the peroxisome.

Catalase-less microbodies, by contrast, are found in certain filamentous fungi, including Neurospora crassa (Kunau et al., 1987), and are endowed with a β-oxidation system that is not programmed for the production of H_2O_2. To what extent, if any, these microbodies are related to authentic peroxisomes remains to be established.

First isolated from the endosperm tissue of germinating castor beans (Breidenbach and Beevers, 1967), glyoxysomes, or "specialized plant peroxisomes" (Ettinger and Harada, 1990), are found in higher plants, certain nematodes and protozoa and house, in addition to catalase and those enzymes responsible for the β-oxidation of fatty acids (Cooper and Beevers, 1969), the glyoxylate cycle, a five enzyme variant of the tricarboxylic acid cycle (Beevers, 1982). Two enzymes of the glyoxylate cycle, isocitrate lyase and malate synthase, are localized exclusively to glyoxysomes and are, therefore, used as markers for this organelle (Ettinger and Harada, 1990). Contrary to the tricarboxylic acid cycle, in which acetyl-CoA is completely oxidized to CO_2, the glyoxylate cycle, in

each turn, converts two molecules of acetyl-CoA into one molecule each of succinate and oxaloacetate (Tolbert, 1981, 1982). In germinating seedlings, when fat is being converted to carbohydrate, "glyoxysomes", which are closely associated with lipid droplets and contain both an active β-oxidation system and glyoxylate cycle, process the fat stored in the cotyledons into glyoxylate cycle products. As the cotyledons mature into green leaves, another microbody, the "peroxisome", which is physically associated with chloroplasts, develops and is specialized to import glycolate produced by photorespiration in the chloroplast and convert it, via glyoxylate, into glycine. Glycine is then shuttled into the mitochondria, where two molecules of glycine condense to serine; serine subsequently returns to the peroxisome and is converted to glycerate, which serves the carbon reduction cycle in chloroplasts (Tolbert, 1981, 1982). Interestingly, the aforementioned physical associations have been visualized in electron micrographs of germinating seedlings, where glyoxysomes and peroxisomes are found wedged between mitochondria and lipid droplets, and mitochondria and chloroplasts, respectively (de Duve, 1983). During the intermediate phase between cotyledon and green leaf stages, both sets of enzymatic activities are present in the cotyledons. Of this, Beevers (1979) postulated the two-population (or replacement) model which suggests that two separate particle populations exist during the

transition and that this transition is accomplished by the destruction of glyoxysomes and simultaneous synthesis of microbodies containing peroxisomal activities. Alternatively, Trelease et al. (1971) hypothesized that the transition occurred within a single population of existing particles via a change in enzyme complement. This model, termed the one-population (or repackaging) model, predicts that glyoxysomal and peroxisomal enzymes will co-exist within individual microbodies during the transition period and that a substantial turnover of microbody membranes and proteins (eg. catalase) need not occur. Subsequent cytochemical, biochemical and morphometric analyses (Burke and Trelease, 1975; Schopfer et al., 1976; Köller and Kindl, 1978) supported the latter model; similarly, immunoelectron microscopy, using two sizes of protein A-gold particles, unequivocally demonstrated that both types of enzymes are present within the same organelle during the transition, verifying the existence of a single population of microbodies exhibiting dual function (Titus and Becker, 1985). These and other studies (Kindl, 1987; Behrends et al., 1990) further suggested that plant peroxisomes occur in more than one biochemically definable form. From this, evolved the concept that glyoxysomes/peroxisomes are organelles whose metabolic function and protein composition depend mainly on the supply of protein precursors present in the cytosol; this, in turn, is dependent primarily upon the

quality and quantity of concomitant gene expression (Titus and Becker, 1985; Kindl, 1987; Behrends et al., 1990). It has also been proposed that during the transition state the enzymes for the preceding process are subject to preferential degradation by endoproteinases (Behrends et al., 1990). In this manner, then, glyoxysomes undergo a continuous, functional transformation into leaf-type peroxisomes (Kindl, 1987; Behrends et al., 1990).

Of the three aforementioned names used in reference to "peroxisomes", microperoxisome is the least encountered in the literature. As a morphological term, originally defined by Novikoff et al. (1973), microperoxisome describes a cellular "particle" exhibiting a positive cytochemical [ie. diaminobenzidine (DAB)] reaction for catalase, which is characteristically small (approximately 0.2 μm in size), or coreless, or located near the endoplasmic reticulum (ER) (often characterized by continuity between its own delimiting membrane and that of the ER), or some combination of the former. Those microperoxisomes which have been investigated biochemically [eg. in the intestine (Small et al., 1980), adrenal cortex (Russo and Black, 1982), heart (Connock and Perry, 1983) and perisinusoidal fat-storing cells of liver (Gorgas, 1987)] were found to be authentic peroxisomes, often possessing β-oxidation activity.

1.2 PROPERTIES OF PEROXISOMES

1.2.1 General Characteristics

1.2.1.1 Abundance and Morphology

Peroxisomes, defined for their metabolism of hydrogen peroxide (see section 1.1.1), are found in animals, plants, fungi, bacteria and protozoa as single bilayer membrane-bounded subcellular organelles, which house a granular matrix of soluble enzymes (Tolbert, 1981; Lazarow and Fujiki, 1985). They exhibit a nearly ubiquituous eukaryotic cell and tissue distribution, such that they are present in most, if not all, mammalian cells (excluding mature erythrocytes); nonetheless, their abundance and morphology varies considerably (Lazarow and Fujiki, 1985; Lazarow, 1987a, 1987b). The organelle, which exhibits an abundance from one to hundreds per cell depending on cell type (Rodricks and Turnball, 1987), is most often spherical, ranging in size from a mean diameter of 0.2 μm in the human intestine (Lazarow, 1987a) to 1.5 μm in certain methanol-grown yeasts (eg. Hansenula polymorpha; Roggenkamp et al., 1989). Peroxisomes forming elongated, tubular structures have been observed in mouse liver (Gorgas, 1985), however, and in the mouse preputial gland, a sebaceous tissue that synthesizes ether glycerolipids and waxes (Sansone and Hamilton, 1969), peroxisomes do not exist as individual entities but rather form extensive interconnections between one another such that they create a single contorted

compartment varying in diameter, which is terminated by an enlarged cup-shaped structure (Gorgas, 1982, 1984). Peroxisomal profiles in individual sections taken across these interconnections are varied, with shapes ranging from small and round to either rod-shaped, dumbbells or ring-like cisternae (Gorgas, 1982, 1984).

The peroxisomal matrix is usually described as moderately electron-dense and finely granular or flocculent, although there are no actual granules (Goldfischer and Reddy, 1984). This homogeneous matrix helps to distinguish peroxisomes from mitochondria, which contain cristae, and lysosomes, which often contain vacuoles, lipids or myelin-like sheaths (Goldfischer and Reddy, 1984). In certain cell types [eg. hepatocytes or yeast (C. boidinii and H. polymorpha)], peroxisomes often contain a paracrystalline core structure [constituted by urate oxidase (Hruban and Swift, 1964) or alcohol oxidase (possibly, in combination with catalase) (Osumi et al., 1979; Veenhuis et al., 1981), respectively] which enhances their recognition in the electron microscope. Contrastingly, in other tissues [brain (nervous tissue), heart, skeletal muscle and epithelial cells of the intestine], peroxisomes are small (and are, therefore, usually referred tc as microperoxisomes) and difficult to detect without a cytochemical staining reaction (Tolbert, 1981; Lazarow and Fujiki, 1985; Lazarow, 1987a). Due to the presence of catalase activity within these

organelles, however, their cytochemical detection is enhanced by the H_2O_2-induced peroxidation of DAB to its conjugated unsaturated structure, which binds osmium tetraoxide (Goldfischer, 1969).

Tissues in which peroxisomes are particularly abundant include the liver and kidney; collectively, peroxisomes may constitute upwards of 3% of the total liver cell cytoplasm (Rodricks and Turnball, 1987). In other tissues (eg. nervous tissue), peroxisomes are relatively small and varied in number and their abundance may decrease with age (Rodricks and Turnball, 1987). Comparitively speaking, peroxisomes are more abundant in glia, satellite and Schwann cells than in neurons (Holtzmann, 1982). Among neurons, peroxisomes are more numerous in catecholaminergic cells, compared to Purkinje cells or cortical neurons, and more abundant during the early postnatal period rather than in the adult animal, as neurons of the cerebrum and cerebellum contain peroxisomes during the first two postnatal weeks, but only rarely after the third week (Moser, 1988). In oligodendrocytes, peroxisomes are often located in the cytoplasm adjacent to the forming myelin sheaths (Holtzmann, 1982); this is consistent with the notion that peroxisomes are required for the synthesis of at least one myelin component, the plasmalogens (see section 1.3.1.2).

1.2.1.2 Enzymic Content

As previously mentioned (section 1.1.1), the development of buoyant sucrose density gradient centrifugation in zonal rotors allowed the isolation of purified peroxisomal fractions, thereby permitting their enzymatic characterization (Beaufay, 1966; Leighton et al., 1968). Peroxisomes, for which catalase (and sometimes urate oxidase) serves as the marker enzyme, equilibrate at a relatively high density of ≈1.23 g/cm^3 on sucrose gradients due to their high protein content, their low lipid content from the single surrounding membrane and their rapid loss of water to the sucrose gradient (Tolbert, 1981; Lazarow and Fujiki, 1985). Subsequent to sucrose density gradient purification of the organelle from homogenate, the specific activities of peroxisomal enzymes may be increased 50- to 100-fold (Tolbert, 1981; Lazarow and Fujiki, 1985).

Enzymatically, peroxisomes can be defined (in very general terms) as respiratory organelles containing one or more flavin oxidases, which catabolize a variety of substrates to produce H_2O_2, and catalase, which catalyzes the decomposition of H_2O_2 into oxygen and water (de Duve and Baudhuin, 1966; Tolbert, 1981; Lazarow and Fujiki, 1985). More specifically, the more than forty enzymes known to be localized to peroxisomes may be grouped into the following categories: (i) catalase; (ii) oxidases; (iii) dehydrogenases; (iv) acyltransferases; (v)

aminotransferases; (vi) the glyoxylate cycle (most often in plants) and (vii) certain enzymes involved in the synthesis of plasmalogens (ether phospholipids), cholesterol and bile acids (Tolbert, 1981; Lazarow and Fujiki, 1985). Common substrates for the aforementioned oxidases include α-hydroxy acids, L- and D-amino acids, lactate, glycolate, urate, fatty acyl-CoAs, polyamines, purines, methanol and ethanol, the peroxisomal oxidation of which yields an appreciable proportion [estimates range between 10% (Mannaerts et al., 1979) and 32% (Kondrup and Lazarow, 1982)] of the total cellular respiration (Masters and Crane, 1985; Lazarow, 1987a). These reactions do not conserve the energy of the oxidations as ATP, however, and are, therefore, thermogenic (Tolbert, 1981; Lazarow and Fujiki, 1985). Since the metabolic pathways in peroxisomes are variable depending on the tissue, substrate availability and stage of development, peroxisomes may contain a large number of different enzymes related to functionality (see section 1.2.2); it is unlikely, however, that peroxisomes from a single tissue would possess the full complement of functional activities ascribed to peroxisomes in general (Tolbert, 1981; Lazarow and Fujiki, 1985). In this respect, mammalian peroxisomes, in contrast with plant peroxisomes, possess a complement of enzymes, more than half of which are related to lipid metabolism (Masters and Crane, 1985). Peroxisomal enzymes, particularly the oxidases, exhibit pH optima of ≈8.5, which

compares with lower pH profiles in other subcellular compartments (Tolbert, 1981). The terminal oxidases, which appear not to be regulated but may be rate-limiting due to their low activity, have as their prosthetic group, either FMN (eg. glycolate oxidase; Cromartie and Walsh, 1975) or FAD (eg. fatty acyl-CoA oxidase; Lazarow, 1978). Except for the peroxisomal integral membrane protein, acyl-CoA synthetase, the catalytic site of which faces the cytosol (Mannaerts et al., 1982), all peroxisomal enzymes investigated thus far are confined to the interior of the organelle. Here they exist either as soluble matrix proteins (most peroxisomal enzymes), core proteins or membrane proteins, the catalytic sites of which face the matrix (Tolbert, 1981; Lazarow and Fujiki, 1985; Mannaerts and Van Veldhoven, 1990). Interestingly, most peroxisomal enzymes have counterparts in other cellular organelles (Tolbert, 1981; Lazarow and Fujiki, 1985; Mannaerts and Van Veldhoven, 1990).

1.2.1.3 Membrane Properties

Peroxisomal membranes have been isolated from purified organelles by osmotic shock (Donaldson et al., 1972; Huang and Beevers, 1973), which yields "ghosts" with apparently sealed membranes and some residual content material (eg. the urate oxidase core in rat liver peroxisomal membranes; Alexson et al., 1985), by treatment

with Na_2CO_3 (pH 11.5) (Fujiki et al., 1982b; Hardeman et al., 1990; Nuttley et al., 1990; Bodnar and Rachubinski, 1991), which yields unsealed (flattened or whorled) membrane sheets stripped of peripheral membrane proteins but retaining phospholipids and integral membrane proteins, and by Triton X-114 phase separation (Hardeman et al., 1990), which preserves the enzymatic activity of most of the rat liver integral membrane proteins. As revealed by electron microscopy, the membrane, which exhibits a trilaminar appearance, ranges from 6.5 to 7.0 nm in thickness for rat liver peroxisomes (Lazarow, 1984) to ≈8.1 nm in thickness for C. tropicalis peroxisomes (Nuttley et al., 1990); this is less than the thickness of the lysosomal membrane, plasma membrane and membranes of most other single membrane-bound cell inclusions (Moser, 1987). Although Fujiki et al. (1982a) reported that both the phospholipid to protein ratio in rat liver peroxisomal membranes and the percentage of membrane protein to total rat liver peroxisomal protein are uniquely low compared to that of other biomembranes (eg. mitochondria and microsomes), subsequent investigations (Appelkvist et al., 1981; Crane and Masters, 1986; Hardeman et al., 1990) have suggested otherwise. Accordingly, estimates for phospholipid to protein ratios and membrane protein to total rat liver peroxisomal protein ratios range from ≈204 nmol/mg (Fujiki et al., 1982a) to 1054 nmol/mg (Hardeman et al., 1990) and from 12% (Fujiki et al., 1982a)

to 26% (Hardeman et al., 1990), respectively. By comparison, peroxisomal membranes from C. tropicalis exhibit a phospholipid to protein ratio of ≈430 nmol/mg and contain between 2.2 and 3.1% of the total peroxisomal protein (Nuttley et al., 1990). With respect to phospholipid content, phosphatidyl choline and phosphatidyl ethanolamine are the principal phospholipids in the peroxisomal membranes of rat liver (Fujiki et al., 1982a) and C. tropicalis (Nuttley et al., 1990) and the glyoxysomal membranes of castor bean (Donaldson and Beevers, 1977). These phospholipids are synthesized in the ER and are presumed to be transferred to the peroxisomal membrane by phospholipid carrier proteins (Lazarow, 1987a).

Based upon the osmotic behaviour of isolated peroxisomes in sucrose solutions and the equilibrium density of these organelles in sucrose gradients, de Duve and Baudhuin (1966) concluded that isolated peroxisomes are permeable to sucrose, an unusual property that is shared by the mitochondrial outer membrane, but not by most other subcellular membranes (Benz, 1985). Certain peroxisomal enzymes, including D-amino acid oxidase and L-α-hydroxy acid oxidase A, do not exhibit latency similarly indicating that their substrates can readily permeate through the membrane of isolated peroxisomes (Beaufay et al., 1964; Baudhuin, 1969). This concept of a freely permeable peroxisomal membrane is also consistent with the absence of cofactor

pools in isolated peroxisomes (Van Broekhoven et al., 1981; Mannaerts et al., 1982). Extending these observations, and paralleling those established for mitochondria, chloroplasts and Gram-negative bacteria (Benz, 1985), Van Veldhoven et al. (1987) suggested that the ≈22 kDa integral membrane protein of rat liver peroxisomes is a non-selective pore-forming porin which creates a channel ≈1.7 nm in diameter (Lemmens et al., 1989), large enough to allow the free diffusion of hydrophilic substrates, products and cofactors with molecular mass of up to 800 Da. Consistent with this, Labarca et al. (1986) and Lemmens et al. (1989) described the presence of a weakly cation-selective, large conductance channel in rat liver peroxisomal membrane fragments fused with planar lipid bilayers or Triton X-100-solubilized integral membrane proteins reconstituted into liposomes, respectively.

The premise of a peroxisomal membrane which is freely permeable to low molecular weight molecules is, however, in direct contrast with results obtained in the yeast Hansenula polymorpha. Based upon ^{31}P-NMR experiments, Nicolay et al. (1987) postulated the H. polymorpha peroxisomes possess an acidic interior. From the chemical shift of the phosphate pool presumed to be present in the peroxisomes of intact, methanol-induced cells, Nicolay et al. (1987) estimated that, in vivo, yeast peroxisomes have an internal pH of between 5.8 and 6.0. In a complementary

study, Douma et al. (1987) presented evidence for the association of an ATPase with the peroxisomal membrane in H. polymorpha, suggesting that this ATPase, the properties of which resemble those of the mitochondrial F_1F_0-ATPase of H. polymorpha, is responsible for the acidification of peroxisomes. Paralleling these investigations, del Valle et al. (1988) and Wolvetang et al. (1990) have reported the presence of an ATPase activity associated with peroxisome-enriched fractions of rat liver. This putative ATPase appears to be immunologically distinct from the mitochondrial F_1F_0-ATPase and biochemically distinct from both mitochondrial and lysosomal ATPases (Wolvetang et al., 1990). At present, the substrate specificity of the peroxisomal membrane-associated ATPase is not known.

1.2.1.4 Inducibility

Although not realized at the time, the work of Lazarow and de Duve (1976), which demonstrated the presence of a fatty acyl-CoA oxidizing system in rat liver peroxisomes, the activity of which could be enhanced approximately one order of magnitude by clofibrate treatment, described what has turned out to be one of the most striking (and, for research purposes, utilized) features of peroxisome biology: the inducibility of this organelle. Currently, it is well established that several structurally dissimilar peroxisome proliferators [including

the hypolipidemic drugs (eg. clofibrate, ciprofibrate, fenofibrate, nafenopin, Wy-14643 and tibric acid) and certain phthalate ester plasticizers (eg. di(2-ethylhexyl)phthalate)] induce marked changes in peroxisomal enzyme activities in the livers of rodents, especially rats and mice (reviewed by Reddy et al., 1982). In general, large doses of these substances cause enhanced proliferation of the entire organelle (and a concomitant increase in associated enzyme activities), in a manner similar to that observed for the proliferation of adrenal cortex peroxisomes upon stimulation by adrenocorticotrophic hormone (Lazarow, 1987a) and the proliferation of brown adipose tissue peroxisomes upon cold adaptation (Nedergaard et al., 1980; Cannon et al., 1982). By contrast, however, peroxisome induction may be selective, influencing only certain enzymes within peroxisomes without appreciable proliferation of the organelle itself. Examples of the latter include: the induction of renal D-amino acid oxidase when germ-free mice are exposed to bacteria (Lazarow, 1987a), the three-fold increase in hepatic peroxisomal β-oxidation enzymes caused by low doses of hypolipidemic drugs (Grasso, 1985; Lake and Gray, 1985) and the increase in peroxisomal β-oxidation caused by diet [eg. starvation or high fat diets (particularly, partially hydrogenated fish oil)] alone (Osmundsen, 1982; Suga et al., 1982; Nilsson et al., 1984; Berge et al., 1988; Flatmark et al., 1988).

Compared to the rodent systems, the proliferation and metabolic function of peroxisomes in certain yeasts is largely prescribed by the growth conditions. Simply changing the composition of the growth medium to include various carbon and energy sources (eg. methanol, fatty acids or n-alkanes, depending on the yeast) will induce peroxisome proliferation in several different species of yeast and cause the introduction of newly synthesized enzymes into these organelles for the catalysis of newly adopted metabolic functions (Osumi et al., 1974; Anthony, 1982; Tanaka et al., 1982; Veenhuis et al., 1983; Veenhuis and Harder, 1987). This phenomenon may be observed in the yeast Hansenula polymorpha during adaptation to growth on methanol; under certain cultivation conditions, as much as 80% of the cytoplasmic volume may eventually be occupied by peroxisomes which contain the key enzymes of methanol metabolism, methanol oxidase and dihydroxyacetone synthase (Douma et al., 1985; Veenhuis and Harder, 1987). Similarly, peroxisomes proliferate strikingly when different strains of yeast from the genus Candida, including C. tropicalis, C. albicans, C. boidinii, C. lipolytica, C. guilliermondii and C. intermedia, are grown on n-alkanes or fatty acids; the induced peroxisomes contain a very active β-oxidation system which enables these yeast to metabolism these compounds (Osumi et al., 1974; Kawamoto et al., 1978; Fukui and Tanaka, 1979; Tanaka et al., 1982; Dommes et al., 1983;

Goodman, 1985). In Saccharomyces cerevisiae, where peroxisomes are difficult to identify under ordinary (ie. unsupplemented) growth conditions (due to their presence in only small numbers; Parish, 1975), it has been reported recently that peroxisomes proliferate to a significant number when cells are cultured on oleic acid, exhibiting a concomitant induction of the activities of the β-oxidation enzymes (Veenhuis et al., 1987; McCammon et al., 1990a). As will be discussed (section 1.4), these means of inducing peroxisome proliferation have been exploited in several studies related to peroxisome biogenesis.

The mechanism surrounding the organelle proliferation, in both rat liver and yeast systems, involves a significant increase (to as much as 30-fold) in the mRNA levels corresponding to several genes encoding peroxisomal proteins; this appears to be attributable to an increased rate of mRNA transcription (Lazarow and de Duve, 1976; Osumi et al., 1985; Rachubinski et al., 1985, 1987; Fujiki et al., 1986; Nuttley et al., 1988; Thangada et al., 1989). Enhanced mRNA transcription culminates in increased rates of synthesis of the corresponding peroxisomal proteins and their eventual intraperoxisomal accumulation to new steady state levels; associated protein degradation rates change only modestly or not at all (Lazarow and Fujiki, 1985; Lazarow, 1987a).

1.2.1.5 Other Properties

Paralleling the situation established for mitochondria (Hay et al., 1984; Tzagoloff and Myers, 1986) and chloroplasts (Cashmore et al., 1985; Smeekens et al., 1990), initial electron microscopic (Osumi, 1976) and physicochemical studies (Osumi et al., 1978) postulated the presence of DNA within the peroxisomes of the yeast, Candida tropicalis. Further investigation in this regard, using highly purified peroxisomes from oleate-grown C. tropicalis, yielded a very small amount of DNA; however, restriction endonuclease mapping revealed that this was the result of contaminating mitochondrial DNA (Kamiryo et al., 1982). Peroxisomal DNA was not detected in intact organelles, nor upon CsCl density gradient centrifugation of lysed organelles, subsequent to staining with ethidium bromide or 4',6-diamidino-2-phenylindole (a fluorescent dye which preferentially binds A/T-rich DNA) (Kamiryo et al., 1982). Similarly, the attempted incorporation of [^{3}H]adenine into peroxisomal DNA yielded negative results, confirming the absence of DNA in C. tropicalis peroxisomes (Kamiryo et al., 1982). Since complementary results have been obtained for putative glyoxysomal DNA in castor bean endosperm (Douglas et al., 1973) and peroxisomal DNA in rat liver (Leighton et al., 1968), it is well established that DNA is not present within these organelles.

Conflicting reports represent the statis of whether

or not peroxisomal proteins are glycosylated. Initial studies with malate synthetase suggested that this glyoxysomal enzyme exists as a glycoprotein in both cucumber and castor beans (Mellor et al., 1978; Riezman et al., 1980); however, it has subsequently been found to lack any sugar (Bergner and Tanner, 1981; Kruse and Kindl, 1983). Similarly, cucumber isocitrate lyase is reportedly both glycosylated (Frevert and Kindl, 1978) and non-glycosylated (Riezman et al., 1980). Certainly, catalase from rat liver (Sugita et al., 1982), S. cerevisiae (Ammerer et al., 1981) and cucumber (Becker et al., 1982), carnitine acetyltransferase from C. tropicalis (Ueda et al., 1984) and malate dehydrogenase from cucumber (Becker et al., 1982) are not glycosylated. Furthering potential controversy, Volkl and Lazarow (1982) failed to detect any glycoproteins among the major rat liver peroxisomal proteins, based on their potential binding to lectin columns, while Harson et al. (1983) and Lord and Roberts (1983) postulated the existence of glyoxysomal glycoproteins based on the labeling of unidentified proteins bands with radioactive sugars. As noted by Lazarow and Fujiki (1985), however, the extreme sensitivity of the radiolabeling method may have permitted the detection of trace contaminant glycoproteins from other cell fractions. Taken collectively, this amounts to only very limited evidence for the existence of glycoproteins within peroxisomes/glyoxysomes and as a consequence, the

glycosylation of peroxisomal proteins is thought not to occur.

1.3 FUNCTIONS OF PEROXISOMES

Subsequent to its discovery, the peroxisome has received much attention due to its functions related to lipid metabolism, including fatty acid β-oxidation, plasmalogen and cholesterol biosynthesis, bile acid formation and α-oxidation. Equally significant, however, are the roles this organelle plays in amino acid and polyamine catabolism, in the degradation of purines and in several of the reactions comprising the pentose phosphate pathway and the glyoxylate cycle. Since the enzyme complement, and thus the functionality, of peroxisomes varies with cell type, substrate availability and stage of development, and because this heterogeneity exists not only between species but between different tissues within the same species (see section 1.2.1.2), the following discussion attempts to establish a general understanding of peroxisomal function(s), while reinforcing the species and tissue specificity of each function.

1.3.1 Lipid Metabolism

1.3.1.1 Fatty Acid β-Oxidation

Almost twenty-five years ago, Breidenbach and Beevers (1967) reported that germinating fatty seedlings

house an organelle (the glyoxysome) which contains all of the enzymes of the glyoxylate cycle; glyoxysomes were later found to catalyze the β-oxidation of stored fatty acids to acetyl-CoA, which is utilized for the synthesis of glucose via the glyoxylate pathway (Cooper and Beevers, 1969). Several years later, a β-oxidation pathway for fatty acyl-CoA was similarly localized to the peroxisomes of Euglena (Graves and Becker, 1974) and Tetrahymena (Blum, 1973; Hryb and Hogg, 1976). In 1976, Lazarow and de Duve described, for the first time, the presence of a cyanide-insensitive, fatty acid β-oxidation system in rat liver peroxisomes, the activity of which could be markedly induced by the administration of the hypolipidemic drug, clofibrate (Lazarow and de Duve, 1976; Lazarow, 1978). Thereafter, a similar cyanide-insensitive, fatty acid β-oxidation system was discovered in the yeast Candida tropicalis, the activity of which was significantly increased upon growth of this yeast in an n-alkane-containing medium (Kawamoto et al., 1978). Since these initial observations, the peroxisomal β-oxidation of fatty acids has been established as a nearly ubiquitous process among eukaryotes, including humans (Bronfman et al., 1979), and one of the major functions of this organelle.

Similar to mitochondria, peroxisomes in many diverse cell types are endowed with an active system for the β-oxidation of fatty acids and possess the auxiliary enzymes

required for the oxidation of unsaturated fatty acids (Tolbert, 1981; Lazarow and Fujiki, 1985; Hovik and Osmundsen, 1987; Osmundsen and Hovik, 1988). The coexistence of mitochondrial and peroxisomal β-oxidation in higher animals is, however, representative of two functionally distinct systems for the catabolism of fatty acids which differ with respect to location, substrate specificity, enzymes, end products and control mechanisms [reviewed by Tolbert (1981)]. In general, isolated mitochondria oxidize short-, medium-, and long-chain fatty acids, medium-chain dicarboxylic fatty acids and the carboxy side chain of prostaglandins (Mannaerts and Van Veldhoven, 1990). By contrast, isolated peroxisomes oxidize a wider spectrum of fatty acids and related derivatives including medium-, long- and very long-chain fatty acids (Hashimoto, 1987; Singh et al., 1987; Singh and Poulos, 1988), medium- and long-chain dicarboxylic fatty acids (Kolvraa and Gregerson, 1986; Suzuki et al., 1989), the carboxy side chains of prostaglandins (Diczfalusy and Alexson, 1988, 1990; Schepers et al., 1988) and xenobiotics (Yamada et al., 1987) and the side chain of the bile acid intermediates, di- and tri-hydroxycoprostanic acid (Pederson et al., 1987). Isolated peroxisomes do not exhibit the ability to oxidize short chain fatty acids; isolated mitochondria, on the other hand, do not oxidize bile acid intermediates, and very long-chain fatty acids and the side-chains of xenobiotics serve

as poor substrates for mitochondrial β-oxidation (Tolbert, 1981; Mannaerts and Van Veldhoven, 1990). In vivo, peroxisomes are responsible for the shortening of very-long-chain fatty acids (Singh et al., 1984); shortened fatty acids are subsequently transported to mitochondria via a carnitine shuttle and further oxidized, along with other short-, medium- and long-chain fatty acids, thereby contributing to the major source of metabolic fuel (Masters and Crane, 1984). Cerdan et al. (1988) have also provided strong evidence supporting a predominant role for the peroxisomal β-oxidation system in the overall oxidation of dicarboxylic acids in vivo. The contribution of peroxisomes to the β-oxidation of the side chains of prostaglandins, other eicosanoids and xenobiotics, in vivo, is less well understood.

Prerequisite to entering the β-oxidation pathway is the activation of the fatty acid (Figure 1.3.1). Similar to the ER (Kornberg and Pricer, 1953) and the mitochondrial outer membrane (Garland et al., 1970), the peroxisomal membrane possesses a long-chain (ie. 12 to 22 carbons in length) ATP-dependent acyl-CoA synthetase for this purpose (Mannaerts et al., 1982). Further, there is evidence to suggest that peroxisomes contain an additional synthetase which is active toward very-long-chain (ie. greater than 22 carbons in length) fatty acids (Singh et al., 1987; Singh and Poulos, 1988; Wanders et al., 1988).

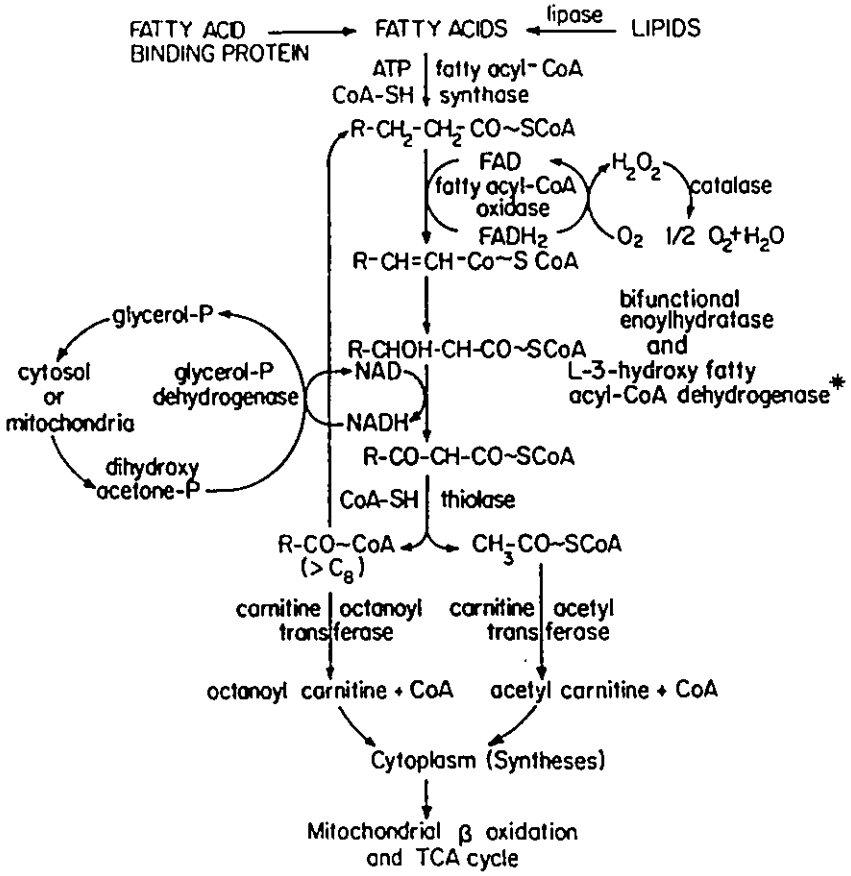


Figure 1.3.1. The fatty acid β-oxidation pathway of liver peroxisomes. The asterisk denotes that in peroxisomes of fungal systems, bifunctional 2-enoyl-CoA hydratase/L-3-hydroxyacyl-CoA dehydrogenase is replaced by a trifunctional enzyme possessing a third Δ^3, Δ^2-enoyl CoA isomerase activity (see section 1.3.1.1). The figure was reproduced from Tolbert (1981).

Mitochondria appear to be deficient in this enzyme activity, which may explain why very-long-chain fatty acids are oxidized preferentially by peroxisomes (Singh et al., 1987; Singh and Poulos, 1988). After activation by an acyl-CoA synthetase, fatty acids must traverse the peroxisomal membrane to encounter the β-oxidation pathway. Initially, fatty acid transport into peroxisomes was thought to be mediated by fatty acid-binding protein (Appelkvist and Dallner, 1980). It has recently been established, however, that fatty acid-binding protein is not involved in the transport of acyl-CoA through the peroxisomal membrane, but may transfer fatty acids to peroxisomes (Reubsaet et al., 1990).

In animal mitochondria, the dehydrogenation of saturated acyl-CoA esters is catalyzed by flavoprotein acyl-CoA dehydrogenases and the electrons resulting from this reduction are accepted by the mitochondrial respiratory chain; the process is mediated by the electron transfer flavoprotein and the electron transfer flavoprotein-ubiquinone-oxidoreductase (Frerman, 1988). This oxidation of long-chain fatty acids by mitochondria is carnitine-dependent, involves a large number of individual enzymes and proceeds, via a repetitive cycle, toward the complete degradation of fatty acids into acetyl units (Saggerson and Zammit, 1986). Conversely, peroxisomal β-oxidation does not shuttle electrons directly into a respiratory chain that is

coupled to a phosphorylation system, is not carnitine dependent [although short- and medium-chain carnitine acyltransferases are present in rat liver peroxisomes (Miyazawa et al., 1983; Bronfman and Leighton, 1984)] and does not degrade fatty acids completely (Tolbert, 1981; Lazarow and Fujiki, 1985). Rather, peroxisomal β-oxidation functions as a fatty acyl chain-shortening system, thereby providing substrates for mitochondrial β-oxidation (Osmundsen et al., 1987). In addition, the enzymes of peroxisomal β-oxidation differ from those in mitochondria, with multifunctional enzymes being characteristic components of peroxisomal (and bacterial; Yang and Schulz, 1983) β-oxidation and structurally independent and monofunctional enzymes being characteristic of mitochondrial β-oxidation (Tolbert, 1981; Hashimoto, 1987; Kunau et al., 1987). The primary dehydrogenation step in peroxisomal (ie. rat liver and yeast peroxisomal and seed glyoxysomal) β-oxidation, which is rate-limiting, is catalyzed by a flavoprotein (FAD)-linked acyl-CoA oxidase (not a dehydrogenase). Accordingly, acyl-CoA oxidase transfers the electrons resulting from the reduction to O_2 to form H_2O_2, which is, in turn, decomposed by catalase (Figure 1.3.1; Hashimoto, 1982; Tanaka et al., 1982). The aforementioned enzymes (ie. acyl-CoA oxidase versus acyl-CoA dehydrogenase), in addition to possessing entirely different molecular structures, catalyze different reactions. Furthermore, the absence of

the respiratory chain from peroxisomes suggests that even their mode of electron transfer is different (Frerman, 1988; Kunau et al., 1988). In peroxisomes from rat liver (Hashimoto, 1982) and human liver (Reddy et al., 1987), the next two reactions of the β-oxidation pathway, in contrast with mitochondrial β-oxidation, are catalyzed by a bifunctional protein comprised of 2-enoyl-CoA hydratase and L-3-hydroxyacyl-CoA dehydrogenase activities (Figure 1.3.1) [It should be noted that a third functionality, a Δ^3, Δ^2-enoyl-CoA isomerase activity, has recently been ascribed to the rat liver enzyme (Palosaari and Hiltunen, 1990).]. Alternatively, for the five fungal systems investigated thus far, including Candida tropicalis (Tanaka et al., 1982; Moreno de la Garza et al., 1985), Candida lipolytica (Tanaka et al., 1982), Neurospora crassa (Kunau et al., 1987), Aspergillus tamarii (Kunua et al., 1988) and Saccharomyces cerevisiae (Kunau et al., 1988), and for the glyoxysomes of cucumber seeds (Behrends et al., 1988), the bifunctional enzyme of peroxisomal β-oxidation is replaced by a trifunctional enzyme possessing an additional 3-hydroxyacyl-CoA epimerase activity.

The fourth enzyme of the peroxisomal β-oxidation pathway, 3-ketoacyl-CoA thiolase (Figure 1.3.1), also appears to be characteristically different from its counterpart in the mitochondrial β-oxidation cycle. In mammalian cells, two distinct types of thiolases are known

to be present in mitochondria (Middleton, 1973; Staack et al., 1978). One of these, acetoacetyl-CoA thiolase, is specific to acetoacetyl-CoA and functions in ketone body metabolism (Middleton, 1973). The other, 3-ketoacyl-CoA thiolase, possesses a tetrameric structure (Miyazawa et al., 1981), has a broad substrate specificity for chain length and catalyzes the last step of the fatty acid β-oxidation cycle (Staack et al., 1978). Alternatively, peroxisomal 3-ketoacyl-CoA thiolases from several different sources, while exhibiting a broad substrate specificity (Miyazawa et al., 1980, 1981), exist as homodimers in their native state and are not immunologically related to their mitochondrial counterpart (Miyazawa et al., 1981; Kunau et al., 1988). Furthermore, in the n-alkane utilizing yeast, Candida tropicalis, for which a mitochondrial β-oxidation system has not been detected (Ueda et al., 1985), two different thiolases are localized to peroxisomes (Kurihara et al., 1988, 1989). The acetoacetyl-CoA-specific thiolase (Thiolase I) is composed of six identical subunits which are immunologically and proteolytically distinguishable from the peptide components of the 3-ketoacyl-CoA thiolase (Thiolase III), which exists as a homodimer (Kurihara et al., 1989). Since kinetic studies have shown that substrates possessing long chains are degraded exclusively by Thiolase III, it has been suggested that in C. tropicalis the β-oxidation of fatty acids is carried out completely and efficiently by

peroxisomes (Kurihara et al., 1989). Interestingly, enzymes similar to the 3-ketoacyl-CoA thiolase have been localized to the peroxisomes of other eukaryotic cells (Tolbert, 1981; Lazarow and Fujiki, 1985), however, the presence of a peroxisomal acetoacetyl-CoA-specific thiolase in any other organism has not been documented as yet.

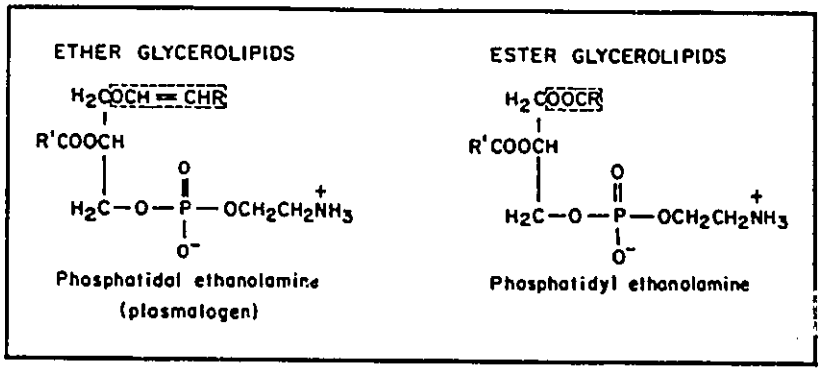
Enzymatically, peroxisomal 3-ketoacyl-CoA thiolase catalyzes the degradation of long-chain acyl-CoAs into acetyl-CoA and an acyl-CoA possessing a chain length of two less carbons than the starting molecule (Hashimoto, 1987). Long-chain ($>C_8$) acyl-CoAs can then re-enter the peroxisomal β-oxidation pathway whereas acetyl-CoA and medium-chain ($\leq C_8$) acyl-CoAs serve as substrates for the peroxisomal matrix enzymes, carnitine acetyltransferase (Bieber et al., 1982; Ueda et al., 1984) and carnitine octanoyltransferase (Farrell et al., 1984; Chatterjee et al., 1988), respectively (Figure 1.3.1). The carnitine esters which result are permeable to the mitochondrial membrane (acetyl-CoA and medium-chain acyl-CoAs are not) and accordingly, are further degraded by the mitochondrial β-oxidation pathway (Tolbert, 1981; Farrell et al., 1984; Bieber and Fiol, 1986; Zammit, 1986).

1.3.1.2 Plasmalogen Biosynthesis

Of the known anabolic functions ascribed to mammalian peroxisomes, one of particular significance is

their involvement in the initial reactions of plasmalogen biosynthesis. Plasmalogens are "specialized" ether glycerolipids that contain a 1,2-unsaturated, long-chain alcohol in vinyl ether linkage with the C-1 position of the glycerol backbone of a phospholipid; this replaces the ester-linked fatty acid found in conventional phospholipids [Figure 1.3.2 (A); Lazarow, 1987a; Moser, 1988]. The involvement of peroxisomes in plasmalogen biosynthesis was first described by Hajra et al. (1979) and Hajra and Bishop (1982) who demonstrated that in rodent liver the first two enzymes of the plasmalogen biosynthetic pathway, acyl-CoA:dihydroxyacetonephosphate (DHAP) acyltransferase and alkyl-DHAP synthase, were localized to peroxisomes. Subsequently, Heymans et al. (1983, 1984) discovered that tissues derived from patients afflicted with Zellweger (also called cerebro-hepato-renal) syndrome [an autosomal recessive disorder, considered to be the prototype of human peroxisomal disorders, which is characterized by an absence of functional peroxisomes (Talwar and Swaiman, 1987; Zellweger, 1987)], were severely deficient in plasmalogens. Extending these observations, Schrakamp et al. (1985a) demonstrated that Zellweger fibroblasts were impaired in the <u>de novo</u> synthesis of ether glycerolipids at a position along the biosynthetic pathway at or before the introduction of the glycero-ether bond. It was further determined that both peroxisomal enzymes involved in this process, acyl-CoA:DHAP

A



B

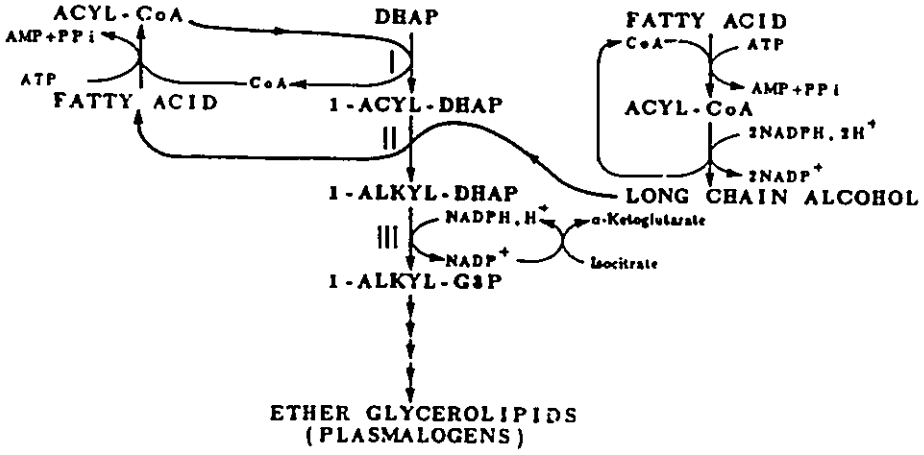


Figure 1.3.2. Plasmalogen biosynthesis. A, Comparison of the structure of plasmalogens ("specialized" ether-linked glycerolipids which contain a double bond between the first two carbons of the long-chain alcohol) and their more common ester-linked counterparts. B, The three steps of ether glycerolipid synthesis that occur in liver peroxisomes and the corresponding enzymes involved: I, acyl-CoA:DHAP acyltransferase; II, alkyl-DHAP synthase; III, acyl-DHAP:NADPH oxidoreductase. The figures were reproduced from Moser (1988).

acyltransferase (Schutgens et al., 1984) and alkyl-DHAP synthase (Schrakamp et al., 1985b) [Figure 1.3.2 (B)] were deficient in fibroblasts and tissues derived from Zellweger patients. Similar observations have since been extended to cultured skin fibroblasts from other human peroxisomal disorders including rhizomelic chondrodysplasia punctata, neonatal adrenoleukodystrophy and infantile Refsum disease (Schrakamp et al., 1988).

Within the cell, glycerolipid biosynthesis is initiated by the acylation of either glycerol-3-phosphate, which occurs in the ER and accounts for the majority of glycerolipid synthesis (ie. ester glycerolipid synthesis), or DHAP, which occurs in peroxisomes and comprises the initial reaction in the biosynthesis of the less abundant ether glycerolipids (Moser, 1988; Mannaerts and Van Veldhoven, 1990). Peroxisomes acylate exclusively DHAP, in a reaction catalyzed by acyl-CoA:DHAP acyltransferase (the first of three steps in ether glycero-lipid synthesis that are localized to peroxisomes), and subsequently convert 1-acyl-DHAP to 1-alkyl-DHAP, by replacing the esterified fatty acid with a fatty alcohol to form an ether bond, in a reaction catalyzed by alkyl-DHAP synthase; 1-alkyl-DHAP is then converted, within the peroxisome, to 1-alkyl-3-glycerophosphate, in a reaction catalyzed by acyl-DHAP:NADPH oxidoreductase [Figure 1.3.2 (B); Moser, 1988; Mannaerts and Van Veldhoven, 1990]. The remaining reactions required for

the complete synthesis of ether glycerolipids (plasmalogens) occur in the ER (Declercq et al., 1984). As yet, the function(s) of plasmalogens, which constitute 5 to 20% of the phospholipids present in most mammalian cell membranes, 15 to 20% of the total phospholipids in the brain and as much as 80 to 90% of the ethanolamine phospholipid class in white matter and myelin (Lazarow, 1987a; Moser, 1988), are not well characterized. It is known, however, that ether glycerolipids serve as precursors of platelet-activating factor (Mannaerts and Van Veldhoven, 1990) and plasmalogens have been purported to play a role in the protection of cell membranes by scavenging oxygen radicals (Zoeller et al., 1988).

1.3.1.3 Cholesterol and Dolichol Biosynthesis

The mevalonate pathway (Figure 1.3.3) consists of a series of enzymes which transform acetate into farnesyl pyrophosphate, the common intermediate in the biosynthesis of cholesterol, ubiquinone and dolichol (Hemming, 1983). The initial part of the pathway, which is localized to the cytoplasm, involves the three-step conversion of acetyl-CoA to 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA). Subsequently, in a reaction considered to be the rate-limiting step in the biosynthesis of polyisoprenoids, HMG-CoA is reduced to mevalonate by HMG-CoA reductase, a transmembrane glycoprotein thought previously (see below) to

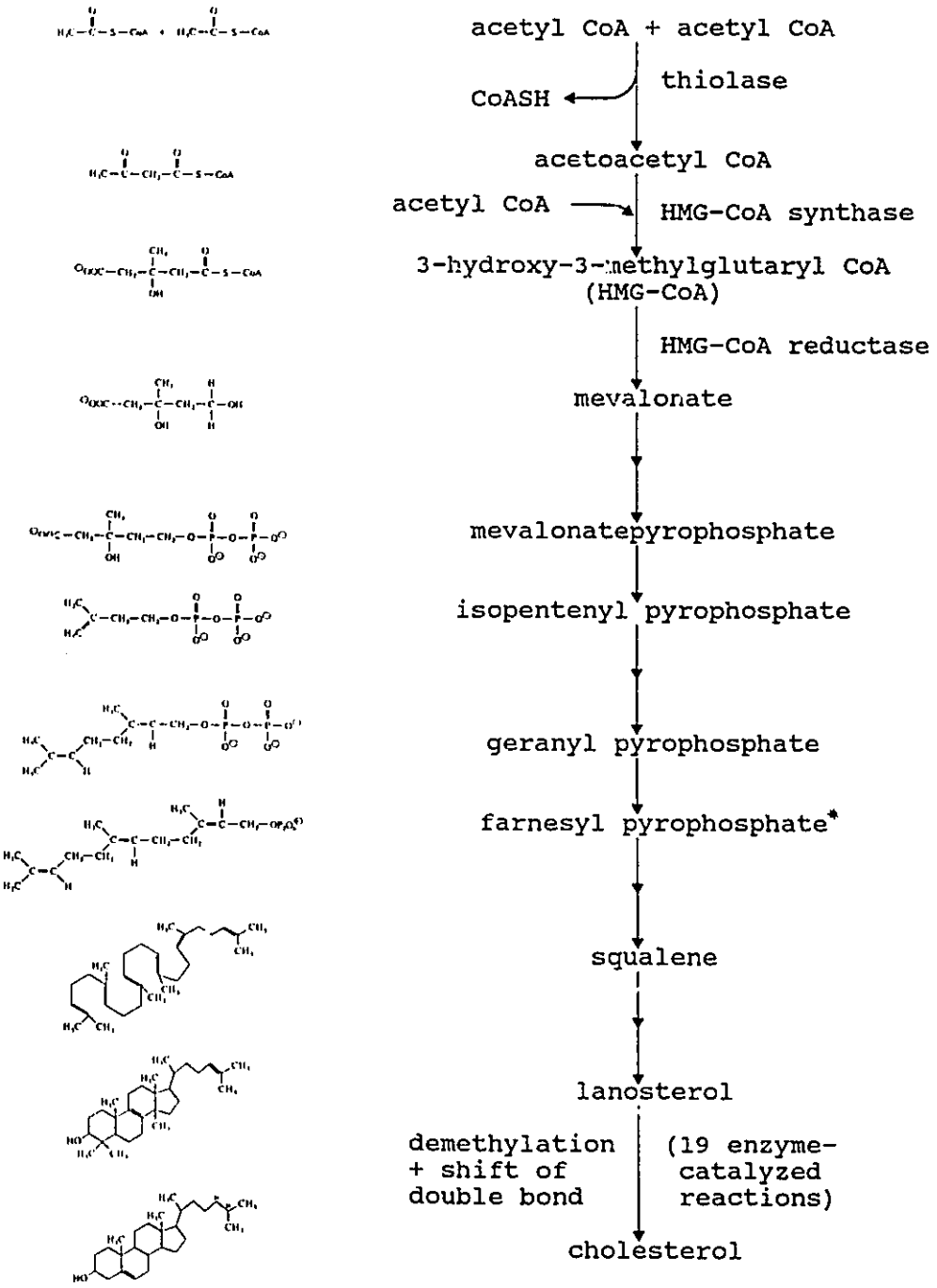


Figure 1.3.3. Cholesterol biosynthesis in mammals via the mevalonate pathway. The asterisk signifies that farnesyl pyrophosphate is a branch point in the biosynthesis of cholesterol, ubiquinone and dolichol (see section 1.3.1.3). The figure is adapted from Rawn (1989).

be exclusively microsomal (Liscum et al., 1983). Following this, several distinct integral membrane-bound polypeptides of the ER catalyze a series of phosphorylation reactions that activate and decarboxylate mevalonate to yield isopentenyl pyrophosphate prior to the sequential synthesis of geranyl pyrophosphate, farnesyl pyrophosphate and squalene, which then cyclizes to form lanosterol (Gaylor, 1981). Nineteen discrete reactions localized to the ER constitute the final stages of sterologenesis in which the removal of three methyl groups from lanosterol and the migration and reduction of double bonds yield cholesterol (Thompson et al., 1987). Alternatively, as a branch point from farnesyl pyrophosphate, the isoprenoid chain may be formed by the sequential addition of five-carbon cis-isoprene units and the synthesis of dolichol completed by a condensation reaction resulting in the free alcohol and the subsequent saturation of the α-isoprene unit (Ekström et al., 1987).

As alluded to above, HMG-CoA reductase, considered at one time to be exclusively microsomal (Liscum et al., 1983), has been localized to the peroxisomal matrix in normal rat liver cells (Keller et al., 1985, 1986). In an effort to determine a role for peroxisomal HMG-CoA reductase, Appelkvist (1987) and Thompson et al. (1987) demonstrated that highly-purified rat liver peroxisomes, supplemented with a cytosolic fraction, convert mevalonic

acid to cholesterol *in vitro*. This implies that sterol biosynthesis does occur in this organelle, however, peroxisomes, like microsomes, apparently possess only those enzymes of cholesterol biosynthesis subsequent to the steps involving farnesyl pyrophosphate (Appelkvist, 1987; Thompson et al., 1987). Furthering this notion, Appelkvist et al. (1990) have recently localized several enzymes related to cholesterol biosynthesis, including dihydrolanosterol oxidase, steroid-14-reductase, steroid-8-isomerase and steroid-3-ketoreductase, to the membrane fraction of isolated rat liver peroxisomes. With reference to the related synthesis of dolichol, Appelkvist and Dallner (1987) and Appelkvist and Kalén (1989) have also shown that dolichol biosynthesis occurs both in peroxisomes (from farnesyl pyrophosphate onwards) and the ER, but the biosynthetic processes at these two locations have different properties.

At present, the physiological importance of cholesterol biosynthesis within peroxisomes to cholesterol homeostasis within the cell is not well understood. Although the contribution of peroxisomes is small relative that of the ER (Mannaerts and Van Veldhoven, 1990), the aforementioned observations, in conjunction with the localization of both sterol carrier protein 2 (SCP-2), which is known to stimulate the conversion of lanosterol to cholesterol (Noland et al., 1980) and activate several other

reactions related to cholesterol metabolism (Thompson et al., 1987), and a thiolase activity, which is responsible for the initial reaction in cholesterol biosynthesis (ie. the condensation of acetyl-CoA into acetoacetyl-CoA), to rat liver peroxisomes (Van der Krift et al., 1985; Keller et al., 1989; and Thompson and Krisans, 1990, respectively), suggest that peroxisomes may play an important role in overall cholesterol metabolism. Perhaps, the peroxisomal synthesis of cholesterol represents a reserve capacity for the cell under stress-related or pathological conditions. Alternatively, cholesterol may be required for the synthesis of peroxisomal membranes or peroxisomes may supply other membranes (eg. the plasma membrane) with this lipid (Appelkvist et al., 1990).

1.3.1.4 Bile Acid Formation

Bile acid formation (predominantly cholic and chenodeoxycholic acid formation; Figure 1.3.4) constitutes the major pathway of cholesterol catabolism in mammals, allowing cholesterol to be excreted from the body in the form of water-soluble bile acid conjugates (Pederson et al., 1987). Mechanistically, the conversion of cholesterol to bile acids involves modification of both the ring and the side chain moieties by enzymes assumed previously (see below) to be localized only to the ER, cytosol and mitochondria (Pederson et al., 1987). To this end, two possible pathways for bile

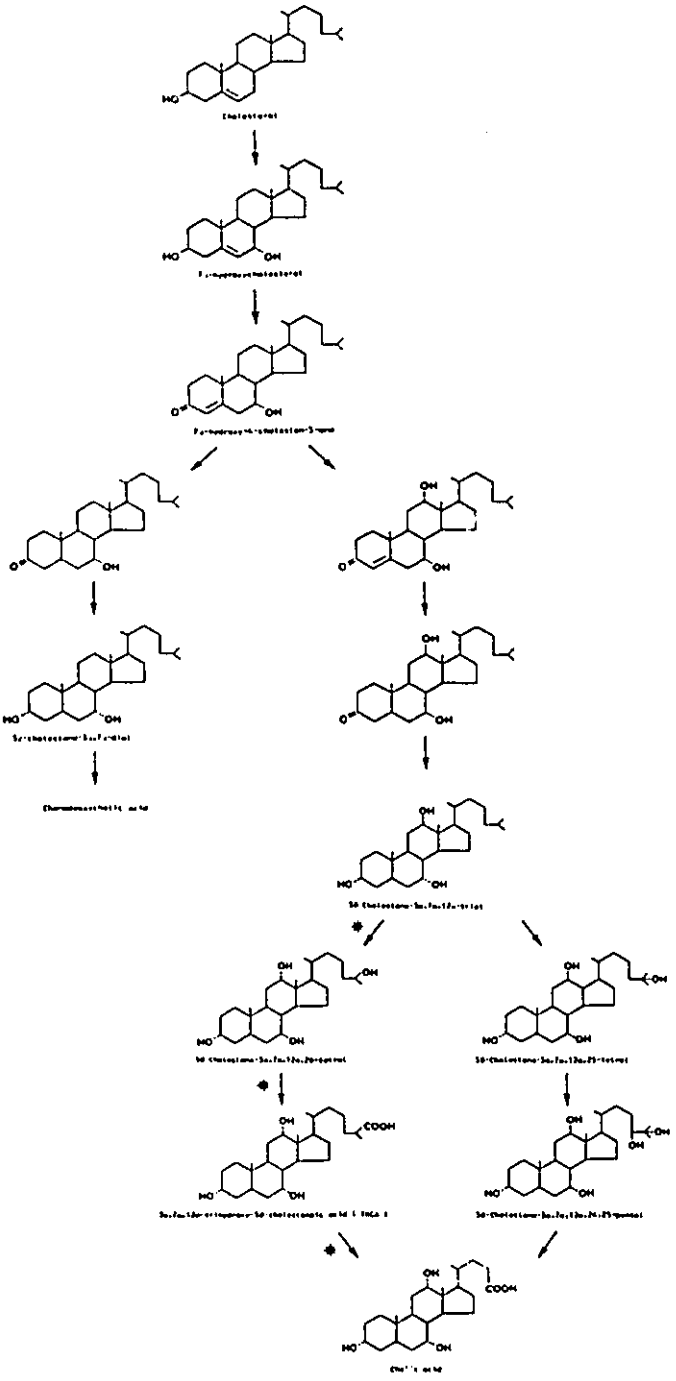


Figure 1.3.4. Reaction sequence of the major pathway leading to the formation of cholic acid and chenodeoxycholic acid from cholesterol. Asterisks denote the predominant pathway for the formation of cholic acid from 5β-cholestane-3α,7α,12α-triol (see section 1.3.1.4). The figure was adapted from Björkhem et al. (1985).

acid formation have been established. The minor pathway necessitates that the C-26 hydroxylation and side chain oxidation of cholesterol, by mitochondria, precede the steroid hydroxylations (Björkhem et al., 1985; Pederson et al., 1987). Conversely, the major, ER-initiated, pathway involves ER and cytosolic enzyme-catalyzed alterations to the steroid nucleus, followed sequentially by the hydroxylation of C-26 (by either ER or mitochondrial hydroxylases) and side chain oxidation in mitochondria to produce a C-24 bile acid (Björkhem et al., 1985; Pederson et al., 1987). Along this pathway, 5β-cholestane-3α,7α,12α-triol and 5β-cholestane-3α,7α-diol are thought to be important intermediates in the synthesis of cholic and chenodeoxycholic acid, respectively (Björkhem et al., 1985; Figure 1.3.4). 5β-Cholestane-3α,7α,12α-triol can be oxidized to yield the C-24 steroid side chain by one of two mechanisms in man (Figure 1.3.4, lower portion). The predominant pathway (Figure 1.3.4, asterisks) involves the attack of one of two terminal methyls by a mitochondrial 26-hydroxylase to yield 5β-cholestane-3α,7α,12α,26-tetrol, which is further oxidized to 3α,7α,12α-trihydroxy-5β-cholestanoic acid (THCA). Propionic acid is subsequently removed from the activated CoA ester of THCA to form the CoA ester of cholic acid; this is then coupled to glycine or taurine to yield the corresponding conjugates to be excreted (Björkhem et al., 1985; Pederson et al., 1987). The

alternative pathway involves sequential 25- and 24-hydroxylation of 5β-cholestane-3α,7α,12α-triol by microsomal hydroxylases to yield 5β-cholestane-3α,7α,12α,24,25-pentol. After oxidation of the 24-hydroxyl group, acetone is removed to yield cholic acid (Björkhem et al., 1985; Pederson et al., 1987).

The initial functional link between peroxisomes and bile acid formation was provided by Pederson and Gustafsson (1980) who demonstrated that enriched peroxisomal fractions from rat liver could efficiently catalyze the conversion of THCA into cholic acid. Consistent with this, Hagey and Krisans (1982) showed that rat liver peroxisomal fractions (70 to 75% pure) oxidized cholesterol to an unidentified C-24 bile acid and propionic acid, the expected 3-carbon fragment produced by either the major or minor pathways for bile acid formation. The peroxisomal conversion of THCA to cholic acid (ie. the major pathway to bile acid formation) was subsequently shown to be dependent on the presence of ATP, CoA, NAD^+ and Mg^{2+} in the incubation medium and thought to involve the intermediary formation of 3α,7α,12α,24-tetrahydroxy-5β-cholestanoic acid (Pederson and Gustafsson, 1980; Kase et al., 1983). It was further determined, using ^2H_2O and $^{18}O_2$, that the introduction of the 24-hydroxyl group in 3α,7α,12α,24-tetrahydroxy-5β-cholestanoic acid was the combined result of both a desaturase and hydratase activity, suggesting a reaction mechanism analogous to that

of the peroxisomal fatty acid β-oxidation system (Björkhem et al., 1984, 1985). Rat liver peroxisomes have since been shown to possess 7α-hydroxylase, which catalyzes the first step in the major pathway to bile acid formation (Appelkvist et al., 1990) and at least one other hydroxylase activity capable of converting 3α,7α,12α-trihydroxy-5β-cholestane (the major substrate for 26-hydroxylation) to cholestanetetrol (Thompson and Krisans, 1985).

The minor pathway to bile acid formation (Figure 1.3.5), which involves C-26 hydroxylation and side chain modification of cholesterol prior to steroid hydroxylations (see above), also appears to be localized to rat liver peroxisomes since they exhibit the capacity to metabolize 26-hydroxycholesterol (the product of the 26-hydroxylation of cholesterol) to 3β-hydroxy-5-cholenoic acid (a C-24 bile acid) (Krisans et al., 1985). Concomitant with the synthesis of cholic acid, Prydz et al. (1986) have determined that rat liver peroxisomes similarly catalyze the side chain cleavage of 3α,7α-dihydroxy-5β-cholestanoic acid (DHCA) to yield chenodeoxycholic acid. Further, the recent localization of bile acid-CoA:amino acid N-acetyltransferase to rat liver peroxisomes (Kase and Björkhem, 1989) has necessitated that the conjugation of newly formed CoA esters of cholic and chenodeoxycholic acid to taurine or glycine be added to the growing list of peroxisomal functions related to bile acid synthesis.

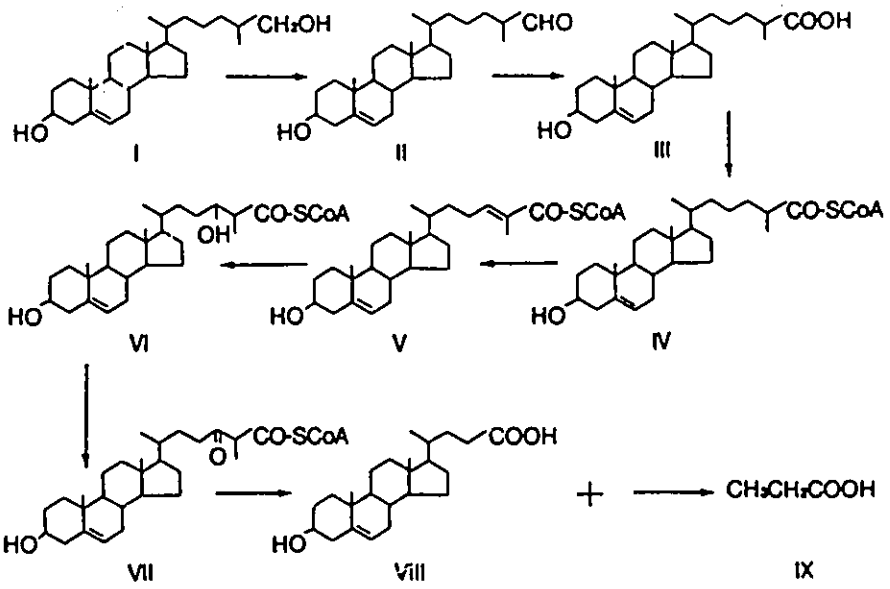


Figure 1.3.5. Reaction sequence of the minor peroxisomal pathway for the oxidation of 26-hydroxycholesterol to 3β-hydroxy-5-cholenoic acid (a C-24 bile acid). Reaction products include: I, 26-hydroxycholesterol; II, 3β-hydroxy-5-cholestenal; III, 3β-hydroxy-5-cholestanoic acid; IV, 3β-hydroxy-5-cholestenoyl coenzyme A; V, 3β-hydroxy-24-ene-5-cholestanoyl coenzyme A; VI, 3β-24-dihydroxy-5-cholestanoyl coenzyme A; VII, 3β-hydroxy-24-keto-5-cholestanoyl coenzyme A; VIII, 3β-hydroxy-5-cholenoic acid; IX, propionic acid. The figure was reproduced from Krisans et al. (1985).

Even though liver peroxisomes from rat and man exhibit the highest capacity of all cell fractions to convert THCA and DHCA into bile acids *in vitro* (Pederson et al., 1987), it has been difficult to evaluate their relative role *in vivo*. Björkhem et al. (1985) and Pederson et al. (1987) have, however, provided evidence that THCA, and various polar metabolites of THCA, accumulate in patients afflicted with Zellweger syndrome, implying that the 26-hydroxylase pathway (in which THCA is an intermediate), not the 25-hydroxylase pathway, is the most important for the *in vivo* formation of cholic acid in man. Despite this, the precise physiological contribution of peroxisomes to the total cellular synthesis of bile acids remains to be established.

1.3.1.5 *α-Oxidation*

Phytanic acid (3,7,11,15-tetramethylhexadecanoic acid) is a highly branched fatty acid derived from dietary sources (phytanic acid is formed in the organism from phytol present, in esterified form, in chlorophyll) which possesses a methyl substitution at the β-position, thereby preventing degradation via β-oxidation (Poulos et al., 1984). Consequently, the mechanism for the catabolism of this fatty acid involves oxidative decarboxylation (ie. α-oxidation) to yield the 19-carbon lower homologue, pristanic acid (2,6,10,14-tetramethylpentadecanoic acid); pristanic acid

can then be further degraded by β-oxidation (Mannaerts and Van Veldhoven, 1990). Since a typical clinical observation in several human peroxisomal disorders, including Refsum's disease (Talwar and Swaiman, 1987), Zellweger syndrome (Poulos et al., 1984), infantile phytanic acid storage disease (Scotto et al., 1982), neonatal adrenoleukodystrophy (Stokke et al., 1984) and rhizomelic chondrodysplasia punctata (Heymans et al., 1985), is the accumulation of phytanic acid in serum and tissues, it has been hypothesized that phytanic acid oxidase activity, and thus the process of α-oxidation, is localized to peroxisomes in humans (Poulos et al., 1984; Mannaerts and Van Veldhoven, 1990). However, Watkins et al. (1990) suggested that the α-oxidation of [1-^{14}C]phytanic acid is a mitochondrial, rather than peroxisomal, process in primates (specifically, human and monkey liver and human fibroblasts) due to the inhibitory effect of two mitochondrial electron transport chain poisons, antimycin A and rotenone, on $^{14}CO_2$ production. Consistent with this, Skjeldal and Stokke (1987) demonstrated that the oxidation of phytanic acid in rat liver occurs in mitochondria, not peroxisomes. Whether the α-oxidation of phytanic acid is a peroxisomal process in certain species, and a mitochondrial process in others [thereby mimicking the situation established for the oxidation of L-pipecolic acid, which is localized to peroxisomes in human (Wanders et al., 1989) and monkey liver

(Mihalik and Rhead, 1989) and localized to mitochondria in rabbit liver (Mihalik and Rhead, 1989)], remains to be determined.

1.3.2 Amino Acid and Polyamine Catabolism

As is detailed below, several observations lend credence to the premise that peroxisomes function in amino acid catabolism. D-amino acid oxidase, which catalyzes the oxidative deamination of the D-isomers of neutral and basic amino acids, and D-aspartate oxidase, which catalyzes the same reaction for the D-isomers of acidic amino acids, are both localized to peroxisomes (de Duve and Baudhuin, 1966; Zaar et al., 1989). Though these enzymes are found in various mammalian tissues (most abundantly in the liver and proximal tubules of the kidney), in direct correlation with the abundance of peroxisomes (Zaar et al., 1989), their physiological significance remains uncertain. It has been suggested that D-amino acid oxidase and D-aspartate oxidase regulate the turnover and metabolism of D-amino acids of microbial origin; alternatively, considering their abundant level in the brain, these enzymes may function as effectors of nervous tissue metabolism [possibly exerting a protective function by degrading extrinsic D-isomers of dicarboxylic amino acids (Hamilton et al., 1979; Zaar et al., 1989)].

In mammals, the degradation of L-lysine to α-aminoadipate occurs via two, initially distinct, pathways

(White et al., 1978). The saccharopine pathway [Figure 1.3.6 (A)] predominates, except in the brain where the activity of this pathway is negligible (Hutzler and Dancis, 1975). Within brain tissue, L-lysine catabolism proceeds via the L-pipecolate pathway [Figure 1.3.6 (B)], a second mechanism in which L-pipecolate, not saccharopine, is the intermediate (Chang, 1978). Furthering the role of peroxisomes in amino acid catabolism, L-pipecolate oxidase [Figure 1.3.6 (I)], which generates H_2O_2 during lysine catabolism via the L-pipecolate pathway, is localized to peroxisomes in human (Wanders et al., 1989) and monkey liver (Mihalik and Rhead, 1989). Moreover, glutaryl-CoA oxidase [Figure 1.3.6 (II)], which converts glutaryl-CoA to glutaconyl-CoA, a common catalytic step in the terminal reactions of the degradative pathways of lysine, hydroxylysine and tryptophan, is similarly localized to peroxisomes (Vamecq et al., 1985). That L-pipecolate accumulates in patients afflicted with certain peroxisomal disorders (eg. Zellweger syndrome) further substantiates the localization of the L-pipecolate pathway to peroxisomes in man (Goldfischer et al., 1986) and the role of peroxisomes in amino acid catabolism.

In a species-dependent manner, peroxisomes are known to be involved in amino acid catabolism in another important respect. Hepatic L-alanine:glyoxylate aminotransferase 1 (AGT1), which catalyzes the irreversible transamination of

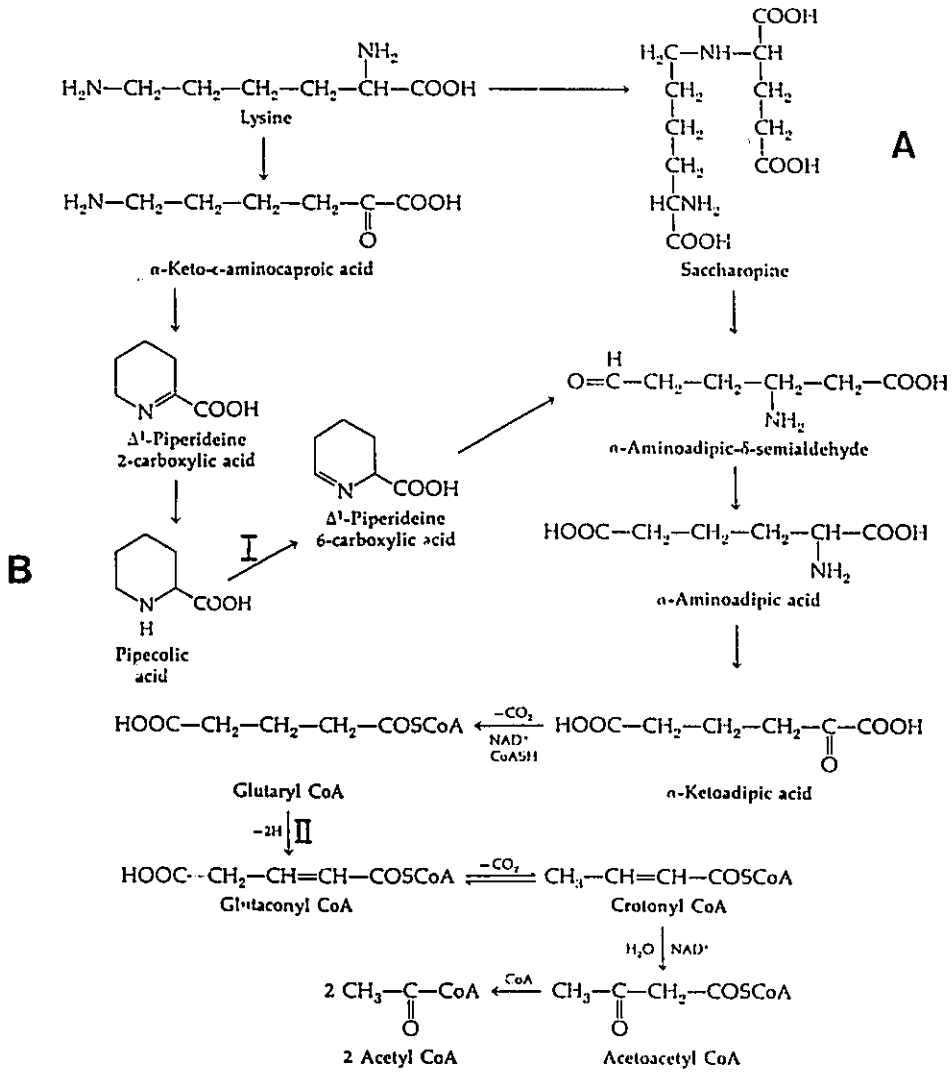


Figure 1.3.6. The saccharopine (A) and L-pipecolate (B) pathways for the catabolism of lysine to α-aminoadipate (and ultimately, acetyl-CoA) in mammals. Two different enzymes, L-pipecolic acid oxidase (I) of the latter pathway and glutaryl-CoA oxidase (II), which catalyzes a common step in the terminal reactions of both of these degradative pathways, are localized to peroxisomes (see section 1.3.2). The figure was reproduced from White et al. (1978).

glyoxylate with L-alanine (the amino-group donor) resulting in the formation of glycine and pyruvate (Noguchi and Takada, 1978), is localized to peroxisomes in primates (including humans) and lagomorphs (Noguchi, 1987; Takada et al., 1990). Conversely, AGT1 is localized to mitochondria in carnivores and to both peroxisomes and mitochondria in rodents (Noguchi, 1987; Takada et al., 1990). In humans, the peroxisomal localization of AGT1 appears to be essential to the efficient detoxification of glyoxylate (Takada et al., 1984). This is evidenced in primary hyperoxaluria type 1 (PH1), a lethal, autosomal recessive, peroxisomal disorder characterized by a deficiency of hepatic peroxisomal AGT1 due either to a lack of enzyme activity or mistargeting of the enzyme to mitochondria (Danpure et al., 1989). The AGT1 deficiency allows accumulated glyoxylate within peroxisomes to be oxidized to oxalate by another peroxisomal enzyme, L-α-hydroxyacid oxidase (also called glycolate oxidase) (Vaughn, 1989; Mannaerts and Van Veldhoven, 1990). The resultant excess of calcium oxalate, due to its very low solubility, is responsible for the pathological symptoms characteristic of PH1 (Williams and Wandzilak, 1989).

Preliminary evidence suggests that in addition to amino acid catabolism, peroxisomes also participate in polyamine catabolism in animal liver. In this regard, a polyamine oxidase purified from rat liver has been shown to possess FAD as a cofactor, generate H_2O_2 upon the oxidation

of spermine and spermidine to 3-aminopropionaldehyde and reside in peroxisomes (Hölttä, 1977; Lazarow, 1987a).

1.3.3 Purine Catabolism

Urate oxidase (uricase; urate:oxygen oxidoreductase), which catalyzes the oxidation of uric acid to allantoin (Usuda et al., 1988), occupies a pivotal position in the degradative pathway for purines, a chain of several enzymes responsible for the sequential conversion of uric acid to allantoin, allantoic acid, urea and ultimately, ammonia and carbon dioxide (Figure 1.3.7; Keilin, 1959). Interestingly, the end product of purine catabolism varies from species to species, with the degradation of purines to urate being common to all animal species and the subsequent degradation of urate less complete in higher animals (Usuda et al., 1988). In bacteria and some marine invertebrates, purines are completely degraded to ammonia and carbon dioxide due to the presence of urease (Noguchi et al., 1986). With phylogenetic evolution, however, the reaction scheme necessary for the degradation of purines to their potential, final metabolic product(s) has become progressively truncated through the successive loss of urease, allantoicase, allantoinase and urate oxidase (Keilin, 1959). Thus, certain fish [eg. teleost fish (carp)] possess allantoicase and excrete urea and glyoxylate, whereas most fish and amphibians possess only

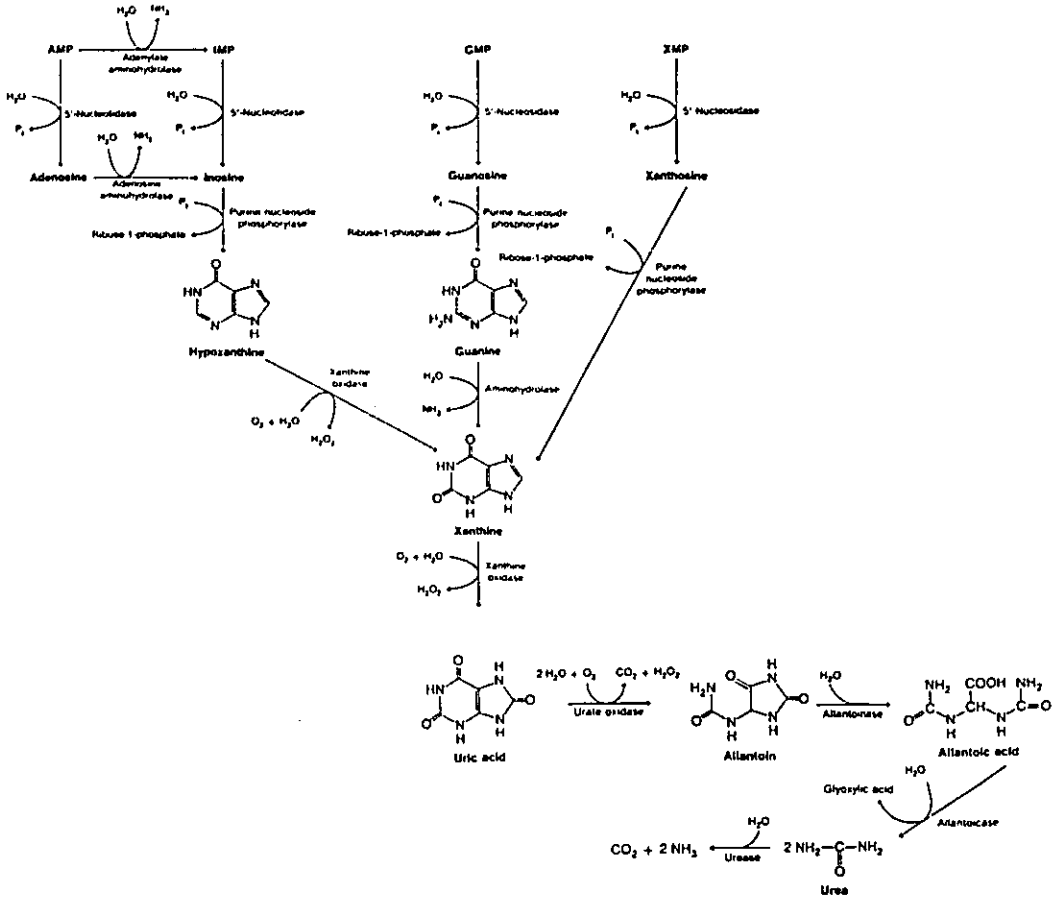


Figure 1.3.7. The major pathways for the degradation of purines to uric acid and the subsequent catabolism thereof. The extent to which uric acid is metabolized varies between species due to the evolutionary loss in higher animals of certain peroxisomal enzymes, including urate oxidase, allantoinase and allantoicase (see section 1.3.3). The figure was reproduced from Zubay (1988).

allantoinase and excrete allantoic acid (Noguchi et al., 1986). Alternatively, since most mammals are known to contain only urate oxidase, they excrete allantoin as the end product of purine catabolism, whereas man, anthropoid apes, some genera of New World monkeys, birds and certain reptiles, which are devoid of urate oxidase, excrete uric acid (Usuda et al., 1988).

Although urate oxidase has been purified from several sources (Völkl et al., 1988), the subcellular fractionation of rat liver has best established the intracellular location of this enzyme (Hruban and Swift, 1964; de Duve and Baudhuin, 1966; Völkl et al., 1988; Usuda et al., 1988). Within this tissue (and that of further evolved mammals possessing this enzyme), urate oxidase is localized to peroxisomes, most often associated with the paracrystalline core or nucleoid present in these organelles (Hruban and Swift, 1964; de Duve and Baudhuin, 1966; Antonenkov and Panchenko, 1978; Völkl et al., 1988; Usuda et al., 1988). Urate oxidase has, however, been localized to the peroxisomal membrane in carp liver (Goldenberg, 1977) and the peroxisomal matrix in fish liver from different sources (Noguchi et al., 1979), amphibian liver (Fujiwara et al., 1987) and soybean nodules (Nguyen et al., 1985). Two additional enzymes involved in urate degradation, allantoinase and allantoicase, have similarly been localized to liver peroxisomes in amphibians, fish and marine

invertebrates (Noguchi et al., 1986). Ureidoglycollate lyase, a metalloenzyme native to the urate degradation pathway in marine fish liver (eg. that of sardine and mackerel), which catalyzes the breakdown of ureidoglycollate to glyoxylate and urea, is also found in peroxisomes (Takada and Noguchi, 1986). Consequently, it has been suggested that those enzymes involved in the degradation of purines to urate are located in the cell cytosol, whereas enzymes catalyzing the subsequent degradation of urate to ammonia and carbon dioxide (or some "precursor" theretofore, depending on the extent of evolutionary conservation in that species) are localized to peroxisomes (Noguchi et al., 1979, 1986).

1.3.4 The Pentose Phosphate Pathway

The pentose phosphate pathway (also called the phosphogluconate pathway; Figure 1.3.8) represents an alternative scheme for carbohydrate (ie. glucose) oxidation in mammalian cells, serving as a principle source of reducing equivalents for the reduction of $NADP^+$ to NADPH and as a mechanism for the synthesis and disposal of pentoses (White et al., 1978). The oxidative steps of this pathway involve the conversion of glucose-6-phosphate to 6-phosphogluconate (catalyzed by glucose-6-phosphate dehydrogenase), followed by the conversion of 6-phosphogluconate to ribulose-5-phosphate (catalyzed by 6-

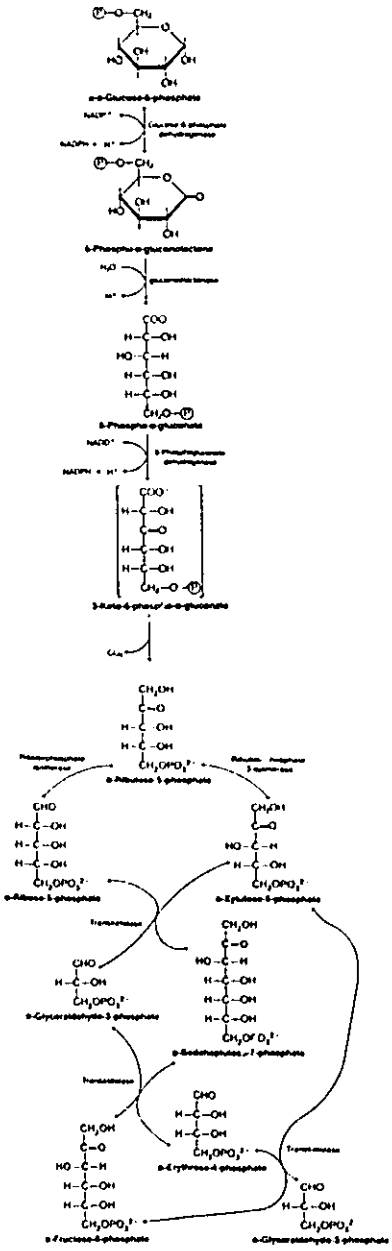


Figure 1.3.8. The pentose phosphate (or phosphogluconate) pathway of carbohydrate oxidation. Two enzymes of this pathway, glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase, are localized to peroxisomes (see section 1.3.4). The figure was reproduced from Zubay (1988).

phosphogluconate dehydrogenase); each of these reactions is accompanied by the synthesis of one molecule of the reduced coenzyme NADPH. In the subsequent, non-oxidative segment of the pathway, ribulose-5-phosphate serves as the substrate for a series of pentose interconversions resulting in the biosynthesis of three-, four-, five-, six- and seven-carbon sugar phosphates (White et al., 1978).

Although it is generally accepted that the oxidation of carbohydrate occurs in the cell cytosol where the enzymes of glycolysis and the hexose monophosphate shunt are located (White et al., 1978), preliminary work has indicated that the key enzymes of the pentose phosphate pathway are also detectable in total particulate fractions (containing mitochondria, peroxisomes and lysosomes) isolated from rat liver (Beitner and Naor, 1972; Hori and Yonezawa, 1972; Zaheer and McLean, 1972). Similarly, in plants, pentose phosphate cycle enzymes are localized not only to the soluble fraction but also in plastids, including chloroplasts (Schnarrenberger and Oeser, 1984). Recently, Panchenko and Antonenkov (1984) and Antonenkov (1989) further investigated the subcellular distribution of the pentose phosphate cycle enzymes in rat liver and detected the activities of the $NADP^+$-dependent dehydrogenases (glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase) in the purified peroxisomal fraction (as well as in the cytosol); within the peroxisome, both

dehydrogenases were localized to the matrix. Conversely, the enzymes of the non-oxidative portion of this pathway (transketolase, transaldolase, triosephosphate isomerase and glucosephosphate isomerase) were present only in the cytosol (Antonenkov, 1989). Antonenkov (1989) further suggested that the biological significance of the location of pentose phosphate pathway dehydrogenases within peroxisomes may be related to the intraperoxisomal generation of $NADP^+$ necessary for lipid synthesis within this organelle, particularly during the initial steps of triglyceride and cholesterol biosynthesis.

1.3.5 The Glyoxylate Cycle

In contrast to the oxidative nature of the pentose phosphate pathway toward carbohydrate, "specialized peroxisomes" called glyoxysomes (see section 1.1.2) are responsible for the operation of the glyoxylate cycle, a five enzyme variant of the tricarboxylic acid cycle which permits the net conversion of fat to carbohydrate (Figure 1.3.9, enzymes labeled I through V; Tolbert, 1981; Lazarow, 1987a). Courtesy of each turn of the glyoxylate cycle, two moles of acetyl-CoA produced by β-oxidation are converted to one mole each of succinate and glyoxylate; succinate may then be metabolized to malate within the mitochondria and used for gluconeogenesis (Beevers, 1982). Malate formed within the glyoxysome (from glyoxylate and acetyl-CoA) is

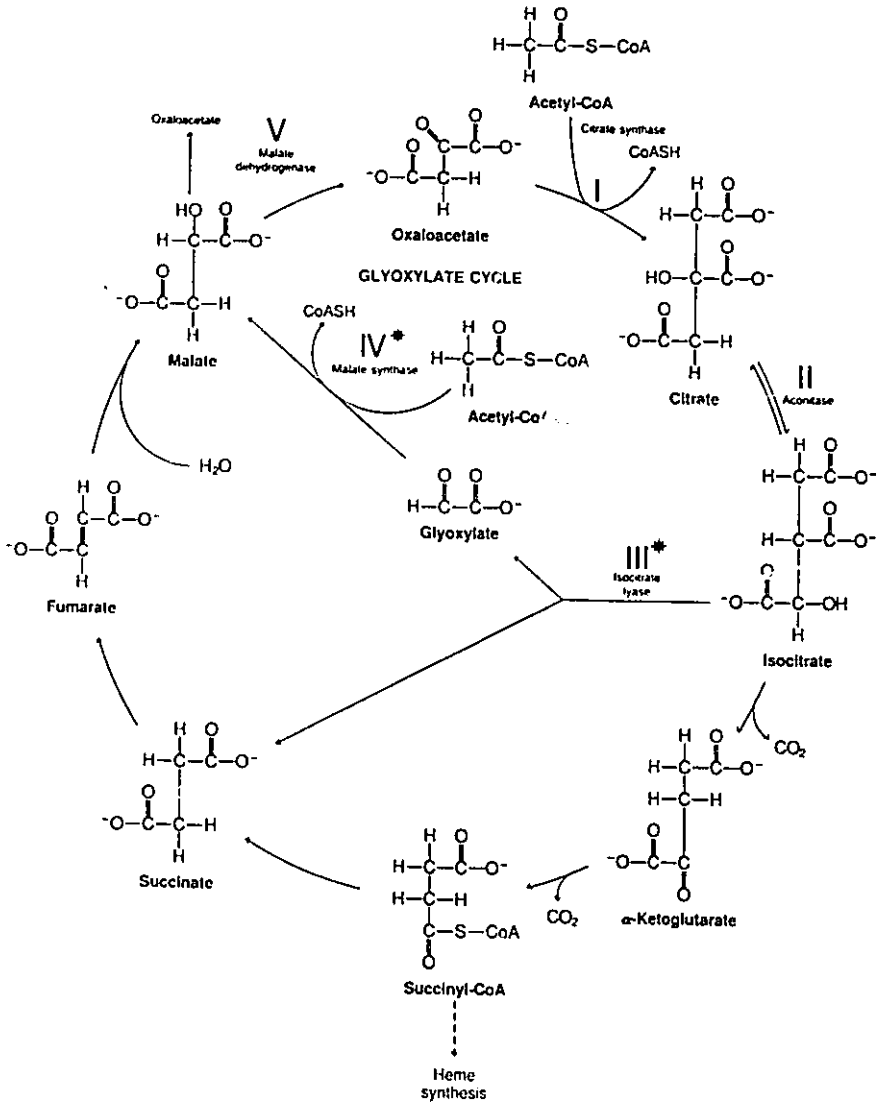


Figure 1.3.9. Comparison of the glyoxylate and tricarboxylic acid cycles. The glyoxylate cycle, which is localized to "specialized peroxisomes" called glyoxysomes, is a five enzyme variant of the tricarboxylic acid cycle that permits the net conversion of fat to carbohydrate (see section 1.3.5). The five enzymes of the glyoxylate cycle are indicated (labeled I through V); the two glyoxysome-specific (ie. marker) enzymes of this pathway, isocitrate lyase and malate synthase, are denoted by asterisks. The figure was reproduced from Zubay (1988).

either metabolized to oxaloacetate and used to reinitiate the glyoxylate cycle or diffuses from the organelle to malate pools in the cytosol and mitochondria (Tolbert, 1981, 1982).

Glyoxylate cycle enzymes (glyoxysomes) were first discovered in the endosperm tissue of germinating castor beans (Cooper and Beevers, 1969), subsequent to which stemmed the realization that glyoxylate cycle reactions are, to a significant degree, independent of those of the tricarboxylic acid cycle occurring in mitochondria (Beevers, 1982). Both glyoxysomes and mitochondria are known to contain specific isozymes of three of the five glyoxylate cycle enzymes, including citrate synthase, aconitase and malate dehydrogenase (Figure 1.3.9, enzymes labeled I, II and V, respectively), however, only glyoxysomes contain isocitrate lyase and malate synthase, the remaining two unique, and therefore used as marker, enzymes of the glyoxylate cycle (Figure 1.3.9, enzymes labeled III and IV, respectively; Beevers, 1982; Ettinger and Harada, 1990). All five glyoxysomal enzymes are found in higher plants, certain nematodes and protozoa (Lazarow, 1987a; Davis et al., 1989) whereas glyoxysomes from yeast, Tetrahymena, Euglena and various fungi contain only isocitrate lyase and malate synthase (Tolbert, 1981; Lazarow and Fujiki, 1985). Generally, it is believed that the tissues of higher organisms are not capable of converting the carbon of

acetyl-CoA into sugar (White et al., 1978). Nevertheless, glyoxylate cycle enzymes, and the corresponding ability to synthesize carbohydrate from lipid, have been identified in toad bladder (Goodman et al., 1980; Lazarow, 1987a), fetal guinea pig liver (Jones, 1980) and rat liver (Davis et al., 1989). It should be noted, however, that the subcellular location of the glyoxylate cycle enzymes in higher animals has not been determined.

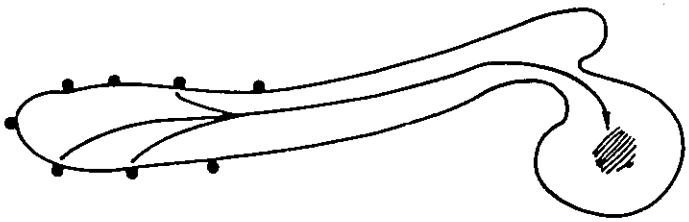
1.4 BIOGENESIS OF PEROXISOMES

1.4.1 Theories of Peroxisome Biogenesis

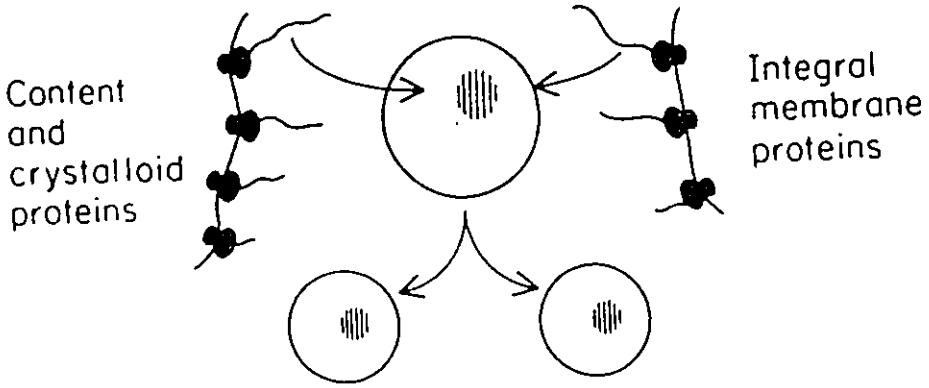
Contrary to the wealth of information present in the literature concerning the properties and functions of the peroxisome, the biogenesis of this organelle has been characterized to a lesser degree. Consequently, the ultrastructural basis of peroxisome biogenesis and the exact mechanism(s) involved in the transfer of proteins to peroxisomes remain debatable. Early biochemical studies demonstrating the apparent localization of newly-synthesized catalase to rat liver microsomes (Higashi and Peters, 1963), as well as ultrastructural observations of direct membrane continuities between the ER and peroxisomes (Novikoff and Shin, 1964; Novikoff et al., 1973; Novikoff and Novikoff, 1982; reviewed in Lazarow and Fujiki, 1985), implicated the ER as the place of origin of this organelle. From this, stemmed the classical model of peroxisome biogenesis which

envisioned the sequential synthesis of peroxisomal proteins on membrane-bound polyribosomes, the co-translational sequestering of these polypeptides within the ER (similar to secretory proteins), and subsequent to the passage of these proteins through the ER into certain peripheral segments (buds), the pinching off of these ER 'outpouchings' to form peroxisomes (Figure 1.4.1, A; Lazarow et al., 1982).

Later morphological investigations challenged this model, however, by reporting the absence of any direct connections between the ER and peroxisomes, thereby negating the notion of luminal continuity between these organelles (Legg and Wood, 1970; Rigatuso et al., 1970; Gorgas, 1985; Zaar et al., 1987). Numerous other studies (morphological, biochemical and molecular in nature) have since reinforced this position (extensively reviewed in Lazarow and Fujiki, 1985; Yamamoto and Fahimi, 1987; Zaar et al., 1987) resulting in a second model of peroxisome biogenesis which proposes that the synthesis of peroxisomal proteins on free polyribosomes is followed by posttranslational entry of these polypeptides into preexisting, "immature" peroxisomes. Upon maturation (ie. upon progressive enlargement of the organelle), new peroxisomes are formed by fission of preexisting ones (Figure 1.4.1, B; Lazarow and Fujiki, 1985; Lazarow, 1987a). In an attempt to account for the tail-like extensions and ring- or hook-shaped membraneous structures occasionally found attached to peroxisomes in electron



A: Classical model of peroxisome biogenesis



B: Current model of peroxisome biogenesis

Figure 1.4.1. Models of peroxisome biogenesis. Figures A and B were reproduced from Lazarow et al. (1982) and Lazarow (1987a), respectively.

micrographs [which were incorrectly interpreted at the time to be connections between peroxisomes and the smooth ER (reviewed in Lazarow and Fujiki, 1985)], it has been proposed that peroxisomes may be interconnected, either permanently or transiently, forming a reticular network (Lazarow et al., 1980). Consistent with this concept of a peroxisomal reticulum, three-dimensional reconstruction studies of mouse liver hepatocytes (Gorgas, 1985) and regenerating rat liver hepatocytes (Yamamoto and Fahimi, 1987) have revealed that interconnecting bridges do exist between two, and often several, peroxisomes. This implies that new peroxisomes must form by breaking away from the peroxisomal reticulum, which is consistent with the budding phenomenon responsible for peroxisome formation in the yeast Hansenula polymorpha (Veenhuis et al., 1978). Under conditions of rapid growth, larger peroxisomes within this yeast are known to form small buds, which later detach and migrate into daughter cells giving rise to a new generation of peroxisomes (Veenhuis et al., 1978). Furthering the reticulum concept, Baumgart et al. (1989) have also demonstrated the presence of a peroxisomal membrane system within rat liver hepatocytes which may act as a membrane reservoir for the proliferation of peroxisomes and for the expansion of the peroxisomal compartment (reticulum).

As a compromise between the first two models of peroxisome biogenesis, a third model proposes that at least

one peroxisomal membrane protein [possibly a receptor(s) necessary for the import of peroxisomal content proteins] is synthesized on membrane-bound polyribosomes and co-translationally inserted into the rough ER. Here, it is selectively sorted from other integral membrane polypeptides to form a membrane segment that buds from the ER forming an "immature" peroxisome. Other peroxisomal membrane and matrix proteins, synthesized on free polyribosomes, are subsequently inserted/imported into the immature organelle yielding a mature peroxisome (Goldman and Blobel, 1978; Köster et al., 1986). At present, there is very little evidence supporting this model, however, the recent localization of a 50 kDa integral membrane polypeptide of rat liver peroxisomes to the ER, and the preferential synthesis of this polypeptide on membrane-bound polyribosomes, may be suggestive of some involvement of the ER in peroxisome biogenesis (Bodnar and Rachubinski, 1991).

Although three models of peroxisome biogenesis have been proposed, reviews of the relevant literature (Lazarow and Fujiki, 1985; Borst, 1986, 1989) acknowledge that the biogenesis of peroxisomes is most consistent with the scheme proposed by the second (or current) model (Figure 1.4.1, B). This consensus stems from a multitude of independent studies that have either cloned and sequenced genes or cDNAs encoding peroxisomal matrix and membrane proteins (see Table 1.4.1), analyzed the import of proteins into peroxisomes

both *in vitro* and *in vivo* (see section 1.4.3.3) or in some other morphological, biochemical or molecular fashion, generated some insight into the "life cycle" (ie. biosynthesis, maturation and turnover) of this organelle (Lazarow and Fujiki, 1985; Borst, 1986, 1989). Concisely summarized, our current understanding of peroxisome biogenesis proposes that: 1) peroxisomal proteins are encoded by nuclear genes; 2) peroxisomal matrix, crystalloid core and membrane proteins are synthesized on free polyribosomes (with only one reported exception; see above), diffuse through the cytosol with typical half-times of 1 to 15 minutes (Lazarow, 1987a) and are imported posttranslationally into preexisting peroxisomes (or inserted posttranslationally into preexisting peroxisomal membranes in the case of membrane proteins); 3) with a few acknowledged exceptions (see section 1.4.3.3), most peroxisomal proteins are synthesized at their mature size and do not undergo any posttranslational modification (eg. proteolytic processing or glycosylation); 4) new peroxisomes originate by fission of preexisting peroxisomes (ie. peroxisomes are thought not to arise from the ER or *de novo*).

1.4.2 Evolutionary Origin of Peroxisomes (Microbodies)

While it is generally accepted that peroxisomes arise from the growth and division of preexisting

peroxisomes, the origin of the first peroxisome remains speculative. Similar to the ancestral origin of mitochondria and chloroplasts, one hypothesis suggests that the peroxisome originally entered a protoeukaryote as an endosymbiont. Subsequently, during the course of evolution, the peroxisomal ancestor has transferred its entire genome to the host cell nucleus (de Duve, 1983; Lazarow and Fujiki, 1985; Borst, 1986, 1989). This complete transfer of DNA to the nucleus is in direct contrast to the situation established for mitochondria and chloroplasts, which have each retained a percentage of their original genome and thus separate organelle systems for DNA replication, mRNA transcription and protein synthesis (de Duve, 1983; Borst, 1986). Alternatively, it has been hypothesized that the peroxisome may represent the birth of a unique compartment within eukaryotic cells that originated from another cellular membrane and has become independent and specialized through evolution (de Duve, 1983; Borst, 1986, 1989). Comparing the peroxisomal β-oxidation systems in fungi, plants and animals, Kunau et al. (1988) revealed a high degree of uniformity in the overall organization and subunit structure of the β-oxidation enzymes, which is significantly different from that of enzymes native to the mitochondrial β-oxidation systems. While this is compatible with a common origin of peroxisomes from a single event, the enzymic composition of microbodies, which is more diverse than that

of mitochondria and chloroplasts due to its organism and tissue specificity, remains to be explained (Borst, 1989). Further, this does not distinguish between the endosymbiont and evolutionary theories of microbody origin.

Two explanations are conceivable for the present-day enzymic diversity of microbodies (Borst, 1986). If acquired endosymbiotically, microbodies have as their ancestor a functionally versatile prokaryote that has, through the differential loss of enzymic capabilities, yielded the spectrum of microbodies currently known to function in several different organisms (Borst, 1986). Alternatively, independent of their origin, the enzymic diversity of microbodies could reflect the redistribution (ie. both gain and loss) of enzymatic activities that has occurred with evolution (Borst, 1986). The second explanation is thought to be the more plausible, however, it is difficult to envision how complex metabolic pathways, found normally in other cellular locations (eg. the cytosol), were transferred to microbodies (Borst, 1986, 1989). Possibly, this transfer of metabolic processes to membrane-bound organelles was driven by selective advantages related to reaction efficiency (eg. shortened substrate diffusion times, optimal pH or the partitioning of products to preferred metabolic fates) and/or the local inactivation of toxic metabolites (eg. H_2O_2) (Borst, 1986, 1989).

The major prediction of the endosymbiont hypothesis

is that genes encoding organelle proteins, independent of their location in either the organellar genome or DNA within the nucleus, will show clear homology with their prokaryotic counterparts (Borst, 1986, 1989). Thus, to verify endosymbiotic ancestory, the sequences of microbody proteins must be compared with their homologues in prokaryotes and in the eukaryotic cytosol. The glycosomes of T. brucei prove valuable in this respect because glycolytic enzymes are highly conserved and present in all prokaryotes and eukaryotes (Borst, 1986, 1989). Most informative have been the T. brucei glyceraldehyde-phosphate dehydrogenases (GAPDHs), which are, by virtue of their sequences, very different from one another (Borst, 1989). Interestingly, the cytosolic enzyme shares the same high degree of homology of all eukaryotic GAPDHs to the E. coli enzyme, whereas the glycosomal enzyme exhibits approximately equal sequence identity to both eukaryotic and prokaryotic GAPDHs (Michels et al., 1986; Misset et al., 1987). While this is suggestive of a different lineage for the cytosolic and glycosomal enzymes, the nature of these lineages is debatable (Borst, 1989) and significantly more work is required to substantiate any theory concerning the origin of microbodies.

1.4.3 Intracellular Protein Targeting

Protein sorting is a fundamental property of all

living cells. In prokaryotes (eg. bacteria), mechanisms exist to address proteins to a minimum of five different locations: the cytoplasm, inner membrane, periplasmic space, outer membrane and extracellular medium (reviewed by Lee and Beckwith, 1986; Saier et al., 1989; Bieker and Silhavy, 1990). Eukaryotes, by contrast, have achieved functional complexity by division of their cytosol into specialized compartments (ie. organelles), with each compartment harbouring a unique set(s) of proteins that enables it to carry out its specific function(s). As a result, eukaryotes possess a greater number of intracellular targets for proteins, further emphasizing that the accurate and efficient delivery of each protein to its correct compartment is an important step in the maintenance and development of the cell. Although relatively little is known about protein targeting to peroxisomes (or microbodies in general), the targeting of proteins to other intracellular destinations, including the secretory pathway, mitochondria, chloroplasts and nuclei, is better characterized (reviewed by Verner and Schatz, 1988; von Heijne, 1988; Pugsley, 1989; Lazdunski and Benedetti, 1990). To provide a background with which to compare certain aspects of protein targeting to peroxisomes, the following discussion presents the mechanisms involved in the translocation of proteins into other organelles. For the sake of brevity, only the targeting of proteins to

nonsecretory organelles (mitochondria, chloroplasts and nuclei), which occurs posttranslationally and parallels protein targeting to peroxisomes, will be discussed.

1.4.3.1 Protein Targeting to Nonsecretory Organelles (Mitochondria, Chloroplasts and Nuclei)

1.4.3.1.1 Mitochondria

Although mitochondria possess their own genetic system and synthesize some polypeptides (13 in humans), greater than 90% of the different mitochondrial proteins are encoded by nuclear genes (Pfanner and Neupert, 1990; Baker and Schatz, 1991). In general, these proteins are synthesized on free polysomes in the cytoplasm as larger precursors possessing N-terminal extensions (Hallermayer et al., 1977). These N-terminal extensions vary in the number, sequence and content of amino acids, generally exhibiting a deficiency of acidic amino acids and possessing a few hydrophobic, and a high content of basic and hydroxylated, residues (von Heijne, 1988). Structurally, a number of these mitochondrial presequences form amphiphilic α-helices, with their basic and hydrophobic amino acids sequestered on opposite faces of the helix (von Heijne, 1988). Functionally, these prepieces contain the information required to target the proteins to mitochondria where they are posttranslationally imported into one of four subcompartments: the outer membrane, intermembrane space,

inner membrane or matrix (Hartl et al., 1989). Thus, the binding and import of mitochondrial precursor proteins is, in general, not dependent on a coupled translational process, however, it has been suggested that under certain physiological conditions *in vivo*, mitochondrial precursors may be imported in a co-translational manner (Hartl et al., 1989). It is also important to note that certain mitochondrial proteins [eg. yeast porin and 70 kDa outer membrane proteins (Hase et al., 1984; Kleene et al., 1987) and *Neurospora crassa* apocytochrome c (Stuart et al., 1987)] do not contain N-terminal presequences and are not proteolytically processed upon import. Rather, the specific targeting information for these proteins resides within the mature polypeptide, although it is not known whether this targeting information is the result of the primary structure itself or stems from a critical conformational arrangement of the protein (Stuart et al., 1988).

The initial recognition of mitochondrial protein precursors is achieved by proteinaceous components (import receptors) localized to the outer surface of the outer mitochondrial membrane; two such receptors, MOM19 and MOM72 (for mitochondrial outer membrane proteins of 19 and 72 kDa, respectively) have been identified (Söllner et al., 1989, 1990). Upon receptor recognition, the precursor proteins are inserted into the outer mitochondrial membrane through a common membrane insertion site termed the 'general insertion

protein' (GIP) (Söllner et al., 1989). Transport across both mitochondrial membranes then occurs in a single step at contact sites (stable structures whose existence is independent of the presence of the precursor protein) between the outer and inner membranes (Pfanner et al., 1990). Insertion into the contact site is dependent upon an electrochemical potential ($\Delta\psi$) across the inner membrane, possibly because the $\Delta\psi$ produces an electrophoretic driving force on the positively charged portions of the precursor protein which serves to mediate penetration through the contact site; completion of translocation across the inner mitochondrial membrane can occur in the absence of $\Delta\psi$ (Pfanner et al., 1990). In addition to $\Delta\psi$, protein import requires energy derived from the hydrolysis of ATP (or GTP); presumably, this energy confers a translocation-competent conformation to the precursor protein (Eilers and Schatz, 1988; Pfanner et al., 1988). The maintainance of mitochondrial protein precursors in a loosely-folded, translocation-competent state is also achieved through interaction with a heat shock protein (HSP), HSP70, which acts as a cytosolic 'chaperone' by selectively binding and stabilizing partly folded proteins (Ellis, 1987; Deshaies et al., 1988).

During or immediately following translocation through the contact site, the N-terminal presequences of precursor proteins are proteolytically processed by a metal-

dependent processing peptidase located in the matrix; two proteins have been determined to be constituents of this processing activity: the mitochondrial processing peptidase (MPP) and a processing enhancing protein (PEP) (Hartl et al., 1989; Hartl and Neupert, 1990). In an ATP-dependent process that modulates protein folding, imported proteins may then interact transiently with a matrix-located chaperone, HSP60 (Ostermann et al., 1989; Baker and Schatz, 1991), prior to being sorted to their appropriate intramitochondrial location where they are functionally assembled into oligomeric complexes (Pfanner and Neupert, 1990). Generally, the sorting of precursors is accomplished by direct transfer to their correct submitochondrial compartment. In select cases (eg. the precursor to Fe/S protein of the ubiquinol-cytochrome c reductase complex, cytochrome b_2 and cytochrome b_1; Hartl et al., 1987), however, sorting occurs by a more sophisticated mechanism in which precursors are initially directed to the matrix (thereby following the conserved sorting and assembly pathways that were established in the prokaryotic ancestors of mitochondria), where the first part of a bipartite presequence is removed by the processing peptidase. An intermediate-sized product is then redirected back across the inner membrane to the intermembrane space where the second portion of the bipartite presequence is proteolytically removed (Hartl et al., 1987; 1989).

In summary, the targeting of proteins to mitochondria involves several consecutive steps which effect the efficient and specific translocation of proteins to their correct submitochondrial destination. For numerous mitochondrial protein constituents, many of these steps are common to the import process, however, the overall targeting pathway often varies significantly between different proteins. To gain a comprehensive understanding of protein translocation into mitochondria, the following recent reviews should be consulted: Stuart et al., 1988; von Heijne, 1988; Hartl et al., 1989; Hartl and Neupert, 1990; Neupert et al., 1990; Pfanner and Neupert, 1990; Pfanner et al., 1990; Baker and Schatz, 1991.

1.4.3.1.2 Chloroplasts

Although chloroplasts possess their own genome and are competent in protein synthesis, most chloroplast proteins, like those of mitochondria, are encoded by nuclear genes and synthesized as precursors on free cytosolic ribosomes (Smeekens et al., 1990). Functionally, the N-terminal extensions of chloroplast precursor proteins serve to target proteins to this organelle with high specificity, where they are posttranslationally imported by a process thought to be more complicated than that of mitochondria since chloroplasts harbour three different membranes (the outer and inner envelope membranes and thylakoid membrane)

defining three distinct intraorganellar spaces (the inner membrane space, chloroplast stroma and thylakoid lumen) (von Heijne, 1988; Keegstra, 1989). Chloroplast presequences are of variable length (approximately 30 to 80 residues) and, when compared to other targeting peptides (eg. secretory signal sequences or mitochondrial transit peptides), lack obvious structural consensus features (ie. no highly conserved regions of homology) beyond an enrichment for serine and threonine residues, a relatively low content of acidic amino acids and a semi-conserved pattern of residues immediately adjacent to the presequence cleavage site (von Heijne et al., 1989; Gavel and von Heijne, 1990). Common structural domains among stromal and thylakoid presequences include an uncharged amino-terminal portion, a central non-amphiphilic segment and a carboxy-terminal amphiphilic β-strand in close proximity to the cleavage site (von Heijne et al., 1989).

The first step in the overall process of translocation across the chloroplast membrane is binding of the precursor to the outer membrane. Binding is, in general, energy independent and is thought to be receptor-mediated (von Heijne, 1988; Keegstra, 1989; Flügge, 1990). Similar to mitochondria, subsequent translocation of chloroplast preproteins across the outer and inner membranes may occur simultaneously at envelope membrane junctions (Pain et al., 1988). Interestingly, ATP is the only energy

source required for protein transport into chloroplasts; an electrochemical gradient, which is necessary for mitochondrial and bacterial protein transport, is not required (Theg et al., 1989). During, or immediately after, translocation into the stroma, stromal preproteins are processed to their mature size by a highly specific, chelator-sensitive stromal peptidase (Robinson and Ellis, 1984). Alternatively, proteins destined for the thylakoid lumen (eg. plastocyanin) are subject to two successive processing steps (Kirwin et al., 1989) similar to those observed for mitochondrial proteins destined for the intermembrane space (see section 1.4.3.1.1). The first step, catalyzed by the stromal processing peptidase, yields an intermediate-sized precursor which is redirected to the thylakoid lumen where it is processed to its mature form by a thylakoid processing peptidase localized to the luminal surface of the thylakoid membrane (Flügge, 1990). Interestingly, the transit peptides of these precursors contain, in addition to the three elements of the N-terminal envelope transfer domain (see above), a C-terminal apolar region (or thylakoid transfer domain) (von Heijne et al., 1989).

Further paralleling protein targeting to mitochondria, a translocation-competent conformation of the precursor protein is essential to import into chloroplasts (Flügge, 1990) and soluble (ATP-dependent) cytosolic factors

are hypothesized to assist in this regard (Waegemann et al., 1990). It is also thought that soluble stromal factors participate in protein translocation into chloroplasts by either conferring translocation competence to proteins that are imported into the stroma but are destined for the thylakoid lumen or by stabilizing those imported proteins that are to be assembled into multimeric protein complexes (Flügge, 1990). To achieve the observed high specificity of chloroplast protein import, it has been suggested that some consensus secondary structure must exist among chloroplast transit peptides (Keegstra, 1989), however, von Heijne and Nishikawa (1991) postulate that these presequences are 'perfect random coils' and a series of interactions with cytosolic and chloroplast chaperones may be all that is necessary for chloroplast targeting and import.

For greater detail concerning the intricacies of protein targeting to chloroplasts, the following recent reviews should be consulted: von Heijne (1988); Keegstra (1989); Flügge (1990); Smeekens et al. (1990); von Heijne and Nishikawa (1991).

1.4.3.1.3 Nuclei

Structurally, the nucleus is comprised of an internal aqueous compartment (the nucleoplasm) which is delimited by the nuclear envelope (a double bilayer membrane separated by the perinuclear space) (Roberts, 1989; Wagner

et al., 1990). Nuclear pores (water-filled channels within large proteinaceous complexes) traverse the nuclear membrane at sites where the inner and outer membranes are thought to be fused, thereby providing access to the nucleoplasm (Roberts, 1989; Wagner et al., 1990). Although the perinuclear space and outer nuclear membrane are continuous with the ER lumen and ER membrane, respectively, the targeting of proteins to the nucleus is not related to the secretory pathway (Wagner et al., 1990). Rather, access to the nucleoplasm is gained via the nuclear pore complex which allows the passive diffusion of small proteins or molecules (of less than 60 kDa or 4.5 nm in diameter), and mediates the active transport of larger proteins, across the nuclear envelope (Silver and Goodson, 1989; Wagner et al., 1990; Silver, 1991).

The currently accepted model for nuclear import contends that nuclear proteins, which are synthesized on free cytosolic polysomes and imported posttranslationally, contain inherent sequence information [ie. a nuclear localization signal (NLS)] which dictates their selective translocation into the nucleus (Silver, 1991). There is no consensus sequence among NLSs, however, they are typically short sequences of 8 to 10 amino acids that: 1) contain a high proportion of positively charged amino acids (lysine and arginine) often associated with a proline residue; 2) may reside in any exposed region of the protein; 3) are not

proteolytically processed following import; and 4) may occur more than once in the protein (Wagner et al., 1990). Nuclear localization signals mediate protein translocation across the nuclear envelope via a two step (at least) process that includes: 1) a recognition event involving protein binding to a receptor located in the cytoplasm, on the nuclear envelope and/or at the nuclear pore complex (Goldfarb et al., 1986; Breeuwer and Goldfarb, 1990) and 2) the subsequent ATP- and physiological temperature-dependent translocation of the protein through the nuclear pore complex (Newmeyer and Forbes, 1988; Wagner et al., 1990). Cellular factors (NLS-binding proteins) are known to interact with NLSs but their role in nuclear translocation is not well characterized; NLS-binding proteins, located in the cytoplasm and/or at the nuclear envelope, may recognize NLSs and deliver proteins to the nuclear pore complex (Adam et al., 1990; Newmeyer and Forbes, 1990; Silver, 1991). In contrast with protein import into mitochondria and chloroplasts, there is evidence to indicate that proteins can traverse the nuclear membrane in their native conformation (Silver and Goodson, 1989).

A more elaborate discussion of the characteristics of protein targeting to the nucleus can be found in the following recent reviews: Roberts, 1989; Silver and Goodson, 1989; Wagner et al., 1990; Silver, 1991.

1.4.3.2 Protein Targeting to Microbodies Other Than Peroxisomes

As highly specialized microbodies unique to Trypanosomes and other Kinetoplastida (see section 1.1.2), glycosomes are characterized predominantly by their complement of enzymes related to glycolysis and glycerol metabolism, but also harbour several enzymes contributing to other metabolic processes (eg. purine salvage, pyrimidine biosynthesis, carbon dioxide fixation, ether-lipid biosynthesis and fatty acid β-oxidation; Opperdoes, 1987a, 1987b). Despite only a slight resemblance to peroxisomes with respect to functionality, glycosomes strongly resemble peroxisomes both physically and morphologically (Opperdoes et al, 1984; reviewed in Borst, 1986, 1989). As a consequence of their shared properties, glycosomes and peroxisomes (microbodies) are thought to be evolutionarily related and inferences gained from studying protein import into glycosomes may provide clues concerning the mechanism(s) of protein import into peroxisomes (Opperdoes, 1988).

It is well documented that glycosomal enzymes are encoded by nuclear genes, synthesized on free polyribosomes in the cytosol and imported into the glycosome via a rapid posttranslational event that, contrary to the situation established for certain mitochondrial, chloroplast and nuclear proteins (section 1.4.3.1), does not involve protein

modification or peptide cleavage (Clayton, 1987; Hart et al., 1987; Opperdoes, 1987a, 1987b). By inference, the topogenic signal specifying protein localization to glycosomes must therefore reside in the sequence or structure of the mature enzyme(s). Difficulties in establishing a competent *in vitro* import assay or *in vivo* trypanosomatid genetic transformation system have, however, dictated that the transfer of glycolysis from its normally cytosolic location to an alternative location within a microbody serve as the primary "tool" with which to study changes in cytosolic enzymes that facilitate translocation into glycosomes (Wierenga et al., 1987; Swinkels et al., 1988; Kendall et al., 1990). As a result, speculation concerning specific translocation signals has originated from sequence comparisons between glycosomal enzymes and cytosolic isoenzymes from the same parasite or cytosolic homologues found in other organisms (Borst, 1986, 1989).

To this end, most glycosomal enzymes have been found to exhibit high isoelectric points (ranging between 9 and 10), with an excess of basic amino acids relative to their cytosolic counterparts in *T. brucei* and other organisms being confirmed by the sequence analysis of cloned genes (Borst, 1989; Sommer et al., 1990). One such enzyme, phosphoglycerate kinase (PGK), has been quite informative with respect to signals involved in microbody import since trypanosomes harbour both a cytosolic and glycosomal PGK

which differ in only 7% of their amino acid sequences (Osinga et al., 1985). This minimal sequence divergence translates into two important differences between the mature enzymes: at neutral pH, glycosomal PGK possesses a high net positive charge compared to a non-charged cytosolic PGK and the glycosomal enzyme contains a 20 amino acid C-terminal extension which is absent from the cytosolic enzyme (Osinga et al., 1985). Either of these differences, or a combination of both, are thought to affect the different intracellular locations of the two PGKs (Osinga et al., 1985). Furthering this notion, a comparison of the three-dimensional structures of three glycosomal enzymes (triosephosphate isomerase, glyceraldehyde-phosphate dehydrogenase and phosphoglycerate kinase) between each other and their cytosolic counterparts led to the hypothesis that two clusters of positively charged amino acids (termed 'hot spots'), separated by ≈40 Å on the protein surface and possessing two basic amino acids displaced from one another by ≈7 Å, may be important to glycosomal localization (Wierenga et al., 1987). The sequence of glycosomal aldolase, for which no homologous three-dimensional information is available, also appears to be compatible with the presence of similar surface hot spots. Further substantiating their hypothesis, Wierenga et al. (1987) cited that one hot spot was always present within a unique peptide insertion in each glycosomal enzyme and that both

hot spots were absent from corresponding positions within the cytosolic isoenzymes and cytosolic homologues of other organisms. The hot spot model does not hold, however, for other glycosomal enzymes. Sequence data for glycosomal and cytosolic PGK isoenzymes from another trypanosomatid, Crithidia fasciculata, reveal that apart from two conservative N-terminal amino acid substitutions, these enzymes, which exhibit the same overall charge, differ only by the presence (in glycosomal PGK) of a 38 amino acid C-terminal extension (Swinkels et al., 1988). Accordingly, Swinkels et al. (1988) conclude that the C-terminal extension of glycosomal PGK is responsible for its glycosomal location, while surface hot spots may be involved in intraglycosomal substrate interactions. Molecular modeling of both T. cruzi and T. brucei glyceraldehyde-phosphate dehydrogenase enzymes similarly discredits the hot spot model of glycosomal import since the T. cruzi enzyme lacks one of the positively-charged motifs and neither structure fulfils the necessary spatial requirements (Kendall et al., 1990). These data imply that glycosomal protein import is not dependent upon a single, universal topogenic signal (Opperdoes, 1988; Borst, 1989).

Originally described in trichomonad flagellates, and until recently considered a member of the family of organelles called microbodies (Keller et al., 1991), the hydrogenosome is a functionally distinct organelle endowed

with an important role in energy metabolism (Müller, 1975, 1988). Biochemically, these organelles serve as sites for the oxidative decarboxylation of pyruvate, which is accompanied by the production of acetyl-CoA; the energy of acetyl-CoA is subsequently conserved by substrate level phosphorylation of ADP (Müller, 1988). Compared to microbodies (ie. peroxisomes, glyoxysomes and glycosomes), only very limited information is available concerning protein targeting to hydrogenosomes. Sequence analysis of the gene encoding hydrogenosomal ferredoxin from the anaerobic protist Trichomonas vaginalis predicts the presence of an N-terminal octapeptide that is absent from the mature protein (Johnson et al., 1990). While its ability to function as a leader peptide remains to be proven, the several similarities between this octapeptide and the general characteristics of mitochondrial presequences, including the presence of the overrepresented amino acids leucine, serine and arginine, the general hydrophobic nature of the peptide which is devoid of negatively charged amino acids and the potential to form an, albeit small, amphiphilic α-helix, are noteworthy (Johnson et al., 1990). Whether this implies that protein import into hydrogenosomes mimics, to some extent, the situation observed for the translocation of certain mitochondrial, chloroplast and/or nuclear proteins or whether the N-terminal octapeptide of hydrogenosomal ferredoxin serves a function independent of

protein translocation (similar to that proposed for select peroxisomal/glyoxysomal proteins possessing N-terminal presequences; section 1.4.3.3) remains to be determined. Related to this, Keller et al. (1991) have recently reported that hydrogenosomes do not contain proteins that cross-react with anti-microbody targeting signal antibody, suggesting a different mechanism(s) for protein targeting to hydrogenosomes, which appear to be unrelated to microbodies and may be the anaerobic equivalent of mitochondria.

1.4.3.3 Protein Targeting to Peroxisomes

As discussed in section 1.4.1, peroxisomal proteins are, in general, synthesized on free cytosolic polysomes at their mature size and posttranslationally imported into pre-existing organelles. The information necessary for targeting proteins to peroxisomes is therefore hypothesized to reside within the mature amino acid sequence, however, only limited information is presently available concerning sequences which are important for addressing proteins to peroxisomes.

Both *in vivo* and *in vitro* expression systems for cloned genes encoding peroxisomal proteins have been used to gain insight into the phenomenon of peroxisomal targeting. *In vivo* expression of firefly luciferase in monkey kidney (CV-1) cells revealed that localization of luciferase to peroxisomes (Keller et al., 1987) is the result of a peroxisomal targeting signal (PTS) present within the

extreme C-terminal twelve amino acids of this protein (Gould et al., 1987). Using the same system, Gould et al. (1988, 1989) subsequently identified PTSs (thought to be variants of the sequence ser-lys/his-leu) located at or near the C-terminus of five peroxisomal proteins (human catalase, porcine D-amino acid oxidase, rat acyl-coenzyme A oxidase, rat hydratase dehydrogenase and the 20 kDa peroxisomal membrane protein of Candida boidinii). Continued investigation into the nature of the luciferase PTS revealed that a conservative variation of the C-terminal tripeptide, ser/ala/cys-lys/his/arg-leu, is both necessary and sufficient for peroxisomal targeting, indicating that import into peroxisomes may be dependent on a conserved amino acid sequence (Gould et al., 1989). Interestingly, a search of several peroxisomal protein sequences (deduced from DNA sequence data) reveals that conserved variants of this tripeptide are common to many such proteins (Table 1.4.1). In addition, antibodies specific for the PTS of luciferase recognize multiple rat liver peroxisomal proteins, further demonstrating the conserved nature of this PTS (Gould et al., 1990b). This degree of conservation may be extended to at least one mechanism of protein translocation into this organelle since the import of firefly luciferase into peroxisomes occurs not only in mammals, but in yeast, plants and insects (Gould et al., 1990a). It should be noted, however, that certain peroxisomal proteins from different

Table 1.4.1. Occurrence of the conserved S/A/C-K/H/R-L tripeptide in select peroxisomal proteins.

Peroxisomal protein	Total no. amino acids	Conserved amino acids	COOH terminal	Internal location	Reference
rat acyl-CoA oxidase	661	659 661 ser-lys-leu	+	-	Miyazawa et al. (1987)
porcine D-amino acid oxidase	347	345 347 ser-his-leu	+	-	Jacobs et al. (1987)
S. cerevisiae citrate synthase 2	460	458 460 ser-lys-leu	+	-	Rosenkrantz et al. (1986); Lewin et al. (1990)
spinach glycolate oxidase	369	367 369 ala-arg-leu	+	-	Volokita and Somerville (1987)
rat hydratase dehydrogenase	722	720 722 ser-lys-leu	+	-	Osumi et al. (1985)
P. pyralis luciferase	550	548 550 ser-lys-leu	+	-	de Wet et al. (1987)
C. boidinii PMP20	167	165 167 ala-lys-leu	+	-	Garrard and Goodman (1989)
human/rat sterol carrier protein 2[a]	143	141 143 ala-lys-leu	+	-	Yamamoto et al. (1991); Morris et al. (1988)
rat uricase	303	301 303 ser-arg-leu	+	-	Alvares et al. (1989)
C. tropicalis acyl-CoA oxidase (PXP-4)	709	554 556 ser-lys-leu	-	+[b]	Okazaki et al. (1986)
S. cerevisiae/ C maltosa acyl-CoA oxidase	748 709	241 243 210 212 ala-arg-leu	- -	+[b] +[b]	Dmochowska et al. (1990); Hill et al. (1988)
H. polymorpha amine oxidase	692	681 683 ser-arg-leu	-	+	Bruinenberg et al. (1989)
rat carnitine octanoyltransferase	523	31 33 cys-arg-leu	-	+	Chatterjee et al. (1988)
C. tropicalis catalase (PXP-9)	485	331 333 ser-arg-leu	-	+	Okada et al. (1987)
human catalase	527	517 519 ser-his-leu	-	+[b]	Quan et al. (1986)
S. cerevisiae catalase A/ catalase T	515 562	350 352 346 348 ala-arg-leu	- -	+[b] +[b]	Cohen et al. (1988); Hartig and Ruis (1986)

C. tropicalis HDE[c]	906	29 31 ala-lys-ile	-	+[b]	Nuttley et al. (1988)
cotton isocitrate lyase	576	190 192 cys-lys-leu	-	+[b]	Turley et al. (1990)
rat 3-ketoacyl-CoA thiolase 1 and 2[a]	398 434	140 142 176 178 ser-arg-leu	- -	+[b] +[b]	Hijikata et al. (1987, 1990); Bodnar and Rachubinski (1990)
watermelon malate dehydrogenase	356	16 18 ala-his-leu	-	+	Gietl (1990)
H. polymorpha malate synthase	555	467 469 ser-arg-leu	-	+[b]	Bruinenberg et al. (1990)
H. polymorpha methanol oxidase	664	592 594 ala-arg-leu	-	+	Ledeboer et al. (1985)
human 3-oxoacyl-CoA thiolase[a]	424	166 168 ser-arg-leu	-	+[b]	Bout et al. (1988)
S. cerevisiae Pas1p	1043	294 296 ser-his-leu	-	+[b]	Erdmann et al. (1991)
CHO cell PMP35	305	218 220 ala-lys-leu	-	+	Tsukamoto et al. (1991)
C. boidinii PMP47	423	320 322 ser-lys-leu	-	+	McCammon et al. (1990b)
rat PMP70	659	65 67 ser-arg-leu	-	+[b]	Kamijo et al. (1990)
D. melanogaster xanthine dehydrogenase (rosy)	1335	56 58 ala-lys-leu	-	+	Keith et al. (1987)
human AGT1	392	NP	-	-	Purdue et al. (1990)
rat catalase	527	NP	-	-	Furuta et al. (1986)
H. polymorpha dihydroxyacetone synthase	702	NP	-	-	Janowicz et al. (1985)
C. tropicalis isocitrate lyase	549	NP	-	-	Atomi et al. (1990)
C. tropicalis nonspecific lipid-transfer protein (PXP-18)	127	NP	-	-	Szabo et al. (1989); Tan et al. (1990)

[a] The precursor polypeptide possesses an N-terminal presequence.
[b] There is more than one occurrence of a conserved tripeptide internally.
[c] The PTS is known to be the C-terminal sequence, ala-lys-ile (Aitchison et al., 1991).

organisms do not contain a variant of the ser/ala/cys-lys/his/arg-leu tripeptide at their extreme C-terminus, while others do not contain this tripeptide at any location within their sequence (Table 1.4.1). This suggests that the context of the tripeptide within the protein may be important to translocation (eg. surface exposure may be required to allow interaction with a putative receptor) and those cellular proteins that contain a 'buried' tripeptide are not targeted to peroxisomes (Gould et al., 1988; Borst, 1989). Alternatively, more than one type of PTS may act to direct proteins to peroxisomes (Borst, 1989). Consistent with these notions, certain deletions of the NH_2-terminus of luciferase abolish import independent of the presence of the C-terminal ser-lys-leu tripeptide (Gould et al., 1987). The *in vivo* expression and translocation of *C. tropicalis* fatty acyl-CoA oxidase into peroxisomes of the yeast *C. maltosa* similarly suggests that there are at least two regions of peroxisomal targeting information in this protein, with the primary information being internal (Kamiryo et al., 1989). Evidence for more than one type of PTS also stems from the fact that certain peroxisomal proteins possess cleavable N-terminal presequences (Table 1.4.1); for rat liver 3-ketoacyl-CoA thiolase, the 11 amino acid prepeptide comprises a novel PTS with the ability to direct a normally cytosolic protein, chloramphenicol acetyltransferase, to the peroxisomes of CV-1 cells (Bodnar, 1991; Swinkels et al.,

1991). Aitchison et al. (1991), who show that the extreme C-terminal tripeptide, ala-lys-ile, of C. tropicalis hydratase-dehydrogenase-epimerase (which is similar to, yet distinct from, the luciferase PTS) is essential for the in vivo targeting of this protein to peroxisomes of C. albicans and S. cerevisiae, further demonstrate that the targeting of proteins to peroxisomes is not simply the result of a universal topogenic signal, although the C-terminal targeting signal motif, and thus at least one aspect of the mechanism of import, appears to be conserved across distinct phylogenetic groups (Barns et al., 1991).

In vitro import studies have similarly postulated the existence of more than one type of peroxisomal translocation signal. Import of C. tropicalis fatty acyl-CoA oxidase into isolated C. tropicalis peroxisomes identified two internal targeting sequences of this protein (amino acids 1 to 118 and 309 to 427 of 709 residues in total), either of which can mediate import into peroxisomes (Small et al., 1988). By contrast, in vitro import into isolated rat liver peroxisomes (Miyazawa et al., 1989) suggests that the PTS of rat liver acyl-CoA oxidase resides in the C-terminal five (or fewer) amino acids, the last three of which are ser-lys-leu, similar to the luciferase PTS of Gould et al. (1989). Miyazawa et al. (1989) also suggest, however, that a C-terminal region of rat liver acyl-CoA oxidase plays an important role in efficient

import, possibly through a conformational effect.

Peroxisomal proteins have been shown to bind specifically to peroxisomes and not to mitochondria *in vitro*, supporting the notion that import is receptor-mediated (Imanaka et al., 1987; Small et al., 1987). Investigations concerned with the *in vitro* import of rat liver fatty acyl-CoA oxidase have further suggested that translocation into peroxisomes requires ATP hydrolysis but not a membrane potential (Imanaka et al., 1987). To the contrary, Bellion and Goodman (1987) found that proton ionophores prevent the import of alcohol oxidase into peroxisomes in *C. boidinii* spheroplasts, but admit that this may be due to the decrease in intracellular ATP concentrations caused by the ionophores, rather than a membrane potential requirement. Thus, energy appears to be necessary for protein import into peroxisomes, however, it is not known whether ATP is used to drive protein translocation or to unfold import-incompetent proteins (Pugsley, 1989).

Studies of cell lines derived from patients afflicted with various peroxisomal disorders and mutants defective in peroxisome biogenesis have yielded interesting observations concerning peroxisomal protein import. Zellweger syndrome fibroblasts have been shown to possess peroxisomal membrane 'ghosts' (ie. intact peroxisomal membranes lacking soluble matrix proteins), indicating that

the pathways for membrane and matrix protein import into peroxisomes may be different and the latter may be the primary defect in this disease (Santos et al., 1988). Complementation experiments with cell lines established from several peroxisomal disorders have led to the identification of at least six different complementation groups, and thus at least six different gene products, which are required for peroxisome biogenesis (Brul et al., 1988a, b; McGuinness et al., 1990). Characterization of peroxisome assembly (pas) mutants in S. cerevisiae identified PAS1 as a gene encoding an ≈117 kDa hydrophilic polypeptide (Pas1p), possessing two putative ATP-binding domains, which is required for peroxisome biogenesis (Erdmann et al., 1991). Pas1p may be a soluble component of the import machinery but its precise function remains to be determined (Erdmann et al., 1991). The investigation of similar mutants obtained in Chinese hamster ovary (CHO) cells allowed the identification of peroxisome assembly factor 1 (PAF-1), a 35 kDa peroxisomal membrane protein that restores peroxisome biogenesis in peroxisome-deficient CHO cells (Tsukamoto et al., 1991). Although its function was not determined experimentally, Tsukamoto et al. (1991) speculate that PAF-1 may be a component of a putative receptor or the transmembrane import machinery of peroxisomes. A cytoplasmic component essential to peroxisome biogenesis in CHO cells has been reported elsewhere, however, its identity was not established (Allen

et al., 1989).

1.5 SCOPE OF THE PROJECT

Prior to the onset of this project, preliminary studies concerned with the biogenesis of peroxisomes in both yeast and mammalian systems were suggestive of the following generalities (reviewed by Lazarow and Fujiki, 1985): 1) peroxisomal proteins are synthesized in the cytosol from mRNAs encoded by the nuclear genome; 2) peroxisomal proteins, including a 22 kDa integral membrane protein (Fujiki et al., 1984), are synthesized exclusively on free polyribosomes and posttranslationally imported into pre-existing organelles; 3) distinct from the biogenesis of other organelles for which proteins are synthesized as larger precursors possessing cleavable transit peptides, peroxisomal proteins are, with a few noted exceptions (see Lazarow and Fujiki, 1985), synthesized at their mature size. Thus, the information necessary for the targeting of proteins specifically to peroxisomes was (and still is) hypothesized to reside within their mature amino acid sequence. Knowledge of the potential targeting sequences of peroxisomal proteins was, however, very limited. Within this framework, the objectives of this project were: 1) subsequent to the construction of yeast genomic DNA and/or cDNA libraries, subclone and sequence at least one full-length cDNA and/or gene encoding a peroxisomal protein; 2)

construct a series of overlapping deletion mutants of this cDNA/gene; 3) develop an *in vitro* and/or *in vivo* expression and translocation system for the full-length and truncated polypeptides encoded by these deletion mutants to allow the amino acid sequence(s) comprising the peroxisomal addressing signal of this protein to be determined.

For several reasons, the yeast *Candida tropicalis* pK233 was a favourable organism for this investigation. Peroxisome proliferation and the activity of the peroxisomal β-oxidation system in this yeast are strikingly induced simply by growth in medium supplemented with *n*-alkanes or oleic acid (Osumi et al., 1974, 1975; Tanaka et al., 1982); these growth conditions induce twelve abundant mRNAs, nine of which are thought to encode peroxisomal proteins (Fujiki et al., 1986). Conversely, when *C. tropicalis* is grown on glucose as the sole source of carbon, peroxisomes are few in number and the β-oxidation activity is low or undetectable (Tanaka et al., 1982; Dommes et al., 1983). In addition, even though peroxisomes can be induced in mammalian (eg. rat liver) systems, they cannot be repressed to the same extent as is possible with *C. tropicalis* (Reddy et al., 1982). These qualities, and the associated relative ease of peroxisome isolation and purification from this yeast (Kamiryo et al., 1982), have allowed *C. tropicalis* to serve as a useful organism for previous studies related to the characterization of peroxisome structure, function and

morphology, the cloning and sequencing of cDNAs and/or genes encoding peroxisomal proteins and the biogenesis of peroxisomes in general (reviewed in Tolbert, 1981; Kindl and Lazarow, 1982; de Duve, 1983; Lazarow and Fujiki, 1985). Thus, C. tropicalis pK233 represented a sufficiently well characterized system amenable to further investigations of this nature.

Efforts made to attain the abovementioned objectives, using Candida tropicalis pK233 as the system for study, resulted in the following: having achieved the cloning, sequencing and characterization of the gene encoding C. tropicalis fatty acyl-CoA oxidase, and both the full-length cDNA and gene encoding C. tropicalis catalase, a series of N-terminal, C-terminal, N-/C-terminal and internal deletion mutants of the catalase gene was constructed. Subsequent to the expression of these mutants both in vitro and in vivo, efforts were made to establish competent systems for the translocation of full-length and truncated catalase polypeptides into isolated C. tropicalis peroxisomes in vitro and the peroxisomes of S. cerevisiae in vivo. The rationale for the approach used in each of these efforts is presented at the beginning of the appropriate section.

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Chemicals and Reagents

The following sources were used for critical reagents:

ampicillin	Sigma Chemical Company (St. Louis, Missouri, USA)
antipain (microbial source)	Sigma
aprotinin (bovine lung)	Boehringer Mannheim Canada Ltée (Dorval, Quebec)
bacto-agar	Difco Laboratories (Detroit, Michigan, USA)
bacto malt extract	Difco
bacto-peptone	Difco
bacto tryptone	Difco
bacto yeast extract	Difco
bacto yeast nitrogen base (w/o amino acids)	Difco
Bio-Rad protein assay dye reagent concentrate	Bio-Rad Laboratories (Canada) Ltd. (Mississauga, Ontario)
Bluo-gal (halogenated indolyl-β-D-galactoside)	Gibco/BRL Canada (Burlington, Ontario)
bovine serum albumin (BSA) - Cohn fraction V	Sigma
bovine serum albumin (BSA) - DNase-free	Pharmacia (Canada) Inc. (Baie D'Urfé, Quebec)
chymostatin (microbial source)	Sigma
corn steep liquor	Sigma
cytochrome-c (horse heart)	Sigma
n-decane	Sigma
deoxycholic acid	Sigma
deoxyribonucleotides (dA/dG/dC/dTTP)	Pharmacia
diguanosine triphosphate [G(5')ppp(5')G]	Pharmacia
n-dodecane	Sigma
$(dT)_{12-18}$	Collaborative Research Incorporated (Bedford, Massachusetts, USA)
Ficoll 400	Pharmacia
Freund's adjuvant (complete)	Sigma
Gigapack Gold in vitro	Stratagene Cloning Systems

packaging extract	(La Jolla, California, USA)
hemoglobin (bovine blood)	Sigma
hydrogen peroxide (30%)	Sigma
IPTG (isopropyl β-D-thiogalactoside)	Gibco/BRL
leupeptin	Sigma
M13 cloning system	Gibco/BRL
M13 sequencing system	Gibco/BRL
$m^7G(5')ppp(5')G$	Pharmacia
molecular weight standards for	
(a) DNA:	
(i) 1 kbp DNA ladder (75-12,216 bp)	Gibco/BRL
(ii) φX174-RF DNA Hae III digest (72-1,358 bp)	Pharmacia
(iii) Lambda DNA - Hind III fragments (125 - 23,100 bp)	Gibco/BRL
(iv) DNA size standards - high range (125 - 23,130 bp)	Bio-Rad
(v) DNA size standards - low range (42-1,746 bp)	Bio-Rad
(b) protein:	
(i) Dalton Mark VII-L molecular weight markers for SDS-PAGE (14,000-66,000 Da)	Sigma
(ii) β-galactosidase (*E. coli*; 116,000 Da)	Sigma
(iii) phosphorylase b (rabbit muscle; 97,400 Da)	Sigma
(iv) prestained SDS-PAGE molecular weight standards (low range ≈15,000-110,000 Da; high range ≈45,000-210,000 Da)	Bio-Rad
nick translation reagent kit	Gibco/BRL
nicotinamide adenine dinucleotide (NAD)	Sigma
nitrocellulose (pore size - 0.45 μm)	Schleicher & Schuell, Inc. (Keene, New Hampshire, USA)
Nonidet P-40 (NP-40)	Sigma
Nycodenz [5-(N-2,3-dihydroxypropylacetamido)-2,4,6-tri-iodo-N,N'-bis(2,3-dihydroxypropyl)isophthal-	Nycomed AS Diagnostica (Oslo, Norway)

amide]	
oleic acid (clear)	Fisher Scientific (Unionville, Ontario)
oligo(dT)-cellulose/type 3	Collaborative Research
ovalbumin	Miles Laboratories (Pty) Ltd. (Rep. of S. Africa)
Pansorbin (Staphylococcus aureus cells, inactivated)	Calbiochem Corporation (San Diego, California, USA)
pepstatin A (microbial source)	Sigma
PMSF (phenylmethylsulfonyl fluoride)	Sigma
polyethylene glycol 3350 (formerly 4000)	Sigma
polyethylene glycol 8000 (formerly 6000)	BDH Inc. (Toronto, Ontario)
polyoxyethylene 23-lauryl ether (Brij 35)	Sigma
polyvinylpyrrolidone	Sigma
rabbit reticulocyte lysate (nuclease-treated)	Promega Corporation (Madison, Wisconsin, USA)
random primers DNA labeling system	Gibco/BRL
ribonucleotides (A/G/C/UTP)	Pharmacia; Promega
RNA guard RNase inhibitor	Pharmacia
RNasin RNase inhibitor	Promega
Sephadex G-25 (medium; fine)	Pharmacia
Sephadex G-50 (medium)	Pharmacia
Sequenase/Version 2.0 DNA sequencing kit	United States Biochemical Corporation (Cleveland, Ohio, USA)
Spectra/Por molecular porous membrane tubing (molecular weight cutoff 12,000-14,000)	Fisher Scientific
spermidine	Sigma
tetracycline	Sigma
titanium oxysulfate hydrate	Aldrich Chemical Company, Inc. (Milwaukee, Wisconsin, USA)
TLCK (Nα-p-tosyl-L-lysine chloromethyl ketone)	Sigma
n-tridecane	Sigma
Triton X-100	BioRad
tRNA (calf liver)	Boehringer Mannheim
Tween 40 (polyoxyethylene-sorbitan monopalmitate)	Sigma
Tween 80 (polyoxyethylene-sorbitan monooleate)	Sigma
n-undecane	Sigma
wheat germ (unprocessed)	General Mills Inc.

	(Vallejo, California, USA; Minneapolis, MN, USA)
X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside)	Gibco/BRL

All other chemicals and reagents were of the highest quality available.

2.1.2 Radiochemicals

^{14}C-methylated β-galactosidase (2.6 μCi/mg protein)	Sigma
^{14}C-methylated phosphorylase b (5.3 μCi/mg protein)	Sigma
^{14}C-methylated proteins (molecular weight range 14,000-70,000 Da) (5-50 μCi/mg proteins)	Sigma
^{125}I-protein A (>30 mCi/mg total protein A, 0.1 μCi/μl)	Amersham Canada Limited (Oakville, Ontario)
[γ-^{32}P]ATP(3,000 Ci/mmol, 10 μCi/μl)	Amersham; Dupont/NEN Canada Inc. (Mississauga, Ontario)
[α-^{32}P]dATP(3,000 Ci/mmol, 10 μCi/μl)	Amersham; ICN Biomedicals, Inc. (Irvine, California, USA)
[α-^{32}P]dCTP(410 Ci/mmol, 10 μCi/μl)	Amersham
L-[^{35}S]-methionine (1,100 Ci/mmol, 10 μCi/μl)	Dupont/NEN
tran^{35}S-label (1,200 Ci/mmol, 10 μCi/μl)	ICN

2.1.3 Enzymes

Alkaline phosphatase (calf intestinal mucosa)	Pharmacia
Creatine kinase (rabbit muscle)	Boehringer Mannheim
DNA ligase (*E. coli*)	Amersham
DNA polymerase I (*E. coli*)	Pharmacia
DNA polymerase I Klenow fragment (*E. coli*)	New England Biolabs, Inc. (Beverly, Massachusetts, USA); Pharmacia; Gibco/BRL
DNase I (bovine pancreas)	Sigma
Exonuclease III (*E. coli*)	Gibco/BRL
β-glucuronidase (*Helix promatia*)	Sigma (G 2887)
Lysozyme (egg white)	Pharmacia
Micrococcal nuclease	Pharmacia

(Staphylococcus aureus)	
Modified T7 DNA polymerase/ Sequenase	USBC
Mung bean nuclease (mung bean sprouts)	NEB
Nuclease S1 (Aspergillus oryzae)	Pharmacia
Pronase protease (Streptomyces griseus)	Calbiochem
Proteinase K (Tritirachium album)	Boehringer Mannheim
Restriction endonucleases	Gibco/BRL; NEB; Pharmacia
Reverse transcriptase (avian myeloblastosis virus)	Pharmacia
RNase I "A" (bovine pancreas)	Pharmacia
RNase H (E. coli)	Amersham
SP6 RNA polymerase	Gibco/BRL; NEB; Promega
T4 DNA ligase	Gibco/BRL; Pharmacia
T4 polynucleotide kinase	Amersham; Pharmacia
T7 RNA polymerase	Pharmacia; NEB; Promega
Terminal deoxynucleotidyl transferase (calf thymus)	Gibco/BRL
Zymolyase-100T (Arthrobacter luteus)	Seikagaku Kogyo Co. Ltd. (Tokyo, Japan); ICN ImmunoBiologicals (Lisle, Illinois, USA)

In general, enzymes were used according to the manufacturers' product data sheets. Section 2.2 presents any deviations from these prescribed methods.

2.1.4 Antisera

Rabbit anti-C. tropicalis catalase serum was prepared and characterized in our laboratory by John D. Aitchison. This entailed the isolation of purified catalase by preparative SDS-PAGE (section 2.3.2) of sucrose density gradient-purified peroxisomes (obtained from oleic acid-grown C. tropicalis), followed by excision and electroelution of the polypeptide. The electroeluted

polypeptide was extensively dialyzed against 2 x 1 L of 50 mM ammonium bicarbonate, lyophilized, resuspended in 0.1 ml of distilled H_2O and emulsified with an equal volume of Freund's complete adjuvant. The antigen-adjuvant mixture (0.2 ml containing ≈100 µg catalase) was directly injected into the popliteal lymph nodes of both hind legs of an anaesthetized New Zealand White rabbit (Goudie et al., 1966). Thereafter, boosts were administered at ≈2-week intervals by 3 subcutaneous injections of 200 µg total polypeptide. Two weeks after the final boost, the rabbit was bled and the serum tested for antibody specificity to C. tropicalis catalase by immunoblot analysis (section 2.3.3).

Rabbit anti-firefly luciferase serum was a gift from Dr. S.J. Gould of Dr. S. Subramani's laboratory, Department of Biology, University of California, La Jolla, California, USA.

Rabbit anti-E. coli chloramphenicol acetyltransferase (CAT) was a gift from Dr. J.P. Capone, Department of Biochemistry, McMaster University, Hamilton, Ontario.

2.1.5 Cloning Vectors and Host Strains

pBR322 (Sutcliffe, 1978), used for the construction of a genomic library of C. tropicalis DNA, and PstI-digested dG-tailed pBR322, used for the construction of cDNA libraries of C. tropicalis poly(A)$^+$RNA, were purchased from Gibco/BRL. Host cells for these libraries, library

efficiency E. coli HB101 and E. coli RR1 competent cells, respectively, were also obtained from Gibco/BRL. The M13mp18/19 replicative form (RF) DNA vectors (Norrander et al., 1983) used for subcloning genomic DNA encoding C. tropicalis fatty acyl-CoA oxidase and cDNA encoding catalase were purchased from Gibco/BRL as an M13 cloning system kit (M13mp18 RF DNA, M13mp19 RF DNA, T4 DNA ligase, ligase dilution and reaction buffer, bacterial alkaline phosphatase, Bluo-gal, IPTG, E. coli strain JM101 and JM107 host cells). An additional M13mp18/19 host strain, E. coli JM109, was a gift from Dr. J.P. Capone, Department of Biochemistry, McMaster University, Hamilton, Ontario. The lambda EMBL3 vector (Frischauf et al., 1983, 1987), used to construct a second genomic library of C. tropicalis DNA, and the appropriate host strains, E. coli NM539 and E. coli LE392 cells, were purchased from Promega. The expression vectors, pGEM-5Zf(+) and pGEM-7Zf(+), which were used to subclone and sequence genomic DNA encoding catalase, as well as in vitro transcribe full-length and truncated catalase mRNA, were obtained from Dr. J.P. Capone (original source was Promega). The E. coli MV1193 host cells that were transformed with various pGEM-5Zf/7Zf(+) constructs were obtained from Dr. A.B. Futcher, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA (original source was Dr. J. Vieira, Department of Biochemistry,

University of Minnesota, Minneapolis, Minnesota, USA). An alternative expression vector, pSP65, which was used for the *in vitro* expression of dihydrofolate reductase (DHFR), was purchased from Promega. The yeast expression vector, pMK22, and accompanying host, *C. albicans* strain SGY243 (Kurtz et al., 1987; Kelly et al., 1987), used for the *in vivo* expression of catalase constructs, were obtained from Dr. M.B. Kurtz of Merck, Sharp and Dohme Research Laboratories, Rahway, New Jersey, USA. An alternative vector, YEp13 (Broach et al., 1979), used for the *in vivo* expression of catalase constructs in the yeast *Saccharomyces cerevisiae* strain DL1 (Veenhuis et al., 1987), was also obtained from Dr. A.B. Futcher. Vectors used for routine subcloning, pUC18/19 (Yanisch-Perron et al., 1985) and pUC118/119 (Vieira and Messing, 1987), were purchased from Pharmacia and obtained from Dr. A.B. Futcher (original source was Dr. J. Vieira), respectively. Routine transformations were carried out using subcloning efficiency *E. coli* DH5α competent cells purchased from Gibco/BRL. Growth media for transformed *E. coli* cells included LB [Maniatis et al., 1982; 1% (w/v) bacto tryptone, 0.5% (w/v) bacto yeast extract, 1% (w/v) NaCl, pH 7.5], YT [Schleif and Wensink, 1981; 0.8% (w/v) bacto tryptone, 0.5% (w/v) bacto yeast extract, 0.5% (w/v) NaCl] or S.O.C. [Hanahan, 1983; 2% (w/v) bacto tryptone, 0.5% (w/v) bacto yeast extract, 10 mM NaCl,

2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose] supplemented with ampicillin (100 µg/ml) or tetracycline (12.5 µg/ml) where appropriate. For plating transformed E. coli cells, selective media were solidified with 1.5% (w/v) bacto-agar. E. coli cultures were incubated at 37°C with or without shaking (≈225 rpm) as appropriate.

2.1.6 In Vivo/Vitro Expression Constructs

The pPGK and YEp13PGK-luciferase constructs (Gould et al., 1990a) used for the in vivo expression of full-length or truncated catalase proteins and luciferase, respectively, in the yeast S. cerevisiae strain DL1, were a gift from Dr. S.J. Gould of Dr. S. Subramani's laboratory, Department of Biology, University of California, La Jolla, California, USA. Plasmid TP408, which was used for the in vitro expression of E. coli CAT, was obtained from Dr. D.W. Andrews, Department of Biochemistry, McMaster University, Hamilton, Ontario. Plasmid pRP1, from which the gene encoding DHFR was excised and subsequently cloned into pSP65 to allow in vitro expression (using plasmid pSP65DHFR-1), was obtained from Dr. K.B. Freeman, Department of Biochemistry, McMaster University, Hamilton, Ontario (original source was Dr. G.A. Reid, Department of Microbiology, School of Agriculture, Edinburgh, Scotland, UK).

2.1.7 Yeast Strains

Candida tropicalis Berkhout strain pK233 (Tanabe et al., 1966) was originally purchased from the American Type Culture Collection (ATCC 20336), Rockville, Maryland, USA. Candida albicans strain SGY243 (Kelly et al., 1987; genotype: ade2/ade2 Δura3::ADE2/Δura3::ADE2; relevant phenotype: ADE^+ URA^-; denomination as per the Squibb Culture Collection) was provided by Dr. M.B. Kurtz of Merck, Sharp and Dohme Research Laboratories, Rahway, New Jersey, USA. Saccharomyces cerevisiae strain DL1 (Veenhuis et al., 1987; relevant phenotype: HIS^- LEU^- URA^-) was provided by Dr. S.J. Gould of Dr. Suresh Subramani's laboratory, Department of Biology, University of California, La Jolla, California, USA.

2.1.8 Oligonucleotide Probes/Primers

The oligonucleotides listed in Table 2.1.1 were synthesized using an Applied Biosystems 381A Automated DNA Synthesizer (in the laboratory of Dr. H.P. Ghosh, Department of Biochemistry, McMaster University, Hamilton, Ontario), obtained from the Central Facility of the Institute of Molecular Biology and Biotechnology, McMaster University, Hamilton, Ontario, or in the case of universal primers and $(dT)_{12-18}$, purchased from commercial sources (Pharmacia, Gibco/BRL, USBC, Collaborative Research).

Table 2.1.1. List of oligonucleotides used as sequence-specific probes and/or primers.

Probe/Primer	Strand (+)or(-)	Oligonucleotide Sequence	Annealing Positions
AOx A1	(-)	5'-AACCAGTGTATTCCTT-3'	1498 - 1483
A2	(-)	5'-TATCTCAAGGCAATGG-3'	962 - 947
A4	(-)	5'-CATTGGTCAATGGCTT-3'	1280 - 1265
A7	(-)	5'-GGTGTCGTTCTTTTCA-3'	1175 - 1160
A8	(-)	5'-ATCTTGGGATTCTGAC-3'	802 - 787
A30	(-)	5'-GAGCAGCACCACCAA-3'	601 - 587
A31	(-)	5'-CAATCTGGCGTAGACA-3'	636 - 621
A33	(-)	5'-AACCGAGGTTGACAC-3'	364 - 350
A34	(-)	5'-CCTTTTGGATGGATGA-3'	67 - 52
A5	(+)	5'-GTGCCCTCAAGGTTGA-3'	1102 - 1117
A6	(+)	5'-ACCAAAAGAGATTGTT-3'	1052 - 1067
A9	(+)	5'-ATCAGATTATCCCAAG-3'	1582 - 1597
A10	(+)	5'-TGACTTTCAACGTTCC-3'	1808 - 1823
A11	(+)	5'-TCACAACAATGTCTTG-3'	1371 - 1386
A32	(+)	5'-CCATGACTGCTTTGGC-3'	1835 - 1850
A35	(+)	5'-AAATTCCAAAGTTGT-3'	1859 - 1873
Cat C1	(+)	5'-TCCATGCCTCTGCTG-3'	1337 - 1351
C2	(+)	5'-AATCTTGGAATTGTCTCCAA-3'	1431 - 1450
C3	(-)	5'-CTTCCAGTGGAATGGAGTAG-3'	1242 - 1223
C188	(-)	5'-TTGAAATCAGCTGGGGTAGC-3'	1262 - 1243
C189	(-)	5'-TCGAAGTGAGCCAAGGAGTC-3'	128 - 109
C231	(-)	5'-CAAAGGACCGTGTTGGCCAA-3'	84 - 65
UP (M13/pUC/pGEM)			
forward primer (-20)	-	5'-GTAAAACGACGGCCAGT-3'	(vector-specific)
forward primer (-40)	-	5'-GTTTTCCCAGTCACGAC-3'	(vector-specific)
reverse primer	-	5'-CAGGAAACAGCTATGAC-3'	(vector-specific)
(dT)12-18	(-)	5'-TTTTTTTTTTTT(TTTTTT)-3'	....

Primers/probes complementary to the DNA sequences encoding C. tropicalis fatty acyl-CoA oxidase (AOx) and catalase (Cat) are listed below AOx and Cat subheadings, respectively. UP denotes those universal primers used for sequencing. Strand complementarity is indicated by (+) or (-) for complementarity to the non-coding (template) or coding (non-template/mRNA) strands, respectively. Numbers indicative of oligonucleotide annealing positions represent sequences along the coding (non-template/mRNA) strand. Underlined nucleotides indicate mismatches with the annealing sequence.

2.2 PEROXISOME INDUCTION AND ISOLATION

2.2.1 Yeast Cell Culture

C. tropicalis pK233 was routinely maintained at 30°C on 3% (w/v) bacto malt extract:1.5% (w/v) bacto-agar slants. To obtain starting suspension cultures (40 ml) of C. tropicalis, cells were looped from slants into 3% (w/v) bacto malt extract medium and grown overnight at 30°C with shaking (≈180 rpm). To induce peroxisome proliferation, cells from starting cultures were pelleted by centrifugation at maximum speed (4,000 rpm) for 2 to 3 min in an IEC table-top clinical centrifuge (model CL; International Equipment Company, Needham Hts., Massachusetts, USA), washed twice with room-temperature, sterile, distilled H_2O, transferred into 1 L of YPB medium [0.3% (w/v) bacto yeast extract, 0.5% (w/v) bacto-peptone, 0.5% (w/v) K_2HPO_4, 0.5% (w/v) KH_2PO_4, 1% (w/v) Brij 35] or YPBO medium [YPB medium plus 1% (w/v) oleic acid] and grown at 30°C for 14 to 18 h with shaking.

C. albicans SGY243 was maintained at 30°C on complete medium (YEPD) plates [2% (w/v) bacto-peptone, 1% (w/v) bacto yeast extract, 2% (w/v) glucose, 2% (w/v) bacto-agar]. SGY243 transformants, selected initially by plating on C. albicans rescue plates [1 M sorbitol, 0.67% (w/v) bacto yeast nitrogen base without amino acids, 2% (w/v) glucose, 2% (w/v) bacto-agar], were maintained on minimal medium (YNBD) plates [0.67% (w/v) bacto yeast nitrogen base

without amino acids, 2% (w/v) glucose, 2% (w/v) bacto-agar]. Starting suspension cultures (10 ml) of wild-type C. albicans and transformants were obtained by looping cells into complete and minimal medium, respectively, followed by overnight incubation at 30°C in a rotating wheel. As described for C. tropicalis, starting cultures were pelleted, washed, transferred into 1 L of YPBO medium and grown at 30°C for 14 to 18 h with shaking to induce peroxisome proliferation.

S. cerevisiae DL1 was maintained at 30°C on complete medium plates, however, DL1 transformants were selected and maintained at 30°C on minimal medium plates supplemented with L-histidine (20 µg/ml) and uracil (20 µg/ml) to satisfy auxotrophic requirements. To obtain starting suspension cultures (10 ml) of wild-type S. cerevisiae and transformants, cells were looped from plates into complete medium and minimal medium supplemented with 20 µg/ml L-histidine and uracil, respectively, and grown overnight at 30°C in a rotating wheel (cells were not allowed to reach saturation without subculturing). Peroxisome proliferation in S. cerevisiae was induced by transfer (post-pelleting and washing) of overnight non-saturated starting cultures into 500 ml of induction medium (IM) [0.67% (w/v) bacto yeast nitrogen base without amino acids, 0.5% (w/v) bacto yeast extract, 0.5% (w/v) bacto-peptone, 0.1% (w/v) glucose, 1%

(w/v) oleic acid, 0.5% (w/v) polyoxyethylenesorbitan monopalmitate (Tween 40), L-histidine (20 µg/ml), uracil (20 µg/ml)] followed by growth at 30°C for 16 to 20 h with shaking.

Incubation periods for peroxisome proliferation were such that cells would be harvested during mid- to late-logarithmic phase of growth (2 to $3x10^7$ cells/ml) for each yeast strain. Between experiments, slants or plates of cells were stored at 4°C; however, actively-growing cells (those which were restreaked and grown overnight at 30°C) were used to begin each experiment. Yeast strains were maintained at 4°C for no greater than 3 to 4 weeks without restreaking onto fresh slants or plates.

2.2.2 Yeast Cell Fractionation and Peroxisome Purification

Peroxisomes were purified from C. tropicalis, C. albicans and S. cerevisiae by differential centrifugation followed by equilibrium sucrose or Nycodenz density gradient centrifugation as previously described (Kamiryo et al., 1982; Small et al., 1987), with modifications as detailed below.

2.2.2.1 Preparation of a Particulate Fraction

C. tropicalis and C. albicans cells, grown to mid- to late-logarithmic phase in YPB or YPBO medium, were pelleted by centrifugation at 4,000 x g_{max} in a Sorvall GSA

rotor (5,000 rpm; Dupont/NEN) or, alternatively, in a Beckman JA10 rotor (4,750 rpm; Beckman Instruments, Inc., Palo Alto, California, USA) for 10 min at 20°C. Cell pellets were resuspended and washed twice with room temperature, distilled H_2O and pelleted by centrifugation as before. Washed pellets (7 to 8 g wet weight of C. tropicalis from 1 L of YPB medium, 11 to 12 g wet weight of C. tropicalis and C. albicans from 1 L of YPBO medium) were resuspended in 4 ml per g of original cells of zymolyase digestion buffer [0.5 M KCl, 5 mM 3-(N-morpholino)propane-sulfonic acid (MOPS) (pH 7.2), 10 mM Na_2SO_3] containing 0.25 mg/ml of Zymolyase-100T and subsequently converted to spheroplasts by incubation at 30°C for 40 to 50 min with very gentle rotation (<60 rpm). Further manipulations, including density gradient purification of peroxisomes, were conducted between 0 and 4°C. Spheroplasts were pelleted by centrifugation at 2,300 x g_{max} in a Sorvall HB-4 rotor (3,750 rpm; Dupont/NEN) or in a Beckman JA-20 rotor (4,200 rpm; Beckman Instruments) for 8 min. Spheroplast pellets were resuspended in 3 ml of F buffer [5% (w/v) Ficoll 400, 0.6 M sorbitol, 2.5 mM MOPS (pH 7.2), 1 mM EDTA (pH 7.2)] per g of original cells, protease inhibitors [leupeptin (5.6 µg/ml), chymostatin (2.8 µg/ml), antipain (5.6 µg/ml), pepstatin (2.8 µg/ml), aprotinin (170 Kallikrein inhibitor units/ml), Nα-p-tosyl-L-lysine chloromethyl ketone (TLCK)

(90 µg/ml) and phenylmethylsulfonyl fluoride (PMSF) (80 µg/ml; freshly made)] were added and spheroplasts broken by 10 up-and-down homogenization strokes using a Cole-Parmer lab stirrer (model 4376-00; Cole-Parmer Instrument Company, Chicago, Illinois, USA), speed setting 3.4, equipped with a piston-type Teflon pestle homogenizer and borosilicate glass grinding vessel (sizes B and BB; clearance 0.13-0.18 mm; Thomas Scientific, Swedesboro, New Jersey, USA). The homogenate was centrifuged at 1,000 x g_{max} in the HB-4 (2,500 rpm) or JA-20 (3,000 rpm) rotor for 10 min. The supernatant was kept on ice and the pellet resuspended in 1.5 ml of F buffer per g of original cells and homogenized as before. The homogenate obtained was centrifuged as above and the combined supernatant (called the post-nuclear supernatant or PNS) further centrifuged at 20,000 x g_{max} in the HB-4 (11,250 rpm) or JA-20 (13,000 rpm) rotor for 20 min. The resulting supernatant (called the 20k x g supernatant) was withdrawn and the resulting particulate fraction/pellet (called the 20k x g pellet) resuspended in 0.5 ml of F buffer per g of original cells.

The preparation of a particulate fraction (20k x g pellet) for S. cerevisiae differed from that described for C. tropicalis and C. albicans in the following respects. Washed pellets (3 to 4 g wet weight from 500 ml of induction medium) were resuspended in 10 ml of zymolyase digestion

buffer containing 0.25 mg/ml of Zymolyase-100T. Spheroplasts and pellets therefrom were obtained in the usual manner; however, these pellets were subsequently resuspended in 10 ml of S. cerevisiae homogenization buffer [Erdmann et al., 1989; 5 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer (pH 6.0), 0.5 mM EDTA, 0.6 M sorbitol, 1 mM PMSF (freshly made)] and broken by one session of 15 up-and-down homogenization strokes. The homogenate was centrifuged as above and the supernatant withdrawn and similarly recentrifuged. Subsequent to the 20,000 x g_{max} spin, the resultant 20k x g supernatant was removed and retained; the resultant 20k x g pellet was resuspended in 0.2 ml of S. cerevisiae homogenization buffer.

2.2.2.2 Sucrose Density Gradient Centrifugation

Peroxisomes were further purified by isopycnic sucrose density gradient centrifugation by layering the 20k x g pellets isolated from YPB- or YPBO-grown C. tropicalis or C. albicans onto discontinuous gradients consisting of 4.67, 7, 14 and 7 ml of 25, 35, 42 and 53% (w/w) sucrose solutions, respectively, containing 2.5 mM MOPS (pH 7.2), 0.5 mM EDTA (pH 7.2) and 0.1% (v/v) absolute ethanol. Alternatively, C. tropicalis 20k x g pellets were fractionated on continuous (linear) sucrose gradients constituted by 15 ml each of 35 and 57% (w/w) sucrose solutions which were mixed and gently poured (dropwise) from

a BRL gradient former (model 150; Gibco/BRL) onto 3 ml of a 66% (w/w) sucrose solution cushion; all sucrose solutions contained 2.5 mM MOPS (pH 7.2), 0.5 mM EDTA (pH 7.2) and 0.1% (v/v) absolute ethanol. Both continuous and discontinuous sucrose gradients were centrifuged at 84,300 x g_{av} in a Beckman VTi50 rotor (32,000 rpm; Beckman Instruments) for 70 min at 4°C. Gradients were punctured at the top, a 50 µl glass capillary tube inserted to the bottom of the gradient and fractions (≈2 ml) collected using a Buchler Instruments polystaltic pump (Buchler Instruments, Fort Lee, New Jersey, USA). When only a purified peroxisomal (or mitochondrial) fraction was required, the centrifuge tube was punctured at the side, below the banded organelle(s), and the band withdrawn using a 10 ml syringe outfitted with a 16-gauge needle.

2.2.2.3 Nycodenz Density Gradient Centrifugation

In addition to sucrose density gradient centrifugation, peroxisomes from YPB-grown C. tropicalis were also purified by isopycnic Nycodenz density gradient centrifugation, essentially as previously described (Small et al., 1987; Nuttley et al., 1990). For this purpose, 20k x g pellets were layered onto discontinuous gradients composed of 4, 8, 12 and 6 ml of 17, 23, 28 and 35% (w/v) Nycodenz solutions, respectively, containing either buffer A [5% (w/v) Ficoll 400, 0.6 M sorbitol, 2.5 mM MOPS (pH 7.2),

0.5 mM EDTA (pH 7.2), 0.1% (v/v) absolute ethanol] or buffer B [5% (w/v) Ficoll 400, 0.6 M sorbitol, 5 mM Tris(hydroxymethyl)aminomethanehydrochloride (Tris-HCl) (pH 7.5), 3 mM KCl, 0.3 mM EDTA (pH 7.5), 0.1% (v/v) absolute ethanol]. Gradients were centrifuged at 101,000 x g_{av} in a Beckman VTi50 rotor (35,000 rpm; Beckman Instruments) for 70 min at 4°C. Fractions, or purified peroxisomes (or mitochondria) alone, were collected in a manner similar to that described for sucrose density gradients (section 2.2.2.2).

Nycodenz density gradient centrifugation was also used to purify peroxisomes from IM-grown S. cerevisiae. To this end, 20k x g pellets were layered onto discontinuous gradients consisting of 0.2, 0.5, 0.75 and 0.3 ml of 17, 28, 35 and 50% (w/v) Nycodenz solutions, respectively, containing 5 mM MES buffer (pH 6.0), 0.6 M sorbitol, 0.5 mM EDTA and 0.1% (v/v) absolute ethanol. Gradients were centrifuged at 122,000 x g_{av} in a Beckman TLV-100 rotor (60,000 rpm; Beckman Instruments) for 60 min at 4°C and fractions (0.13 ml) collected from the bottom of the gradient as described above.

2.2.2.4 Analysis of Gradient Fractions: Assay of Enzymes

Each of the fractions isolated from the sucrose (section 2.2.2.2) or Nycodenz gradients (section 2.2.2.3) was analyzed for its refractive index, protein concentration

and catalase and cytochrome c oxidase activities. To this end, the refractive index of 5 µl of each fraction was determined using an Abbe 3L tabletop refractometer (Bausch and Lomb, Rochester, New York, USA), protein concentrations were determined according to the method of Bradford (1976) (section 2.3.1) and catalase and cytochrome c oxidase activities were determined as described below.

2.2.2.4.1 Catalase Assay

Subcellular fractions (of no greater than 50 µl) were incubated in the presence of 1% (w/v) Triton X-100 (100 µl final volume) for 2 to 3 min at 0°C. Subsequently, 1 ml of ice-cold catalase substrate solution [20 mM imidazole buffer, 0.1% (w/v) bovine serum albumin and 0.01% (v/v) H_2O_2] was added and the incubations continued for 10 to 15 min at 0°C. Reactions were terminated by the addition of 1 ml of a saturated solution of $TiOSO_4$ [titanium (IV) oxysulfate] in 2N H_2SO_4 (kept at room temperature). After ≈15 min, the remaining H_2O_2 was determined spectrophotometrically (OD_{410}) as a yellow (peroxy-titanium sulfate) complex (Baudhuin et al., 1964). Catalase latency was established by comparison of the catalase activities exhibited by subcellular fractions incubated in the presence, versus absence, of 1% (w/v) Triton X-100.

2.2.2.4.2 Cytochrome c Oxidase Assay

Initially, 2.5 ml of horse heart cytochrome c solution [0.38 mg/ml in 0.03 M NH_4OAc (pH 7.4)], 0.25 ml of 10% (w/v) Triton X-100 and a few crystals of sodium hydrosulfite (used to reduce the cytochrome c) were added to a 3 ml cuvette and the solution (which was inverted once to mix) used to establish a baseline absorbance at OD_{550}. Each solution was then mixed vigorously (five to six inversions) to remove the excess hydrosulfite and the baseline re-established. No greater than 50 µl of the subcellular fraction was added, the solution mixed and the rate of oxidation of cytochrome c by cytochrome c oxidase determined on a chart recorder as a decrease in OD_{550} (Copperstein and Lazarow, 1951).

2.3 ANALYSIS OF PROTEINS

2.3.1 Quantitation

Protein concentrations were determined spectrophotometrically according to the method of Bradford (1976), using ovalbumin as the standard. Manipulations involved the addition of 0.1 ml of sample to 1.0 ml of diluted Bio-Rad protein assay dye reagent concentrate [0.01% (w/v) Coomassie Brilliant Blue G-250, 4.7% (w/v) ethanol, 8.5% (w/v) phosphoric acid] followed by vortexing, 15 min for colour development and measurement of OD_{595} versus reagent blank. For each assay, a standard curve was

established using duplicate samples of ovalbumin at concentrations of 1.0, 2.5, 5.0, 10.0, 15.0 and 20.0 µg/ml. Protein determinations for unknowns were performed in duplicate for two different sample dilutions containing between 1 and 15 µg of protein.

2.3.1.1 Trichloroacetic acid (TCA)-Precipitation of Protein

Dilute samples of protein were concentrated by TCA-precipitation using a modified method of that described by Arola et al. (1977). This procedure entailed the addition of an equal volume of 20% (v/v) TCA to each sample followed by incubation on ice for 1 h. The mixtures were subsequently centrifuged (14,000 rpm) in an Eppendorf microcentrifuge (model 5415; Brinkman Instruments (Canada) Ltd., Rexdale, Ontario) for 10 min at 4°C [larger volumes (>1.5 ml) were centrifuged at 20,000 x g_{max} in a Sorvall SS-34 rotor (13,000 rpm; Dupont/NEN) for 20 min at 4°C], the supernatants discarded and the pellets washed twice with excess ice-cold ether. After centrifugation as above, pellets were dried and resuspended in SDS-PAGE sample buffer (section 2.3.2).

2.3.2 Electrophoresis

One dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Laemmli (1970) using slab gels with an

accompanying discontinuous buffer system. Separating gels [12 cm in length for the Hoefer vertical slab gel unit (model SE400; Hoefer Scientific Instruments, San Francisco, California, USA); 5 cm in length for the Bio-Rad mini-protean II apparatus (Bio-Rad)], constituted by 0.37 M Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 0.05% (v/v) N,N,N',N'-tetramethylethylenediamine (TEMED), 0.2% (w/v) ammonium persulfate (AP) and 30:0.8 acrylamide:N,N'-methylene-bis-acrylamide (w/w) diluted to achieve the appropriate final concentration of acrylamide, were cast as 10 or 15% continuous gels or 10 to 20% linear gradient gels of 0.75 or 1.5 mm thickness. For gradient gels, which were poured with the aid of a Hoefer gradient maker (Model SG15 or SG30; Hoefer Scientific), the 20% (w/v) acrylamide component of the separating gel contained 15% (w/v) sucrose to increase its density. Stacking gels (2 cm in length for the Hoefer unit; 1 cm in length for the mini-protean II apparatus), containing 60 mM Tris-HCl (pH 6.8), 0.1% (w/v) SDS, 0.1% (v/v) TEMED, 0.1% (w/v) AP and 30:0.8 acrylamide:N,N'-methylene-bis-acrylamide (w/w) diluted to achieve a final concentration of 3% (w/v) acrylamide, were poured over the polymerized separating gels and 10-, 15-, or 20-well combs inserted. Samples, of no more than 250 µg protein, were dissolved in sample buffer [62.5 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 10% (w/v) sucrose, 10 mM dithiothreitol (DTT),

0.001% (w/v) bromophenol blue] and boiled for 5 to 10 min just prior to loading. Loaded samples were overlaid with electrophoresis running buffer [0.4 M glycine, 50 mM Tris base, 0.1% (w/v) SDS, pH 8.8] and electrophoresis carried out at 200 V for approximately 6 h or 50 to 60 V overnight until the bromophenol blue marker reached the bottom of the separating gel. After electrophoresis, gels were stained for total protein in 0.1% (w/v) Coomassie Brilliant Blue R-250, 10% (v/v) acetic acid, 35% (v/v) methanol, diffusion-destained by repeated washing in a 10% (v/v) acetic acid, 35% (v/v) methanol solution and dried under vacuum at 80°C using a Bio-Rad slab gel dryer (Model 483; Bio-Rad). The molecular weight standards which were used for SDS-PAGE are presented in section 2.1.1.

2.3.2.1 Fluorography

Fluorographic detection (signal enhancement) of radiolabeled protein subjected to SDS-PAGE was conducted in a manner similar to that described by Bonner and Laskey (1974), with modifications as described below. After SDS-PAGE, proteins were fixed in the acrylamide matrix by gentle agitation of gels in 20 volumes of "destaining" solution [10% (v/v) acetic acid, 35% (v/v) methanol] for 40 min, during which time there was one solution change (initial staining of gels with Coomassie blue was not necessary). Subsequently, gels were dehydrated during agitation in the

presence of 20 volumes of dimethylsulfoxide (DMSO) for 30 min, followed by a second 30 min immersion in fresh DMSO. Gels were then placed in 10 volumes of 22.2% (w/v) 2,5-diphenyloxazole (PPO) in DMSO, gently agitated for 1.5 h and the PPO/DMSO solution decanted and replaced by 20 volumes of distilled H_2O. After agitation for 5 to 10 min, the gels were dried under vacuum at 60°C and exposed to x-ray film (usually Kodak X-Omat AR-5 film; Eastman Kodak Company, Rochester, New York, USA) at -70°C. The radiolabeled molecular weight standards which were used for SDS-PAGE/fluorography are listed in section 2.1.2.

2.3.3 Immunoblot Analysis

Samples subjected to immunoblot analysis included 20k x g supernatants and pellets, purified peroxisomes and cell lysates isolated from C. tropicalis, C. albicans and/or S. cerevisiae. Sample preparation procedures for the 20k x g supernatants, 20k x g pellets and peroxisomes are presented in section 2.2. Yeast cell lysates were prepared according to a modified method of that described by Ausubel et al. (1988). This involved overnight growth of non-saturated ($\approx 10^8$ cells/ml) starting cultures in appropriate medium (section 2.2.1), followed by centrifugation at maximum speed in an IEC table-top clinical centrifuge, two cell pellet washes (one with distilled H_2O, the other with breakage buffer (50 mM Tris base, 50 mM NaCl, 0.1 mM EDTA,

0.1 mM $ZnCl_2$, 15 mM PMSF (freshly made), pH 8.0)] and resuspension in an equal volume (≈0.2 ml) of breakage buffer. Glass beads (425 to 600 µm) were added to the level of the meniscus, the mixture cooled on ice and then disrupted by vigorous vortexing for 2 x 1 min pulses interspersed by 2 min intervals on ice. Lysates were microcentrifuged at 14,000 rpm for ≈3 s and the supernatant removed and stored on ice or frozen at -20°C.

Electrophoretic transfer of polypeptides from SDS-polyacrylamide gels to nitrocellulose was performed in a manner similar to that described by Towbin et al. (1979), as modified by Burnette (1981). To this end, samples/polypeptides, prepared as described above, were first subjected to SDS-PAGE (section 2.3.2). After electrophoresis, a "gel:nitrocellulose sandwich" was constructed such that there was direct contact (no bubbles present) between the gel and nitrocellulose membrane, both of which were housed between two sheets of Whatman 3 mm chromatography paper (Whatman International Ltd., Maidstone, England, UK), two Scotch-Brite pads and the plastic holder supplied with the Bio-Rad Trans-Blot cell (Bio-Rad). All components in contact with the slab gel were presoaked in transfer buffer [20 mM Tris base, 150 mM glycine, 20% (v/v) methanol]. With the "sandwich" immersed in transfer buffer (nitrocellulose toward the anode), transfer to

nitrocellulose was carried out using the Trans-Blot cell at 100 mA for 24 to 48 h or 1 A for 2.5 h. The extent of protein transfer was monitored by the migration of prestained SDS-PAGE molecular weight standards (see section 2.1.1) to the nitrocellulose membrane and by gel staining/destaining post-transfer. If necessary, the protein profile of the nitrocellulose blot was visualized post-transfer by staining with 0.1% (w/v) amido black, 45% (v/v) methanol, 10% (v/v) acetic acid and subsequent destaining with 90% (v/v) methanol, 2% (v/v) acetic acid (Schaffner and Weissman, 1973). Post-transfer, the nitrocellulose blot was immersed in 50 to 100 ml of blocking solution [Fujiki et al., 1984; 1% (w/v) equine hemoglobin in Tris-buffered saline (10 mM Tris-HCl (pH 7.4), 0.9% (w/v) NaCl)], incubated at 37°C for 30 min with gentle agitation and then sealed in a plastic bag with 5 to 10 ml of fresh blocking solution containing an appropriate (predetermined) dilution of antiserum (1:300 dilution for rabbit anti-C. tropicalis catalase serum; 1:2000 dilution for rabbit anti-firefly luciferase serum; 1:1000 dilution for rabbit anti-CAT serum). After gentle rocking for 90 min at room temperature using a Tek-Pro tube rocker (American Dade, Miami, Florida, USA), the nitocellulose membrane was washed once with Tris-buffered saline for 10 min, twice with Tris-buffered saline containing 0.05% (w/v) Nonidet P-40 (NP-40)

for 20 min and once again with Tris-buffered saline for 10 min. Subsequently, the blot was resealed in a plastic bag with 5 to 10 ml of blocking solution containing ^{125}I-protein A (0.05 µCi/ml of blocking solution) and gently rocked for 30 min at room temperature. The nitrocellulose blot was then washed as described above, dried and exposed to X-Omat AR-5 film at -70°C in the presence of a Dupont Cronex lightning plus intensifying screen (Picker International Canada Inc., Brampton, Ontario).

2.3.4 Immunoprecipitation of Cell-Free Translation Products

In vitro translation products (section 2.6.1.3) were immunoprecipitated according to the method of Fujiki et al. (1984). Accordingly, translation mixtures, normalized to ≈50,000 cpm above background, were adjusted to contain 10 mM Tris-HCl (pH 7.4), 1% (w/v) NP-40, 150 mM NaCl, 10 mM methionine, leupeptin (25 µg/ml), antipain (25 µg/ml), chymostatin (12.5 µg/ml), pepstatin (12.5 µg/ml). These diluted translation mixtures were centrifuged at 145,000 x g_{max} in a Beckman Type 65 rotor (40,000 rpm; Beckman Instruments) for 1 h at 4°C. The resultant post-ribosomal supernatants were removed, and SDS and EDTA (pH 7.5) added to yield final concentrations of 0.1% (w/v) and 2 mM, respectively. Post-ribosomal supernatants were subsequently incubated with 5 µl of rabbit anti-C. tropicalis catalase

serum for 90 min at room temperature and then overnight at 4°C, using a Labquake shaker (Labindustries Laboratory and Research Instruments, Berkeley, California, USA) to maintain continuous gentle rolling. The following day, 30 µl of a 10% (w/v) solution of inactivated Staphylococcus aureus cells (Pansorbin), prewashed with binding buffer [10 mM Tris-HCl (pH 7.4), 1% (w/v) NP-40, 150 mM NaCl, 10 mM methionine, 0.1% (w/v) SDS, 2 mM EDTA, leupeptin (25 µg/ml), antipain (25 µg/ml), chymostatin (12.5 µg/ml), pepstatin (12.5 µg/ml)], were used to adsorb antigen-antibody complexes during a 90 min incubation at room temperature with gentle rolling. S. aureus cells were then pelleted by microcentrifugation at 14,000 rpm for 1 min and washed sequentially with 1 ml each of binding buffer (3 washes), binding buffer without 0.1% (w/v) SDS and 2 mM EDTA (pH 7.5) containing only 0.05% (w/v) NP-40 (2 washes), binding buffer without 1% (w/v) NP-40, 0.1% (w/v) SDS and 2 mM EDTA (pH 7.5) (1 wash) and binding buffer without 1% (w/v) NP-40, 0.1% (w/v) SDS, 2 mM EDTA (pH 7.5) and 150 mM NaCl (1 wash). Resultant pellets were resuspended in SDS-PAGE sample buffer and the immunoprecipitates dissociated by boiling for 5 to 10 min. S. aureus cells were subsequently removed by microcentrifugation for 5 min and the supernatants analyzed by electrophoresis (10 to 20% linear gradient gels; section 2.3.2), fluorography and exposure to X-Omat AR-5 film at

-70°C.

2.4 ANALYSIS OF NUCLEIC ACIDS

2.4.1 Quantitation

Nucleic acid concentrations of pure, concentrated samples [those obtained subsequent to equilibrium cesium chloride (CsCl) density gradient centrifugation] were determined by spectrophotometric measurement at OD_{260}. For this purpose, 50 µl of sample was diluted to 1 ml in distilled H_2O, the OD_{260} and OD_{280} determined and the nucleic acid concentration calculated assuming an OD_{260} of 1 corresponds to ≈50 µg/ml for double-stranded DNA, 40 µg/ml for single-stranded DNA and RNA and 20 µg/ml for oligonucleotides (Sambrook et al., 1989). Ratios of OD_{260}/OD_{280} provided estimates of sample purity with ideal values for DNA and RNA being 1.8 and 2.0, respectively (Sambrook et al., 1989). Dilute or impure samples of nucleic acid were quantitated by the ethidium bromide fluorescence method of Morgan et al. (1979). To this end, a Sequoia-Turner digital fluorometer (model 450, Sequoia-Turner Corporation, Mountain View, California, USA) was calibrated with 1.0 µg of a standard RNA solution set to a value of 100 in pH 8.1 buffer. After determining fluorescence readings of individual samples (1 µl each), 1 µl of RNase A (1 mg/ml) was added and the solutions mixed

and incubated at 37°C for 5 min. The resultant loss of fluorescence of the samples relative to the standard was used as a measure of the amount of RNA. DNA was quantified similarly using fluorescence measurements pre- and post-DNase-treatment of samples and a standard DNA solution (0.5 µg set to 100 in pH 11.8 buffer).

2.4.2 Purification and Concentration of DNA from Aqueous Solutions

Aqueous solutions of DNA contaminated with protein were purified by repeated extraction with H_2O-saturated phenol:chloroform:isoamyl alcohol. Typically, an equal volume (50 µl minimum) of H_2O-saturated phenol:chloroform:isoamyl alcohol (50:50:1, v/v/v) was added to the DNA solution to be purified, the contents vortexed vigorously until an emulsion formed and this mixture microcentrifuged (14,000 rpm) at room temperature for 20 to 30 s to obtain a well-defined phase separation (spins of 2 to 3 min duration were required for concentrated DNA solutions or those contaminated with large amounts of protein). If a white precipitate collected at the aqueous/organic interface, the aqueous phase was removed and re-extracted as detailed above until a "clean" interface was obtained (usually two extractions were sufficient). The protein-deficient aqueous phase was subsequently removed, similarly extracted once with an equal volume of

chloroform:isoamyl alcohol (50:1, v/v) and the DNA recovered by ethanol precipitation (see below).

Dilute solutions of DNA were concentrated by ethanol precipitation as described by Maniatis et al. (1982). To this end, either a 1/10 volume of 3 M NaOAc (pH 5.2) or an equal volume of 4 M NH_4OAc (pH 5.5) was added to the DNA solution, followed by vortexing and a further addition of 2 to 2.5 volumes (calculated after salt addition) of ice-cold absolute ethanol. The DNA was precipitated in a dry ice:95% (v/v) ethanol bath for 20 min and pelleted by microcentrifugation (14,000 rpm) for 10 min at 4°C [small DNA fragments (<500 bp) were subjected to longer periods of precipitation and centrifugation - usually 30 min each]. DNA pellets were washed once with 70% (v/v) ethanol, recentrifuged, dried under vacuum for 10 min using a Savant Speed Vac concentrator (model SVC100-H; Savant Instruments, Inc., Hicksville, New York, USA) and dissolved in sterile, distilled H_2O at the desired concentration.

2.4.3 Electrophoresis

The standard method employed to separate, identify and purify DNA fragments was electrophoresis through agarose gels in a manner similar to that described by Aaij and Borst (1972) and Sharp et al. (1973), with modifications as detailed below. Agarose gels [ranging in size from that of

the Hoefer minnie horizontal agarose submarine unit (6.5 x 10 cm; model HE33; Hoefer Scientific) to that of the Hoefer max horizontal agarose submarine unit (14.5 x 25 cm; model HE99; Hoefer Scientific)] were prepared by heating a mixture of 0.8 to 2.0% (w/v) agarose in electrophoresis (TBE) buffer [89 mM Tris base, 89 mM boric acid, 2 mM EDTA (pH 8.0)] until the agarose dissolved. The solution was cooled to <50°C, ethidium bromide added to 0.5 μg/ml and the mixture poured into a gel casting tray. A sample well comb (8- to 16-wells for the "minnie" gel, 10- to 20-wells for the "max" gel) was inserted into the molten agarose and the gel allowed to solidify for 30 to 45 min at room temperature. Once the gel had solidified, the comb was removed and samples, prepared in gel-loading dye #1 [Maniatis et al., 1982; 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 40% (w/v) sucrose], were loaded into the preformed wells. The amount of DNA per sample that could be loaded into each well, without significant loss of resolution, was dependent upon the dimensions of each well, the number of fragments per sample and their sizes. Buffer chambers were subsequently filled with TBE until the gel was submerged to a depth of ≈2 mm and electrophoresis carried out at 5 to 6 V/cm until the bromophenol blue tracking dye approached the bottom of the gel. DNA embedded in ethidium bromide-stained agarose gels was visualized by exposure to midrange

ultraviolet light (wavelength = 300 nm) using a Fotodyne UV 300 "plus" transilluminator (Bio/Can Scientific Inc., Mississauga, Ontario) and photographed with the aid of a Fotodyne FCR-10 DNA photographic system (Bio/Can Scientific).

An alternative to agarose gel electrophoresis, used principally for purifying DNA fragments, was nondenaturing polyacrylamide gel electrophoresis. Physically assembled as described for SDS-PAGE (section 2.3.2), nondenaturing polyacrylamide gels were constituted by 29:1 acrylamide:N,N'-methylene-bis-acrylamide (w/w) diluted to the appropriate final concentration of acrylamide [usually 3.5 to 5% (w/v)] in TBE containing 0.063% (w/v) AP. This solution was degassed for 15 min and 0.04% (v/v) TEMED added. The mixture was then poured between glass plates, a 10- or 15-well comb inserted and polymerization allowed to occur for 1 h at room temperature. Samples (not greater than 1 µg of DNA per band) were prepared in gel loading dye #1 and electrophoresis carried out at 50 to 150 V until the tracking dye reached the bottom. The gel was removed, stained for 30 min in TBE containing 0.5 µg/ml ethidium bromide and visualized/photographed as described above. The molecular weight standards which were used for agarose and polyacrylamide gel electrophoresis of DNA are listed in section 2.1.1.

2.4.3.1 Recovery of DNA Fragments from Agarose or Nondenaturing Polyacrylamide Gels

Quantitative recovery of DNA fragments from agarose gels was accomplished exactly as described in the operator's manual accompanying the International Biotechnologies unidirectional electroeluter, analytical (model UEA; International Biotechnologies, Inc., New Haven, Connecticut, USA). By contrast, isolation of DNA fragments from nondenaturing polyacrylamide gels was achieved as described by Maxam and Gilbert (1977) and modified by Maniatis et al. (1982). This entailed excision and chopping into pieces of the desired gel slice (DNA fragment), followed by transfer of the pieces into a 1.5 ml microcentrifuge tube containing 1.5 volumes of elution buffer [0.5 M NH_4OAc, 1 mM EDTA (pH 8.0)]. The mixture was incubated at 37°C overnight in a rotating wheel and the eluant collected by centrifugation (speed setting 5 in an IEC table-top clinical centrifuge) through a 1 ml syringe plugged with sterile siliconized glass wool. The pelleted polyacrylamide was washed with 0.5 volumes of elution buffer, the eluants combined, 2 volumes of absolute ethanol added and the DNA precipitated at -20°C for 20 min. After microcentrifugation (14,000 rpm) for 10 min at 4°C, the pellet was dissolved in 200 µl of TE buffer [10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)], 25 µl of 3M NaOAc (pH 5.2) added and the DNA reprecipitated and

pelleted. The pellet was washed once with 70% (v/v) ethanol, dried under vacuum and dissolved in a minimum volume of distilled H_2O. Using this method, recovery of a 1 kbp fragment from a 3.5% (w/v) nondenaturing polyacrylamide gel was estimated to be approximately 50%.

2.4.4 Standard Techniques Used for the Enzymatic Manipulation of DNA

Standard techniques used for the enzymatic manipulation of DNA included:

(i) restriction endonuclease cleavage of DNA

(ii) conversion of recessed 3'-ends of double-stranded DNA to blunt ends using DNA polymerase I Klenow fragment

(iii) conversion of 3'-protruding termini of double-stranded DNA to blunt ends using mung bean nuclease

(iv) dephosphorylation of terminal 5'-phosphates of double-stranded DNA using calf intestinal phosphatase

(v) homopolymeric tailing of terminal 3'-hydroxyls of double-stranded DNA using terminal deoxynucleotidyl transferase

(vi) blunt-end or complementary-end DNA fragment ligation using T4 DNA ligase

In general, these techniques were performed according to manufacturers' product data sheets, Maniatis et

al. (1982) or Davis et al. (1986). The information presented below represents only specific modifications or additions that were made to pre-existing protocols.

With reference to restriction endonuclease cleavage of DNA, buffer compatibility for double restriction enzyme digestions was achieved in the appropriate concentration of potassium glutamate buffer [KGB; 100 mM potassium glutamate, 25 mM Tris-acetate (pH 7.6), 10 mM $Mg(OAc)_2$, 50 µg/ml nuclease-free BSA, 0.5 mM 2-mercaptoethanol] according to Hanish and McClelland (1988).

Conversion of 3'-protruding termini of double-stranded DNA to blunt ends with mung bean nuclease was accomplished using a procedure obtained from Dr. D.W. Andrews, Department of Biochemistry, McMaster University, Hamilton, Ontario. Accordingly, purified DNA (up to 10 µg) was dissolved in 100 µl of buffer A [6 mM Tris-HCl (pH 8.0), 6 mM NaCl, 0.2 mM EDTA] and then 10 µl of buffer B [300 mM NaOAc (pH 4.7), 500 mM NaCl, 10 mM $ZnCl_2$] was added. This mixture was incubated at 37°C for 2 min, 1 µl of mung bean nuclease [diluted to 4 units/µl in buffer B containing 50% (v/v) glycerol] added and incubation continued at 37°C for 10 min. After adding 1 µl of 20% (w/v) SDS, the mixture was incubated at 65°C for 5 min followed by the addition of 4 µl of 2 M Tris-HCl (pH 8.9) and 6 µl of 12 M LiCl. Following two H_2O-saturated phenol:chloroform:isoamyl alcohol and one

chloroform:isoamyl alcohol extraction, the DNA was ethanol precipitated, pelleted, washed, dried and dissolved in sterile, distilled H_2O.

Prior to ligation of DNA fragments, residual restriction endonuclease activity was destroyed by heat inactivation as specified by the manufacturer or, if necessary, by sample extraction and precipitation (section 2.4.2). Complementary-end ligations were carried out as detailed in the cloning strategies section of the International Biotechnologies catalogue. Blunt-end ligations, while mimicking these instructions, were routinely conducted overnight at 16°C with molar ratios of insert:vector greater than 1:1, usually 2:1 and 5:1. Ligations routinely involved two different T4 DNA ligase reaction buffers, either the T4 DNA ligase buffer supplied by Gibco/BRL [50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% (w/v) polyethylene glycol (PEG)-8000] or 0.5 x KGB supplemented with 1.0 mM ATP (Hanish and McClelland, 1988). The highest efficiency ligations (greatest percentage of recombinants) were achieved using gel-purified DNA fragments.

2.4.5 Manipulation and Analysis of DNA

2.4.5.1 Radiolabeling of DNA Probes

Radiolabeling the 5'-ends of desalted

oligonucleotides with T4 polynucleotide kinase was conducted according to a procedure obtained from Dr. C.B. Harley, Department of Biochemistry, McMaster University, Hamilton, Ontario. Reactions were carried out at 37°C for 40 min in 28 µl (final volume) of T4 polynucleotide kinase buffer [50 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT] containing 6 pmol of oligonucleotide [33 ng (0.0017 OD_{260} units) assuming 1 OD_{260} = 20 µg of oligonucleotide; Sambrook et al., 1989], 10 µCi of [γ-^{32}P]ATP (3,000 Ci/mmol, 10 µCi/µl) and 3 to 5 units of T4 polynucleotide kinase. Occasionally, the incorporation of $^{32}PO_4$ was monitored by electrophoresis and autoradiography of denaturing polyacrylamide gels for DNA [section 2.4.3; 12% (w/v) polyacrylamide-TBE gels containing 7 M urea]. Routinely, free [γ-^{32}P]ATP was separated from labeled oligonucleotide by centrifugation through a Sephadex G-25 spin column (Maniatis et al., 1982). To this end, 0.9 ml of Sephadex G-25, swollen in TE (pH 8.0), was poured into a 1 ml disposable syringe plugged with sterile, siliconized glass wool. The column was equilibrated with two TE (pH 8.0) washes (0.1 ml/wash) by centrifugation for 4 min at speed setting 5 in an IEC table-top clinical centrifuge. The labeling reaction, diluted to 100 µl with TE (pH 8.0), was applied and the column recentrifuged as above, collecting the effluent in a decapped 1.5 ml microcentrifuge tube. Quantitation of the incorporated counts per 0.5 µl of

effluent revealed typical specific activities for 5'-end-labeled oligonucleotides to be approximately 10^8 cpm/µg.

DNA fragments, purified by agarose or nondenaturing polyacrylamide gel electrophoresis (section 2.4.3), were radiolabeled using one of three different methods: nick translation (Rigby et al., 1977), the random primer method of Feinberg and Vogelstein (1983, 1984) or DNA polymerase I Klenow fragment fill-in of the recessed 3'-ends of double-stranded DNA.

Nick translation was carried out exactly as described by the protocol accompanying the Gibco/BRL nick translation reagent kit. To this end, the final reaction mixture (50 µl volume) included: 20 µM of each dNTP (excluding the radionucleotide), 50 mM Tris-HCl (pH 7.8), 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 10 µg/ml nuclease-free BSA, 156 pmoles of [α-^{32}P]dCTP (410 Ci/mmol, 10 µCi/µl) and 1.0 µg gel-purified DNA fragment dissolved in sterile, distilled H_2O. The reaction, initiated by the addition of enzyme [2 units DNA polymerase I, 200 pg DNase I, 5 mM Tris-HCl (pH 7.5), 0.5 mM $Mg(OAc)_2$, 0.1 mM 2-mercaptoethanol, 10 µM PMSF, 5% (v/v) glycerol, 10 µg/ml nuclease-free BSA] was allowed to proceed for 1 h at 15°C and subsequently stopped by the addition of EDTA (pH 8.0) to 27.3 mM. Free [α-^{32}P]dCTP was separated from labeled DNA by centrifugation through a Sephadex G-50 spin column as described above.

Typical specific activities obtained for nick translated DNA fragments were approximately 10^7 cpm/μg.

Random primer labeling of DNA fragments was accomplished following the protocol supplied with the Gibco/BRL random primers DNA labeling system. Accordingly, 20 to 40 ng of gel-purified DNA fragment was heat denatured by boiling for 5 min and quick-cooled on ice. The denatured template was then added to a 50 μl (final volume) reaction mixture (pH 6.8) containing 20 μM of each dNTP (excluding the radionucleotide), 200 mM N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES), 50 mM Tris-HCl, 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 0.4 mg/ml BSA, 50 μCi of [α-^{32}P]dATP (3,000 Ci/mmol, 10 μCi/μl) and 5.4 OD_{260} units/ml of oligodeoxyribonucleotide primers (hexamer fraction). The reaction was initiated by adding 3 units of DNA polymerase I Klenow fragment, allowed to proceed at room temperature for 1 to 24 h and stopped by the addition of EDTA (pH 7.5) to 18 mM. Free [α-^{32}P]dATP was separated from radiolabeled DNA by centrifugation through a Sephadex G-50 spin column. Typical specific activities obtained for random primer labeled DNA fragments were approximately 10^8 cpm/μg.

DNA polymerase I Klenow fragment fill-in of the recessed 3'-ends of double-stranded DNA was routinely used to radiolabel a 77 bp EcoR I, Pvu II-digested DNA fragment derived from the partial cDNA recombinant encoding catalase,

RU1:87 (Rachubinski et al., 1987; section 4). For this purpose, the following reaction mixture (20 µl final volume) was employed: 7 mM Tris-HCl (pH 7.4), 50 mM NaCl, 7 mM $MgCl_2$, 10 µCi of [α-^{32}P]dATP (3,000 Ci/mmol, 10 µCi/µl) and 15 to 20 ng of gel-purified DNA fragment. The reaction, initiated by adding 3 to 4 units of enzyme, was allowed to proceed at 37°C for 15 min, after which time a Sephadex G-50 spin column was used to separate free isotope from radiolabeled DNA. Typically, DNA polymerase I Klenow fragment fill-in reactions yielded specific activities of approximately 10^8 cpm/µg.

All ^{32}P-labeled oligonucleotides and DNA fragments were heat denaturated by boiling for 5 to 10 min prior to storage at -20°C.

2.4.5.2 Detection of Specific DNA Sequences by Blotting and Hybridization

The detection of specific DNA sequences by blotting and hybridization to nitrocellulose membranes was accomplished using three different, but related, techniques: Southern blotting, colony hybridization and DNA dot-blotting.

Southern blotting, or the capillary transfer of denatured DNA fragments (separated by agarose gel electrophoresis) to solid-support membranes at high salt concentrations (Southern, 1975), was carried out as

described by Maniatis et al. (1982) with modifications as indicated below. For Southern blots of genomic DNA, agarose [0.8% (w/v)] gel electrophoresis was conducted with 5.0 µg of restriction endonuclease-digested, total yeast DNA isolated from C. tropicalis (section 2.5.1). However, recombinant plasmids harbouring cloned DNA sequences of interest were easily detected post-agarose [1.0% (w/v)] gel electrophoresis and blotting of 0.1 to 0.2 µg of DNA. Prior to Southern transfer, agarose-embedded DNA fragments were denatured by soaking the gel in 3 volumes of 0.4 M NaOH, 0.8 M NaCl for 1 h at room temperature with gentle agitation and subsequently neutralized by soaking the gel in 3 volumes of 0.5 M Tris-HCl (pH 7.6), 1.5 M NaCl for 1 h at room temperature with gentle agitation. DNA transfer was accomplished by capillary action as outlined by Maniatis et al. (1982) using 6 x SSC (1 x SSC = 150 mM NaCl, 15 mM Na_3·citrate, pH 7.0) as the transfer solution. Transfer was allowed to proceed for 12 to 24 h using 1 h of transfer time per 1 kbp of fragment length as a guideline. After disassembly, the nitrocellulose membrane was gently agitated in 6 x SSC for 15 min, air-dried and either baked *in vacuo* for 2 h at 80°C using a Napco vacuum oven (Model 5831; Fisher Scientific) or photo-crosslinked by ultraviolet irradiation (0.12 J) using a Stratagene UV Stratalinker 1800 (Stratagene Cloning Systems).

The two alternative blotting methods used for the identification of recombinant plasmids (cloned DNA fragments) harbouring a specific gene, namely, colony hybridization (Grunstein and Hogness, 1975; Hanahan and Meselson, 1980) and DNA dot-blotting (Kafatos et al., 1979), were performed as detailed below.

Prior to colony hybridization, transformed E. coli (section 2.4.5.3) were grown overnight at 37°C on antibiotic:LB medium:agar plates to a desired size (≈1 mm in diameter). The following day, dry nitrocellulose disks were placed directly on the bacterial colonies, marked, and when completely wetted (2 to 3 min), removed and placed colony-side up onto a minimal volume (2 ml) of 0.5 M NaOH (freshly prepared and aliquoted onto plastic wrap). After 3 min, the disks were blotted briefly on a paper towel and the process repeated once with 0.5 M NaOH, twice with 2 ml of 1 M Tris-HCl (pH 7.6) and twice with 2 ml of 0.5 M Tris-HCl (pH 7.4), 1.5 M NaCl. Blotted disks were subsequently placed onto 5 ml of 2 x SSC for 5 min, after which time colony debris was removed by gently rubbing the nitrocellulose membrane. The excess fluid was then drained and the disks air-dried and processed by baking in vacuo or UV photo-crosslinking.

Prior to DNA dot-blotting, recombinant plasmid was isolated from transformed E. coli using the boiling method of Holmes and Quigley (1981) (section 2.4.5.4). Plasmid

pellets were dissolved in 20 µl of 0.5 M NaOH and the DNA denatured for 20 min at room temperature. After the addition of 30 µl of neutralizing solution [1 M NaCl, 50 mM Tris-HCl (pH 8.0), 50 mM EDTA], the mixture was diluted to 200 µl (final volume) with 50 µl of 1 M Tris-HCl (pH 6.8) and 100 µl of sterile, distilled H_2O. One-twentieth (10 µl) of each volume (≈200 ng of DNA) was consecutively aspirated onto a nitrocellulose membrane (presoaked in 3 x SSC) using a Schleicher and Schuell minifold microsample filtration manifold (model SRC-96/0; Mandel Scientific Company Ltd., Rockwood, Ontario). Upon disassembly of the manifold, the nitrocellulose sheet was air-dried and baked in vacuo for 2 h at 80°C.

Those conditions used for the hybridization of ^{32}P-labeled DNA fragments or oligonucleotides to DNA fixed to either Southern blots, blots arising from colony hybridization or DNA dot-blots are presented below.

To reduce nonspecific hybridization with the probe, nitrocellulose blots were initially prehybridized in the presence of a solution containing 0.625 x SSC, 0.08 x Denhardt's solution [1 x Denhardt's solution = 0.02% (w/v) Ficoll 400, 0.02% (w/v) polyvinylpyrrolidone, 0.02% (w/v) BSA (Cohn fraction V)], 2 µg denatured salmon sperm DNA/ml, 0.0005% (w/v) SDS and 10 mM sodium phosphate buffer (pH 7.0) for 2 to 3 h at 65°C, in the absence of any probe. The

prehybridization solution was then drained and replaced by hybridization solution [2 x prehybridization solution supplemented with 30% (v/v) deionized formamide] containing 1 x 10^5 to 1 x 10^6 cpm of ^{32}P-labeled DNA fragment or oligonucleotide/ml of hybridization mix. For ^{32}P-labeled DNA fragments, hybridization was carried out overnight at between 42 and 55°C, depending upon fragment length [shorter fragments (<100 nucleotides) necessitated the lower temperature]. The subsequent washing of blots (4 washes of ≈10 min each with moderate agitation) was similarly conducted at between 42 and 55°C in the presence of 1 x SSC. The appropriate hybridization temperatures for ^{32}P-labeled oligonucleotides were estimated using the formula, melting temperature (T_m) = 4°C x (each G or C residue of the probe) + 2°C x (each A or T residue of the probe) - 10°C (Itakura et al., 1984; Sambrook et al., 1989). For most oligonucleotides, this dictated overnight hybridization at a temperature between 30 and 42°C. Subsequent washing of blots (4 washes of 5 to 10 min each with moderate agitation) was routinely conducted at the same temperature as the hybridization. After washing, nitrocellulose was briefly blotted on Whatman paper and while still moist, wrapped in plastic film and exposed to Kodak X-Omat AR-5 film at -70°C in the presence of an intensifying screen. Moist blots could be rewashed at higher stringency (increased

temperature or decreased concentrations of SSC) to remove excessive background, if necessary.

2.4.5.3 Transformation of E. coli with Plasmid DNA

Transformation of competent E. coli RR1, HB101 and DH5α cells purchased from Gibco/BRL was conducted exactly as described by the accompanying product data sheets. Alternatively, the method used for the preparation of competent E. coli RR1 and HB101 cells was the calcium chloride procedure of Mandel and Higa (1970), as modified below. LB or YT medium (100 ml) was inoculated with 1 ml of an overnight bacterial culture and grown at 37°C with shaking to a density of ≈5 x 10^7 cells/ml (requires 2 to 4 h; OD_{550} should be 0.2 and 0.5 for RR1 and HB101 cells, respectively, such that cells are present in early- to mid-logarithmic phase of growth). The cell suspension was subsequently placed on ice for 10 min, pelleted by centrifugation at 1,100 x g_{max} in a Sorvall SS-34 rotor (3,000 rpm) for 5 min at 4°C, the supernatant discarded and the pellet gently resuspended in 20 ml (1/5 of the original volume) of ice-cold transformation medium [5 mM Tris-HCl (pH 7.6), 5 mM $MgCl_2$, 0.25 M KCl, 0.1 M $CaCl_2$]. The resuspended cells were left on ice for 15 min, recentrifuged as above and the pellet gently resuspended in 1 ml (1/100 of the original volume) of ice-cold transformation medium containing 15% (v/v) glycerol. Cells were stored at 4°C for

12 to 24 h prior to transformation or quick-frozen in a dry ice:ethanol bath and stored for longer periods at -70°C. The transformation protocol used for competent E. coli strains prepared in this manner was identical to that suggested by the Gibco/BRL product data sheets.

2.4.5.4 Plasmid DNA Isolation and Purification

Two different procedures were employed for the rapid, small-scale isolation of plasmid DNA. One procedure paralleled the boiling method of Holmes and Quigley (1981), with modifications as described below. Microcentrifugation (14,000 rpm) of 1.5 ml of an overnight bacterial culture yielded a pellet which was dispersed in 350 µl of STET medium [8% (w/v) sucrose, 10 mM Tris-HCl (pH 8.0), 50 mM EDTA (pH 8.0), 0.5% (w/v) Triton X-100]. Subsequent to the addition of 25 µl of lysozyme solution (10 mg/ml in sterile, distilled H_2O), the mixture was vortexed briefly, boiled for 40 s and microcentrifuged for 10 min at room temperature. The gelatinous pellet was removed using a toothpick, 200 µl of ice-cold 5 M NH_4OAc (pH 5.5) and 1 ml of ice-cold 2-propanol added and the mixture placed at -20°C for 30 min. After microcentrifugation for 10 min at 4°C, the supernatant was discarded and the pellet washed with 70% (v/v) ethanol and recentrifuged. The washed pellet was dried under vacuum and dissolved as appropriate. The second procedure routinely used for the small-scale isolation of plasmid DNA

was the alkaline extraction method of Birnboim and Doly (1979) as modified by Maniatis et al. (1982). Accordingly, microcentrifugation of 1.5 ml of an overnight bacterial culture yielded a pellet which was dispersed in 100 µl of lysozyme solution [25 mM Tris-HCl (pH 8.0), 50 mM glucose, 10 mM EDTA, 4 mg/ml lysozyme] and incubated for 5 min at room temperature. Subsequent to the addition of 200 µl of 0.2 M NaOH, 1% (w/v) SDS, the solution was mixed by inversion and placed on ice for 5 min. To this, 150 µl of an ice-cold 3 M potassium, 5 M acetate solution (pH 4.8) was added, the mixture gently vortexed and then placed on ice for 5 min. After microcentrifugation for 5 min at 4°C, the supernatant was extracted with an equal volume of H_2O-saturated phenol:chloroform:isoamyl alcohol and the plasmid precipitated, for 5 min at room temperature, by the addition of 2 volumes of absolute ethanol. Microcentrifugation for 5 min at room temperature yielded a pellet which was washed with 70% (v/v) ethanol, dried under vacuum and dissolved as appropriate.

The large-scale isolation of plasmid DNA was accomplished as described by Maniatis et al. (1982) using a proportionally scaled-up alkaline extraction method (Birnboim and Doly, 1979). Initially, 100 to 500 ml of antibiotic-containing medium was inoculated with a 1:100 dilution of an overnight bacterial culture and allowed to

grow to saturation at 37°C overnight. Plasmid DNA, isolated in a similar, but scaled-up manner to that discussed above for the alkaline extraction mini-preparation procedure, was purified by centrifugation to equilibrium in CsCl-ethidium bromide density gradients (Radloff et al., 1967) as indicated below. One gram of solid CsCl was added to each ml of plasmid DNA [dissolved in TE (pH 8.0)], followed by the addition of 0.7 ml of ethidium bromide (10 mg/ml) per 10 ml of CsCl solution. Gradients were established at 101,000 x g_{av} in a Beckman Type 65 rotor (40,000 rpm) for 48 h at 20°C or at 338,000 x g_{av} in a Beckman TLV-100 rotor (100,000 rpm) for a minimum of 3 h at 20°C. Banded, covalently closed, circular plasmid DNA was withdrawn and extracted 4 to 6 times with H_2O-saturated 1-butanol to remove the residual ethidium bromide (Maniatis et al., 1982). Cesium chloride was removed by overnight dialysis [using Spectra/Por molecular porous membrane tubing (molecular weight cutoff 12,000-14,000)] at 4°C against 2 x 2 L of TE (pH 7.5) or by ethanol precipitation of plasmid DNA. This entailed the addition of 2 volumes of TE (pH 8.0) and 6 volumes of absolute ethanol, followed by the precipitation of plasmid DNA on ice for 20 min. After microcentrifugation for 10 min at 4°C, the pellet was dissolved in 0.5 ml of 0.3 M NaOAc (pH 5.2). Two volumes of absolute ethanol were then added and the plasmid DNA reprecipitated on ice for 20 min.

Subsequent to microcentrifugation, the DNA was washed with 70% (v/v) ethanol, dried under vacuum and dissolved in TE (pH 8.0). Plasmid DNA isolated by equilibrium CsCl density gradient centrifugation was quantitated by measurement of OD_{260} and checked for purity by measurement of OD_{260}/OD_{280} (section 2.4.1).

2.5 GENE CLONING AND SEQUENCING

2.5.1 Isolation and Size Fractionation of Nuclear DNA from C. tropicalis

Nuclear DNA was isolated from C. tropicalis by alcohol precipitation from a 3 M LiCl solubilization of total nucleic acids (Maccecchini et al., 1979), followed by purification on CsCl (Maniatis et al., 1982) as described below. Initially, a starting suspension culture of C. tropicalis was transferred to 1 L of n-alkane medium [Osumi, 1974; 0.5% (w/v) $NH_4H_2PO_4$, 0.25% (w/v) KH_2PO_4, 0.1% (w/v) $MgSO_4$, 0.002% (w/v) $FeCl_3$, 0.1% (w/v) corn steep liquor, 0.05% (v/v) Tween 80, 0.24% (v/v) n-decane, 0.47% (v/v) n-undecane, 0.26% (v/v) n-dodecane, 0.03% (v/v) n-tridecane, pH 5.2] and cultured for ≈15 h at 30°C with shaking. Cell pellets (≈20 g wet weight) were washed twice with H_2O, frozen in liquid nitrogen and ground with a porcelain mortar and pestle. The resulting yeast "powder" was extracted at room temperature for 20 min in a mixture (100 ml per 10 g powder) of 1 volume of extraction buffer [150 mM NaCl, 50 mM

Tris-HCl (pH 7.5), 5 mM EDTA, 5% (w/v) SDS] and 1 volume of phenol [saturated in 10 mM EDTA (pH 7.4)]:chloroform:isoamyl alcohol (50:50:1, v/v/v). The phases were separated by centrifugation at 13,000 x g_{av} in a Sorvall GSA rotor (≈9,000 rpm) for 20 min at 20°C, the aqueous phase removed and the organic phase re-extracted as before. The resultant aqueous phases were pooled, extracted three times with 1 volume of EDTA-saturated phenol:chloroform:isoamyl alcohol and total nucleic acids precipitated overnight at -20°C upon the addition of a 1/10 volume of 2 M KOAc (pH 5.5) and 2.5 volumes of absolute ethanol. Nucleic acids were pelleted by centrifugation at speed setting 5 in an IEC table-top clinical centrifuge for 10 min at 4°C and the pellet dissolved and precipitated with 3 M LiCl for 24 to 48 h on ice. The precipitate containing the single-stranded total RNA was pelleted as above, washed twice with 70% (v/v) ethanol, dried under vacuum and dissolved in sterile, distilled H_2O. The supernatant containing the nuclear DNA was decanted, 2.5 volumes of absolute ethanol added and the DNA precipitated overnight at -20°C. After centrifugation at 12,000 x g_{max} in a Sorvall SS-34 rotor (10,000 rpm) for 20 min at 4°C, the DNA pellet was washed twice with 70% (v/v) ethanol, dried under vacuum and dissolved in TE (pH 8.0). Solid CsCl was added to yield a concentration of 1.0 g/ml, 0.7 ml of ethidium bromide (10 mg/ml) was then added

per 10 ml of CsCl solution and the DNA banded by centrifugation at 101,000 x g_{av} in a Beckman Type 65 rotor (40,000 rpm) for 48 h at 20°C. CsCl-purified DNA was subsequently withdrawn, extracted and dialyzed as described in section 2.4.5.4. Genomic DNA isolated in this manner was greater than 35 kbp in length, as estimated by agarose [0.4% (w/v)] gel electrophoresis. Total RNA and DNA were quantitated and checked for purity by measurement of OD_{260} and OD_{260}/OD_{280}, respectively (section 2.4.1).

C. tropicalis genomic DNA (100 µg) was partially digested with Sau3A I (0.04 units/µg) for 1 h at 37°C and the reaction stopped by the addition of 0.5 M EDTA (pH 8.0) to 20 mM. After extraction and precipitation, the Sau3A I-digested DNA was heated to 68°C for 10 min, cooled to 20°C and loaded onto a 10.4 ml, 10-40% (w/v) linear sucrose gradient [sucrose solutions contained 1 M NaCl, 20 mM Tris-HCl (pH 8.0), 5 mM EDTA (pH 8.0)]. The DNA was size-fractionated by centrifugation at 86,200 x g_{av} in a Sorvall TST 41.14 rotor (26,000 rpm; Dupont/NEN) for 24 h at 20°C and 175 µl fractions collected from the bottom of the gradient. Those sets of 3 fractions containing ≈8, 15 or 20 kbp fragments were independently combined, dialyzed overnight at 4°C against 2 x 2 L of TE (pH 7.5), NaOAc:absolute ethanol precipitated and dissolved in sterile, distilled H_2O.

2.5.2 Construction of C. tropicalis Genomic DNA Libraries

Two independent libraries of C. tropicalis genomic DNA were constructed using size-fractionated, Sau3A I-digested DNA fragments (section 2.5.1).

The initial library was constructed by ligation of the ≈8, 15 or 20 kbp Sau3A I fragments into BamH I-digested pBR322 that had been pretreated with calf intestinal phosphatase (0.01 units/pmol of 5'-ends; section 2.4.4). Ligations which employed DNA concentrations of 5 to 10 µg/ml and vector:insert ratios of 2:1 were carried out overnight at 4°C in 20 µl of ligation buffer containing 0.5 units of T4 DNA ligase. Ligated DNA was introduced into E. coli HB101 cells by transformation with $CaCl_2$ (Mandel and Higa, 1970; section 2.4.5.3) and insertional activation [ampicillin resistance (Amp^r), tetracycline sensitivity (Tet^s)] used as a means to select those transformants harbouring inserts. A partial (≈1.7 kbp) cDNA fragment known to encode C. tropicalis fatty acyl-CoA oxidase (Kochubinski et al., 1985), isolated by polyacrylamide gel electrophoresis and radiolabeled by nick translation (sections 2.4.3 and 2.4.5.1), was used to screen DNA dot-blots (section 2.4.5.2) of this pBR322-based library for a full-length genomic recombinant encoding fatty acyl-CoA oxidase from C. tropicalis.

The second library was constructed using a modified protocol of that originally described by Frischauf et al. (1983) for library construction and recombinant selection using lambda EMBL3 as the cloning vector. To this end, ≈20 kbp Sau3A I fragments of C. tropicalis genomic DNA were ligated into phosphatased (0.01 units/pmol of 5'-ends; section 2.4.4) BamH I, EcoR I-digested lambda EMBL3 arms. Ligations which employed DNA concentrations of ≈50 µg/ml and vector:insert ratios of 3:1 were carried out overnight at 15°C in 10 µl of ligation buffer containing 1 unit of T4 DNA ligase. The following day, ligated DNA (≈0.2 µg) was packaged exactly as described by the product data sheets accompanying the Gigapack Gold in vitro packaging extract of lambda DNA. To screen the lambda EMBL3 library, E. coli NM539 cells (0.1 ml), grown overnight at 37°C in LBMM medium [LB medium supplemented with 10 mM $MgSO_4$ and 0.2% (w/v) maltose], were added to a mixture of 5 µl of packaged phage and 95 µl of phage dilution buffer [100 mM NaCl, 10 mM Tris-HCl (pH 7.9), 10 mM $MgSO_4$], followed by incubation for 20 min at room temperature. LB medium (2.5 ml) containing 0.75% (w/v) agar, maintained at 47°C, was subsequently added, the mixture plated on LB medium:agar plates and the plates incubated at 37°C until visible plaques appeared (≈8 h). Plaques were overlaid with dry nitrocellulose disks, marked, and when completely wetted, removed and placed

phage-side up in denaturing solution (0.5 M NaOH, 1.5 M NaCl). After 1 min, the disks were removed, placed in neutralizing solution [0.5 M Tris-HCl (pH 8.0), 1.5 M NaCl] for 8 min, rinsed in 3 x SSC for 5 min, air-dried and baked *in vacuo* for 2 h at 80°C. The partial (≈1.1 kbp) cDNA fragment of catalase recombinant C188-53 (section 4.1.2), isolated by agarose gel electrophoresis and radiolabeled using random primers (sections 2.4.3 and 2.4.5.1), was used to screen these plaque lifts for a full-length genomic recombinant encoding catalase from *C. tropicalis*.

2.5.3 Isolation of Poly(A)$^+$RNA from C. tropicalis

As detailed in section 2.5.1, total RNA was isolated from *n*-alkane-grown *C. tropicalis* by 3 M LiCl solubilization of total nucleic acids and subsequent precipitation on ice of single-stranded nucleic acids (Maccecchini et al., 1979). Poly(A)$^+$RNA, which served as the template for the primer-directed synthesis of cDNA, was isolated from total RNA by oligo(dT)-cellulose chromatography (Aviv and Leder, 1972) essentially as described by Maniatis et al. (1982). For this purpose, a 1 ml column of oligo(dT)-cellulose/type 3 equilibrated in binding buffer [10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1 mM EDTA, 0.5% (w/v) SDS] was poured and prewashed sequentially with 30 ml each of binding buffer, elution buffer [10 mM Tris-HCl (pH 7.5), 0.2% (w/v) SDS] and binding buffer again. Total RNA was adjusted to 1.0 mg/ml in 0.5 M

NaCl, 1 mM EDTA, 0.5% (w/v) SDS, heated to 65°C for 5 min, quick-cooled to room temperature and applied to the column. The effluent was collected, reheated to 65°C, quick-cooled to room temperature and reapplied to the column, twice. Subsequently, the column was washed with 30 ml of binding buffer and the bound poly(A)$^+$RNA eluted with 3 ml of elution buffer and precipitated overnight at -20°C upon the addition of a 1/10 volume of 2 M KOAc (pH 5.5) and 2.5 volumes of absolute ethanol. Precipitated poly(A)$^+$RNA was pelleted by centrifugation at 16,300 x g_{max} in a Sorvall HB-4 rotor (10,000 rpm) for 15 min at 4°C, washed with 70% (v/v) ethanol, dried under vacuum and dissolved in sterile, distilled H_2O. OD_{260} and OD_{280} measurements were used to assess the quantity, purity and yield (relative to starting amounts of total RNA) of poly(A)$^+$RNA. The usual yield of poly(A)$^+$RNA from total RNA was approximately 1%.

2.5.4 Construction of C. tropicalis Complementary DNA (cDNA) Libraries

Three independent cDNA libraries were constructed by primer extension of poly(A)$^+$RNA isolated from n-alkane-grown C. tropicalis (section 2.5.3). The first library, constructed by primer extension with $(dT)_{12-18}$, was screened in an attempt to isolate full-length cDNA recombinants encoding fatty acyl-CoA oxidase and catalase. The second and third libraries, constructed by mRNA extension with

primers C188 and C189 (Table 2.1.1), respectively, were used to isolate the complete 5'-end of the catalase gene from C. tropicalis. For each of these libraries, the primer-directed synthesis of cDNA was accomplished following the procedure of Gubler and Hoffman (1983) with modifications as detailed below.

Prior to the synthesis of first-strand cDNA, an annealing reaction (12.5 µl final volume) was set up containing 50 mM Tris-HCl (pH 8.3), 10 mM $MgCl_2$, 4 µg of poly(A)$^+$RNA and 25 µg/ml (≈55 pmol) of primer C188 or C189 [$(dT)_{12-18}$ was used at 50 µg/ml (≈250 pmol)]. This mixture was heated to 70°C for 5 to 10 min and allowed to cool slowly to 42°C. First-strand cDNA synthesis was subsequently carried out by increasing the annealing reaction to 75 µl (final volume) containing 50 mM Tris-HCl (pH 8.3), 10 mM $MgCl_2$, 10 mM DTT, 4 mM Na·pyrophosphate, 100 mM KCl, 1.25 mM each of dATP, dGTP and dTTP, 0.5 mM of dCTP and the annealed primer:template. The reaction was initiated by the addition of 22.5 units of avian myeloblastosis virus (AMV) reverse transcriptase, incubated at 42°C for 2 h and stopped by the addition of EDTA (pH 8.0) to 20 mM. Reaction products were extracted with H_2O-saturated phenol:chloroform:isoamyl alcohol and precipitated with ethanol from 2 M NH_4OAc (pH 5.5) as described by Okayama and Berg (1982). Primer extension was monitored

using a radioactive side reaction in which 1 μCi of [α-^{32}P]dCTP (410 Ci/mmol, 10 μCi/μl) was substituted for 0.5 mM dCTP in a scaled-down (1/10 volume) first-strand synthesis reaction. Radiolabeled cDNA was then analyzed by electrophoresis on a denaturing 5% (w/v) polyacrylamide-TBE gel (Maniatis et al., 1982). Successful first-strand synthesis was followed by second-strand cDNA synthesis in a final reaction volume of 100 μl containing 20 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$, 100 mM KCl, 50 μg/ml nuclease-free BSA, 100 μM each of dATP, dCTP, dGTP and dTTP, 0.15 mM nicotinamide adenine dinucleotide (NAD), 8.5 units/ml of E. coli RNase H, 10 units/ml of E. coli DNA ligase and the mRNA:cDNA hybrid. The reaction, initiated by the addition of 230 units/ml of E. coli DNA polymerase I, was incubated sequentially for 1 h at 12°C, followed by 1 h at 22°C and subsequently stopped by the addition of EDTA (pH 8.0) to 20 mM. Double-stranded cDNA was extracted and precipitated as described for first-strand cDNA and subjected to homopolymeric tailing, in the presence of dCTP, using terminal deoxynucleotidyl transferase (Villa-Komaroff et al., 1978). Accordingly, a 50 μl reaction was established in Gibco/BRL DNA tailing buffer (100 mM potassium cacodylate, 2 mM $CoCl_2$, 0.2 mM DTT) containing 200 μM dCTP and the double-stranded cDNA. The reaction, initiated by the addition of 400 units/ml of terminal

deoxynucleotidyl transferase, was incubated at 15°C for 15 min, stopped by the addition of EDTA (pH 8.0) to 10 mM and subsequently extracted and precipitated as indicated above. Equimolar amounts of dC-tailed, double-stranded cDNA and Pst I-digested, dG-tailed pBR322 were mixed at a final concentration of 1 µg/ml and annealed, in the presence of annealing buffer [Maniatis et al., 1982; 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM EDTA], by sequential incubation at 65°C for 5 min, 57°C for 2 h and 42°C for 1 h. Annealed DNA was introduced into E. coli RR1 cells by transformation with $CaCl_2$ (Mandel and Higa, 1970; section 2.4.5.3) and insertional inactivation (Amp^S, Tet^R) used to select those transformants harbouring inserts. Transformants of the primer C188-directed cDNA library were screened by DNA dot-blotting (section 2.4.5.2) using the 77 bp EcoR I, Pvu II-digested DNA fragment of catalase cDNA recombinant RU1:87 (Rachubinski et al., 1987; section 4) that had been ^{32}P-labeled by fill-in of the recessed 3'-end with DNA polymerase I Klenow fragment (section 2.4.5.1). Transformants of the primer C189-directed cDNA library were screened by colony hybridization (section 2.4.5.2) using the ^{32}P-end-labeled (section 2.4.5.1) oligonucleotide C231 (Table 2.1.1).

2.5.5 M13 Cloning and 2',3'-Dideoxynucleoside 5'-Triphosphate (ddNTP) Sequencing

Having screened the pBR322-based libraries (section 2.5.2 and 2.5.4) and identified potential recombinants harbouring genomic and cDNA fragments encoding C. tropicalis fatty acyl-CoA oxidase and catalase, respectively, the M13 cloning/sequencing system (Messing et al., 1977, 1981; Messing and Vieira, 1982; Yanisch-Perron et al., 1985) was employed to determine the DNA sequences of these genes. The relevant protocols are detailed below; however, more explicit procedural information may be found in the Gibco/BRL M13 Cloning/Dideoxy Sequencing Instruction Manual.

2.5.5.1 M13mp18/19 Cloning

Standard ligations (section 2.4.4) of restriction endonuclease-digested DNA fragments with compatibly-digested M13mp18/19 vector DNA (usually 10 ng of vector DNA per ligation such that DNA concentrations were 2 to 10 µg/ml and insert:vector ratios were at least 1:1) were carried out overnight at 4°C in 5 to 10 µl of ligation buffer containing 0.5 units of T4 DNA ligase. Ligations were stopped by the addition of EDTA (pH 8.0) to 20 mM, diluted to a final volume of 30 µl and the ligated M13 DNA introduced into E. coli JM 109 cells (occasionally JM 101 or JM 107 host cells were used) by transformation with $CaCl_2$ (section 2.4.5.3) as described below. Initially, an overnight culture of M13 host cells was diluted 1:100 into 100 ml of YT medium and grown at 37°C with shaking. When an OD_{660} of 0.3 to 0.4 was

obtained (requires ≈2 h), host cells were pelleted by centrifugation at 7,300 x g_{max} in a Sorvall SS-34 rotor (7,800 rpm) for 5 min at 4°C. Concomitantly, an exponentially growing culture of host cells was established by inoculating 4 ml of YT medium with 40 µl of the original overnight culture, followed by growth for 2 to 3 h at 37°C with shaking. Pelleted cells were resuspended in 50 ml of ice-cold 50 mM $CaCl_2$ (1/2 of the original volume), placed on ice for 20 min and recentrifuged. Cell pellets were resuspended in 10 ml of ice-cold 50 mM $CaCl_2$ (1/10 of the original volume) and stored at 4°C for 12 to 24 h prior to transformation. Transformation of host cells with M13 DNA was achieved by mixing 5.0 µl of the diluted ligation reaction with 0.3 ml of competent cells, followed by incubation on ice for 40 min and heat shock for 2 min at 42°C. Subsequently, 100 µl of 100 mM IPTG, 50 µl of 2% (w/v) Bluo-gal or X-gal (dissolved in dimethylformamide), 0.2 ml of exponentially growing host cells and 3 ml of YT medium:top agar [YT medium containing 0.75% (w/v) agar kept liquified at 45°C] were added, mixed and poured directly onto YT medium:agar plates (prewarmed to 37°C). The plates were allowed to solidify at room temperature and then incubated at 37°C overnight.

2.5.5.2 Preparation of M13 Phage DNA for Sequencing

Transformation of *E. coli* host cells with

recombinant M13mp18/19 DNA produces white plaques upon plating and growth on indicator plates (section 2.5.5.1). Single-stranded DNA templates for sequencing were prepared from these white plaques by small- or large-scale phage growth and subsequent DNA extraction. For small-scale isolations, distinct white plaques were picked into 4 ml of YT medium and grown for ≈8 h at 37°C with shaking. Cells were pelleted by centrifugation for 10 to 15 min at maximum speed in an IEC table-top clinical centrifuge and the supernatant decanted, recentrifuged and retained (to be used for the isolation of phage DNA or as a stock of phage particles for large-scale preparations). To 1.2 ml of supernatant, 160 µl of 20% (w/v) PEG 8000, 2.5 M NaCl was added, followed by incubation for 15 min at room temperature and microcentrifugation (14,000 rpm) for 30 min. The PEG-containing supernatant was completely discarded; the pellet was resuspended in 100 µl of TES buffer [20 mM Tris-HCl (pH 7.5), 100 mM EDTA, 10 mM NaCl], vortexed and extracted with 1 volume of H_2O-saturated phenol:chloroform:isoamyl alcohol and 1 volume of chloroform:isoamyl alcohol. The aqueous phase was removed, 10 µl of 3 M NaOAc (pH 5.2) and 250 µl of absolute ethanol added and the DNA precipitated in a dry ice:ethanol bath for 20 min. Precipitated DNA was pelleted by microcentrifugation for 10 min at 4°C, washed, dried under vacuum, dissolved in sterile, distilled H_2O and

analyzed by agarose [0.7%(w/v)] gel electrophoresis. For large-scale phage DNA isolations, an exponentially growing culture of host cells was established by inoculating 4 ml of YT medium with 40 µl of an overnight culture, followed by growth for 2 h at 37°C with shaking. Stock M13 recombinant phage (1 to 2 µl of the small-scale supernatant) was then added and growth continued for 3 h at 37°C. Subsequently, 100 µl of this culture and 9 ml of YT medium were added to 1 ml of "fresh" exponentially growing host cells (grown for 2 h only). This mixture was grown for 6 h at 37°C with shaking, the cells pelleted by centrifugation for 15 min at maximum speed in a table-top clinical centrifuge and the supernatant decanted and recentrifuged. To each 1.2 ml of supernatant, 225 µl of 20% (w/v) PEG 8000, 2.5 M NaCl was added, followed by incubation at room temperature for 20 min and centrifugation at 12,000 x g_{max} in a Sorvall SS-34 rotor (10,000 rpm) for 30 min at 20°C. Resultant pellets were resuspended in TES buffer, extracted and NaOAc:ethanol precipitated in a scaled-up manner similar to that described for small-scale phage DNA isolations. Typically, 1.2 ml of supernatant resulting from a large-scale preparation yielded 3 to 4 µg of single-stranded M13 phage DNA.

2.5.5.3 DideoxyNTP Sequencing of M13 Single-Stranded Phage DNA Using DNA Polymerase I Klenow Fragment

Sequencing the single-stranded M13 phage DNA was

accomplished using the 2',3'-dideoxynucleoside 5'-triphosphate chain-termination procedure of Sanger et al. (1977) in conjunction with DNA polymerase I Klenow fragment. Accordingly, an annealing reaction was established in 12.4 µl (final volume) of polymerase reaction buffer [10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 30 mM NaCl] containing 0.5 to 1.0 µg of single-stranded template DNA and 4 ng (≈0.7 pmol) of universal or template-specific primer (Table 2.1.1). The annealing mixture was incubated at 90°C for 5 min and allowed to cool slowly (over a period of 30 to 45 min) to room temperature. Sequencing reactions were subsequently carried out exactly as described by the Gibco/BRL M13 Cloning/Dideoxy Sequencing Instruction Manual using 1 unit of DNA polymerase I Klenow fragment per annealing reaction and [α-^{32}P]dCTP (410 Ci/mmol, 10 µCi/µl) as the radiolabel. Reactions were terminated by the addition of 5 µl of dideoxy stop buffer [0.1% (w/v) xylene cyanol, 0.1% (w/v) bromophenol blue, 10 mM EDTA, 95% (v/v) deionized formamide] and denatured by heating at 90°C for 3 to 5 min.

Standard sequencing gels (30 x 40 x 0.04 cm) were constituted by 19:1 acrylamide:N,N'-methylene-bis-acrylamide (w/w) diluted to a final concentration of 6 or 8% (w/v) acrylamide in TBE containing 8 M urea, 0.13% (w/v) AP and 0.013% (v/v) TEMED. Post-polymerization (requires ≈2 h), sequencing gels were pre-electrophoresed, using a BRL

sequencing gel electrophoresis system (model S0; Gibco/BRL) at 60 watts (≈40 mA) for 15 to 30 min. Routinely, 2.5 to 3.0 µl of heat-denatured sequencing reaction were loaded per sample well and electrophoresis, at 60 watts, monitored by the migration of the marker dyes. When appropriate, the apparatus was disassembled and the gel transferred from the glass plates to Whatman 3 mm chromatography paper. After drying under vacuum at 80°C for 30 to 40 min, the gel was exposed to Kodak X-Omat K-1 film at room temperature overnight.

2.5.6 Lambda EMBL3 DNA Isolation and DideoxyNTP Sequencing

2.5.6.1 Isolation and Purification of Lambda EMBL3 DNA

Having screened the lambda EMBL3 library of *C. tropicalis* genomic DNA (section 2.5.4) and identified potential recombinants harbouring genomic fragments encoding catalase, a large-scale preparation of recombinant lambda EMBL3 phage DNA was carried out according to the following procedure modified from Maniatis et al. (1982) and Davis et al. (1986). Initially, an exponentially growing culture of *E. coli* LE392 cells was established by inoculating 10 ml of LBMM medium with 50 µl of an overnight culture, followed by growth for 1 h at 37°C with shaking. Single plaques, picked from plaque-purified plates, were added directly to the exponential culture of host cells and allowed to grow for 6

h at 37°C with shaking. Chloroform (0.5 ml) was then added, the mixture shaken for 10 to 15 min at 37°C and the cells pelleted by centrifugation at 7,700 x g_{max} in a Sorvall SS-34 rotor (8,000 rpm) for 15 min at 20°C. The supernatant containing the phage particles was subsequently titered by plating a series of 10-fold dilutions onto LB:agar plates. Two plates containing $\approx 10^5$ plaques/plate were used for phage elution by adding 3 ml of SM medium [50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 8 mM $MgSO_4$, 0.01% (w/v) gelatin] to one plate, followed by agitation for 1 h at room temperature and subsequent withdrawal, reapplication and reagitation of that same 3 ml on the second plate. Eluted phage particles (100 µl) were added to a "fresh" exponentially growing (≈1 h) culture of E. coli LE392 cells (8 ml) and grown at 37°C with shaking until cell lysis occurred (≈7 h). After chloroform treatment and centrifugation as above, the supernatant was retained as a high titre ($\approx 10^9$ plaque forming units/ml) stock of phage particles. To begin the large-scale isolation, 500 ml of LBMM medium was inoculated with 1 ml of an overnight culture of LE392 cells and grown at 37°C with shaking until turbidity was just visible (45 to 60 min). Subsequently, 1.5 ml of stock phage particles were added and the culture grown at 37°C with shaking until lysis occurred (7 to 8 h). Lysis was completed by the addition of 10 ml of chloroform, followed by shaking at 37°C for 15 min, after

which time the bacterial debris was pelleted by centrifugation at 6,500 x g_{max} in a Sorvall GSA rotor (6,300 rpm) for 15 min at 20°C. Crystalline PEG 8000 and NaCl were dissolved in the resulting supernatant to yield final concentrations of 10% (w/v) and 1 M, respectively, and the lambda EMBL3 phage precipitated overnight on ice. Precipitated phage were pelleted by centrifugation at 11,500 x g_{max} in a Sorvall GSA rotor (8,400 rpm) for 30 min at 4°C and the phage pellet resuspended in a final volume of 32.8 ml of SM medium. Cesium chloride was added to yield a final concentration of 4.13 M and a gradient established by centrifugation at 132,000 x g_{av} in a Beckman VTi50 rotor (40,000 rpm) for 20 h at 20°C. Banded phage particles were withdrawn and dialyzed at room temperature against 2 x 1 L of 25 mM Tris-HCl (pH 8.0), 10 mM NaCl, 10 mM $MgCl_2$ for a total of 2 h. Phage DNA was then extracted exactly as described by Maniatis et al. (1982) and concentrated in the usual manner (section 2.4.2). If necessary, further purification was achieved by passage of the lambda EMBL3 phage DNA over a Gibco/BRL NACS PREPAC mini-column following the protocol specified by the manufacturer's instruction manual. Typically, a large-scale isolation (500 ml culture) yielded 500 µg of lambda EMBL3 phage DNA suitable for further manipulation.

2.5.6.2 Preparation of Double-Stranded DNA Templates for Sequencing

Subsequent to the isolation of recombinant lambda EMBL3 phage DNA, that ≈3 kbp BamH I, Sal I genomic DNA fragment identified by Southern blotting as encompassing the entire C. tropicalis catalase gene (section 4.3.1) was subcloned into the EcoR V site of the expression vector, pGEM-5Zf(+), in both orientations. Unidirectional digestion of these recombinant plasmids with E. coli exonuclease III was used to create nested sets of deletion mutants suitable for subsequent DNA sequencing of both strands. To achieve this, the protocol of Henikoff (1984) was followed, with modifications as detailed below. Initially, 10 µg of CsCl-purified, Apa I, Nco I-digested plasmid DNA was dissolved in 100 µl of buffer A [66 mM Tris-HCl (pH 8.0), 0.66 mM $MgCl_2$], 650 units (10 µl) of exonuclease III added and the mixture incubated at 37°C. Aliquots of 2.5 µl (≈250 ng of DNA) were removed at 30 s intervals, mixed with 7.5 µl of buffer B [0.2 M NaCl, 5 mM EDTA (pH 8.0)] and incubated at 70°C for 10 min to inactivate the enzyme. After the addition of 30 µl of absolute ethanol, the DNA was precipitated in a dry ice:ethanol bath, pelleted by microcentrifugation for 15 min at 4°C, washed with 70% (v/v) ethanol and dried under vacuum. Each pellet was dissolved in 50 µl of S1 nuclease buffer [50 mM NaOAc (pH 4.5), 200 mM NaCl, 1 mM $ZnCl_2$], 4.5 units (1 µl) of S1 nuclease added and the reaction allowed

to proceed for 30 min at room temperature. Reactions were terminated by the addition of 6 µl of buffer C [0.5 M Tris-HCl (pH 8.0), 0.125 M EDTA], extracted once with 50 µl of H_2O-saturated phenol:chloroform:isoamyl alcohol, precipitated in a dry ice:ethanol bath by the addition of 140 µl of absolute ethanol, microcentrifuged, washed and dried. Pellets were dissolved in 10 µl of buffer D [20 mM Tris-HCl (pH 8.0), 7 mM $MgCl_2$] and 2 µl of the sample analyzed by agarose [1.0% (w/v)] gel electrophoresis. One unit of DNA polymerase I Klenow fragment (in 1 µl) was added to the remainder of the resuspension, the mixture incubated at 37°C for 2 min, 1 µl of a solutior containing 0.125 mM of each dNTP added and incubation continued at 37°C for another 2 min. Subsequent to the addition of 10 µl of 5 x T4 DNA ligase buffer, 30 µl of sterile, distilled H_2O and 1 unit of T4 DNA ligase, blunt-end ligations were allowed to proceed overnight at room temperature. Competent E. coli MV1193 host cells were prepared and transformed with ligated DNA as described in sections 2.5.5.1 and 2.4.5.3, respectively, using a maximum of 20 µl of ligation mixture per 200 µl of competent cells. Plasmid DNA was isolated from transformants using the alkaline extraction mini-preparation procedure (section 2.4.5.4) and analyzed by restriction endonuclease digestion and agarose [1% (w/v)] gel electrophoresis. If necessary, plasmid DNA was further

purified before sequencing by treatment with 50 µg/ml of RNase A for 2 h at 37°C, followed by two sequential extractions with 1 volume each of H_2O-saturated phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol. Extracted plasmid DNA was then centrifuged through a Sephadex G-50 spin column (section 2.4.5.1), NH_4OAc:absolute ethanol precipitated (section 2.4.2) and dissolved in 54 µl of sterile, distilled H_2O (a sufficient volume for three double-stranded DNA sequencing reactions). Certain mini-preparations of plasmid DNA and all large-scale, CsCl-purified plasmids were used as is for subsequent double-stranded DNA sequencing reactions.

2.5.6.3 DideoxyNTP Sequencing of Double-Stranded DNA Using Modified T7 DNA Polymerase (Sequenase)

Sequencing double-stranded DNA templates was accomplished using the ddNTP chain-termination procedure of Sanger (1977) in conjunction with modified T7 DNA polymerase (Sequenase; Tabor and Richardson, 1987). Initially, double-stranded DNA templates were denatured, prior to primer:template annealing, according to the method of Zhang et al. (1988) as modified below. For this purpose, 2 to 3 µg of prepared template DNA (section 2.5.6.2) were denatured for 5 min at room temperature in 20 µl of denaturing solution [0.2 M NaOH, 0.2 mM EDTA (pH 8.0)] and then neutralized by the addition of 2 µl of NH_4OAc (pH 5.5).

After the addition of 60 µl of absolute ethanol, the denatured DNA was precipitated for 20 min in a dry ice:ethanol bath, microcentrifuged (14,000 rpm) for 10 min at 4°C, washed with 70% (v/v) ethanol and dried under vacuum. DNA pellets were dissolved in 7 µl of sterile, distilled H_2O and mixed with 1 µl of 0.5 pmol (≈2.5 ng)/µl of universal or template-specific primer (Table 2.1.1) and 2 µl of sequencing buffer [40 mM Tris-HCl (pH 7.5), 20 mM $MgCl_2$, 50 mM NaCl]. The primer:template annealing mixture was heated to 65°C for 2 min and allowed to cool slowly (over a period of 30 min) to <30°C. Subsequent labeling and termination reactions, in which [α-^{32}P]dATP (3,000 Ci/mmol, 10 µCi/µl) served as the radiolabel, were performed exactly as indicated in the USBC instruction manual entitled "Step-By-Step Protocols For DNA Sequencing With Sequenase". Termination reactions were stopped by the addition of 4 µl of stop solution [0.05% (w/v) xylene cyanol, 0.05% (w/v) bromphenol blue, 20 mM EDTA, 95% (v/v) formamide] and denatured by heating at 75°C for 2 min. The set up, electrophoresis and processing of sequencing gels was as described in section 2.5.5.3.

2.5.7 Computer Analysis of Sequence Data

The Beckman MicroGenie (Versions 6.0 and 7.0) Sequence Analysis Program (Queen and Korn, 1984; Beckman Instruments) was used for the preliminary analysis of DNA

sequence data. This routinely included: DNA sequence entry, translation and comparisons; the determination of restriction endonuclease sites and of nucleotide, amino acid and codon frequencies in a particular DNA sequence; and protein sequence comparisons.

2.6 PROTEIN TARGETING STUDIES

2.6.1 Expression of Catalase Sequences

2.6.1.1 Constructs Used for the In Vitro Expression of Catalase Sequences

Those constructs used for the *in vitro* expression of catalase sequences were generated by enzymatic deletion of various portions of the catalase gene from *C. tropicalis*. Subsequent *in vitro* transcription and translation of this family of constructs yielded catalase peptides which were either truncated progessively from the N- or C-terminus, truncated simultaneously from both the N- and C-terminus or truncated "internally" such that the fusion of different portions of the N- and C-terminus maintained the original reading frame. A description of the methods used to generate these constructs follows; however, details delineating the various constructs and their characteristics are presented in section 5.1.

As described in section 2.5.6.2, the DNA fragment encompassing the entire *C. tropicalis* catalase gene was

subcloned into the EcoR V site of the expression vector, pGEM-5Zf(+), in both orientations. That orientation in which the transcription of the catalase gene was under the control of the T7 RNA polymerase promoter was subsequently used for the *in vitro* expression of catalase sequences. Unidirectional digestion of this recombinant plasmid with *E. coli* exonuclease III (section 2.5.6.2) was used to create two constructs, pGEMC2-10b and -11b, possessing 68 and 28 bp of 5'-noncoding sequence, respectively, which would dictate the synthesis of full-length catalase (M_r 54944) upon *in vitro* transcription and translation (sections 2.6.1.2, 2.6.1.3 and 2.6.1.4). Five other constructs were similarly generated by sequential deletion from the 5'-end of the catalase gene to endogenous, in-frame ATG codons, which would ultimately function as initiator methionines upon *in vitro* transcription and translation and direct the synthesis of progressively smaller catalase peptides truncated from the N-terminus. These constructs, designated pGEMC2-16b, -16n, -14Af, -14Ab and -16Ac, generated peptides of M_r 37291, 35310, 22465, 18647 and 13626, respectively. Alternatively, restriction endonuclease digestion of the full-length construct, pGEMC2-11b, with either Aha II, EcoR I or Pvu II, followed by *in vitro* transcription from linearized DNA templates and subsequent translation of "fall-off" mRNA transcripts, resulted in the synthesis of

three progressively smaller catalase peptides truncated from the C-terminus. These constructs, designated pGEMC2-11bAha II, EcoR I and Pvu II, generated peptides of M_r 17504, 30613 and 47215, respectively. That construct which yielded the catalase peptide truncated simultaneously from both the N- and C-terminus, called pGEMC2-16bAH, was generated by gel-purification of an Aha II, Hinc II fragment derived from the parental recombinant, pGEMC2-16b, followed by conversion to blunt ends using DNA polymerase I Klenow fragment and subsequent ligation into Aat II, Nsi I-digested, mung bean nuclease-treated pGEM-7Zf(+). Upon <u>in vitro</u> transcription and translation, a catalase peptide of M_r 22446 resulted from initiation at methionine-159 and termination at the TAG codon inherent to the Nsi I site of pGEM-7Zf(+) (ie. the C-terminal residue of this peptide was valine-353). Those constructs yielding catalase peptides which were truncated "internally", such that the correct reading frame was maintained upon fusion of various portions of the N- and C-terminus, were generated by cleavage of the full-length construct, pGEMC2-11b, with restriction enzymes recognizing a minimum of two sites within the catalase gene. Accordingly, digestion of pGEMC2-11b with either EcoR I, to release an internal 360 bp fragment (followed by re-ligation with T4 DNA ligase), Kpn I, to release an internal 489 bp fragment (followed by re-ligation with T4 DNA ligase) or

Hinc II, EcoR I, to release an internal 1,139 bp fragment (followed by conversion to blunt ends using DNA polymerase I Klenow fragment and re-ligation with T4 DNA ligase) generated constructs pGEMC2-11bE$_D$, K$_D$ and HE$_D$, respectively. Upon in vitro transcription and translation, these constructs dictated the synthesis of catalase peptides of M_r 41409, 36338 and 11926, respectively.

All constructions were confirmed by double-stranded ddNTP sequencing (section 2.5.6.3) and immunoprecipitation of the cell-free products (section 2.3.4) resulting from the in vitro transcription and translation of each construct.

2.6.1.2 In Vitro Transcription

As indicated in section 2.6.1.1, those pGEMC2 constructs in which correct transcript synthesis was directed by the T7 RNA polymerase promoter were used for the in vitro transcription of catalase sequences. Attempts at in vitro transcription utilizing T7 RNA polymerase, the aforementioned catalase constructs as DNA templates, and procedures obtained from Promega, Pharmacia, Dr. D.W. Andrews (Department of Biochemistry, McMaster University, Hamilton, Ontario) and the work of Chamberlin and Ring (1973), led to the establishment of the following protocol for the consistent synthesis of full-length RNA transcripts. Reactions (usually 50 µl final volume) were carried out in transcription buffer [40 mM Tris-acetate (pH 8.1), 2 mM

spermidine, 50 ng/µl nuclease-free BSA] containing 3 mM each of ATP, CTP and UTP, 10 µM GTP, 20 mM $Mg(OAc)_2$, 10 mM 2-mercaptoethanol, 0.016 A_{260} units/µl of 5'-cap [G(5')ppp(5')G or m^7G(5')ppp(5')G], 0.2 µg/µl of calf liver tRNA, 1 unit/µl of placental RNase inhibitor (RNA guard or RNasin) and 0.2 µg/µl of supercoiled or linearized plasmid DNA. For subsequent rabbit reticulocyte lysate translations, the presence of Cl^- ions (ie. Tris-HCl or $MgCl_2$) and the use of either 5'-cap in the transcription reaction is allowable. For subsequent wheat germ extract translations, however, acetate-based reagents must be used for RNA synthesis since Cl^- ion concentrations >70 to 80 mM are inhibitory to the wheat germ system (Erickson and Blobel, 1983) and the use of G(5')ppp(5')G allows approximately 10-fold greater translation of message relative to m^7G(5')ppp(5')G (see section 5.2.2). Transcription components were mixed at room temperature to avoid the precipitation of spermidine and the pH of the complete mixture adjusted with 2 M Tris base to be in the range of 7.7 to 8.3, which is required for the optimal activity of T7 RNA polymerase. The reaction was initiated by the addition of T7 RNA polymerase to 0.4 units/µl and incubated at 38 to 40°C (optimum temperature) for 10 min, during which time the limiting concentration of GTP serves to increase the 5'-capping efficiency of each transcript.

The GTP concentration was then increased to 2 mM (to ensure transcript elongation) and the reaction further incubated for 80 min at 38 to 40°C. Forty minutes into this incubation period, an additional aliquot of T7 RNA polymerase (10 units) was added to increase the yield of RNA transcript and upon completion of the reaction, mRNA synthesis was quantitated as described in section 2.4.1. Following this procedure, 10 to 20 µg of full-length transcript per 50 µl reaction could be synthesized from each catalase construct. Prior to in vitro translation using the Promega rabbit reticulocyte lysate system (section 2.6.1.3), in vitro transcribed mRNA was purified by precipitation with NH_4OAc:absolute ethanol (section 2.4.2) and dissolved in sterile, distilled H_2O at a concentration of 1 µg/µl. By contrast, in vitro transcribed mRNA was used as is for subsequent in vitro translation in the S23 wheat germ extract system (section 2.6.1.4).

A similar protocol was also used to facilitate the in vitro expression of chloramphenicol acetyltransferase (CAT) and dihydrofolate reductase (DHFR), from plasmids TP408 and pSP65DHFR-1 (section 2.1.6), respectively, the transcription of which was under the control of an SP6 RNA polymerase promoter (Melton et al., 1984). For these constructs, the in vitro transcription reaction was also initiated by the addition of 0.4 units/µl of SP6 RNA

polymerase, however, a second addition of enzyme was not performed.

2.6.1.3 In Vitro Translation Using A Nuclease-Treated Rabbit Reticulocyte Lysate System

Micrococcal nuclease-treated cell-free translation lysates derived from New Zealand White rabbit reticulocytes (Pelham and Jackson, 1976) were purchased from Promega. Typically, translation assays were carried out in a final volume of 25 to 100 µl containing 70% (v/v) nuclease-treated reticulocyte lysate [which had been optimized with respect to concentrations of creatine phosphate, creatine kinase, hemin, calf liver tRNA, KOAc and $Mg(OAc)_2$] and the following translation components: 20 µM each of 19 unlabeled amino acids excluding methionine, 1 µCi/µl of L-[^{35}S]-methionine (1,100 Ci/mmol, 10 µCi/µl) and 0.5 to 1.0 µg of yeast total RNA, yeast poly(A)$^+$RNA or purified, *in vitro* transcribed catalase mRNA (section 2.6.1.2), which had been preheated to 60°C for 5 min and quick-cooled on ice (Mortensen et al., 1984). Translations were incubated at 30°C for 60 min, terminated by placement on ice and the extent of incorporation of radiolabeled methionine into acid-precipitable protein determined using a modification of the method of Mans and Novelli (1961). To this end, 2.5 µl of completed translation was diluted to 50 µl (1:20) in NP-40 mix [10 mM Tris-HCl (pH 8.5), 150 mM NaCl, 200 mM

methionine, 0.02% (w/v) NaN_3, 1% (w/v) NP-40], duplicate samples (20 µl each) of the diluted mixture spotted onto 2 cm x 2 cm disks of Whatman 3 mm chromatography paper and the disks placed into ice-cold 10% (w/v) TCA containing 20 mM methionine. After 30 min on ice for the precipitation of protein, the disks were transferred to boiling 5% (w/v) TCA containing 20 mM methionine and periodically stirred during the subsequent 20 min boiling period. The disks were then washed four times with 5% (w/v) TCA and twice with absolute ethanol, placed in boiling ether for 5 min, air-dried and counted in aqueous counting scintillant. The protein profiles resulting from the reticulocyte lysate translation of various RNA transcripts were subsequently examined by SDS-PAGE [10 or 15% (w/v) acrylamide gels; section 2.3.2] of approximately equal cpm from each translation (to a maximum of 2 µl of lysate), followed by fluorography (section 2.3.2.1).

2.6.1.4 In Vitro Translation Using a Nuclease-Treated Wheat Germ Extract System

Cell-free translation extracts of wheat germ embryos were prepared essentially as described by Erickson and Blobel (1983) with modifications as per a procedure obtained from Dr. D.W. Andrews, Department of Biochemistry, McMaster University, Hamilton, Ontario. Initially, solvent floatation was used to enrich the unprocessed wheat germ

(preferably, freshly harvested) for viable, intact embryos. For this purpose, 2.5 volumes of carbon tetrachloride (600 ml) and 1 volume of cyclohexane (240 ml) were mixed until Schlieren lines were no longer visible. Wheat germ (≈40 g) was then added to the solvent, mixed in with a glass rod and damaged embryos and contaminating endosperm fragments allowed to settle away from intact, floating embryos for 2 to 3 min. Optimally, 20 to 40% of the wheat germ should float (Erickson and Blobel; 1983); if the recovery is less than this, the solvent ratio should be adjusted by adding carbon tetrachloride until ≈30% of the wheat germ does float. Floating wheat germ was collected in a Büchner funnel, allowed to dry under vacuum for 30 min and either processed immediately or stored at -80°C.

Subsequent preparation of the wheat germ extract was conducted in a minimum amount of time at 4°C employing only autoclaved buffers and oven-baked utensils. Accordingly, 6 g of floated wheat germ was ground in a total of 10 ml of homogenization buffer [40 mM HEPES (pH 7.6), 100 mM KOAc, 1 mM $Mg(OAc)_2$, 2 mM $CaCl_2$, 4 mM DTT], using a small mortar and pestle, by sequentially adding 3 ml of buffer and grinding 30 s, then adding another 3 ml of buffer and grinding for 60 s and finally adding 4 ml of buffer and grinding for 1 to 2 min to obtain a consistent slurry. This slurry was scraped into a 15 ml Corex tube, centrifuged at 64,000 x g_{max} in a

Beckman JA-20 rotor (23,000 rpm) for 10 min at 4°C and the supernatant collected and recentrifuged as before. The supernatant recovered after the second centrifugation, known as the S23 wheat germ extract, was applied (4.5 ml maximum) to a 17 ml Sephadex G-25 (fine) desalting column (prepared in a 20 ml disposable syringe), which had been pre-equilibrated with 25 ml of column buffer [40 mM HEPES (pH 7.6), 100 mM KOAc, 5 mM $Mg(OAc)_2$, 4 mM DTT] by centrifugation at 480 x g_{max} in the JA-20 rotor (2,000 rpm) for 2 x 3 min intervals at 4°C. Once the S23 extract permeated the column, centrifugation was at 530 x g_{max} in the JA-20 rotor (2,100 rpm) for 3 min at 4°C. The desalted extract obtained was subsequently centrifuged at 64,000 x g_{max} in the JA-20 rotor (23,000 rpm) for 15 min at 4°C and 1 ml aliquots of the final S23 supernatant frozen in liquid nitrogen and stored at -80°C.

Micrococcal nuclease-treatment of thawed 1 ml aliquots of the extract was initiated by the sequential addition of $CaCl_2$ to 1 mM and 75 units (10 µl) of micrococcal nuclease [dissolved in 10 mM HEPES (pH 7.5)]. This mixture was incubated at 25°C for 5 min, after which time the nuclease was inactivated by the addition of EGTA to 2.2 mM and 200 µl aliquots were frozen and stored as above.

The optimum concentration of nuclease-treated, S23 wheat germ extract required for the efficient translation of

in vitro transcribed, full-length catalase mRNA was determined to be 20% (v/v) of the total translation mixture [the range tested was between 10 and 35% (v/v)]. Consequently, a typical 50 µl translation contained 20% (v/v) nuclease-treated, S23 wheat germ extract (10 µl), 20% (v/v) *in vitro* transcription reaction (10 µl) and 60% (v/v) wheat germ translation master mix (30 µl). The major constituents of each of these three components were: S23 wheat germ extract - 40 mM HEPES (pH 7.6), 100 mM KOAc, 5 mM $Mg(OAc)_2$, 4 mM DTT; *in vitro* transcription reactions - 40 mM Tris-HCl (pH 8.1), 2 mM spermidine, 20 mM $Mg(OAc)_2$, 10 mM 2-mercaptoethanol, 0.2 to 0.4 µg/µl of catalase mRNA; and wheat germ translation master mix - 6.7 mM HEPES (pH 7.6), 200 mM KOAc, 0.17 mM EDTA, 8.3 mM ATP and GTP, 16.7 mM creatine phosphate, 67 µM each of 19 unlabeled amino acids excluding methionine, 0.33 mCi of L-[^{35}S]-methionine or tran^{35}S-label (≈1,200 Ci/mmol, 10 µCi/µl), 6.7 mM Tris base, 0.17 µg/µl calf liver tRNA, 0.7 units/µl placental RNase inhibitor (RNA guard or RNasin), 0.07 µg/µl creatine kinase [dissolved in 10 mM Tris-acetate (pH 7.5), 50% (v/v) glycerol]. Therefore, each wheat germ extract translation (final pH ≈7.5) contained 12 mM HEPES, 12 mM Tris, 140 mM KOAc, 5 mM $Mg(OAc)_2$, 0.1 mM EDTA, 5 mM ATP, 5 mM GTP, 10 mM creatine phosphate, 0.8 mM DTT, 2.0 mM 2-mercaptoethanol, 0.4 mM spermidine, 40 µM each of 19 unlabeled amino acids

excluding methionine, 0.2 mCi of ^{35}S-methionine, 0.1 µg/µl calf liver tRNA, 0.4 units/µl placental RNase inhibitor, 40 ng/µl creatine kinase and was programmed with 2 to 4 µg of catalase mRNA. Translations were initiated by the addition of S23 extract, incubated at 24°C for 90 min and terminated by placement on ice. SDS-PAGE [10 or 15% (w/v) acrylamide gels; section 2.3.2] of 1 µl samples of each wheat germ extract translation, followed by fluorography (section 2.3.2.1), was used to determine the extent of incorporation of radiolabeled methionine into total protein, as well as examine the protein profiles resulting from the translation of different catalase mRNA transcripts.

2.6.1.5 Constructs Used for the In Vivo Expression of Catalase Sequences

The in vivo expression of catalase sequences was conducted in two different yeast strains, Candida albicans strain SGY243 (Kelly et al., 1987) and Saccharomyces cerevisiae strain DL1 (Veenhuis et al., 1987). To achieve this, those truncated catalase constructs generated for the in vitro expression of catalase sequences (section 2.6.1.1) were subcloned into the vector, pPGK (Gould et al., 1990a), such that their expression was placed under the control of the phosphoglycerate kinase (PGK) promoter (Dobson et al., 1982; Ogden et al., 1986). This was facilitated by agarose gel-purification of an Mse I fragment derived from each in

vitro construct (or an Mse I, Ban II fragment in the case of constructs pGEMC2-16bAH, pGEMC2-11bESS and pGEMC2-11bESL, see section 5.4.1), followed by conversion to blunt ends of each fragment using DNA polymerase I Klenow fragment and subsequent ligation with Bgl II-digested pPGK that had been similarly converted to blunt ends using DNA polymerase I Klenow fragment. pPGK-catalase recombinants were digested with Hind III, thereby liberating DNA fragments, composed of PGK promoter sequences upstream of catalase encoding sequences, which were agarose gel-purified and subsequently ligated into one of two different yeast expression vectors: either pMK22 (Kurtz et al., 1987), for expression in *C. albicans*, or YEp13 (Broach et al., 1979), for expression in *S. cerevisiae*. These fragments were converted to blunt ends with DNA polymerase I Klenow fragment prior to ligation with Sca I-digested pMK22 or ligated as is into Hind III-digested YEp13.

For each construct, the proper orientation of catalase sequences downstream of the PGK promoter was confirmed by restriction endonuclease digestion of the pPGK:catalase recombinant.

2.6.1.6 *Transformation of Yeast with Plasmid DNA*

Transformation of *C. albicans* strain SGY243 with each *in vivo* expression construct (section 2.6.1.5) was accomplished using the protoplast transformation procedure

of Beggs (1978), as modified by Kurtz et al. (1986). Accordingly, an overnight culture of C. albicans SGY243 cells, grown in complete medium (YEPD), was subcultured 1:100 into 50 ml of fresh YEPD medium and grown at 30°C for ≈5 h with shaking to achieve a cell density of 1 to 2 x 10^7 cells/ml (allow at least 2 population doublings). Cells were harvested by centrifugation at 3,000 rpm in an IEC table-top clinical centrifuge for 5 min at room temperature and the supernatant discarded. The cell pellet was then resuspended in 5 ml of SED [1 M sorbitol, 25 mM EDTA (pH 8.0), 50 mM DTT], gently rocked at 30°C for 10 min using a Tek-Pro tube rocker and recentrifuged as above. After discarding the supernatant, the cells were 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-glucuronidase (100 units/µl at pH 5.0) was added and the mixture gently rocked at 30°C for 30 min (at which time 80 to 90% of the cells should be converted to spheroplasts). Spheroplasts were subsequently pelleted at 1,000 rpm in the table-top clinical centrifuge for 5 min at 4°C and the spheroplast pellet resuspended in 5 ml of CaS solution [1 M sorbitol, 10 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$] and similarly recentrifuged. After resuspending the pellet in 0.5 ml of CaS solution, the spheroplasts were divided into 3 aliquots (≈165 µl each), 2 to 5 µg of transforming, CsCl-purified, plasmid DNA (section

2.6.1.5) were added to each aliquot and the solution gently mixed and allowed to stand for 15 min at room temperature. Subsequent to the addition of 1 ml of PEG solution [20% (w/v) PEG 3350, 10 mM $CaCl_2$, 10 mM Tris-HCl (pH 7.5)], the mixture was kept at room temperature for 15 min and then pelleted by centrifugation at 3,000 rpm for 5 min using the table-top clinical centrifuge. After total removal of the supernatant, the spheroplast pellet was resuspended in 100 µl of SOS medium [1 M sorbitol, 33% (v/v) YEPD, 6.5 mM $CaCl_2$] and the culture incubated at 30°C for 40 min with gentle rocking. SGY243 transformants were then selected on C. albicans rescue plates (section 2.2.1) by incubation at 30°C for 2 to 4 d.

Transformation of S. cerevisiae strain DL1 with each in vivo expression construct (section 2.6.1.5) was accomplished using the lithium acetate transformation procedure of Ito et al. (1983), as modified below. For this purpose, an overnight culture of S. cerevisiae DL1 cells, grown in complete medium (YEPD), was subcultured 1:25 into 50 ml of fresh YEPD medium and grown at 30°C for 3 to 4 h with shaking to achieve a cell density of ≈5 x 10^7 cells/ml (allow at least 2 population doublings). Cells were harvested by centrifugation at 3,000 rpm in a table-top clinical centrifuge for 5 min at room temperature, washed once with sterile, distilled H_2O, recentrifuged and the

supernatant removed completely. The cell pellet was resuspended in 0.5 ml (1/100 of the original volume) of LiOAc solution [10 mM Tris-HCl, 1 mM EDTA, 0.1 M LiOAc, pH 7.5], incubated at 30°C for 60 min with gentle rocking and divided into 5 aliquots (0.1 ml each). Subsequent to the addition of 50 µg of sonicated, salmon sperm, carrier DNA and 1 µg of transforming, CsCl-purified plasmid DNA (section 2.6.1.5) to each aliquot, the mixture was incubated at 30°C for 30 min without agitation, 0.7 ml of PEG/LiOAc solution [40% (w/v) PEG 3350, 0.1 M LiOAc, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 7.5)] added and the incubation at 30°C, without agitation, continued for another 30 min. The cells were then pelleted by microcentrifugation (14,000 rpm) for 2 to 3 s at room temperature, the supernatant discarded and the cell pellets washed twice with TE (pH 7.5). Washed pellets were resuspended in 100 µl of TE (pH 7.5), plated on minimal medium plates supplemented with 20 µg/ml L-histidine and uracil (section 2.2.1) and incubated at 30°C for 2 to 3 d to select DL1 transformants.

2.6.2 In Vitro Targeting of Catalase to Peroxisomes Isolated from C. tropicalis

Studies concerned with the *in vitro* targeting of catalase to peroxisomes isolated from *C. tropicalis* were conducted exactly as described by Small et al. (1987, 1988) and Small and Lazarow (1987) or, alternatively, as detailed

below. Initially, ^{35}S-labeled catalase and CAT [a negative (cytosolic) control for import] peptides were translated *in vitro* using a nuclease-treated wheat germ extract system (section 2.6.1.4). Concomitantly, peroxisomes were isolated from *C. tropicalis* and purified by sucrose or Nycodenz equilibrium density gradient centrifugation (section 2.2). Upon withdrawal from the sucrose or Nycodenz gradient, previously banded peroxisomes (or mitochondria) were diluted with 4 volumes of import buffer [Small et al., 1987; 2.5 mM MOPS (pH 7.2), 0.5 mM EDTA (pH 7.2), 0.5 M sucrose, 50 mM KCl] and then pelleted by centrifugation at 20,400 x g_{max} in a Sorvall SS-34 rotor (13,000 rpm) for 20 min at 4°C. After gentle resuspension of the pellets in 0.4 ml of ice-cold import buffer, the protein concentration, determined using the method of Bradford (section 2.3.1), was adjusted to 10 mg/ml with import buffer and the integrity of the peroxisomal membrane determined by assaying catalase latency (section 2.2.2.4). Typically, *C. tropicalis* peroxisomes exhibited latency values greater than 70%. Subsequently, 100 µg (10 µl) of peroxisomal protein were combined with 10,000 to 20,000 cpm each of ^{35}S-labeled catalase and CAT peptides (ie. 1 to 2 µl of each translation to yield approximately equivalent cpm) and the mixture diluted with import buffer to a final reaction volume of 40 µl (which contained ATP at a concentration of 1 mM). Five such

complete import reactions and two mock import reactions (without peroxisomes), similarly assembled for each catalase peptide, were incubated at 26°C for 30 min to allow protein association with or import into peroxisomes and then analyzed to determine the extent of association/import of each *in vitro* translation product with/into *C. tropicalis* peroxisomes.

The extent of the association of each *in vitro* translation product with purified peroxisomes was assessed according to the method of Andrews (1989). To this end, ice-cold import buffer or import buffer containing 8 M urea was added to two of the five complete import reactions to yield 68 µl final reaction volumes containing either 0 or 3 M urea, respectively. Each 68 µl volume was then layered onto 100 µl sucrose cushions of similar composition (ie. import buffer with or without urea), but containing 0.75 M sucrose. Gradients were subsequently centrifuged at 180,000 x g_{av} in a Beckman A-100/30 rotor (28 psi; Beckman Instruments) for 15 min at 4°C. After centrifugation, the gradients were collected as three 82 µl fractions; two fractions, called the top and middle fractions, resulted from removal of the sucrose cushion and the third fraction, called the bottom fraction, resulted from the addition of 82 µl of 10 mM Tris-HCl (pH 9.0), 1% (w/v) SDS to the pellet, followed by heating at 65°C for 10 min and resuspension.

Equal volumes of each fraction were subsequently analyzed by SDS-PAGE [10 to 15% (w/v) acrylamide gels; section 2.3.2], followed by fluorography (section 2.3.2.1).

The extent of import of each *in vitro* translation product into *C. tropicalis* peroxisomes was assessed by differential treatment of the remaining mock and complete import reactions. Mock import reactions 1 and 2 were differentially treated in that only reaction 2 was proteolyzed with 5 µg of proteinase K (5 mg/ml dissolved in H_2O) for 15 min on ice (reaction 1 was kept on ice only). Complete import reactions 3, 4 and 5 were differentially treated in that reaction 3 received no further treatment (similar to reaction 1 in that it is kept on ice only, but contains peroxisomes), reaction 4 was treated with 5 µg of proteinase K for 15 min on ice (similar to reaction 2 but contains peroxisomes) and reaction 5 was treated with 5 µg of proteinase K, 0.8% (w/v) deoxycholic acid (DOC) and 0.8% (w/v) Triton X-100 for 15 min on ice. Protease treatment was terminated by the sequential addition of 1 µl of PMSF (8 mg/ml, freshly made) and an equal volume of 2 x SDS-PAGE sample buffer (kept at 95 to 100°C), followed by the immediate boiling of each sample for 10 min. To maintain consistency, all samples, including those without externally added protease, were similarly treated. Equal volumes of each reaction were subsequently analyzed by SDS-PAGE and

fluorography, as indicated for the aforementioned association assay.

2.6.3 In Vivo Targeting of Catalase from C. tropicalis to the Peroxisomes of C. albicans and S. cerevisiae

Having generated those constructs necessary for the in vivo expression of catalase sequences (and controls-CAT, luciferase) and the corresponding C. albicans SGY243 and S. cerevisiae DL1 transformants (sections 2.6.1.5 and 2.6.1.6, respectively), in vivo PGK promoter-induced expression of catalase sequences was confirmed by growth of C. albicans and S. cerevisiae transformants in minimal medium and minimal medium supplemented with 20 µg/ml L-histidine and uracil (section 2.2.1), respectively, to induce the PGK promoter, followed by immunoblot analysis of cell lysates (section 2.3.3) to verify expression.

Subsequent studies concerned with the in vivo targeting of catalase from C. tropicalis (and controls) to the peroxisomes of C. albicans and S. cerevisiae were instigated by the growth of C. albicans and S. cerevisiae transformants in minimal medium and minimal medium supplemented with 20 µg/ml L-histidine and uracil (section 2.2.1), respectively, to induce the PGK promoter, followed by further growth of these starting cultures of C. albicans and S. cerevisiae transformants in YPBO medium and IM medium (section 2.2.1), respectively, to induce peroxisome

proliferation. Peroxisomes were then isolated from C. albicans and S. cerevisiae transformants and purified by sucrose and Nycodenz equilibrium density gradient centrifugation as described in section 2.2. During the peroxisome isolations, 20k x g supernatants were withdrawn and retained in their entirety; 20k x g pellets were resuspended as appropriate (section 2.2.2.1), 10 to 15% of the resuspension retained and the remainder purified on sucrose or Nycodenz gradients. The purified peroxisomes, 20k x g pellets and 20k x g supernatants obtained were subjected to SDS-PAGE [10 to 15% (w/v) acrylamide gels; section 2.3.2], followed by immunoblot analysis (section 2.3.3), to verify the intracellular location of *in vivo* expressed CAT (cytosolic; Gould et al., 1987, 1988) and luciferase (peroxisomal; Gould et al., 1990a) and determine the intracellular location of full-length and truncated C. tropicalis catalase peptides.

RESULTS AND DISCUSSION

3. CLONING AND SEQUENCING OF THE GENE ENCODING PEROXISOMAL ACYL-COENZYME A OXIDASE FROM THE YEAST CANDIDA TROPICALIS pK233

Candida tropicalis pK233 is capable of assimilating fatty acids or alkanes as the sole source of carbon and energy (see section 1.3.1.1). When grown on n-alkanes (C_{10}-C_{13}) or oleic acid, peroxisome proliferation in C. tropicalis is accompanied by the increased expression of those enzymes which constitute the peroxisomal β-oxidation pathway (Osumi et al., 1975; Kawamoto et al., 1978). In contrast, when this yeast is grown on glucose as the carbon source, peroxisomes are few in number and β-oxidation activity is low or undetectable (Tanaka et al., 1982; Dommes et al., 1983). As a consequence of this finding, efforts toward the cloning of those genes coding for peroxisomal proteins have involved the screening of C. tropicalis cDNA and genomic DNA libraries by differential DNA dot blot hybridization using [^{32}P]-labeled cDNA complementary to mRNA isolated from glucose- and alkane-grown yeast (Kamiryo and Okazaki, 1984; Rachubinski et al., 1985; Okazaki et al., 1986, 1987). In this manner, Rachubinski et al. (1985) identified several alkane-inducible cDNA recombinants; one partial recombinant, designated RU1:18, was further shown by hybridization-selection translation and immunoprecipitation

to encode acyl-coenzyme A oxidase (AOx), the first enzyme of the peroxisomal β-oxidation pathway. To extend this work, present efforts were directed toward cloning and sequencing the entire open reading frame and flanking regions of the gene encoding AOx. Having achieved this, a detailed analysis of the structure of the gene and the deduced primary structure of the enzyme was conducted.

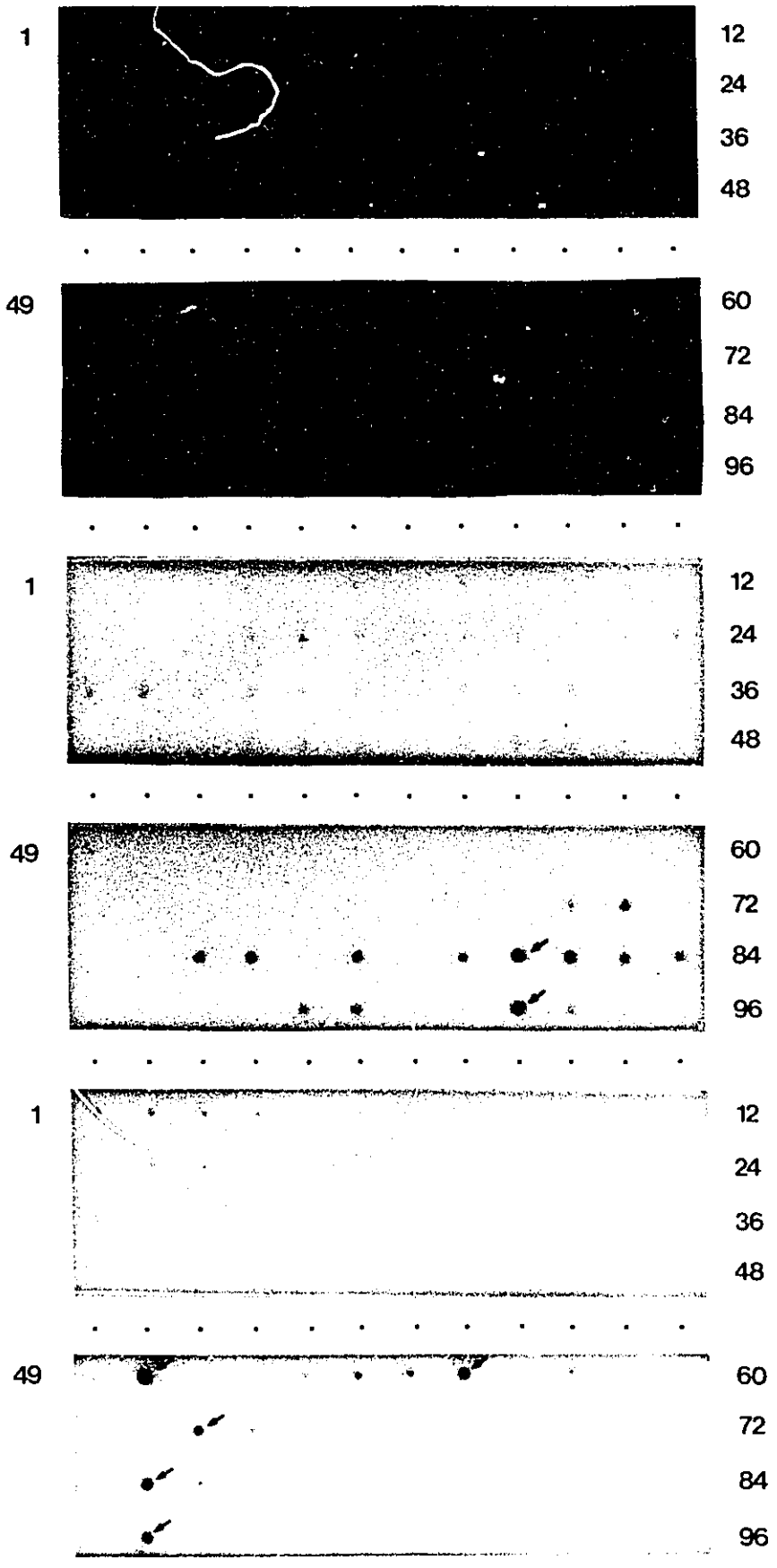
3.1 MOLECULAR CLONING OF THE GENE ENCODING AOx

3.1.1 Selection of a Genomic Clone Encoding AOx

The strategy employed to select a genomic clone encoding AOx entailed screening a genomic DNA library of n-alkane grown C. tropicalis using the partial ≈1.7 kbp Pst I insert of cDNA recombinant RU1:18 (Rachubinski et al., 1985) as the hybridization probe. Construction of the initial genomic DNA library involved ligation of size-fractionated (≈20 kbp), Sau3A I-digested DNA fragments into phosphatased, BamH I-digested pBR322 (section 2.5.2; Figure 3.1.3, upper portion), followed by transformation of competent E. coli HB101 cells. Selection of transformants by insertional inactivation (Amp^r, Tet^s) indicated that ligations performed at vector:insert ratios of 2:1 yielded approximately 70% recombinants. Assuming the total genome of C. tropicalis to be approximately 15,000 kbp in size (Kamiryo and Okazaki, 1984), and knowing that the recombinants harboured ≈20 kbp

genomic DNA inserts, implied that the screening of ≈1700 independent clones was necessary to obtain the particular sequence of interest at 99% probability (Clarke and Carbon, 1976). Nevertheless, DNA dot blot analysis of plasmid isolated from only 282 randomly selected clones of C. tropicalis genomic DNA (Figure 3.1.1), using nick translated, [α-^{32}P]dCTP-labeled, Pst I fragment of cDNA recombinant RU1:18 as the probe (henceforth called the AOx probe), was sufficient to select the recombinant harbouring the AOx gene. As seen in Figure 3.1.1, an intense hybridization signal was obtained with one recombinant, designated A76. The positive control for the hybridization, cDNA recombinant RU1:18 itself, generated a similarly strong signal (not shown) whereas the negative control, pBR322 plasmid alone, did not yield any significant hybridization signal (Figure 3.1.1, panels A, B and C, positions 1 and 96). Interestingly, weaker signals were observed with plasmid DNA isolated from a number of other genomic clones (Figure 3.1.1, arrows). The identities of these clones have not been determined; however, two explanations for the observed hybridization are plausible. Subsequent to sequencing most of the Cla I, Pst I fragment of cDNA recombinant RU1:18 (section 3.2.1, Figure 3.2.3), comparison of the sequence of the AOx cDNA probe to that of POX2, 4 and 5 (Okazaki et al., 1986, 1987) revealed nucleic acid

Figure 3.1.1. DNA dot blot analysis of plasmid isolated from 282 randomly selected clones of C. tropicalis genomic DNA. Plasmid (≈200 ng of DNA) from genomic clones, a pBR322 control (Panels A, B and C, positions 1 and 96) and AOx cDNA clone RU1:18 (Rachubinski et al., 1985; not shown) was isolated using the method of Holmes and Quigley (1981), denatured, neutralized and spotted onto nitrocellulose. Hybridization (42°C for 24 h) was with 6.3 x 10^5 cpm of nick translated AOx probe (specific activity = 2.1 x 10^7 cpm/μg cDNA). Washes were conducted at 42°C in 1 x SSC. Panels A, A-series genomic clones 2 to 95; Panels B, B-series genomic clones 2 to 95; Panels C, C-series genomic clones 2 to 95. Arrows indicate those weaker hybridization signals discussed in the text (section 3.1.1). For blots of panels A, exposure was for 18 h at -70°C with Kodak XAR-5 film and 2 intensifying screens. For blots of panels B and C, exposure was for 20 h at -70°C with Kodak XRP-1 film and 1 intensifying screen.

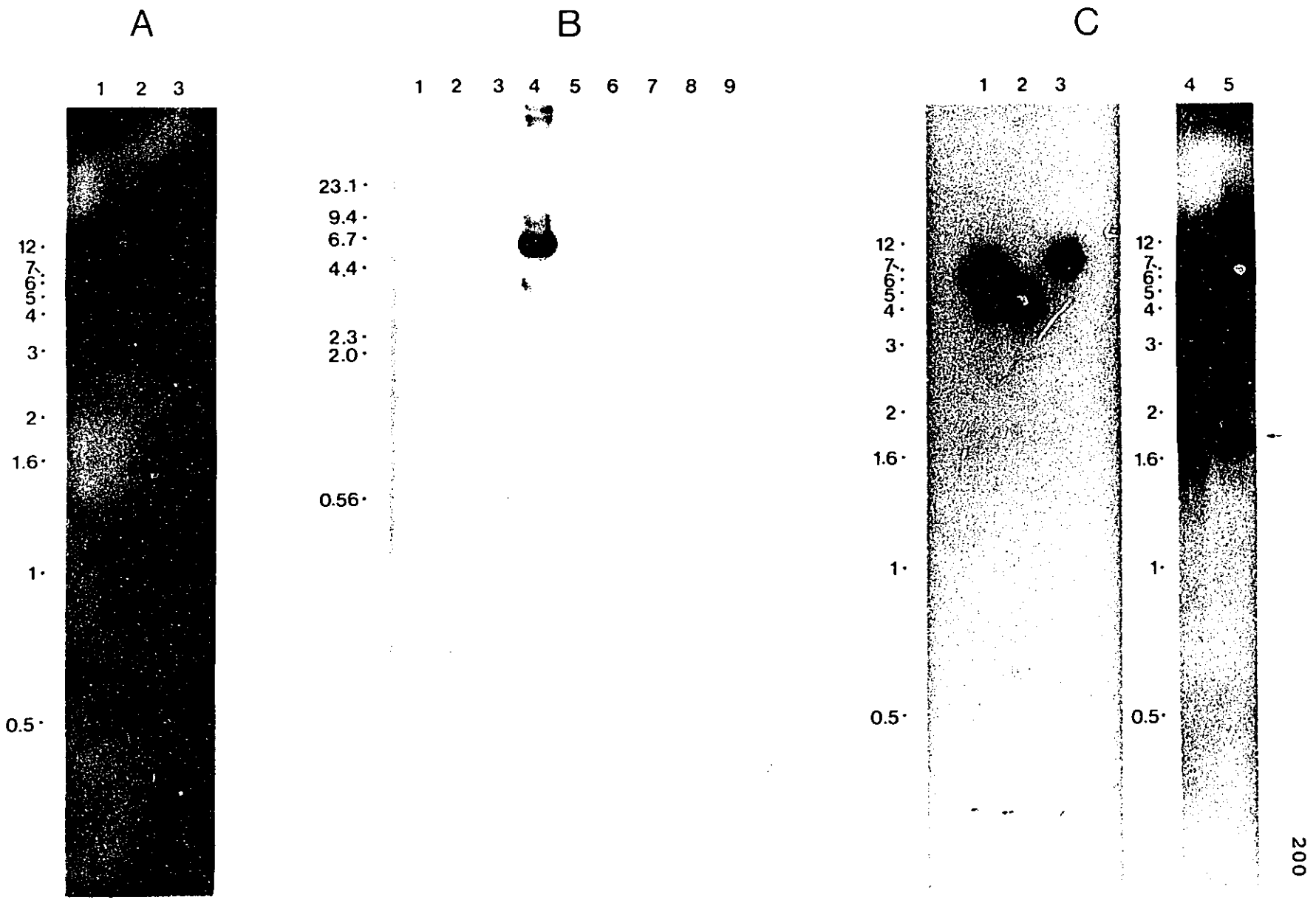

1
12
24
36
48
A
49
60
72
84
96
A
1
12
24
36
48
B
49
60
72
84
96
B
1
12
24
36
48
C
49
60
72
84
96
C

sequence identities of 61.2%, 97.0% and 61.4%, respectively. Consequently, certain of these signals may be the result of cross-hybridization with sequences encoding these other members of the peroxisomal acyl-coenzyme A oxidase multigene family of C. tropicalis (Okazaki et al., 1986, 1987). Alternatively, since these signals occur at relatively high frequency and persist under high stringency hybridization conditions (ie. in the presence of 45% formamide), a common sequence element may exist among certain C. tropicalis genes which results in cross-hybridization. However, sequencing most of the Cla I, Pst I fragment of the AOx cDNA probe (section 3.2.1, Figure 3.2.3) further revealed that this fragment contained AOx coding sequence only. This decreases the likelihood of the latter explanation accounting for the observed cross-hybridization.

3.1.2 Characterization of the Genomic Clone Encoding AOx

To confirm the identity of genomic clone A76 and to generate a partial restriction endonuclease cleavage map for the A76 recombinant plasmid, Southern blot analysis of A76 plasmid DNA and C. tropicalis genomic DNA was performed (Figure 3.1.2). As seen in Figure 3.1.2 (panel B, lane 4), the AOx probe hybridized to a single restriction fragment of ≈6 kbp upon Hind III-digestion of A76 plasmid DNA. Consistent with the observations derived from Figure 3.1.1,

Figure 3.1.2. Southern blot hybridization of the AOx probe with C. tropicalis genomic DNA (Panel A) and plasmid isolated from putative C. tropicalis genomic clones encoding AOx (Panel B) and AOx genomic clone, A76 (Panel C). Hybridization (55°C for 24 h) was with 2.0 x 10^6 cpm of nick translated AOx probe (specific activity = 2.5 x 10^7 cpm/µg cDNA). Washes were conducted at 55°C in 1 x SSC. Panel A, 5 µg of C. tropicalis genomic DNA/lane digested with Hind III (lane 1), Hind III and Cla I (lane 2) or Cla I (lane 3). Exposure was for 68 h at -70°C with Kodak XRP-1 film and 1 intensifying screen. Marker sizes in kbp (1 kbp DNA ladder) are indicated to the left. Panel B, ≈250 ng of plasmid DNA from several independent genomic clones digested with Hind III. Lanes: 1, a pBR322 control; 2-7, plasmid DNA from genomic clones designated A95, A87, A76, A56, A54 and A40, respectively; 8, molecular weight markers - ϕX 174-RF DNA Hae III digest; 9, molecular weight markers - lambda DNA Hind III fragments. Exposure was for 20 h at -70°C with Kodak XRP-1 film and 1 intensifying screen. Marker sizes in kbp (lambda DNA Hind III fragments) are indicated to the left. Panel C, ≈200 ng of A76 plasmid DNA digested with Hind III (lane 1), Hind III and Cla I (lane 2) or Cla I (lane 3). Exposure was for 16 h at -70°C with Kodak XRP-1 film and 1 intensifying screen. Lanes 4 and 5 represent a 15 x longer exposure of lanes 2 and 3, respectively. The arrow indicates the position of the ≈1.8 kbp Cla I restriction fragment that hybridizes with the AOx probe (see text, section 3.1.2). Marker sizes in kbp (1 kbp DNA ladder) are indicated to the left.


A
1 2 3
12
7
6
5
4
3
2
1.6
1
0.5
B
1 2 3 4 5 6 7 8 9
23.1
9.4
6.7
4.4
2.3
2.0
0.56
C
1 2 3
4 5
12
7
6
5
4
3
2
1.6
1
0.5
200

longer exposure (47 h) of the Southern blot of panel B indicated that recombinants A40, A54, A56, A87 and A95, which exhibited weaker hybridization signals upon DNA dot blot analysis, similarly possessed single, Hind III restriction fragments which weakly hybridized with the AOx probe (not shown). Pst I-, BamH I- or EcoR I-digestion of A76 plasmidDNA also generated a single genomic fragment to which the AOx probe hybridized; however, these fragments were all significantly larger [greater than 12 kbp for Pst I- or BamH I-digestion of A76 plasmid DNA and ≈8 kbp for EcoR I-digestion of A76 plasmid DNA (not shown)] than that generated by Hind III-digestion of A76 plasmid DNA and therefore, were less convenient for subcloning for the purpose of DNA sequencing.

Hybridization of the AOx probe with Hind III-digested C. tropicalis genomic DNA revealed the presence of two distinct bands (Figure 3.1.2, panel A, lane 1) ranging in size from between 6 and 7 kbp (similar to that observed with Hind III-digested A76 plasmid DNA) and greater than 12 kbp. Similarly, the AOx probe hybridized with two distinct restriction fragments of Hind III, Cla I-digested C. tropicalis DNA; however, these fragments possessed significantly smaller sizes of approximately 4 kbp and between 8 and 9 kbp (Figure 3.1.2, panel A, lane 2). Upon digestion of C. tropicalis DNA with Cla I, the AOx probe

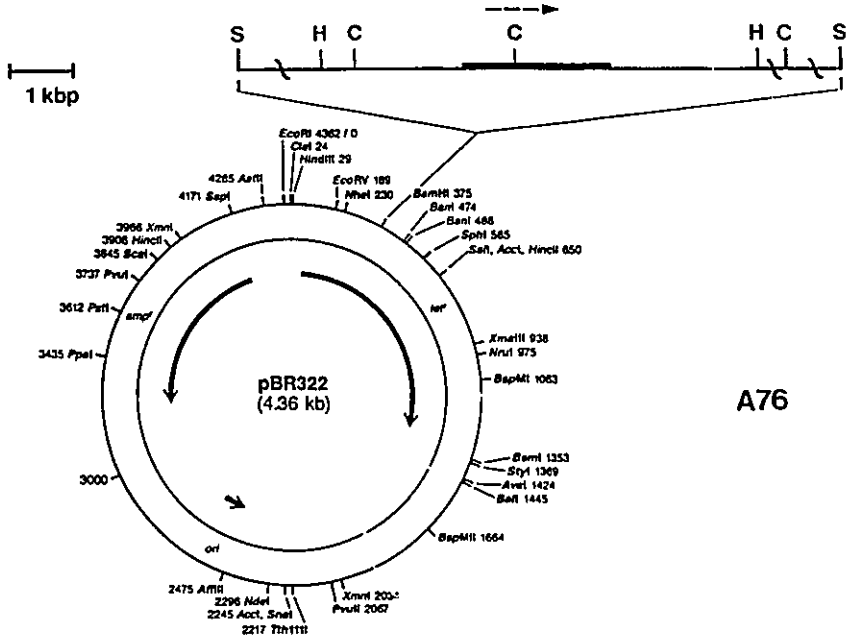
hybridized to a single fragment of between 7.5 and 8.5 kbp (Figure 3.1.2, panel A, lane 3). By comparison, digestion of A76 plasmid DNA with Hind III alone, Hind III and Cla I or Cla I alone generated a hybridization pattern which resembled that observed for similar digests of C. tropicalis genomic DNA. As seen in Figure 3.1.2 (panel C), Hind III-, Hind III, Cla I- or Cla I-digestion of A76 plasmid DNA yielded single fragments of between 6 and 7 kbp, ≈4 kbp or between 7.5 and 8.5 kbp, respectively (lanes 1, 2 and 3, respectively); however, the larger bands observed with Hind III- or Hind III, Cla I- digested C. tropicalis genomic DNA were not seen for similar digests of A76 plasmid DNA. At present, the origin of the greater than 12 kbp Hind III fragment and the 8 to 9 kbp Hind III, Cla I fragment of C. tropicalis genomic DNA which hybridized to the AOx probe is uncertain. Hybridization of the AOx probe with sequences encoding other members of the peroxisomal acyl-coenzyme A oxidase multigene family of C. tropicalis (Okazaki et al., 1986, 1987), which was similarly postulated for the cross-hybridization observed upon DNA dot blot analysis with the AOx probe (section 3.1.1), is the favoured explanation. Longer exposure (22.5 h) of the Southern blot of panel C indicated that Hind III, Cla I- and Cla I-digestion of A76 plasmid DNA generated an ≈1.8 kbp Cla I restriction fragment which also hybridized with the AOx probe (Figure 3.1.2,

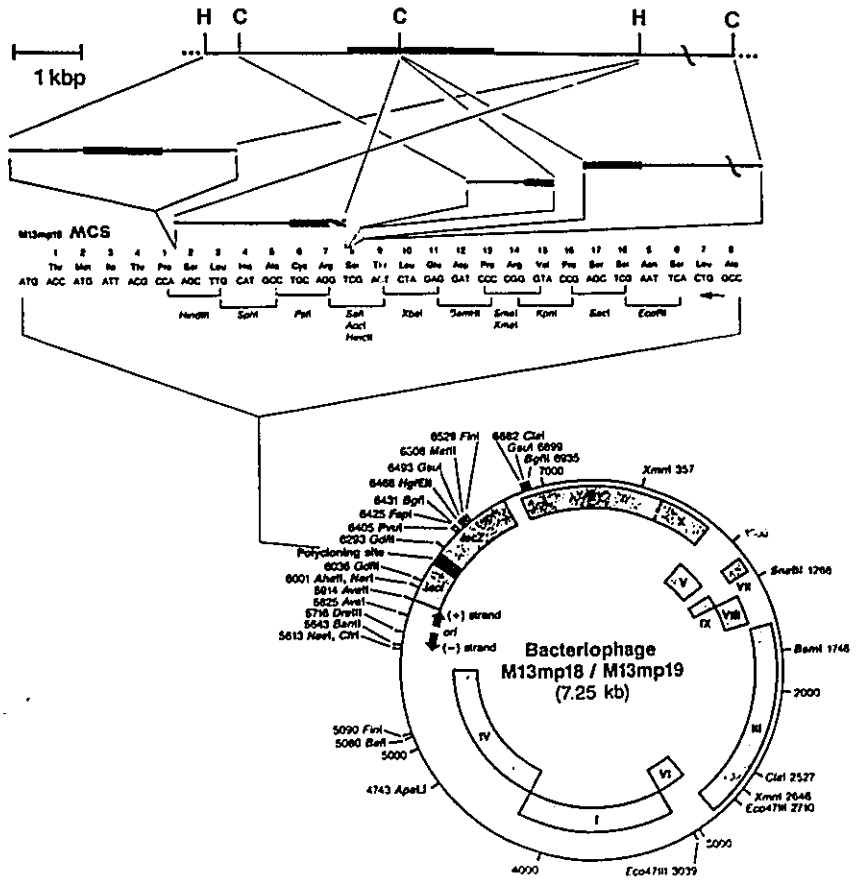
panel C, lanes 4 and 5, respectively, arrow). Thus, Cla I cleaved, at least once, a Hind III fragment of C. tropicalis genomic DNA which encompassed the entire AOx gene. This finding was suggestive of the partial restriction endonuclease cleavage map for recombinant A76 presented in the upper portion of Figure 3.1.3 (the orientation illustrated is correct but was fully confirmed only after having sequenced the AOx gene). Further, it was consistent with the partial restriction endonuclease cleavage map established by Rachubinski et al. (1985) for AOx cDNA recombinant RU1:18.

3.1.3 Strategy Used to Subclone the AOx Gene and cDNA Insert of Recombinant RU1:18

Having established that Cla I was a unique restriction endonuclease cleavage site within a Hind III fragment of A76 plasmid DNA which encompassed the AOx gene, the subcloning of four principal fragments of A76 plasmid DNA into the M13mp18/19 vectors (Norrander et al., 1983), for the purpose of DNA sequencing, was attempted (Figure 3.1.3, lower portion). Shotgun ligation of those fragments resulting from Cla I-digestion of A76 plasmid DNA into the compatible AccI site of M13mp19 resulted in subcloning the ≈1.8 kbp Cla I fragment, presumed to harbour the 5'-end of the AOx gene, in both orientations (Figure 3.1.3, lower portion); however, the 7.5 to 8.5 kbp Cla I fragment,

Figure 3.1.3. Schematic representation of A76 plasmid DNA and the subcloning strategy used to sequence the AOx gene. The upper portion of the figure represents: 1) construction of the C. tropicalis genomic DNA library by ligation of size-fractionated, Sau3A I-digested DNA fragments into phosphatased, BamH I-digested pBR322 (Sutcliffe, 1978) and 2) A76 plasmid DNA. The partial restriction endonuclease cleavage map established for recombinant A76 (see text, section 3.1.2) is illustrated, with the open reading frame of the AOx gene indicated by the thick line. Restriction endonuclease sites included are: C, Cla I; H, Hind III; S, Sau3A I. Deviations from the size scale of the map (\) are indicated. The dashed arrow represents the direction of AOx gene polarity. The orientation of the Sau3A I genomic DNA fragment within the BamH I site of pBR322 was not determined. The lower portion of the figure represents the strategy used to subclone fragments of A76 plasmid DNA into the multiple cloning site (MCS) of M13mp18/19 (Norrander et al., 1983) for the purpose of sequencing (see text, section 3.1.3). Those fragments subcloned, and their orientation within the M13mp19 MCS only, are illustrated [the M13mp18 MCS was used but is not shown; it possesses the opposite orientation to that of the M13mp19 MCS (ie. EcoR I-->Hind III)]. The arrow shown at the 3'-end of the MCS represents the direction of sequencing. Physical maps of pBR322 and M13mp18/19 were reproduced from Sambrook et al. (1989).


S H C C H C S
1 kbp
pBR322
(4.36 kb)
A76


H C C H C
1 kbp
MCS
Bacteriophage
M13mp18 / M13mp19
(7.25 kb)

presumed to harbour the 3'-end of the AOx gene, was not subcloned in either orientation. Similarly, attempts to subclone the ≈6 kbp Hind III fragment of A76 plasmid DNA, which encompassed the entire AOx gene, into Hind III-digested M13mp18, by either shotgun ligation of a Hind III-digest of A76 plasmid DNA or by fragment isolation from a nondenaturing polyacrylamide gel prior to ligation, were unsuccessful. Even though M13 is a filamentous bacteriophage and therefore, should not possess tight constraints on the size of the DNA that can be packaged efficiently (Messing et al., 1977), it is possible that an ≈6 kbp (or larger) DNA fragment approaches a size above which packaging of the circular, recombinant M13 DNA into the viral protein coat does not occur easily. Ultimately, the required subcloning, in both orientations, of the ≈4 kbp Cla I, Hind III fragment of A76 plasmid DNA, presumed to harbour the 3'-end of the AOx gene (Figure 3.1.3, lower portion), was achieved by fragment isolation from a non-denaturing polyacrylamide gel, followed by forced-orientation ligation into Acc I, Hind III-digested M13mp18 and 19. Having attained this, the AOx gene and flanking regions could be sequenced in their entirety along both DNA strands. Those sequence anomalies or ambiguities encountered were deciphered by subcloning (and sequencing) alternative fragments of A76 plasmid DNA. These included:

1) the ≈0.1 kbp Taq I fragment [used to confirm the nucleotide deletion which created a sequence frameshift in the AOx gene relative to the POX-4 gene (see section 3.2.3)] and 2) the ≈0.5 kbp Taq I fragment [used to confirm the nucleotide substitution which compensated for the AOx sequence frameshift (see section 3.2.3)].

To confirm the nucleotide sequence data generated for the AOx gene between nucleotides 757 and 1900, the ≈1.5 kbp Cla I, Pst I cDNA fragment of recombinant RU1:18 was subcloned (not shown) and sequenced. This was accomplished by fragment isolation (as above), followed by forced-orientation ligation into AccI, Pst I-digested M13mp18 and 19 to allow sequence determination in both directions. Taq I fragments of the cDNA insert of recombinant RU1:18, which were identical in size to those subcloned from A76 plasmid DNA (ie. ≈0.1 kbp and ≈0.5 kbp, respectively), were similarly subcloned into Acc I-digested M13mp19 and sequenced in an effort to substantiate the observed AOx sequence frameshift (section 3.2.3).

3.2 SEQUENCE DETERMINATION AND ANALYSIS OF THE AOx CODING REGION

3.2.1 Nucleotide Sequence of the Gene Encoding AOx

The physical map and sequencing strategy for the AOx gene and its flanking regions are presented in Figure 3.2.1 (upper portion, A76). Those A76 plasmid DNA fragments

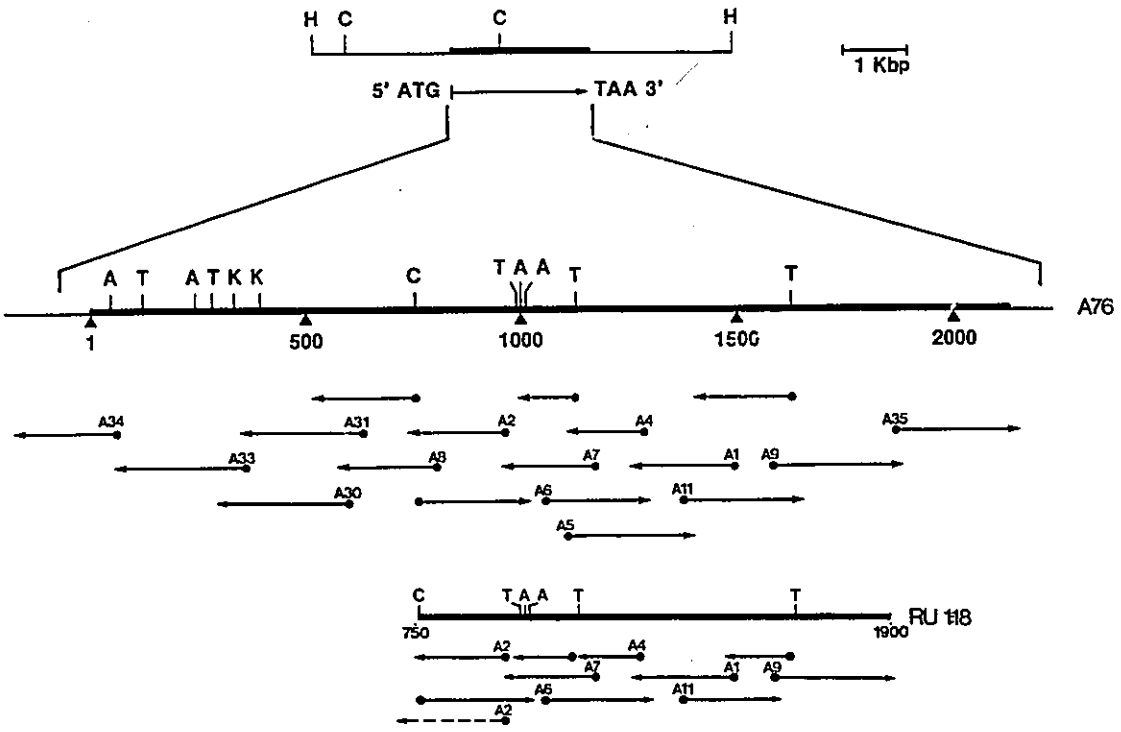


Figure 3.2.1. Physical map and sequencing strategy for the AOx gene and cDNA insert of recombinant RU1:18. Restriction endonuclease sites included are: A, Sau3A I; C, Cla I; H, Hind III; K, Kpn I; T, Taq I. The direction of AOx gene polarity is shown. The AOx coding region of recombinant A76 is indicated by the thick line. That portion of recombinant 1:18 which was sequenced represented AOx coding region only and is also indicated by a thick line. The base pair numbering of recombinants A76 and RU1:18 is presented in Figures 3.2.2 and 3.2.3, respectively. The start points (closed circles), direction and extent of single-stranded dideoxy sequence determinations are indicated by continuous arrows. Sequence data were generated using specific AOx(A) primers (labeled start points) or M13 universal, forward (-20 or -40) primers (non-labeled start points). Exact annealing positions and sequences of the AOx(A) sequence primers are compiled in Table 2.1.1 (sequences of the universal primers are also presented in Table 2.1.1). The double-stranded dideoxy sequencing which was conducted "across" the Cla I site of the cDNA insert of recombinant RU1:18 (see text, section 3.2.1) is represented by the dashed arrow.

subcloned into the M13mp18/19 vectors (section 3.1.3) were sequenced using the single-stranded ddNTP chain-termination method of Sanger et al. (1977). This involved the use of M13 universal forward (-20 or -40) sequencing primers or specific AOx (A) sequencing primers (Figure 3.2.1) synthesized complementary to a previously determined nucleotide sequence of the AOx gene [see Table 2.1.1 for the AOx (A) primer annealing positions and sequences]. Collectively, 78% of the entire sequence was determined in the 3' to 5' (noncoding) direction whereas 60% of the entire sequence was determined in the 5' to 3' (coding) direction, with each nucleotide sequenced at least twice. Those remaining nucleotides, which were sequenced in only one direction (ie. from one strand only), were determined twice and then compared to both the POX4 DNA sequence (Okazaki et al., 1986), which was found to exhibit an overall nucleic acid identity of 98.4% with that of the AOx gene (see below), and the sequence of the cDNA insert of recombinant RU1:18 (Figure 3.2.3). Any discrepancies encountered were resolved by resequencing at least one additional time. Double-stranded ddNTP sequencing, involving primer A2 (Figure 3.2.1, dashed arrow), was conducted "across" the Cla I site of the cDNA insert of recombinant RU1:18 to ensure that a small Cla I restriction fragment(s) had not been unknowingly excluded during the subcloning of A76 plasmid

DNA fragments.

Translation in the three possible reading frames revealed the existence of a single open reading frame (ORF), without intervening sequences, of 2,127 nucleotides (Figure 3.2.2) stretching from the putative ATG initiation codon to a TAA termination codon (nucleotides 2128 to 2130). The length of the ORF was identical to that of C. tropicalis POX4 (Okazaki et al., 1986), which has similarly been shown (along with C. tropicalis POX2 and POX5) to be devoid of any intervening sequence(s) (Okazaki et al., 1986, 1987). Comparison of the sequence of the ORF with that of POX2, 4 and 5 revealed nucleic acid sequence identities of 63.7%, 98.4% and 69.3%, respectively, which confirmed its AOx-encoding capacity. Moreover, the 2,127 nucleotide ORF was consistent with the ≈2.2 kb size of the AOx-encoding mRNA identified by Northern blot analysis (Rachubinski et al., 1985). The partial nucleotide sequence of the cDNA insert of recombinant RU1:18 (Figure 3.2.3), generated according to the sequencing strategy shown in Figure 3.2.1 (lower portion), was identical to that of nucleotides 757 to 1,900 of the ORF of the AOx gene with only one exception; the nucleotide at position 807 of the cDNA was a C, whereas that same position of the gene was a T. Translation of the cDNA sequence (Figure 3.2.3) revealed a partial, single ORF of identical amino acid sequence (residues 253 to 633) to

```
                                  -150                                                                    -92
                                  TATAAAATCAACTCCATTTCCCTAAAATCTCCCCAGATTCTAAAAACAACTTCTTCTCTTC
                                                                                                          -1
TGCTTTTCCTTTTTTTTTGTTATATTTATTTACCATCCCTTTTTTTTGAATAGTTATTCCCCACTAACATTGTTCAAATCTTCACGACATA
                                                                                                          69
ATG ACT TTT ACA AAG AAA AAC GTT AGT GTA TCA CAA GGT CCT GAC CCT AGA TCA TCC ATC CAA AAG GAA
Met Thr Phe Thr Lys Lys Asn Val Ser Val Ser Gln Gly Pro Asp Pro Arg Ser Ser Ile Gln Lys Glu
                                                                                                         138
AGA GAC AGC TCC AAA TGG AAC CCT CAA CAA ATG AAC TAC TTC TTG GAA GGC TCC GTC GAA AGA AGT GAG
Arg Asp Ser Ser Lys Trp Asn Pro Gln Gln Met Asn Tyr Phe Leu Glu Gly Ser Val Glu Arg Ser Glu
                                                                                                         207
TTG ATG AAG GCT TTG GCC CAA CAA ATG GAA AGA GAC CCA ATC TTG TTC ACA GAC GGC TCC TAC TAC GAC
Leu Met Lys Ala Leu Ala Gln Gln Met Glu Arg Asp Pro Ile Leu Phe Thr Asp Gly Ser Tyr Tyr Asp
                                                                                                         276
TTG ACC AAG GAC CAA CAA AGA GAA TTG ACC GCC GTC AAG ATC AAC AGA ATC GCC AGA TAC AGA GAA CAA
Leu Thr Lys Asp Gln Gln Arg Glu Leu Thr Ala Val Lys Ile Asn Arg Ile Ala Arg Tyr Arg Glu Gln
                                                                                                         345
GAA TCC ATC GAC ACT TTC AAC AAG AGA TTG TCC TTG ATT GGT ATC TTT GAC CCA CAG GTC GGT ACC AGA
Glu Ser Ile Asp Thr Phe Asn Lys Arg Leu Ser Leu Ile Gly Ile Phe Asp Pro Gln Val Gly Thr Arg
                                                                                                         414
ATT GGT GTC AAC CTC GGT TTG TTC CTT TCT TGT ATC AGA GGT AAC GGT ACC ACT TCC CAA TTG AAC TAC
Ile Gly Val Asn Leu Gly Leu Phe Leu Ser Cys Ile Arg Gly Asn Gly Thr Thr Ser Gln Leu Asn Tyr
                                                                                                         483
TGG GCT AAC GAA AAG GAA ACC GCT GAC GTT AAA GGT ATC TAC GGT TGT TTC GGT ATG ACC GAA TTG GCC
Trp Ala Asn Glu Lys Glu Thr Ala Asp Val Lys Gly Ile Tyr Gly Cys Phe Gly Met Thr Glu Leu Ala
                                                                                                         552
CAC GGT TCC AAC GTT GCT GGT TTG GAA ACC ACC GCC ACA TTT GAC AAG GAA TCT GAC GAG TTT GTC ATC
His Gly Ser Asn Val Ala Gly Leu Glu Thr Thr Ala Thr Phe Asp Lys Glu Ser Asp Glu Phe Val Ile
                                                                                                         621
AAC ACC CCA CAC ATT GGT GCC ACC AAG TGG TGG ATT GGT GGT GCT GCT CAC TCC GCC ACC CAC TGT TCT
Asn Thr Pro His Ile Gly Ala Thr Lys Trp Trp Ile Gly Gly Ala Ala His Ser Ala Thr His Cys Ser
                                                                                                         690
GTC TAC GCC AGA TTG ATT GTT GAC GGT CAA GAT TAC GGT GTC AAG ACT TTT GTT GTC CCA TTG AGA GAC
Val Tyr Ala Arg Leu Ile Val Asp Gly Gln Asp Tyr Gly Val Lys Thr Phe Val Val Pro Leu Arg Asp
                                                                                                         759
TCC AAC CAC GAC CTC ATG CCA GGT GTC ACT GTT GGT GAC ATT GGT GCC AAG ATG GGT AGA GAT GGT ATC
Ser Asn His Asp Leu Met Pro Gly Val Thr Val Gly Asp Ile Gly Ala Lys Met Gly Arg Asp Gly Ile
                                                                                                         828
                                                                  (C)
GAT AAC GGT TGG ATT CAA TTC TCC AAC GTC AGA ATC CCA AGA TTC TTT ATG TTG CAA AAG TTC TGT AAG
Asp Asn Gly Trp Ile Gln Phe Ser Asn Val Arg Ile Pro Arg Phe Phe Met Leu Gln Lys Phe Cys Lys
                                                                                                         897
GTT TCT GCT GAA GGT GAA GTC ACC    CCA CCT TTG GAA CAA TTG TCT TAC TCC GCC TTG TTG GGT GGT
Val Ser Ala Glu Gly Glu Val Thr Leu Pro Pro Leu Glu Gln Leu Ser Tyr Ser Ala Leu Leu Gly Gly
                                                                                                         966
AGA GTC ATG ATG GTT TTG GAC TCC TAC AGA ATG TTG GCT AGA ATG TCT ACC ATT GCC TTG AGA TAC GCC
Arg Val Met Met Val Leu Asp Ser Tyr Arg Met Leu Ala Arg Met Ser Thr Ile Ala Leu Arg Tyr Ala
                                                                                                        1035
ATT GGT AGA AGA CAA TTC AAG GGT GAT AAT GTC GAT CCA AAC GAT CCA AAT GCT TTG GAA ACC CAA TTG
Ile Gly Arg Arg Gln Phe Lys Gly Asp Asn Val Asp Pro Asn Asp Pro Asn Ala Leu Glu Thr Gln Leu
                                                                                                        1104
ATA GAT TAC CCA TTG CAC CAA AAG AGA TTG TTC CCA TAC TTC GTG CCG CCT ATG TCG TCT CCG CTG GTG
Ile Asp Tyr Pro Leu His Gln Lys Arg Leu Phe Pro Tyr Phe Val Pro Pro Met Ser Ser Pro Leu Val
                                                                                                        1173
CCC TCA AGG TTG AAC ACA CCA TCG AGA CCA CCT TGG CTG AAT TGG ACG CTG CCG TTG AAA AGA ACG ACA
Pro Ser Arg Leu Asn Thr Pro Ser Arg Pro Pro Trp Leu Asn Trp Thr Leu Pro Leu Lys Arg Thr Thr
```

```
                                                                    1242
CCA AGG CTA* ATC TTT AAG TCT ATT GAC GAC ATG AAG TCA TTG TTT GTT GAC TCT GGT TCC TTG AAG TCC
Pro Arg Leu Ile Phe Lys Ser Ile Asp Asp Met Lys Ser Leu Phe Val Asp Ser Gly Ser Leu Lys Ser
                                                                    1311
ACT GCC ACT TGG TTG GGT GCT GAA GCC ATT GAC CAA TGT AGA CAA GCC TGT GGT GGT CAC GGT CAC TCT
Thr Ala Thr Trp Leu Gly Ala Glu Ala Ile Asp Gln Cys Arg Gln Ala Cys Gly Gly His Gly His Ser
                                                                    1380
TCC TAC AAC GGT TTC GGT AAG GCC TAC AAC GAT TGG GTT GTC CAA TGT ACC TGG GAA GGT GAC AAC AAT
Ser Tyr Asn Gly Phe Gly Lys Ala Tyr Asn Asp Trp Val Val Gln Cys Thr Trp Glu Gly Asp Asn Asn
                                                                    1449
GTC TTG GGC ATG AGT GTT GGT AAG CCA ATT GTC AAG CAA GTC ATC AGC ATT GAA GAT GCC GGC AAG ACC
Val Leu Gly Met Ser Val Gly Lys Pro Ile Val Lys Gln Val Ile Ser Ile Glu Asp Ala Gly Lys Thr
                                                                    1518
GTC AGA GGC TCC ACT GCT TTC TTG AAC CAA TTG AAG GAA TAC ACT GGT TCC AAC AGC TCC AAG GTT GTT
Val Arg Gly Ser Thr Ala Phe Leu Asn Gln Leu Lys Glu Tyr Thr Gly Ser Asn Ser Ser Lys Val Val
                                                                    1587
TTG AAC ACT GTT GCT GAC TTA GAC GAC ATC AAG ACT GTC ATC AAG GCT ATT GAA GTT GCC ATC ATC AGA
Leu Asn Thr Val Ala Asp Leu Asp Asp Ile Lys Thr Val Ile Lys Ala Ile Glu Val Ala Ile Ile Arg
                                                                    1656
TTA TCC CAA GAA GCT GCT TCT ATT GTC AAG AAG GAA TCT TTC GAC TAT GTC GGT GCT GAA TTG GTT CAA
Leu Ser Gln Glu Ala Ala Ser Ile Val Lys Lys Glu Ser Phe Asp Tyr Val Gly Ala Glu Leu Val Gln
                                                                    1725
CTC TCC AAG TTG AAG GCT CAC CAC TAC TTG TTG ACT GAA TAC ATC AGA AGA ATT GAC ACC TTT GAC CAA
Leu Ser Lys Leu Lys Ala His His Tyr Leu Leu Thr Glu Tyr Ile Arg Arg Ile Asp Thr Phe Asp Gln
                                                                    1794
AAG GAG TTG GCT CCA TAC TTG ATA ACC CTC GGT AAG TTG TAC GCT GCC ACT ATT GTC TTG GAC AGA TTT
Lys Glu Leu Ala Pro Tyr Leu Ile Thr Leu Gly Lys Leu Tyr Ala Ala Thr Ile Val Leu Asp Arg Phe
                                                                    1863
GCT GGT GTC TTT TTG ACT TTC AAC GTT GCC TCC ACC GAA GCC ATC ACT GCT TTG GCC TCT GTG CAA ATT
Ala Gly Val Phe Leu Thr Phe Asn Val Ala Ser Thr Glu Ala Ile Thr Ala Leu Ala Ser Val Gln Ile
                                                                    1932
CCA AAG TTG TGT GCT GAA GTC AGA CCA AAC GTT GTT GCT TAC ACC GAC TCC TTC CAA CAA TCC GAC ATG
Pro Lys Leu Cys Ala Glu Val Arg Pro Asn Val Val Ala Tyr Thr Asp Ser Phe Gln Gln Ser Asp Met
                                                                    2001
ATT GTC AAT TCT GCT ATT GGT AGA TAC GAT GGT GAC ATC TAT GAG AAC TAC TTT GAC TTG GTC AAG TTG
Ile Val Asn Ser Ala Ile Gly Arg Tyr Asp Gly Asp Ile Tyr Glu Asn Tyr Phe Asp Leu Val Lys Leu
                                                                    2070
CAG AAC CCA CCA TCC AAG ACC AAG GCT CCT TAC TCT GAT GCT TTG GAA GCC ATG TTG AAC AGA CCA ACC
Gln Asn Pro Pro Ser Lys Thr Lys Ala Pro Tyr Ser Asp Ala Leu Glu Ala Met Leu Asn Arg Pro Thr
                                                                    2136
TTG GAC GAA AGA GAA AGA TTT CAA AAG TCT GAT GAA ACC GCT GCT ATC TTG TCC AAG TAA GAT AAG
Leu Asp Glu Arg Glu Arg Phe Gln Lys Ser Asp Glu Thr Ala Ala Ile Leu Ser Lys End
```

Figure 3.2.2. Complete nucleotide sequence of the AOx gene and the deduced amino acid sequence of AOx. Nucleotides are numbered in the 5' to 3' direction from the ATG start codon. The 5'- and 3'-flanking regions are shown; nucleotides 5' to position 1 are indicated by negative numbers. Within the 5'-noncoding region, a putative 'TATA box' sequence is boxed and in-frame stop codons are underlined. The single nucleotide difference between the sequence of the AOx gene and cDNA (T vs. C, respectively) is bracketed above nucleotide 807. Positions at which a nucleotide deletion/insertion exists relative to the POX4 sequence, creating the AOx sequence frameshift (section 3.2.3), are denoted by *. The stop codon of the ORF is signified by 'End'. The deduced amino acid sequence of AOx is presented in the standard three-letter code. Those NH_2-terminal amino acids of PXP-4 which were confirmed by Edman degradation (Okazaki et al., 1986) are underlined. Tripeptides denoted by double-underlining are variants of the PTS identified by Gould et al. (1989).

```
                                                                  (T)
    ATC GAT AAC GGT TGG ATT CAA TTC TCC AAC GTC AGA ATC CCA AGA TTC TTC ATG TTG CAA   816
253 Ile Asp Asn Gly Trp Ile Gln Phe Ser Asn Val Arg Ile Pro Arg Phe Phe Met Leu Gln
    AAG TTC TGT AAG GTT TCT GCT GAA GGT GAA GTC ACC TTG CCA CCT TTG GAA CAA TTG TCT   876
273 Lys Phe Cys Lys Val Ser Ala Glu Gly Glu Val Thr Leu Pro Pro Leu Glu Gln Leu Ser
    TAC TCC GCC TTG TTG GGT GGT AGA GTC ATG ATG GTT TTG GAC TCC TAC AGA ATG TTG GCT   936
293 Tyr Ser Ala Leu Leu Gly Gly Arg Val Met Met Val Leu Asp Ser Tyr Arg Met Leu Ala
    AGA ATG TCT ACC ATT GCC TTG AGA TAC GCC ATT GGT AGA AGA CAA TTC AAG GGT GAT AAT   996
313 Arg Met Ser Thr Ile Ala Leu Arg Tyr Ala Ile Gly Arg Arg Gln Phe Lys Gly Asp Asn
    GTC GAT CCA AAC GAT CCA AAT GCT TTG GAA ACC CAA TTG ATA GAT TAC CCA TTG CAC CAA  1056
333 Val Asp Pro Asn Asp Pro Asn Ala Leu Glu Thr Gln Leu Ile Asp Tyr Pro Leu His Gln
                *
    AAG AGA TTG TTC CCA TAC TTC GTG CCG CCT ATG TCG TCT CCG CTG GTG CCC TCA AGG TTG  1116
353 Lys Arg Leu Phe Pro Tyr Phe Val Pro Pro Met Ser Ser Pro Leu Val Pro Ser Arg Leu
    AAC ACA CCA TCG AGA CCA CCT TGG CTG AAT TGG ACG CTG CCG TTG AAA AGA ACG ACA CCA  1176
373 Asn Thr Pro Ser Arg Pro Pro Trp Leu Asn Trp Thr Leu Pro Leu Lys Arg Thr Thr Pro
          *
    AGG CTA ATC TTT AAG TCT ATT GAC GAC ATG AAG TCA TTG TTT GTT GAC TCT GGT TCC TTG  1236
393 Arg Leu Ile Phe Lys Ser Ile Asp Asp Met Lys Ser Leu Phe Val Asp Ser Gly Ser Leu
    AAG TCC ACT GCC ACT TGG TTG GGT GCT GAA GCC ATT GAC CAA TGT AGA CAA GCC TGT GGT  1296
413 Lys Ser Thr Ala Thr Trp Leu Gly Ala Glu Ala Ile Asp Gln Cys Arg Gln Ala Cys Gly
    GGT CAC GGT CAC TCT TCC TAC AAC GGT TT  GGT AAG GCC TAC AAC GAT TGG GTT GTC CAA  1356
433 Gly His Gly His Ser Ser Tyr Asn Gly Phe Gly Lys Ala Tyr Asn Asp Trp Val Val Gln
    TGT ACC TGG GAA GGT GAC AAC AAT GTC TTG GGC ATG AGT GTT GGT AAG CCA ATT GTC AAG  1416
453 Cys Thr Trp Glu Gly Asp Asn Asn Val Leu Gly Met Ser Val Gly Lys Pro Ile Val Lys
    CAA GTC ATC AGC ATT GAA GAT GCC GGC AAG ACC GTC AGA GGC TCC ACT GCT TTC TTG AAC  1476
473 Gln Val Ile Ser Ile Glu Asp Ala Gly Lys Thr Val Arg Gly Ser Thr Ala Phe Leu Asn
    CAA TTG AAG GAA TAC ACT GGT TCC AAC AGC TCC AAG GTT GTT TTG AAC ACT GTT GCT GAC  1536
493 Gln Leu Lys Glu Tyr Thr Gly Ser Asn Ser Ser Lys Val Val Leu Asn Thr Val Ala Asp
    TTA GAC GAC ATC AAG ACT GTC ATC AAG GCT ATT GAA GTT GCC ATC ATC AGA TTA TCC CAA  1596
513 Leu Asp Asp Ile Lys Thr Val Ile Lys Ala Ile Glu Val Ala Ile Ile Arg Leu Ser Gln
    GAA GCT GCT TCT ATT GTC AAG AAG GAA TCT TTC GAC TAT GTC GGT GCT GAA TTG GTT CAA  1656
533 Glu Ala Ala Ser Ile Val Lys Lys Glu Ser Phe Asp Tyr Val Gly Ala Glu Leu Val Gln
    CTC TCC AAG TTG AAG GCT CAC CAC TAC TTG TTG ACT GAA TAC ATC AGA AGA ATT GAC ACC  1716
553 Leu Ser Lys Leu Lys Ala His His Tyr Leu Leu Thr Glu Tyr Ile Arg Arg Ile Asp Thr
    TTT GAC CAA AAG GAG TTG GCT CCA TAC TTG ATA ACC CTC GGT AAG TTG TAC GCT GCC ACT  1776
573 Phe Asp Gln Lys Glu Leu Ala Pro Tyr Leu Ile Thr Leu Gly Lys Leu Tyr Ala Ala Thr
    ATT GTC TTG GAC AGA TTT GCT GGT GTC TTT TTG ACT TTC AAC GTT GCC TCC ACC GAA GCC  1836
593 Ile Val Leu Asp Arg Phe Ala Gly Val Phe Leu Thr Phe Asn Val Ala Ser Thr Glu Ala
    ATC ACT GCT TTG GCC TCT GTG CAA ATT CCA AAG TTG TGT GCT GAA GTC AGA CCA AAC GTT  1896
613 Ile Thr Ala Leu Ala Ser Val Gln Ile Pro Lys Leu Cys Ala Glu Val Arg Pro Asn Val
    GTT G                                                                         1900
633 Val
```

Figure 3.2.3. Nucleotide sequence of the cDNA insert of recombinant RU1:18 and the corresponding deduced amino acid sequence. The ≈1.5 kbp Cla I, Pst I fragment of cDNA recombinant RU1:18 (Rachubinski et al., 1985) was subcloned into Acc I, Pst I-digested M13mp18 and 19 and partially sequenced (1144 nucleotides; Figure 3.2.1, lower portion). Nucleotides (in bold) are numbered in the 5' to 3' direction from the Cla I site; the numbering system is identical to that used in Figure 3.2.2. The nucleotide difference identified between the AOx cDNA and gene (C vs. T, respectively) is indicated in brackets above nucleotide 807. Those positions where a nucleotide deletion/insertion exists relative to the POX-4 sequence, creating the AOx sequence frameshift (section 3.2.3), are denoted by *. The deduced amino acid of AOx (ile-253 to val-633) is presented in the three-letter code.

that deduced from the AOx gene.

Limited sequencing of the 5'-noncoding region of the AOx gene indicated the presence of a consensus 'TATA box' sequence (Breathnach and Chambon, 1981) at position -152 (Figure 3.2.2, boxed). Comparison of this 152 nucleotides of upstream sequence with that of POX4 (Okazaki et al., 1986) revealed perfect nucleic acid identity between these regions. The A + T content of the 5'-flanking region (70.4%) was significantly higher than that of the coding region (54.1%). Within the "leader" region of the AOx gene, taken to be upstream nucleotides -52 to -1 [which is the average size of yeast leader regions (Cigan and Donahue, 1987)], a nucleotide composition of A: 28.8%, T: 42.3%, G: 7.7%, C: 21.2% was not in compliance with that found for several other leader regions of yeast genes, which tend to be rich in adenosine nucleotides (Cigan and Donahue, 1987).

The proposed ATG start codon of the ORF (Figure 3.2.2) was the only possible choice from the obtained sequence. No other ATG codons (in- or out of frame) existed between the putative initiator codon (nucleotides 1 to 3) and the upstream, in-frame TAA and TGA termination codons (nucleotides -27 to -25 and -45 to -43, respectively, Figure 3.2.2, underlined). Also, the sequences adjacent to this start codon, 5'-CAUA<u>AUG</u>ACU-3', concurred with the generalized, strong consensus sequence for translation

initiation in yeast, 5'-$\begin{matrix}A\\Y\end{matrix}A\begin{matrix}A\\U\end{matrix}A\underline{AUG}UCU$-3' (where Y = pyrimidine; Cigan and Donahue, 1987), except for the presence of an adenosine nucleotide at the +4 position. On average, 29% of yeast genes exhibit an adenosine nucleotide at the +4 position whereas 39% of yeast genes have a uridine nucleotide at this same position (Cigan and Donahue, 1987). Consistent with most other yeast genes, those sequences adjacent to the start codon of the AOx gene did not reflect the higher eukaryotic consensus sequence for translation initiation, 5'-C$\begin{matrix}A\\G\end{matrix}CC\underline{AUG}$G-3' (Kozak, 1987), except for the cytidine and adenosine nucleotide preferences at the -4 and -3 positions, respectively.

The nucleotide frequency of the AOx gene, A + T: 55.2%, C + G: 44.8% (Table 3.2.1), was representative of several other peroxisomal genes from C. tropicalis including POX2 (A + T: 54.4%; Okazaki et al., 1987), POX4 (A + T: 53.9%; Okazaki et al., 1986), POX5 (A + T: 52.2%; Okazaki et al., 1986), catalase (A + T: 55.7%; Table 4.4.1; Okada et al., 1987), HDE (A + T: 54.8%; Nuttley et al., 1988), isocitrate lyase (A + T: 55.3%; Atomi et al., 1990) and POX18 (A + T: 56.7%; Szabo et al., 1989; Tan et al., 1990). Nevertheless, a significant difference existed between the observed A + T nucleotide frequency of the AOx gene (55.2%) and the estimated A + T composition of C. tropicalis genomic DNA (65.1%; Stenderup and Bak, 1968) derived from thermal

Table 3.2.1. Nucleotide frequency and deduced amino acid composition of the Candida tropicalis AOx gene and protein.

NUCLEOTIDE FREQUENCY	A - 626 (27.4)					
	C - 538 (23.5)					
	G - 486 (21.2)					
	T - 638 (27.9)					
	A + T - 1264 (55.2)					
	C + G - 1024 (44.8)					
DINUCLEOTIDE FREQUENCY	AA - 201 (8.8)			GA - 167 (7.3)		
	AC - 167 (7.3)			GC - 76 (3.3)		
	AG - 124 (5.4)			GG - 101 (4.4)		
	AT - 134 (5.9)			GT - 141 (6.2)		
	CA - 176 (7.7)			TA - 82 (3.6)		
	CC - 148 (6.5)			TC - 147 (6.4)		
	CG - 63 (2.8)			TG - 198 (8.7)		
	CT - 151 (6.6)			TT - 211 (9.2)		
AMINO ACID COMPOSITION	Ala	51	(7.2)	Leu	64	(9.0)
	Arg	38	(5.4)	Lys	42	(5.9)
	Asn	32	(4.5)	Met[a]	17	(2.4)
	Asp	45	(6.3)	Phe	29	(4.1)
	Cys	8	(1.1)	Pro	32	(4.5)
	Gln	31	(4.4)	Ser	54	(7.6)
	Glu	35	(4.9)	Thr	45	(6.3)
	Gly	48	(6.8)	Trp	10	(1.4)
	His	10	(1.4)	Tyr	26	(3.7)
	Ile	43	(6.1)	Val	49	(6.9)
	Acidic	(Asp+Glu)			80	(11.3)
	Basic	(Arg+Lys)			80	(11.3)
	Aromatic	(Phe+Trp+Tyr)			65	(9.2)
	Hydrophobic	(Aromatic+Ile +Leu+Met+Val)			238	(33.6)
	M_r[a] = 79155			Calculated pI = 6.11		

Nucleotide, dinucleotide and amino acid frequencies were calculated using the Beckman MicroGenie (Version 6.0) Sequence Analysis Program (Queen and Korn, 1984). Numbers in parentheses indicate the % of the total number of "residues" (nucleotides, dinucleotides or amino acids) in AOx represented by each "residue".

[a] includes the N-terminal (initiator) methionine residue

denaturation data. Possibly, this difference is the result of a strong tendency toward A + T richness within the noncoding regions of the C. tropicalis genome (as noted above for the 5'-flanking sequence of the AOx gene). Analysis of the dinucleotide frequencies of the AOx gene indicated a possible suppression of the dinucleotides, CpG (2.8%), GpC (3.3%) and TpA (3.6%) (Table 3.2.1). Those peroxisomal genes from C. tropicalis mentioned above may similarly exhibit a general suppression of CpG dinucleotides, since the range in frequency of occurrence is from 2.0 to 3.7%; however, there is probably no suppression of GpC and TpA dinucleotides since the ranges in frequency of occurrence are from 3.5 to 4.2% and 2.9 to 5.0%, respectively. Validation of CpG suppression, by calculation of the expected probability of occurrence of CpG dinucleotides within the C. tropicalis genome, relative to their observed frequency of occurrence, was not possible since the individual C or G nucleotide composition of the genome is not known. Should CpG suppression exist within the C. tropicalis genome, the degree is less severe than that observed for mammalian genomes. In mammalian genomes, the CpG dinucleotide is five-fold rarer than expected from G + C content (Max, 1984) which implies that the frequency of occurrence of CpG in a mammalian gene similar in size to the AOx gene would be less than 1.0%.

The sequence of the 3'-noncoding region of the AOx gene was not extended significantly beyond the TAA termination codon. Consequently, the positions of "secondary" in-frame termination codons, putative polyadenylation signals and transcription termination signals, all of which have been reported for C. tropicalis POX4 and 5 (Okazaki et al., 1986), were not determined.

3.2.2 Deduced Amino Acid Sequence of AOx

Based upon the sequence of the 2,127 nucleotide ORF of the AOx gene, a polypeptide of 709 amino acids (M_r 79155; Table 3.2.1) was predicted as the primary structure of AOx (Figure 3.2.2). Residues 2 to 18 of the deduced polypeptide were identical to the NH_2-terminal sequence of PXP-4, thr-phe-thr-lys-lys-asn-val-ser-val-ser-gln-gly-(asp or pro)-asp-pro-arg-ser, determined by Edman degradation (Okazaki et al., 1986). Consistent with other yeast proteins (Huang et al., 1987), the initiator methionines of PXP-2, -4 and -5 appeared to be removed during "maturation" of these polypeptides (Okazaki et al., 1986, 1987) and may be similarly absent from "mature" AOx. If so, the calculated M_r of "mature" AOx (ie. without the methionine) is 79006, which is in agreement with the M_r of 76000 estimated by hybridization-selection translation, followed by SDS-PAGE (Rachubinski et al., 1985) and quite similar to that deduced for "mature" PXP-4 (M_r = 78554, Okazaki et al., 1986).

The amino acid composition of AOx is presented in Table 3.2.1. Consistent with its peroxisomal matrix location (Yokota and Fahimi, 1982), AOx was found to be comprised of 33.6% apolar residues (238 aromatic + ile + leu + met + val amino acids of 708 total). This compared closely with that of other acyl-coenzyme A oxidase polypeptides [C. tropicalis PXP-2, 32.7% (Okazaki et al., 1987); C. tropicalis PXP-4, 33.6% (Okazaki et al., 1986); C. tropicalis PXP-5, 33.4% (Okazaki et al., 1986); C. maltosa AOX1, 33.3% (Hill et al., 1988); S. cerevisiae POX1 protein, 34.9% (Dmochowska et al., 1990); rat liver AOx TypeI, 34.6% (Miyazawa et al., 1987)], and that of other C. tropicalis matrix proteins [catalase, 29.7% (Table 4.4.1; Okada et al., 1987); HDE, 31.9% (Nuttley et al., 1988); PXP-18, 33.1% (Szabo et al., 1989; Tan et al., 1990); isocitrate lyase, 30.2% (Atomi et al., 1990)]. The ratio of charged to apolar residues in AOx (ratio = 0.67) suggested that this polypeptide adopted a generally globular form such that the apolar groups were easily accommodated within the centre of the globular unit (Cohen and Parry, 1986). This appears to be true of peroxisomal acyl-coenzyme A oxidases from several species, with appropriate ratios being 0.78 and 0.73 for PXP-2 and -5, respectively, of C. tropicalis (Okazaki et al., 1986, 1987), 0.65 for AOX1 of C. maltosa (Hill et al., 1988), 0.67 for POX1 protein of S. cerevisiae (Dmochowska et

al., 1990) and 0.61 for AOx TypeI of rat liver (Miyazawa et al., 1987). However, a protein characterized by a ratio of charged to apolar residues of 0.75 or greater (eg. PXP-2), probably adopts a slightly more extended conformation than that of a more compact, globular protein of lower ratio (Cohen and Parry, 1986). Overall, the number of acidic (asp + glu) and basic (arg + lys) residues in AOx was identical (80 each, Table 3.2.1) with the pI calculated to be 6.11. This compares with the experimental values of 5.5 (Shimizu et al., 1979) and 5.2 (Coudron et al., 1983) determined for PXP-5 (acyl-coenzyme A oxidase I).

Prediction of the secondary structure of AOx (Figure 3.2.4, Alph and Beta graphs), using the algorithm of Garnier et al. (1978), indicated a slight dominance of α-helical content (≈32%), with the most prominent region of α-helix encompassing 56 amino acids (residues 510 to 565; Figure 3.2.4, Alph graph, bar). The remainder of the polypeptide was predicted to be comprised of ≈27% β-sheet structure and equal proportions (≈20%) of both turn and random coil structures. Two different hydropathy profiles for AOx, one calculated using the algorithm of Hopp and Woods (1981) and the other using the algorithm of Kyte and Doolittle (1982), are also illustrated in Figure 3.2.4 (Hyd and Val graphs, respectively). The algorithm of Hopp and Woods (1981) was established to allow identification of protein

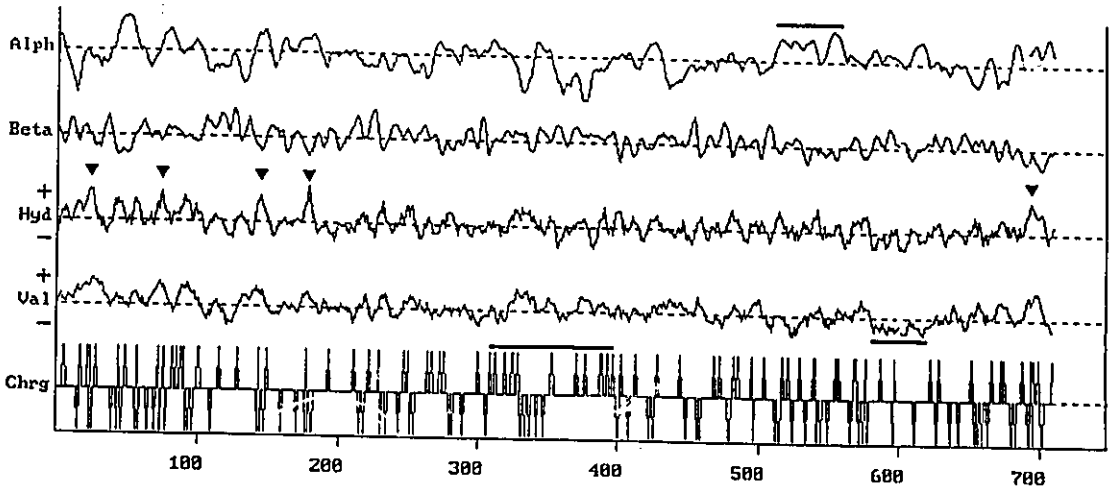


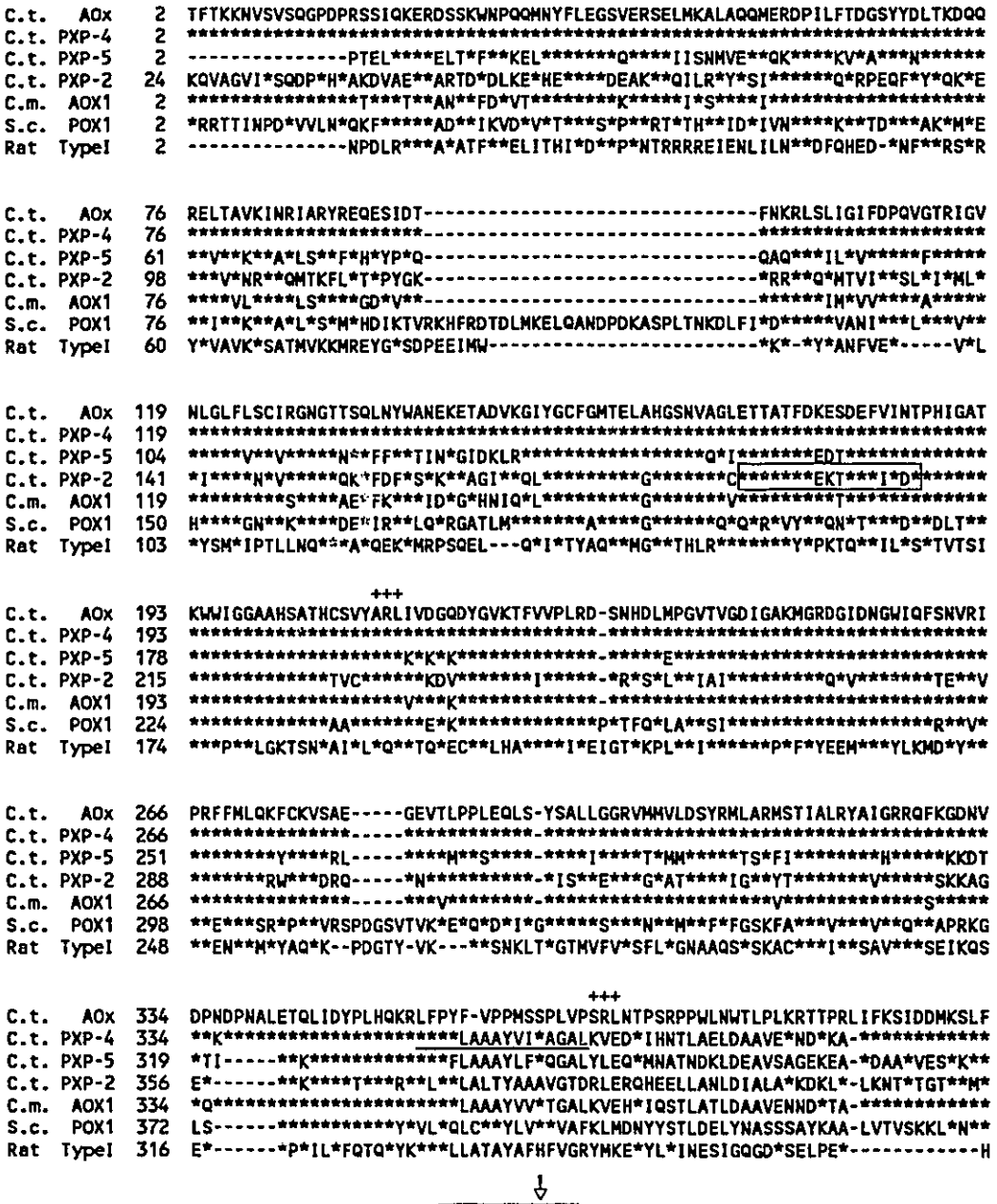
Figure 3.2.4. Predicted secondary structure, hydropathy and charge profile of AOx. Graphs were generated using standard parameters for the Beckman Microgenie (Version 6.0) Sequence Analysis Program (Queen and Korn, 1984). Secondary structure (Alph = α-helix, Beta = β-sheet) was determined using the algorithm of Garnier et al. (1978); the bar above the graph of α-helical content highlights the most prominent region of α-helix within AOx (see text, section 3.2.2). Hydropathy was determined using the algorithms of Hopp and Woods (1981) (Hyd) and Kyte and Doolittle (1982) (Val). For both algorithms, hydrophilicity is indicated by + (above the base line) while hydrophobicity is indicated by - (below the base line); closed triangles above the graph plotted according to Hopp and Woods (1981) indicate possible antigenic determinants within AOx; the bar below the graph plotted according to Kyte and Doolittle (1982) represents the only significant region of hydrophobicity within AOx (see text, section 3.2.2). The charge profile (Chrg) was plotted by assigning values of +1 to R or K residues, -1 to D or E residues and 0 to all other residues; the bar above this graph delineates the "cluster" of basic, followed by acidic, followed by basic, amino acids within AOx discussed in the text (section 3.2.2).

antigenic determinants by analyzing amino acid sequences in order to determine the areas of greatest local hydrophilicity. Under this premise, it may be speculated that AOx possesses at least five major antigenic determinants [residues 20 to 25, 70 to 75, 140 to 145, 175 to 180 and 690 to 695 (Figure 3.2.4, Hyd graph, closed triangles)], four of which are located toward the N-terminus. Alternatively, the algorithm of Kyte and Doolittle (1982) is particularly well suited to identifying those portions of membrane proteins located within the lipid bilayer as relatively large, uninterrupted regions (usually greater than 10 amino acids) of hydrophobicity. As expected for a peroxisomal matrix protein, the corresponding hydropathy profile of AOx did not reveal any potential membrane spanning regions (Figure 3.2.4, Val graph). The only significant region of uninterrupted hydrophobicity (≈30 amino acids, residues 578 to 607) was only minimally hydrophobic according to the algorithm of Kyte and Doolittle (1982) (Figure 3.2.4, Val graph, bar), and even less so according to the algorithm of Hopp and Woods (1981) (Figure 3.2.4, Hyd graph). The algorithm of Kyte and Doolittle (1982) predicted a more hydrophilic nature for the AOx polypeptide than that of Hopp and Woods (1981); the former may be more accurate since the algorithm of Parker et al. (1986) further purported a hydrophilic nature for the AOx

polypeptide (not shown). The charge profile of AOx was, in general, unremarkable (Figure 3.2.4, Chrg graph); however, one internal region of 98 amino acids (residues 309 to 397; Figure 3.2.4, Chrg graph, bar) was found to possess three contiguous stretches of 21, 17 and 27 amino acids (residues 309 to 329, 331 to 347 and 371 to 397, respectively) containing 6 basic, 5 acidic and 6 basic amino acids, respectively. The purpose of this "cluster" of oscillating charge, if any, is unknown.

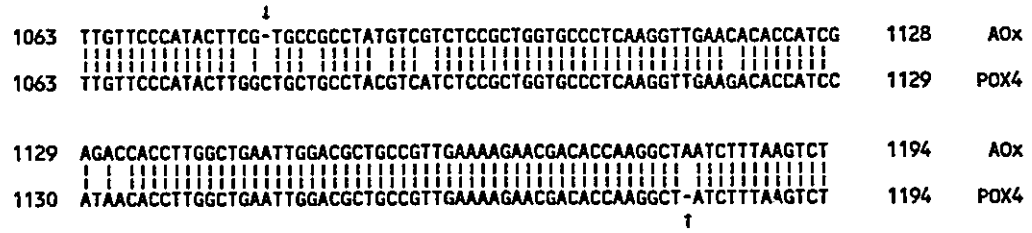
3.2.3 Comparison of the Deduced Amino Acid Sequence of AOx with those of Other Acyl-Coenzyme A Oxidases

Comparison of the deduced amino acid sequence of AOx with C. tropicalis PXP-4 (Okazaki et al., 1986) revealed an overall amino acid identity of 94.0% (Figure 3.2.5). Nevertheless, between residues 360 and 394, these proteins diverged significantly with only 4 of 35 amino acids (11.4%) exhibiting identity [Figures 3.2.5 and 3.2.6 (lower portion)]. Inspection of that segment of the AOx gene encoding this amino acid stretch (nucleotides 1078 to 1182, Figure 3.2.6, upper portion), which was originally sequenced along both DNA strands of both the AOx gene and cDNA (section 3.2.1) and found to exhibit 92.4% nucleotide identity with POX4 (Okazaki et al., 1986), revealed the presence of a nucleotide sequence frameshift, relative to the POX4 sequence, which resulted from the deletion of a C

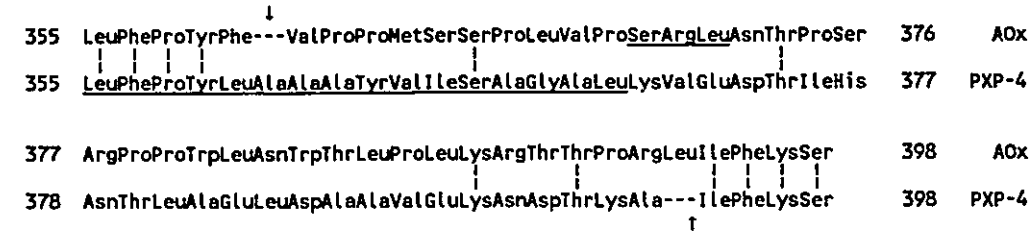

C.t. AOx 2 TFTKKNVSVSQGPDPRSSIQKERDSSKWNPQQMNYFLEGSVERSELMKALAQQMERDPILFTDGSYYDLTKDQQ
C.t. PXP-4 2
C.t. PXP-5 2
C.t. PXP-2 24
C.m. AOX1 2
S.c. POX1 2
Rat TypeI 2
C.t. AOx 76 RELTAVKINRIARYREQESIDT
C.t. PXP-4 76
C.t. PXP-5 61
C.t. PXP-2 98
C.m. AOX1 76
S.c. POX1 76
Rat TypeI 60
C.t. AOx 119 NLGLFLSCIRGNGTTSQLNYWANEKETADVKGIYGCFGMTELAHGSNVAGLETTATFDKESDEFVINTPHIGAT
C.t. PXP-4 119
C.t. PXP-5 104
C.t. PXP-2 141
C.m. AOX1 119
S.c. POX1 150
Rat TypeI 103
+++
C.t. AOx 193 KWWIGGAAHSATHCSVYARLIVDGQDYGVKTFVVPLRD-SNHDLMPGVTVGDIGAKMGRDGIDNGWIQFSNVRI
C.t. PXP-4 193
C.t. PXP-5 178
C.t. PXP-2 215
C.m. AOX1 193
S.c. POX1 224
Rat TypeI 174
C.t. AOx 266 PRFFMLQKFCKVSAE-----GEVTLPPLEQLS-YSALLGGRVMHVLDSYRMLARMSTIALRYAIGRRQFKGDNV
C.t. PXP-4 266
C.t. PXP-5 251
C.t. PXP-2 288
C.m. AOX1 266
S.c. POX1 298
Rat TypeI 248
+++
C.t. AOx 334 DPNDPNALETQLIDYPLHQKRLFPYF-VPPMSSPLVPSRLNTPSRPPWLNWTLPLKRTTPRLIFKSIDDMKSLF
C.t. PXP-4 334
C.t. PXP-5 319
C.t. PXP-2 356
C.m. AOX1 334
S.c. POX1 372
Rat TypeI 316
C.t. AOx 407 VDSGSLKSTATWLGAEAIDQCRQACGGHGHSSYNGFGKAYNDWVVQCTWEGDNNVLGMSVGKPIVKQVI-SIED
C.t. PXP-4 407
C.t. PXP-5 387
C.t. PXP-2 423
C.m. AOX1 407
S.c. POX1 439
Rat TypeI 372

```
C.t.   AOx   480  AGKTVRGSTA-FLNQLK-EYTGSNSSKVVLNTVADL-   ---------KTVIKAIEVAIIRLSQEAASIVKKE
C.t. PXP-4   480  **********-******-D*********************----------*********************
C.t. PXP-5   460  ----------------------QKGL**SS******P----------AKLV**FDH*LSG*ARDIGAVAEDK
C.t. PXP-2   496  -**ELKD**LE***-AAP*--L*KAK*A*IRIRDHV**V----------DR*L***-AGL*SKFSKDLIPISYQ
C.m.  AOX1   480  N*********-****V*-DF****A******NTS**N*----------NK***S********AH***IS*R**
S.c.  POX1   513  GREDNTLDVDS*SY-**PQ*I**VV*GEIKSGLKE*G*----------Y*E*WS*-TL*KL*AHIGTLVE*SR
Rat  TypeI   445  -**L*-*GMVSY**D*PSQRIQPQQVA*WPTM*DINSLEGLTEAYKLRAARLVEIAAKNLQTHVSHR--KS*EV

                          +++
C.t.   AOx   542  SFDYVGAELVQLSKLKAHH-YLLTEYIR---RIDTFDQKELAPYLITLGKLYAATIVLDRFAGVFLTFNVASTE
C.t. PXP-4   542  ********************-*********---********D*V**********************************
C.t. PXP-5   500  G**IT*PS**LV***N**R-F*IDGFFK---**TPEWSEV*R*----**F***DW*-*TN*GAT**QYGIITPD
C.t. PXP-2   555  *W*SI**QR*I****RC*Y-***ETFNE-RLNDKIKAKSPAR*H*ENII***YV*NI*GP*IDE**R*G*I*PQ
C.m.  AOX1   542  *L*FA******I*******-*****FVK---*VGE*EH***V*F*N*I*R**S**V***K**********PQ
S.c.  POX1   574  *I*S*SKL**LV**FH*LRCM*K*Y*DKLNS*DSHISDEITKESMWNVY**FSLYF-I*KHS*E*QQ*KIFTPD
Rat  TypeI   515  AWNLTSVD**RA*EAHC*YVVVKVFSDKL---PKIQ*KAVQ*V-*RN*CL**SLYG-ISQKG*D**EGSIITGA

C.t.   AOx   612  AITALASVQIPKLCAEVRPNVVAYTDSFQQSDNIVNSAIGRYDGDIYENYFDLVKLQNPPSKTKA--PYSDALE
C.t. PXP-4   612  ************************************************************--*******
C.t. PXP-5   565  VSRKIS*EHF*A***K******GL**G*NLT**MT*A*******NV**H**ET**AL***EN***--***K***
C.t. PXP-2   627  VAKYITYEYPQ****NI**Y*IGL******P*NFI**L**K***N**T**LES**DV*D**NY**--***E***
C.m.  AOX1   612  ***D***T**********************VI******K*****V*********QL***KN***--**TA***
S.c.  POX1   647  Q*SKVVQP*LLA*LPI**KDCIGL****ELP*AML**P**YF******H***NE*CRN**VEADG*GK*SYH**L
Rat  TypeI   584  QLSQVNAR-*LE*LTLI***A**LV*A*DFK**TLG*VL*****NV***L*EWA*K-S*LN**EV-HESYHKHL

C.t.   AOx   684  A-MLNPPTLDERERFQKSDETAAILSK          709
C.t. PXP-4   684  *-**************E***********          709
C.t. PXP-5   637  D-*****D*EV***GE**E*A*E***S          662
C.t. PXP-2   699  *-*****SA*EN***SERGKAA*D****          724
C.m.  AOX1   684  G-I****S*EA***YE***********          709
S.c.  POX1   721  SS**G*--GF*FDQKLGGAAN*E****INK       748
Rat  TypeI   655  KPLQSKL                             661
```

Figure 3.2.5. Comparison of the deduced amino acid sequence of AOx with those of other acyl-coenzyme A oxidases. C.t. AOx, C. tropicalis AOx (from Figure 3.2.2); C.t. PXP-4, C. tropicalis PXP-4 (Okazaki et al., 1986); C.t. PXP-5, C. tropicalis PXP-5 (Okazaki et al., 1986); C.t. PXP-2, C. tropicalis PXP-2 (Okazaki et al., 1987); C.m. AOX1, C. maltosa AOX1 (Hill et al., 1988); S.c. POX1, S. cerevisiae POX1 (Dmochowska et al., 1990); Rat TypeI (Miyazawa et al., 1987). Homologies between the polypeptides (excluding AOx) were identified by Dmochowska et al. (1990) using the FASTP program (Lipman and Pearson, 1985) and reproduced here. AOx was subsequently aligned based upon its homology with C. tropicalis PXP-4 (Okazaki et al., 1986). Amino acids are represented by the standard one-letter code. Gaps (-), inserted at appropriate positions to maximize homology between sequences, are counted as mismatches. Asterisks (*) indicate amino acid identity between AOx and the other acyl-coenzyme A oxidases. The bar above amino acids 430 to 442 of AOx highlights a possible FMN-binding sequence common to each acyl-coenzyme A oxidase (see text, section 3.2.3); the arrow denotes the tyrosine moiety (residue 439 of AOx) within this region which is essential for stabilizing the FMN prosthetic group. The 17 hydrophilic amino acids of PXP-2 which are boxed (residues 192 to 208) represent an internal PTS postulated by Okazaki et al. (1987) based upon their high homology with other acyl-coenzyme A oxidases and ability to form a stable α-helical conformation. The underlined segment of 16 hydrophobic amino acids of PXP-4 (residues 355 to 370) represents the internal signal for the peroxisomal location of PXP-4 postulated by Okazaki et al. (1986) and Kamiryo et al. (1989). Tripeptide variants of the PTS identified by Gould et al. (1989) are denoted by +++ (see text, section 3.3). The N-terminal 23 amino acids of C. tropicalis PXP-2 are not included in the alignment.



Nucleotide sequence identity between residues 1078 and 1182 (allowing two gaps) = 92.4%



Amino acid sequence identity between residues 360 and 394 (allowing two gaps) = 11.4%

Figure 3.2.6. Nucleotide and deduced amino acid sequence alignment of the C. tropicalis AOx and POX4 genes and encoded proteins. The upper portion of the figure illustrates the nucleotide sequence frameshift of the AOx gene, relative to the POX4 gene (Okazaki et al., 1986), which resulted from the deletion of a C at position 1079 of the POX4 sequence and a compensating insertion of an A at position 1182 of the AOx sequence (arrows). The lower portion of the figure illustrates the minimal amino acid identity between the AOx and PXP-4 proteins which resulted due to the nucleotide sequence frameshift (arrows). Nucleotide and amino acid identities (represented by vertical lines between residues) were maximized by the insertion of gaps (-) and calculated to be 92.4 and 11.4%, respectively. The numbering systems for the AOx and POX4 genes and encoded proteins are presented in Figure 3.2.2 and reported by Okazaki et al. (1986), respectively. The underlined tripeptide within the AOx amino acid sequence represents a variant of the PTS identified by Gould et al. (1989) which was introduced as a result of the nucleotide sequence frameshift. The underlined segment of 16 hydrophobic amino acids represents the internal signal for the peroxisomal location of PXP-4 postulated by Okazaki et al. (1986) and Kamiryo et al. (1989) (see text, section 3.3).

at position 1079 of the POX4 sequence and a compensating insertion of an A at position 1182 of the AOx sequence (Figure 3.2.6, upper portion, arrows). This, in turn, generated an extremely proline-rich stretch (25.7% proline versus 3.4% for the remainder of AOx) of 35 amino acids which exhibited no greater than 11.4% amino acid identity with any of the other acyl-coenzyme A oxidases from C. tropicalis, C. maltosa, S. cerevisiae or rat (Figure 3.2.5). The corresponding nucleotide stretch was similarly determined to be abundant (≈40%) in anomalous codons, relative to the remainder of the AOx coding region, the significance of which will be discussed in more detail elsewhere (section 3.2.4). Other differences discovered between the amino acid sequences of AOx and PXP-4 (Figure 3.2.5) included: 1) conservative substitutions - residue 496 (glu versus asp, respectively); residue 577 (glu versus asp, respectively) and 2) non-conservative substitutions- residue 336 (asn versus lys, respectively); residue 359 (phe versus leu, respectively); residue 436 (his versus tyr, respectively); residue 463 (gly versus ala, respectively); residue 579 (ala versus val, respectively); residue 698 (gln versus glu, respectively). These observed differences in amino acid sequence between AOx and PXP-4 (Okazaki et al., 1986) may be the result of allelic heterogeneity within C. tropicalis pK233. This would mimic that previously reported

for the gene encoding C. maltosa AOX1 (Hill et al., 1988). Further, it would be in accordance with that reported by Okazaki et al. (1986), who isolated a recombinant, very similar in sequence, but not identical, to that encoding C. tropicalis PXP4, which was postulated to be the genomic counterpart of the cDNA encoding AOx identified by Rachubinski et al. (1985). In lieu of the fact that the AOx sequence was confirmed by cDNA sequencing, the AOx gene is being transcribed within the yeast cell. Consequently, the functional and/or structural significance, if any, of the aforementioned, "variant" amino acid stretch inherent to AOx, remains to be resolved.

Comparison of AOx with each of the other acyl-coenzyme A oxidases revealed amino acid identities of 51.4% with C. tropicalis PXP-2 (Okazaki et al., 1987), 57.7% with C. tropicalis PXP-5 (Okazaki et al., 1986), 77.9% with C. maltosa AOX1 (Hill et al., 1988), 43.6% with S. cerevisiae POX1 (Dmochowska et al., 1990) and 28.4% with rat AOx TypeI (Miyazawa et al., 1987). Moreover, the hydropathy profiles of each of these acyl-coenzyme A oxidases, constructed according to the algorithm of Kyte and Doolittle (1982), were quite similar (not shown), which reflects a consensus higher order structure amongst these proteins. Thus, AOx displayed significant overall homology with each of the other acyl-coenzyme A oxidases; this may imply that the AOx

genes of C. tropicalis, C. maltosa, S. cerevisiae and rat were derived from a common ancestor. Nevertheless, the degree of homology found with each of the other acyl-coenzyme A oxidases was not constant throughout the entire amino acid sequence. Generally, N-terminal portions of these proteins exhibited greater amino acid identities than C-terminal portions. A large segment of the N-terminal "half" of each acyl-coenzyme A oxidase (residues 152 to 357 of AOx, Figure 3.2.5) exhibited greater than 81% amino acid identity upon comparison of a minimum of four of the seven acyl-coenzyme A oxidase sequences to a consensus sequence generated by multiple sequence alignment within this region. This segment could be further subdivided into six highly homologous, smaller regions (residues 152 to 176, 180 to 212, 218 to 230, 242 to 272, 316 to 328 and 342 to 357 of AOx, Figure 3.2.5), which exhibited greater than 88, 85, 92, 87, 93 and 81% amino acid identity, respectively, upon comparison of a minimum of five of the seven acyl-coenzyme A oxidases to the derived consensus sequence for each region. By comparison, only one segment of the C-terminal "half" of each acyl-coenzyme A oxidase (residues 428 to 462 of AOx, Figure 3.2.5), was found to exhibit greater than 88% amino acid identity upon comparison of a minimum of six of the seven acyl-coenzyme A oxidases to the consensus sequence generated for this region. Undoubtedly, these regions

impart important common functions to the acyl-coenzyme A oxidases of each organism (possibly, substrate- or cofactor-binding sites, signal(s) for intracellular transport and/or structural requirements). In this respect, Miyazawa et al. (1987) searched, by sequence comparison, for a characteristic flavin-binding site (Shimizu et al., 1979; Coudron et al., 1983) within the amino acid sequences PXP-4, -5 and rat AOx TypeI. They discovered significant homology between these acyl-coenzyme A oxidases and the proposed FMN-binding sequences of Desulfovibro vulgaris flavodoxin (Dubourdieu and Fox, 1977) and NADPH-cytochrome P-450 oxidoreductase (Porter and Kasper, 1985) within a region [residues 395 to 407 of rat AOx TypeI, residues 430 to 442 of PXP-4 (and AOx), Figure 3.2.5] which overlaps the highly homologous, C-terminal segment (residues 428 to 462 of AOx) found amongst all seven acyl-coenzyme A oxidases (see above). Interestingly, within this region (amino acids 430 to 442 of AOx), a tyrosine moiety, shown by x-ray crystallographic studies to be essential for stabilizing the FMN prosthetic group of D. vulgaris flavodoxin through stacking interactions between the tyrosine aromatic side chain and the flavin isoalloxazine ring (Watenpaugh et al., 1973), is conserved in each acyl-coenzyme A oxidase except for rat AOx TypeI (residue 439 of AOx, Figure 3.2.5, arrow). A corresponding tyrosine moiety (tyr-157) is similarly

conserved within the FMN binding domain C. tropicalis NADPH-cytochrome P-450 oxidoreductase, recently reported by Sutter et al. (1990).

3.2.4 Codon Usage in the Gene Encoding AOx

At present, only a limited discussion regarding codon usage in the gene encoding AOx will be introduced. A more elaborate disquisition, including generalities of codon usage within several C. tropicalis genes, similarities/differences in codon usage between these genes and those of other unicellular organisms and possible correlations between codon bias and the extent of gene expression, will follow subsequent to the analysis of the gene encoding catalase from C. tropicalis (see section 4.4.4).

In E. coli and S. cerevisiae, the expression of a gene is correlated with the degree of nonrandom codon usage (Gouy and Gautier, 1982; Sharp et al., 1986; Andersson and Kurland, 1990). Moreover, genes which are expressed to a high degree are strongly biased toward those codons which correspond to the major tRNA isoacceptor species within the cell (Bennetzen and Hall, 1982b; Ikemura, 1985; Holm, 1986; Bulmer, 1987). Representative of such genes are: 1) those encoding outer membrane lipoprotein (lpp, Nakamura et al., 1980), elongation factor TU-A and -B [tufA (Yokota et al., 1980) and tufB (An and Friesen, 1980)] and outer membrane

protein A (ompA, Movva et al., 1980) of E. coli and 2) those encoding the glycolytic and gluconeogenic enzymes of S. cerevisiae including hexokinase PI (HXK1, Kopetzki et al., 1985) and PII (HXK2, Fröhlich et al., 1985), glyceraldehyde-3-phosphate dehydrogenase I (G3PDH-I, Holland and Holland, 1979) and II (G3PDH-II, Holland and Holland, 1980), alcohol dehydrogenase I (ADH-I, Bennetzen and Hall, 1982a) and II (ADH-II, Russell et al., 1983), enolase I (E-I) and II (E-II) (Holland et al., 1981), pyruvate kinase (PK, Burke et al., 1983) and 3-phosphoglycerate kinase (PGK, Hitzeman et al., 1982). Relative to these genes, codon usage in the gene encoding AOx was similarly nonrandom (Table 3.2.2). Several codons (10 of 61 possible), including GCA and GCG for ala, CGT, CGC, CGA and CGG for arg, TGC for cys, GGA and GGG for gly and CAT for his, were never used. Other codons (18 of 61 possible), including AGG for arg, CAG for gln, ATA for ile, CTT, CTC, CTA, CTG and TTA for leu, AAA for lys, CCC for pro, TCA, TCG, AGT and AGC for ser, ACG for thr, TAT for tyr and GTA and GTG for val, were used to encode less than 10% of the total number of each corresponding amino acid in AOx. Consequently, only 23 of a possible 61 codons (note that met and trp codons are not considered to be preferential since they do not exhibit degeneracy) were preferentially used to encode 82.7% of the amino acids in AOx (Table 3.2.2, underlined codons). This percentage is

Table 3.2.2. Codon usage in the gene encoding AOx

Amino Acid	Codon[a]	Number of codons used[b] (%)	Amino Acid	Codon[a]	Number of codons used[b] (%)
Ala	GCT	28 (3.9)	Lys	AAA	4 (0.6)
	GCC	23 (3.2)		AAG	38 (5.4)
	GCA	0 (0.0)	Met	ATG	17 (2.4)
	GCG	0 (0.0)	Phe	TTT	13 (1.8)
Arg	CGT	0 (0.0)		TTC	16 (2.3)
	CGC	0 (0.0)	Pro	CCT	7 (1.0)
	CGA	0 (0.0)		CCC	1 (0.1)
	CGG	0 (0.0)		CCA	21 (3.0)
	AGA	36 (5.1)		CCG	3 (0.4)
	AGG	2 (0.3)	Ser	TCT	16 (2.3)
Asn	AAT	5 (0.7)		TCC	26 (3.7)
	AAC	27 (3.8)		TCA	4 (0.6)
Asp	GAT	12 (1.7)		TCG	2 (0.3)
	GAC	33 (4.7)		AGT	3 (0.4)
Cys	TGT	8 (1.1)		AGC	3 (0.4)
	TGC	0 (0.0)	Thr	ACT	15 (2.1)
Gln	CAA	29 (4.1)		ACC	23 (3.2)
	CAG	2 (0.3)		ACA	5 (0.7)
Glu	GAA	31 (4.4)		ACG	2 (0.3)
	GAG	4 (0.6)	Trp	TGG	10 (1.4)
Gly	GGT	43 (6.1)	Tyr	TAT	2 (0.3)
	GGC	5 (0.7)		TAC	24 (3.4)
	GGA	0 (0.0)	Val	GTT	19 (2.7)
	GGG	0 (0.0)		GTC	26 (3.7)
His	CAT	0 (0.0)		GTA	1 (0.1)
	CAC	10 (1.4)		GTG	3 (0.4)
Ile	ATT	20 (2.8)			
	ATC	21 (3.0)			
	ATA	2 (0.3)	Sum aa[c]		709.0
Leu	CTT	1 (0.1)	Mf aa[d]		586.0
	CTC	4 (0.6)	% Mf aa[e]		82.7
	CTA	1 (0.1)			
	CTG	3 (0.4)			
	TTA	2 (0.3)			
	TTG	53 (7.5)			

[a] The 23 preferentially used codons in the AOx gene are underlined. Met and trp codons, which show no degeneracy, were not considered to be preferentially used codons.

[b] Numbers in parentheses represent the percentage of all amino acids in AOx encoded by a particular codon.

[c] Sum aa, total number of amino acids in AOx (including the initiator methionine).

[d] Mf aa, total number of amino acids encoded by the 23 preferentially used codons.

[e] % Mf aa, percentage of Mf aa as related to Sum aa.

greater than the codon selectivities of the S. cerevisiae genes encoding HXK1 (71.7%), HXK2 (79.0%) and ADH-II (79.6%) but less than the percentages for E. coli genes encoding ompA (84.4%), tufB (85.5%), lpp (85.9%) and tufA (87.3%) and S. cerevisiae genes encoding ADH-I (91.4%), PGK (92.0%), PK (92.2%), E-I (92.4%), E-II (93.3%), G3PDH-II (93.5%) and G3PDH-I (94.5%) (see Table 4.4.5). Thus, consistent with these E. coli and S. cerevisiae genes, the high degree of preferential codon usage observed in the AOx gene may be directly correlated with its level of expression. For several E. coli and S. cerevisiae genes, the degree of bias in sense codon usage has also been correlated with selective differences among translation termination codons (Sharp and Bulmer, 1988). In these species, there is a very strong bias in favour of TAA (over TGA and TAG) in genes where sense codon usage is highly biased (Sharp and Bulmer, 1988). This generality may similarly apply to the highly biased genes of C. tropicalis since the ORF of the AOx gene is also terminated by the codon, TAA (Figure 3.2.2).

Relative to the codon usage found throughout most of the AOx gene (nucleotides 1 to 1077 and 1183 to 2127), that nucleotide stretch (residues 1078 to 1182, Figure 3.2.6) constituting the AOx sequence frameshift (section 3.2.3) was determined to be abundant (≈40%) in anomalous codons (ie. codons which do not encode even a single amino acid

elsewhere in AOx) (Table 3.2.3, asterisks). Within this region, 14 of 35 amino acids were encoded by anomolous codons, including AGG (twice) for arg, CTA (once) and CTG (thrice) for leu, CCC (once) and CCG (thrice) for pro, TCG (twice) for ser and ACG (twice) for thr. Two other codons, ACA (twice) for thr and GTG (twice) for val, occurred at significantly higher frequencies within the translated region of the AOx sequence frameshift than throughout the remainder of the AOx protein. When compared to other C. tropicalis genes encoding peroxisomal proteins and other known C. tropicalis genes, the collective frequency of occurrence of the anomalous codons in these genes, relative to that within the region of the AOx sequence frameshift, was extremely low (Table 3.2.3): AGG, 3 occurrences in 7,073 codons versus 2 occurrences in 35 codons, respectively; CTA, 0 occurrences in 7,073 codons versus 1 occurrence in 35 codons, respectively; CTG, 2 occurrences in 7,073 codons versus 3 occurrences in 35 codons, respectively; CCC, 2 occurrences in 7,073 codons versus 1 occurrence in 35 codons, respectively; CCG, 4 occurrences in 7,073 codons versus 3 occurrences in 35 codons, respectively; TCG, 14 occurrences in 7,073 codons versus 2 occurrences in 35 codons, respectively; ACG, 3 occurrences in 7,073 codons versus 2 occurrences in 35 codons, respectively. As yet, the significance of this "variant"

Table 3.2.3. Codon usage between nucleotides 1078 and 1182 of the AOx gene: Comparison with the remainder of the AOx gene and with several other genes of Candida tropicalis.

Amino Acid	Codon[a]	Gene(s)			
		AOx (nucleotides 1078 to 1182 inclusive)	AOx (nucleotides 1 to 1077 and 1183 to 2127)	Other genes encoding peroxisomal proteins[b]	Other C. tropicalis genes[c]
		Number of codons used[d]			
Arg	AGA	2 (5.7)	34 (5.0)	219 (4.1)	76 (4.3)
	AGG*	2 (5.7)	0 (0.0)	2 (0.04)	1 (0.1)
Asn	AAT	1 (2.9)	4 (0.6)	25 (0.5)	43 (2.5)
	AAC	1 (2.9)	26 (3.9)	232 (4.4)	34 (2.1)
Leu	CTA*	1 (2.9)	0 (0.0)	0 (0.0)	0 (0.0)
	CTG*	3 (8.6)	0 (0.0)	2 (0.04)	0 (0.0)
	TTG	2 (5.7)	51 (7.6)	38[illegible] (7.2)	97 (5.5)
Lys	AAA	1 (2.9)	3 (0.4)	54 (1.0)	70 (4.0)
Pro	CCT	2 (5.7)	5 (0.7)	35 (0.7)	9 (0.5)
	CCC*	1 (2.9)	0 (0.0)	1 (0.02)	1 (0.1)
	CCA	3 (8.6)	18 (2.7)	187 (3.5)	64 (3.7)
	CCG*	3 (8.6)	0 (0.0)	3 (0.06)	1 (0.1)
Ser	TCT	1 (2.9)	15 (2.2)	84 (1.6)	42 (2.4)
	TCA	1 (2.9)	3 (0.4)	15 (0.3)	15 (0.9)
	TCG*	2 (5.7)	0 (0.0)	8 (0.15)	6 (0.6)
Thr	ACA	2 (5.7)	3 (0.4)	12 (0.2)	9 (0.5)
	ACG*	2 (5.7)	0 (0.0)	1 (0.02)	2 (0.1)
Val	GTG	2 (5.7)	1 (0.1)	12 (0.2)	4 (0.2)

a Of the 35 codons used between nucleotides 1078 and 1182 of the AOx gene, only 32 are listed [the single met and two trp codons, which show no degeneracy, are excluded]. The preferentially used codons in the entire AOx gene are underlined (see Table 3.2.2). Anomalous codons (see text, section 3.2.4) are denoted by *.

b Includes peroxisomal genes of C. tropicalis encoding AOx (nucleotides 1 to 1077 and 1183 to 2127), catalase (cDNA and gene, section 4.4.4), POX2 (Okazaki et al., 1987), POX4 (Okazaki et al., 1986), POX5 (Okazaki et al., 1986), POX18 (Szabo et al., 1989; Tan et al., 1990), HDE (Nuttley et al., 1988) and isocitrate lyase (Atomi et al., 1990).

c Includes genes of C. tropicalis encoding cytochrome P450 (Sanglard and Loper, 1989), cytochrome P450 lanosterol 14α-demetnylase (Chen et al., 1988) and NADPH-cytochrome P-450 oxidoreductase (Sutter et al., 1990).

d Numbers in parentheses represent the percentage of all amino acids in the gene(s), or portions thereof, encoded by a particular codon.

stretch within AOx, with respect to possible effects on its structure, function and/or level of expression, remains to be resolved.

3.3 PEROXISOMAL TARGETING OF ACYL-COENZYME A OXIDASES

Recently, Gould et al. (1989) showed that a conservative variation of the tripeptide, ser/ala/cys-lys/his/arg - leu, is capable of acting as a peroxisomal targeting signal (PTS) when located at the extreme C-terminus of the peroxisomal protein, firefly luciferase. Furthering this work, Gould et al. (1990a) demonstrated the ability of the C-terminal tripeptide of luciferase, ser-lys-leu, to function as a PTS in several eukaryotic systems, including yeast, plants, insects and mammals, suggesting that protein translocation into peroxisomes has, at least in part, been conserved throughout eukaryotic evolution. However, the potential for any variant of the C-terminal tripeptide sequence, ser/ala/cys - lys/his/arg - leu, to play a universal role in peroxisomal targeting is questionable since this putative PTS is simultaneously: 1) a common C-terminal feature of several peroxisomal proteins (Gould et al., 1989; Table 1.4.1), 2) an uncommon C-terminal feature of other peroxisomal proteins but nevertheless is found internally, at various and/or numerous locations, within these proteins (Gould et al., 1989; Table 1.4.1), 3) not found at any location, C-terminal or internal, within

certain other peroxisomal proteins (Table 1.4.1), and 4) found at the C-terminus of, or internally within, cytoplasmic, mitochondrial and chloroplast proteins (Gould et al., 1988; Borst, 1989). Within this framework, studies concerned with the in vitro import of rat AOx TypeI into purified rat liver peroxisomes revealed that the PTS of this protein resides in the C-terminal five (or fewer) amino acids, consistent with the in vivo studies of Gould et al. (1988) which localized the PTS of rat AOx Type1 to its C-terminal 15 amino acids. Interestingly, the C-terminal three amino acids of rat AOx Type1 are ser-lys-leu (Miyazawa et al., 1989), similar to the luciferase PTS (Gould et al., 1989). By comparison, AOx from C. tropicalis does not possess a C-terminal variant of this tripeptide PTS (Figure 3.2.2); however, a "related" sequence motif, ile-leu-ser-lys, is highly conserved at the extreme C-terminus of C. tropicalis AOx, PXP-4, PXP-2 and C. maltosa AOX1 and is localized within the seven C-terminal amino acids of S. cerevisiae POX1 (Figure 3.2.5). From this, it may be speculated that the critical component of the acyl-coenzyme A oxidase PTS is the presence of C-terminal positive charge; this would not only parallel the presence of positively charged residues within the N-terminal presequences of certain secretory (von Heijne, 1990; Lazdunski and Benedetti, 1990), mitochondrial (Hartl and Neupert, 1990;

Skerjanc, 1990), chloroplast (Flügge, 1990; Smeekens et al., 1990; Lazdunski and Benedetti, 1990) and nuclear (Roberts, 1989; Silver and Goodson, 1989) proteins but may, more appropriately, parallel the postulated necessity for the presence of positive charge within the internal targeting sequences of proteins addressed to the peroxisome-like microbody, the glycosome (Opperdoes, 1987a,b, 1988). However, not without complication, it should be noted that C. tropicalis PXP-5 does not possess a positively charged amino acid at its extreme C-terminus (Figure 3.2.5). Interestingly, three variants of the tripeptide PTS of Gould et al. (1989) can be found at internal locations within AOx: ala-arg-leu (residues 210 to 212), ser-arg-leu (residues 370 to 372) and ser-lys-leu (residues 554 to 556) (Figure 3.2.2). Universal involvement of these tripeptides in the peroxisomal targeting of each of the acyl-coenzyme A oxidases is debatable, however, since the ala-arg-leu and ser-arg-leu tripeptides, while being highly conserved within the yeast proteins, are not present in rat AOx TypeI (Figure 3.2.5) and the ser-lys-leu tripeptide is unique to AOx from C. tropicalis due to the AOx sequence frameshift (section 3.2.3, Figures 3.2.5 and 3.2.6).

Initial attempts at the in vitro import of acyl-coenzyme A oxidase (PXP-4) into purified C. tropicalis peroxisomes tentatively identified two internal targeting

sequences of the polypeptide (amino acids 1 to 118 and 309 to 427), either of which could mediate import into peroxisomes (Small et al., 1988). Consistent with this, the in vivo expression and import of C. tropicalis PXP-4 into the peroxisomes of C. maltosa (Kamiryo et al., 1989) suggested the existence of three regions of peroxisomal targeting information, with the primary information found at two internal locations (amino acids 258 to 462 and 463 to 616) and secondary information located within the amino terminal one-third of the polypeptide. Within amino acids 258 to 462, a stretch of 16 hydrophobic residues (leu-355 to leu-370, Figure 3.2.5), punctuated by nonhomologous regions but found to be well-conserved between PXP-4, -5 and C. maltosa AOX1, was suggested to be a PTS (Okazaki et al., 1986; Kamiryo et al., 1989); however, this region is conserved to a lesser degree in PXP-2 and rat AOx TypeI (Figure 3.2.5) and is not conserved in AOx due to the location of the AOx sequence frameshift (residues 360 to 394, Figure 3.2.5 and 3.2.6). Interestingly, the second internal region (residues 463 to 616) is, in general, nonhomologous between the acyl-coenzyme A oxidases but contains the tripeptide, ser-lys-leu, which is, as already discussed, quite well conserved, at least among the yeast enzymes (Figure 3.2.5). As yet, the targeting component of the amino terminal segment has not been speculated upon, but

consistent with this notion, firefly luciferase similarly possesses an amino-terminal component, that when ablated, abolishes import (Gould et al., 1987).

Two alternative regions of the C. tropicalis acyl-coenzyme A oxidases have been deemed possible PTSs (Okazaki et al., 1987). Based upon the assumption that a targeting signal should be exposed on the protein surface to allow interaction with a putative receptor, Okazaki et al. (1987) suggested that the acyl-coenzyme A oxidase PTS would be localized to a hydrophilic region exhibiting reasonably high homology with other acyl-coenzyme A oxidases and having the potential to form a stable conformation (eg. α-helix). Two segments were subsequently found to match these criteria: one internal (residues 192 to 208 of PXP-2, Figure 3.2.5, box) and one at the extreme C-terminus (residues 694 to 724 of PXP-2). Although the C-terminus of rat AOx TypeI is dissimilar in sequence and size to the yeast counterparts (Figure 3.2.5), limited potential for targeting has been previously discussed (see above) with respect to a possible requirement for a C-terminal positively charged residue (or residues). Regarding the internal region (residues 192 to 208 of PXP-2), its potential as a peroxisomal addressing signal relies upon the accuracy of the, as yet, unsubstantiated assumption made by Okazaki et al. (1987) and the nature of the selective pressure which has maintained

the extremely high degree of conservation observed in each of the seven acyl-coenzyme A oxidases within this region (Figure 3.2.5).

4. CLONING AND SEQUENCING OF THE cDNA AND GENE ENCODING PEROXISOMAL CATALASE FROM THE YEAST CANDIDA TROPICALIS pK233

As discussed in the preceding chapter, Rachubinski et al. (1985), upon screening a C. tropicalis cDNA library by differential DNA dot blot hybridization using [^{32}P]-labeled cDNA complementary to mRNA isolated from glucose- and alkane-grown yeast, identified several alkane-inducible cDNA recombinants, one of which encoded acyl-coenzyme A oxidase. Further screening of this library revealed that ≈9% of the total recombinants encoded alkane-induced sequences. Eight of these hybridization positives were analyzed (Rachubinski et al., 1987) and one partial recombinant, designated RU1:87, was shown by hybridization-selection translation and immunoprecipitation to encode C. tropicalis catalase (Cat), the tetrameric heme-containing enzyme responsible for the dismutation of H_2O_2 (produced as a bi-product of peroxisomal β-oxidation) into oxygen and water. As an extension of this work, present efforts were directed toward cloning and sequencing the entire open reading frame and flanking regions of the gene encoding Cat. Having achieved this for both the catalase cDNA and gene, a detailed analysis of the cDNA/gene structure and the predicted primary sequence of the enzyme was conducted. Subsequently, similarities/differences in gene structure between AOx, Cat, other C. tropicalis genes and genes from

other unicellular organisms (E. coli, S. cerevisiae) were examined.

4.1 MOLECULAR CLONING OF THE cDNA ENCODING Cat

4.1.1 Selection of Recombinants Harbouring the 5'-End of the Cat cDNA

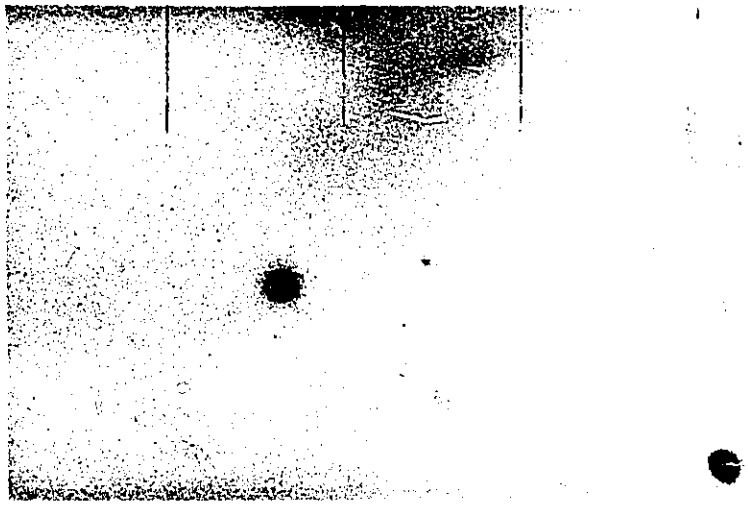
The strategy employed to select recombinants harbouring the 5'-end of the Cat cDNA entailed the construction and screening of a catalase specific, primer-directed, cDNA library of poly(A)$^+$RNA isolated from n-alkane-grown C. tropicalis (section 2.5.3). To this end, after having determined the nucleotide sequence of the partial cDNA insert of recombinant RU1:87 (also called CdT-1, section 4.2.2; illustrated in Figure 4.2.1), a 20-nucleotide synthetic primer, oligodeoxynucleotide C188, constructed complementary to poly(A)$^+$RNA encoding Cat, served as the primer for first strand cDNA synthesis (section 2.5.4) thereby enriching for 5' Cat cDNA sequence [see Table 2.1.1 for the sequence and annealing position of C188; Figures 4.2.1 and 4.2.2 (C188-box) illustrate the C188 annealing position (nucleotides 1262 to 1243) along the Cat mRNA]. Double-stranded cDNA was then synthesized from C188-primed, first strand cDNA according to the method of Gubler and Hoffman (1983), with modifications as described (section 2.5.4). Equimolar amounts of dC-tailed, double-stranded DNA and Pst I-digested, dG-tailed pBR322 were subsequently

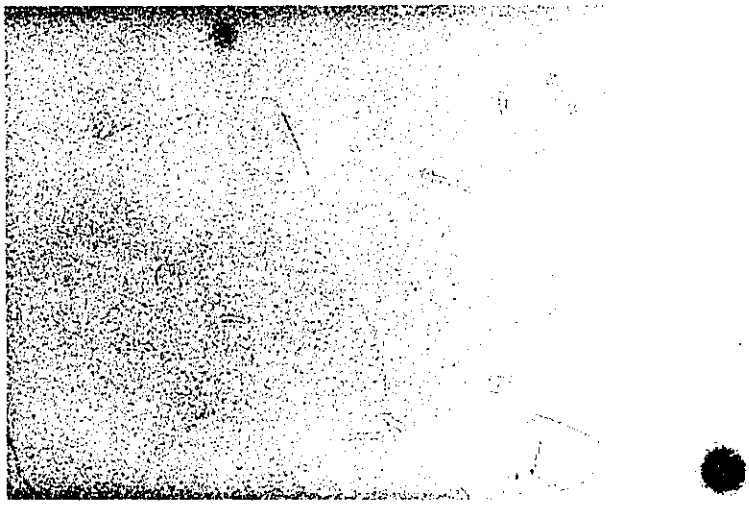
annealed (section 2.5.4; illustrated in Figure 4.1.6) and the C188-primed, cDNA library established by transformation of competent E. coli RR1 cells. Selection of transformants by insertional inactivation (Amp^S, Tet^r) indicated that annealing reactions performed at vector:insert ratios of 1:1 yielded approximately 80% recombinants. Recombinant plasmids of the C188-primed cDNA library were screened by DNA dot blot analysis (Figure 4.1.1) using the 77 bp EcoR I, Pvu II fragment of catalase cDNA recombinant RU1:87 (nucleotides 1172 to 1252, Figure 4.2.2), previously radiolabeled by DNA polymerase I Klenow fragment fill-in of the recessed 3'-end with [α-^{32}P]dATP (section 2.4.5.1), as the probe (probe EP, Figure 4.2.1). Of the 282 recombinants screened initially, two hybridization positives, C188-53 and C188-104, were detected (Figure 4.1.1). As expected, the positive control for the hybridization, cDNA recombinant RU1:87 itself, exhibited a strong hybridization signal (Figure 4.1.1, positions 96, 196 and 296) whereas the negative control for the hybridization, plasmid pBR322 alone, did not yield any signal (Figure 4.1.1, positions 95, 195 and 295). Consequently, both recombinants C188-53 and C188-104 were retained for further analysis.

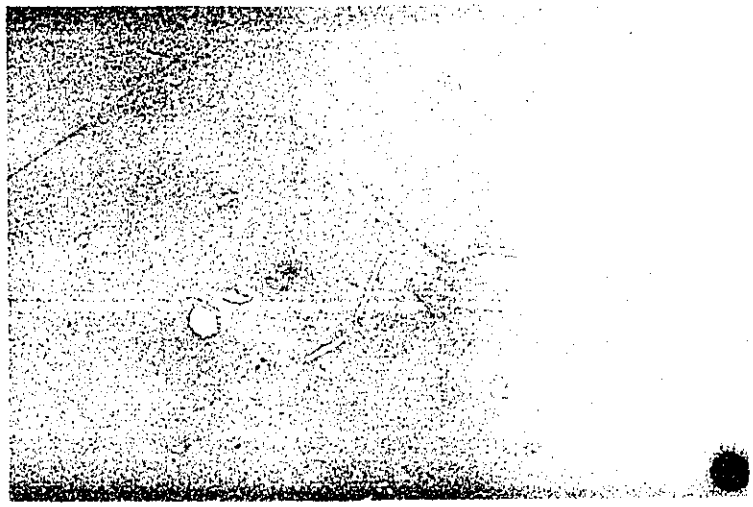
4.1.2 Characterization of the Primer C188-Directed cDNA Recombinants Encoding the 5'-End of the Cat cDNA

To confirm the identity of the two hybridization

Figure 4.1.1. DNA dot blot analysis of plasmid isolated from 282 randomly selected clones of the primer C188-directed cDNA library to poly(A)$^+$RNA from C. tropicalis. Plasmid (≈200 ng of DNA) from cDNA clones (Panel A, clones 1 to 94; Panel B, clones 101 to 194; Panel C, clones 201 to 294), a pBR322 control (positions A95, B195 and C295) and Cat cDNA clone RU1:87 (Rachubinski et al., 1987; positions A96, B196 and C296) was isolated using the method of Holmes and Quigley (1981), denatured, neutralized and spotted onto nitrocellulose. Hybridization (42°C for 24 h) was with 3.3 x 10^5 cpm of [α-^{32}P]dATP-labeled probe EP (see text, section 4.1.1; illustrated in Figure 4.2.1) (specific activity = 2.2 x 10^8 cpm/μg). Washes were conducted at 42°C in 1 x SSC. Exposure was for 25 h at -70°C with Kodak XAR-5 film and 2 intensifying screens.


1
49
85
12
24
36
48
60
72
84
96
A


101
149
185
112
124
136
148
160
172
184
196
B


201
249
285
212
224
236
248
260
272
284
296
C

positive recombinants identified from the C188-primed cDNA library, plasmids from clones C188-53 and C188-104 were subjected to restriction endonuclease digestion and Southern blot analysis. As seen in Figure 4.1.2, probe EP, specific for the 3'-end of the Cat cDNA (Figure 4.2.1), hybridized to a single restriction fragment of ≈500 bp upon Pst I-digestion of plasmid from cDNA recombinant RU1:87 (Figure 4.1.2, lane 3, arrowhead), consistent with that reported by Rachubinski et al. (1987). Pst I-digested pBR322 alone exhibited no hybridization with probe EP (Figure 4.1.2, lane 2). Three other randomly-chosen recombinants, C188-113, C188-222 and C188-253, which appeared negative upon DNA dot blot analysis (Figure 4.1.1), but were included for the sake of completeness, were similarly negative upon Southern blot hybridization with probe EP (Figure 4.1.2, lanes 6, 7 and 8, respectively). As anticipated, both recombinants C188-53 and C188-104 possessed cDNA inserts which exhibited strong hybridization signals with probe EP (Figure 4.1.2, lanes 4 and 5, respectively, arrowheads). Unfortunately, Pst I-digestion of plasmid C188-104 indicated that the cDNA insert was less than 300 bp in length (Figure 4.1.2, lane 5, arrowhead). This was significantly smaller (≈200 bp) than that of recombinant RU1:87 and for this reason, analysis of C188-104 was discontinued. Fortunately, Pst I-digestion of C188-53 demonstrated that this plasmid possessed an insert

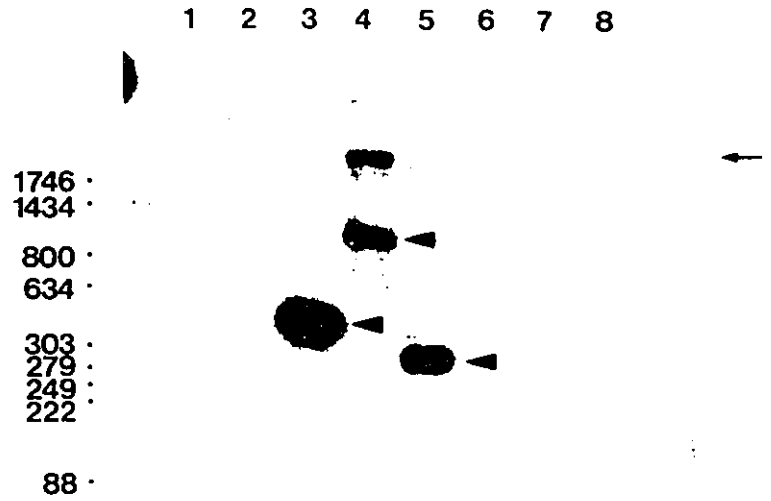


Figure 4.1.2. Southern blot analysis of putative primer C188-directed cDNA recombinants encoding the 5'-end of the Cat cDNA. Plasmid (≈250 ng of DNA) from several independent clones was isolated using the method of Birnboim and Doly (1979), digested with Pst I and subjected to Southern transfer (section 2.4.5.2). Lanes: 1, molecular weight markers - ϕX174-RF DNA Hae III digest; 2, a pBR322 control; 3, cDNA recombinant RU1:87 (Rachubinski et al., 1987); 4-8, cDNA recombinants designated C188-53, C188-104, C188-113, C188-222 and C188-253. Arrowheads indicate the migration of the cDNA inserts harboured by each recombinant plasmid. The arrow indicates the migration of the vector, dG-tailed pBR322, from each recombinant plasmid. Hybridization (42°C for 20 h) was with 5.0 x 10^5 cpm of [α-^{32}P]dATP-labeled probe EP (see text, section 4.1.1) (specific activity = 2.2 x 10^8 cpm/μg). Washes were conducted at 42°C in 1 x SSC. Exposure was for 19 h at -70°C with Kodak XAR-5 film and 2 intensifying screens. Marker sizes in bp (ϕX174-RF DNA Hae III digest) are indicated to the left.

of ≈1.1 kbp in length (Figure 4.1.2, lane 4, arrowhead), thereby exhibiting the potential to extend the Cat cDNA sequence ≈1 kbp 5' of the sequence of recombinant RU1:87. The varying degrees of "hybridization" which appeared to occur between probe EP and vector DNA (dG-tailed pBR322) for recombinants RU1:87, C188-53 and C188-104 (Figure 4.1.2, lanes 3, 4 and 5, respectively, arrow) were probably the result of catalase cDNA insert entrapped within, and therefore co-migrating with, the vector DNA band, since this was not observed with pBR322 alone (Figure 4.1.2, lane 2).

Although recombinant C188-53 was expected to extend the Cat cDNA sequence a minimum of ≈1 kbp farther 5' of the sequence obtained from recombinant RU1:87, subcloning (section 4.1.4) and sequencing (section 4.2.2) of the extreme 5'-end of C188-53 revealed that it was insufficient to encode the entire 5'-region of the Cat cDNA (illustrated in Figure 4.2.1). Subsequent rescreening of the C188-primed cDNA library (≈400 recombinants) by DNA dot blot analysis with a different probe, ^{32}P-end-labeled oligodeoxynucleotide C189, which was synthesized complementary to the extreme 5'-end of the Cat mRNA, did not yield any positive recombinants (not shown) [see Table 2.1.1 for the sequence and annealing position of C189; Figures 4.2.1 and 4.2.2 (C189-box) illustrate the C189 annealing position (nucleotides 128 to 109) along the Cat mRNA]. Consequently, the construction of

a C189-primed cDNA library of poly(A)$^+$RNA from C. tropicalis was necessitated.

4.1.3 Selection and Characterization of Primer C189-Directed cDNA Recombinants Encoding the Extreme 5'-End of the Cat cDNA

Prior to the construction of the C189-primed, cDNA library, primer extension analysis was conducted to determine the distance between the annealing position of primer C189 and the 5'-end of the Cat mRNA. To this end, ≈2 pmoles of ^{32}P-end-labeled primer C189 was annealed to 5 µg of poly(A)$^+$RNA (isolated from n-alkane-grown C. tropicalis) at 42°C for 15 min prior to the addition of 30 units of AMV reverse transcriptase. All other conditions for primer extension were as described for first strand cDNA synthesis according to the method of Gubler and Hoffman (1983) (section 2.5.4). As seen in Figure 4.1.3 (arrow), the major primer C189 extension product was ≈185 nucleotides in length, which corresponded to an approximate transcription start point (tsp) at nucleotide -57 (ie. ≈57 nucleotides upstream of the ATG start codon). A second, minor C189 extension product was ≈175 nucleotides in length (Figure 4.1.3, arrowhead), which corresponded to a second, minor tsp at ≈47 nucleotides upstream of the initiation codon. Although an exact assessment of these tsp would require S1 nuclease mapping, which was not performed, this result is suggestive of at least two tsp for the Cat gene (one major,

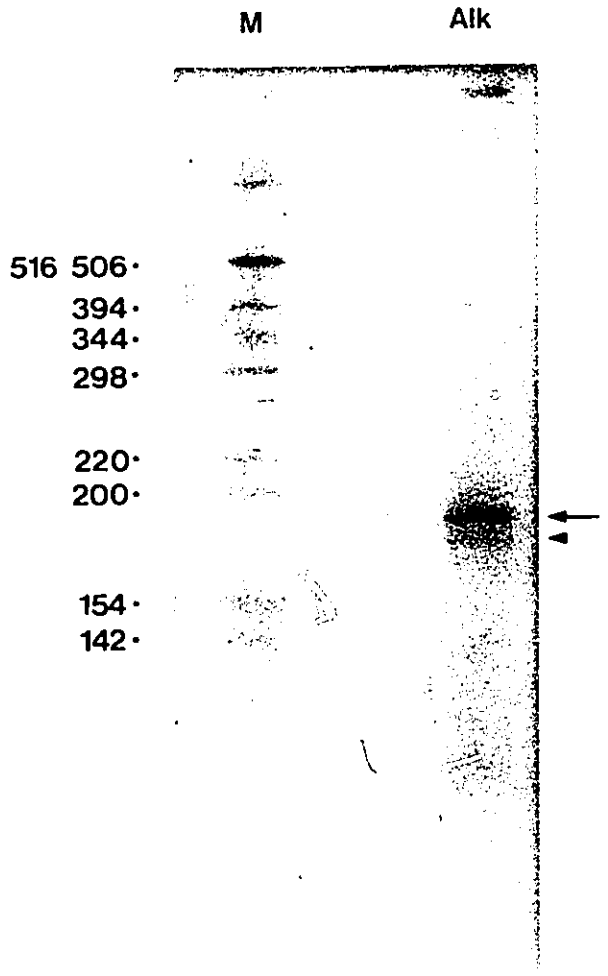


Figure 4.1.3. Determination of the 5'-end of Cat mRNA by primer extension analysis. Approximately 2 pmoles of ^{32}P-end-labeled primer C189 was annealed to 5 µg of poly(A)$^+$RNA (isolated from n-alkane-grown C. tropicalis) at 42°C for 15 min prior to the addition of 30 units of AMV reverse transcriptase. All other reaction conditions were as described for first strand cDNA synthesis according to the method of Gubler and Hoffman (1983) (section 2.5.4). Products were electrophoresed on an 8% polyacrylamide/7M urea gel and analyzed by autoradiography for 20 h at room temperature with Kodak XAR-5 film. The sequence and annealing position of primer C189 are presented in Table 2.1.1 and illustrated in Figures 4.2.1 and 4.2.2. The major primer C189 extension product is denoted by an arrow (lane Alk); a putative minor primer C189 extension product is denoted by an arrowhead. Size markers (^{32}P-end-labeled 1 kbp DNA ladder) were electrophoresed in lane M; their corresponding sizes in bp are indicated to the left.

one minor) localized within the upstream region of nucleotides -70 to -40. By comparison, only a single tsp has been mapped at or near positions -64 and -65 for the C. tropicalis POX4 and 5 genes, respectively (Okazaki et al., 1986); however, multiple tsp have similarly been reported for the genes encoding S. cerevisiae catalase T (CTT1, Spevak et al., 1986) and several mammalian peroxisomal enzymes including human catalase (Quan et al., 1986), rat acyl-CoA oxidase (Osumi et al., 1987), rat bifunctional enzyme (Ishii et al., 1987) and rat catalase (Nakashima et al., 1989). For the S. cerevisiae CTT1 gene, S1 nuclease mapping and primer extension identified two major tsp at positions -25 and -19 (Spevak et al., 1986). With reference to the rat catalase gene, S1 mapping revealed the existence of eight tsp between nucleotide positions -105 and -66 (Nakashima et al., 1989). For the human catalase gene, multiple tsp have been reported, but not definitively assigned (Quan et al., 1986). Interestingly, genes containing multiple tsp often possess upstream regions which are devoid of a TATA box sequence but harbour a 'GC box' upstream from the tsp (Crouse et al., 1985; Ishii et al., 1985). In this respect, the absence of a consensus TATA box sequence and the presence of upstream GC boxes, which are characteristic of the promoters for those genes encoding the aforementioned mammalian peroxisomal enzymes, contrast with

the features found within the upstream region of the Cat gene from C. tropicalis, which contains a TATA box sequence, a GC box and a 'CCAAT box' (discussed in section 4.4.1; Okada et al., 1987), and that of the CTT1 gene of S. cerevisiae, which contains, at minimum, a TATA box sequence (Hartig and Ruis, 1986).

Having established that primer C189 annealed ≈185 nucleotides from the 5'-end of the Cat mRNA, successful screening of a C189-primed cDNA library to obtain the extreme 5'-end of the Cat cDNA necessitated the selection of a recombinant harbouring an insert greater than 185 bp in length (allowing for the dC-tailing). To construct the C189-primed cDNA library, first strand synthesis conditions mimicked those used for primer C189 extension analysis (see above). Double-stranded cDNA was subsequently synthesized, dC-tailed and annealed with Pst I-digested, dG-tailed pBR322 (vector:insert ratios of 1:1 and 1:4) in the same manner as described for the C188-primed cDNA library (section 4.1.1; illustrated in Figure 4.1.6) The C189-primed cDNA library was then established by transformation of competent E. coli RR1 cells. Four hundred, tetracycline-resistant transformants were screened by colony hybridization using a ^{32}P-end-labeled synthetic probe, oligodeoxynucleotide C231, which was synthesized complementary to the extreme 5'-end of the coding strand of cDNA recombinant C188-53 [see Table

2.1.1 for the sequence and annealing position of C231; Figure 4.2.1 illustrates the C231 annealing position (nucleotides 84 to 65) along recombinant C188-53]. Twenty-two, putative, positive clones were detected; however, only one of these, C189-296, exhibited an intense hybridization signal with probe C231 (not shown). Consequently, to confirm or negate their identities as C231-hybridization positive clones, plasmid from these transformants was subjected to restriction endonuclease digestion and Southern blot analysis. As seen in Figure 4.1.4, probe C231 hybridized to a single restriction fragment of ≈200 bp upon Pst I-digestion of plasmid from only one cDNA clone, C189-296 (Figure 4.1.4, lane 11, arrowhead). This, in conjunction with the similarly strong hybridization signal observed between probe C231 and the ≈1.1 kbp cDNA insert of Cat recombinant C188-53 (Figure 4.1.4, lane 15), implied that recombinant C189-296 should harbour sufficient sequence to encode the entire 5'-end of the Cat cDNA. Plasmid from the remaining 21 clones, only 11 of which are presented in the Southern blot of Figure 4.1.4 (lanes 2-10, 12 and 13), exhibited no hybridization with probe C231, as was the case for Pst I-digested pBR322 alone (Figure 4.1.4, lane 14). An explanation accounting for "false" hybridization observed between the probe and vector DNA (Figure 4.1.4, lane 11, arrow) was previously presented in section 4.1.2.

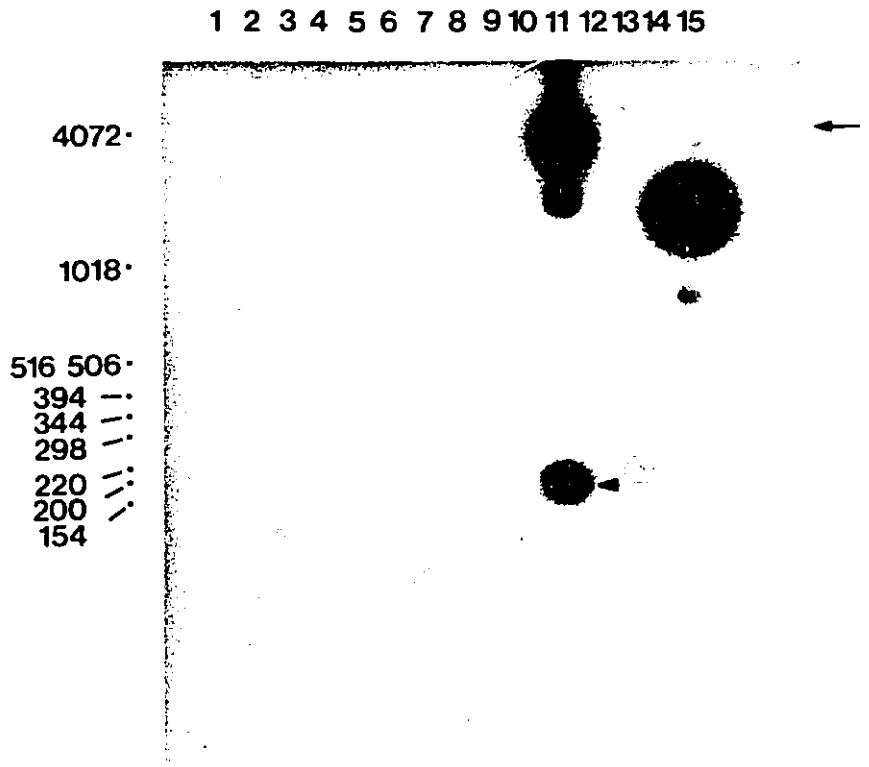


Figure 4.1.4. Southern blot analysis of putative primer C189-directed cDNA recombinants encoding the extreme 5'-end of the Cat cDNA. Plasmid (≈150 ng) from several independent clones was isolated using the method of Birnboim and Doly (1979), digested with Pst I and subjected to Southern transfer (section 2.4.5.2). Lanes: 1, molecular weight markers - 1 kbp DNA ladder; 2-13, cDNA recombinants designated C189-158, -206, -226, -227, -228, -238, -278, -281, -287, -296, -297 and -328; 14, a pBR322 control; 15, cDNA insert of Cat recombinant C188-53. The arrowhead indicates the migration of the cDNA insert of recombinant C189-296. The arrow indicates the migration of the parent vector, dG-tailed pBR322. The cDNA insert of Cat recombinant C188-53 was loaded after significant migration of the other fragments had occurred and has therefore migrated anomalously relative to its ≈1.1 kbp size. Hybridization (37°C for 16 h) was with 1.0 x 10^6 cpm of ^{32}P-end-labeled probe C231 (specific activity = 3.2 x 10^8 cpm/µg) (the sequence and annealing position of probe C231 are presented in Table 2.1.1 and illustrated in Figure 4.2.1). Washes were conducted at 37°C in 1 x SSC. Exposure was for 14 h at -70°C with Kodak XAR-5 film and 2 intensifying screens. Marker sizes in bp (1 kbp DNA ladder) are indicated to the left.

Subsequent subcloning (section 4.1.4) and sequencing (section 4.2.2) of the cDNA insert from recombinant C189-296 generated sufficient DNA sequence (≈150 nucleotides) to encode the desired 5'-end of the Cat cDNA (illustrated in Figure 4.2.1); unfortunately, however, a sequence anomaly appeared to exist (see section 4.2.2) which resulted in an in-frame TGA termination codon (nucleotides 42 to 44, Figure 4.2.2) within the correct ORF encoding catalase, previously established from the sequence of recombinants C188-53 and CdT-1 (RU1:87) by translation and comparison of the deduced amino acid sequences (from three different reading frames) to those reported for several other catalases (presented in Figure 4.4.5). In an attempt to resolve this problem, an additional ≈2000 tetracycline-resistant transformants from the C189-primed cDNA library were screened with ^{32}P-end-labeled probe C231 (not shown). Plasmid was isolated from each of the 21 possible C231-hybridization positive clones identified and subjected to Pst I-digestion and Southern blot analysis. As revealed by Figure 4.1.5 (Panels A and B), 9 recombinants, including C189-1.1, -2.1, -2.2, -3.1, -4.1, -6.1, -6.2, -7.5 and -8.1, were confirmed to encode at least some portion of the extreme 5'-end of the Cat cDNA. Collectively, the cDNA inserts of these recombinants ranged in size from ≈140 bp [or smaller, since the insert of recombinant C189-2.2 (Figure 4.1.5, Panel A, lane 4, which

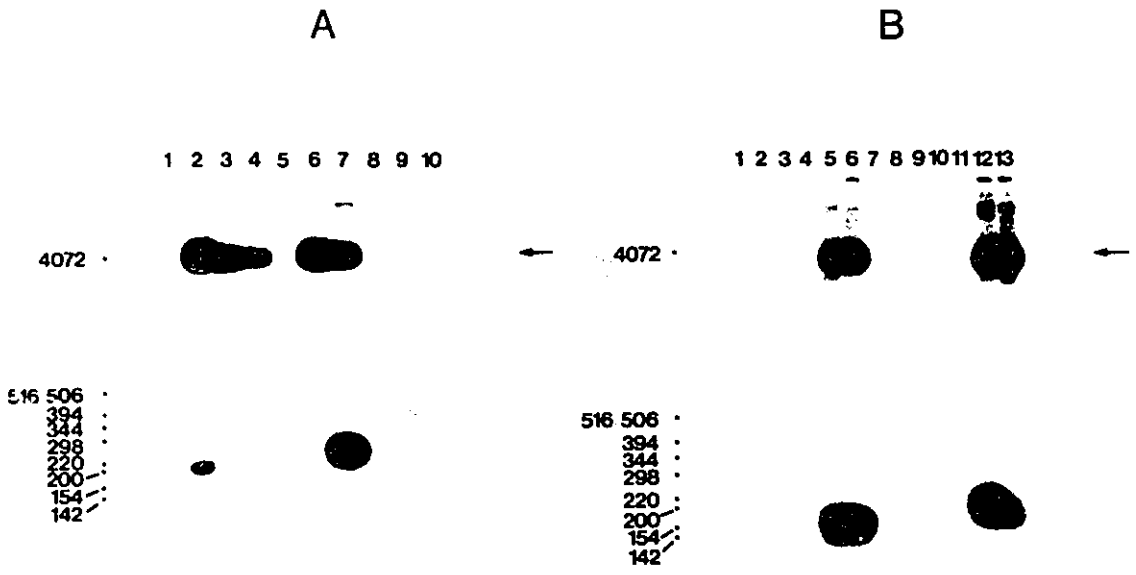


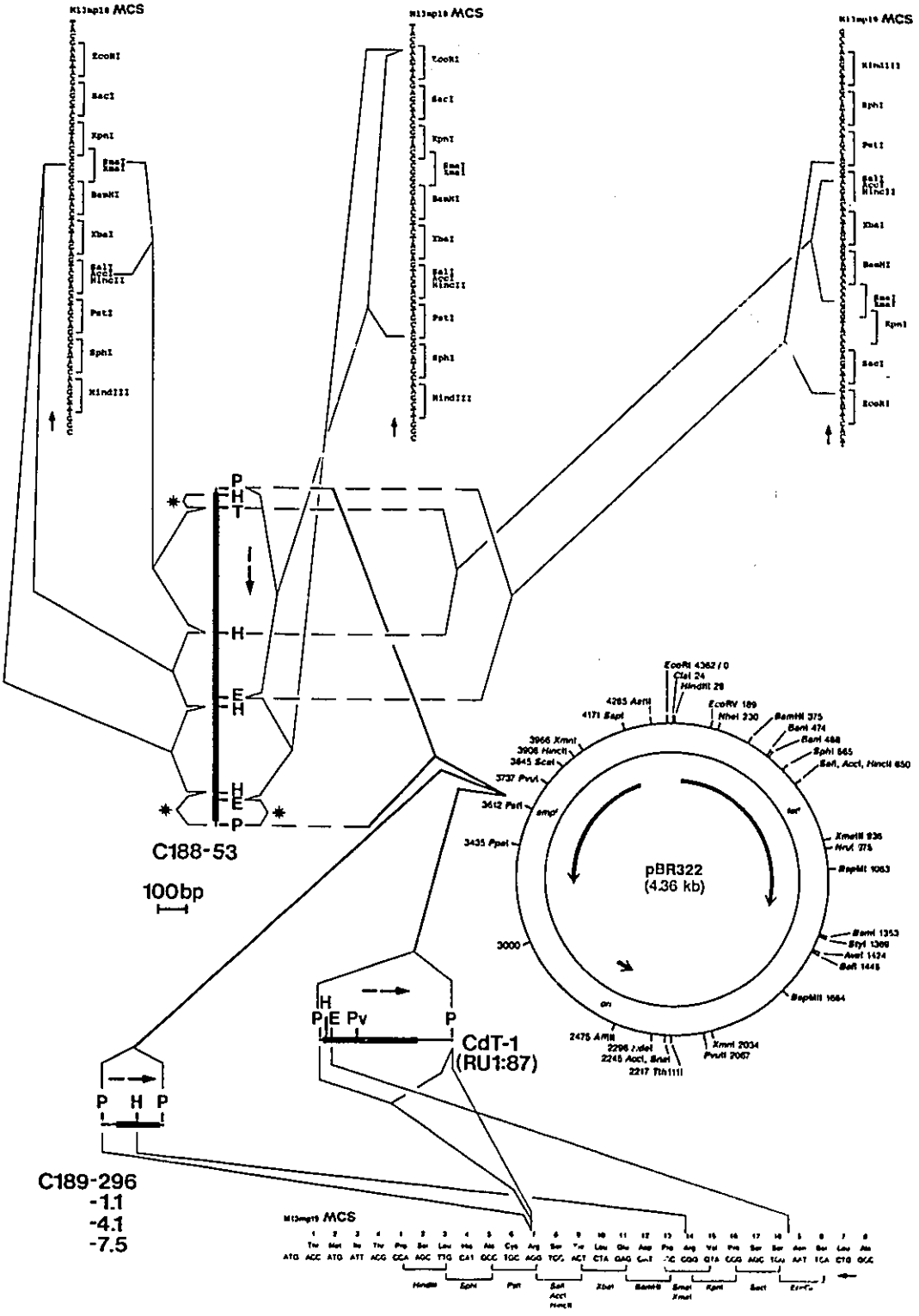
Figure 4.1.5. Southern blot analysis of additional primer C189-directed cDNA recombinants encoding the extreme 5'-end of the Cat cDNA. Plasmid (≈150 ng) from several independent clones was isolated using the method of Birnboim and Doly (1979), digested with Pst I and subjected to Southern transfer (section 2.4.5.2). Lanes: Panel A: 1, molecular weight markers - 1 kbp DNA ladder; 2-10, cDNA recombinants designated C189-1.1, -2.1, -2.2, -2.3, -3.1, -4.1, -4.2, -4.3 and -4.4; Panel B: 1, molecular weight markers - 1 kbp DNA ladder; 2-13, cDNA recombinants designated C189-5.2, -5.3, -5.4, -6.1, -6.2, -6.3, -7.1, -7.2, -7.3, -7.4, -7.5 and -8.1. For both Panels A and B, the arrow indicates the migration of the vector, dG-tailed pBR322, from each recombinant plasmid. Hybridization (37°C for 9 h) was with 6.9×10^5 cpm of ^{32}P-end-labeled probe C231 (specific activity = 3.2×10^8 cpm/μg). Washes were conducted at 37°C in 1 x SSC. Exposure was for 13 h at -70°C with Kodak XAR-5 film and 2 intensifying screens. Marker sizes in bp (1 kbp DNA ladder) are indicated to the left of each panel.

appeared positive due to the "false" hybridization of C231 with fragments entrapped within the vector DNA band (Figure 4.1.5, Panels A and B, arrows), was not detectable] to greater than 220 bp. Consequently, three independent recombinants, C189-1.1, -4.1 and -7.5, which possessed the largest cDNA inserts (≈200 bp or greater), were retained for DNA sequencing purposes. Single- and double-stranded sequencing of each of these cDNA inserts yielded the entire 5'-end of the Cat cDNA, and various portions of 5'-flanking sequence (Figure 4.2.1), without the occurrence of the previously observed DNA sequence anomaly (see section 4.2.2).

4.1.4 Strategy Used to Subclone the cDNA Fragments Encoding Catalase

To confirm the previously reported sequence of the 3'-coding and -flanking sequences of the Cat cDNA (recombinant RU1:87; Rachubinski et al., 1987), and generate the sequence data necessary for primer synthesis and subsequent extension of Cat mRNA to obtain the 5'-sequences of the Cat cDNA, both the ≈450 bp EcoR I, Pst I and the ≈500 bp Pst I cDNA fragments of recombinant CdT-1, isolated by fragment elution from nondenaturing polyacrylamide gels (sections 2.4.3.1), were ligated into compatibly-digested M13mp19 [Figure 4.1.6, recombinant CdT-1 (RU1:87)]. From the sequence data obtained (section 4.2.2),

Figure 4.1.6. Schematic representation of the Cat cDNA recombinants and the subcloning strategy used to sequence the Cat cDNA. The annealing of dC-tailed, double-stranded cDNA with Pst I-digested, dG-tailed pBR322 is illustrated for Cat recombinants CdT-1 (RU1:87; Rachubinski et al., 1987), C188-53 and C189-296, -1.1, -4.1 and -7.5. The partial restriction endonuclease cleavage maps established for these recombinants (see text, section 4.1.4) are presented with that portion of the ORF of the Cat cDNA harboured by each recombinant indicated by a thick line. Flanking sequences [5'-sequence for recombinants C189-296, -1.1, -4.1 and -7.5 and 3'-sequence for recombinant CdT-1 (RU1:87)] and dC-tailing are represented by adjoining thin lines. Restriction endonuclease sites included are: E, EcoR I; H, Hind III; P, Pst I; Pv, Pvu II; T, Taq I (the Pst I sites are not inherent to the cDNA sequence but are generated upon annealing with Pst I-digested, dG-tailed pBR322). Those cDNA fragments subcloned into the MCS of M13mp18 and/or 19 (Norrander et al., 1983), and their orientations (see text, section 4.1.4), are illustrated (where possible). Fragments denoted by asterisks were also subcloned into the MCS of M13mp18 and/or 19 (section 4.1.4) but, for simplicity, this is not illustrated. The arrow shown at the 3'-end of each MCS represents the direction of sequencing. Dashed arrows indicate the direction of Cat cDNA polarity. The orientation of each of the cDNA inserts within the Pst I site of pBR322 was not determined. The plasmid map of pBR322 was reproduced from Sambrook et al. (1989).


M13mp18 MCS
EcoRI
SacI
KpnI
SmaI
XmaI
BamHI
XbaI
SalI
AccI
HincII
PstI
SphI
HindIII
P
H
T
E
C188-53
100bp
pBR322
(4.36 kb)
CdT-1
(RU1:87)
Pv
C189-296
-1.1
-4.1
-7.5

oligodeoxynucleotide C188 was synthesized and used to construct a synthetic primer-directed cDNA library from which recombinant C188-53, presumed to encode the remainder of the 5'-end of the Cat cDNA, was identified (section 4.1.2). Pst I-digestion of C188-53 revealed that this recombinant harboured an ≈1.1 kbp cDNA insert lacking any internal Pst I site(s) (Figure 4.1.2). Upon cleavage with EcoR I, the ≈1.1 kbp cDNA insert yielded three fragments of ≈700 bp, 300 bp and 100 bp. Realizing that the ≈100 bp fragment was the equivalent of the EcoR I, Pst I fragment of recombinant CdT-1, and that EcoR I-digestion of C188-53 generated fragments of ≈4.5 kbp, 800 bp and 300 bp, suggested that the partial restriction map of the ≈1.1 kbp insert of C188-53 was comprised of a 5', ≈700 bp, Pst I, EcoR I fragment, an internal, ≈300 bp, EcoR I fragment and a 3', ≈100 bp, EcoR I, Pst I fragment (Figure 4.1.6, recombinant C188-53). For sequencing purposes, each of these fragments was subcloned, subsequent to fragment elution from nondenaturing polyacrylamide gels, as follows (illustrated in Figure 4.1.6): the ≈700 bp Pst I, EcoR I fragment was ligated (forced-orientation) into compatibly-digested M13mp18 and 19, the ≈300 bp EcoR I fragment was ligated into EcoR I-digested M13mp18 (in both orientations) and the ≈100 bp EcoR I, Pst I fragment was ligated (forced-orientation) into compatibly-digested M13mp18 and 19. To

ensure that smaller restriction fragments were not unknowingly excluded from these ligations, a second library of overlapping sequencing fragments was established by Taq I, Hae III-digestion of the Pst I-liberated cDNA insert of C188-53. This digestion yielded six fragments which, when resolved according to their 5' to 3' orientation within the C188-53 insert (Figure 4.1.6, recombinant C188-53), included a small (non-agarose gel resolvable) Pst I, Hae III fragment (which constituted the 5'-tail only), an ≈60 bp Hae III, Taq I fragment, an ≈450 bp Taq I, Hae III fragment, an ≈250 bp Hae III fragment, an ≈300 bp Hae III fragment and an ≈90 bp Hae III, Pst I fragment. For sequencing purposes, each of these fragments (excluding the small Pst I, Hae III fragment) was subcloned, after fragment elution as above, in the following manner (illustrated in Figure 4.1.6): the ≈60 bp Hae III, Taq I fragment was ligated (forced-orientation) into Sma I, Acc I-digested M13mp18 only, the ≈450 bp Taq I, Hae III fragment was ligated (forced-orientation) into Acc I, Sma I-digested M13mp18 and 19, both the ≈250 bp and ≈300 bp Hae III fragments were ligated into Sma I-digested M13mp18 (in both orientations) and the ≈90 bp Hae III, Pst I fragment was ligated (forced-orientation) into Sma I, Pst I-digested M13mp19 only. Unfortunately, sequencing the 5'-end of the ≈700 bp Pst I, EcoR I fragment of the cDNA insert of C188-53 (section 4.2.2) revealed that C188-primed extension

of Cat mRNA had not been sufficient to encode the entire 5'-end of the Cat cDNA. Consequently, a second extension of Cat mRNA, primed by oligonucleotide C189, was conducted (section 4.1.3). From the C189-primed cDNA library, four recombinants, C189-296, -1.1, -4.1 and -7.5, were selected as the most likely to encode the extreme 5'-end of the Cat cDNA (section 4.1.3). Since primer C189 was constructed such that it annealed along the Cat mRNA at a position 3' to the most 5' Hae III site of the cDNA insert of recombinant C188-53 (Figure 4.2.1), each of the cDNA inserts of the aforementioned C189-primed recombinants was susceptible to digestion with Hae III. Thus, for sequencing purposes, Pst I-liberated cDNA inserts, isolated from recombinants C189-296, -1.1, -4.1 and -7.5, were digested with Hae III and subcloned by ligation (forced-orientation) into Pst I, Sma I-digested M13mp19 only (Figure 4.1.6, recombinants C189-296, -1.1, -4.1 and -7.5).

4.2 SEQUENCE DETERMINATION OF THE COMPOSITE cDNA ENCODING Cat

4.2.1 Sequencing Strategy of the cDNAs Encoding Cat

As described in section 4.1, cDNA recombinants encoding three distinct regions of Cat from C. tropicalis were selected such that determination of their overlapping nucleotide sequences would allow reconstitution of the entire ORF of the Cat cDNA. Accordingly, the cDNA inserts

of recombinants CdT-1 (encoding the 3'-end of the Cat cDNA), C188-53 (encoding the internal majority of the Cat cDNA) and C189-296, -1.1, -4.1 and -7.5 (each encoding the extreme 5'-end of the Cat cDNA) were subcloned into the M13mp18/19 vectors (section 4.1.4; illustrated in Figure 4.1.6) and sequenced, using the single-stranded ddNTP chain-termination method of Sanger et al. (1977), according to the strategy outlined in Figure 4.2.1. For the most part, sequence data were generated using M13 universal, forward (-20 to -40) sequencing primers; however, three Cat (C) specific primers (C1, C2 and C3, Figure 4.2.1), synthesized complementary to a previously determined nucleotide sequence of the cDNA insert of CdT-1, were also used to confirm and extend the sequence data initially obtained for CdT-1 using the M13 universal primers [Figure 4.2.1, see Table 2.1.1 for the Cat (C) primer annealing positions and sequences]. In addition, to validate the sequence data generated by single-stranded sequencing of the cDNA inserts of recombinants C189-1.1, -4.1 and -7.5 (Figure 4.2.1, recombinants 24:1-6), double-stranded ddNTP sequencing (Zhang et al., 1988) of these plasmids was conducted using either primer C189 or C231 (Figure 4.2.1, recombinants 23:1-3).

Collectively, 92% of the entire DNA sequence depicted by the partial restriction map of the Cat cDNA and flanking sequences (Figure 4.2.1) was determined in the 5'

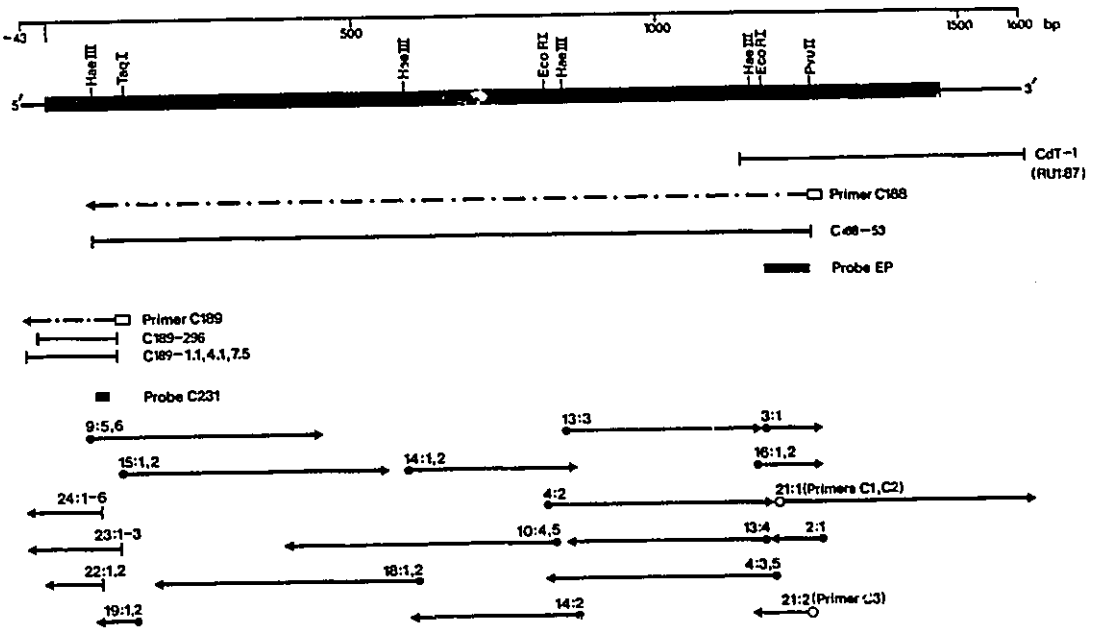


Figure 4.2.1. Physical map and sequencing strategy of the cDNAs coding for Cat from C. tropicalis. The direction of Cat cDNA polarity is shown; the accompanying base pair numbering of the composite Cat cDNA is presented in Figure 4.2.2. The protein coding region is indicated by the thick line. The start points, direction and extent of primer-extended cDNAs are indicated by open rectangles (primers C188 and C189) and dash-dot arrows. Recombinants CdT-1 (RU1:87), C188-53 and C189-296, -1.1, -4.1 and -7.5 (shown as thin lines terminated by short vertical bars) were the result of primer extension from oligo(dT), primer C188 and primer C189, respectively. Recombinant CdT-1 (RU1:87) was shown to encode catalase by hybridization-selection translation and immunoprecipitation (Rachubinski et al., 1987). Probes EP and C231 (indicated by solid rectangles) were used to identify recombinant C188-53 and recombinants C189-296, -1.1, -4.5 and -7.5, respectively. Only those restriction endonuclease sites used to subclone DNA fragments for the purpose of sequencing are shown. The start points, direction and extent of sequence determinations are indicated by continuous arrows: the sequence of CdT-1, which confirms the sequence reported by Rachubinski et al. (1987) for recombinant RU1:87, is represented by open circles; closed circles indicate sequencing of recombinant C188-53; short vertical bars represent single- or double-stranded sequencing of recombinants C189-296, -1.1, -4.1 and -7.5 (see text, section 4.2.1). Numbers above the sequencing start points denote the different M13mp18/19 recombinants sequenced in each case. Sequence data were generated using M13 universal, forward (-20 and -40) primers; three Cat(C) specific primers (C1, C2 and C3) were also used to sequence parts of recombinant CdT-1 (see Table 2.1.1 for the annealing positions and sequences of these primers). The map does not contain poly (dG·dC) tails but is otherwise drawn to scale.

to 3' (coding) direction, with each nucleotide sequenced at least twice [except for the sequence of CdT-1 which was sequenced only once but confirmed by comparison to that sequence published for recombinant RU1:87 (Rachubinski et al., 1987)]. Those discrepancies encountered upon sequence comparison of CdT-1 and RU1:87 (Figure 4.2.2, bracketed nucleotides) were resolved by both resequencing in the 5' to 3' direction and by sequencing that region, at least once, in the 3' to 5' (noncoding) direction. By comparison, 98% of the sequence depicted in Figure 4.2.1 (excluding that of CdT-1 which was, as mentioned above, confirmed independently) was determined in the 3' to 5' direction, with each nucleotide sequenced at least twice. Those remaining nucleotides, which were determined in one direction only, were confirmed by sequencing four independent templates (Figure 4.2.1).

4.2.2 Composite Nucleotide Sequence of the cDNA Encoding Cat

Of those recombinants selected to encode the entire ORF of the Cat cDNA, the insert of the initial recombinant, CdT-1 (RU1:87), harboured 462 bp of the 3'-end of the Cat cDNA (nucleotides 1129 to 1590, Figure 4.2.2), independent of the poly(A) tract. This included 132 bp of 3'-noncoding sequence (nucleotides 1459 to 1590, Figure 4.2.2) and extended 330 bp in the 5' direction (nucleotides 1129 to

Figure 4.2.2. Composite nucleotide sequence of the Cat cDNA and deduced amino acid sequence of Cat from C. tropicalis. Nucleotides are numbered in the 5' to 3' direction beginning with the first residue of the ATG start codon. Portions of the 5'- and 3'-flanking regions are shown; nucleotides 5' to position 1 are designated by negative numbers. Within the 5' noncoding region, an in-frame TAA termination codon is denoted by 'stop'. Boxes labeled C188 or C189 represent the annealing positions of oligonucleotides C188 and C189, synthesized complementary to these sequences (see Table 2.1.1), which were used to primer extend the Cat mRNA. Nucleotide differences between the Cat cDNA sequence reported here and that of Rachubinski et al. (1987) are bracketed above the appropriate positions. The small 't' present between nucleotides 12 and 13 represents the position of the "extra" thymidine nucleotide discovered upon sequencing the cDNA insert of recombinant C189-296 (see section 4.2.2). Within the 3'-noncoding region, a putative polyadenylation sequence is boxed and possible transcription termination sequences are double-underlined. The deduced amino acid sequence of Cat is presented below the nucleotide sequence in the standard three-letter code. Those NH_2-terminal amino acids of Cat which were sequenced by Ueda et al. (1987) and Okada et al. (1987) are dash-dot underlined. Underlined amino acids represent those residues of bovine liver catalase (Schroeder et al., 1982), in contact with the heme moiety, which are conserved in Cat. Double asterisks indicate those amino acids of bovine catalase, participating directly in catalysis, which are conserved in Cat. The boxed tripeptide denotes a variant of the PTS identified by Gould et al. (1989).

```
                                                                   -1
                         AAACAATCAACATCACACCACACCCCCATACTAATTATTCAAT
                                                              stop     60
   ATG GCT CCT ACC TTC ACA AAC TCA AAC GGT CAA CCT ATC CCT GAA CCT TTC GCC ACT CAA
 1 Met Ala Pro Thr Phe Thr Asn Ser Asn Gly Gln Pro Ile Pro Glu Pro Phe Ala Thr Gln
                                                                    C189      120
   AGA GTT GGC CAA CAC GGT CCT TTG TTG TTG CAA GAC TTC AAC TTG ATT GAC TCC TTG GCT
21 Arg Val Gly Gln His Gly Pro Leu Leu Leu Gln Asp Phe Asn Leu Ile Asp Ser Leu Ala
                                                                              180
   CAC TTC GAC AGA GAA AGA ATC CCA GAA AGA GTC GTC CAC GCT AAA GGC TCC GGT GCT TAC
41 His Phe Asp Arg Glu Arg Ile Pro Glu Arg Val Val His Ala Lys Gly Ser Gly Ala Tyr
                                                                              240
   GGT GTT TTT GAA GTC ACC GAT GAC ATC ACC GAC GTT TGT GCC GCA AAG TTT TTG GAC ACT
61 Gly Val Phe Glu Val Thr Asp Asp Ile Thr Asp Val Cys Ala Ala Lys Phe Leu Asp Thr
                                                                              300
   GTC GGT AAG AAA ACC AGA ATT TTC ACT AGA TTC TCC ACC GTC GGT GGT GAA TTG GGT TCC
81 Val Gly Lys Lys Thr Arg Ile Phe Thr Arg Phe Ser Thr Val Gly Gly Glu Leu Gly Ser
                                                                              360
   GCT GAC ACC GCC AGA GAC CCA AGA GGT TTT GCT ACC AAG TTT TAC ACC GAA GAA GGT AAC
101 Ala Asp Thr Ala Arg Asp Pro Arg Gly Phe Ala Thr Lys Phe Tyr Thr Glu Glu Gly Asn
                                                                              420
   TTG GAC TTG GTT TAC AAC AAC ACC CCA GTC TTT TTC ATT AGA GAC CCT TCT AAG TTC CCT
121 Leu Asp Leu Val Tyr Asn Asn Thr Pro Val Phe Phe Ile Arg Asp Pro Ser Lys Phe Pro
                                                                              480
   CAC TTC ATC CAC ACC CAA AAG AGA AAC CCA GAA ACT CAC TTG AAG GAC GCC AAC ATG TTC
141 His Phe Ile His Thr Gln Lys Arg Asn Pro Glu Thr His Leu Lys Asp Ala Asn Met Phe
                                                                              540
   TGG GAC TAC TTG ACC AGC AAC GAA GAA TCC GTC CAC CAA GTT ATG GTT TTG TTC TCC GAC
161 Trp Asp Tyr Leu Thr Ser Asn Glu Glu Ser Val His Gln Val Met Val Leu Phe Ser Asp
                                                                              600
   AGA GGT ACC CCA GCT TCC TAC AGA GAA ATG AAT GGT TAC TCT GGC CAC ACC TAC AAG TGG
181 Arg Gly Thr Pro Ala Ser Tyr Arg Glu Met Asn Gly Tyr Ser Gly His Thr Tyr Lys Trp
                                                                              660
   TCC AAC AAC AAG GGT GAA TGG TTC TAT GTT CAA GTC CAT TTC ATC AGT GAC CAA GGT ATC
201 Ser Asn Asn Lys Gly Glu Trp Phe Tyr Val Gln Val His Phe Ile Ser Asp Gln Gly Ile
                                                                              720
   AAG ACC TTG ACC AAC GAA GAA GCT GGC TCC TTG GCT GGT TCC AAC CCA GAC TAT GCT CAA
221 Lys Thr Leu Thr Asn Glu Glu Ala Gly Ser Leu Ala Gly Ser Asn Pro Asp Tyr Ala Gln
                                                                              780
   GAA GAT TTG TTC AAG AAC ATT GCT GCT GGT AAC TAC CCA TCC TGG ACT TGT TAC ATC CAA
241 Glu Asp Leu Phe Lys Asn Ile Ala Ala Gly Asn Tyr Pro Ser Trp Thr Cys Tyr Ile Gln
                                                                              840
   ACC ATG ACT GAA GCT CAA GCT AAA GAA GCC GAA TTC TCT GTC TTT GAC TTG ACC AAG GTC
261 Thr Met Thr Glu Ala Gln Ala Lys Glu Ala Glu Phe Ser Val Phe Asp Leu Thr Lys Val
                                                                              900
   TGG CCA CAC GGA AAG TAC CCA ATG AGA AGA TTT GGT AAG TTT ACT TTG AAC GAA AAC CCA
281 Trp Pro His Gly Lys Tyr Pro Met Arg Arg Phe Gly Lys Phe Thr Leu Asn Glu Asn Pro
                                                                              960
   AAG AAC TAC TTT GCC GAA GTT GAA CAA GCT GCT TTC TCC CCA GCC CAC ACC GTC CCA CAC
301 Lys Asn Tyr Phe Ala Glu Val Glu Gln Ala Ala Phe Ser Pro Ala His Thr Val Pro His
                                                                             1020
   ATG GAA CCA TCC GCT GAC CCA GTC TTG CAA TCT AGA TTG TTC TCT TAT GCT GAT ACT CAC
321 Met Glu Pro Ser Ala Asp Pro Val Leu Gln Ser Arg Leu Phe Ser Tyr Ala Asp Thr His
                                                                             1080
   AGA CAC AGA TTG GGT ACC AAC TAC ACC CAA ATC CCA GTC AAC TGC CCA GTC ACT GGT GCT
341 Arg His Arg Leu Gly Thr Asn Tyr Thr Gln Ile Pro Val Asn Cys Pro Val Thr Gly Ala
                                                                   (G)       1140
   GTT TTC AAC CCA CAC ATG AGA GAC GGT GCC ATG AAT GTC AAT GGT AAC TTG GGC AAC CAT
361 Val Phe Asn Pro His Met Arg Asp Gly Ala Met Asn Val Asn Gly Asn Leu Gly Asn His
                            (A) (A)(A)           (A)     (C)                 1200
   CCA AAC TAC TTG GCC TCT GAC AAG CCT GTT GAA TTC AAG CAG TTC TCT CTC CAA GAA GAC
381 Pro Asn Tyr Leu Ala Ser Asp Lys Pro Val Glu Phe Lys Gln Phe Ser Leu Gln Glu Asp
                                                                  C188       1260
   CAA GAA GTC TGG CAC GGT GCT GCT ACT CCA TTC CAC TGG AAG GCT ACC CCA GCT GAT TTC
401 Gln Glu Val Trp His Gly Ala Ala Thr Pro Phe His Trp Lys Ala Thr Pro Ala Asp Phe
                                                                             1320
   AAG CAA GCC ACT GAA TTG TGG AAG GTG TTG AAG AAA TAT CCA AAC CAA CAA GAA CAC TTG
421 Lys Gln Ala Thr Glu Leu Trp Lys Val Leu Lys Lys Tyr Pro Asn Gln Gln Glu His Leu
                                                                             1380
   GCC CAC AAT GTT GCT GTC CAT GCC TCT GCT GCC GAT GCT CCA ATC CAA GAC AGA GTC ATT
441 Ala His Asn Val Ala Val His Ala Ser Ala Ala Asp Ala Pro Ile Gln Asp Arg Val Ile
                                                                             1440
   GCT TAC TTT ACC AAG GTC CAC CCA GAC TTG GGA GAC CTC ATC AAG AAG GAA ATC TTG GAA
461 Ala Tyr Phe Thr Lys Val His Pro Asp Leu Gly Asp Leu Ile Lys Lys Glu Ile Leu Glu
                                                                             1513
   TTG TCT CCA AGA AAG TAA AGAAGATTCTGATTTGATATTGGAGTGCTAAAGGTCTTAGTTATGTTTGTTTCTT
481 Leu Ser Pro Arg Lys End

                                                                             1592
   TTCTTTTTTCTTATTAATATTTATAGCTGTTGTAATGCGTTTTACAGTAAATTGAATGGTTATATTTGATATATCTGAA

      1600
   AAAAAAAA
```

1458, Figure 4.2.2) from the putative TAA termination codon, which was established by the presence of only a single ORF exhibiting deduced amino acid identity to other catalases (Figure 4.4.5; Rachubinski et al., 1987). Primer extension of Cat mRNA to obtain the remainder of the 5'-end of the Cat cDNA yielded C188-53 as the largest recombinant (section 4.1.2), the cDNA insert of which was 1,186 bp in length (nucleotides 58 to 1243, Figure 4.2.2), not including dC-tailing. Even though this generated an additional 1,071 bp of Cat cDNA sequence (nucleotides 58 to 1128, Figure 4.2.2), this appeared insufficient to encode the entire 5'-end of the Cat cDNA since translation of this sequence, from the first, in-frame ATG codon of the reading frame established by translation of the sequence of CdT-1 (ie. met-159 of Figure 4.2.2), yielded a polypeptide of M_r ≈37000, the size of which was significantly less than that previously reported for purified catalase from C. tropicalis (M_r 54000; Yamada et al., 1982). Consequently, a second primer extension was conducted (section 4.1.3), from which one initial positive recombinant, C189-296, expected to encode the extreme 5'-end of the Cat cDNA, was obtained. The cDNA insert of C189-296 was determined to be 150 bp in length (not including dC-tailing), thereby generating 99 nucleotides of novel 5'-sequence, which should have been sufficient to encode the extreme 5'-end of the Cat cDNA.

Upon translation of this sequence, however, in the correct Cat-encoding reading frame previously established from the sequence of recombinants C188-53 and CdT-1, an in-frame TGA termination codon (nucleotides 42 to 44, Figure 4.2.2) was encountered at the 5'-end of the cDNA. In an attempt to resolve this problem, three additional, independent, 5'-recombinants, C189-1.1, -4.1 and -7.5, selected upon rescreening (section 4.1.3), were subjected to both single- and double-stranded sequencing (section 4.2.1, Figure 4.2.1). Each of these inserts encoded the remaining 57 bp of the 5'-end of the Cat cDNA, as well as a maximum of 43 bp (derived from recombinant C189-4.1) of 5'-noncoding sequence (nucleotides -43 to 57, Figure 4.2.2). Contrary to the situation observed with the cDNA insert of C189-296, translation of each of these sequences in the previously established Cat ORF (ie. from the first ATG codon encountered, met-1 of Figure 4.2.2) allowed continuous readthrough of a composite sequence compiled from the cDNA inserts of recombinants C189-1.1 or -4.1 or -7.5, through C188-53, to the TAA stop codon of CdT-1 (Figure 4.2.2). Subsequent comparison of the sequence of the cDNA inserts from C189-1.1, -4.1 and -7.5 to that of C189-296 revealed the presence of an "extra" thymidine nucleotide within the C189-296 sequence (Figure 4.2.2, small 't' between nucleotide positions 12 and 13). Upon translation, this

anomalous nucleotide dictated a sequence frameshift which introduced a TGA stop codon at a position ten codons farther downstream of the inserted nucleotide (Figure 4.2.2). At present, the occurrence of this sequence anomaly remains unexplained.

Once compiled, translation in the three possible reading frames of the composite sequence of the Cat cDNA revealed the existence of a single ORF of 1,455 nucleotides (Figure 4.2.2) stretching from the putative ATG initiation codon to a TAA termination codon (nucleotides 1456 to 1458). As expected, the length of the ORF of the Cat cDNA was identical to that reported for the catalase gene, POX9 (section 4.4.1, Okada et al., 1987). Comparison of the sequence of this ORF with that of human catalase (Quan et al., 1986), rat catalase (Furuta et al., 1986) and S. cerevisiae catalases T (CTT1; Hartig and Ruis, 1986) and A (CTA1; Cohen et al., 1988) revealed nucleic acid sequence identities of 52.3%, 53.5%, 48.0% and 61.0%, respectively, which was strongly suggestive of a Cat-encoding capacity for this composite cDNA. Moreover, the 1,643 bp of cDNA sequence [independent of the poly(A) tract] (Figure 4.2.2) was consistent with the 1.7 to 1.8 kb size of the Cat-encoding mRNA identified upon Northern blot analysis by Rachubinski et al. (1987), as well as the 1.85 kb predicted size of the Cat mRNA similarly identified by Okada et al.

(1987). Interestingly, comparison of the partial cDNA sequences encoding Cat from recombinants CdT-1 (sequenced herein) and RU1:87 (Rachubinski et al., 1987), which were initially perceived to be identical, revealed slightly less than perfect nucleotide sequence identity (Figure 4.2.2). Discrepancies were consistently found (upon resequencing and sequencing in the opposite direction; Figure 4.2.1, recombinants CdT-1 and C188-53) at positions 1129 (T versus G, respectively), 1164 (G versus A, respectively), 1167 (T versus A, respectively), 1168 (G versus A, respectively), 1182 (G versus A, respectively) and 1188 (T versus C, respectively). As yet, an explanation for these differences has not been found.

The 5'-noncoding region of the Cat cDNA (ie. the leader region; Cigan and Donahue, 1987) consisted of 43 nucleotides, which was consistent with the size of the primer C189 extension product(s) of poly(A)$^+$RNA (≈175 to 185 nucleotides, section 4.1.3, Figure 4.1.3), but was slightly smaller than the average size of 52 nucleotides determined for the leader region of several yeast cDNAs (Cigan and Donahue, 1987). This compared with 70 and 83 nucleotides of 5'-untranslated region found in the cDNAs encoding human kidney catalase (Bell et al., 1986) and rat liver catalase (Furuta et al., 1986), respectively, the sequences of which showed no homology with the 5'-untranslated region of the

Cat cDNA.

The A + T content of the 5'-untranslated region of the Cat cDNA (65.1%) was significantly higher than that of the coding region (52.8%), analogous to the situation determined for the 5'-flanking region of the AOx gene (section 3.2.1). This was, however, contrary to that determined for the leader regions of the human kidney (Bell et al., 1986) and rat liver catalase cDNAs (Furuta et al., 1986) which exhibit A + T contents of 36.1% and 37.1%, respectively. Compared to the general characteristics established for the leader regions of several yeast genes (Cigan and Donahue, 1987), the leader region of the Cat cDNA complied in that: (i) a nucleotide composition of A:44.2%, T:20.9%, G:0.0%, C:34.9% agreed with the observed richness in adenosine nucleotides, (ii) the nucleotide composition was such that it was void of significant secondary structure and (iii) the nucleotide composition dictated that the first AUG from the 5'-end of the Cat mRNA most likely served as the start codon for translation (see below).

The proposed ATG initiation codon (Figure 4.2.2) was the only possible choice from the obtained sequence. No other ATG codons (in- or out of frame) existed upstream of the putative start codon (nucleotides 1 to 3), while one upstream, in-frame TAA termination codon was found (nucleotides -12 to -10, Figure 4.2.2, highlighted by

'stop'). In addition, those sequences adjacent to this codon, 5'-CAAUAUGGCU-3' (Figure 4.2.2), concurred with the generalized, strong consensus sequence for translation initiation in yeast, 5'-$^{A}_{Y}A^{A}_{U}$AAUGUCU-3' (where Y = pyrimidine; Cigan and Donahue, 1987), except for the uridine and guanosine nucleotides at the -1 and +4 positions, respectively. These nucleotides occur with a prevalence of 13% and 24%, respectively, at these same positions within several yeast genes, relative to the 52% and 39% frequency of occurrence of the most prevalent nucleotides, A and U, respectively (Cigan and Donahue, 1987). Interestingly, within the highly expressed glycolytic genes of S. cerevisiae, the most prevalent residue at the +4 position, with a frequency of occurrence of 60%, is a guanosine nucleotide, similar to that seen for the Cat cDNA (Cigan and Donahue, 1987). Contrary to the situation observed for the AOx gene (section 3.2.1), and most other yeast genes (Cigan and Donahue, 1987), those sequences adjacent to the start codon of the Cat cDNA also reflected the higher eukaryotic consensus sequence for translation initiation, 5'-C$^{A}_{G}$CCAUGG-3' (Kozak, 1987), including the highly conserved adenosine nucleotide at the -3 position; however, the adenosine and uridine nucleotide preferences at the -2 and -1 positions, respectively, of the Cat cDNA, were noted exceptions. These nucleotides occur with a frequency of 27% and 9%,

respectively, at these same positions within several vertebrate mRNAs, relative to the 49% and 55% frequency of occurrence of the most prevalent nucleotide, a C, at both the -2 and -1 positions (Kozak, 1987). Finally, the presence of an alanine residue immediately C-terminal to the first AUG codon encountered from the 5'-end of the Cat mRNA (Figure 4.2.2) suggested further that its identity was that of the initiation codon since an identical amino acid exists at this position of rat liver catalase (Furuta et al., 1986), human kidney and lymphoblast catalase (Bell et al., 1986; Quan et al., 1986), bovine liver catalase (Schroeder et al., 1982) and Drosophila catalase (Orr et al., 1990) (Figure 4.4.5).

The 3'-noncoding region of the Cat cDNA consisted of 135 nucleotides (residues 1456 to 1590, Figure 4.2.2) terminated by a stretch of ten consecutive adenosine nucleotides, probably representing the poly(A) tract of the Cat mRNA. Typical of the comparatively short 3'-flanking regions of yeast genes (Henikoff and Cohen, 1984), this compared with ≈631 and 831 nucleotides of 3'-untranslated region found in the cDNAs encoding human kidney catalase (Bell et al., 1986) and rat liver catalase (Furuta et al., 1986), respectively. A possible polyadenylation sequence, AGTAAA (nucleotides 1559 to 1564, Figure 4.2.2, box), which is a functional derivative of the eukaryotic consensus

polyadenylation signal, AATAAA (Birnstiel et al., 1985), was located 27 nucleotides upstream of the poly(A) tract; this was within the appropriate distance of 10 to 30 nucleotides upstream of the poly(A) tract generally observed for eukaryotic mRNAs (Birnstiel et al., 1985). Similar to the situation observed for the genes encoding POX4 and 5 from C. tropicalis (Okazaki et al., 1986), two sequences proposed to be transcription termination signals in S. cerevisiae, the repeat TTTTTATA (Henikoff and Cohen, 1984) and the tripartite signal TAG...TA(T)GT...TTT (Zaret and Sherman, 1982), or derivatives thereof, were also found in the 3'-noncoding region of the Cat cDNA (Figure 4.2.2, double underlining). This finding reaffirms the postulated similarity of gene organization between C. tropicalis and S. cerevisiae discussed in sections 3.2.4 and 4.4.4.

To minimize repetition, a discussion of several other features of the Cat cDNA, including nucleotide frequencies, the deduced amino acid composition and secondary structure of Cat and codon usage, will ensue subsequent to the following section describing the molecular cloning of the C. tropicalis gene encoding Cat, which complements data presented here for the Cat cDNA.

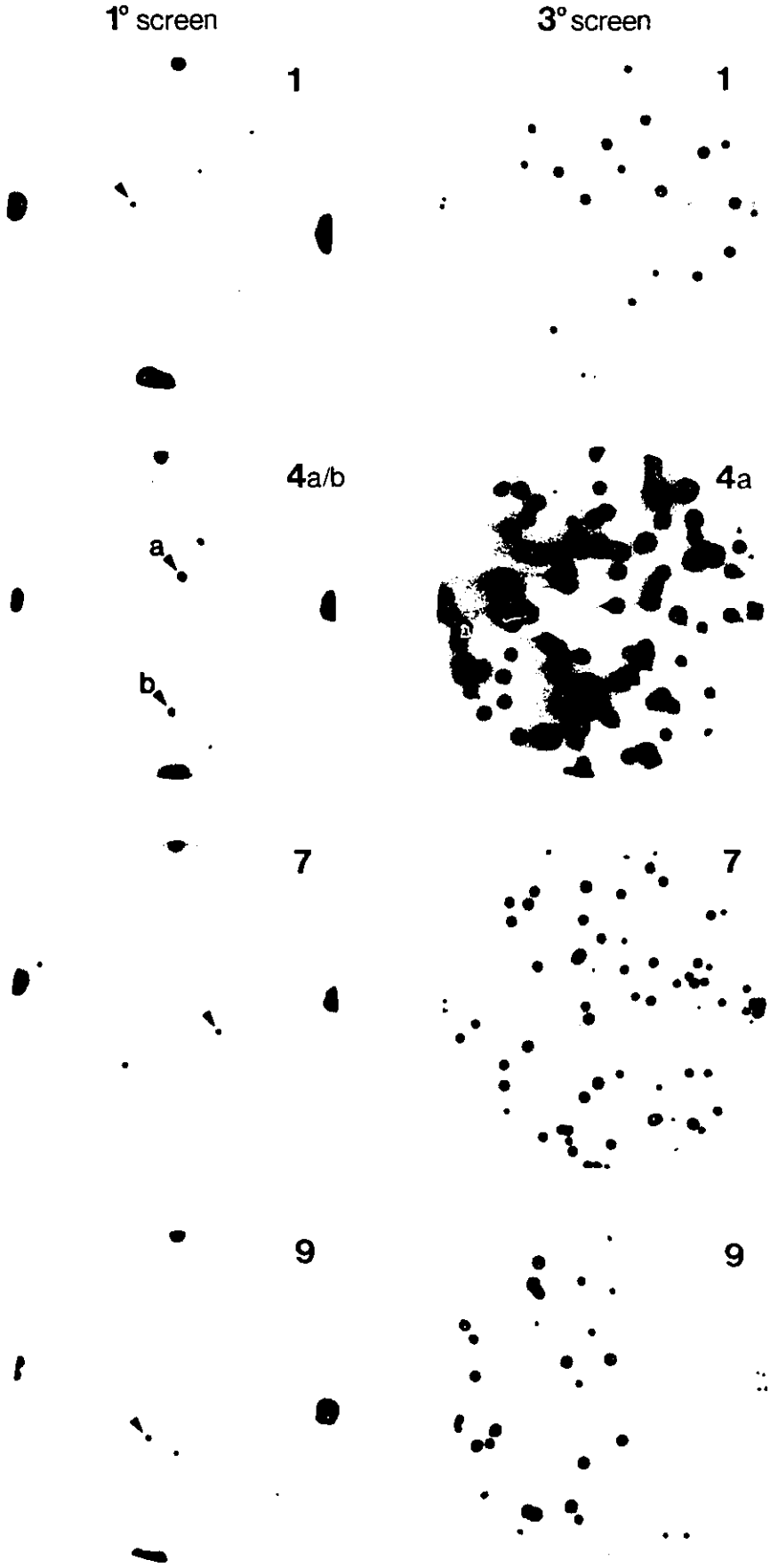
4.3 MOLECULAR CLONING OF THE GENE ENCODING Cat

4.3.1 Selection and Characterization of Genomic Clones Encoding Cat

The strategy employed to select genomic clones encoding Cat entailed the construction and screening of a second library of genomic DNA isolated from n-alkane grown C. tropicalis, using the ≈1.1 kbp cDNA insert of recombinant C188-53 (section 4.1.2) as the hybridization probe (henceforth called the Cat probe). Construction of this library involved the ligation of size-fractionated (≈20 kbp), Sau3A I-digested genomic DNA fragments into phosphatased, BamH I, EcoR I-digested lambda EMBL3 arms (Figure 4.3.3, upper portion), followed by packaging of the ligated DNA and transfection of the restrictive host, E. coli NM539 (section 2.5.2). Of the ≈10,000 plaques screened, between 20 and 30 potential positives were identified (not shown); this was consistent with an estimated number of 26 calculated for a C. tropicalis diploid genome of ≈15,000 kbp (Kamiryo and Okazaki, 1984) and ≈10,000 plaques screened, each of which contained an ≈20 kbp genomic fragment. Of these, 5 plaques (ie. areas containing plaques) which exhibited strong hybridization signals upon primary screening with random primer labeled Cat probe, were processed and subsequent to tertiary screening, purified to homogeneity, which allowed positive identification of Cat genomic clones 1, 4a, 4b, 7 and 9 (Figure 4.3.1, arrowheads).

Prior to phage DNA isolation from these clones,

Figure 4.3.1. Detection and purification of clones harbouring the Cat gene from the lambda EMBL3 library of C. tropicalis genomic DNA. For the primary (1°) screen, plaque lifts onto nitrocellulose (≈1000 plaques/dish) were processed as described in section 2.5.2 and hybridized (42°C for ≈16 h) with 2.2 x 10^6 cpm of random primer labeled cDNA insert from Cat recombinant C188-53 (specific activity = 2.2 x 10^7 cpm/µg). Conditions for the tertiary (3°) screen were identical to those of the 1° screen except that plaque lifts were conducted at a density of <100 plaques/dish. Arrowheads indicate those hybridization positives, including Cat genomic clones 1, 4a, 4b, 7 and 9, identified upon 1° screening which were picked, processed and subsequent to 3° screening, purified to homogeneity. Washes were conducted at 55°C in 1 x SSC. Exposure was for 16 h at -70°C with Kodak XAR-5 film and 1 intensifying screen.


1° screen
3° screen
1
1
4a/b
4a
a
b
7
7
9
9

Southern blot analysis of C. tropicalis genomic DNA was performed in an attempt to identify a smaller than 20 kbp, single DNA fragment harbouring the Cat gene which would be suitable to subcloning for the purpose of sequencing. Since the composite cDNA encoding Cat contained two EcoR I cleavage sites (section 4.1.4), digestion of C. tropicalis genomic DNA with EcoR I was included as a control for the hybridization of the Cat probe with the Cat gene. As seen in Figure 4.3.2 (Panel A, lane 3), hybridization of the Cat probe with EcoR I-digested C. tropicalis genomic DNA revealed the presence of two distinct bands, the expected ≈300 bp EcoR I fragment found internal to the Cat cDNA/gene (nucleotides 812 to 1175 of the Cat cDNA, Figure 4.2.2) and the ≈5 kbp EcoR I fragment representative of the 5'-end of the Cat gene. Hybridization of the Cat probe to the 3'-end of the Cat gene was not detected, probably due to the minimal overlap (<70 bp) between the probe and target sequence. Xba I-digestion of C. tropicalis genomic DNA was inappropriate to subcloning the Cat gene since the Cat probe hybridized to a single restriction fragment significantly greater than 7 kbp in size (Figure 4.3.2, Panel A, lane 5). Similarly, Pst I-digestion of C. tropicalis genomic DNA was excluded because the probe hybridized to two distinct fragments of ≈3.5 kbp and much greater than 7 kbp (Figure 4.3.2, Panel A, lane 4). Those restriction endonuclease

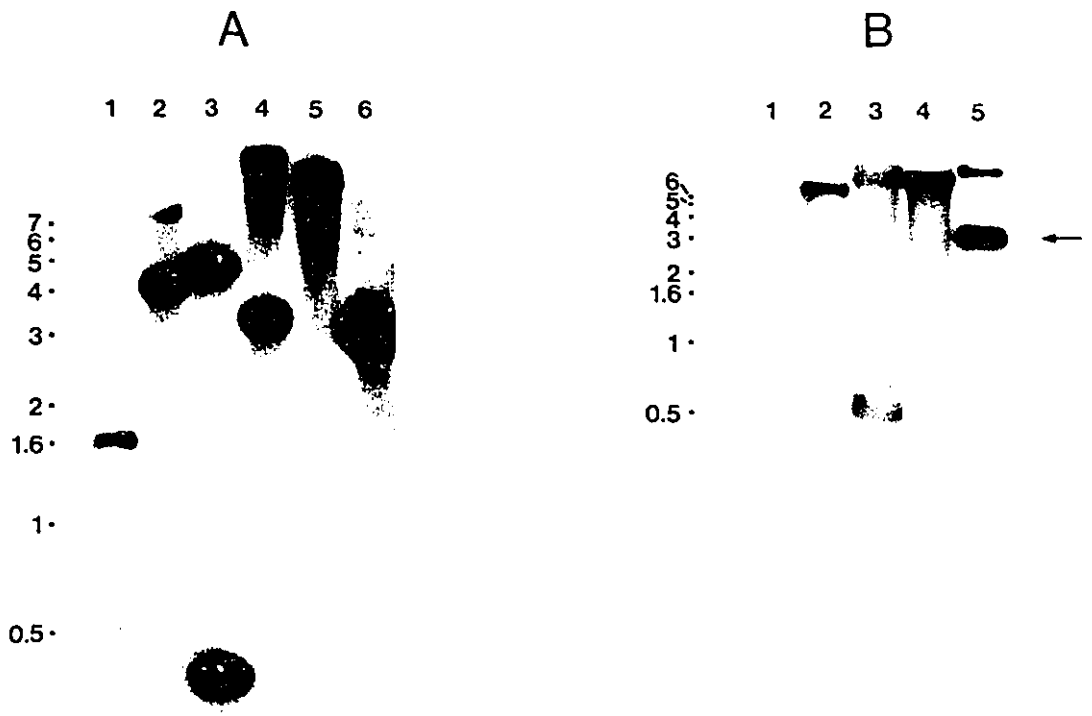


Figure 4.3.2. Southern blot hybridization of the Cat cDNA probe with C. tropicalis genomic DNA (Panel A) and lambda EMBL3 phage DNA isolated from the putative Cat genomic clone, 4a (Panel B). Hybridization (55° for 16 h) was with 1.5 x 10^6 cpm of random primer labeled cDNA insert from Cat recombinant C188-53 (specific activity = 1.1 x 10^8 cpm/µg). Washes were conducted at 52°C in 1 x SSC. Panel A, 5 µg of C. tropicalis genomic DNA/lane was digested with the following restriction endonucleases: Lanes: 1, molecular weight markers - 1 kbp DNA ladder; 2, BamH I; 3, EcoR I; 4, Pst I; 5, Xba I; 6, BamH I and Pst I. Exposure was for 20 h at -70°C with Kodak XAR-5 film and 1 intensifying screen. Marker sizes in kbp (1 kbp DNA ladder) are indicated to the left. The fortuitous "hybridization" of the probe to the ≈1.6 kbp molecular weight marker (in Panel A only) is unexplained. Panel B, ≈100 ng/lane of phage DNA isolated from Cat genomic clone 4a was digested with the following restriction endonucleases: Lanes: 1, molecular weight markers - 1 kbp DNA ladder; 2, Cla I; 3, Kpn I; 4, Pst I; 5, BamH I and Sal I. Exposure was for 20 h at room temperature with Kodak XK-1 film. The arrow indicates the position of the ≈3 kbp BamH I, Sal I fragment of C. tropicalis genomic DNA which harbours the Cat gene (see text, section 4.3.1). Marker sizes in kbp (1 kbp DNA ladder) are indicated to the left.

digestions of C. tropicalis genomic DNA which generated suitable fragments for subcloning the Cat gene included: 1) BamH I alone, which yielded a single, ≈4 kbp fragment complementary to the Cat probe [Figure 4.3.2, Panel A, lane 2; the lower intensity "band" observed at >7 kbp was presumed to be an incomplete digestion product since BamH I, Pst I-digestion (Panel A, lane 6) generated a single, smaller fragment only] and 2) combined BamH I, Pst I-digestion, which generated a single, ≈3 kbp fragment complementary to the Cat probe (Figure 4.3.2, Panel A, lane 6). Unfortunately, twenty-eight Pst I cleavage sites exist within the lambda EMBL3 sequence (Ausubel et al., 1988). Thus, the simplest route to subcloning the Cat gene was the isolation of an ≈4 kbp BamH I restriction fragment from the previously identified Cat genomic clones, since no vector-related fragments would be liberated by BamH I-digestion (other than the intact EMBL3 arms), contrary to the multiple vector-related fragments which would be generated upon BamH I, Pst I-digestion. However, subsequent to large scale phage DNA isolations (section 2.5.6.1) from Cat genomic clones 4a, 4b and 9, digestion of phage 4a DNA (not shown) revealed the existence of only a single BamH I site (ie. no ≈4 kbp BamH I fragment was liberated). As an explanation, it was reasoned that ligation of the ≈20 kbp Sau3A I fragment into BamH I, EcoR I-digested lambda EMBL3 arms did

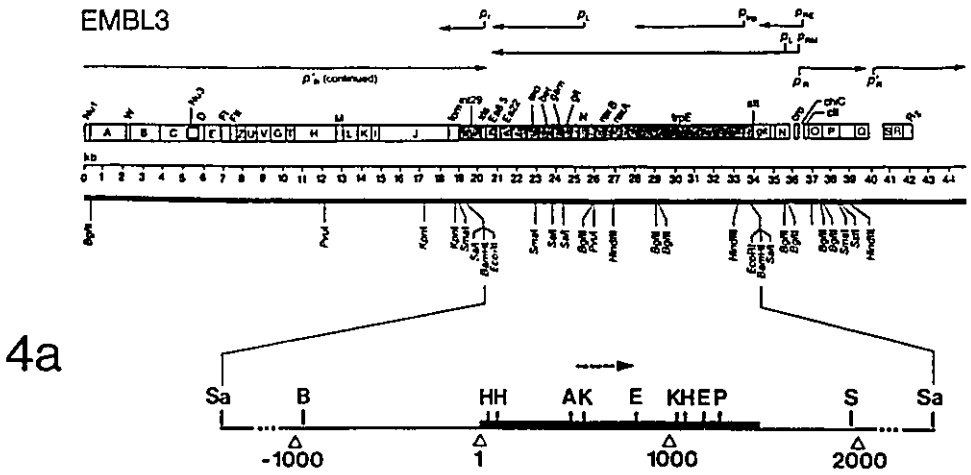
not regenerate either BamH I site. Further, the single BamH I site recognized was one of two flanking the Cat gene, the other of which was excluded upon subcloning. Realizing this possibility, BamH I, Sal I-digestion of phage 4a DNA was used to liberate the Cat gene since no Sal I site existed within the Cat cDNA sequence (Figure 4.2.2) and the Sal I sites which flanked each EMBL3 MCS (Figure 4.3.3, upper portion) guaranteed fragment excision independent of the sequences which flanked the Cat gene (although excision by either of two mechanisms, BamH I, Sal I-digestion or Sal I-digestion alone, was possible). The accompanying Southern blot of phage DNA isolated from Cat genomic clone 4a is shown in Figure 4.3.2 (Panel B). For this blot, Kpn I-digestion of phage 4a DNA [an alternative to EcoR I-digestion (Panel A)] served as the control for hybridization of the Cat probe with the Cat gene since the composite cDNA encoding Cat contained two Kpn I cleavage sites (nucleotides 544 to 549 and 1033 to 1038, Figure 4.2.2). As seen in the figure (Panel B, lane 3), hybridization of the Cat probe with Kpn I-digested phage DNA revealed the expected ≈500 bp Kpn I fragment found internal to the Cat cDNA/gene and, most likely, two, nonresolvable (in this system), >6 kbp fragments which consisted of the 5'- and 3'-ends of the Cat gene ligated to one or the other EMBL3 arm. Although Cla I- and Pst I-digestions of phage 4a DNA were included, they

proved useless since hybridization to the Cat probe was exhibited by >6 kbp fragment(s) only (Panel B, lanes 2 and 4, respectively). BamH I, Sal I-digestion of phage DNA from clone 4a liberated a DNA fragment which hybridized with the Cat probe (as expected) of a suitable size for subcloning, ≈3 kbp (Panel B, lane 5, arrow). Although not included in the Southern blot of Figure 4.3.2 (Panel B), neither BamH I- nor Sal I-digestion of phage 4a DNA generated the ≈3 kbp restriction fragment observed upon combined BamH I, Sal I-digestion, thereby establishing that the ≈3 kbp DNA fragment harbouring the Cat gene was BamH I, Sal I-flanked.

4.3.2 Strategy Used to Subclone the Cat Gene and Generate Exonuclease III-Deleted Templates for Sequencing

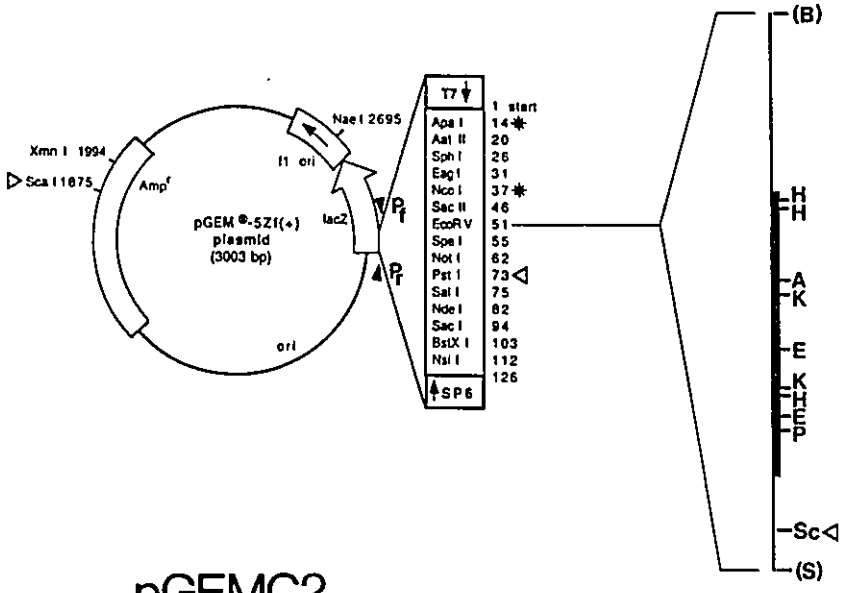
Having established that the ≈20 kbp Sau3A I fragment of phage DNA from clone 4a contained the Cat gene as an ≈3 kbp, BamH I, Sal I-excisable fragment (Figure 4.3.3, upper portion), subcloning entailed fragment isolation from nondenaturing polyacrylamide gels, conversion to blunt ends using DNA polymerase I Klenow fragment and ligation into the MCS (EcoR V site) of the *in vitro* expression vector, pGEM-5Zf(+) (Figure 4.3.3, lower portion). Subclones were obtained for both possible orientations; pGEMC2 contained the BamH I, Sal I fragment in that orientation which directed transcription of Cat mRNA from the T7 RNA polymerase promoter (illustrated in Figure 4.3.3, lower

Figure 4.3.3. Schematic representation of phage DNA from Cat genomic clone, 4a, and the recombinants used to sequence the Cat gene, pGEMC1 and pGEMC2. The upper portion of the figure represents: 1) construction of a second C. tropicalis genomic DNA library by ligation of size-fractionated, Sau3A I-digested DNA fragments into phosphatased, BamH I, EcoR I-digested lambda EMBL3 arms (Frischauf et al., 1983) and 2) phage DNA from genomic clone 4a. The partial restriction endonuclease cleavage map established for the ≈3 kbp BamH I, Sal I fragment of C. tropicalis genomic DNA harbouring the Cat gene is illustrated, with the ORF of the Cat gene indicated by the thick line. Restriction endonuclease sites included are: A, Aha II; B, BamH I; E, EcoR I; H, Hinc II; K, Kpn I; P, Pvu II; S, Sal I; Sa, Sau3A I; Sc, Sca I [bracketed sites, (B) and (S), were converted to blunt ends]. Deviations from the size scale of the genomic DNA fragment (•••) are indicated. The dashed arrow represents the direction of Cat gene polarity. The orientation of the Sau3A I genomic DNA fragment within the EMBL3 stuffer segment was not determined. The lower portion of the figure represents the strategy used to subclone the BamH I, Sal I fragment of genomic DNA into the MCS of pGEM-5Zf(+), in both orientations, for the purpose of sequencing (section 4.3.2). The orientation of the blunt-ended BamH I, Sal I fragment within clone pGEMC2 is illustrated; pGEMC1 (not shown) contains this fragment in the opposite orientation within the EcoR V site of the MCS. Those restriction endonuclease sites used to double-digest plasmids pGEMC1 and C2 prior to unidirectional digestion with exonuclease III (section 4.3.2) are denoted by asterisks. Plasmid digestion at restriction endonuclease sites denoted by was used to screen for overlapping, exonuclease III-deleted, sequencing templates (section 4.3.2). The arrowheads shown at the 5'- and 3'-ends of the MCS represent the direction of sequencing from the M13 universal, forward (-20 or -40) primers (P_f) and reverse primer (P_r), respectively. Physical maps of lambda EMBL3 and pGEM-5Zf(+) were reproduced from Sambrook et al. (1989) and the Promega 1989/90 catalogue, respectively.


EMBL3

4a

Sa
B
HH
AK
E
KHEP
S
Sa
-1000
1
1000
2000


Nae I 2695
Xmn I 1994
Sca I 1675
Amp^r
f1 ori
pGEM®-5Zf(+)
plasmid
(3003 bp)
lacZ
ori
T7
SP6
1 start
Apa I 14
Aat II 20
Sph I 26
Eag I 31
Nco I 37
Sac II 46
EcoR V 51
Spe I 55
Not I 62
Pst I 73
Sal I 75
Nde I 82
Sac I 94
BstX I 103
Nsi I 112
126
(B)
H
H
A
K
E
K
H
E
P
Sc
(S)

pGEMC2

portion) whereas the complementary subclone, pGEMC1, contained the BamH I, Sal I fragment in the opposite orientation, thereby directing Cat mRNA transcription from the SP6 RNA polymerase promoter. For sequencing purposes, exonuclease III deletion of the BamH I, Sal I fragment of both pGEMC1 and C2 was directed by Apa I, Nco I-digestion of the pGEM-5Zf MCS (neither Apa I nor Nco I cleaved the Cat genomic fragment). This created a 3', exonuclease III-resistant overhang (Apa I) adjacent to the vector DNA, which maintained the integrity of the M13 universal, forward primer (P_f) binding site, and a 5', exonuclease III-sensitive overhang (Nco I) adjacent to the BamH I, Sal I fragment, which dictated unidirectional digestion through the entire Cat gene and flanking regions (Figure 4.3.3, lower portion). For pGEMC1 and C2 deletions, visual inspection of ethidium bromide-stained agarose gels was used to monitor both the extent and rate of exonuclease III-digestion of the BamH I, Sal I genomic DNA fragment. Using those conditions described in section 2.5.6.2, exonuclease III-digestion of pGEMC1 and C2 progressed at a rate of 150 to 200 nucleotides/min. Subsequent to transformation of competent E. coli MV1193 cells with these exonuclease III-deleted constructs, overlapping sequencing templates were identified by Sca I, Pst I-digestion of isolated plasmid DNA (Figure 4.3.3, lower portion). Plasmids estimated to be

progressively smaller by increments of ≈250 bp were retained as the nested sets of deletion mutants required for sequencing both DNA strands of the Cat gene and flanking regions.

4.4 SEQUENCE DETERMINATION AND ANALYSIS OF THE Cat CODING REGION

4.4.1 Nucleotide Sequence of the Gene Encoding Cat

The physical map and sequencing strategy of the Cat gene and its 5'- and 3'-flanking sequences (ie. the entire BamH I, Sal I genomic DNA fragment initially isolated from phage 4a DNA) are presented in Figure 4.4.1. Having established nested sets of pGEMC1 and C2 deletion mutants (section 4.3.2), double-stranded plasmid isolated from each of these constructs was sequenced essentially as described by Zhang et al. (1988) (section 2.5.6.3) using the ddNTP chain termination method (Sanger et al., 1977). For the most part, the labeled sequence determinations of Figure 4.4.1 represent pGEMC1 (eg. 1-1a) or pGEMC2 (eg. 2-1a) exonuclease III deletion constructs which were sequenced using the M13 universal, forward (-20 or -40) sequencing primers. Those exceptions included the sequencing of pGEMC2 constructs 2-11bE$_D$, 2-11bK$_D$ and 2-16bAH (see section 5.1), which was accomplished using the M13 reverse sequencing primer (P$_r$, Figure 4.3.3), and the 3' to 5' sequence determination used to confirm the sequence at the

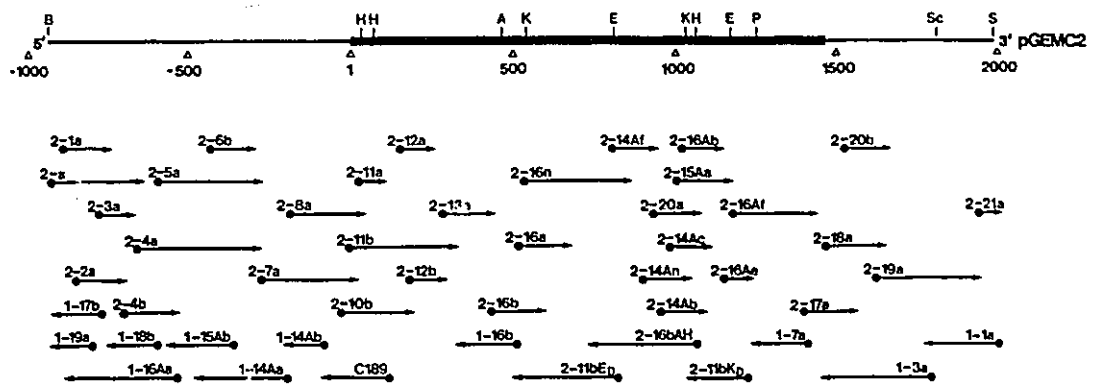
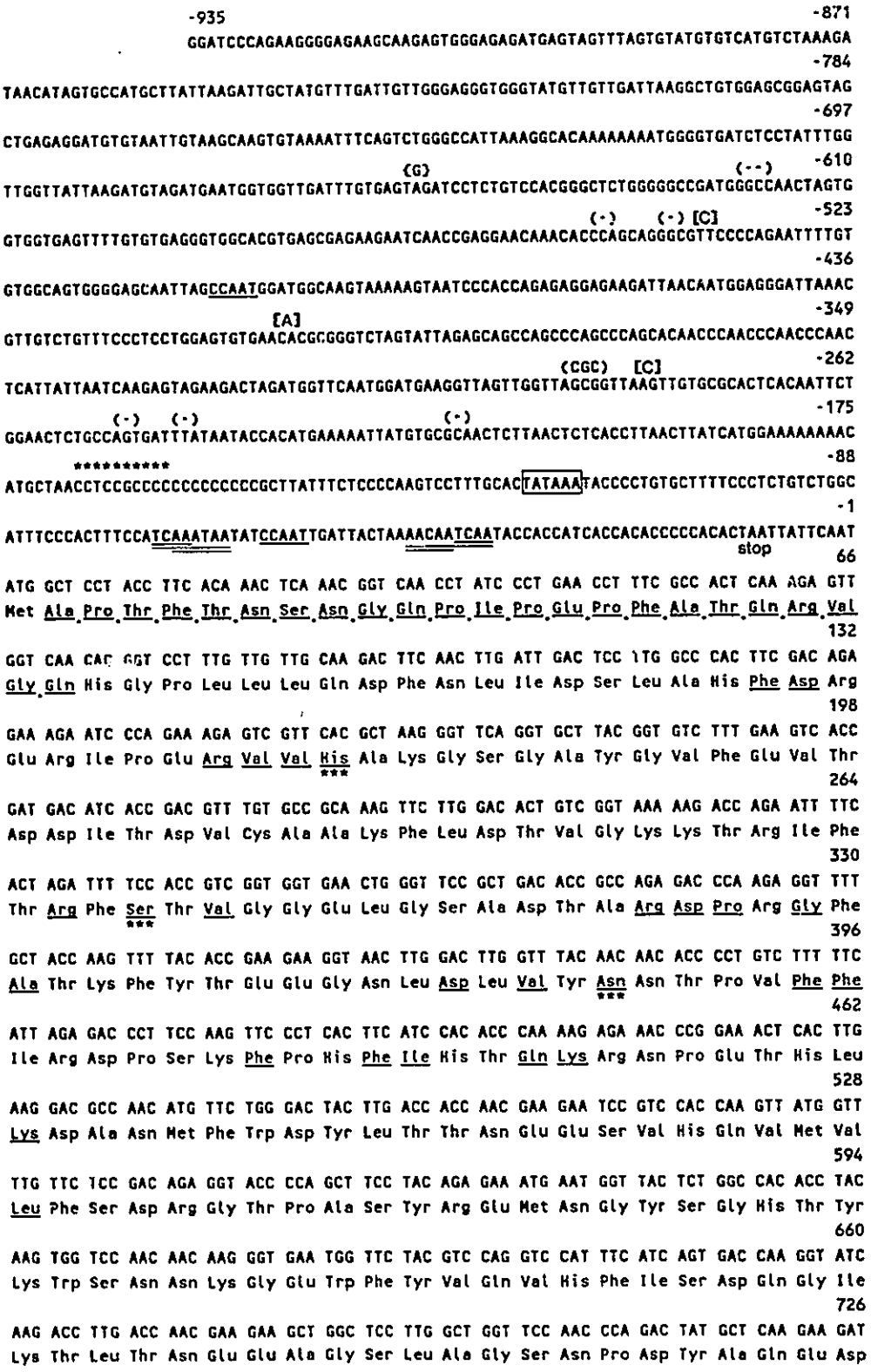


Figure 4.4.1. Physical map and sequencing strategy of the Cat gene and 5′- and 3′-flanking regions. Restriction endonuclease sites included are: A, Aha II; B, BamH I; E, EcoR I; H, Hinc II; K, Kpn I; P, Pvu II; S, Sal I; Sc, Sca I. The direction of Cat gene polarity is shown; the accompanying base pair numbering of the ⅓3 kbp BamH I, Sal I fragment of C. tropicalis genomic DNA is presented in Figure 4.4.2. The protein coding region of clone pGEMC2 is indicated by the thick line. Unidirectional digestion of Apa I, Nco I-digested pGEMC1 and C2 with exonuclease III was used to create the nested sets of deletion mutants required for sequencing (section 4.3.2). The start points (closed circles), direction and extent of double-stranded dideoxy sequence determinations are indicated by continuous arrows. Numbers above the sequencing start points denote the different pGEMC1 (eg. 1-1a) or pGEMC2 (eg. 2-1a) deletion mutants or constructs sequenced in each case. Sequence data were generated using M13 universal, forward (-20 and -40) and reverse primers, as appropriate (section 4.4.1); only one Cat(C) specific primer (C189) was used to confirm the sequence of the 5′-end of the Cat ORF (see Table 2.1.1 for the sequences and annealing positions of these primers).

5'-end of the Cat ORF [labeled C189 (Figure 4.4.1)], for which the Cat(C) specific primer, C189, was used (see Table 2.1.1 for the primer annealing position and sequence). Collectively, the entire sequence depicted by the physical map of Figure 4.4.1 was determined in the 5' to 3' (coding) direction, while 92% of this sequence was determined in the 3' to 5' (noncoding) direction, with each nucleotide sequenced at least twice. Those nucleotides which were sequenced in the 5' to 3' direction only were determined twice and compared to both the Cat cDNA sequence (section 4.2.2, Figure 4.2.2) and the POX9 DNA sequence (Okada et al., 1987) for confirmation. Any discrepancies encountered were resolved by resequencing, at least once, using an independent construct (if possible).

Translation of the 2,921 nucleotides which constituted the BamH I, Sal I fragment of C. tropicalis genomic DNA, in the three possible reading frames, revealed the existence of a single ORF of 1,455 nucleotides, without intervening sequences, which was flanked by 938 and 528 nucleotides of 5'- and 3'-noncoding sequence, respectively (Figure 4.4.2). For reasons similar to those cited for the Cat cDNA (section 4.2.2), including the presence of an upstream, in-frame TAA termination codon (nucleotides -12 to -10, Figure 4.2.2, highlighted by 'stop') and sequences adjacent to an AUG codon, 5'-CAAUAUGGCU-3' (nucleotides -4

Figure 4.4.2. Nucleotide sequence of the Cat gene and flanking regions and the deduced amino acid sequence of Cat. Nucleotides are numbered in the 5' to 3' direction beginning with the first residue of the ATG start codon. Portions of the 5'- and 3'-flanking regions are shown; nucleotides 5' to position 1 are designated by negative numbers. Symbols above the Cat sequence represent differences found in the sequence reported by Okada et al. (1987): (G), nucleotide substitution creating the BamH I site polymorphism; (-), nucleotide is absent; [N], nucleotide is inserted; (N), nucleotide is substituted (see section 4.4.1). Within the 5'-noncoding region, an in-frame TAA stop codon is highlighted by 'stop', 'CCAAT box' sequences are underlined, sequences for transcription initiation are double-underlined, a putative 'TATA box' sequence is enclosed by a box and a 'GC box' derivative is highlighted by *•••*. Within the 3'-noncoding region, a putative polyadenylation signal is boxed, possible transcription termination signals are underlined and an arrow denotes the approximate beginning of the poly(A) tract of the Cat mRNA (section 4.4.1). The deduced amino acid sequence of Cat is presented below the nucleotide sequence in the standard three-letter code. Those NH_2-terminal amino acids of Cat which were sequenced by Ueda et al. (1987) and Okada et al. (1987) are dash-dot underlined. Underlined amino acids represent those residues of bovine liver catalase (Schroeder et al., 1982), in contact with the heme moiety, which are conserved in Cat. Asterisks (***) indicate those amino acids of bovine catalase, participating directly in catalysis, which are conserved in Cat. The boxed tripeptide denotes a variant of the PTS identified by Gould et al. (1989).


-935
-871
GGATCCCAGAAGGGGAGAAGCAAGAGTGGGAGAGATGAGTAGTTTAGTGTATGTGTCATGTCTAAAGA
-784
TAACATAGTGCCATGCTTATTAAGATTGCTATGTTTGATTGTTGGGAGGGTGGGTATGTTGTTGATTAAGGCTGTGGAGCGGAGTAG
-697
CTGAGAGGATGTGTAATTGTAAGCAAGTGTAAAATTTCAGTCTGGGCCATTAAAGGCACAAAAAAAATGGGGTGATCTCCTATTTGG
-610
(G)
(--)
TTGGTTATTAAGATGTAGATGAATGGTGGTTGATTTGTGAGTAGATCCTCTGTCCACGGGCTCTGGGGGCCGATGGGCCAACTAGTG
-523
(-) (-) [C]
GTGGTGAGTTTTGTGTGAGGGTGGCACGTGAGCGAGAAGAATCAACCGAGGAACAAACACCCAGCAGGGCGTTCCCCAGAATTTTGT
-436
GTGGCAGTGGGGAGCAATTAGCCAATGGATGGCAAGTAAAAAGTAATCCCACCAGAGAGGAGAAGATTAACAATGGAGGGATTAAAC
-349
[A]
GTTGTCTGTTTCCCTCCTGGAGTGTGAACACGCGGGTCTAGTATTAGAGCAGCCAGCCCAGCCCAGCACAACCCAACCCAACCCAAC
-262
(CGC) [C]
TCATTATTAATCAAGAGTAGAAGACTAGATGGTTCAATGGATGAAGGTTAGTTGGTTAGCGGTTAAGTTGTGCGCACTCACAATTCT
-175
(-) (-) (-)
GGAACTCTGCCAGTGATTTATAATACCACATGAAAAATTATGTGCGCAACTCTTAACTCTCACCTTAACTTATCATGGAAAAAAAAC
-88
****
ATGCTAACCTCCGCCCCCCCCCCCCCGCTTATTTCTCCCCAAGTCCTTTGCACTATAAATACCCCTGTGCTTTTCCCTCTGTCTGGC
-1
ATTTCCCACTTTCCATCAAATAATATCCAATTGATTACTAAAACAATCAATACCACCATCACCACACCCCCACACTAATTATTCAAT
stop
66
ATG GCT CCT ACC TTC ACA AAC TCA AAC GGT CAA CCT ATC CCT GAA CCT TTC GCC ACT CAA AGA GTT
Met Ala Pro Thr Phe Thr Asn Ser Asn Gly Gln Pro Ile Pro Glu Pro Phe Ala Thr Gln Arg Val
132
GGT CAA CAC GGT CCT TTG TTG TTG CAA GAC TTC AAC TTG ATT GAC TCC TTG GCC CAC TTC GAC AGA
Gly Gln His Gly Pro Leu Leu Leu Gln Asp Phe Asn Leu Ile Asp Ser Leu Ala His Phe Asp Arg
198
GAA AGA ATC CCA GAA AGA GTC GTT CAC GCT AAG GGT TCA GGT GCT TAC GGT GTC TTT GAA GTC ACC
Glu Arg Ile Pro Glu Arg Val Val His Ala Lys Gly Ser Gly Ala Tyr Gly Val Phe Glu Val Thr
***
264
GAT GAC ATC ACC GAC GTT TGT GCC GCA AAG TTC TTG GAC ACT GTC GGT AAA AAG ACC AGA ATT TTC
Asp Asp Ile Thr Asp Val Cys Ala Ala Lys Phe Leu Asp Thr Val Gly Lys Lys Thr Arg Ile Phe
330
ACT AGA TTT TCC ACC GTC GGT GGT GAA CTG GGT TCC GCT GAC ACC GCC AGA GAC CCA AGA GGT TTT
Thr Arg Phe Ser Thr Val Gly Gly Glu Leu Gly Ser Ala Asp Thr Ala Arg Asp Pro Arg Gly Phe
***
396
GCT ACC AAG TTT TAC ACC GAA GAA GGT AAC TTG GAC TTG GTT TAC AAC AAC ACC CCT GTC TTT TTC
Ala Thr Lys Phe Tyr Thr Glu Glu Gly Asn Leu Asp Leu Val Tyr Asn Asn Thr Pro Val Phe Phe
***
462
ATT AGA GAC CCT TCC AAG TTC CCT CAC TTC ATC CAC ACC CAA AAG AGA AAC CCG GAA ACT CAC TTG
Ile Arg Asp Pro Ser Lys Phe Pro His Phe Ile His Thr Gln Lys Arg Asn Pro Glu Thr His Leu
528
AAG GAC GCC AAC ATG TTC TGG GAC TAC TTG ACC ACC AAC GAA GAA TCC GTC CAC CAA GTT ATG GTT
Lys Asp Ala Asn Met Phe Trp Asp Tyr Leu Thr Thr Asn Glu Glu Ser Val His Gln Val Met Val
594
TTG TTC TCC GAC AGA GGT ACC CCA GCT TCC TAC AGA GAA ATG AAT GGT TAC TCT GGC CAC ACC TAC
Leu Phe Ser Asp Arg Gly Thr Pro Ala Ser Tyr Arg Glu Met Asn Gly Tyr Ser Gly His Thr Tyr
660
AAG TGG TCC AAC AAC AAG GGT GAA TGG TTC TAC GTC CAG GTC CAT TTC ATC AGT GAC CAA GGT ATC
Lys Trp Ser Asn Asn Lys Gly Glu Trp Phe Tyr Val Gln Val His Phe Ile Ser Asp Gln Gly Ile
726
AAG ACC TTG ACC AAC GAA GAA GCT GGC TCC TTG GCT GGT TCC AAC CCA GAC TAT GCT CAA GAA GAT
Lys Thr Leu Thr Asn Glu Glu Ala Gly Ser Leu Ala Gly Ser Asn Pro Asp Tyr Ala Gln Glu Asp

792
TTG TTC AAA AAC ATT GCT GCT GGT AAC TAC CCA TCC TGG ACC TGT TAC ATC CAA ACC ATG ACT GAA
Leu Phe Lys Asn Ile Ala Ala Gly Asn Tyr Pro Ser Trp Thr Cys Tyr Ile Gln Thr Met Thr Glu
858
GCT CAA GCT AAA GAA GCT GAA TTC TCC GTC TTT GAC TTG ACC AAG GTC TGG CCA CAC GGA AAG TAC
Ala Gln Ala Lys Glu Ala Glu Phe Ser Val Phe Asp Leu Thr Lys Val Trp Pro His Gly Lys Tyr
924
CCA ATG AGA AGA TTT GGT AAG TTT ACT TTG AAC GAA AAC CCA AAG AAC TAC TTT GCC GAA GTT GAA
Pro Met Arg Arg Phe Gly Lys Phe Thr Leu Asn Glu Asn Pro Lys Asn Tyr Phe Ala Glu Val Glu
990
CAA GCT GCT TTC TCC CCA GCC CAC ACC GTT CCA CAC ATG GAA CCA TCT GCT GAC CCA GTC TTG CAA
Gln Ala Ala Phe Ser Pro Ala His Thr Val Pro His Met Glu Pro Ser Ala Asp Pro Val Leu Gln
1056
TCC AGA TTG TTC TCT TAT GCT GAT ACT CAC AGA CAC AGA TTG GGT ACC AAC TAC ACC CAA ATC CCA
Ser Arg Leu Phe Ser Tyr Ala Asp Thr His Arg His Arg Leu Gly Thr Asn Tyr Thr Gln Ile Pro
1122
GTC AAC TGC CCA GTC ACT GGT GCT GTT TTC AAC CCA CAC ATG AGA GAC GGT GCC ATG AAT GTC AAT
Val Asn Cys Pro Val Thr Gly Ala Val Phe Asn Pro His Met Arg Asp Gly Ala Met Asn Val Asn
1188
GGT AAC TTG GGC AAC CAT CCA AAC TAC TTG GCC TCT GAC AAA CCA ATT GAA TTC AAG CAA TTC TCC
Gly Asn Leu Gly Asn His Pro Asn Tyr Leu Ala Ser Asp Lys Pro Ile Glu Phe Lys Gln Phe Ser
1254
CTC CAA GAA GAC CAA GAA GTC TGG CAC GGT GCT GCT ACT CCA TTC CAC TGG AAG GCT ACC CCA GCT
Leu Gln Glu Asp Gln Glu Val Trp His Gly Ala Ala Thr Pro Phe His Trp Lys Ala Thr Pro Ala
1320
GAT TTC AAG CAA GCC ACT GAA TTG TGG AAG GTG TTG AAG AAA TAT CCA AAC CAA CAA GAA CAC TTG
Asp Phe Lys Gln Ala Thr Glu Leu Trp Lys Val Leu Lys Lys Tyr Pro Asn Gln Gln Glu His Leu
1386
GCC CAC AAT GTT GCT GTC CAT GCC TCT GCT GCC GAT GCT CCA ATC CAA GAC AGA GTC ATT GCT TAC
Ala His Asn Val Ala Val His Ala Ser Ala Ala Asp Ala Pro Ile Gln Asp Arg Val Ile Ala Tyr
1452
TTT ACC AAG GTC CAC CCA GAC TTG GGA GAC CTC ATC AAG AAG GAA ATC TTG GAA TTG TCT CCA AGA
Phe Thr Lys Val His Pro Asp Leu Gly Asp Leu Ile Lys Lys Glu Ile Leu Glu Leu Ser Pro Arg
1537
AAG TAA AGAAGATTCTGATTTGATATTGGAGTGCTAAAGGTCTTAGTTATGTTTGTTTCTTTTCTTTTTTCTTATTAATATTTAT
Lys End
1624
AGCTGTTGTAATGCGTTTTACAGTAAATTGAATGGTTATATTTGATATATCTGACAGTGGAAGTGTAGAAGCAATGGTGGTGGTTGT
1711
GGTTATTAAATGCTACCCGTTGTTCGGGATGAAATTCTCATTCATTGGCGCGCAAAGCAAAAATAACAGAAGATGGGCGAGGAGTTA
1798
[A] [G]
AAAAAAAAAAAAAAAGATTTAGCTCAAAAATAAAAAAAAAATTTTTCTCACTTACTCACAGTACTGAAAATATTGAAAACCATCATCA
1885
TCATCACCAACAACAACAACAACAACAACAACCAAAACCGCAGCCATCATAGCACCACGATACCTATAACATAACCCACACAGCTAT
1972
CCATCTGTACCAGTCACTAACAAAACAACAGAGACAGAGAGAGAAGGAAAGAGAAAGAGAAAGACATCTTGGAATTTACAAATCATG
1983
ACAGCGTCGAC

to 6, Figure 4.4.2), which concurred with the generalized, strong consensus sequences for translation initiation in yeast (Cigan and Donahue, 1987) and higher eukaryotes (Kozak, 1987), the proposed ATG initiation codon (Figure 4.4.2) was the only possible choice from the obtained sequence. The length of the corresponding ORF, which stretched from the putative ATG initiation codon to a TAA termination codon (nucleotides 1456 to 1458, Figure 4.4.2), was identical to that determined for both the Cat cDNA (section 4.2.2) and C. tropicalis POX9, which was similarly shown to be devoid of any intervening sequences (Okada et al., 1987). Comparison of the sequence of the ORF with that of C. tropicalis POX9 revealed perfect nucleotide sequence identity; however, a similar comparison with the ORF of the Cat cDNA revealed thirty-one mismatches, hypothesized to be the result of allelic heterogeneity within the diploid genome of C. tropicalis (see below), which yielded a nucleotide sequence identity of only 97.9% (Figure 4.4.3). Of the thirty-one mismatches identified, two occurred at the +1 position of the corresponding codons [nucleotides 292 and 1168 (Figure 4.4.3)], one of which resulted in a conservative ile-390 to val-390 substitution, one occurred at the +2 position of the corresponding codon (nucleotide 497, Figure 4.4.3), which created a conservative thr-166 to ser-166 substitution and twenty-eight occurred at the +3

```
Cat cDNA                          -43     AAACAATCAACATCACACCACACCCCCATACTAATTATTCAAT    -1
Cat gene                          -43     AATCAATACCACCATCACCACACCCCCACACTAATTATTCAAT    -1
                                            *    ********            *

ATG GCT CCT ACC TTC ACA AAC TCA AAC GGT CAA CCT ATC CCT GAA CCT TTC GCC ACT CAA AGA GTT    66
ATG GCT CCT ACC TTC ACA AAC TCA AAC GGT CAA CCT ATC CCT GAA CCT TTC GCC ACT CAA AGA GTT    66

Gly                                                                 Ala
GGC CAA CAC GGT CCT TTG TTG TTG CAA GAC TTC AAC TTG ATT GAC TCC TTG GCI CAC TTC GAC AGA   132
GGI CAA CAC GGT CCT TTG TTG TTG CAA GAC TTC AAC TTG ATT GAC TCC TTG GCC CAC TTC GAC AGA   132
Gly                                                                 Ala
                              Val         Lys Gly Ser               Val
GAA AGA ATC CCA GAA AGA GTC GTC CAC GCT AAA GGC TCC GGT GCT TAC GGT GTI TTT GAA GTC ACC   198
GAA AGA ATC CCA GAA AGA GTC GTI CAC GCT AAG GGI TCA GGT GCT TAC GGT GTC TTT GAA GTC ACC   198
                              Val         Lys Gly Ser               Val
                                          Phe                   Lys Lys
GAT GAC ATC ACC GAC GTT TGT GCC GCA AAG TTI TTG GAC ACT GTC GGT AAG AAA ACC AGA ATT TTC   264
GAT GAC ATC ACC GAC GTT TGT GCC GCA AAG TTC TTG GAC ACT GTC GGT AAA AAG ACC AGA ATT TTC   264
                                          Phe                   Lys Lys
       Phe                        Leu
ACT AGA TTC TCC ACC GTC GGT GGT GAA ITG GGT TCC GCT GAC ACC GCC AGA GAC CCA AGA GGT TTT   330
ACT AGA TTI TCC ACC GTC GGT GGT GAA CTG GGT TCC GCT GAC ACC GCC AGA GAC CCA AGA GGT TTT   330
       Phe                        Leu
                                                                  Pro
GCT ACC AAG TTT TAC ACC GAA GAA GGT AAC TTG GAC TTG GTT TAC AAC AAC ACC CCA GTC TTT TTC   396
GCT ACC AAG TTT TAC ACC GAA GAA GGT AAC TTG GAC TTG GTT TAC AAC AAC ACC CCI GTC TTT TTC   396
                                                                  Pro
              Ser                                                 Pro
ATT AGA GAC CCT TCI AAG TTC CCT CAC TTC ATC CAC ACC CAA AAG AGA AAC CCA GAA ACT CAC TTG   462
ATT AGA GAC CCT TCC AAG TTC CCT CAC TTC ATC CAC ACC CAA AAG AGA AAC CCG GAA ACT CAC TTG   462
              Ser                                                 Pro
                                        Ser
AAG GAC GCC AAC ATG TTC TGG GAC TAC TTG ACC AGC AAC GAA GAA TCC GTC CAC CAA GTT ATG GTT   528
AAG GAC GCC AAC ATG TTC TGG GAC TAC TTG ACC ACC AAC GAA GAA TCC GTC CAC CAA GTT ATG GTT   528
                                        Thr

TTG TTC TCC GAC AGA GGT ACC CCA GCT TCC TAC AGA GAA ATG AAT GGT TAC TCT GGC CAC ACC TAC   594
TTG TTC TCC GAC AGA GGT ACC CCA GCT TCC TAC AGA GAA ATG AAT GGT TAC TCT GGC CAC ACC TAC   594

                                     Tyr Val Gln
AAG TGG TCC AAC AAC AAG GGT GAA TGG TTC TAI GTI CAA GTC CAT TTC ATC AGT GAC CAA GGT ATC   660
AAG TGG TCC AAC AAC AAG GGT GAA TGG TTC TAC GTC CAG GTC CAT TTC ATC AGT GAC CAA GGT ATC   660
                                     Tyr Val Gln

AAG ACC TTG ACC AAC GAA GAA GCT GGC TCC TTG GCT GGT TCC AAC CCA GAC TAT GCT CAA GAA GAT   726
AAG ACC TTG ACC AAC GAA GAA GCT GGC TCC TTG GCT GGT TCC AAC CCA GAC TAT GCT CAA GAA GAT   726

       Lys                                       Thr
TTG TTC AAG AAC ATT GCT GCT GGT AAC TAC CCA TCC TGG ACI TGT TAC ATC CAA ACC ATG ACT GAA   792
TTG TTC AAA AAC ATT GCT GCT GGT AAC TAC CCA TCC TGG ACC TGT TAC ATC CAA ACC ATG ACT GAA   792
       Lys                                       Thr
```

```
                    Ala         Ser
GCT CAA GCT AAA GAA GCC GAA TTC TCT GTC TTT GAC TTG ACC AAG GTC TGG CCA CAC GGA AAG TAC   858
GCT CAA GCT AAA GAA GCT GAA TTC TCC GTC TTT GAC TTG ACC AAG GTC TGG CCA CAC GGA AAG TAC   858
                    Ala         Ser

CCA ATG AGA AGA TTT GGT AAG TTT ACT TTG AAC GAA AAC CCA AAG AAC TAC TTT GCC GAA GTT GAA   924
CCA ATG AGA AGA TTT GGT AAG TTT ACT TTG AAC GAA AAC CCA AAG AAC TAC TTT GCC GAA GTT GAA   924

                                    Val                     Ser
CAA GCT GCT TTC TCC CCA GCC CAC ACC GTC CCA CAC ATG GAA CCA TCC GCT GAC CCA GTC TTG CAA   990
CAA GCT GCT TTC TCC CCA GCC CAC ACC GTT CCA CAC ATG GAA CCA TCT GCT GAC CCA GTC TTG CAA   990
                                    Val                     Ser
Ser
TCT AGA TTG TTC TCT TAT GCT GAT ACT CAC AGA CAC AGA TTG GGT ACC AAC TAC ACC CAA ATC CCA  1056
TCC AGA TTG TTC TCT TAT GCT GAT ACT CAC AGA CAC AGA TTG GGT ACC AAC TAC ACC CAA ATC CCA  1056
Ser

GTC AAC TGC CCA GTC ACT GGT GCT GTT TTC AAC CCA CAC ATG AGA GAC GGT GCC ATG AAT GTC AAT  1122
GTC AAC TGC CCA GTC ACT GGT GCT GTT TTC AAC CCA CAC ATG AGA GAC GGT GCC ATG AAT GTC AAT  1122

                                                Lys Pro Val             Gln     Ser
GGT AAC TTG GGC AAC CAT CCA AAC TAC TTG GCC TCT GAC AAG CCT GTT GAA TTC AAG CAG TTC TCT  1188
GGT AAC TTG GGC AAC CAT CCA AAC TAC TTG GCC TCT GAC AAA CCA ATT GAA TTC AAG CAA TTC TCC  1188
                                                Lys Pro Ile             Gln     Ser

CTC CAA GAA GAC CAA GAA GTC TGG CAC GGT GCT GCT ACT CCA TTC CAC TGG AAG GCT ACC CCA GCT  1254
CTC CAA GAA GAC CAA GAA GTC TGG CAC GGT GCT GCT ACT CCA TTC CAC TGG AAG GCT ACC CCA GCT  1254

GAT TTC AAG CAA GCC ACT GAA TTG TGG AAG GTG TTG AAG AAA TAT CCA AAC CAA CAA GAA CAC TTG  1320
GAT TTC AAG CAA GCC ACT GAA TTG TGG AAG GTG TTG AAG AAA TAT CCA AAC CAA CAA GAA CAC TTG  1320

GCC CAC AAT GTT GCT GTC CAT GCC TCT GCT GCC GAT GCT CCA ATC CAA GAC AGA GTC ATT GCT TAC  1386
GCC CAC AAT GTT GCT GTC CAT GCC TCT GCT GCC GAT GCT CCA ATC CAA GAC AGA GTC ATT GCT TAC  1386

TTT ACC AAG GTC CAC CCA GAC TTG GGA GAC CTC ATC AAG AAG GAA ATC TTG GAA TTG TCT CCA AGA  1452
TTT ACC AAG GTC CAC CCA GAC TTG GGA GAC CTC ATC AAG AAG GAA ATC TTG GAA TTG TCT CCA AGA  1452

AAG TAA AGAAGATTCTGATTTGATATTGGAGTGCTAAAGGTCTTAGTTATGTTTGTTTCTTTTCTTTTTTCTTATTAATATTTAT  1537
AAG TAA AGAAGATTCTGATTTGATATTGGAGTGCTAAAGGTCTTAGTTATGTTTGTTTCTTTTCTTTTTTCTTATTAATATTTAT  1537

AGCTGTTGTAATGCGTTTTACAGTAAATTGAATGGTTATATTTGATATATCTGA     1591              Cat cDNA
AGCTGTTGTAATGCGTTTTACAGTAAATTGAATGGTTATATTTGATATATCTGA     1591              Cat gene
```

Figure 4.4.3. Comparison of the nucleotide sequences of the Cat cDNA and gene including the 5'- and 3'- flanking regions. The base pair numbering system of the Cat cDNA is presented in Figure 4.2.2; that of the Cat gene is presented in Figure 4.4.2. Mismatched nucleotides identified between the open reading frames are underlined; those identified between the 5'-flanking sequences are denoted by asterisks (no mismatches were identified between the 3'-flanking sequences). The amino acids encoded by the thirty-one mismatched codons are indicated above and below the nucleotide sequences of the Cat cDNA and gene, respectively, with the two mismatched amino acids highlighted by bold type (see text, section 4.4.1).

position of the corresponding codons (Figure 4.4.3), none of which altered the amino acid sequence of Cat. Further comparison of the sequence of this ORF with that of human catalase (Quan et al., 1986), rat catalase (Furuta et al., 1986) and S. cerevisiae catalases T (CTT1; Hartig and Ruis, 1986) and A (CTA1; Cohen et al., 1988) revealed nucleotide sequence identities of 52.5%, 52.9%, 49.3% and 61.0%, respectively, which reaffirmed the Cat-encoding capacity of this gene.

As previously mentioned, the 5'-noncoding region of the Cat gene consisted of 938 nucleotides; this compared with 650 nucleotides of 5'-flanking sequence reported by Okada et al. (1987) for a similar BamH I, Sal I fragment of C. tropicalis genomic DNA harbouring the POX9 gene. Sequencing revealed that this difference resulted from a G (position -650 of the POX9 sequence, Figure 4.4.2, (G)) to A (position -654 of the Cat gene, Figure 4.4.2) transition which eliminated the BamH I restriction site upstream of the POX9 gene. Further, it was determined that the 5'-flanking BamH I site used to subclone the BamH I, Sal I genomic fragment of phage 4a DNA in this work was present at position -938. Consequently, a polymorphic BamH I site exists 5' to the catalase gene of C. tropicalis, the absence of which generated 284 nucleotides of novel upstream sequence (residues -655 to -938, Figure 4.4.2) relative to

that reported for the POX9 gene (Okada et al., 1987). Similar restriction fragment polymorphisms have been reported to exist in the C. maltosa genomic DNA encompassing POX4 (Hill et al., 1988), suggesting allelic heterogeneity in the C. maltosa genome. Allelic heterogeneity within the C. tropicalis genome was not only found in those sequences which flanked the Cat gene, but was also observed upon comparison of the AOx and POX4 genes (Okazaki et al., 1986) of this yeast (section 3.2.3). Other nucleotide differences that existed between the 5'-noncoding region of the Cat gene and that of POX9 (Okada et al., 1987) included (Figure 4.4.2): 1) insertions at positions -215(C), -243(T), -249(G), -541(G), -548(C), -619(C) and -620(G), 2) deletions between positions -282/-283(C), -406/-407(A) and -537/-538(C) and 3) substitutions at positions -288(G for C), -289(C for G) and -290(G for C). Comparison of the 5'-noncoding region of the Cat gene to that of the Cat cDNA revealed less than perfect nucleotide sequence identity, similar to that observed upon comparison of the ORF of the Cat gene and cDNA. Of the 43 nucleotides of 5'-flanking sequence determined for the Cat cDNA, ten mismatches with that of the Cat gene were found (Figure 4.4.3, asterisks), thereby yielding a nucleotide sequence identity within this region of only 76.7%. The consequence(s) of these differences, with respect to transcription and/or

translation of the Cat gene or mRNA, if any, are not known.

Within the 5'-noncoding regions of most yeast genes, yeast promoters, which are necessary for the regulation, amount and accuracy of transcription initiation, contain three basic DNA sequence elements: upstream activating, TATA and initiation elements (for reviews, see Guarente, 1987, 1988; Struhl, 1987, 1989). Upstream activating sequences (UASs), located between 100 and 1500 bp upstream of the mRNA initiation site, determine the particular regulatory properties of a given promoter. They are required, in addition to the TATA box, for transcription to occur; when ablated, transcription is abolished. Certain other upstream promoter elements, including CCAAT boxes and GC boxes, which facilitate transcription by specifically binding CCAAT- and Sp1 transcription activators, respectively, are localized to many different eukaryotic promoters (Dynan and Tjian, 1985), including those of yeast genes (Schena, 1989). By comparison, TATA elements (consensus sequence = TATAAA; Breathnach and Chambon, 1981), found at a distance of 40 to 120 bp upstream of the mRNA initiation site, are necessary but not sufficient for the transcriptional initiation of most yeast genes. Deletion of these elements greatly reduces the mRNA level (Struhl, 1987, 1989). Conversely, initiator elements, which are located near the actual mRNA start site, exert minimal effects on the mRNA level but are

the primary determinants of where transcription begins (Struhl, 1987, 1989). Several consensus sequences have been determined for yeast initiation sites; two of these, TC(G/A)A and RRYRR (where R=purine, Y=pyrimidine), account for greater than 50% of the known initiation sites (Guarante, 1987). Recognizing these inherent characteristics of most yeast promoters, the 938 nucleotides of the 5'-noncoding region of the Cat gene were analyzed for similar promoter elements. A consensus TATA box sequence, TATAAA, was found at position -121 (Figure 4.4.2, boxed), which was identical to its location upstream of the POX9 gene (Okada et al., 1987) and corresponded to a position 50 to 80 nucleotides upstream of the approximate Cat mRNA initiation site(s) (-70 to -40 nucleotides upstream of the start codon) determined by primer extension analysis (section 4.1.3). Within this -70 to -40 region, two copies of each of the aforementioned consensus sequences for transcription initiation, TC(G/A)A and RRYRR, were found; TCAA was located at both the -72 and -41 positions whereas AATAA and AACAA were located at the -69 and -46 positions, respectively (Figure 4.4.2, double underlining). This finding may substantiate the previously postulated existence of two independent tsp for the Cat mRNA (section 4.1.3). Interestingly, at a position ≈100 nucleotides upstream of the putative mRNA transcription start site, the

decanucleotide sequence, CCTCCGCCCC, exists (position -167, Figure 4.4.2, highlighted by *••••*), which differs by only one nucleotide from the consensus GC box sequences, (G/T)GGGCGG(G/A)(G/A)(C/T), and its inverse complement, (G/A)(C/T)(C/T)CCGCCC(C/A), proposed for strong Sp1 binding sites (Dynan et al., 1986). Upstream of this putative GC box, a consensus CCAAT box, was found at position -501 (Figure 4.4.2, underlining). A second CCAAT box, located at position -61, was also found; however, functionality of this sequence, with respect to the activation of transcription, is questionable based upon its "aberrant" location downstream of the TATA box, within the region of transcription initiation. The expression of S. cerevisiae catalases A and T (ie. CTA1 and CTT1 genes) is controlled by heme (Hörtner et al., 1982; Richter et al., 1980) and accordingly, the upstream regions of both genes contain consensus binding sites for the HAP1 protein (Cohen et al., 1988), an important mediator of heme control (Pfeifer et al., 1987). By comparison, similar consensus sequences could not be localized to the 938 nucleotides of 5'-noncoding sequence determined for the Cat gene. This implies that Cat gene expression in C. tropicalis may be regulated by an alternative method to that employed by S. cerevisiae, should a HAP1 binding site(s) not exist farther upstream of the 5'-noncoding sequence determined here.

Of the 2,921 nucleotides which constituted the BamH I, Sal I fragment of C. tropicalis genomic DNA, 528 nucleotides (residues 1456 to 1983, Figure 4.4.2) was 3'-noncoding sequence which flanked the Cat gene. Differences relative to the 530 nucleotides of 3'-flanking sequence reported for the POX9 gene (Okada et al., 1987) were minor; they included nucleotide deletions between positions 1725/1726(A) and 1776/1777(G) (Figure 4.4.2). Comparison of the 3'-noncoding sequence of the Cat gene to the 135 nucleotides of 3'-noncoding sequence of the Cat cDNA [excluding the poly(A) tract] revealed perfect nucleotide sequence identity (Figure 4.4.3). Similar to the Cat cDNA (section 4.4.2), a possible polyadenylation sequence, AGTAAA (nucleotides 1559 to 1564, Figure 4.4.2, boxed), was located 27 nucleotides upstream of the polyadenylation site (Figure 4.4.2, arrow). A second consensus sequence for polyadenylation, AATAAA, was found between nucleotides 1739 and 1744 (Figure 4.4.2), however, its position downstream of the Cat mRNA polyadenylation site negated functionality for this sequence. Two signals for transcription termination, TAG...TATGT...TTT (Zaret and Sherman, 1982) and TATTTATA (Henikoff and Cohen, 1984), were located, in tandem, upstream of the AGTAAA polyadenylation sequence (Figure 4.4.2, underlining), at identical positions to those determined for the 3'-flanking region of the Cat cDNA

(Figure 4.2.2). Two A-rich regions [nucleotides 1710 to 1727 (16 of 17 residues = A) and 1736 to 1751 (15 of 16 residues = A), Figure 4.4.2] and a region of eight 'CAA' repeats (nucleotides 1806 to 1829, Figure 4.4.2) were noted downstream of the polyadenylation site; prescribed functions for these sequences, if any exist, are not known.

The nucleotide frequency of the Cat gene/cDNA, A + T: 55.7%/55.1%, C + G: 44.3%/44.9% (Table 4.4.1), approximated that of several other peroxisomal genes from C. tropicalis including AOx (A + T: 55.2%; Table 3.2.1), POX2 (A + T: 54.4%; Okazaki et al., 1987), POX4 (A + T: 53.9%; Okazaki et al., 1986), POX5 (A + T: 52.2%; Okazaki et al., 1986), HDE (A + T: 54.8%; Nuttley et al., 1988), isocitrate lyase (A + T: 55.3%; Atomi et al., 1990) and POX18 (A + T: 56.7%; Szabo et al., 1989; Tan et al., 1990). Nevertheless, as previously mentioned (section 3.2.1), these A + T nucleotide frequencies were inconsistent with the estimated A + T composition of C. tropicalis genomic DNA (65.1%; Stenderup and Bak, 1968). It was postulated that this difference could be accounted for by a strong tendency toward A + T-richness within the noncoding regions of the C. tropicalis genome, as was determined for the 5'-flanking sequence of the AOx gene (A + T content= 70.4%, section 3.2.1). While an A + T composition of 55.1% for the 938 nucleotides of 5'-sequence flanking the Cat gene is

Table 4.4.1. Nucleotide frequency and deduced amino acid composition of the Candida tropicalis Cat gene and protein.

NUCLEOTIDE FREQUENCY	A - 891/479 (30.5/29.2)				
	C - 685/422 (23.5/25.7)				
	G - 608/315 (20.8/19.2)				
	T - 737/427 (25.2/26.0)				
	A + T -1628/906 (55.7/55.1)				
	C + G -1293/737 (44.3/44.9)				
DINUCLEOTIDE FREQUENCY	AA - 315/162	(10.8/9.9)	GA - 194/109	(6.6/6.6)	
	AC - 209/134	(7.2/8.2)	GC - 106/58	(3.6/3.5)	
	AG - 195/102	(6.7/6.2)	GG - 142/63	(4.9/3.8)	
	AT - 172/80	(5.9/4.9)	GT - 166/85	(5.7/5.2)	
	CA - 249/145	(8.5/8.8)	TA - 133/62	(4.6/3.8)	
	CC - 209/127	(7.2/7.7)	TC - 161/103	(5.5/6.3)	
	CG - 57/33	(2.0/2.0)	TG - 213/117	(7.3/7.1)	
	CT - 169/117	(5.8/7.1)	TT - 230/145	(7.9/8.8)	
AMINO ACID COMPOSITION	Ala 41	(8.5)	Leu 30	(6.2)	
	Arg 20	(4.1)	Lys 28	(5.8)	
	Asn 29	(6.0)	Met[a] 9	(1.9)	
	Asp 28	(5.8)	Phe 32	(6.6)	
	Cys 3	(0.6)	Pro 32	(6.6)	
	Gln 21	(4.3)	Ser 24/25	(4.9/5.2)	
	Glu 29	(6.0)	Thr 35/34	(7.2/7.0)	
	Gly 29	(6.0)	Trp 8	(1.6)	
	His 22	(4.5)	Tyr 18	(3.7)	
	Ile 17/16	(3.5/3.3)	Val 30/31	(6.2/6.4)	
	Acidic	(Asp+Glu)	57	(11.8)	
	Basic	(Arg+Lys)	48	(9.9)	
	Aromatic	(Phe+Trp+Tyr)	58	(12.0)	
	Hydrophobic	(Aromatic+Ile +Leu+Met+Val)	144	(29.7)	

M_r[a]= 54944/54916 Calculated pI = 6.10

Nucleotide, dinucleotide and amino acid frequencies were calculated using the Beckman MicroGenie (Version 6.0) Sequence Analysis Program (Queen and Korn, 1984). Numbers in parentheses indicate the % of the total number of "residues" (nucleotides, dinucleotides or amino acids) in Cat represented by each "residue". The two sets of values shown represent the different frequencies obtained for the Cat gene/cDNA, respectively.

[a] includes the N-terminal (initiator) methionine residue

inconsistent with this supposition, an A + T nucleotide frequency of 65.5% for the 528 nucleotides of 3'-flanking sequence lends support to this premise. Alternatively, a combined A + T content of only 61.6% for the 5'- and 3'-flanking regions of all of the aforementioned peroxisomal genes from C. tropicalis (including AOx, Cat, POX2, POX4, POX5, POX18, HDE and isocitrate lyase) may be indicative of an excessive estimate for the A + T composition of the C. tropicalis genome derived from the thermal denaturation data of Stenderup and Bak (1968). It should be noted, however, that the C. tropicalis gene encoding NADPH-cytochrome P-450 oxidoreductase (Sutter et al., 1990) possesses an A + T content of 64.9%, which implies that the A + T composition of at least one gene (and possibly more) from C. tropicalis parallels the 65.1% A + T content purported for the entire genome. Consistent with that determined for the AOx gene (section 3.2.1), the frequency of occurrence within the Cat gene of the dinucleotide CpG was only 2.0% (Table 4.4.1), which further supported the possibility of CpG suppression within the C. tropicalis genome (section 3.2.1).

4.4.2 Deduced Amino Acid Sequence of Cat

Based upon the sequence of the 1,455 nucleotide ORF of the Cat gene/cDNA, a polypeptide of 485 amino acids (M_r 54944/54916; Table 4.4.1) was predicted as the primary structure of Cat (Figures 4.4.2/4.2.2, respectively).

Residues 2 to 24 of the deduced polypeptide were identical to the NH_2-terminal 23 amino acids of PXP-9, ala-pro-thr-phe-thr-asn-ser-asn-gly-gln-pro-ile-pro-glu-pro-phe-ala-thr-gln-arg-val-gly-gln, which were determined by sequencing purified peroxisomal catalase (Ueda et al., 1987; Okada et al., 1987). Consistent with several yeast proteins, including C. tropicalis peroxisomal proteins PXP-2, -4 and -5 (Okazaki et al., 1986, 1987), the initiator methionine of PXP-9 is deleted from the "mature" polypeptide (Okada et al., 1987). Thus, the calculated M_r of "mature" Cat (ie. without the methionine) is 54795, which agrees with the M_r of 57000 estimated by hybridization-selection translation, followed by SDS-PAGE (Rachubinski et al., 1987) and compares favourably with the M_r of 54000 estimated for the catalase subunit, subsequent to SDS-PAGE of the purified enzyme isolated from C. tropicalis (Yamada et al., 1982). The amino acid composition of Cat (Table 4.4.1), deduced from both the Cat gene and cDNA, was typical of other catalases from diverse sources in having relatively low amounts of cysteine and tryptophan (Loewen and Switala, 1987). Consistent with its peroxisomal matrix location (Yamada et al., 1982), Cat was found to be comprised of only 29.7% hydrophobic residues (144 aromatic + ile + leu + met + val amino acids of 485 total). This compared well with values determined for catalase polypeptides, of either cytosolic or

peroxisomal location, from several species, including mammals [human lymphoblast catalase, 30.4% (Quan et al., 1986)], plants [cottonseed catalase, 32.1% (Ni et al., 1990)], yeast [S. cerevisiae catalase A (peroxisomal), 29.9% (Cohen et al., 1988); S. cerevisiae catalase T (cytosolic), 31.7% (Hartig and Ruis, 1986)] and Drosophila [Drosophila catalase, 30.6% (Orr et al., 1990)]. Further, it closely resembled the percentages calculated for other C. tropicalis matrix proteins, including AOx (33.6%, Table 3.2.1), PXP-2 (32.7%, Okazaki et al., 1987), PXP-4 (33.6%, Okazaki et al., 1986), PXP-5 (33.4%, Okazaki et al., 1986), PXP-18 (33.1%, Szabo et al., 1989; Tan et al., 1990), HDE (31.9%, Nuttley et al., 1988) and isocitrate lyase (30.2%, Atomi et al., 1990).

As described by Cohen and Parry (1986), proteins exhibiting a high ratio (0.9 to 1.4) of charged to apolar residues generally adopt a rod-shaped form, consistent with the presence of an extended conformation, which aids structural stability by decreasing both the surface charge density and internal volume. Alternatively, proteins with a low ratio (0.3 to 0.6) of charged to apolar residues adopt a globular shape since the apolar groups are more easily accomodated internally. The ratio of charged to apolar residues in Cat (0.73) suggested that this polypeptide adopted a generally globular form, with the hydrophobic

residues internalized and the charged moieties exposed at the surface. This appears to be true of catalases, cytosolic or peroxisomal, from several species, with appropriate ratios being 0.74 for human lymphoblast catalase (Quan et al., 1986), 0.74 for cottonseed catalase (Ni et al., 1990), 0.69 for S. cerevisiae catalase A (Cohen et al., 1988), 0.72 for S. cerevisiae catalase T (Hartig and Ruis, 1986) and 0.75 for Drosophila catalase (Orr et al., 1990). C. tropicalis AOx (ratio of charged: apolar residues = 0.67, section 3.2.2) and Cat polypeptides most likely adopt similar, generally globular conformations, possessing only minimal or short-range regularity in secondary structure, with the Cat polypeptide exhibiting a slightly less compact form (Cohen and Parry, 1986).

Contrary to that determined for AOx (Table 3.2.1), the number of acidic (asp + glu) residues (57) in Cat exceeded the number of basic (arg + lys) residues (48) (Table 4.4.1); however, a calculated pI of 6.10 for Cat was identical to that of AOx (section 3.2.2). Although an experimentally determined value of the pI of Cat could not be found, the calculated value of 6.10 compares with those of 5.2 and 5.7 determined experimentally for plant catalases from Zantedeschia aethiopica (Trindade et al., 1988) and maize (Scandalios et al., 1972), respectively.

Prediction of the secondary structure of Cat (Figure

4.4.4, Alph and Beta graphs), using the algorithm of Garnier et al. (1978), indicated a slight dominance of α-helical content (≈33%), similar to AOx (section 3.2.2), with the most prominent, extended region of α-helix, encompassing C-terminal amino acids 419 to 483 (Figure 4.4.4, Alph graph, bar), interrupted by only small segments of turn/random coil and β-sheet structures. The remainder of the Cat polypeptide was predicted to be comprised predominantly (≈28%) of random coils, which were interspersed with ≈21% turn and ≈19% β-sheet structures. The higher portion of turn and random coil structures within Cat, which may adopt a more loosely packed structure than ordered α-helices or β-sheets, would be consistent with the aforementioned less compact, globular conformation of Cat relative to AOx. Two different hydropathy profiles for Cat, one calculated according to the algorithm of Hopp and Woods (1981) and the other according to the algorithm of Kyte and Doolittle (1982) are also illustrated in Figure 4.4.4 (Hyd and Val graphs, respectively). Using the algorithm of Hopp and Woods (1981), which aids the determination of protein antigenic determinants by identifying those areas of greatest local hydrophilicity, it may be speculated that Cat possesses five major antigenic determinants [residues 45 to 50, 100 to 105, 145 to 150, 263 to 269 and 470 to 480 (Figure 4.4.4, Hyd graph, closed triangles)], three of which

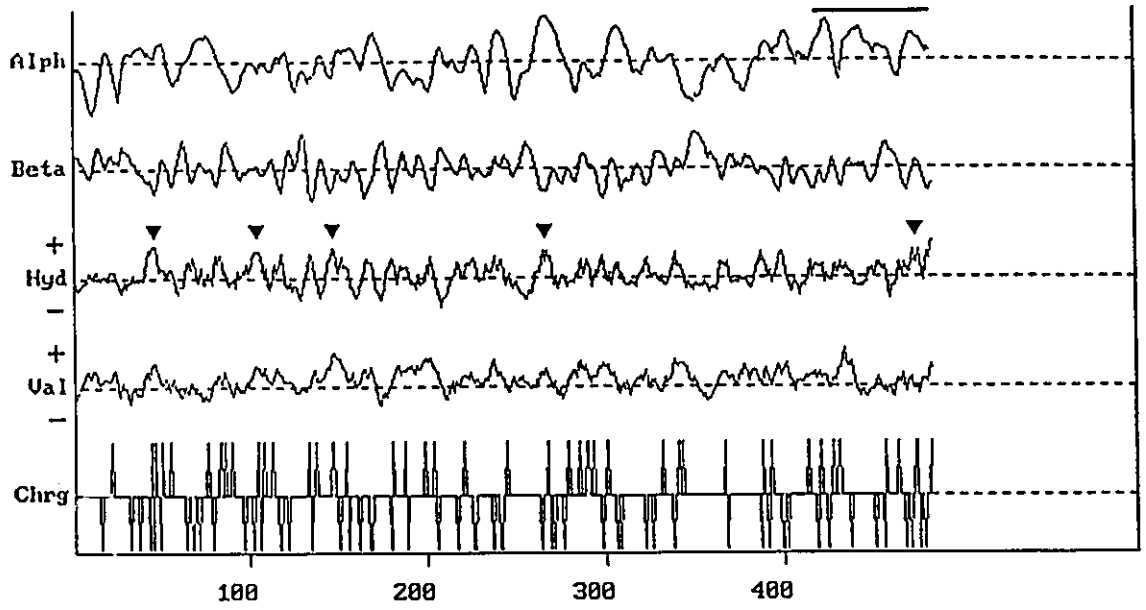


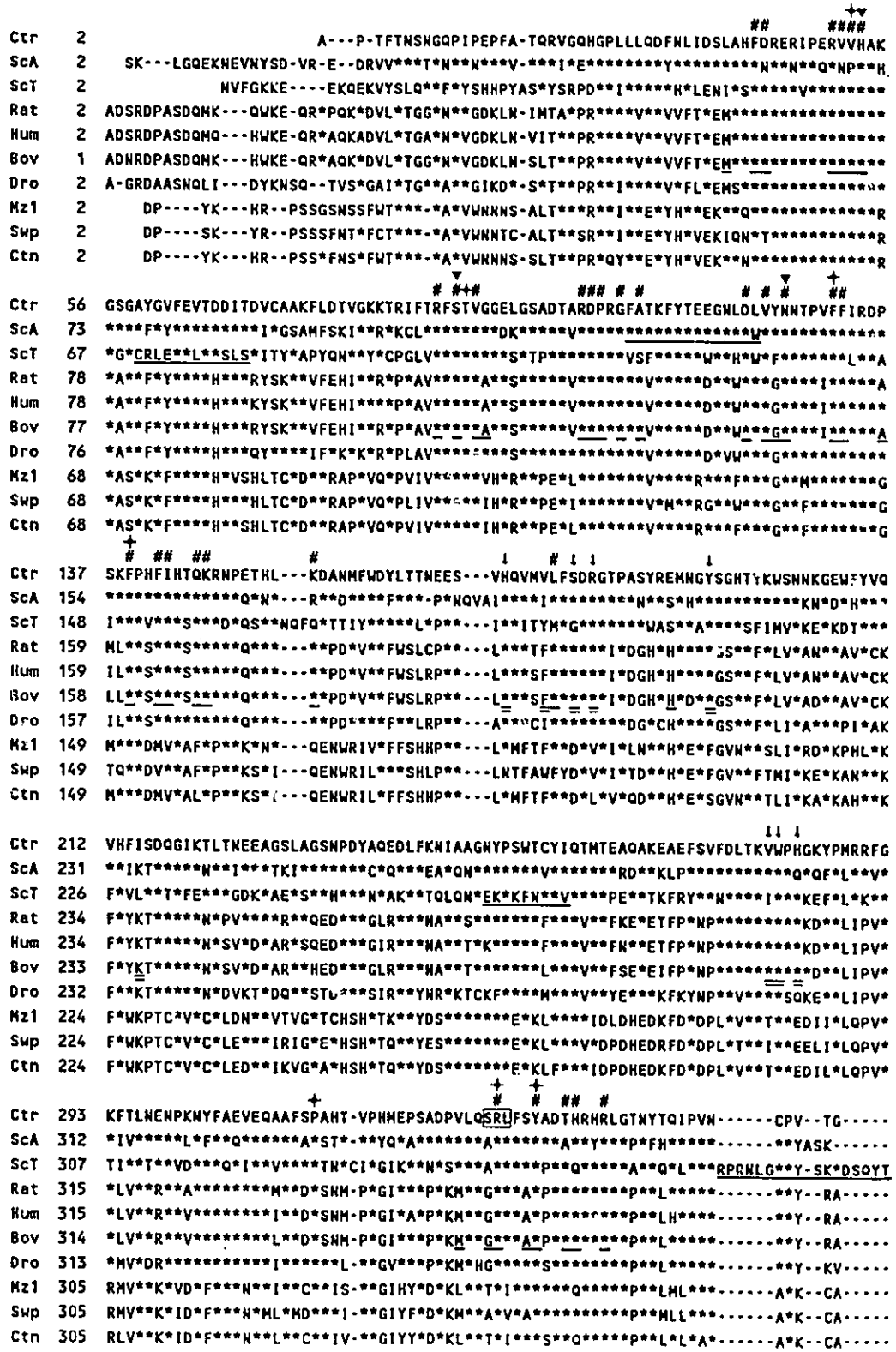
Figure 4.4.4. Predicted secondary structure, hydropathy and charge profile of Cat. Graphs were generated using standard parameters for the Beckman MicroGenie (Version 6.0) Sequence Analysis Program (Queen and Korn, 1984) and the amino acid sequence for Cat deduced from the nucleotide sequence of the Cat gene. Secondary structure (Alph = Á-helix, Beta = Ã-sheet) was determined using the algorithm of Garnier et al. (1978); the bar above the graph of Á-helical content highlights the most prominent region of Á-helix within Cat (see text, section 4.4.2). Hydropathy was determined using the algorithms of Hopp and Woods (1981) (Hyd) and Kyte and Doolittle (1982) (Val). For both algorithms, hydrophilicity is indicated by + (above the base line) while hydrophobicity is indicated by - (below the base line); closed triangles above the graph plotted according to Hopp and Woods (1981) indicate possible antigenic determinants within Cat (see text, section 4.4.2). The charge profile (Chrg) was plotted by assigning values of +1 to R or K residues, -1 to D or E residues and 0 to all other residues.

were located toward the N-terminus. For the most part, this postulation is consistent with that observed upon immunoprecipitation of N- or C-terminal-deleted, in vitro translated Cat polypeptides (see section 5.2.2). Paralleling the location of Cat within the peroxisomal matrix (Yamada et al., 1982), the hydropathy profile corresponding to the algorithm of Kyte and Doolittle (1982) (Figure 4.4.4, Val graph) did not reveal any significant regions of uninterrupted hydrophobicity (ie. potential membrane spanning domains). In fact, this algorithm like that of Parker et al. (1986) (not shown), predicted an overall hydrophilic nature for the Cat polypeptide, more so than the algorithm of Hopp and Woods (1981). The charge profile of Cat was, in general, nondescript, characterized simply by regularly interspersed acidic and basic amino acids (Figure 4.4.4, Chrg graph). Nevertheless, in light of the C-terminal tripeptide PTS, ser/ala/cys-lys/his/arg (ie. basic amino acid)-leu, identified by Gould et al. (1989), it may be of some consequence that both the Cat and AOx polypeptides (Figure 3.2.4) possess C-terminal α-helical (ie. stable) structures which terminate with at least one basic amino acid.

4.4.3 Comparison of the Deduced Amino Acid Sequence of Cat with those of Other Catalases

Comparison of the deduced amino acid sequence of the

Cat gene with that of C. tropicalis POX9 (Okada et al., 1987) revealed perfect amino acid identity between these polypeptides; however, comparison of the deduced amino acid sequences of Cat derived from both the Cat gene and cDNA revealed slightly less than perfect identity (99.6%) since the thirty-one nucleotide sequence mismatches found between the Cat gene and cDNA (section 4.4.1) translated into two conservative amino acid substitutions, thr-166 to ser-166 and ile-390 to val-390, respectively (Figure 4.4.3). Using the polypeptide sequence deduced from the Cat gene, Cat was compared to several other catalases from diverse species. As seen in Figure 4.4.5, mammalian catalases, including rat liver catalase (Furuta et al., 1986), human lymphoblast catalase (Quan et al., 1986) and bovine liver catalase (Schroeder et al., 1982), are highly homologous to each other, exhibiting amino acid identities greater than 87.8% throughout the entire polypeptide. Plant catalases, including maize (Zea mays L.) CAT-1 (Redinbaugh et al., 1988), sweet potato catalase (Sakajo et al., 1987) and cottonseed catalase (Ni et al., 1990), are slightly less homologous as a group, exhibiting amino acid identities ranging between 74.0% (maize vs. sweet potato catalase) and 84.8% (maize vs. cottonseed catalase). Yeast catalases, including C. tropicalis Cat (Figure 4.4.2; Okada et al., 1987) and S. cerevisiae catalases A (Cohen et al., 1988) and



```
Ctr  360  AVFNP-----HMRDGAMNVNGNLGNHPNY----L-------------ASDK---PI--EFKQ--FS--LQE-DQ---EVW
ScA  380  -F***-----AI***P******F*SE*T*----*-------------*N**---SY--TYI*QDRP--I*Q-H*---***
ScT  387  *EQC*FKAVNFQ***P*-SYY*F*PE***-ISS*PNQTLKFKNEVNDEV***FKGIVLD*VTE--V*VRK**Q**IRN*HI
Rat  382  R*A*------YQ***P*CMHD*Q*GA***Y---------------PN*FS--A*---*-Q*--G*--AL*-HH---SQC
Hum  382  R*A*------YQ***P*CMQD*Q*GA***Y---------------PN*FG--A*---*-Q*--P*--AL*-HS---IQY
Bov  381  R*A*------YQ***P*CMHD*Q*GA***Y---------------PN*FS--A*---*-H*--P*--AL*-HR---THF
Dro  380  KIE*------FQ*******TD*QDGA***F---------------PN*FN--G*------*ECPRARALSSCC---P*T
Mz1  372  HHN*H-----*--**F**F-MHRDEEV**F---------------P-*---------K*DP--AR--HA*-KV---PIP
Swp  372  HHN*S-----Y--**Y**F-VHRDEEVD*F---------------P-*---------K*DN--TR--NA*-RF---PTP
Ctn  372  HHN*H-----*--E*F**F-MHRDEEI**F---------------P-*---------RYDP--VR--HA*-MF---PIP

Ctr  405  HGA---ATPFHWK---AT-P--ADFKQ-ATEL-W-KVLKKYPNQQEHL-AHN--VAVHA-SAADAPIQDRVIAYFTKVHPD
ScA  426  N*P---*I*Y**----**S*GDV**V*-*RN*-Y-R**G*Q*G**KN*-*Y*--IGI*V-EG*CPQ**Q**YDH*AR*DKG
ScT  464  VD*KINQYYYVYG---IS-*--L**E*-PRA*----YE*V*NDE*KK*FV**--*VC**CKIK*PKVKK**TQ**GLLNE*
Rat  426  S-*---DVKRFNS---*N-E--DNVT*VR*FY-T-***--NEEERKR*-CE*--I*N*L-KD*QLF**RKAVKN**D****
Hum  426  S*E---VRR*-NT---*--N--D*NVT-QVRA-F-Y*NVLNEE*RKR*-CE*--I*G*L-KD*QIF**KKAVKN**E****
Bov  425  S*D---VQR*-NS---*--N--D*NVT-QVRT-F-YLKVLNEE*RKR*-CE*--I*G*L-KD*QLF**KKAVKN*SD***E
Dro  427  -*D---VYRY-------SS-GDTEDNFGQVTDF*VH**D*CA--KKR*-VQ*--I*G*L-*N*SQFL*E*AVKN**Q**A*
Mz1  410  PRV---L*GCRE*CIIQK-E--NN***-*G*R-Y-RSFD--*AR*DRF-IQRWVD*LTD-PRVTHEHRTIW*S*WSQCDAA
Swp  410  LRI---V*GQRD*CVIEK-E--NN***-PGDR-Y-RSWA--*DR*DRF-INRWVK*LSE-PRVTHE*RSTW*S*L*QADRS
Ctn  410  PAV---C*GRRE*CIIEK-E--NN***-PG*R-Y-RSWA--ADR**RF-ICRWVD*LSD-PRVTHE*RSIW*S*WSQADKS

Ctr  470  LGDLIKKEILELSPRK                                  485
ScA  493  *SEA***VAEAKHASELSSNSKF                           515
ScT  532  **KV*-A*G*-GV*WE**DLEG-YAKTWS-IAS---AN            562
Rat  489  Y*ARVQALLDQYNSQ*PKNAIHTYVQAGSHIAAKGKANL          527
Hum  489  Y*SH*QALLDKYNAE*PKNAIHTFVQSGSHLAAREKANL          527
Bov  488  Y*SR*QALLDKYNEE*PKN                               506
Dro  490  F*RHL-T*E*N**--A*---SSKF                          506
Mz1  478  **QKL-PSR*N*K*SM                                  492
Swp  478  **QKV-ASR*NIR*TM                                  492
Ctn  478  V*QKL-ASL*NVR*SI                                  492
```

Figure 4.4.5. Comparison of the deduced amino acid sequence of Cat with those of other catalases. Ctr, C. tropicalis Cat (from Figure 4.4.2; Okada et al., 1987); ScA, S. cerevisiae catalase A (Cohen et al., 1988); ScT, S. cerevisiae catalase T (Hartig and Ruis, 1986); Rat, rat liver catalase (Furuta et al., 1986); Hum, human lymphoblast catalase (Quan et al., 1986); Bov, bovine liver catalase (Schroeder et al., 1982); Dro, Drosophila catalase (Orr et al., 1990); Mz1, maize (Zea mays L.) CAT-1 (Redinbaugh et al., 1988); Swp, sweet potato catalase (Sakajo et al., 1987); Ctn, cottonseed catalase (Ni et al., 1990). Homologies between the polypeptides were identified using standard parameters for the multiple sequence alignment procedure of the Beckman MicroGenie (Version 7.0) Sequence Analysis Program (Queen and Korn, 1984). Amino acids are represented by the standard one-letter code. Gaps(-), inserted to maximize homology between sequences, are counted as mismatches. Asterisks(*) indicate amino acid identity between Cat and the other catalases. Those amino acids of bovine liver catalase (Schroeder et al., 1982) which participate directly in catalysis are denoted by ▼. Amino acids of bovine catalase which are in contact with the heme moiety are underlined; of these, those residues which are conserved in Cat are highlighted by #. Amino acids located on the proximal or distal side of the heme group of bovine catalase are denoted by +. Those amino acids constituting the NADP environment of bovine catalase are double-underlined; of these, those residues which are conserved in Cat are highlighted by ↓. A tripeptide variant of the PTS identified by Gould et al. (1989) is boxed (see text, section 4.5). Underlined amino acids of S. cerevisiae catalases A and T and human lymphoblast catalase may play a role in peroxisomal targeting (see text, section 4.5).

T (Hartig and Ruis, 1986), are the least homologous group, displaying amino acid identities ranging between 38.5% (S. cerevisiae catalase A vs. T) and 63.2% (C. tropicalis Cat vs. S. cerevisiae catalase A). Comparing Cat from C. tropicalis to each of the catalases from these species revealed significant amino acid identity with the mammalian (48.3, 48.0 and 50.1% with rat liver, human lymphoblast and bovine liver catalase, respectively), plant (40.0, 39.1 and 39.9% with maize CAT-1, sweet potato and cottonseed catalase, respectively) and yeast catalases (63.2 and 44.2% with S. cerevisiae catalases A and T, respectively), as well as 53.4% amino acid identity with Drosophila catalase (Orr et al., 1990). Moreover, the hydropathy profiles of each of these catalases, constructed according to the algorithm of Kyte and Doolittle (1982) using appropriate gaps to maximize homology, were quite similar, reflecting a consensus higher order structure amongst these polypeptides. Thus, Cat displayed significant overall homology with each of the other catalases, implying that these genes may have diverged, to a greater or lesser extent, from a common ancestral gene. Nevertheless, the degree of homology observed between Cat and each of the other catalases was not constant throughout the entire amino acid sequence. Generally, N-terminal portions of these proteins exhibited greater amino acid identities than C-terminal portions,

similar to that observed upon comparison of several different peroxisomal acyl-coenzyme A oxidases (section 3.2.3). In this respect, comparison of the N-terminal two-thirds of each of these catalases (ie. from the first amino acid of each protein to the residue aligned with asn-354 of Cat, Figure 4.4.5) to a consensus sequence generated by multiple sequence alignment within this region (consensus was defined as occurrence in a minimum of 5 of the 10 catalases) revealed between 46.7% (for S. cerevisiae catalase T) and 70.5% (for bovine liver catalase) amino acid identity. Conversely, comparison of the C-terminal one-third of each of these catalases (ie. from each residue aligned with cys-355 of Cat to the C-terminal amino acid of each protein, Figure 4.4.5) to a consensus sequence generated in a similar manner for this region revealed between 16.7% (for S. cerevisiae catalase T) and 37.9% (for bovine liver catalase) amino acid identity only. Within both of these regions, however, the extreme N-terminal (ie. from the first amino acid of each protein to the residue aligned with phe-4 of Cat, Figure 4.4.5) and C-terminal (ie. from each residue aligned with glu-477 of Cat to the C-terminal amino acid of each protein, Figure 4.4.5) portions of these catalases were highly divergent. As best exemplified by Cat from C. tropicalis, which is comprised of the least number of amino acids (485), this resulted from

the "elimination" of the extreme N- and C-terminal residues of the larger proteins [ie. S. cerevisiae catalase T (562 residues) and rat liver and human lymphoblast catalases (527 residues each)], with the number of residues eliminated dependent upon the particular catalase (Figure 4.4.5).

Although the amino acid identity between Cat and bovine liver catalase was determined to be only 50.1%, the amino acid residues essential for the structure, function and catalytic activity of bovine liver catalase were highly conserved in Cat. Those amino acids of the bovine liver catalase sequence (his-74, ser-113 and asn-147, Figure 4.4.5, denoted by ▼) which have been suggested to participate directly in catalysis (Fita and Rossman, 1985) were conserved not only in Cat from C. tropicalis (his-53, ser-9? and asn-126, Figures 4.2.2, 4.4.2 and 4.4.5), but in each of the other catalases as well (Figure 4.4.5, denoted by ▼). Of the 39 amino acids proposed to be in contact with the heme moiety of each bovine liver catalase subunit (Murthy et al., 1981), including residues numbered 60, 63, 64, 71 to 74, 111, 113, 115, 116, 126 to 128, 130, 132, 143, 145 to 147, 152, 153, 157, 160, 163, 164, 167, 168, 176, 198, 210, 349, 352, 353, 356, 357, 360, 361 and 364 (Figure 4.4.5, underlined), 31 were conserved in Cat (Figures 4.2.2 and 4.4.2, underlined amino acids; Figure 4.4.5, highlighted by #). Relative to the near perfect conservation observed between the mammalian catalases in this respect (the only

discrepancy was pro-157 of human lymphoblast catalase), 36 amino acids of Drosophila catalase, 28 amino acids of the plant catalases (considered collectively) and 27 and 26 amino acids of S. cerevisiae catalases T and A, respectively, corresponded with those residues of bovine liver catalase in contact with the heme ligand (Figure 4.4.5). Amino acids located on the proximal or distal side of the heme group in bovine liver catalase [pro-335, arg-353 and tyr-357 (proximal); val-73, thr-114, phe-152 and phe-160 (distal), Figure 4.4.5, denoted by +], which have been postulated to be important for the structure and function of the active site (Murthy et al., 1981; Fita and Rossman, 1985), were similarly conserved in Cat and each of the other catalases (except for pro-69 of S. cerevisiae catalase A) (Figure 4.4.5, denoted by +). A high affinity NADP-binding site has also been shown to exist in mammalian catalases (Kirkman and Gaetani, 1984; Fita and Rossman, 1985). Of the 13 amino acids constituting the NADP environment of bovine liver catalase (Fita and Rossman, 1985), including his-193, phe-197, ser-200, arg-202, tyr-214, lys-236, val-301, trp-302, his-304, gln-441, thr-444, phe-445 and val-449 (Figure 4.4.5, double-underlining), 7 residues were conserved in Cat (Figure 4.4.5, highlighted by arrows). For the remaining catalase sequences, conservation of the NADP environment native to bovine liver catalase was highest among the

mammalian enzymes (69.2 and 92.3% for rat liver and human lymphoblast catalase, respectively), less for the yeast (38.5 and 61.5% for S. cerevisiae catalases T and A, respectively) and Drosophila enzymes (61.5%) and least for the plant enzymes (30.8% for maize CAT-1 and cottonseed catalase; 15.4% for sweet potato catalase) (Figure 4.4.5).

The significant degrees of homology recognized between several domains of each of the mammalian, yeast, plant and Drosophila catalases infer the following conclusions. First, the catalytically relevant regions of each of these enzymes have maintained a high degree of sequence conservation throughout evolution, while other, less important, domains (ie. the extreme N- and C-termini) have diverged significantly. Consequently, the folding of each catalase subunit has been conserved over the majority of the polypeptide chain. This further implies that, similar to bovine liver catalase (Murthy et al., 1981; Fita and Rossman, 1985), each of these catalase subunits possesses a tertiary structure comprised of four distinct functional domains: a semi-conserved, extended (non-globular), N-terminal "arm" (the length of which is species-specific), which stabilizes the tetrameric quaternary structure (residues 1 to 75 of bovine catalase); a highly conserved, β-barrel domain, which provides the residues on the distal side of the heme moiety (residues 76 to 320 of

bovine catalase); a highly conserved, subunit "wrapping" domain, which is catalytically essential due to the presence of the proximal heme ligand, tyr-357 (residues 321 to 436 of bovine catalase); and a least conserved, C-terminal α-helical domain, which helps form a hydrophobic channel limiting access to the active site (residues 437 to 506 of bovine catalase). Therefore, each of these catalases, from diverse species, should function according to the same basic mechanism.

4.4.4 Codon Usage in the Gene Encoding Cat

In general, the codon composition of unicellular organisms reflects the base composition of the genome; however, there are subsets of genes from both unicellular prokaryotic (E. coli) and eukaryotic (S. cerevisiae) organisms for which codon usage is strongly biased toward a group of "optimal" or "preferred" codons (for reviews, see Bennetzen and Hall, 1982b; Gouy and Gautier, 1982; Ikemura, 1985; Holm, 1986; Sharp et al., 1986; Bulmer, 1987; Andersson and Kurland, 1990). Representative of such genes are those encoding the outer membrane proteins [lpp (Nakamura et al., 1980) and ompA (Movva et al., 1980)] and elongation factors [tufA (Yokota et al., 1980) and tufB (An and Friesen, 1980)] of E. coli and the glycolytic and gluconeogenic enzymes of S. cerevisiae [HXK1 (Kopetzki et al., 1985), HXK2 (Fröhlich et al., 1985), G3PDH-I (Holland

and Holland, 1979), G3PDH-II (Holland and Holland, 1980), ADH-I (Bennetzen and Hall, 1982a), ADH-II (Russell et al., 1983), E-I (Holland et al., 1981), E-II (Holland et al., 1981), PK (Burke et al., 1983) and PGK (Hitzeman et al., 1982)]. For these genes, from these organisms, the following characteristics of codon usage have been established (see aforementioned reviews): (1) choice between synonymous codons is biased for most amino acids, with easily discernable similarities of choice existing among genes of each organism, (2) the degree of nonrandom codon usage (ie. the extent of codon bias) for individual genes of either organism is correlated with the level of expression (or protein production) of that gene. Thus, highly expressed genes, which exhibit extreme codon bias, use one or a few synonymous codons to the nearly complete exclusion of all others. Alternatively, moderately and lowly expressed genes exhibit the same trend in codon bias (ie. preferred codons are still preferred), however, the extent is moderated in that synonymous codons are used at significant levels, (3) those genes which are expressed to a high degree are strongly biased toward those codons which correspond to the major tRNA isoacceptor species of that organism and (4) for certain amino acids, codon choices are different between organisms; this organism-specific codon choice has been termed the "dialect" of that organism.

Relative to the characteristics of these genes, codon usage in the gene/cDNA encoding Cat from C. tropicalis was similarly nonrandom (Table 4.4.2). Several codons [17 (for the gene)/18 (for the cDNA) of 61 possible], including GCG for ala, CGT, CGC, CGA, CGG and AGG for arg, GAG for glu, GGG for gly, ATA for ile, CTT, CTA, CTG (Cat cDNA only) and TTA for leu, CCC and CCG (Cat cDNA only) for pro, TCG and AGC (Cat gene only) for ser, ACG for thr and GTA for val, were never used. Other codons [10 (for the gene)/9 (for the cDNA) of 61 possible], including GCA for ala, CAG for gln, GGA for gly, CTC and CTG (Cat gene only) for leu, CCG (Cat gene only) for pro, TCA, AGT and AGC (Cat cDNA only) for ser, ACA for thr and GTG for val, were used to encode less than 10% of the total number of each corresponding amino acid in Cat. Consequently, only 23 of a possible 61 codons (note that met and trp codons are not considered to be preferential since they do not exhibit degeneracy) were preferentially used to encode 85.2% of the amino acids in Cat (Table 4.4.2, underlined codons). This percentage is significantly greater than that of the S. cerevisiae genes encoding HXK1 (71.1%), HXK2 (79.0%) and ADH-II (79.6%), comparable to that of the E. coli genes encoding ompA (84.4%), tufB (85.5%), lpp (85.9%) and tufA (87.3%) and significantly less than that of the S. cerevisiae genes encoding ADH-I (91.4%), PGK (92.0%), PK

Table 4.4.2. Codon usage in the cDNA and gene encoding Cat.

Amino Acid	Codon[a]	Number of codons used[b] (%) cDNA	genomic DNA	Amino Acid	Codon[a]	Number of codons used[b] (%) cDNA	genomic DNA
Ala	GCT	27 (5.6)	27 (5.6)	Lys	AAA	4 (0.8)	5 (1.0)
	GCC	13 (2.7)	13 (2.7)		AAG	24 (4.9)	23 (4.7)
	GCA	1 (0.2)	1 (0.2)	Met	ATG	9 (1.9)	9 (1.9)
	GCG	0 (0.0)	0 (0.0)	Phe	TTT	10 (2.1)	10 (2.1)
Arg	CGT	0 (0.0)	0 (0.0)		TTC	22 (4.5)	22 (4.5)
	CGC	0 (0.0)	0 (0.0)	Pro	CCT	8 (1.6)	8 (1.6)
	CGA	0 (0.0)	0 (0.0)		CCC	0 (0.0)	0 (0.0)
	CGG	0 (0.0)	0 (0.0)		CCA	24 (4.9)	23 (4.7)
	AGA	20 (4.1)	20 (4.1)		CCG	0 (0.0)	1 (0.2)
	AGG	0 (0.0)	0 (0.0)	Ser	TCT	9 (1.9)	6 (1.2)
Asn	AAT	4 (0.8)	4 (0.8)		TCC	13 (2.7)	15 (3.1)
	AAC	25 (5.2)	25 (5.2)		TCA	1 (0.2)	2 (0.4)
Asp	GAT	5 (1.0)	5 (1.0)		TCG	0 (0.0)	0 (0.0)
	GAC	23 (4.7)	23 (4.7)		AGT	1 (0.2)	1 (0.2)
Cys	TGT	2 (0.4)	2 (0.4)		AGC	1 (0.2)	0 (0.0)
	TGC	1 (0.2)	1 (0.2)	Thr	ACT	11 (2.3)	10 (2.1)
Gln	CAA	20 (4.1)	20 (4.1)		ACC	22 (4.5)	24 (4.9)
	CAG	1 (0.2)	1 (0.2)		ACA	1 (0.2)	1 (0.2)
Glu	GAA	29 (6.0)	29 (6.0)		ACG	0 (0.0)	0 (0.0)
	GAG	0 (0.0)	0 (0.0)	Trp	TGG	8 (1.6)	8 (1.6)
Gly	GGT	22 (4.5)	24 (4.9)	Tyr	TAT	4 (0.8)	3 (0.6)
	GGC	5 (1.0)	3 (0.6)		TAC	14 (2.9)	15 (3.1)
	GGA	2 (0.4)	2 (0.4)	Val	GTT	11 (2.3)	10 (2.1)
	GGG	0 (0.0)	0 (0.0)		GTC	19 (3.9)	19 (3.9)
His	CAT	3 (0.6)	3 (0.6)		GTA	0 (0.0)	0 (0.0)
	CAC	19 (3.9)	19 (3.9)		GTG	1 (0.2)	1 (0.2)
Ile	ATT	5 (1.0)	6 (1.2)				
	ATC	11 (2.3)	11 (2.3)				
	ATA	0 (0.0)	0 (0.0)	Sum aa[c]		485.0	485.0
Leu	CTT	0 (0.0)	0 (0.0)	Mf aa[d]		413.0	413.0
	CTC	2 (0.4)	2 (0.4)	% Mf aa[e]		85.2	85.2
	CTA	0 (0.0)	0 (0.0)				
	CTG	0 (0.0)	1 (0.2)				
	TTA	0 (0.0)	0 (0.0)				
	TTG	28 (5.8)	27 (5.6)				

[a] The 23 preferentially used codons in the Cat gene are underlined. Met and trp codons, which show no degeneracy, were not considered to be preferentially used codons.

[b] Numbers in parentheses represent the percentage of all amino acids in Cat encoded by a particular codon.

[c] Sum aa, total number of amino acids in Cat (including the initiator methionine).

[d] Mf aa, total number of amino acids encoded by the 23 preferentially used codons.

[e] % Mf aa, percentage of Mf aa as related to Sum aa.

(92.2%), E-I (92.4%), E-II (93.3%), G3PDH-II (93.5%) and G3PDH-I (94.5%) (see Table 4.4.5). Thus, consistent with these E. coli and S. cerevisiae genes, the high degree of preferential codon usage observed in the Cat gene may be directly correlated with its level of expression. Moreover, a codon bias index (see below for definition) of 0.80 calculated for the Cat gene (Table 4.4.5), compared to values of 0.11 and 0.31 calculated for the S. cerevisiae CTA1 (Cohen et al., 1988) and CTT1 genes (Hartig and Ruis, 1986), respectively (Table 4.4.5), is suggestive of a significantly higher level of expression of the Cat gene in C. tropicalis relative to that of the CTA1 and CTT1 genes in S. cerevisiae.

As previously discussed in section 3.2.4, for several E. coli and S. cerevisiae genes, there is a strong bias toward the translation termination codon TAA (over TGA and TAG) where sense codon usage is highly biased (Sharp and Bulmer, 1988). Since the ORF of the Cat gene/cDNA, like that of the AOx gene (section 3.2.4), is terminated by a TAA codon, the degree of sense codon usage in C. tropicalis genes may similarly be correlated with selective differences among translation termination codons (discussed in greater detail below).

To further compare and/or contrast codon usage in C. tropicalis with that of E. coli and S. cerevisiae, codon

usage in the AOx (section 3.2.4) and Cat genes was compared with that of several genes from C. tropicalis. Any generalities established concerning codon usage in C. tropicalis were subsequently related to those determined for the several E. coli and S. cerevisiae genes mentioned above.

Accordingly, Table 4.4.3 presents the codon selectivity exhibited by several genes of C. tropicalis. Consistent with the codon usage profiles determined for the AOx (section 3.2.4) and Cat genes (or Cat cDNA), codon usage in each of the C. tropicalis genes listed in Table 4.4.3 is distinctly biased. Twenty-eight codons (of 61 possible), including GCA and GCG for ala, CGT, CGC, CGA, CGG and AGG for arg, TGC for cys, CAG for gln, GAG for glu, GGC, GGA and GGG for gly, ATA for ile, CTT, CTC, CTA and CTG for leu, CCT and CCG for pro, TCA, TCG, AGT and AGC for ser, ACA and ACG for thr and GTA and GTG for val, are used to encode less than 10% of the total number of each corresponding amino acid in all of the genes combined. For those genes encoding peroxisomal proteins, codon usage is highly biased, with the same subset of 23 preferred codons identified for the AOx and Cat genes encoding 82.7% (AOx; section 3.2.4), 85.2% (Cat; Okada et al., 1987), 85.2% (POX2; Okazaki et al., 1987), 86.6% (POX4; Okazaki et al., 1986), 84.4% (POX5; Okazaki et al., 1986), 85.0% (POX18; Szabo et al., 1989; Tan et al., 1990), 86.0% (HDE, Nuttley et al., 1988) and 88.9%

Table 4.4.3. Comparison of the codon selectivity exhibited by several genes of Candida tropicalis.

Amino Acid	Codon[a]	Genes[b]											
		AOx	Cat (cDNA)	Cat	POX2	POX4	POX5	POX18	HDE	Isocitrate lyase	P_{450}Alk	CPR	14DH
		Number of codons used											
Ala	GCT	28	27	27	27	32	23	4	37	40	20	23	23
	GCC	23	13	13	20	28	27	10	47	23	3	11	3
	GCA	0	1	1	5	0	1	0	5	2	7	9	1
	GCG	0	0	0	0	0	1	0	0	0	2	0	0
Arg	CGT	0	0	0	5	0	5	0	3	1	2	2	2
	CGC	0	0	0	2	0	0	0	0	0	0	0	4
	CGA	0	0	0	0	0	0	0	0	0	0	0	0
	CGG	0	0	0	0	0	0	0	0	0	0	0	0
	AGA	36	20	20	36	34	29	1	23	22	24	28	24
	AGG	2	0	0	1	0	0	0	1	0	1	0	0
Asn	AAT	5	4	4	4	3	2	0	4	0	11	27	5
	AAC	27	25	25	27	28	26	9	48	18	7	12	15
Asp	GAT	12	5	5	2	11	5	3	22	7	21	44	28
	GAC	33	23	23	45	39	36	7	44	27	5	9	6
Cys	TGT	8	2	2	9	8	8	0	1	3	4	4	5
	TGC	0	1	1	0	0	1	0	1	0	0	0	0
Gln	CAA	29	20	20	23	28	18	4	23	22	23	20	17
	CAG	2	1	1	9	2	5	0	1	2	2	1	0
Glu	GAA	31	29	29	32	34	34	5	56	37	32	51	28
	GAG	4	0	0	11	3	9	1	2	2	8	2	0
Gly	GGT	43	22	24	46	43	46	7	68	33	22	32	28
	GGC	5	5	3	1	5	1	1	3	2	7	2	1
	GGA	0	2	2	1	0	2	1	2	1	3	10	1
	GGG	0	0	0	1	0	0	0	0	0	0	1	0
His	CAT	0	3	3	2	1	1	0	1	0	7	8	9
	CAC	10	19	19	13	9	12	0	16	15	5	5	5
Ile	ATT	20	5	6	26	19	17	1	25	10	24	27	17
	ATC	21	11	11	23	25	19	4	31	26	6	6	9
	ATA	2	0	0	0	1	1	0	0	0	1	1	2
Leu	CTT	1	0	0	2	1	3	0	2	0	12	6	3
	CTC	4	2	2	3	5	5	0	1	0	2	0	0
	CTA	1	0	0	0	0	0	0	0	0	0	0	0
	CTG	3	0	1	0	0	1	0	0	0	0	0	0
	TTA	2	0	0	2	0	1	2	2	0	20	21	17
	TTG	53	28	27	64	56	48	12	61	37	34	33	30
Lys	AAA	4	4	5	5	5	6	2	16	8	18	31	21
	AAG	38	24	23	46	40	37	20	55	38	18	17	15
Met	ATG	17	9	9	13	16	17	4	13	12	8	5	16
Phe	TTT	13	10	10	7	12	9	2	15	6	26	26	13
	TTC	16	22	22	18	16	24	3	27	14	10	18	19

Amino Acid	Codon[a]	Genes[b]											
		AOx	Cat (cDNA)	Cat	POX2	POX4	POX5	POX18	HDE	Isocitrate lyase	P_{450}Alk	CPR	14DM
		Number of codons used											
Pro	CCT	7	8	8	2	5	3	0	3	1	2	4	3
	CCC	1	0	0	0	0	0	0	0	1	1	0	0
	CCA	21	24	23	27	18	23	4	33	17	20	20	24
	CCG	3	0	1	1	0	0	0	1	0	1	0	0
Ser	TCT	16	9	6	12	13	4	0	16	9	14	13	15
	TCC	26	13	15	24	28	25	1	20	19	3	10	8
	TCA	4	1	2	1	3	1	1	3	0	2	10	3
	TCG	2	0	0	2	1	3	0	2	0	6	0	0
	AGT	3	1	1	1	3	4	0	3	1	1	9	6
	AGC	3	1	0	0	3	3	1	3	0	1	2	1
Thr	ACT	15	11	10	21	15	8	0	22	9	20	27	21
	ACC	23	22	24	20	26	32	3	31	23	10	12	11
	ACA	5	1	1	2	3	0	0	1	1	6	3	0
	ACG	2	0	0	1	0	0	0	0	0	1	1	0
Trp	TGG	10	8	8	10	8	9	1	11	12	6	7	9
Tyr	TAT	2	4	3	8	2	3	0	3	1	18	16	16
	TAC	24	14	15	22	26	22	0	30	21	7	6	12
Val	GTT	19	11	10	23	23	16	3	33	13	17	38	33
	GTC	26	19	19	13	26	25	10	32	13	7	6	3
	GTA	1	0	0	0	1	0	0	0	0	3	2	0
	GTG	3	1	1	3	1	1	0	3	1	2	2	0
Sum aa[c]		709	485	485	724	709	662	127	906	550	543	680	528
Hf aa[d]		586	413	413	617	614	559	108	779	489	335	428	371
% Hf aa[e]		82.7	85.2	85.2	85.2	86.6	84.4	85.0	86.0	88.9	61.7	62.9	70.3

[a] The 23 preferentially used codons in the AOx and Cat genes are underlined. Met and trp codons, which show no degeneracy, were not considered to be preferentially used codons.

[b] C. tropicalis genes included are: AOx (Table 3.2.2), Cat (Table 4.4.2; Okada et al., 1987), POX2 (Okazaki et al., 1987), POX4 (Okazaki et al., 1986), POX5 (Okazaki et al., 1986), POX18 (Szabo et al., 1989; Tan et al., 1990), HDE (Nuttley et al., 1988), isocitrate lyase (Atomi et al., 1990), P_{450}Alk (Sanglard and Loper, 1989), CPR (Sutter et al., 1990) and 14DM (Chen et al., 1988).

[c] Sum aa, total number of amino acids encoded by each gene (including the initiator methionine).

[d] Hf aa, total number of amino acids encoded by the 23 preferentially used codons.

[e] % Hf aa, percentage of Hf aa as related to Sum aa.

(isocitrate lyase; Atomi et al., 1990) of the total number of amino acids comprising each of these proteins (Table 4.4.3). By comparison, the remaining C. tropicalis genes, including three members of the cytochrome P450 family of C. tropicalis proteins, display less bias in codon usage, with the previously defined subset of preferred codons encoding only 61.7%, 62.9% and 70.3% of the total number of amino acids comprising cytochrome P_{450} (P_{450}Alk; Sanglard and Loper, 1989), NADPH-cytochrome P-450 oxidoreductase (CPR; Sutter et al., 1990) and cytochrome P450 lanosterol 14α-demethylase (14DM; Chen et al., 1988), respectively (Table 4.3.3). This decreased codon bias results from a "redistribution" of codon usage within these genes relative to those encoding the peroxisomal proteins. As seen in Table 4.4.3, six formerly preferred codons are replaced by different "optimal" codons for that amino acid, including AAT for AAC (encoding asn; for the P_{450}Alk and CPR genes only), GAT for GAC (encoding asp), CAT for CAC (encoding his), AAA for AAG (encoding lys), TTT for TTC (encoding phe) and TAT for TAC (encoding tyr). Further, the degree of codon bias for other amino acids, most notably leucine, but also alanine and serine, is lessened by the increased use of synonymous degenerate codons (Table 4.4.3).

Interestingly, the 23 preferred codons common to the C. tropicalis genes encoding peroxisomal proteins are

identical to the subset of 23 preferred codons determined initially for three highly expressed genes of S. cerevisiae, including G3PDH-I, G3PDH-II and ADH-I (Table 4.4.4; Bennetzen and Hall, 1982b), and later confirmed by the analysis of codon usage among 110 highly and lowly expressed yeast genes, predominantly of S. cerevisiae (Sharp et al., 1986). It should be noted, however, that the extent of the use of preferred codons in the S. cerevisiae genes is higher than that exhibited by the C. tropicalis genes, except for the preferred triplet encoding aspartic acid (Table 4.4.4). Alternatively, the codon selectivity exhibited by the three cytochrome P450-related genes is representative of that subset of "preferred" codons characteristic of the lowly expressed genes of S. cerevisiae (not shown; Sharp et al., 1986). By comparison, four highly expressed genes of E. coli, including ompA (Movva et al., 1980), lpp (Nakamura et al., 1980), tufA (Yokota et al., 1980) and tufB (An and Friesen, 1980), display a codon usage profile which is highly biased toward, but not identical to, the same subset of preferred codons established for C. tropicalis and S. cerevisiae (Table 4.4.4). The differences in codon choice (ie. dialect) are, for the most part, limited to six amino acids for which a completely different codon from that used in the highly expressed S. cerevisiae genes is preferred to encode the same amino acid in E. coli

[including arginine (CGT for AGA), cysteine (TGC for TGT), glutamine (CAG for CAA), leucine (CTG for TTG), lysine (AAA for AAG) and proline (CCG for CCA) (Table 4.4.4)]. Relative to C. tropicalis and S. cerevisiae, the extent of the use of preferred codons for a particular amino acid in E. coli is codon-specific, ranging from a greater bias for some codons (eg. GAC for asp), to an intermediate bias (of S. cerevisiae> E. coli> C. tropicalis) for others (eg. AAC for asn and TTC for phe), to yet a lesser bias for different codons (eg. GAA for glu and TAC for tyr).

Since the codon preferences manifested by the genes encoding peroxisomal proteins parallel those established for the highly expressed genes of S. cerevisiae, two immediate inferences are possible. First, because the preferred codons of S. cerevisiae are highly homologous to the major yeast isoacceptor tRNAs (Bennetzen and Hall, 1982b), C. tropicalis most likely uses the same major isoacceptor tRNAs for the expression of its highly expressed genes. Second, the empirical rules established by Bennetzen and Hall (1982b) for preferred codon usage in S. cerevisiae should apply, with slight modifications, to preferred codon usage in C. tropicalis. These rules may be summarized as follows: 1) for serine and four of the six amino acids exhibiting 3- or 4-fold coding [illegible] acy (including ala, ile, thr and val), NNT and NNC are used with approximately equal

Table 4.4.4. Comparison of preferred codon usage between the AOx and Cat genes of *Candida tropicalis* and highly expressed genes of *Saccharomyces cerevisiae* (G3PDH-I, G3PDH-II and ADH-I) and *Escherichia coli* (ompA, lpp, tufA and tufB).

Amino Acid	Preferred codon(s) and the extent of their use in[a]			Use of all other codons for that amino acid in[b]			Extent of the use of preferred codon(s) in[c]		
	C. tropicalis AOx and Cat[d]	*S. cerevisiae* G3PDH-I, II and ADH-I[e]	*E. coli* ompA, lpp, tufA and tufB[f]	AOx and Cat[d]	G3PDH-I,II and ADH-I[e]	ompA, lpp, tufA and tufB[f]	AOx and Cat[d]	G3PDH-I,II and ADH-I[e]	ompA, lpp, tufA and tufB[f]
Ala	GCT(55), GCC(36)	GCT(68), GCC(32)	GCT(63); NCP[g]	1	0	49	98.9	100.0	56.3
Arg	AGA(56)	AGA(30)	CGT(54)	2	0	9	96.6	100.0	85.7
Asn	AAC(52)	AAC(37)	AAC(38)	9	0	1	85.2	100.0	97.4
Asp	GAC(56)	GAC(44)	GAC(64)	17	17	15	76.7	72.1	81.0
Cys	TGT(10)	TGT(12)	TGC(6)	1	0	3	90.9	100.0	66.7
Gln	CAA(49)	CAA(20)	CAG(36)	3	0	2	94.2	100.0	94.7
Glu	GAA(60)	GAA(49)	GAA(70)	4	0	16	93.8	100.0	81.4
Gly	GGT(67)	GGT(90)	GGT(64), GGC(56)	10	3	7	87.0	96.8	94.5
His	CAC(29)	CAC(26)	CAC(22)	3	1	5	90.6	96.3	81.5
Ile	ATT(26), ATC(32)	ATT(25), ATC(35)	ATC(69)	2	0	7	96.7	100.0	90.8
Leu	TTG(80)	TTG(60)	CTG(82)	14	6	4	85.1	90.1	95.3
Lys	AAG(61)	AAG(69)	AAA(35)	9	7	11	87.1	90.8	76.1
Phe	TTC(38)	TTC(29)	TTC(33)	23	0	4	62.3	100.0	89.2
Pro	CCA(44)	CCA(32)	CCG(53)	20	4	6	71.0	88.9	89.8
Ser	TCT(22), TCC(41)	TCT(38), TCC(33)	TCT(21), TCC(16)	15	0	7	80.8	100.0	84.1
Thr	ACT(25), ACC(47)	ACT(77), ACC(34)	ACT(39), ACC(43)	8	0	-	90.0	100.0	94.3
Tyr	TAC(39)	TAC(33)	TAC(33)	5	0	5	88.6	100.0	86.8
Val	GTT(29), GTC(45)	GTT(64), GTC(44)	GTT(66), GTA(30)	5	0	11	93.7	100.0	89.7

[a] The 23 preferentially used codons in the AOx and Cat genes are listed. The codon usage figures represent the sum of the use of these codons in the indicated genes (see footnotes d, e and f).

[b] The codon usage figures represent the sum of the use of all other (i.e. non-preferred) codons for each amino acid in the indicated genes.

[c] Preferred codon(s) used, relative to total codons used, for a particular amino acid in the indicated genes, expressed as a percentage.

[d] *C. tropicalis* genes included are: AOx (Table 3.2.2) and Cat (Table 4.4.2; Okada et al., 1987).

[e] *S. cerevisiae* genes included are: G3PDH-I (Holland and Holland, 1979), G3PDH-II (Holland and Holland, 1980) and ADH-I (Bennetzen and Hall, 1982a).

[f] *E. coli* genes included are: ompA (Movva et al., 1980), lpp (Nakamura et al., 1980), tufA (Yokota et al., 1980) and tufB (An and Friesen, 1980).

[g] NCP= no clear preference between other codons.

probability, whereas NNA and NNG are avoided (ie. purines in the wobble position are avoided), 2) for the two amino acids exhibiting 4-fold coding degeneracy which do not obey rule 1 (including gly and pro), the predominant codon choice(s) is that which prevents the codon from being either 100% GC, 100% purine or 100% pyrimidine (ie. GGT for gly and CCA for pro), 3) for five of the six amino acids exhibiting 2-fold coding degeneracy with a pyrimidine in the wobble position (including asn, asp, his, phe and tyr), NNC is preferred over NNT. The sixth amino acid exhibiting 2-fold coding degeneracy with a pyrimidine in the wobble position (cysteine), reverses the codon preference (ie. NNT is preferred over NNC) to avoid side-by-side GC base pairs in the codon-anticodon interaction (Bennetzen and Hall, 1982b), 4) for leucine, arginine and the three amino acids exhibiting 2-fold coding degeneracy with a purine in the wobble position (including gln, glu and lys), one degenerate codon is used to the almost complete exclusion of the other(s) (ie. AGA for arg, CAA for gln, GAA for glu, TTG for leu and AAG for lys). To further this comparison of codon usage between C. tropicalis and S. cerevisiae, that set of preferred codons inferred from the codon usage profiles of several C. tropicalis genes encoding peroxisomal proteins was used to calculate the codon bias index (CBI) of these genes, in a manner similar to that previously conducted for

many S. cerevisiae and E. coli genes (Bennetzen and Hall, 1982b; Sharp et al., 1986). By definition, the CBI measures the distance between any given mRNA sequence and the "preferred" mRNA sequence that could code for that particular protein (ie. the fraction of codon choices which is biased toward the preferred subset of codons) (Bennetzen and Hall, 1982b). Accordingly, a value of one indicates that only preferred codons are used, whereas a value of zero implies totally random codon choice. As seen in Table 4.4.5, which presents such an analysis for several genes from C. tropicalis, S. cerevisiae and E. coli, the higher the level of expression of a S. cerevisiae or E. coli gene [with G3PDH-I and tufA (or lpp), respectively, being the most highly expressed], the greater the CBI, consistent with the definition of preferred codons being the most used in these genes. Conversely, the lower the level of expression of a S. cerevisiae or E. coli gene (with CTA1 and lacI, respectively, being the least expressed), the lower the CBI, implying that codon choice in these genes is nearly random. By comparison, it may be speculated that C. tropicalis genes encoding peroxisomal proteins, which yield relatively high values for the CBI (ranging from 0.78 for AOx to 0.89 for HDE, Table 4.4.5), are phenotypically highly expressed genes, whereas the cytochrome P450-related genes from C. tropicalis, which yield lower, but significantly greater

Table 4.4.5. The extent of preferred codon usage and the codon bias index determined for several genes from Candida tropicalis, Saccharomyces cerevisiae and Escherichia coli.

	Gene[a]	Extent of preferred codon usage (% MF aa)[b]	Codon bias index (CBI)[c]
C. tropicalis	Isocitrate lyase	88.9	0.89
	POX4	85.6	0.84
	HDE	86.0	0.83
	POX18	85.0	0.83
	Cat	85.2	0.80
	POX2	85.2	0.79
	POX5	84.4	0.79
	AOx	82.7	0.78
	14DM	70.3	0.62
	CPR	62.9	0.48
	P_{450} Alk	61.7	0.42
S. cerevisiae	G3PDH-I	94.5	0.99
	G3PDH-II	93.5	0.98
	E-I	92.4	0.96
	PK	92.2	0.95
	E-II	93.3	0.93
	ADH-I	91.4	0.92
	PGK	92.0	0.91
	HXK2	79.0	0.73
	ADH-II	79.6	0.71
	HXK1	71.7	0.61
	Iso-1 cytochrome c	68.2	0.50
	CTT1	58.5	0.31
	Iso-2 cytochrome c	47.8	0.15
	CTA1	45.8	0.11
E. coli	tufA	87.3	0.84
	lpp	85.9	0.84
	tufB	85.5	0.81
	ompA	84.4	0.78
	RNA polymerase β	67.7	0.53
	lacI	48.6	0.18

[a] Genes included are: isocitrate lyase (Atomi et al., 1990), POX4 (Okazaki et al., 1986), HDE (Nuttley et al., 1988), POX18 (Szabo et al., 1989; Tan et al., 1990), Cat (Table 4.4.2; Okada et al., 1987), POX2 (Okazaki et al., 1987), POX5 (Okazaki et al., 1986), C. tropicalis AOx (Table 3.2.2), 14DM (Chen et al., 1988), CPR (Sutter et al., 1990), P_{450}Alk (Sanglard and Loper, 1989); S. cerevisiae G3PDH-I (Holland and Holland, 1979), G3PDH-II (Holland and Holland, 1980), E-I (Holland et al., 1981), PK (Burke et al., 1983), E-II (Holland et al., 1981), ADH-I (Bennetzen and Hall, 1982a), PGK (Hitzeman et al., 1982), HXK2 (Fröhlich et al., 1985), ADH-II (Russell et al., 1983), HXK1 (Kopetzki et al., 1985), iso-1 cytochrome c (Smith et al., 1979), CTT1 (Hartig and Ruis, 1986), iso-2 cytochrome c (Montgomery et al., 1980), CTA1 (Cohen et al., 1988); and E. coli tufA (Yokota et al., 1980), lpp (Nakamura et al., 1980), tufB (An and Friesen, 1980), ompA (Movva et al., 1980), RNA polymerase β (Delcuve et al., 1980), lacI (Farabaugh, 1978).

[b] See Table 4.4.3, footnote e.

[c] See section 4.4.4 for definition. See Bennetzen and Hall (1982b) for calculation.

than zero, values for the CBI (ranging from 0.42 for P_{450}Alk to 0.62 for 14DM; Table 4.4.5), are phenotypically intermediately expressed genes. Thus, synonymous with S. cerevisiae and E. coli (Bennetzen and Hall, 1982b; Sharp et al., 1986), the extent of codon bias exhibited by the genes of C. tropicalis may be strongly correlated with their level of expression and ultimately, the corresponding intracellular mRNA and protein levels. Further, the degree of bias in sense codon usage in C. tropicalis may also be correlated with selective differences among translation termination codons. In this respect, seven of the eight genes encoding peroxisomal proteins terminate with a TAA codon (only POX2 terminates with a TAG codon) and two of the three cytochrome P450-related genes terminate with a TAA codon (only 14DM terminates with a TAG codon) (not shown). Therefore, paralleling that previously determined for several highly biased genes of S. cerevisiae and E. coli (Sharp and Bulmer, 1988), there may be a strong bias in favour of TAA as a translation termination codon (over TGA and TAG) in the highly expressed genes of C. tropicalis. [Interestingly, but possibly of no significant consequence, the majority of sequenced C. tropicalis genes (ie. eight of ten genes excluding P_{450}Alk and CPR; AOx was not determined (Figure 3.2.2)) possess in-frame, "secondary?", termination codons located at various positions downstream of the natural termination codon.]

Collectively, all of the aforementioned observations indicate that the selective pressures which dictate preferred codon bias in the highly expressed genes of S. cerevisiae and E. coli exert generally similar effects, dictating generally similar trends and/or consequences, in C. tropicalis as well.

4.5 PEROXISOMAL TARGETING OF CATALASES

Recently, a limited number of studies concerned with the in vivo expression and peroxisomal targeting of mammalian (ie. human) and yeast (ie. S. cerevisiae catalase A) catalases have been documented. Initially, Gould et al. (1988) determined that the C-terminal 27 amino acids of human catalase (Figure 4.4.5, underlined) contain sufficient information to target a heterologous protein (ie. a luciferase-catalase fusion protein) to the peroxisomes of CV-1 monkey kidney cells. Within this 27 amino acids, the sequence, ser-his-leu, is located 11 amino acids from the extreme C-terminus (Figure 4.4.5); interestingly, this sequence has since been included as a functional variant of the C-terminal tripeptide PTS, ser/ala/cys-lys/his/arg-leu, shown to be sufficient to address several different proteins to the peroxisomes of diverse eukaryotic systems (discussed in section 3.3; Gould et al., 1987, 1988, 1989, 1990a). It should be noted, however, that larger fusion proteins containing either the C-terminal 145 or 404 amino acids of

human catalase are not imported into the peroxisomes of CV-1 cells, even though they possess the identical C-terminal 27 amino acids of human catalase (Gould et al., 1988). At present, an explanation for this inconsistency is unknown.

To study the peroxisomal targeting of yeast catalases, Hansen and Roggenkamp (1989) initially isolated a catalase-defective mutant of the methylotrophic yeast Hansenula polymorpha which abolished peroxisome function. Subsequent to transformation of mutant cells with the S. cerevisiae catalase A gene (CTA1; Cohen et al., 1988), they demonstrated that the in vivo expression and heterologous transport of catalase A results in functional complementation of catalase-defective peroxisomes. From this, it was concluded that the proper localization of catalase activity is prerequisite to peroxisome function in H. polymorpha and that the import mechanisms in both H. polymorpha and S. cerevisiae are highly conserved despite the functional diversity of their respective peroxisome types (Hansen and Roggenkamp, 1989). In an effort to identify the PTS of S. cerevisiae catalase A, which possesses an extreme N-terminal, but no C-terminal, variant of the tripeptide PTS identified by Gould et al. (1989) (Figure 4.4.5), Hartig et al. (1990) determined the subcellular location of several in vivo expressed, catalase A-mouse DHFR fusion proteins. They demonstrated that

neither the N-terminal 126 amino acids (which included the ser-lys-leu tripeptide), nor the C-terminal 16 amino acids, of catalase A are sufficient to target the corresponding fusion proteins to S. cerevisiae peroxisomes. Alternatively, fusion proteins possessing 140, or more, N-terminal amino acids of catalase A pellet and copurify with peroxisomes upon sucrose or Nycodenz density gradient centrifugation. This indicates that a short internal sequence of 14 amino acids (residues 127 to 140; Figure 4.4.5, underlined) is responsible for the comigration of these fusion proteins with S. cerevisiae peroxisomes. Subsequent analysis by immunofluorescence and electron microscopy revealed, however, that the fusion proteins are not imported into peroxisomes but rather form large aggregates (or protein bodies), lacking a surrounding membrane, within the cytosolic space of the yeast cell (Hartig et al., 1990). Consequently, determination of the exact nature of the internal PTS of S. cerevisiae catalase A, or any of the other yeast catalases, has not been achieved.

Studies concerned with the in vitro expression and import of any of the sequenced mammalian, plant or Drosophila catalases (section 4.4.3) into peroxisomes isolated from these species have not been reported as yet. Recently, however, preliminary work concerning the in vitro

expression and translocation of catalase A into peroxisomes isolated from S. cerevisiae has been documented (Thieringer et al., 1991), although no investigation of the PTS of catalase A was conducted. Nevertheless, speculation regarding the identity of the PTSs of these catalase polypeptides has resulted from comparisons of their deduced amino acid sequences, especially upon comparison with the sequence of S. cerevisiae catalase T, a strictly cytosolic protein (Hartig and Ruis, 1986). In this respect, Okada et al. (1987) and Cohen et al. (1988) suggested that one or more of the major insertions present in the catalase T polypeptide [amino acids 370 to 396 and 415 to 431 (Figure 4.4.5, underlined)], which are absent from the peroxisomal catalases (Figure 4.5.5), may exert a negative effect on transport by disturbing the recognition of catalase T by peroxisomes and/or its translocation into the organelle. In addition, Cohen et al. (1988) proposed that two smaller regions (amino acids 70 to 81 and 265 to 273 of catalase T, Figure 4.4.5, underlined) common to certain peroxisomal catalases [including bovine liver catalase, S. cerevisiae catalase A and C. tropicalis PXP-9 (Cat)], but differing in S. cerevisiae catalase T, may act as topogenic signals for catalase import into peroxisomes. However, the aforementioned work of Hartig et al. (1990) negates any role for N-terminal amino acids 76 to 87 of catalase A (which

correspond to N-terminal amino acids 70 to 81 of catalase T, Figure 4.4.5) in peroxisomal targeting. Further, a more comprehensive comparison of amino acids 265 to 273 (of catalase T) between several different catalases (Figure 4.4.5) reveals decreased homology within this region for the plant and Drosophila catalases, as well as catalase T. Thus, the likelihood of these segments functioning as "universal" PTSs for catalase polypeptides is minimal. Similarly, a universal role for the C-terminal tripeptide PTS of Gould et al. (1989), ser/ala/cys/-lys/his/arg/-leu, in the targeting of catalases to peroxisomes appears unlikely. Even though the C-terminal 27 amino acids of human catalase, which contain the tripeptide ser-his-leu, have been shown to facilitate peroxisomal targeting (see above; Gould et al., 1988), and each of the plant catalases also possesses a related ser-arg-leu tripeptide 9 amino acids from its C-terminus (Figure 4.4.5), conserved variants of this PTS do not exist at or near the extreme C-terminus of any of the other catalases depicted in Figure 4.4.5. [In fact, a comparison of the C-terminal 25 amino acids of several catalases, conducted by Ni et al. (1990), negated the presence of any common structural feature (eg. α-helix, β-sheet, hydropathy or amphiphilic structure) among the C-termini of these enzymes.] Nevertheless, conserved variants of the tripeptide PTS of Gould et al. (1989) are present at

internal locations within each of the mammalian, plant, yeast and Drosophila catalases (excluding rat liver catalase) [eg. the tripeptide ser-arg-leu of Cat from C. tropicalis (residues 331 to 333, Figure 4.4.5, boxed)]; however, any involvement of these tripeptides in peroxisomal targeting remains to be determined experimentally. With reference to Cat from C. tropicalis, any role of the ser-arg-leu tripeptide in translocation is unlikely since this sequence is only minimally conserved in other catalases (Figure 4.4.5) and the observed degree of conservation (ie. the arginine residue at position 332 of Cat, Figure 4.4.5) is probably related only to its role in heme binding (section 4.4.3).

As mentioned in section 3.3, studies concerned with biogenesis of glycosomes in Trypanosoma brucei have suggested that for certain glycosomal enzymes, clusters of positive charge, located at their surface, may be important for translocation (for reviews, see Borst, 1986, 1989; Opperdoes, 1987, 1988). Although clusters of positive charge are not universal to peroxisomal proteins (Sakajo et al., 1987), regions which are rich in positively charged residues do exist, most notably the C-terminal 25 amino acids of several different peroxisomal proteins. Accordingly, for each of the acyl-coenzyme A oxidases presented in Figure 3.2.5, between 4 and 8 positively

charged residues are located within their C-terminal 25 amino acids. Similarly, each of the peroxisomal catalases depicted in Figure 4.4.5 possesses between 3 and 6 positively charged residues within their C-terminal 25 amino acids. The same is true of other peroxisomal proteins, including C. tropicalis PXP-18 (Szabo et al., 1989), HDE (Nuttley et al., 1988) and isocitrate lyase (Atomi et al., 1990), which contain 7, 5 and 3 positively charged residues, respectively, within their C-terminal 25 amino acids. In definite contrast with this seemingly nonrandom occurrence, however, cytosolic S. cerevisiae catalase T possesses only a single positively charged amino acid within the 25 residues located at its C-terminus (Figure 4.4.5). Thus, "appropriately" charged, C-terminal regions of peroxisomal proteins (ie. those possessing significant positive charge) may play an important role in organellar import.

5. TARGETING OF C. TROPICALIS CATALASE TO YEAST PEROXISOMES

Having achieved the cloning, sequencing and characterization of the genes encoding C. tropicalis acyl-coenzyme A oxidase (AOx) and catalase (Cat), efforts were directed toward establishing in vitro and/or in vivo expression and translocation systems to identify the targeting signal(s) responsible for the peroxisomal location of these enzymes. Small et al. (1987) reported yeast growth and fractionation conditions for the isolation of stable, translocation competent, C. tropicalis peroxisomes and, using this system, generated preliminary data suggestive of a PTS located within the C-terminal 40% of C. tropicalis PXP-4 (Small and Lazarow, 1987). Consequently, the present work utilized this in vitro import system to elucidate the nature of the peroxisomal addressing signal of C. tropicalis Cat. To this end, a series of N-terminal, C-terminal, N-/C-terminal and internal deletion mutants of Cat were constructed, expressed in cell-free systems and used in the in vitro import assay of Small et al. (1987) to determine whether the truncated polypeptides retained or lost the translocation competence characteristic of full-length Cat. Unfortunately, the in vitro import system proved to be inadequate for Cat translocation. This necessitated characterization of the system and culminated in attempts to develop a modified procedure for assaying in vitro protein

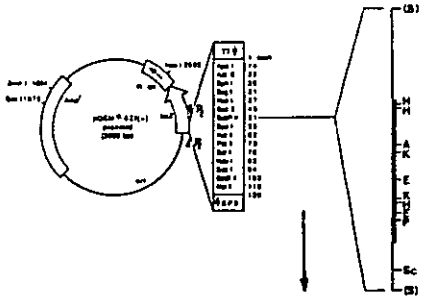
translocation into isolated C. tropicalis peroxisomes. Since the outcome of these in vitro import studies was uncertain, parallel studies concerned with the in vivo targeting of Cat to the peroxisomes of C. albicans and S. cerevisiae were undertaken. Upon achieving the in vivo expression of Cat and several deletion mutants, the subcellular location of each truncated Cat polypeptide was determined in an effort to delineate the region(s) of Cat which is essential for its targeting to yeast peroxisomes.

5.1 CONSTRUCTION OF THE Cat DELETION MUTANTS

5.1.1 N-terminal Deletions of the Cat ORF

As described in section 4.3.2, the BamH I, Sal I fragment of C. tropicalis genomic DNA, which encompassed the entire Cat gene, was subcloned into the EcoR V site of the in vitro expression vector pGEM-5Zf(+) for the purpose of DNA sequencing. Recombinants pGEMC1 and C2, which contained the genomic DNA fragment in opposite orientations, were used to sequence the Cat gene in both directions (ie. along both DNA strands); pGEMC2, which contained the BamH I, Sal I fragment in that orientation in which transcription of Cat mRNA was directed from the T7 RNA polymerase promoter (Figure 5.1.1, upper portion), also served as the template for the construction of each of the N-terminal deletions of the Cat gene. To this end, exonuclease III-directed,

Figure 5.1.1. Construction of the Cat deletion mutants used for the in vitro transcription of full-length and truncated Cat mRNAs. The parental recombinant of these constructs, pGEMC2, is illustrated; the partial restriction endonuclease cleavage map of the Cat gene is included, with the ORF indicated by the thick line. Unidirectional digestion of Apa I, Nco I-digested pGEMC2 with exonuclease III (⟼) generated the full length and N-terminal deletion constructs of the Cat gene. The full-length recombinant, pGEMC2-11b, subsequently served as the template for all of the other constructions (excluding pGEMC2-16bAH). Restriction endonuclease sites included in the Cat ORF are those used to digest pGEMC2-11b and yield the C-terminal, internal (drop-out) and N-/C-terminal deletions. They are: A, Aha II; E, EcoR I; H, Hinc II; K, Kpn I; P, Pvu II. The numbers adjacent to each construct indicate the nucleotide positions at which exonuclease III digestions were terminated (for the N-terminal deletions) or the nucleotide positions of restriction endonuclease cleavage sites used during construction (for all other deletions). Arrows represent the positions of internal, in-frame ATG codons to be used as "initiator" methionines upon in vitro translation of the N- and N-/C-terminal deleted mRNAs. Asterisks denote stop codons contributed by vector sequence. Dashed lines signify 5'- or 3'-nontranslated Cat sequence retained in each construct. All constructions were confirmed by the double-stranded sequencing of plasmid DNA using, for the most part, M13 forward (P_f) and reverse (P_r) sequencing primers. The plasmid map of pGEM-5Zf(+) was reproduced from the Promega 1989/90 catalogue.

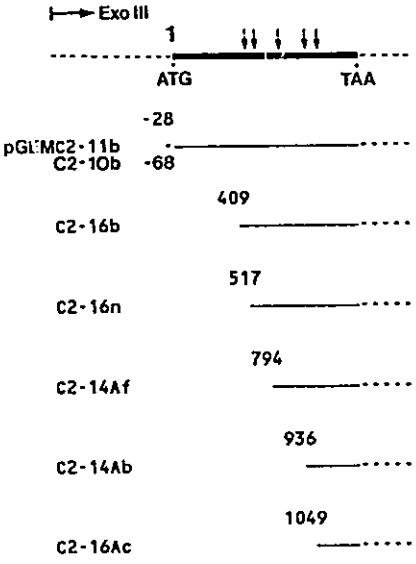


pGEMC2

DELETION CONSTRUCTS:

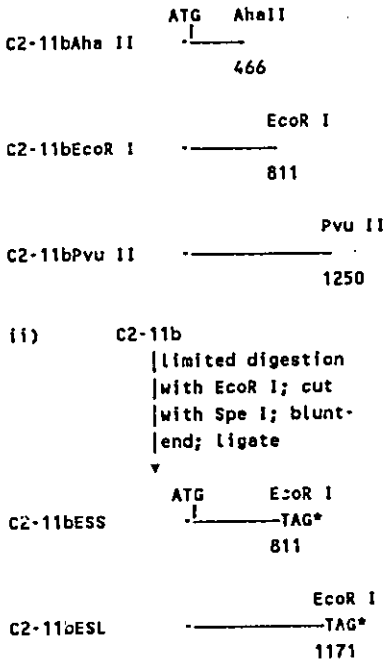
A) N-Terminal:

i) digest pGEMC2 with Apa I + Nco I; digest with exonuclease III to generate:



B) C-terminal:

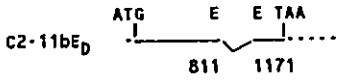
i) linearize C2-11b by digestion with:



ii) C2-11b
limited digestion with EcoR I; cut with Spe I; blunt-end; ligate

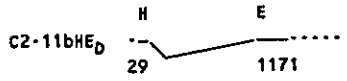
C) Internal (drop-out):

i) digest C2-11b with EcoR I (E); re-circulize:



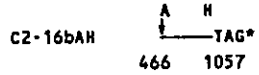
ii) digest C2-11b with Kpn I (K); re-circulize

iii) digest C2-11b with Hinc II (H) + EcoR I; blunt-end; re-circulize:



D) N-/C-terminal:

i) digest C2-16b with Aha II (A) + Hinc II; isolate Cat fragment; blunt-end; ligate with blunt-ended, Aat II, Nsi I-digested pGEM-7Zf(+):



unidirectional deletion of Apa I, Nco I-digested pGEMC2 was used (as described in section 2.5.6.2) to create sequential N-terminal deletions of the Cat ORF. These would, upon expression in vitro, dictate the synthesis of N-terminal truncated Cat polypeptides which could be used in import assays. Ninety-two potential, N-terminal deletion clones were prescreened by Sca I, Pst I-digestion of isolated plasmid DNA, followed by agarose gel electrophoresis, to determine those which comprised a subset of recombinants of sequentially decreasing size. Double-stranded plasmid sequencing of these recombinants was then used to further screen for those Cat constructs which had been deleted to a position 3' of the out of frame ATG codons present in the Cat ORF (of which there are ten in total, Figure 4.4.2), but 5' of the in-frame ATG codons present in the Cat ORF (of which there are nine in total including the start codon, Figure 4.4.2). It was anticipated that this would allow the in-frame ATG codons located at downstream positions to function as "initiation" codons upon in vitro transcription and translation of the N-terminal deletion constructs (see section 5.2). Fifteen pGEMC2 "daughter" recombinants were found to satisfy these criteria; collectively, they accounted for two full-length and five of eight different, N-terminal truncated Cat constructs that were possible (ie. of the eight downstream, in-frame ATG codons in the Cat ORF

that could potentially act as "start" codons, 5 were accounted for).

The two full-length constructs, pGEMC2-10b and -11b, possessed 68 and 28 bp of 5'-noncoding sequence, respectively (Figure 5.1.1, N-terminal deletion constructs) and were expected to dictate the synthesis of the 485 amino acid (ie. full-length) Cat polypeptide (Figure 5.1.2) of M_r 54,944 (Table 5.1.1). By comparison, the five N-terminal deletion constructs, pGEMC2-16b, -16n, -14Af, -14Ab and -16Ac, which exhibited sequentially larger deletions from the N-terminus, respectively (Figure 5.1.1, N-terminal deletion constructs), were characterized as follows:

pGEMC2-16b - Sequencing revealed that nucleotide 409 (a 'T', Figure 4.4.2) was the first residue of the Cat ORF retained in this construct (Figure 5.1.1). This dictated that, upon in vitro transcription and translation, met-159 (nucleotides 475 to 477, Figure 4.4.2) would function as the "initiator" methionine (Figure 5.1.2) for the synthesis of a 327 amino acid Cat polypeptide of M_r 37291 (Table 5.1.1).

pGEMC2-16n - Sequencing revealed that nucleotide 517 (a 'C', Figure 4.4.2) was the first residue of the Cat ORF retained in this construct (Figure 5.1.1). This dictated that, upon in vitro transcription and translation, met-175 (nucleotides 523 to 525, Figure 4.4.2) would function as the "initiator" methionine (Figure 5.1.2) for the synthesis of a 311 amino

Figure 5.1.2. Schematic representation of the truncated Cat polypeptides synthesized upon *in vitro* transcription and translation of the Cat deletion mutants. The full-length Cat polypeptide, synthesized from either pGEMC2-10b or -11b transcripts, is illustrated at the top with the positions of the internal, in-frame methionines of Cat, one of which will function as the "initiator" methionine upon *in vitro* translation of the different N- and N-/C-terminal deleted mRNAs, indicated by the line graph above. The numbers adjacent to the polypeptides indicate their respective N- and C-terminal amino acids; for the internal (drop-out) deletions, those amino acids of Cat which were retained N- and C-terminal to each deleted segment are also indicated. For C-terminal deletions C2-11bAha II, C2-11bEcoR I and C2-11bPvu II, the C-terminal residue of each polypeptide was not precisely determined (see text, section 5.1.2). The partial restriction endonuclease cleavage map of the C2-11b ORF is included to allow a comparison of the nucleotide positions of the restriction endonuclease cleavage sites used to generate the deletions, with the number of residues retained in the resultant polypeptides [this indicates that no "extra" (i.e. non-*C. tropicalis* catalase) amino acids resulted from the DNA manipulations]. Polypeptides C2-11bEcoR I and C2-11bESS, shown to be of identical size by the line graph, differ in size by approximately one amino acid (see text, section 5.1.2).

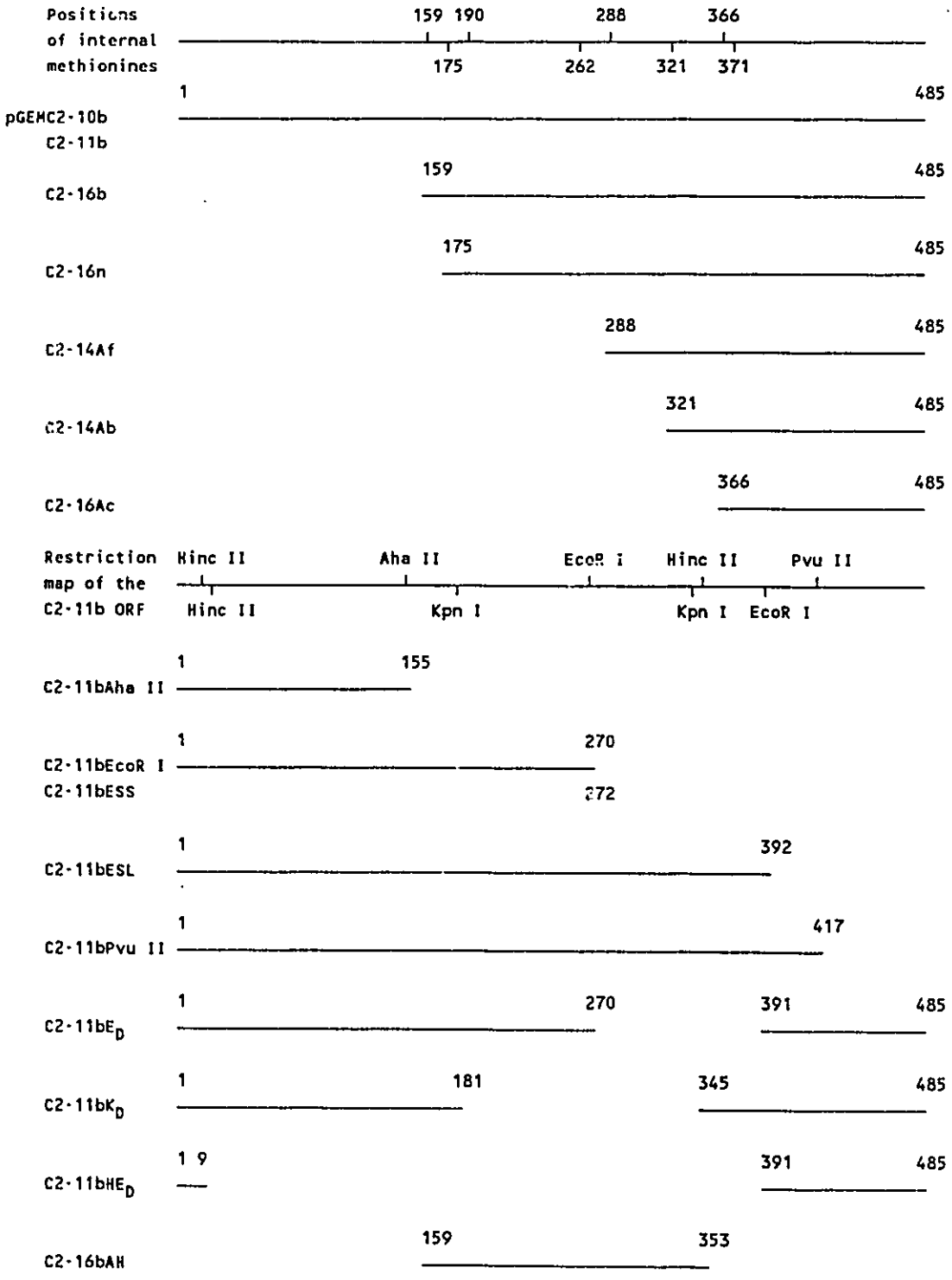

Positions of internal methionines
159 190
288
366
175
262
321
371
pGEMC2-10b
1
485
C2-11b
C2-16b
159
485
C2-16n
175
485
C2-14Af
288
485
C2-14Ab
321
485
C2-16Ac
366
485
Restriction map of the C2-11b ORF
Hinc II
Aha II
EcoR I
Hinc II
Pvu II
Hinc II
Kpn I
Kpn I
EcoR I
C2-11bAha II
1
155
C2-11bEcoR I
1
270
C2-11bESS
272
C2-11bESL
1
392
C2-11bPvu II
1
417
C2-11bE$_D$
1
270
391
485
C2-11bK$_D$
1
181
345
485
C2-11bHE$_D$
1 9
391
485
C2-16bAH
159
353

Table 5.1.1. Characteristics of the various truncated Cat polypeptides.

Construct	Characteristics Relative To Full Length Cat						M_r
	N-terminal amino acid	C-terminal amino acid	Amino acids deleted (D)/retained (R)			Percentage of full length Cat	
			N-terminal	C-terminal	Internal		
N-terminal deletions:							
C2-10b/C2-11b	met-1	lys-485	-	-	-	100.0	54944
C2-16b	met-159	lys-485	158(D)	327(R)	-	67.4	37291
C2-16n	met-175	lys-485	174(D)	311(R)	-	64.1	35310
C2-14Af	met-288	lys-485	287(D)	198(R)	-	40.8	22465
C2-14Ab	met-321	lys-485	320(D)	165(R)	-	34.0	18647
C2-16Ac	met-366	lys-485	365(D)	120(R)	-	24.7	13626
C-terminal deletions:							
C2-11bAha II[a]	met-1	asp-156	156(R)	329(D)	-	32.2	17504
C2-11bEcoR I[a]	met-1	glu-271	271(R)	214(D)	-	55.9	30613
C2-11bESS	met-1	phe-272	272(R)	213(D)	-	56.1	30778
C2-11bESL	met-1	phe-392	392(R)	93(D)	-	80.8	44294
C2-11bPvu II[a]	met-1	pro-417	417(R)	68(D)	-	86.0	47215
Internal (drop-out) deletions:							
C2-11bE_D	met-1	lys-485	270(R)	95(R)	120(D)	75.3	41409
C2-11bK_D	met-1	lys-485	181(R)	141(R)	163(D)	66.4	36338
C2-11bHE_D	met-1	lys-485	9(R)	95(R)	381(D)	21.4	11926
N-/C-terminal deletions:							
C2-16bAH	met-159	val-353	158(D)	132(D)	195(R)	40.2	22446

[a] For C-terminal deletions C2-11bAha II, C2-11bEcoR I and C2-11bPvu II, the C-terminal residue of each polypeptide was not precisely determined (see text, section 5.1.2). Consequently, other characteristics of these truncated polypeptides (including amino acids deleted/retained, percentage of full length Cat and M_r) have been approximated.

acid Cat polypeptide of M_r 35310 (Table 5.1.1).

pGEMC2-14Af - Sequencing revealed that nucleotide 794 (a 'C', Figure 4.4.2) was the first residue of the Cat ORF retained in this construct (Figure 5.1.1). This dictated that, upon <u>in vitro</u> transcription and translation, met-288 (nucleotides 862 to 864, Figure 4.4.2) would function as the "initiator" methionine (Figure 5.1.2) for the synthesis of a 198 amino acid Cat polypeptide of M_r 22465 (Table 5.1.1).

pGEMC2-14Ab - Sequencing revealed that nucleotide 936 (a 'C', Figure 4.4.2) was the first residue of the Cat ORF retained in this construct (Figure 5.1.1). This dictated that, upon <u>in vitro</u> transcription and translation, met-321 (nucleotides 961 to 963, Figure 4.4.2) would function as the "initiator" methionine (Figure 5.1.2) for the synthesis of a 165 amino acid Cat polypeptide of M_r 18647 (Table 5.1.1).

pGEMC2-16Ac - Sequencing revealed that nucleotide 1049 (an 'A', Figure 4.4.2) was the first residue of the Cat ORF retained in this construct (Figure 5.1.1). This dictated that, upon <u>in vitro</u> transcription and translation, met-366 (nucleotides 1096 to 1098, Figure 4.4.2) would function as the "initiator" methionine (Figure 5.1.2) for the synthesis of a 120 amino acid Cat polypeptide of M_r 13626 (Table 5.1.1).

5.1.2 <u>C-terminal Deletions of the Cat ORF</u>

Since exonuclease III-directed deletion of the Cat

gene from its 3'-end would ablate the TAA termination codon, construction of the C-terminal deletions of the Cat ORF involved two alternate approaches to that used for construction of the N-terminal deletions (section 5.1.1). First, consistent with that reported by Hoffman and Gilmore (1988), it was postulated that linearization of recombinant pGEMC2-11b within the Cat ORF, using different restriction endonucleases (eg. Aha II, EcoR I or Pvu II), would dictate T7 RNA polymerase-directed, in vitro transcription of "fall-off" Cat transcripts of different lengths. These transcripts, when translated in vitro, would yield discrete-sized, C-terminal truncated Cat polypeptides due to initiation from the natural AUG start codon of the Cat ORF and termination by transcript length alone (ie. "fall-off" translation). The three constructs generated using this approach, pGEMC2-11bAha II, EcoR I and Pvu II, which exhibited sequentially smaller deletions from the C-terminus, respectively [Figure 5.1.1, C-terminal deletion constructs (i)], were characterized as follows:

pGEMC2-11bAha II - Linearization of pGEMC2-11b at the Aha II site (GACGCC) present at nucleotide 466 of the Cat ORF (Figures 4.4.2 and 5.1.1) would dictate the in vitro transcription of Cat mRNA which was 469 nucleotides in length (maximum*). Upon in vitro translation, this message would encode a 156 amino acid Cat polypeptide (having asp-

156 as the C-terminal residue, Figure 5.1.2) of M_r 17504 (Table 5.1.1).

pGEMC2-11bEcoR I - Linearization of pGEMC2-11b at the most 5' EcoR I site (GAATTC) present at nucleotide 811 of the Cat ORF (Figures 4.4.2 and 5.1.1) would dictate the in vitro transcription of Cat mRNA which was 815 nucleotides in length (maximum*). Upon in vitro translation, this message would encode a 271 amino acid Cat polypeptide (having glu-271 as the C-terminal residue, Figure 5.1.2) of M_r 30613 (Table 5.1.1).

pGEMC2-11bPvu II - Linearization of pGEMC2-11b at the Pvu II site (CAGCTG) present at nucleotide 1250 of the Cat ORF (Figures 4.4.2 and 5.1.1) would dictate the in vitro transcription of Cat mRNA which was 1252 nucleotides in length (maximum*). Upon in vitro translation, this message would encode a 417 amino acid Cat polypeptide (having pro-417 as the C-terminal residue, Figure 5.1.2) of M_r 47215 (Table 5.1.1).

[* These estimates are of the maximum number of nucleotides that could be incorporated into each deleted Cat transcript. The synthesis of shorter transcripts/polypeptides, due to premature displacement of the T7 RNA polymerase or ribosomal complexes during in vitro transcription or translation, respectively, may have occurred, but was not determined.]

Realizing that "fall-off" transcription would not be

applicable to the in vivo expression of Cat sequences, a second approach was undertaken to create additional C-terminal deletion constructs that could be expressed in vitro and in vivo. This involved double digestion of pGEMC2-11b with EcoR I, which, at limiting concentrations of enzyme, linearized pGEMC2-11b within the Cat ORF at only one of two possible GAATTC sites (located at nucleotides 811 and 1171, Figure 4.4.2), and SpeI, which linearized pGEMC2-11b at the ACTAGT site located at position 55 of the pGEM-5Zf(+) MCS (Figure 5.1.1, upper portion), followed by blunt-ending with Klenow fragment and re-ligation. Since a blunt-ended (using Klenow fragment) SpeI site contains an inherent TAG stop codon when in the proper context [eg. when ligated to the blunt-ended (using Klenow fragment) EcoR I site of the Cat ORF], two different C-terminal deletion constructions, complete with termination codons but without extra, non-Cat-encoding sequences, were created. EcoR I digestion of isolated plasmid DNA was subsequently used to verify and distinguish between constructs since, depending on the position of the initial EcoR I cleavage site, re-ligation of the blunt-ended EcoR I and SpeI sites would generate recombinants possessing either one or two EcoR I sites. The resulting constructs, pGEMC2-11bESS and ESL, which exhibited sequentially smaller deletions from the C-terminus, respectively [Figure 5.1.1, C-terminal deletion constructs

(ii)], were characterized as follows:

pGEMC2-11bESS - Restriction endonuclease mapping revealed that the aforementioned DNA manipulations resulted in the ligation of the blunt-ended EcoR I site at nucleotide 811 of the Cat ORF to the blunt-ended Spe I site of the pGEM-5Zf(+) MCS (Figure 5.1.1). This dictated that a 272 amino acid Cat polypeptide (having phe-272 as the C-terminal residue, Figure 5.1.2) of M_r 30778 (Table 5.1.1) would be synthesized upon in vitro transcription and translation.

pGEMC2-11bESL - Restriction endonuclease mapping revealed that the aforementioned DNA manipulations resulted in the ligation of the blunt-ended EcoR I site at nucleotide 1171 of the Cat ORF to the blunt-ended Spe I site of the pGEM-5Zf(+) MCS (Figure 5.1.1). This dictated that a 392 amino acid Cat polypeptide (having phe-392 as the C-terminal residue, Figure 5.1.2) of M_r 44294 (Table 5.1.1) would be synthesized upon in vitro transcription and translation.

5.1.3 Internal (Drop-out) Deletions of the Cat ORF

Having generated several constructs which would yield different N- and C-terminal truncated Cat polypeptides, efforts were taken to complement these mutants with internal deletions of the Cat ORF which would, upon in vitro transcription and translation, maintain the integrity of different portions of both the N- and C-termini of Cat, but delete different portions of those amino acids internal

to Cat. This entailed linearization of recombinant pGEMC2-11b with different restriction endonucleases (eg. EcoR I, Kpn I or Hinc II) that recognized a minimum of two cleavage sites within the Cat ORF (thereby allowing the "drop-out" of a DNA fragment), followed by conversion to blunts ends (if required) and re-ligation of the deleted construct to maintain the original reading frame without introducing extra, non-Cat-encoding sequence. In this manner, three internal (drop-out) deletion constructs, pGEMC2-11bE_D, K_D and HE_D, which exhibited sequentially larger internal deletions, respectively [Figure 5.1.1, internal (drop-out) deletion constructs], were constructed and characterized as follows:

pGEMC2-11bE_D - Digestion of pGEMC2-11b at the two EcoR I sites (GAATTC) present at nucleotides 811 and 1171 of the Cat ORF (Figure 4.4.2 and 5.1.1), followed by re-ligation, created a 360 bp deletion, as confirmed by restriction endonuclease mapping (ie. a single EcoR I site should have been, and was, regenerated) and sequencing. This dictated that a 365 amino acid Cat polypeptide [possessing 270 and 95 residues of the N- and C-terminus of Cat, respectively (Figure 5.1.2)] of M_r 41409 (Table 5.1.1) would be synthesized upon <u>in vitro</u> transcription and translation.

pGEMC2-11bK_D - Digestion of pGEMC2-11b at the two Kpn I sites (GGTACC) present at nucleotides 544 and 1033 of the

Cat ORF (Figure 4.4.2 and 5.1.1), followed by re-ligation, created a 489 bp deletion, as confirmed by restriction endonuclease mapping (ie. a single Kpn I site should have been, and was, regenerated) and sequencing. This dictated that a 322 amino acid Cat polypeptide [possessing 181 and 141 residues of the N- and C-terminus of Cat, respectively (Figure 5.1.2)] of M_r 36338 (Table 5.1.1) would be synthesized upon <u>in vitro</u> transcription and translation.

pGEMC2-11bHE$_D$ - Digestion of pGEMC2-11b at the three Hinc II sites (GTCAAC, located at nucleotides 29, 68 and 1057, Figure 4.4.2) and the two EcoR I sites (GAATTC, located at nucleotides 811 and 1171, Figure 4.4.2) of the Cat ORF, followed by blunt-ending (using Klenow fragment) and re-ligation, created a 1143 bp deletion, as confirmed by sequencing. [By design, this deletion should have been 1140 bp in size: however, 3 additional bp were deleted from the most 5' Hinc II site (possibly by exonuclease activity) to maintain the integrity of the Cat ORF upon re-ligation.] This dictated that a 104 amino acid Cat polypeptide [possessing 9 and 95 residues of the N- and C-terminus of Cat, respectively (Figure 5.1.2)] of M_r 11926 (Table 5.1.1) would be synthesized upon <u>in vitro</u> transcription and translation.

5.1.4 <u>N-/C-terminal Deletions of the Cat ORF</u>

To complete this "family" of Cat deletion mutants,

one construct, pGEMC2-16bAH, was generated in which the 5'- and 3'-sequences of the Cat ORF were simultaneously deleted so that upon *in vitro* transcription and translation, a polypeptide comprised only of the "internal" amino acids of Cat would be synthesized. To this end, an Aha II, Hinc II Cat fragment, isolated from recombinant pGEMC2-16b, was blunt-ended (using Klenow fragment) and ligated with Aat II, Nsi I-digested, blunt-ended (using mung bean nuclease) pGEM-7Zf(+) (Figure 5.1.1). Since mung bean nuclease treatment of the Nsi I site (**ATGCA**TAG) of pGEM-7Zf(+) should expose a TAG codon (underlined), which, in the proper context [eg. when ligated to the 3' (Hinc II-cut)-end of the Aha II, Hinc II Cat fragment of pGEMC2-16b], would function as a termination codon, the pGEMC2-16bAH construct was generated with an in-frame TAG termination codon, but without any extra, non-Cat-encoding sequence (as confirmed by sequencing). Thus, upon *in vitro* transcription and translation, a 195 amino acid Cat polypeptide of M_r of 22446 (Table 5.1.1) would be synthesized since met-159 (nucleotides 475 to 477, Figure 4.4.2) would function as the "initiator" methionine and val-353 (nucleotides 1057 to 1059, Figure 4.4.2) would be the C-terminal residue preceding the TAG stop codon provided by the vector sequence (Figure 5.1.2).

5.2 IN VITRO EXPRESSION OF THE Cat DELETION MUTANTS

5.2.1 In Vitro Transcription of Cat mRNA

Having achieved the construction of several Cat deletion mutants, efforts were directed toward establishing *in vitro* transcription and translation systems for the expression of truncated Cat mRNAs and polypeptides, respectively, to be used in assaying *in vitro* import into isolated *C. tropicalis* peroxisomes. As described in section 5.1 (and illustrated in Figure 5.1.1), each of the Cat deletion mutants was generated such that the *in vitro* transcription of Cat mRNA was directed from the T7 RNA polymerase promoter. The bacteriophage T7 system (Chamberlain and Ring, 1973) was chosen as it is reportedly superior, both quantitatively and qualitatively, to the bacteriophage SP6 (Krieg and Melton, 1987) and T5 (Bujard et al., 1987) systems; quantitatively, T7 RNA polymerase is more active than the other polymerases and qualitatively, it appears to be less sensitive to premature transcription termination (Roitsch and Lehle, 1989). To maximize the efficiency of the T7 transcription system, and the subsequent *in vitro* translation of Cat transcripts, several relevant observations, which are mentioned below, were incorporated into the *in vitro* transcription protocol. Since transcripts possessing "extra" 5'-noncoding sequence upstream of the natural translation start codon exhibit

diminished *in vitro* protein synthesis (Roitsch and Lehle, 1989), due either to the presence of upstream, out of frame AUG codons or the "extra" upstream sequence itself, each N-terminal Cat deletion mutant was purposely selected to possess a minimum of noncoding sequence (ie. sequence located 5' to the putative start codon) which was deficient in out of frame ATG codons. For a similar reason, the full-length Cat construct, pGEMC2-11b (which possesses 28 bp of upstream Cat sequence, section 5.1.1), not pGEMC2-10b (which possesses 68 bp of upstream Cat sequence, section 5.1.1), was used as the template for the subsequent construction of most of the other Cat deletion mutants. Further, because the 5'-capping of *in vitro* synthesized transcripts is essential to translation in wheat germ extracts and increases, by 3 to 4-fold, the translation efficiency of rabbit reticulocyte lysates (Roitsch and Lehle, 1989), *in vitro* 5'-capping of Cat mRNA was achieved cotranscriptionally in the presence of G(5')ppp(5')G (for subsequent wheat germ extract translation) or m^7G(5')ppp(5')G (for subsequent rabbit reticulocyte lysate translation), added in an initial excess (≈10-fold) over GTP (section 2.6.1.2). Additional sequences present at the 3'-end of transcripts, or conversely, the absence of a stop codon from an ORF, are known not to adversely affect *in vitro* translation (Haeuptle et al., 1986; Roitsch and Lehle,

1989). Consequently, "fall-off" transcription was employed to generate certain C-terminal truncated Cat transcripts; however, only those restriction endonucleases creating 5'-protruding termini were used to linearize the template DNA since "run-off" transcription templates possessing 3'-protruding termini are prone to extraneous transcription to as much as 90% of the total transcripts synthesized (Mierendorf and Jendrisak, 1985). Based upon these premises, the procedure outlined in section 2.6.1.2 resulted in the T7 RNA polymerase-directed, in vitro transcription of 10 to 20 µg of full-length or truncated Cat mRNA (per 50 µl reaction; quantitated by fluorometric analysis), independent of the construct used; however, on average, supercoiled template yielded up to twice as much transcript as linearized template. Similarly, SP6 RNA polymerase-directed, in vitro transcription, which was used (as described in section 2.6.1.2) for the synthesis of CAT and DHFR transcripts only, generated approximately equivalent amounts (0.5 to 1.0 µg) of mRNA per unit enzyme.

5.2.2 In Vitro Rabbit Reticulocyte Lysate/Wheat Germ Extract Translation of Cat mRNA

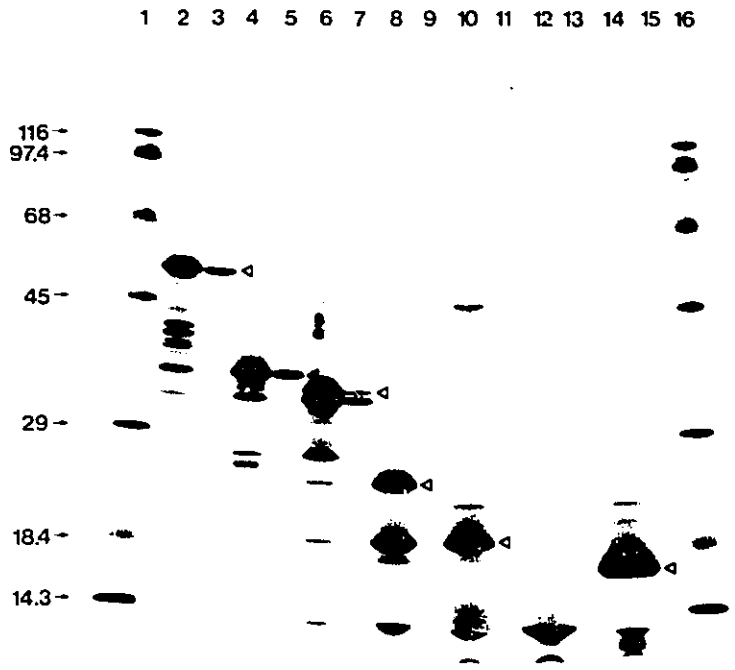
For reasons explained in section 5.3, in vitro transcribed Cat mRNA was translated in two different cell-free systems: either a micrococcal nuclease-treated, rabbit reticulocyte lysate or wheat germ extract translation

system. The reticulocyte lysate system was purchased from Promega Corporation and was, therefore, pre-optimized for in vitro translation. Typically, translation assays were conducted as described in section 2.6.1.3 using 70% (v/v) nuclease-treated reticulocyte lysate programmed with 0.5 to 1.0 µg of full-length, or truncated, in vitro transcribed Cat mRNA. Since the sequences adjacent to the AUG start codon of the Cat ORF closely reflected the higher eukaryotic consensus sequence for translation initiation (see section 4.2.2; Kozak, 1987), the major reticulocyte lysate (or wheat germ extract) translation products, synthesized from Cat transcripts derived from the full-length deletion mutant, pGEMC2-11b, or the C-terminal and internal (drop-out) deletions therefrom, were expected to initiate from the natural start codon and thus, closely approximate the M_r predicted for each "full-length" polypeptide (see Table 5.1.1). With respect to the N-terminal and N-/C-terminal deletion mutants, however, the natural start codon was deleted, leaving downstream in-frame AUG triplets to function as "initiation" codons. Two of these AUG codons, which would encode "initiator" met-159 and -175 of polypeptides C2-16b and C2-16n, respectively [nucleotides 475 to 477 and 523 to 525, respectively (Figure 4.4.2)], were expected to function well in this regard due to the conservation of the critical purine nucleotide at the -3

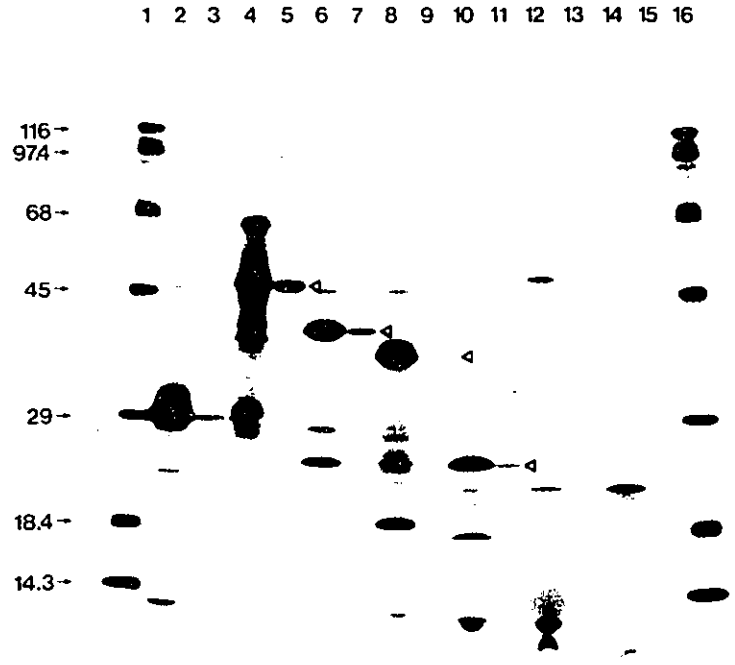
position relative to each AUG codon (Figure 4.4.2; Kozak, 1987, 1989a). By contrast, the other three AUG codons, which would encode "initiator" met-288, -321 and -366 of polypeptides C2-14Af, -14Ab and -16Ac, respectively [nucleotides 862 to 864, 961 to 963 and 1096 to 1098, respectively (Figure 4.4.2)], were less likely to function as "initiation" codons since, in each case, the critical -3 position relative to the AUG codon is a pyrimidine nucleotide (ie. a "C"; Figure 4.4.2, Kozak, 1989a), which occurs naturally at a frequency of only 2.1% (15 occurrences in a sample of 699 vertebrate mRNAs examined; Kozak, 1987). Nevertheless, based on the M_r estimated for the major rabbit reticulocyte lysate translation products synthesized from each full-length or truncated Cat transcript (Figure 5.2.1, Panels A and B, open triangles), it appeared that each Cat polypeptide was initiated from the most 5' AUG codon of the corresponding mRNA, including those deleted from the N-terminus. Transcripts derived from pGEMC2-11b dictated the synthesis of a major polypeptide of ≈55 kDa (Figure 5.2.1, Panel A, lane 2, open triangle), presumed to be full-length Cat, which was immunoprecipitated with rabbit anti-C. tropicalis Cat serum (Figure 5.2.1, Panel A, lane 3, open triangle). As hoped for, the N-terminal truncated transcripts directed the synthesis of the following, progressively smaller, major polypeptides: pGEMC2-16b

Figure 5.2.1. In vitro expression and immunoprecipitation of translation products derived from the Cat deletion mutants. Each deletion construct was in vitro transcribed and translated (using a nuclease-treated rabbit reticulocyte lysate system) as described in sections 2.6.1.2 and 2.6.1.3, respectively. Where possible, approximately equivalent cpm from each translation (to a maximum of 2 µl of lysate) were used for SDS-PAGE [10 to 20% (w/v) acrylamide, linear gradient gels], followed by fluorography. Approximately 50,000 cpm of each translation (where possible), immunoprecipitated with rabbit anti-C. tropicalis catalase serum, were similarly analyzed. The "full-length" polypeptide synthesized upon translation (and immunoprecipitation of translation products) of each deleted Cat mRNA is indicated by an open triangle. The 13.6 and 11.9 kDa translation products of C2-16Ac and C2-11bHE$_D$, respectively (if synthesized) were not resolvable in this gel system. The molecular mass protein standards (in kDa) shown at the left include: E. coli β-galactosidase (116), phosphorylase b (97.4), bovine serum albumin (68), chicken egg albumin (45), carbonic anhydrase (29), β-lactoglobulin (18.4) and egg white lysozyme (14.3). Exposures were for 40 h at -70°C with Kodak XAR-5 film. Panel A: Lanes: 1, ^{14}C-molecular mass protein standards; 2, in vitro translation products (IVTP) from pGEMC2-11b transcripts; 3, immunoprecipitated translation products (ITP) from 2; 4, IVTP from pGEMC2-16b transcripts; 5, ITP from 4; 6, IVTP from pGEMC2-16n transcripts; 7, ITP from 6; 8, IVTP from pGEMC2-14Af transcripts; 9, ITP from 8; 10, IVTP from pGEMC2-14Ab transcripts; 11, ITP from 10; 12, IVTP from pGEMC2-16Ac transcripts; 13, ITP from 12; 14, IVTP from pGEMC2-11bAha II transcripts; 15, ITP from 14; 16, ^{14}C-molecular mass protein standards. Panel B: Lanes: 1, ^{14}C-molecular mass protein standards; 2, in vitro translation products (IVTP) from pGEMC2-11bEcoR I transcripts; 3, immunoprecipitated translation products (ITP) from 2; 4, IVTP from pGEMC2-11bPvu II transcripts; 5, ITP from 4; 6, IVTP from pGEMC2-11bE$_D$ transcripts; 7, ITP from 6; 8, IVTP from pGEMC2-11bK$_D$ transcripts; 9, ITP from 8; 10, IVTP from pGEMC2-16bAH transcripts; 11, ITP from 10; 12, IVTP from pGEMC2-11bHE$_D$ transcripts; 13, ITP from 12; 14, control (minus message) translation products; 15, ITP from 14; 16, ^{14}C-molecular mass protein standards.

A


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
116
97.4
68
45
29
18.4
14.3

B


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
116
97.4
68
45
29
18.4
14.3

transcripts - an ≈36 kDa polypeptide (Figure 5.2.1, Panel A, lane 4, open triangle); pGEMC2-16n transcripts - an ≈34 kDa polypeptide (Figure 5.2.1, Panel A, lane 6, open triangle); pGEMC2-14Af transcripts - an ≈24 kDa polypeptide (Figure 5.2.1, Panel A, lane 8, open triangle) and pGEMC2-14Ab transcripts - an ≈18 kDa polypeptide (Figure 5.2.1, Panel A, lane 10, open triangle). Unfortunately, the C2-16Ac translation product (predicted M_r = 13626; Table 5.1.1), if synthesized, was not resolved in the 10 to 20% (w/v) acrylamide, linear gradient gel system of Figure 5.2.1 (Panel A, lane 12). Immunoprecipitation of the N-terminal deleted Cat polypeptides (Figure 5.2.1, Panel A, lanes 5, 7, 9 and 11, open triangles) suggested that some of the antigenic determinants of Cat were located either toward the N-terminus or internally since progressively larger deletions from the N-terminus significantly reduced the immunoprecipitable protein [to the extent that immunoprecipitated C2-14Af and C2-14Ab polypeptides were detected only after an 18 d exposure (not shown)]. As expected, rabbit reticulocyte lysate translation of pGEMC2-11b Aha II, EcoR I and Pvu II transcripts yielded progressively larger Cat polypeptides of ≈16 kDa (Figure 5.2.1, Panel A, lane 14, open triangle), ≈29 kDa (Figure 5.2.1, Panel B, lane 2, open triangle) and ≈46 kDa (Figure 5.2.1, Panel B, lane 4, open triangle), respectively. This

was confirmed by immunoprecipitation (Figure 5.2.1, Panel A, lane 15 and Panel B, lanes 3 and 5, respectively, open triangles), which further suggested an internal location for the major antigenic determinants of Cat since C2-11bAha II was the least immunoreactive polypeptide (this immunoprecipitate was detected only after a 7 d exposure; not shown). Internal deletion constructs pGEMC2-11bE_D and K_D yielded appropriately-sized, major translation products of ≈40 kDa and ≈35 kDa, respectively (Figure 5.2.1, Panel B, lanes 6 and 8, respectively, open triangles), which were immunoprecipitated with rabbit anti-C. tropicalis Cat serum (Figure 5.2.1, Panel B, lanes 7 and 9, respectively, open triangles); however, similar to the situation observed for the C2-16Ac polypeptide, the C2-11bHE_D translation product [which would possess the largest internal deletion of Cat (predicted M_r = 11926; Table 5.1.1)], if synthesized, was not resolved in this gradient gel system (Figure 5.2.1, Panel B, lane 12). Translation of pGEMC2-16bAH transcripts yielded an N-/C-terminal deleted polypeptide of ≈24 kDa (Figure 5.2.1, Panel B, lane 10, open triangle), which compared favourably with a predicted M_r of 22446 (Table 5.1.1). Consistent with the postulated internal location of the major antigenic determinants of Cat, the larger (more central), internal deletion of the Cat polypeptide (ie. that of C2-11bK_D) significantly reduced the amount of

immunoprecipitable protein (Figure 5.2.1, Panel B, lane 9, open triangle), while immunoprecipitation of the "central" 195 amino acids of Cat (ie. the C2-16bAH polypeptide) yielded a significant recovery (Figure 5.2.1, Panel B, lane 11, open triangle).

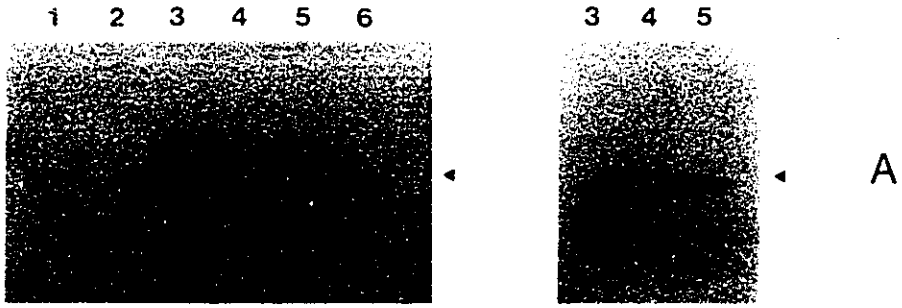
Thus, rabbit reticulocyte lysate translation of each of the pGEMC transcripts resulted in the synthesis of a major translation product, the size of which correlated with its predicted M_r (Table 5.1.1). For several of these transcipts, however, other minor translation products were also synthesized (Figure 5.2.1, Panel A, lanes 2, 4, 6 and 8; Panel B, lanes 4, 6 and 8). These polypeptides were postulated to be the result of translation initiation from both internal in- and out of frame AUG codons since some of these polypeptides were clearly immunoprecipitated with rabbit anti-C. tropicalis Cat serum, while others were not [eg. translation of pGEMC2-16n transcripts yielded two immunoprecipitable and several non-immunoprecipitable polypeptides (Figure 5.2.1, Panel A, compare lanes 6 and 7)]. Further, the migration of the major bands resulting from the translation of progressively smaller N-terminal deletion transcripts (ie. progressively larger N-terminal deletions of Cat) appeared to be identical to that of minor, immunoprecipitable polypeptides which resulted upon translation of transcripts possessing smaller N-terminal

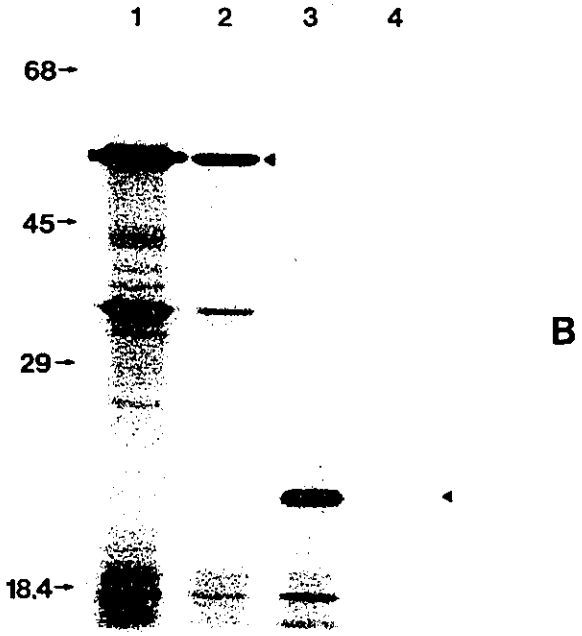
deletions or no deletion at all (ie. the more "full-length" Cat transcripts) (Figure 5.2.1, Panel A, lanes 2 through 10, compare polypeptides denoted by open triangles across preceding lanes). Altering the in vitro K^+ and/or Mg^{2+} concentrations (and possibly the spermidine concentration) for each translation may have increased the fidelity (and/or efficiency) of translation from the different "initiator" AUG codons (Mead et al., 1985; Lightfoot, 1988; Kozak, 1989b, 1989c), thereby dictating the synthesis of predominantly one major Cat polypeptide. This was not attempted, however, since lower fidelity translations presented the opportunity to simultaneously study the in vitro import of more than one, progressively smaller, Cat polypeptide into isolated C. tropicalis peroxisomes while using nonsense polypeptides initiated from out of frame AUG codons as controls indicative of the specificity of catalase import. Interestingly, certain translations yielded polypeptides of greater M_r than that predicted for the "full-length" Cat polypeptide, one of which was immunoprecipitated with rabbit anti-C. tropicalis Cat serum (Figure 5.2.1, Panel B, lanes 4 and 5). The identity of these translation products, which were not present in the control (minus message) translation and immunoprecipitation (Figure 5.2.1, Panel B, lanes 14 and 15), is not known.

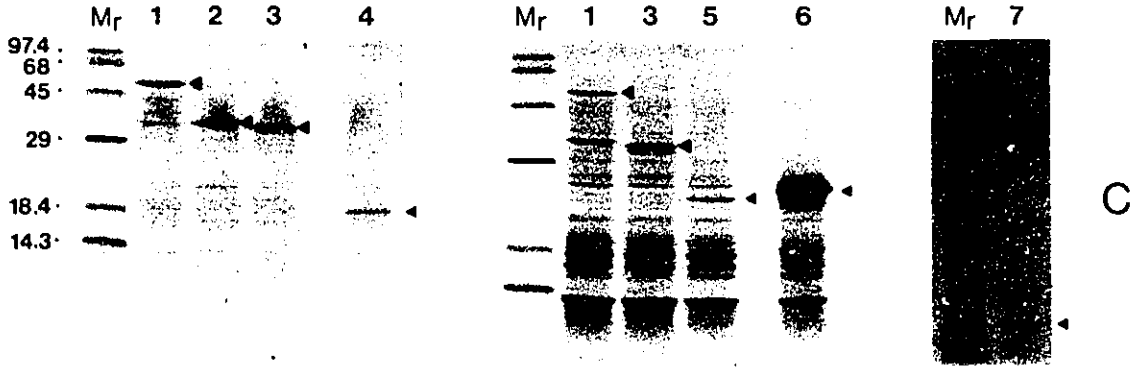
As an alternative to cell-free translation using a

rabbit reticulocyte lysate system, full-length and truncated Cat transcripts were also translated in a nuclease-treated wheat germ extract system. To this end, extracts of wheat germ embryos were prepared essentially as described by Erickson and Blobel (1983) (see section 2.6.1.4); consequently, they were pre-optimized for their K^+ and Mg^{2+} concentrations, both of which are known to have dramatic effects on the fidelity and efficiency of translation, by wheat germ ribosomes, of particular mRNAs (Erickson and Blobel, 1983; Kozak, 1989a, 1989c). Cell-free translations involving the nuclease-treated wheat germ extract were also optimized with respect to the percentage (v/v) of the extract necessary to achieve the maximum overall incorporation of amino acids (as judged by the incorporation of ^{35}S-methionine) upon translation of <u>in vitro</u> transcribed mRNA (eg. 0.5 µg of <u>in vitro</u> transcribed prolactin mRNA; Figure 5.2.2, Panels A). As seen in Figure 5.2.2 (left Panel A), the use of 10%, 15% or 35% (v/v) nuclease-treated wheat germ extract yielded only minimal translation of the prolactin mRNA (lanes 1, 2 and 6, respectively, arrowhead). Conversely, the use of either 20%, 25% or 30% (v/v) nuclease-treated wheat germ extract dictated the synthesis of a constant, significant amount of the prolactin polypeptide (lanes 3, 4 and 5, respectively, arrowhead) [note that : 1) the apparent decrease in translation

Figure 5.2.2. Optimization of the nuclease-treated, wheat germ extract translation system and subsequent <u>in vitro</u> expression of Cat deletion mutants. Panels A, ≈0.5 µg of <u>in vitro</u> transcribed prolactin mRNA were translated (as described in section 2.6.1.4) in the presence of different percentages (v/v) of nuclease-treated wheat germ extract and 1 µl of each translation analyzed by SDS-PAGE and autoradiography. Lanes (for both panels): 1, 10% wheat germ extract; 2, 15%; 3, 20%; 4, 25%; 5, 30%; 6, 35%. The apparent decrease in translation efficiency observed with 25% (v/v) wheat germ extract (left Panel A, lane 4) was artifactual since re-electrophoresis of 0.5 µl of samples 3, 4 and 5 (right Panel A) resolved this anomaly. Exposure was for 12 h (left Panel A) and 36 h (right Panel A) at -70°C with Kodak XAR-5 film. Panel B, ≈1 µg of pGEMC2-11b (lanes 1 and 2) or pGEMC2-14Af (lanes 3 and 4) mRNA, synthesized (as per section 2.6.1.2) in the presence of 0.016 A_{260} units/µl of 5'-cap [G(5')ppp(5')G (lanes 1 and 3) or m^7G(5')ppp(5')G (lanes 2 and 4)], was translated using the nuclease-treated wheat germ extract [20% (v/v)] and 1 µl of each translation analyzed by SDS-PAGE and fluorography. Exposure was for 20 h at -70°C with Kodak XAR-5 film. Panels C, 0.5 to 1.0 µg of several different <u>in vitro</u> transcribed mRNAs (encoding Cat, CAT or DHFR) was translated using the nuclease-treated wheat germ extract and 1 µl of each translation analyzed by SDS-PAGE and fluorography. Lanes: transcripts derived from: 1, pGEMC2-11b; 2, pGEMC2-16b; 3, pGEMC2-16n; 4, pGEMC2-14Ab; 5, pGEMC2-14Af; 6, TP408 (CAT); 7, pSP65DHFR-1 (DHFR). The molecular mass protein standards (in kDa) seen in lanes M_r include: phosphorylase b (97.4), bovine serum albumin (68), chicken egg albumin (45), carbonic anhydrase (29), β-lactoglobulin (18.4) and egg white lysozyme (14.3). Exposures were for 14 h at -70°C with Kodak XAR-5 film. Each respective "full-length" translation product in Panels A, B and C is denoted by a closed triangle.


1 2 3 4 5 6
3 4 5
A


1 2 3 4
68
45
29
18.4
B


Mr 1 2 3 4
97.4
68
45
29
18.4
14.3
Mr 1 3 5 6
Mr 7
C

efficiency observed with 25% (v/v) wheat germ extract (Figure 5.2.2, left Panel A, lane 4) was artifactual since re-electrophoresis of the same samples from lanes 3, 4 and 5 (Figure 5.2.2, right Panel A) resolved this anomaly, and 2) TCA-precipitable counts were not determined for wheat germ extract translations since examination, by SDS-PAGE, of the radiolabeled amino acid incorporated into a specific protein is preferred over establishing the amount of radiolabeled amino acid incorporated into total protein (Erickson and Blobel, 1983)]. Since upon visual inspection there was no appreciable difference between the amount of prolactin synthesized using either 20%, 25% or 30% (v/v) wheat germ extract (Figure 5.2.2, right Panel A, lanes 3, 4 and 5, respectively, arrowhead), all subsequent translations of Cat transcripts were conducted in the presence of only 20% (v/v) nuclease-treated wheat germ extract.

As previously mentioned (section 5.2.1), the 5'-capping of <u>in vitro</u> synthesized transcripts is essential to wheat germ extract translation (Roitsch and Lehle, 1989) and was, therefore, conducted. In an attempt to further optimize the wheat germ extract system for the translation of Cat-specific messages, the <u>in vitro</u> transcription of Cat mRNA was carried out in the presence of two different 5'-caps, either G(5')ppp(5')G or m^7G(5')ppp(5')G, to determine the effect, if any, of different 5'-caps (ie. differently

5'-capped Cat mRNAs) on the efficiency of subsequent wheat germ extract translations. As depicted in Figure 5.2.2 (Panel B), full-length Cat transcripts, 5'-capped with G(5')ppp(5')G (lane 1), were translated 3- to 5-fold more efficiently than the same transcripts which were m^7G(5')ppp(5')G-capped (lane 2). This difference in translation efficiency was further accentuated for pGEMC2-14Af transcripts in that G(5')ppp(5')G-capped messages were translated at least 10-fold more efficiently than m^7G(5')ppp(5')G-capped messages (Figure 5.2.2, Panel B, compare lanes 3 and 4, respectively, arrowhead). Thus, it appeared that the 5'-cap, G(5')ppp(5')G, was most compatible with the in vitro translation of Cat transcripts in the wheat germ extract system and accordingly, all in vitro transcriptions of Cat mRNAs destined for wheat germ extract translation were conducted in the presence of G(5')ppp(5')G.

Subsequent in vitro wheat germ extract [20% (v/v)] translation of full-length and truncated Cat transcripts, 5'-capped with G(5')ppp(5')G, yielded polypeptides of a M_r which paralleled that observed for similar translations using rabbit reticulocyte lysates. Transcripts derived from pGEMC2-11b dictated the synthesis of a major polypeptide of ≈55 kDa (Figure 5.2.2, Panels B and C, lanes 1, arrowheads), presumed to be full-length Cat. By comparison, N-terminal truncated Cat transcripts derived from pGEMC2-16b, -16n,

-14Af and -14Ab directed the synthesis of progressively smaller Cat polypeptides (Figure 5.2.2) of ≈36 kDa (Panels C, lane 2, arrowhead), ≈34 kDa (Panels C, lanes 3, arrowheads), ≈24 kDa (Panel B, lane 3 and Panels C, lane 5, arrowheads) and ≈18 kDa (Panels C, lane 4, arrowhead), respectively. Thus, it appeared that translation of each N-terminal truncated Cat polypeptide was initiated from the most 5' AUG codon of the corresponding mRNA in both the rabbit reticulocyte lysate and wheat germ extract systems, independent of a poor consensus sequence for translation initiation by eukaryotic ribosomes (eg. that possessed by pGEMC2-14Af and -14Ab transcripts, discussed above; Kozak, 1987, 1989a). However, consistent with that observed by others (reviewed in Kozak, 1989a), the fidelity of translation from each AUG "initiation" codon appeared to be higher in the wheat germ extract system since secondary (or internal) translation products (both in- and out of frame), while being present (Figure 5.2.2, Panel B, lanes 1 and 2, Panels C, lanes 1 and 3), were less prominent than observed for similar translations involving the reticulocyte lysate system (Figure 5.2.1). Also, there was no instance in which wheat germ extract translation of Cat transcripts yielded unexplainable polypeptides of greater M_r than that predicted for the "full-length" Cat polypeptide. Further comparison with the reticulocyte lysate system revealed the wheat germ

extract system to be at least as competent for the translation of C-terminal, internal (drop-out) and N-/C-terminal deleted Cat transcripts as well (not shown).

Paralleling the synthesis of appropriately-sized Cat polypeptides, wheat germ extract translation of CAT transcripts [derived from plasmid TP408 (section 2.1.6)] yielded a major polypeptide of ≈25 kDa (Figure 5.2.2, Panels C, lane 6, arrowhead), as expected (the predicted M_r of CAT is 24000; Shaw, 1975). Similarly, wheat germ extract translation of DHFR transcripts [derived from plasmid pSP65DHFR-1 (section 2.1.6)] resulted in the synthesis of a single polypeptide of ≈20 kDa (Figure 5.2.2, Panels C, lane 7, arrowhead), which is consistent with a predicted M_r of 21500 for DHFR (Stone and Phillips, 1977; Nunberg et al., 1980).

5.3 IN VITRO STUDIES CONCERNED WITH THE IMPORT OF CATALASE INTO C. TROPICALIS PEROXISOMES

Having achieved the *in vitro* expression of the full-length and truncated Cat deletion mutants in both rabbit reticulocyte lysate and wheat germ extract systems, efforts were made to establish an assay which was competent for the import of Cat polypeptides into isolated *C. tropicalis* peroxisomes *in vitro*. Since Small et al. (1987), Small and Lazarow (1987) and Small et al. (1988) reported *C. tropicalis* growth and fractionation conditions for the

isolation of stable peroxisomes competent for the *in vitro* import of acyl-coenzyme A oxidase, initial work in this regard evaluated the potential of the aforementioned system to function in the same capacity for the *in vitro* translocation of *C. tropicalis* catalase.

Small et al. (1987) reported that the growth of *C. tropicalis* in oleate-containing medium induced an abundance of large (0.7 to 0.9 μm in diameter) peroxisomes which were too fragile for use in a cell-free import assay. Consequently, they surveyed several different growth media to determine those conditions which would induce peroxisome proliferation but yield smaller peroxisomes exhibiting increased stability. Growth medium containing 1% (w/v) Brij 35 (YPB medium; section 2.2.1) fulfilled this requirement, providing an adequate cell yield per litre from which medium (0.3 to 0.4 μm in diameter)-sized peroxisomes, purified by differential centrifugation followed by equilibrium density centrifugation in either linear sucrose or discontinuous Nycodenz gradients, exhibited a 3-fold increase in catalase activity relative to peroxisomes purified from *C. tropicalis* grown in YP medium alone and a 2- to 3-fold greater stability than that of oleate-induced organelles (Small et al., 1987). Accordingly, peroxisomes used herein as a component of the *in vitro* import assay were isolated from Brij-grown *C. tropicalis*. To achieve a greater separation

between peak peroxisomal and mitochondrial fractions, however, purification involved differential centrifugation followed by equilibrium density centrifugation in discontinuous sucrose gradients (section 2.2.2; see below). As seen in Figure 5.3.1, the distribution pattern of marker enzymes (catalase for peroxisomes; cytochrome c oxidase for mitochondria) for representative fractions collected from the discontinuous sucrose gradients indicates a good separation between the peak fractions containing peroxisomes (fraction 4) and mitochondria (fraction 9). This is better than the separation obtained with linear sucrose gradients in which the peak peroxisomal and mitochondrial fractions are 8 and 11, respectively (Small et al., 1987). The equilibrium densities of peroxisomes and mitochondria purified from Brij-grown C. tropicalis on discontinuous sucrose gradients, determined to be 1.21 g/cc and 1.18 g/cc, respectively (Figure 5.3.1), are slightly less than those of 1.23 g/cc and 1.20 g/cc, respectively, reported by Small et al. (1987); the reason for this difference is unknown. Mitochondrial contamination of the peak peroxisomal fraction was minimal; only 1.7% of the total cytochrome c oxidase activity present in the gradient was localized to fraction 4, similar to the value of 2.3% obtained by Small et al. (1987). The relative stability of the Brij-induced peroxisomes, determined by measuring the latency of catalase

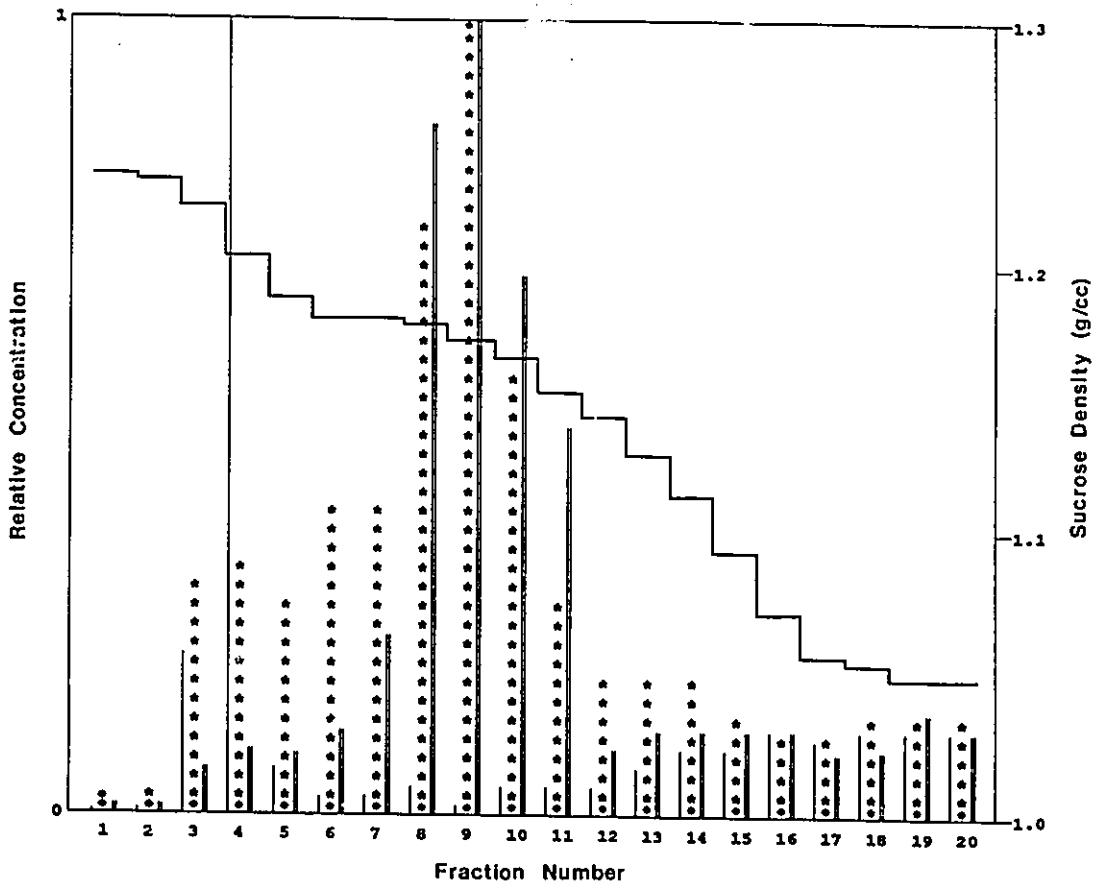


Figure 5.3.1. Distribution patterns of enzymes subsequent to subcellular fractionation of Brij-grown C. tropicalis. A crude peroxisomal fraction was isolated from Brij-grown C. tropicalis (as described in section 2.2.2.1) and purified by sucrose density step-gradient centrifugation (as described in section 2.2.2.2). Fractions (≈2 ml), represented along the abscissa scale in the order of their position in the gradient [ie. collected from bottom (left) to meniscus (right)], were analyzed for their average density (represented by the right axis; decreases ascending the gradient), protein content (vertical asterisks), and two marker enzyme activities, catalase (peroxisomal, single vertical line) and cytochrome c oxidase (mitochondrial, double vertical line). The relative concentration (represented by the left axis) of proteins, catalase and cytochrome c oxidase enzymes in each fraction was calculated as described by Beaufay et al. (1964) and normalized to one.

activity (ie. the relative difference in catalase activity pre- and post-detergent solubilization of the peroxisomal membrane), was high. Typically, Brij-induced peroxisomes, which had been purified in discontinuous sucrose-MOPS-EDTA gradients and concentrated (as described in section 2.6.2) for the purpose of *in vitro* import, exhibited between 65 and 75% catalase latency; this compares with 54% obtained by Small et al. (1987) for Brij-induced organelles purified in linear sucrose-MOPS-EDTA gradients and concentrated therefrom. To ensure that the integrity of the *C. tropicalis* peroxisomes was not compromized in the import buffer used by Small et al. (1987), catalase latency was determined at several intervals subsequent to purification, concentration and resuspension of peroxisomes in this buffer [2.5 mM MOPS (pH 7.2), 0.5 mM EDTA (pH 7.2), 0.5 M sucrose, 50 mM KCl]. Brij-induced peroxisomes were found to maintain 65 to 75% catalase latency for greater than 3 hours after resuspension in import buffer. In an effort to further standardize each import assay, attempts were made to determine suitable buffers systems which would allow Brij-induced peroxisomes to be frozen and thawed, and retain their membrane integrity. For this purpose, freshly-isolated *C. tropicalis* peroxisomes were resuspended in four different buffers, including the abovementioned import buffer of Small et al. (1987) and three other buffers used as components of the *in*

vitro import systems established for rat liver peroxisomes [10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 7.5), 5 mM methionine, 3 mM $MgCl_2$, 50 mM KCl, 0.5 M sucrose; Miyazawa et al., 1989], *Trypanosoma brucei* glycosomes [25 mM Tris-HCl (pH 7.9), 1 mM EDTA (pH 7.5), 1 mM DTT, 0.5 M sucrose; Dovey et al., 1988] and dog pancreatic microsomes (50 mM triethanolamine, 1 mM DTT, 250 mM sucrose; Walter and Blobel, 1983), and stored at -70°C. Unfortunately, after thawing, the catalase latency of Brij-induced peroxisomes was determined to be less than 10% in each of these buffers. This necessitated that the isolation and purification of 'fresh' peroxisomes from Brij-grown *C. tropicalis* be prerequisite to all subsequent import assays.

Post-purification and concentration of Brij-induced *C. tropicalis* peroxisomes, the *in vitro* import assay of Small et al. (1987) requires 300 µg of peroxisomal protein (resuspended in 60 µl of import buffer supplemented with ATP to 1 mM) to be combined with 60 µl of reticulocyte lysate translation mixture containing the ^{35}S-labeled polypeptides (supplemented with sucrose to 0.5 M) and the mixture incubated for 30 minutes at 26°C. The import assay is then divided into three 40 µl aliquots which are, or are not, treated with protease (eg. proteinase K) for 15 minutes (on ice) in the presence, or absence, of detergent (0.8% DOC, 0.8% Triton X-100). Protease treatment is terminated by the

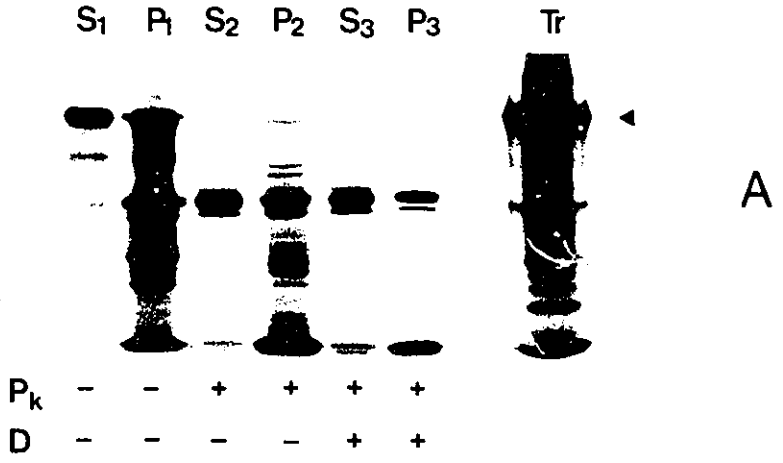
addition of protease inhibitors (Small et al., 1987); supernatants and peroxisome pellets are then separated by microcentrifugation (14,000 rpm) for 15 minutes at 4°C and fractions of each analyzed by SDS-PAGE and flourography. The rationale involved in the identification of imported polypeptides is that internalization within the peroxisome renders them inaccessible to (ie. protected from) externally added protease. After solubilization of the peroxisomal membrane with detergent, however, the previously imported (protected) polypeptides are again susceptible to protease digestion. Thus, subsequent to analysis by SDS-PAGE and flourography, protease treatment alone, compared with protease and detergent treatment combined, would reveal the presence, and absence, respectively, of ^{35}S-labeled, *in vitro* imported, polypeptides.

Initial attempts at the *in vitro* import of full-length catalase into isolated *C. tropicalis* peroxisomes, using a protocol which paralleled that reported by Small and Lazarow (1987) and Small et al. (1988) {see above; including protease treatment with proteinase K [20 µg/40 µl assay; Small et al. (1988)]}, were negative. Subsequent titration of proteinase K against a constant amount (≈100 µg) of peroxisomal protein (the latency of which ranged between 65 and 75%, depending on the experiment) suggested, however, that concentrations of proteinase K greater than 2.0 to 3.0

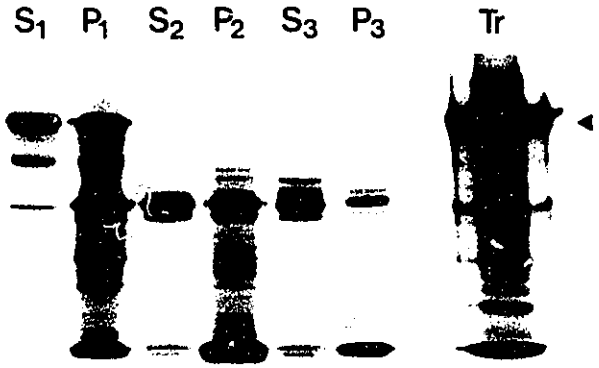
µg/100 µg of peroxisomal protein may be excessive, progressively abolishing the protein profile observed to be characteristic of purified C. tropicalis peroxisomes following resolution of peroxisome-specific proteins by SDS-PAGE. Additional work in light of this revealed that only at low concentrations of proteinase K (eg. 0.1 and 0.3 µg/40 µl assay) did there appear to be "imported" catalase (Figure 5.3.2, panels A and B, respectively, compare lanes P_2 and P_3). Increased proteinase K concentrations (1 µg/40 µl assay or greater) repeatedly abolished any indication of imported protein (Figure 5.3.2, panel C, compare lanes P_2 and P_3). This became troublesome when it was determined that the integrity of the peroxisomal membrane (ie. catalase latency) is maintained in the presence of 15.0 to 20.0 µg of proteinase K/40 µl assay. Further doubt concerning the putative import of full-length catalase into isolated C. tropicalis peroxisomes originated when reproducibility, at low concentrations of proteinase K, could not be attained.

It is conceivable that a full-length, compactly-folded Cat polypeptide is not competent for peroxisomal membrane translocation and, as a consequence, remains "cytosolic" and protease-accessible in vitro. Similar to the situation reported for certain mitochondrial (Eilers et al., 1988; Vestweber and Schatz, 1988; Verner and Lemire, 1989) and chloroplast (della-Cioppa and Kishore, 1988)

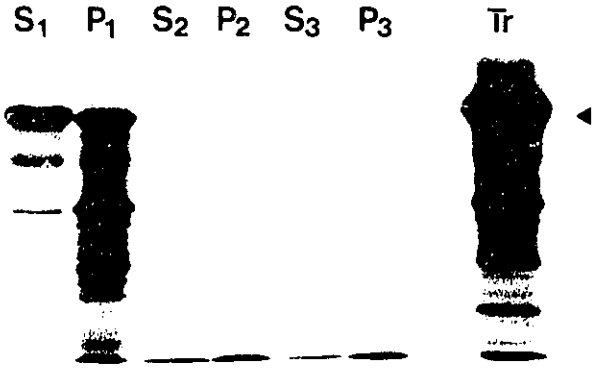
Figure 5.3.2. In vitro studies concerned with the targeting of full-length Cat to peroxisomes isolated from C. tropicalis. Each post-translational import assay was conducted as described by Small et al. (1987, 1988). To this end, 60 µl of rabbit reticulocyte lysate translation products (≈4 x 10^5 cpm; supplemented with sucrose to 0.5 M) were mixed with 300 µg (60 µl) of sucrose density step-gradient purified peroxisomes (resuspended at 5 µg/ml in import buffer containing 2 mM ATP, see section 2.6.2) isolated from Brij-grown C. tropicalis. The mixture was incubated at 26°C for 30 min and subsequently divided into three equal samples (40 µl each). Each sample was (+) or was not (-) treated with proteinase K (P_k) for 15 min on ice, in the presence (+) or absence (-) of detergent (D) [0.8% (w/v) DOC, 0.8% (w/v) Triton X-100]. Proteolysis was terminated by the addition of PMSF to 1 mM. Supernatants [S_1, S_2 or S_3; defined by treatment (see Panel A)] and peroxisome pellets [P_1, P_2 or P_3; defined by treatment (see Panel A)] were then separated by microcentrifugation at 4°C for 15 min. The supernatants were aspirated and retained; peroxisome pellets were washed with import buffer (100 µl) and recentrifuged. One-tenth of the supernatant and one-half of the peroxisome pellet fractions were subsequently analyzed by SDS-PAGE and fluorography. Panels A, B and C represent protease treatment with 0.1, 0.3 and 1.0 µg of proteinase K/40 µl assay, respectively. Lanes Tr indicate the rabbit reticulocyte lysate translation products (1 µl of the translation) synthesized from pGEMC2-11b transcripts. The position of full-length Cat is indicated by a closed triangle. The fluorograms were exposed for 96 h at -70°C with Kodak XAR-5 film.



A



B



C

proteins, a more translocation competent peroxisomal protein may be that which possesses a loosely-folded conformation. It was postulated, therefore, that truncated Cat polypeptides, which may adopt "unfolded" or "denatured" structures relative to the full-length Cat polypeptides, may exhibit higher efficiencies of import into C. tropicalis peroxisomes in vitro. This would parallel that observed for the in vitro import of acyl-coenzyme A oxidase into isolated C. tropicalis peroxisomes (Small et al., 1988). To test this notion, the import efficiencies of several of the truncated Cat polypeptides depicted in Figure 5.1.2 were determined using the in vitro import assay of Small et al. (1988). Unfortunately, the results obtained using the N-terminal deletion Cat polypeptides C2-16b, C2-16n, C2-14Af, C2-14Ab and C2-16Ac, the C-terminal deletion Cat polypeptides C2-11bAha II and C2-11bEcoR I, the internal deletion Cat polypeptides C2-11bE_D and C2-11bK_D and the N-/C-terminal deletion Cat polypeptide C2-16bAH paralleled that obtained with the full-length Cat polypeptide C2-11b; that is, the use of proteinase K concentrations between 1 and 3 µg/40 µl assay revealed an absence of import of truncated Cat polypeptides into peroxisomes isolated from Brij-grown C. tropicalis (as determined by two independent experiments; not shown).

Since the import of full-length or truncated Cat

polypeptides into peroxisomes isolated from C. tropicalis was not operative using the in vitro system reported by Small et al. (1988), efforts were made to modify this import system, through further characterization and appropriate change(s) or through supplementation, in an attempt to achieve some degree of translocation competence for peroxisomal proteins. Although several of the efforts undertaken in this regard yielded negative results (and are, for the most part, not shown), the alternative approaches and/or modifications, and the reasons for which they were explored, are outlined below:

i) Small et al. (1987) reported the specific and temperature-dependent, in vitro association of ten major rabbit reticulocyte lysate translation products (synthesized from C. tropicalis total RNA and immunoprecipitated by antiserum against peroxisomal protein) with peroxisomes isolated from Brij-grown C. tropicalis. As a result, poly(A)$^+$RNA isolated from oleic acid-grown C. tropicalis, which is similarly known to encode several different peroxisomal proteins (Kamiryo and Okazaki, 1984), was translated in a rabbit reticulocyte lysate system and the ^{35}S-labeled polypeptides used in the in vitro import assay of Small et al. (1988). In addition, constructs used for the in vitro expression of full-length and N-terminal truncated C. tropicalis hydratase-dehydrogenase-epimerase

(HDE), as well as full-length C. tropicalis acyl-coenzyme A oxidase (AOx), were transcribed, translated (using a rabbit reticulocyte lysate system) and the corresponding polypeptides used to assay in vitro import. The results of all translocation experiments were negative.

ii) Urea denaturation of certain mitochondrial proteins has been shown to enhance their import efficiency (Eilers et al., 1988; Verner and Lemire, 1989). In a related effort, the urea-induced denaturation of full-length catalase was carried out prior to attempts at importing the unfolded polypeptide. The results indicated an absence of in vitro import.

iii) The growth of C. tropicalis in different media yields purified peroxisomes of various sizes and stabilities which are characterized further by differences in their content and activities of the β-oxidation enzymes (Small et al., 1987). Although reported to exhibit decreased stability (Small et al., 1987), peroxisomes purified from oleic acid-grown C. tropicalis (which yielded average values for catalase latency ranging between 40 and 50%) were used in the in vitro import assay to determine whether they possessed an enhanced ability (relative to peroxisomes purified from Brij-grown C. tropicalis) to import full-length or truncated Cat polypeptides. The results of all experiments were negative.

iv) Lazarow and Fujiki (1985) reported that a high speed supernatant (or cytosolic) fraction (isolated from rat liver) was found to markedly stimulate the in vitro import of ^{35}S-labeled catalase into purified rat liver peroxisomes. For this reason, a C. tropicalis cytosolic fraction (ie. a 20k x g supernatant fraction; section 2.2.2.1) was used to supplement the in vitro import system of Small et al. (1988). Unfortunately, the import system remained inoperative.

v) Although Lazarow and Fujiki (1985) demonstrated that rat liver catalase is imported in vitro into purified rat liver peroxisomes more efficiently than into peroxisomes present in a postnuclear supernatant fraction, the converse was found to be true of rat liver acyl-coenzyme A oxidase. Using a similar import system (ie. an in vitro system based upon the peroxisomal content of a crude postnuclear supernatant fraction isolated from rat liver), Miyazawa et al. (1989) extended this work by demonstrating that the PTS of rat liver acyl-coenzyme A oxidase is contained within its C-terminal five amino acids. To determine whether an in vitro import system for C. tropicalis peroxisomal proteins could be established using crude peroxisomal fractions, attempts were made to import full-length and truncated Cat and HDE polypeptides into postnuclear supernatant and 20k x g pellet fractions isolated from Brij-grown C. tropicalis.

As seen in Figure 5.3.3, proteinase K treatment alone (lanes P_2 in all panels), compared to proteinase K and detergent treatment combined (lanes P_3 in all panels), of the import assays which utilized PNS fractions and either full-length (C2-11b for Cat, 32A for HDE) or the N-terminal truncated (C2-14Af for Cat, 35A for HDE) polypeptides, revealed an absence of imported protein. Although not shown, the results obtained for the import assays which involved the use of 20k x g pellet fractions and either of the aforementioned Cat or HDE polypeptides were similarly indicative of a lack of *in vitro* import.

vi) Small et al. (1987) established a system for the efficient *in vitro* association of translation products, derived from rabbit reticulocyte lysates, with purified *C. tropicalis* peroxisomes. Furthering this work, Small et al. (1988) demonstrated, using an *in vitro* system supplemented with either wheat germ extract or rabbit reticulocyte lysate translation products, that the import of acyl-coenzyme A oxidase into purified *C. tropicalis* peroxisomes could be mediated by two independent internal targeting signals. Since the use of two different protein synthesizing systems was unexplained, both were analyzed to determine that which was better suited to function in the import assay of Small et al. (1988). To this end, mock import assays containing purified peroxisomes (isolated from Brij-grown *C.*

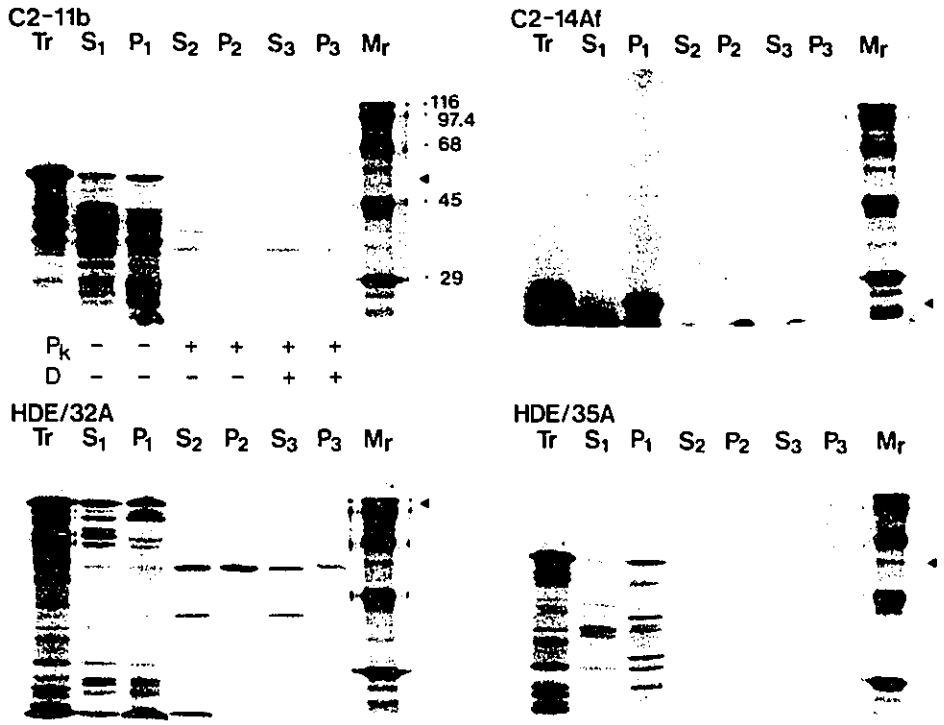


Figure 5.3.3. Studies concerned with the in vitro targeting of Cat and HDE polypeptides to crude peroxisomal fractions isolated from C. tropicalis. Post-translational import assays were conducted essentially as described in the legend of Figure 5.3.2 using one of two different peroxisomal (crude) fractions: either a post-nuclear supernatant (PNS) fraction or a 20k x g pellet fraction (each isolated as described in section 2.2.2.1). The PNS import mixture (120 µl) was adjusted to contain 100 µl of PNS fraction (obtained from the first homogenization of spheroplasts), as well as rabbit reticulocyte lysate translation products (10 µl) and appropriate supplements as described in Figure 5.3.2. The 20k x g pellet import mixture contained 600 µg (60 µl) of the 20k x g pellet as the organellar fraction. All subsequent manipulations were as described in Figure 5.3.2. The fluorograms presented are those resulting from the PNS import assays only; however, they are representative of those generated by similar import studies with 20k x g pellet fractions. Supernatant (S_1, S_2 and S_3) and pellet (P_1, P_2 and P_3) fractions are defined by treatment as indicated in Panel C2-11b. Lanes T_r show the rabbit reticulocyte lysate translation products (1 µl of the translation) synthesized from the C2-11b, C2-14Af, HDE/32A and HDE/35A transcripts, respectively. The position of each "full-length" translation product is indicated by a closed triangle. The molecular mass protein standards (in kDa) seen in lanes M_r include: E. coli β-galactosidase (116), phosphorylase b (97.4), bovine serum albumin (68), chicken egg albumin (45) and carbonic anhydrase (29). Symbols represent: P_k, proteinase K (1 µg/40 µl assay); D, detergent [0.8% (w/v) DOC, 0.8% (w/v) Triton X-100]; +, present; -, absent. Exposures were for 42 h at -70°C with Kodak XAR-5 film.

tropicalis) and various amounts of either rabbit reticulocyte lysate or wheat germ extract, were assayed for catalase latency to determine the effect, if any, of either translation system on peroxisomal integrity. The inclusion of reticulocyte lysate to 50% of the import mixture, the amount suggested by Small et al. (1987, 1988), resulted in the complete lysis (ie. no detectable catalase latency) of C. tropicalis peroxisomes in vitro. This was true of reticulocyte lysates purchased from commercial sources (section 2.1.1) or obtained from Dr. D.W. Andrews (Department of Biochemistry, McMaster University), irrespective of whether the lysate was salt-exchanged over a Sephadex column equilibrated with import buffer. Even small amounts of each of these reticulocyte lysates were found to cause appreciable lysis of peroxisomes; values for the amount of reticulocyte lysate included in the import mixture (in %) relative to the average catalase latency observed (in %) were determined to be 37.5: 0, 12.5: 16, 8.3: 38, 4.2: 52, 1.7: 56 and zero (ie. latency in import buffer alone): 82. This would imply that import assays which involved the use of the N-terminal truncated Cat polypeptides were performed under less than ideal conditions since the inherently lower levels of ^{35}S-methionine incorporation necessitated the inclusion of increased amounts of reticulocyte lysate translation mixture (relative to the 1

or 2 µl used for the full-length Cat polypeptide), thereby compromizing peroxisomal integrity to various extremes. To the contrary, the inclusion of wheat germ extract (purchased from Gibco/BRL or prepared as described in section 2.6.1.4) to as much as 37.5% of the import mixture caused no significant lysis of C. tropicalis peroxisomes in vitro; values for the amount of wheat germ extract included in the import mixture (in %) relative to the average catalase latency observed (in %) were determined to be 37.5: 80, 25: 76, 8.3: 78, 1.7: 78 and zero (ie. latency in import buffer alone): 83. For this reason, subsequent import assays were conducted with ^{35}S-labeled polypeptides derived from wheat germ extract translations (see below).

vii) Investigations concerned with protein targeting to peroxisomes have characterized the in vitro association of rat liver catalase (Lazarow and Fujiki, 1985), rat liver acyl-coenzyme A oxidase (Lazarow and Fujiki, 1985; Miyazawa et al., 1989) and C. tropicalis acyl-coenzyme A oxidase (Small et al., 1987, 1988; Small and Lazarow, 1987) with peroxisomes as specific based upon at least two different criteria: (i) peroxisomal polypeptides do not associate (ie. in the absence of either protease or protease plus detergent treatment, peroxisomal polypeptides do not, subsequent to microcentrifugation, pellet) with the mitochondrial fraction when mitochondria are substituted for peroxisomes in the

import assay and (ii) nonperoxisomal polypeptides do not associate (as defined above) with the peroxisomal fraction. To determine the extent and specificity of the association of catalase polypeptides (synthesized by wheat germ extract translation) with peroxisomes, import assays were conducted using sucrose-step gradient purified peroxisomes or mitochondria (isolated from Brij-grown C. tropicalis) as the organellar fraction; included as a control, was a mock import assay in which the organellar fraction was replaced by import buffer alone. As depicted in Figure 5.3.4, the apparent association of 40 to 50% of the total ^{35}S-labeled catalase used in the import assay with the non-protease/non-detergent-treated peroxisomal fraction (PEROXISOMES panel, lane P_1, arrowhead) was not indicative of a specific association of catalase with C. tropicalis peroxisomes since catalase was also pelleted by microcentrifugation in the absence of either a peroxisomal or mitochondrial fraction (CONTROL panel, Lane P_1, arrowhead). This finding necessitated the use of an alternative procedure to assay the extent and specificity of the in vitro association between wheat germ extract translated polypeptides and purified C. tropicalis peroxisomes. Related studies concerned with ER membrane translocation have demonstrated that for most proteins, excellent physical separation of membrane- and non-membrane-associated species can be

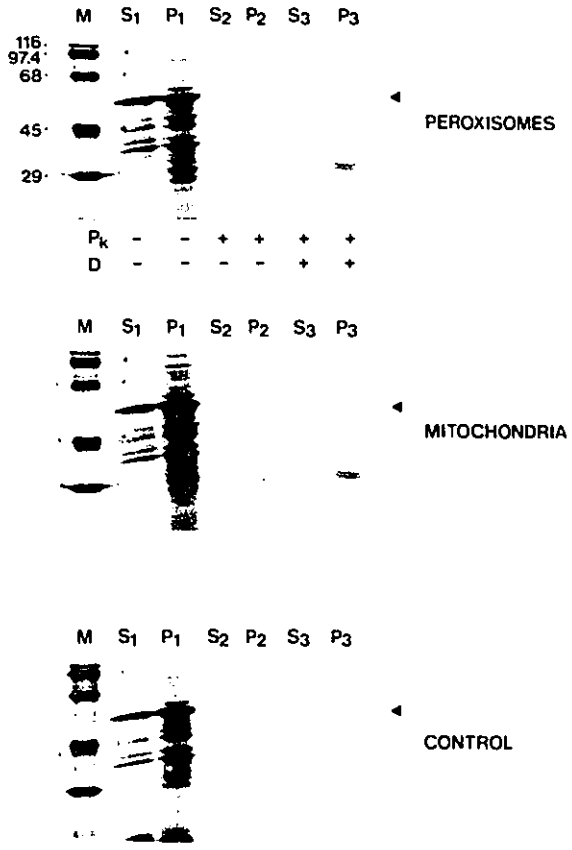
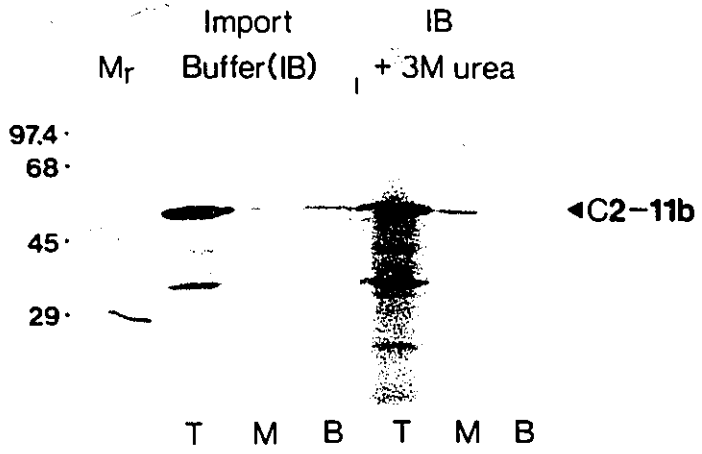


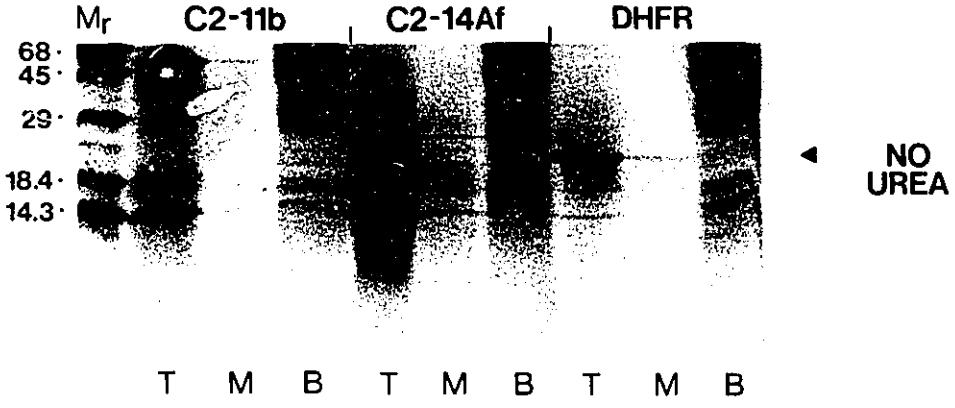
Figure 5.3.4. In vitro studies concerned with the targeting of wheat germ extract translation products to peroxisomes isolated from C. tropicalis. Post-translational import assays were conducted as described in the legend of Figure 5.3.2, with modifications as follows. Full-length Cat translation products (60 µl, supplemented with sucrose to 0.5 M), synthesized using a nuclease-treated, wheat germ extract, were mixed with either 300 µg (60 µl) of sucrose density step-gradient purified peroxisomes or mitochondria (isolated from Brij-grown C. tropicalis and resuspended in import buffer supplemented with ATP to 2 mM) or 60 µl of import buffer alone (control). All subsequent manipulations were as detailed in Figure 5.3.2. Supernatant (S_1, S_2 and S_3) and pellet (P_1, P_2 and P_3) fractions are defined by treatment as indicated for the PEROXISOMES panel. The position of full-length Cat is indicated by a closed triangle. The molecular mass protein standards (in kDa) seen in lanes M include: E. coli β-galactosidase (116), phosphorylase b (97.4), bovine serum albumin (68), chicken egg albumin (45) and carbonic anhydrase (29). Symbols represent: P_k, proteinase K (1 µg/40 µl assay); D, detergent [0.8% (w/v) DOC, 0.8% (w/v) Triton X-100]; +, present; -, absent. Exposures were for 36 h at -70°C with Kodak XAR-5 film.

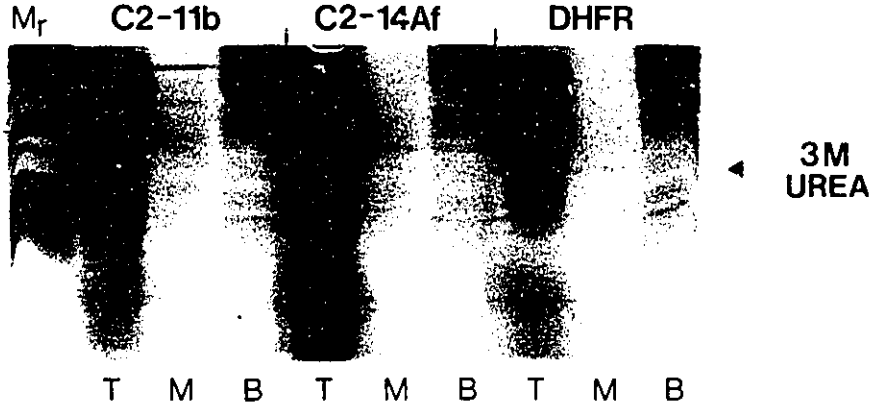
achieved by centrifugation through sucrose cushions (Andrews, 1989). Typically, membrane-association is assayed by centrifugation of translocation assays (which are differentially supplemented with or without 2M urea subsequent to carrying out the assay) through sucrose cushions which contain similar concentrations of urea. In the absence of urea, membrane-associated polypeptides are pelleted (and resuspended as the bottom fraction) whereas non-membrane-associated polypeptides remain in the sucrose cushion (and are removed as the top fraction). In the presence of 2M urea, however, previously membrane-associated proteins are released into the sucrose cushion due to the urea-induced disruption of the ER membrane. The transfer of ^{35}S-labeled polypeptides from the bottom fraction (in the absence of 2M urea) to the top fraction (in the presence of 2M urea) thereby establishes membrane-association for those polypeptides. Prior to investigating the association of C. tropicalis catalase with purified C. tropicalis peroxisomes, it was determined that 3M urea was required for the complete disruption of the peroxisomal membrane (ie. to achieve zero catalase latency). Further, mock import assays (from which peroxisomes were omitted) were conducted to ensure that full-length and N-terminal truncated C. tropicalis catalase polypeptides, when synthesized in vitro by wheat germ extract translation, were not, in the absence of a membrane

fraction, pelleted upon centrifugation through the sucrose cushion. As seen in Figure 5.3.5 (panel C2-11b), mock association assays containing either import buffer alone or import buffer supplemented with 3M urea exhibited the required profiles; that is, the majority of the full-length Cat polypeptides (indicated by the arrowhead) were retained in the top (T) fraction [ie. the initial 82 µl which was removed from the 168 µl (final volume) sucrose cushion] compared to the only minor amounts detected in the middle (M) (ie. the remaining 82 µl of the cushion) and bottom (B) (ie. the resuspended pellet) fractions (see section 2.6.2). The N-terminal truncated Cat polypeptide (C2-14Af) and DHFR (used in the assay as a cytosolic control not expected to associate with purified C. tropicalis peroxisomes) exhibited comparable profiles for the mock association assays (not shown). The association profile exhibited by DHFR in the actual assay was similar to that seen for the mock assays in that the majority of the ^{35}S-labeled DHFR polypeptides were localized to the top (T) fraction of the sucrose cushion, irrespective of the presence of urea (Figure 5.3.5, compare polypeptide profiles indicated by arrowheads in DHFR/no urea and DHFR/3M urea panels). Unfortunately, the profile observed for the full-length Cat polypeptide (C2-11b) was not indicative of a large degree of association with purified C. tropicalis peroxisomes since, similar to DHFR,

Figure 5.3.5. Analysis of the *in vitro* association of Cat polypeptides with peroxisomes isolated from *C. tropicalis*. Each association assay was carried out in the absence (middle panel) or presence (lower panel) of 3 M urea as described in section 2.6.2 (illustrated in Figure 5.3.6), using wheat germ extract translation products synthesized from pGEMC2-11b, pGEMC2-14Af or pSP65DHFR-1 transcripts. Controls for the association assay included assessing the pelletability (i.e. lack thereof) of full-length Cat in import buffer (IB) alone and IB supplemented with 3 M urea (upper panel). Three 82 µl fractions, called top (T), middle (M) and bottom (B), were collected per gradient; equal volumes (10 µl) from each fraction were analyzed by SDS-PAGE and fluorography. The position of each "full-length" translation product is indicated by a closed triangle. The molecular mass protein standards (in kDa) seen in lanes M_r include: phosphorylase b (97.4), bovine serum albumin (68), chicken egg albumin (45), carbonic anhydrase (29), β-lactoglobulin (18.4) and egg white lysozyme (14.3). Exposures were for 17 h (upper panel) and 84 h (middle and lower panels) at -70°C with Kodak XAR-5 film.


Import
Buffer(IB)
IB
+ 3M urea
Mr
97.4
68
45
29
C2-11b
T M B T M B

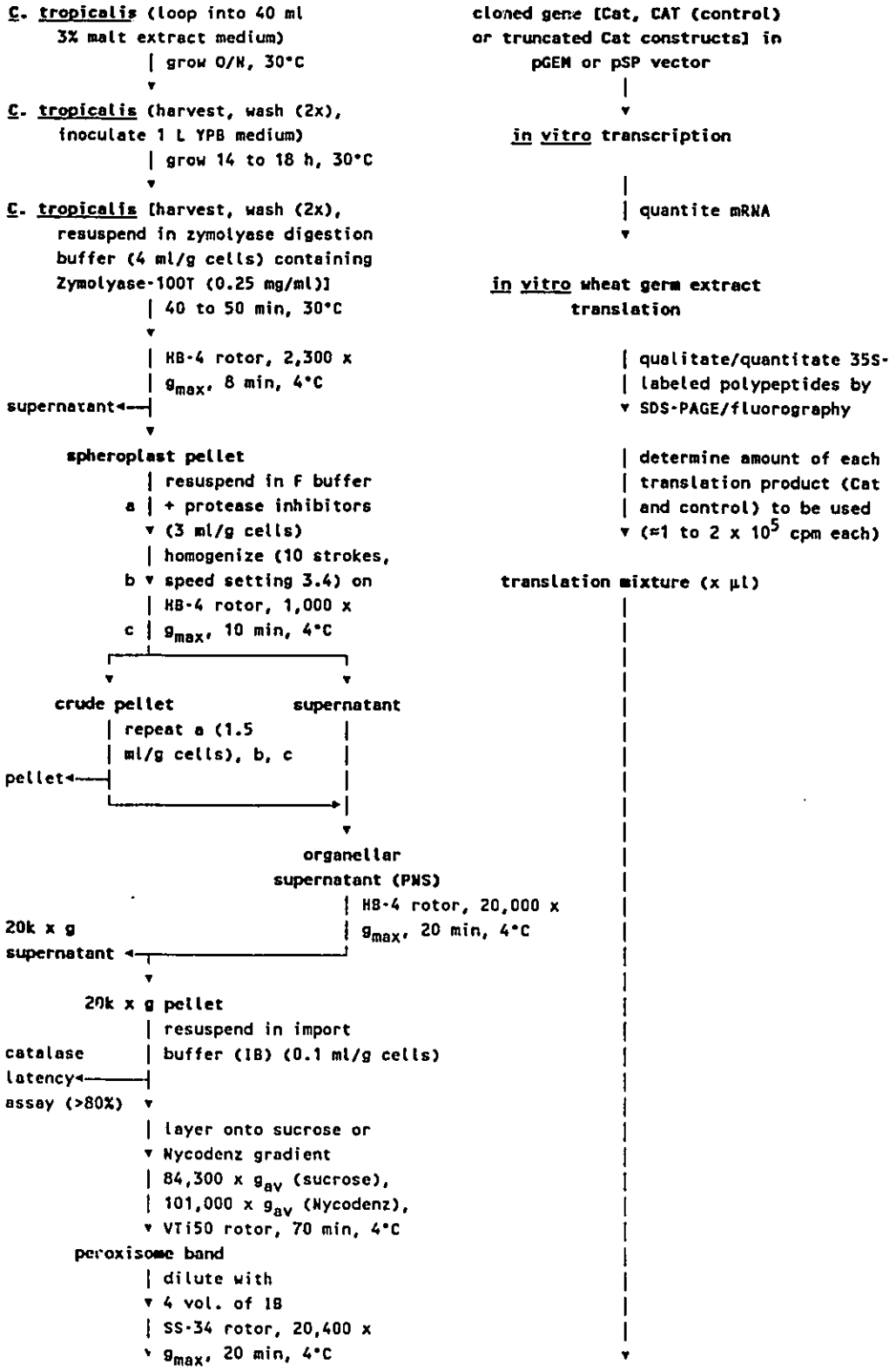

Mr
C2-11b
C2-14Af
DHFR
68
45
29
18.4
14.3
NO UREA
T M B T M B T M B


Mr
C2-11b
C2-14Af
DHFR
3M UREA
T M B T M B T M B

the majority of the labeled polypeptides were retained in the top (T) fraction in the absence of urea (Figure 5.3.5, C2-11b/no urea panel, compare T and B fractions at position of arrowhead). By comparison, however, the N-terminal truncated Cat polypeptide (C2-14Af) appeared to show some (albeit small) degree of association with purified C. tropicalis peroxisomes. This was evidenced by the presence of ^{35}S-labeled Cat (C2-14Af) polypeptides in the bottom (B) fraction of the sucrose cushion without urea compared to their absence from the bottom fraction of the sucrose cushion containing 3M urea (Figure 5.3.5, compare polypeptide profiles indicated by arrowheads in C2-14Af/no urea and C2-14Af/3M urea panels). These preliminary results may imply that the truncated (and possibly, less compactly-folded) Cat polypeptides exhibit a more efficient in vitro association with purified C. tropicalis peroxisomes than do the full-length (more compactly-folded) Cat polypeptides.

Collectively, the aforementioned results demonstrated that the in vitro import system reported by Small et al. (1988) was not functional for the translocation of catalase (and several other C. tropicalis peroxisomal proteins, including HDE and AOx) into peroxisomes isolated from C. tropicalis. Related to this, it was determined further that certain of the assay conditions employed by Small et al. (1988) were less than ideal for maintaining

peroxisomal integrity during import, for assessing the degree of the association of *in vitro* translated polypeptides with purified *C. tropicalis* peroxisomes *in vitro* or for determining the specificity and efficiency of observed import (see below). Consequently, the import assay of Small et al. (1988) was modified to yield an *in vitro* import procedure (Figure 5.3.6) which included the use of: (i) ^{35}S-labeled polypeptides synthesized by wheat germ extract translation only, (ii) a documented membrane association assay (Andrews, 1989) to assess the degree and specificity of the *in vitro* association of wheat germ extract translated polypeptide(s) with *C. tropicalis* peroxisomes (this replaced the analysis of Small et al. (1988) [item (vii) above] in which *in vitro* translated polypeptides that pelleted (upon microcentrifugation alone) with the peroxisomal fraction were considered to be specifically-associated with peroxisomes) and (iii) an assay system consisting of two mock and three import reactions, each of which contained an ^{35}S-labeled polypeptide that served as a negative (or cytosolic) control for import and was differentially treated as indicated in Figure 5.3.6, thereby allowing a direct determination of the efficiency and specificity of observed import. The titration of increasing amounts (from 1 to 30 µg) of proteinase K per 40 µl import reaction revealed that CAT (a known cytosolic


C. tropicalis (loop into 40 ml 3% malt extract medium)
grow O/N, 30°C
C. tropicalis (harvest, wash (2x), inoculate 1 L YPB medium)
grow 14 to 18 h, 30°C
C. tropicalis [harvest, wash (2x), resuspend in zymolyase digestion buffer (4 ml/g cells) containing Zymolyase-100T (0.25 mg/ml)]
40 to 50 min, 30°C
HB-4 rotor, 2,300 x g max, 8 min, 4°C
supernatant
spheroplast pellet
a resuspend in F buffer + protease inhibitors (3 ml/g cells)
b homogenize (10 strokes, speed setting 3.4) on
c HB-4 rotor, 1,000 x g max, 10 min, 4°C
crude pellet
repeat a (1.5 ml/g cells), b, c
pellet
supernatant
organellar supernatant (PNS)
HB-4 rotor, 20,000 x g max, 20 min, 4°C
20k x g supernatant
20k x g pellet
resuspend in import buffer (IB) (0.1 ml/g cells)
catalase latency assay (>80%)
layer onto sucrose or Nycodenz gradient
84,300 x g av (sucrose), 101,000 x g av (Nycodenz),
VTi50 rotor, 70 min, 4°C
peroxisome band
dilute with 4 vol. of IB
SS-34 rotor, 20,400 x g max, 20 min, 4°C
cloned gene [Cat, CAT (control) or truncated Cat constructs] in pGEM or pSP vector
in vitro transcription
quantite mRNA
in vitro wheat germ extract translation
qualitate/quantitate 35S-labeled polypeptides by SDS-PAGE/fluorography
determine amount of each translation product (Cat and control) to be used (≈1 to 2 x 10^5 cpm each)
translation mixture (x µl)

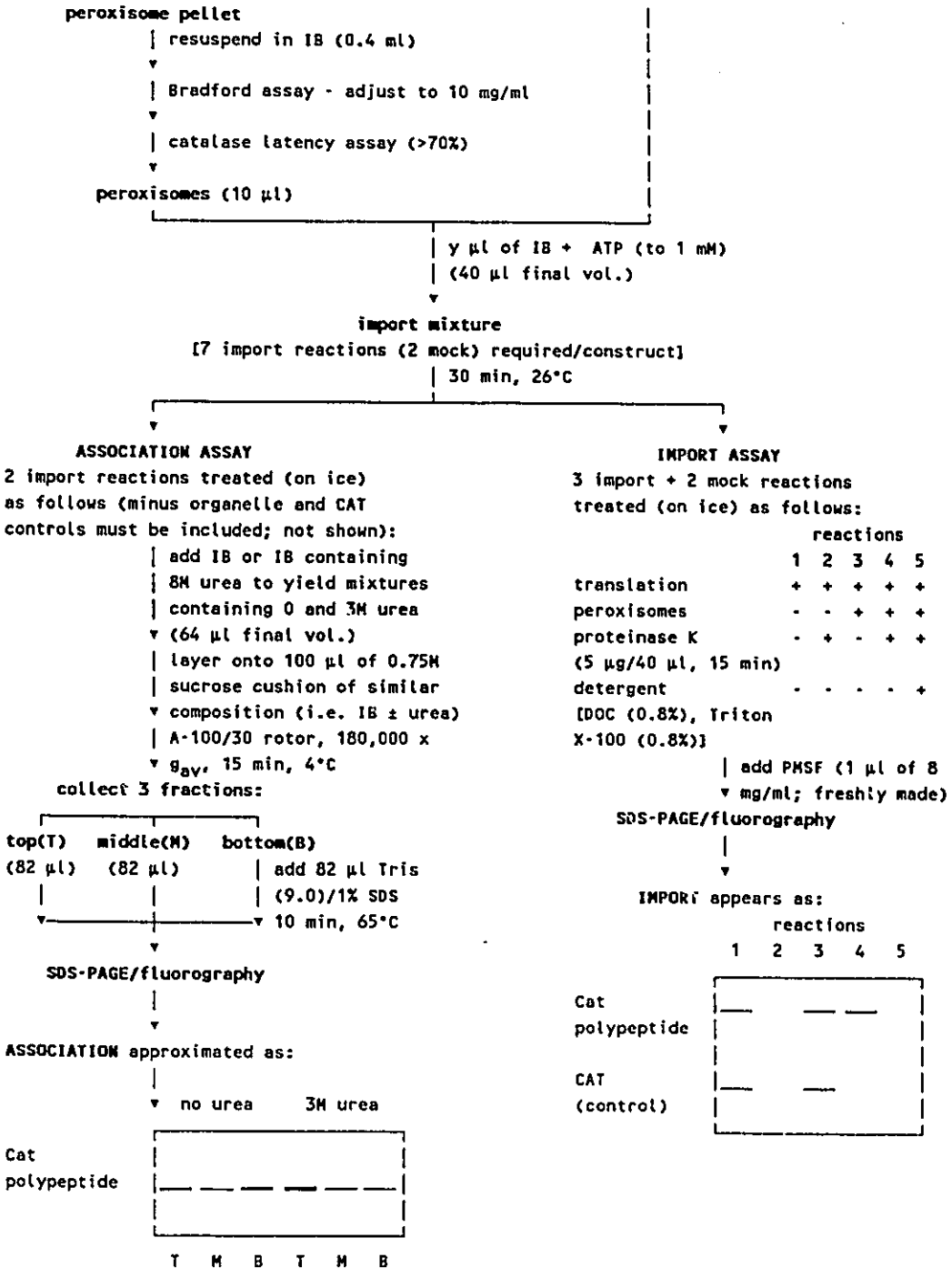


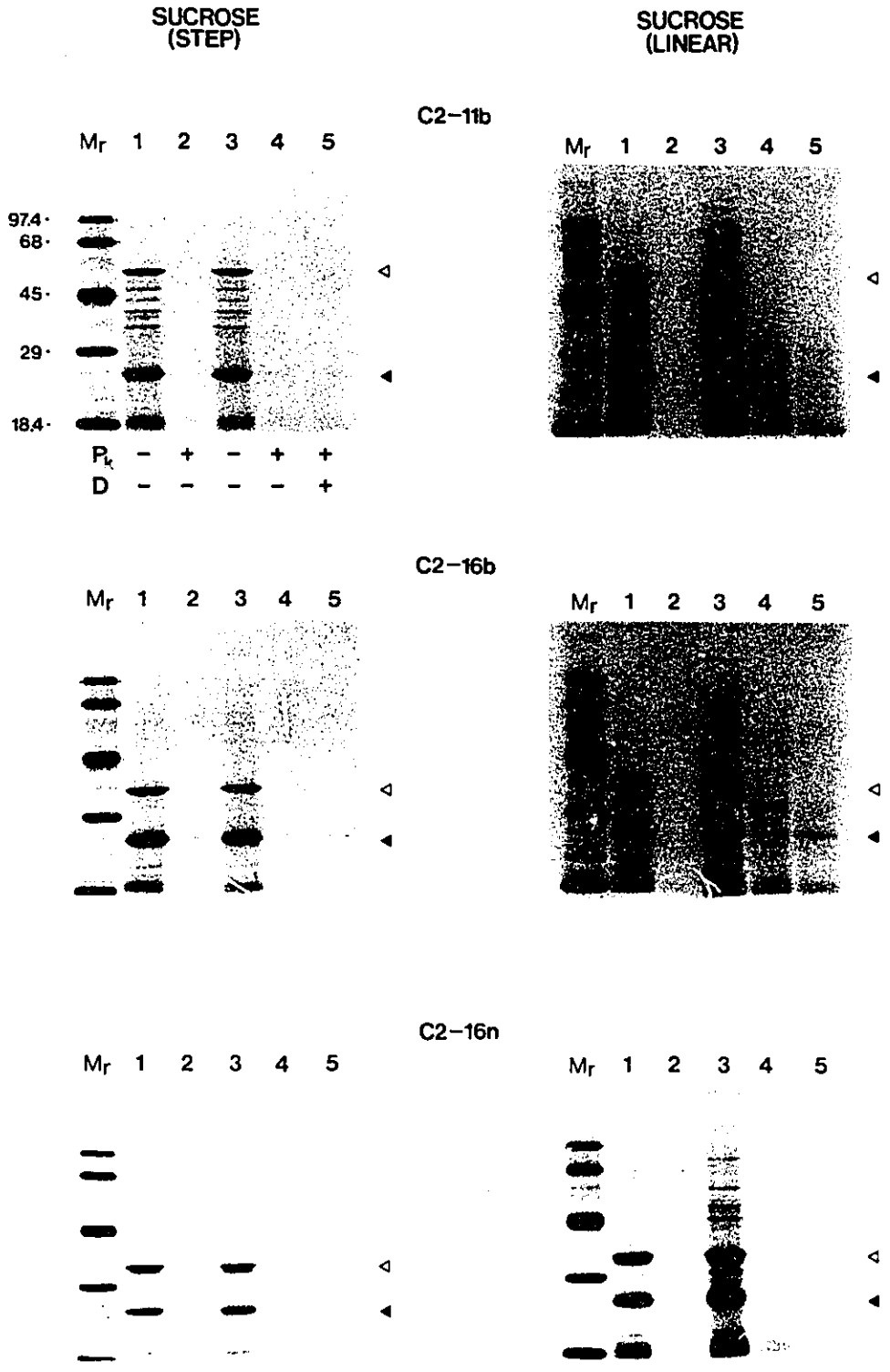
Figure 5.3.6. A modified in vitro procedure to analyze the targeting of proteins to peroxisomes isolated from C. tropicalis. The methods are presented in condensed form; more detailed descriptions are found in the following sections: yeast cell culture (2.2.1), peroxisome isolation (2.2.2), in vitro transcription (2.6.2.1), in vitro translation (2.6.1.4) and association and import assays (2.6.2). Fluorograms depicting protein association with and import into peroxisomes, respectively, are illustrated at the end of the figure.

protein; Gould et al., 1987, 1988) would function as an acceptable control (or cytosolic) polypeptide for import. Since Cat was found to be more susceptible to proteolysis than CAT and because the amounts of radiolabeled Cat and CAT included per reaction were similar (≈1 to 2 x 10^5 cpm), the complete proteolysis of CAT necessarily implied that nonproteolyzed Cat (detected subsequent to proteinase K treatment) would be the result of protection afforded by the peroxisomal membrane (ie. import) (Mouse DHFR was found to be quite resistant to proteinase K treatment and was not ideal for this purpose.). The titration experiment also revealed that only 5 μg of proteinase K was required for the complete proteolysis of the ^{35}S-labeled Cat and CAT polypeptides added to each import reaction, an amount previously shown not to compromize peroxisomal integrity.

Having amended the import assay of Small et al. (1988) to yield the modified protocol depicted in Figure 5.3.6, further efforts were made to attain the import of full-length and truncated Cat polypeptides into isolated C. tropicalis peroxisomes in vitro. To allow for differences observed in the purity of C. tropicalis peroxisomes isolated from gradients of various composition (typically, sucrose or Nycodenz gradients; Small et al., 1987; Nuttley et al., 1990), and address the possibility that the efficiency of in vitro import may be related to the degree of peroxisome

purity [such that a requirement for a cytosolic factor(s) may be more readily met by an "impure" peroxisome preparation], stable peroxisomes isolated from Brij-grown C. tropicalis were purified by equilibrium density centrifugation in one of three different gradients, either discontinuous sucrose or Nycodenz gradients or continuous sucrose gradients. Upon withdrawal from the sucrose or Nycodenz discontinuous (or step-) gradients, previously banded particulate fractions were analyzed for catalase activity and latency (which ranged from 65 to 75% and 40 to 50% for sucrose and Nycodenz step-gradients, respectively) and the peak fractions of catalase activity used for the import assay. Interestingly (but unexplainedly), select Nycodenz step-gradients yielded two "peroxisomal" fractions to be used for import [termed peroxisomes middle (P_M) and peroxisomes bottom (P_B), of discernable density ($P_M<P_B$) but similar in amount (ie. total protein), protein profile (which was determined subsequent to SDS-PAGE and staining with Coomassie blue), catalase activity and latency. Continuous (or linear) sucrose gradients, which did not yield discrete particulate bands, were fractionated and the peak fraction of catalase activity used for import (typical latency values for these peroxisomes ranged from 50 to 60%). Peroxisomal fractions were mixed with wheat germ extract translations of the full-length Cat polypeptide C2-11b or

the N-terminal truncated Cat polypeptides C2-16b and C2-16n (N-terminal truncated polypeptides were chosen because of the small degree of association previously observed [see item (vii) above] between the N-terminal truncated Cat polypeptide C2-14Af and purified C. tropicalis peroxisomes (C2-14Af was not used since its calculated M_r (22465, Table 5.1.1) approximates that of CAT (M_r≈24000; Shaw, 1975)). Import assays were subsequently conducted as depicted in Figure 5.3.6. Unfortunately, proteinase K treatment of each import assay resulted in the complete proteolysis of the ^{35}S-labeled C2-11b, C2-16b or C2-16n polypeptides (Figure 5.3.7, lanes 4 in all panels). This was evident even under conditions in which the cytosolic control polypeptide, CAT, was not completely digested (Figure 5.3.7, Nycodenz (step)/C2-16n panel, lane 4). Thus, irrespective of the use of sucrose step-, sucrose linear or Nycodenz step-gradient purified peroxisomes isolated from Brij-grown C. tropicalis, the modified in vitro import assay remained nonfunctional. As a result, studies concerned with the in vitro import of Cat into purified C. tropicalis peroxisomes were abandoned for an alternative system with which to determine the region(s) of C. tropicalis Cat which is essential for its targeting to yeast peroxisomes. For this purpose, parallel efforts were made to develop a heterologous in vivo expression and targeting system for Cat to the peroxisomes


SUCROSE (STEP)
SUCROSE (LINEAR)
C2–11b
C2–16b
C2–16n
Mr 1 2 3 4 5
97.4
68
45
29
18.4
Pk – + – + +
D – – – – +

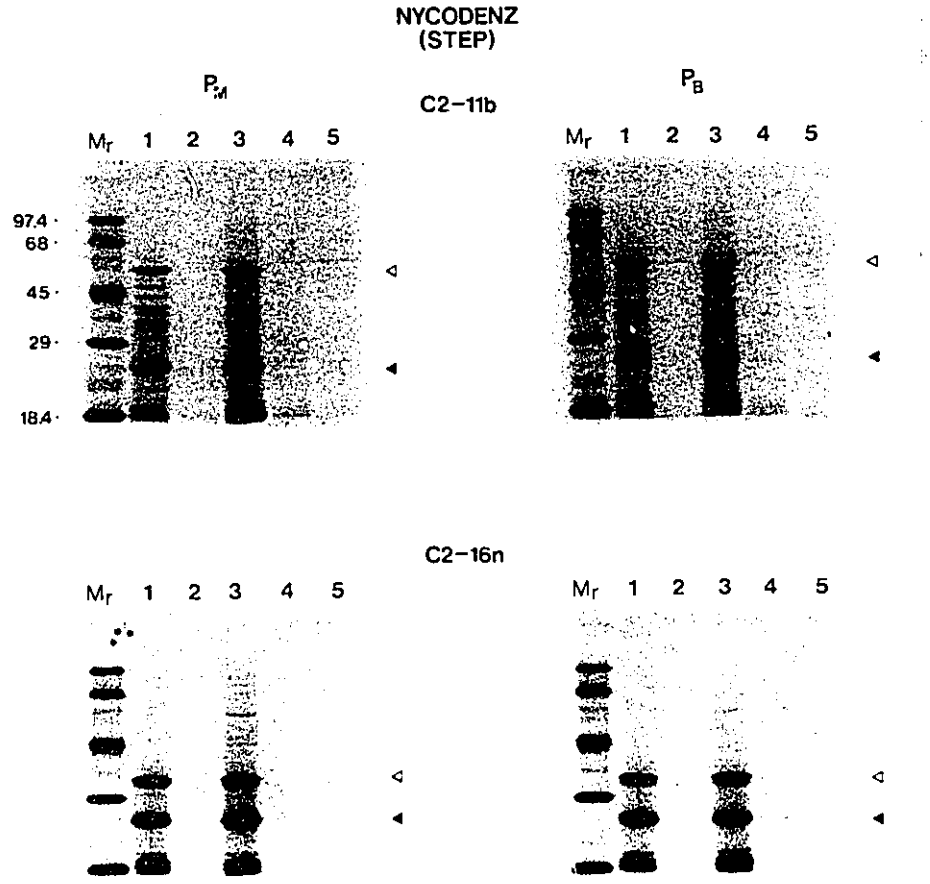


Figure 5.3.7. In vitro studies concerned with the targeting of Cat polypeptides to sucrose step-, sucrose linear- or Nycodenz step-gradient purified peroxisomes isolated from C. tropicalis. Post-translational import assays were conducted as described in section 2.6.2 (illustrated in Figure 5.3.6), using wheat germ extract-translated C2-11b, C2-16b or C2-16n polypeptides and one of four different peroxisomal fractions: either sucrose step-gradient purified peroxisomes, sucrose linear-gradient purified peroxisomes or Nycodenz step-gradient purified peroxisomes [middle (P_M) or bottom (P_B) peroxisomal fractions; see text, section 5.3]. The control (cytosolic) polypeptide for these import assays, CAT, was synthesized using a wheat germ translation extract primed with TP40C transcripts. Lanes 1 through 5 (for each panel) are defined by treatment as indicated for sucrose step-gradient purified peroxisomes, Panel C2-11b. The position of each "full-length" Cat translation product is indicated by an open triangle; the position of full-length CAT is denoted by a closed triangle. The molecular mass protein standards (in kDa) seen in lanes M_r include: phosphorylase b (97.4), bovine serum albumin (68), chicken egg albumin (45), carbonic anhydrase (29) and β-lactoglobulin (18.4). Symbols represent: P_K, proteinase K (5 μg/40 μl assay); D, detergent [0.8% (w/v) DOC, 0.8% (w/v) Triton X-100]; +, present; -, absent. Exposures were for 36 h at -70°C with Kodak XAR-5 film.

of C. albicans and S. cerevisiae. The preliminary results of this work are presented in section 5.4.

Future research that could be conducted in the development of a competent system for import into C. tropicalis peroxisomes in vitro is presented in the section entitled, "Conclusions and Future Directions".

5.4 IN VIVO STUDIES CONCERNED WITH THE TARGETING OF Cat FROM C. TROPICALIS TO THE PEROXISOMES OF C. ALBICANS AND S. CEREVISIAE

As indicated in section 5.3, attempts to develop a functional in vitro system for the import of Cat into purified C. tropicalis peroxisomes were unsuccessful. For this reason, as well as to provide a complementary system with which to compare results obtained in vitro (once a competent system is established), efforts were made to utilize a heterologous yeast system to study the in vivo expression and peroxisomal targeting of Cat from C. tropicalis.

Traditionally, C. tropicalis has served as a useful organism for in vitro studies related to peroxisome biogenesis (Lazarow and Fujiki, 1985; see Introduction), however, the lack of a transformation system for this yeast has limited studies in vivo [noteworthy are the preliminary efforts concerned with the development of a transformation system for C. tropicalis that have been reported recently (Corner and Polter, 1989)]. To circumvent this problem,

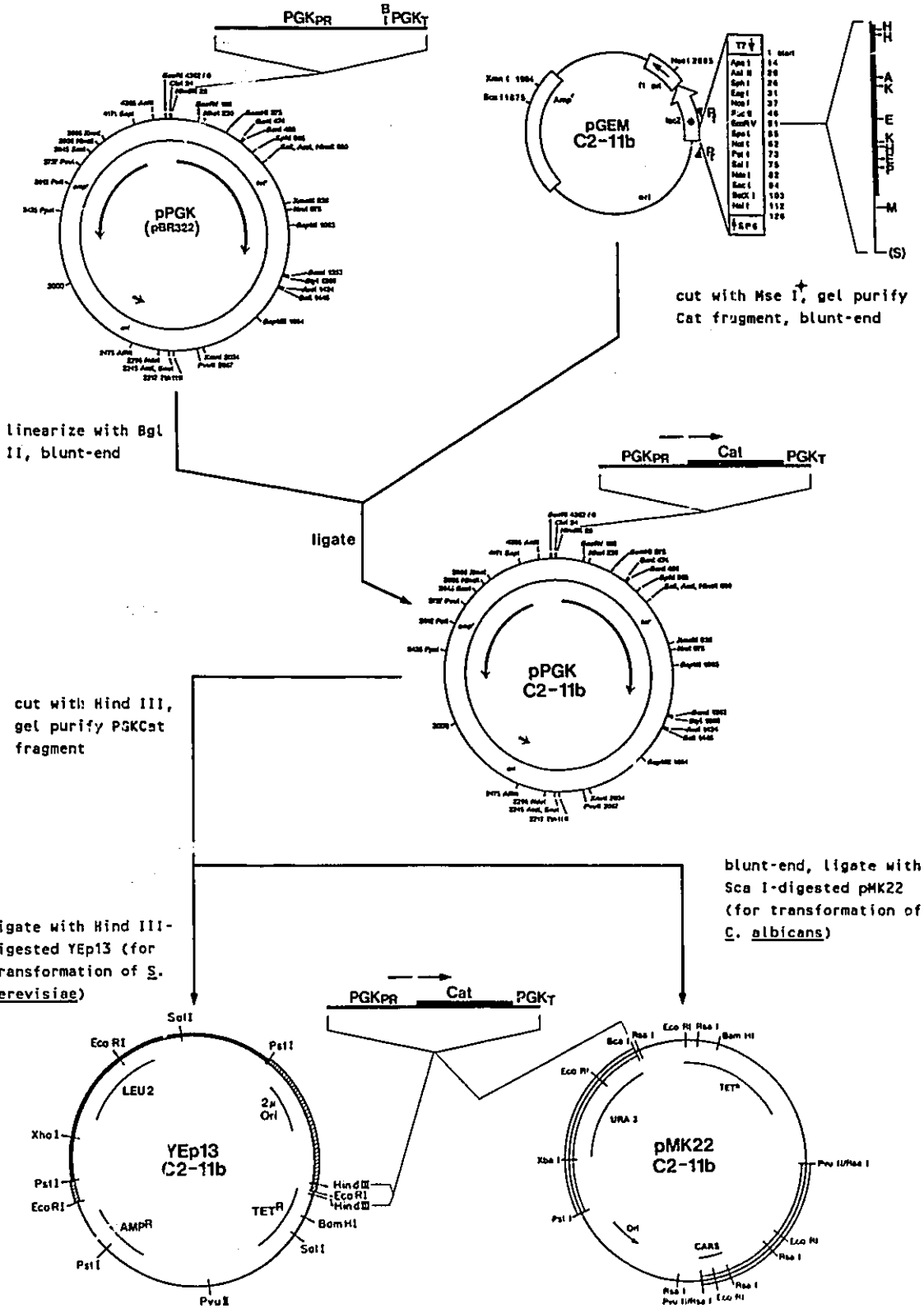
studies concerned with the identification of the targeting signal(s) of C. tropicalis (and other yeast) peroxisomal proteins in vivo have utilized heterologous expression and import systems established for the yeasts Candida albicans (Aitchison and Rachubinski, 1990; Aitchison et al., 1991), Candida maltosa (Kamiryo et al., 1989), Saccharomyces cerevisiae (Distel et al., 1987; Hartig et al., 1990; Aitchison et al., 1991; Binder et al., 1991) and Hansenula polymorpha (Hansen and Roggenkamp, 1989). Since, in our laboratory, the in vivo expression of C. tropicalis HDE, and deletion mutants of HDE, in transformed C. albicans SGY243 and S. cerevisiae DL1 led to the identification of the C-terminal tripeptide, ala-lys-ile (AKI), as the PTS for this protein (Aitchison and Rachubinski, 1990; Aitchison et al., 1991), both C. albicans and S. cerevisiae were considered to be potentially useful systems for the related task of identifying the PTS of C. tropicalis Cat.

5.4.1 Expression of the Cat gene from C. tropicalis in C. albicans and S. cerevisiae

In an effort to ensure the in vivo expression of C. tropicalis Cat in C. albicans, the Cat gene was subcloned into pMK22 (the multicopy yeast-E. coli shuttle plasmid to be used for the transformation of C. albicans; Kurtz et al., 1987) under the transcriptional control of two different promoters, the endogenous promoter of the Cat gene or that

of the phosphoglycerate kinase gene of S. cerevisiae (Gould et al., 1990a; section 2.6.1.5). This dictated that the in vivo expression of Cat in C. albicans (if observed) would either be concomitant with oleic acid-induced peroxisome proliferation (due to transcription from the endogenous Cat gene promoter) or alternatively, induced by the growth of transformants in glucose-containing medium (due to transcription from the PGK promoter); import into peroxisomes would then occur upon the transfer of glucose-grown transformants to oleic acid induction medium (section 2.6.3; Distel et al., 1987; Gödecke et al., 1989). To this end, the ≈3 kbp BamH I, Sal I fragment encompassing the entire Cat gene and its 5'- and 3'-flanking regions (section 4.4.1) was blunt-ended with Klenow fragment and subcloned into Sca I-digested pMK22 to allow expression in vivo from the Cat gene promoter. In addition, as depicted in Figure 5.4.1 and described in detail in section 2.6.1.5, Mse I fragments isolated from the in vitro expression constructs pGEMC2-11b, pGEMC2-16b and pGEMC2-11bE_D and an Mse I, Ban II fragment isolated from pGEMC2-11bESL (section 5.1) were subcloned into plasmid pPGK (Gould et al., 1990a), in the correct orientation 3' to the PGK promoter, and the PGKCat gene segments excised, blunt-ended and subcloned into Sca I-digested pMK22 to allow expression in vivo from the PGK promoter. All recombinant pMKCat or pMKPGKCat (or truncated

Figure 5.4.1. Construction of the recombinant plasmids used for the *in vivo* expression of Cat deletion mutants. Although the diagram illustrates only those recombinants which were constructed for the *in vivo* expression of full-length Cat, it is representative of the construction scheme followed for the expression of each Cat deletion mutant (and for *E. coli* CAT and firefly luciferase). The initial step toward the *in vivo* expression of Cat polypeptides involved the subcloning of those truncated Cat sequences generated for *in vitro* expression into the Bgl II (B) site located between the 5′- and 3′-noncoding sequences of the PGK gene [i.e. between the PGK promoter (PGK_{PR}) and terminator (PGK_T), respectively; top left]. For each of these constructs, the proper orientation of PGK-Cat sequences (dashed arrows) was confirmed by restriction endonuclease mapping. PGK-Cat sequences were then excised and ligated into two different yeast:*E. coli* shuttle vectors: YEp13 (Broach et al., 1979) for expression in *S. cerevisiae* strain DL1 and pMK22 (Kurtz et al., 1987) for expression in *C. albicans* strain SGY243 (the orientation of the PGK-Cat sequences within these vectors was not determined). Yeast transformations were subsequently conducted as described in section 2.6.1.6. The partial restriction endonuclease cleavage map of the Cat gene (top right; thick line indicates the ORF) includes: A, Aha II, E, EcoR I; H, Hinc II; K, Kpn I; M, Mse I (*, indicates the Mse I site of the vector used to excise the Cat fragments; +, note that Cat fragments were liberated from pGEMC2-16bAH, pGEMC2-11bESS and pGEMC2-11bESL constructs by Mse I, Ban II-digestion); P, Pvu II; (S), blunt-ended Sal I site. Plasmid maps of pBR322, pGEM-5Zf(+), YEp13 and pMK22 were reproduced from Sambrook et al. (1989), the Promega 1989/90 catalogue, Broach et al. (1979) and Kurtz et al. (1987), respectively.


PGKPR
PGKT
pPGK
(pBR322)
pGEM
C2-11b
cut with Mse I, gel purify
Cat fragment, blunt-end
linearize with Bgl
II, blunt-end
PGKPR
Cat
PGKT
ligate
pPGK
C2-11b
cut with Hind III,
gel purify PGKCat
fragment
blunt-end, ligate with
Sca I-digested pMK22
(for transformation of
C. albicans)
ligate with Hind III-
digested YEp13 (for
transformation of S.
cerevisiae)
YEp13
C2-11b
LEU2
2μ
Ori
TETR
AMPR
pMK22
C2-11b
URA 3
TETR
CARS
Ori

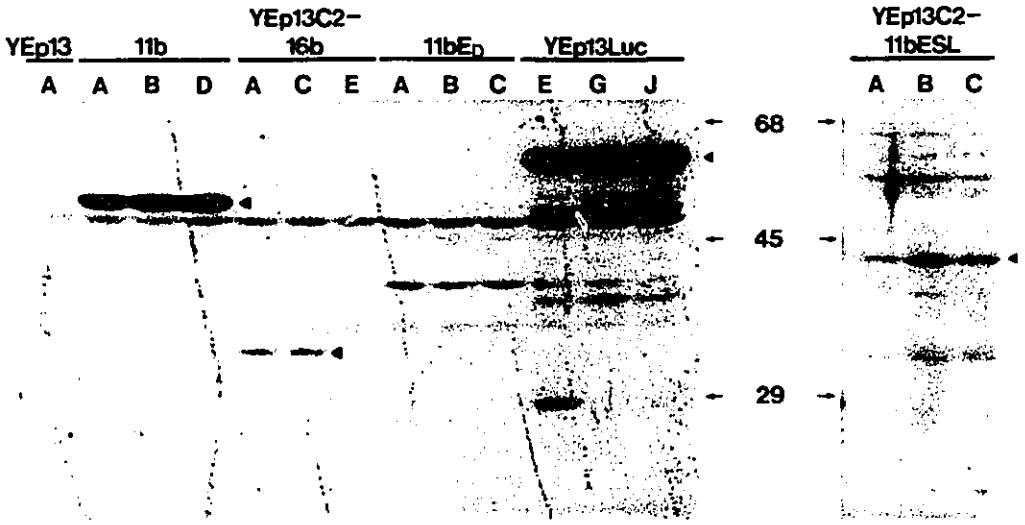
Cat) gene plasmids were subsequently used to transform C. albicans SGY243 to uracil prototrophy (section 2.6.1.6). Cell lysates prepared from C. albicans SGY243, C. tropicalis pK233 and several transformants (grown in media appropriate for the induction of Cat gene expression; section 2.2.1), including pMKCat (harbouring the endogenous Cat promoter) and pMKPGKCat, pMKPGKC2-16b, pMKPGKC2-11bE$_D$ and pMKPGKC2-11bESL (harbouring the PGK promoter), were fractionated by SDS-PAGE and analyzed by immunodetection with rabbit anti-C. tropicalis Cat serum. Since no cross-reacting polypeptides of the appropriate size were detected in any of the cell lysates (not shown), it was hypothesized that the PGK promoter was not functional in C. albicans SGY243. The only cross-reacting polypeptide observed was postulated to be that of the endogenous C. albicans catalase, which exhibited a slightly larger (≈1 to 2 kDa) molecular mass than that of Cat from C. tropicalis (not shown). Unfortunately, the presence of a cross-reacting polypeptide in C. albicans, which was similar in size to that of Cat from C. tropicalis, made it difficult to determine whether the endogenous Cat promoter was functional in C. albicans. For these reasons, subsequent efforts were directed toward determining the potential of S. cerevisiae to serve as a heterologous in vivo expression and import system for Cat from C. tropicalis.

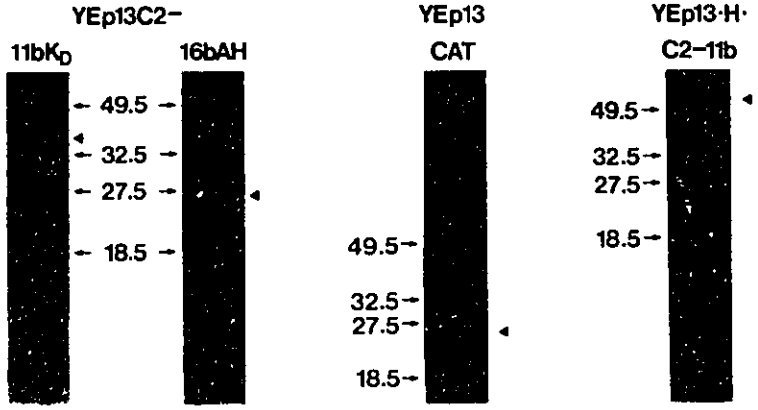
S. cerevisiae is an attractive system for in vivo studies related to Cat expression and import for several reasons (in addition to those mentioned above). First, PGK promoter-directed expression of the cDNA encoding Photinus pyralis (firefly) luciferase in S. cerevisiae results in the import of luciferase into S. cerevisiae peroxisomes (Gould et al., 1990a); this demonstrates the functionality of the system in vivo and provides an excellent control for similar studies involving Cat from C. tropicalis. Second, although S. cerevisiae harbours both a peroxisomal and a cytosolic catalase [catalases A (Cohen et al., 1988) and T (Hartig and Ruis, 1986), respectively], which could potentially cross-react with the rabbit anti-C. tropicalis Cat serum, their relative molecular masses (58490 and 64449, respectively) are considerably larger than that of Cat from C. tropicalis (54944; section 4.4.2) and should be discernable upon SDS-PAGE. Further, since both full-length and truncated pPGKCat constructs had been generated previously in an attempt to achieve the PGK promoter-directed in vivo expression of Cat in C. albicans, several of the YEp13Cat constructs used to transform S. cerevisiae DL1 to leucine prototrophy (section 2.6.1.6), including YEp13C2-11b, YEp13C2-16b, YEp13C2-11bE_D and YEp13C2-11bESL, were generated simply by ligating the previously purified PGKCat sequences into Hind III-digested YEp13 (Figure 5.4.1; section 2.6.1.5). Additional YEp13-

based constructs used to transform S. cerevisiae, and their origins, included: (i) YEp13C2-11bK$_D$, YEp13C2-16bAH and YEp13CAT - which were generated from the corresponding in vitro expression constructs as depicted in Figure 5.4.1, (ii) YEp13·H·C2-11b and YEp13·H·C2-11bK$_D$ - which were constructed such that the expression of Cat sequences was directed by the promoter for the gene encoding C. tropicalis HDE (this allowed the oleic acid-induced expression of Cat to be simultaneous with peroxisome proliferation) and (iii) YEp13PGK-luciferase (YEp13Luc; Gould et al., 1990a) - which was obtained from Dr. S.J. Gould (section 2.1.6).

Subsequent to the growth of transformants in YNBD medium to induce expression from the PGK promoter (except for the transformant YEp13·H·C2-11b, which was grown in S. cerevisiae induction medium to induce Cat expression from the HDE promoter), cells were lysed and the protein extracts analyzed by SDS-PAGE and Western blotting with antiserum raised against C. tropicalis Cat. As seen in Figure 5.4.2 (lane YEp13), the major cross-reacting polypeptide detected in the control lysate (of M_r ≈50000) was significantly smaller in size than either S. cerevisiae catalase A or T (see above); its identity is, therefore, unknown (detection of this polypeptide may have resulted from cross-reaction with the anti-firefly luciferase serum that was present along with the anti-Cat serum during immunoblotting).

Figure 5.4.2. Heterologous expression of full-length and truncated Cat polypeptides in S. cerevisiae. In vivo, PGK promoter-induced expression of full-length and truncated Cat sequences (and sequences encoding CAT and firefly luciferase) was investigated by western blot analysis of crude (total) cell lysates prepared (as per section 2.3.3) from YNBD (supplemented with L-histidine and uracil)-grown S. cerevisiae transformants (note that transformant YEp13·H·C2-11b was grown in S. cerevisiae induction medium to induce Cat expression from the HDE promoter; see text, section 5.4.1). The control cell lysate (lane YEp13A) was prepared from YEp13-transformed S. cerevisiae. Approximately 200 µg of protein from each lysate were for used SDS-PAGE. Post-transfer, immunodetection of Cat, CAT and luciferase polypeptides was conducted using appropriate antisera as described in section 2.3.3. The position of each "full-length" polypeptide is indicated by a closed triangle. The molecular mass protein standards (indicated by arrows) include: Upper panel (in kDa): bovine serum albumin (68), chicken egg albumin (45) and carbonic anhydrase (29); Lower panel (apparent molecular weights in kDa): ovalbumin (49.5), carbonic anhydrase (32.5), soybean trypsin inhibitor (27.5) and lysozyme (18.5). Exposures were as follows: Upper left panel: 36 h at -70°C with Kodak XAR-5 film and 1 intensifying screen; Upper right panel: 84 h at -70°C with Kodak XAR-5 film and 1 intensifying screen; Lower panels: 8 min with Kodak XAR-5 film [using the enhanced chemiluminescence (ECL) western blotting detection system; Amersham Canada Limited].


YEp13
11b
YEp13C2-
16b
11bE$_D$
YEp13Luc
YEp13C2-
11bESL
A A B D A C E A B C E G J
A B C
68
45
29


YEp13C2-
11bK$_D$
16bAH
49.5
32.5
27.5
18.5
YEp13
CAT
49.5
32.5
27.5
18.5
YEp13·H·
C2-11b
49.5
32.5
27.5
18.5

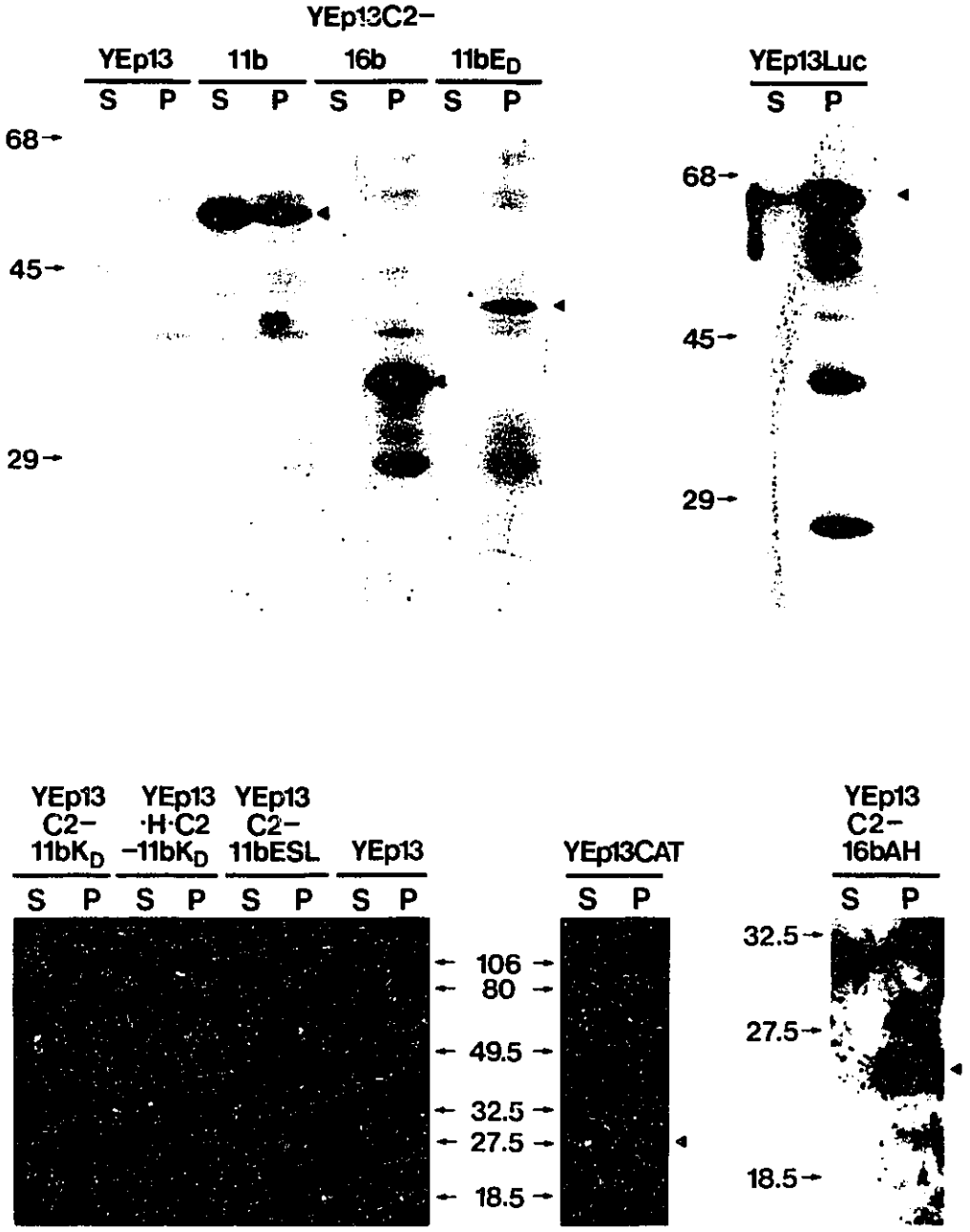
Independent of this, however, the results revealed the presence of an extra polypeptide (relative to the control extract) in at least one lysate of several made from different transformants obtained for each of the YEp13 constructs indicated [full-length Cat and luciferase polypeptides were clearly visible components of transformed cell lysates subsequent to SDS-PAGE and staining with Coomassie blue (not shown), indicating that they comprised a significant percentage of the total cellular protein]. Further, the approximate M_r of each additional polypeptide observed corresponded to that deduced (see Table 5.1.1) for the full-length or truncated polypeptide (either Cat, CAT or luciferase) that was to be expressed. This suggested that S. cerevisiae DL1 would serve as a useful heterologous system with which to study the PGK promoter-directed in vivo expression and import of Cat from C. tropicalis.

5.4.2 Subcellular Location of Full-Length and Truncated Cat Expressed in S. cerevisiae

Having achieved the expression of full-length and truncated Cat polypeptides in S. cerevisiae, efforts were made to determine their subcellular location(s). As a "quick" screening method, the pelletability of each of the Cat polypeptides was determined by isolating a particulate fraction (containing mitochondria and peroxisomes) from homogenates of each S. cerevisiae transformant. For this

purpose, transformants were grown initially in YNBD medium to induce Cat (or CAT or luciferase) expression from the PGK promoter and then transferred to S. cerevisiae induction medium to allow peroxisome proliferation [YEp13·H·C2-11bK_D was grown only in induction (oleic acid-containing) medium since this co-induced the HDE promoter and peroxisome proliferation]. The preliminary steps of peroxisome purification (section 2.2.2.1) were then carried out to obtain 20k x g supernatant (ie. cytosolic) and 20k x g pellet (ie. particulate) fractions. Subsequent to SDS-PAGE, the protein content of these fractions was analyzed by immunoblotting with anti-C. tropicalis Cat (or other appropriate) serum. Consistent with that reported by Gould et al. (1990a), the majority of the luciferase polypeptides resulting from PGK promoter-directed expression in S. cerevisiae pelleted with the particulate fraction (Figure 5.4.3, panel YEp13Luc, lane P), presumably due to import into peroxisomes. Although full-length Cat from C. tropicalis was also pelletable (Figure 5.4.3, panel YEp13C2-11b, lane P), the ratio of the amount of pelletable to cytosolic full-length Cat polypeptide (compare lane P to S) was less than that observed for either luciferase or the truncated Cat polypeptides analyzed (compare lanes P to S for each of these panels). This may be the result of a more compactly folded full-length Cat polypeptide being less

Figure 5.4.3. Analysis of the pelletability of full-length and truncated Cat polypeptides expressed in S. cerevisiae. S. cerevisiae transformants (excluding YEp13·H·C2-11bK$_D$), grown initially in YNBD medium (supplemented with L-histidine and uracil) to induce Cat, CAT or luciferase expression from the PGK promoter, were transferred to and grown in S. cerevisiae induction medium to induce peroxisome proliferation (as per sections 2.2.1 and 2.6.3) (transformant YEp13·H·C2-11bK$_D$ was grown in S. cerevisiae induction medium only to allow co-induction of C2-11bK$_D$ expression from the HDE promoter and peroxisome proliferation; see text, section 5.4.1). Preliminary steps toward peroxisome purification (see section 2.2.2.1) were then carried out to obtain 20k x g supernatant (S) and pellet (P) fractions from each transformant. SDS-PAGE of these fractions was conducted as follows: Upper panels: loaded 0.2% of 20k x g S and 1% (≈150 μg) of 20k x g P; Lower left and middle panels: loaded 0.1% of 20k x g S and 1% of 20k x g P; Lower right panel: loaded 10% of 20k x g S and 10% of 20k x g P. Post-transfer, immunodetection of Cat, CAT and luciferase polypeptides was conducted using appropriate antisera as described in section 2.3.3. The position of each "full-length" polypeptide is indicated by a closed triangle. The molecular mass protein standards (indicated by arrows) include: Upper panels (in kDa): bovine serum albumin (68), chicken egg albumin (45) and carbonic anhydrase (29); Lower panels (apparent molecular weights in kDa): phosphorylase b (106), bovine serum albumin (80), ovalbumin (49.5), carbonic anhydrase (32.5), soybean trypsin inhibitor (27.5) and lysozyme (18.5). Exposures were as follows: Upper panels: 64 h at -70°C with Kodak XAR-5 film and 1 intensifying screen; Lower left and middle panels: 5 sec with Kodak XAR-5 film (using ECL); Lower right panel: 5 min with XAR-5 film (using ECL).


YEp13C2−
YEp13
11b
16b
11bE$_D$
S P S P S P S P
68
45
29
YEp13Luc
S P
68
45
29
YEp13 C2− 11bK$_D$
YEp13 ·H·C2 −11bK$_D$
YEp13 C2− 11bESL
YEp13
S P S P S P S P
106
80
49.5
32.5
27.5
18.5
YEp13CAT
S P
YEp13 C2− 16bAH
S P
32.5
27.5
18.5

translocation competent than truncated Cat polypeptides (which may be loosely folded) or, alternatively, results from an increased rate of degradation of truncated (possibly, malfolded or aberrant) polypeptides in the S. cerevisiae cytosol. Each truncated Cat polypeptide was found to be pelletable (Figure 5.4.3, all truncated Cat construct panels, lanes P), independent of N-terminal (YEp13C2-16b), C-terminal (YEp13C2-11bESL), N-/C-terminal (YEp13C2-16bAH) or internal (YEp13C2-11bE_D and K_D) deletions which collectively removed 95% of the full-length Cat polypeptide (only amino acids 159 to 181 were not deleted, Figure 5.1.2). Unexpectedly, CAT, the acting cytosolic control, was also pelletable (Figure 5.4.3, panel YEp13CAT, lane P), although significant amounts (≈40%) of the polypeptide remained cytosolic (compare lanes S and P). In light of these results, efforts were made to fractionate the 20k x g pellets isolated from transformants into their constituent organelles to determine the intraorganellar location of each of the aforementioned polypeptides.

To establish the conditions appropriate to this end, 20k x g pellets were isolated from oleic acid-grown S. cerevisiae DL1 and the peroxisomes/mitochondria purified by Nycodenz density step-gradient centrifugation (section 2.2.2.3). As depicted in Figure 5.4.4, the distribution pattern of marker enzymes (catalase for peroxisomes,

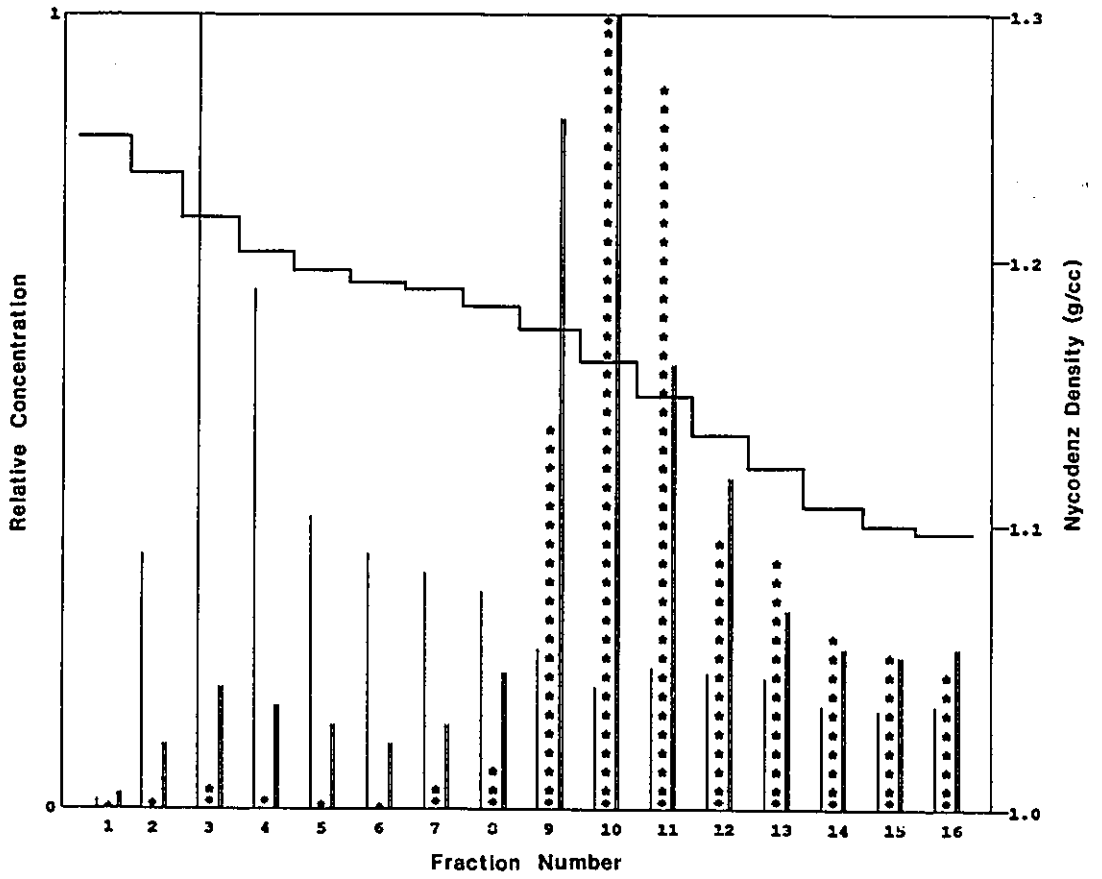


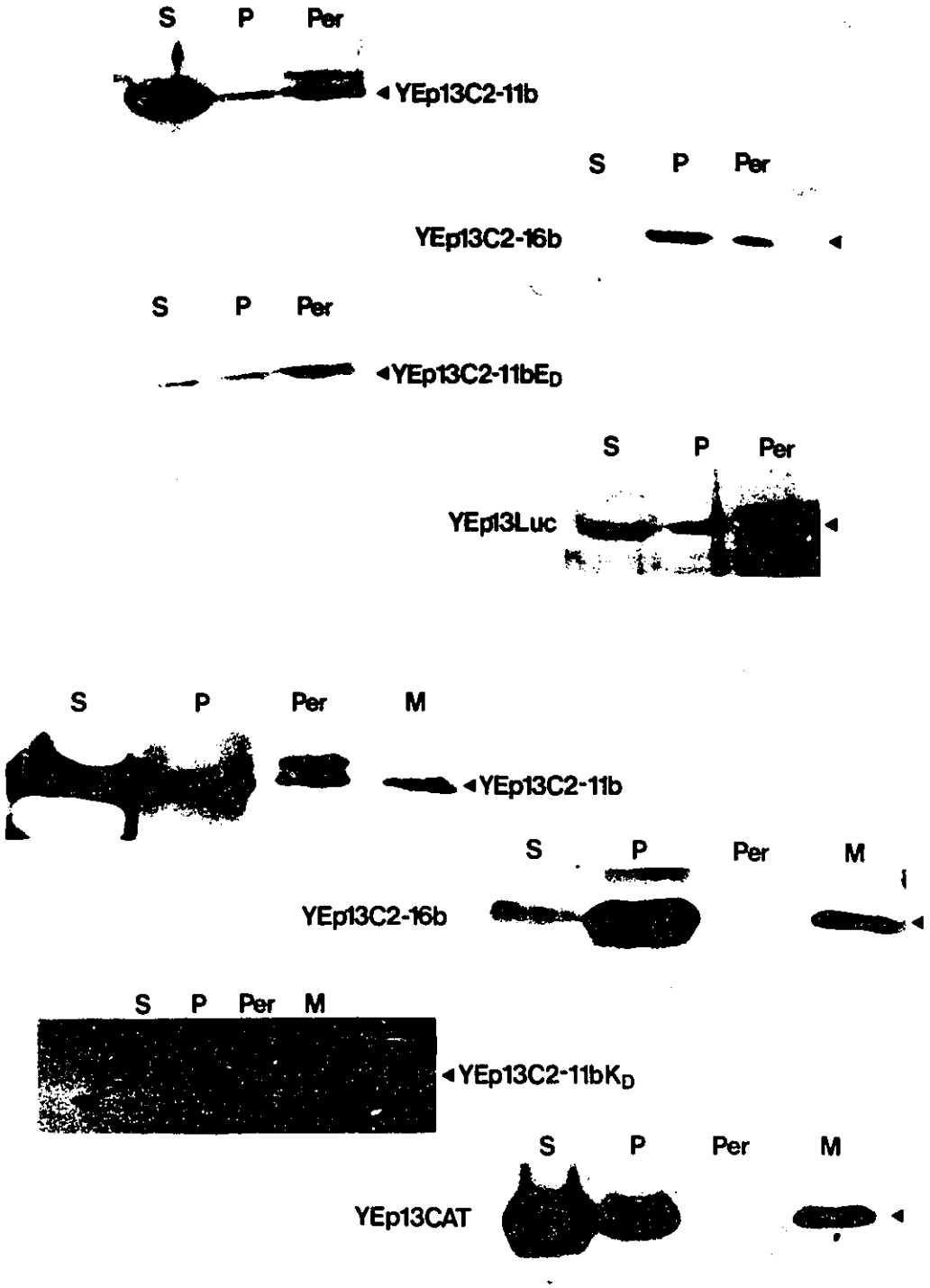
Figure 5.4.4. Distribution patterns of enzymes subsequent to subcellular fractionation of oleic acid-grown S. cerevisiae. A crude peroxisomal fraction was isolated from oleic acid-grown S. cerevisiae (as described in section 2.2.2.1) and purified by Nycodenz density step-gradient centrifugation (as described in section 2.2.2.3). Fractions (≈ 0.13 ml), represented along the abscissa scale in the order of their position in the gradient [ie. collected from bottom (left) to meniscus (right)], were analyzed for their average density (represented by the right axis; decreases ascending the gradient), protein content (double vertical line), and two marker enzyme activities, catalase (peroxisomal, single vertical line) and cytochrome c oxidase (mitochondrial, vertical asterisks). The relative concentration (represented by the left axis) of proteins, catalase and cytochrome c oxidase enzymes in each fraction was calculated as described by Beaufay et al. (1964) and normalized to one.

cytochrome c oxidase for mitochondria) for representative fractions collected from the Nycodenz step-gradients indicates a good separation between the peak fractions containing peroxisomes (fraction 3) and mitochondria (fraction 10). The equilibrium densities of peroxisomes and mitochondria isolated from oleic acid-grown S. cerevisiae DL1 and purified on discontinuous Nycodenz gradients were determined to be 1.22 and 1.17 g/cc, respectively (Figure 5.4.4). These values compare well with those of 1.20 and 1.17 g/cc obtained for peroxisomes and mitochondria, respectively, isolated from S. cerevisiae D273-10B (grown under similar conditions) and purified by linear sucrose density gradient centrifugation (Thieringer et al., 1991). Mitochondrial contamination of the peak peroxisomal fraction was negligible (Figure 5.4.4); only 0.6% of the total cytochrome c oxidase activity present in the gradient was localized to fraction 3. Although peroxisomal contamination of the peak mitochondrial fraction was higher by comparison (3.0%; Figure 5.4.4), a portion of the catalase activity assayed was likely the result of nonmembrane-bound enzyme released upon peroxisome breakage. Related to this, the catalase latency of S. cerevisiae DL1 peroxisomes purified on Nycodenz step-gradients ranged from 50 to 55%, a value which is slightly lower than that of 61 to 72% obtained by Thieringer et al. (1991) for peroxisomes isolated from S.

cerevisiae D273-10B and purified on linear Nycodenz-0.25 M sucrose gradients.

Having established conditions which were suitable for the purification of S. cerevisiae peroxisomes (and mitochondria), the subcellular distribution of the Cat polypeptides expressed in S. cerevisiae transformants was determined by SDS-PAGE of cytosolic and purified organellar fractions, followed by immunoblot analysis. As seen in Figure 5.4.5 (YEp13Luc panel), and consistent with that observed previously (Figure 5.4.3; Gould et al., 1990a), luciferase polypeptides expressed in S. cerevisiae were pelletable (lane P) and found to be localized to peroxisomes (lane Per); luciferase detected in the 20k x g supernatant fraction (lane S) may result from the saturation of the peroxisomal import machinery of S. cerevisiae due to the PGK promoter-induced overexpression of luciferase from the multicopy construct, YEp13Luc. CAT was consistently found to be distributed between the 20k x g supernatant and pellet fractions (Figure 5.4.5, YEp13CAT panel, compare lanes S and P). Interestingly, but unexplainedly (see below), the target organelle for that amount of CAT which was pelletable was the mitochondrion (compare lanes Per and M). As hoped for, full-length Cat from C. tropicalis was targeted to S. cerevisiae peroxisomes (Figure 5.4.5, YEp13C2-11b panels, lanes Per), although the efficiency of import appeared to be

Figure 5.4.5. Determination of the subcellular location of full-length and truncated Cat polypeptides expressed in S. cerevisiae. S. cerevisiae transformants, grown as described in the legend of Figure 5.5.2, were fractionated into 20k x g supernatant (S) and pellet (P) fractions (as per section 2.2.2.1). The 20k x g pellet fractions were further fractionated by centrifugation through Nycodenz step-gradients (as per section 2.2.2.3) to obtain purified peroxisomal (PER) and mitochondrial (M) fractions. SDS-PAGE of these fractions was conducted as follows: Upper panels: loaded equivalent amounts (1%) of 20k x g S (TCA-precipitated) and 20k x g P fractions (≈150 µg), loaded 40 µg of peak peroxisomal fractions; Lower panels: loaded equivalent amounts (5%) of 20k x g S (TCA-precipitated) and 20k x g P fractions, loaded equivalent volumes (10 µl) of peak PER and M fractions. Post-transfer, immunodetection of Cat, CAT and luciferase polypeptides was conducted using appropriate antisera as described in section 2.3.3. The position of each full-length polypeptide is indicated by a closed triangle. Exposures were as follows: Panels A: 64 h at -70°C with Kodak XAR-5 film and 1 intensifying screen; Panels B: 5 sec with Kodak XAR-5 film (using ECL).


S
P
Per
YEp13C2-11b
S
P
Per
YEp13C2-16b
S
P
Per
YEp13C2-11bE$_D$
S
P
Per
YEp13Luc
S
P
Per
M
YEp13C2-11b
S
P
Per
M
YEp13C2-16b
S
P
Per
M
YEp13C2-11bK$_D$
S
P
Per
M
YEp13CAT

low relative to that of luciferase (compare lanes S and Per for the YEp13C2-11b and YEp13Luc panels). This has also been observed for the *in vivo* expression of dihydroxyacetone synthase (DHAS) from *H. polymorpha* in *S. cerevisiae* in that greater than 50% of the DHAS protein was found in the soluble cytoplasmic fraction, with the remainder targeted to peroxisomes (Gödecke et al., 1989). Since the C2-11bE$_D$ polypeptide (Figure 5.1.2, Table 5.1.1) co-purified with peroxisomes from *S. cerevisiae* (Figure 5.4.5, YEp13C2-11bE$_D$ panel, lane Per), amino acids 271 to 390 of Cat appeared to be dispensible for import. Truncated Cat polypeptide C2-16b, which harbours an N-terminal deletion of 158 amino acids (Figure 5.1.2, Table 5.1.1), also appeared to be localized to peroxisomes (Figure 5.4.5, upper YEp13C2-16b panel, lane Per), suggesting that the N-terminal residues of Cat were without function in peroxisomal targeting. This deletion may have influenced the efficiency of peroxisomal targeting, however, since the amount of C2-16b polypeptide associated with *S. cerevisiae* peroxisomes was significantly less than that available for import and ultimately found to co-purify with mitochondria (Figure 5.4.5, lower YEp13C2-16b panel, compare lanes Per and M). By contrast, the truncated Cat polypeptide C2-11bK$_D$, which possesses an internal deletion of amino acids 182 to 344 (Figure 5.1.2, Table 5.1.1) [and was previously observed to be pelletable (Figure

5.4.3)], appeared not to be targeted to peroxisomes (Figure 5.4.5, YEP13C2-11bK$_D$ panel, lane Per), but was found to co-purify with S. cerevisiae mitochondria (lane M). Collectively, these preliminary results suggest that amino acids 182 to 269 of Cat from C. tropicalis may function in the targeting of Cat to the peroxisomes of S. cerevisiae. This is consistent with that reported for S. cerevisiae catalase A in that neither the N-terminal 126 amino acids, nor the C-terminal 15 amino acids, are able to direct DHFR into the peroxisomes of S. cerevisiae (Hartig et al., 1990). Hartig et al. (1990) suggest further that an internal sequence of 14 amino acids (residues 127 to 140) is responsible for the comigration of the DHFR-S. cerevisiae catalase A fusion product with S. cerevisiae peroxisomes upon density gradient centrifugation. It should be noted, however, that the corresponding amino acids of Cat from C. tropicalis (residues 110 to 123, Figure 4.4.5) are deleted from C2-16b polypeptides which appear to be translocated into the peroxisomes of S. cerevisiae, albeit inefficiently (see above). Caution is also required in the interpretation of the results obtained for the targeting of Cat to S. cerevisiae peroxisomes since Hartig et al. (1990) and Binder et al. (1991) have determined, by immunofluorescence and electron microscopy, that the expression of DHFR-S. cerevisiae catalase A fusion products containing 140 or more

amino acids of catalase A does not result in the import of catalase A into the peroxisomes of S. cerevisiae. Instead, this causes the formation of nonmembrane-bound protein aggregates (inclusion bodies) in S. cerevisiae which cannot be resolved from peroxisomes upon sucrose or Nycodenz density gradient centrifugation. Large, electron-dense, cytoplasmic inclusions have also been visualized in S. cerevisiae transformed with multicopy plasmids used for the expression of alcohol oxidase from H. polymorpha (Distel et al., 1987). Whether the overexpression of Cat from C. tropicalis in S. cerevisiae similarly results in inclusion body formation is not known. If so, this may explain the association of full-length and truncated Cat polypeptides (and CAT) with S. cerevisiae mitochondria (Figure 5.4.5, lower panels, lane M), assuming inclusion bodies of different sizes and densities are formed in vivo. This would, however, leave the absence of an association between C2-11bK_D Cat polypeptides and peroxisomes unexplained. If not, it may be that the apparent targeting of Cat to mitochondria in S. cerevisiae results from a nonspecific membrane association or is an artifact created by saturation of the peroxisomal import machinery. Further, and more importantly, this would suggest that at least part of the PTS for Cat from C. tropicalis resides at an internal location within its mature amino acid sequence. Additional

studies that could be conducted in an effort to resolve these issues are discussed in the "Conclusions and Future Directions" section.

CONCLUSIONS AND FUTURE DIRECTIONS

Prior to the work conducted herein, studies in both yeast and mammalian systems (reviewed by Lazarow and Fujiki, 1985) had yielded the following generalities concerning peroxisome biogenesis: 1) new peroxisomes arise from the fission of preexisting peroxisomes; 2) peroxisomal proteins are encoded exclusively by nuclear genes; 3) peroxisomal proteins are synthesized on free polyribosomes in the cell cytosol (with one possible exception; section 1.4.1) and posttranslationally imported into preexisting organelles; 4) distinct from the biogenesis of mitochondria or chloroplasts for which proteins are synthesized as larger precursors and posttranslationally targeted to the organelles via cleavable transit peptides, peroxisomal proteins are, with a few noted exceptions (section 1.4.3.3), synthesized at their mature size and do not undergo posttranslational modification. From this, it was realized that the information responsible for the specificity of protein targeting to peroxisomes must reside within the mature amino acid sequence of the imported protein. Knowledge of the targeting sequences of peroxisomal proteins was, however, very limited. Consequently, using the yeast C. tropicalis pK233 as a model system, this work endeavoured to clone and sequence at least one full-length cDNA or gene encoding a peroxisomal protein so that the

subsequent creation of deletion mutants of this cDNA/gene, together with the development of an *in vitro* and/or *in vivo* expression and translocation system for the corresponding polypeptides, would allow the amino acid sequence(s) comprising the peroxisomal addressing signal of this protein to be identified.

Efforts toward this end resulted in the cloning of two genes (and one composite full-length cDNA; see section 4) encoding different enzymes associated with the β-oxidation pathway of *C. tropicalis* peroxisomes. Sequencing of the gene encoding the first enzyme of the peroxisomal β-oxidation pathway, acyl-coenzyme A oxidase (AOx), revealed a single open reading frame, without intervening sequences, of 2,127 nucleotides which encodes a polypeptide of 709 amino acids ($M_r \approx 79155$). The deduced amino acid sequence of *C. tropicalis* AOx exhibits significant identity and homology with that of several other acyl-coenzyme A oxidases of *C. tropicalis*, as well as those of other yeasts (*C. maltosa* and *S. cerevisiae*) and rat liver. By comparison, sequencing of the *C. tropicalis* gene (or cDNA) encoding catalase (Cat), the enzyme responsible for the breakdown of hydrogen peroxide (produced as a bi-product of acyl-coenzyme A oxidase activity) to oxygen and water, revealed a single open reading frame of 1,455 nucleotides, without intervening sequences, which encodes a polypeptide of 485 amino acids

($M_r \approx 54944$). The deduced amino acid sequence of C. tropicalis Cat shows significant identity and homology with that of several other catalases, including those of the yeast S. cerevisiae, mammals (rat, human and bovine), plants (maize, sweet potato and cottonseed) and Drosophila. Codon usage in the genes encoding C. tropicalis AOx and Cat is nonrandom, exhibiting a strong bias toward a group of preferred codons which is similar to that of several highly expressed genes of S. cerevisiae and E. coli, as well as other C. tropicalis genes encoding peroxisomal proteins. The high degree of preferential codon usage observed in the AOx and Cat genes may be directly correlated with their level of expression and ultimately, the corresponding intracellular mRNA and protein levels in C. tropicalis.

Subsequent to the cloning, sequencing and characterization of the genes encoding C. tropicalis AOx and Cat, attempts were made to use the in vitro translocation system of Small et al. (1987, 1988), which was competent for the import of C. tropicalis PXP-4 into peroxisomes purified from C. tropicalis, to determine the nature of the PTS of Cat. For this purpose, N-terminal, C-terminal, N-/C-terminal and internal deletion mutants of Cat were constructed, expressed in cell-free (rabbit reticulocyte lysate and wheat germ extract) systems and assayed for translocation into purified C. tropicalis peroxisomes. Unfortunately, the import system

of Small et al. (1987, 1988), and several modifications thereof, proved to be inadequate for the translocation of Cat (or AOx or HDE) into purified C. tropicalis peroxisomes. Further work aimed at establishing a functional in vitro system for protein translocation into peroxisomes is, however, warranted. Similar to the understanding which has been gained concerning protein targeting to mitochondria, chloroplasts and nuclei (see section 1.4.3.1), a competent in vitro import system for peroxisomes would contribute greatly to an understanding of the structure and function of PTSs, the components of the translocation apparatus which are responsible for PTS recognition and import specificity [eg. cytosolic chaperonins and/or import receptor(s)] and the internalization process as a whole (eg. the energetics involved). Efforts directed toward the development of a homologous yeast in vitro expression and translocation system might prove fruitful in this regard since this system would mimic the in vivo import process in several important respects. First, import would be allowed to occur in the presence of a cytoplasmic fraction derived from the same cell type as the peroxisomes. Thus, any cytoplasmic factor(s) required for import would be present to interact with the targeted protein and/or peroxisomal import machinery. Further, since import into peroxisomes would occur from a homologous cytoplasmic fraction which also

permits protein synthesis, prerequisite interactions which may be required between cytosolic factors and targeted proteins would be allowed to occur during protein synthesis (ie. prior to complete folding of the polypeptide), thereby assisting translocation [Related to this, catalase import into the peroxisomes of Zellweger cells is inhibited by prior binding to catalase of 3-amino-1,2,4-triazole, which appears to retard unfolding of the protein (Middelkoop et al., 1991).]. Such a homologous assay would also minimize the variability within the system should the required interactions be specific to an extent that they do not occur favourably in heterologous systems (eg. when wheat germ extract translation mixtures are combined with yeast peroxisomes as was done herein). Due to the proteinaceous (and thus, protease-sensitive) nature of the translation and translocation machinery, the development of an efficient *in vitro* import system may also involve purified *S. cerevisiae* DL1 peroxisomes (section 5.4), together with the use of a vacuolar protease-deficient strain of *S. cerevisiae* (eg. ABYS-1; Verner and Weber, 1989) for the preparation of the cytosolic fraction, as a means of maintaining the integrity of the system. Collectively, these adjustments may influence the efficiency of protein import into peroxisomes in a positive manner, dictate that the *in vitro* import system more accurately resembles that process which occurs *in vivo*

and allow the eventual characterization of several components of the peroxisomal import apparatus. Specific to the work described herein, the development of a competent *in vitro* import system could be used to confirm the preliminary results obtained for the *in vivo* targeting of Cat from *C. tropicalis* to the peroxisomes of *S. cerevisiae* [and those obtained for *C. tropicalis* PXP-4 by Small et al. (1987, 1988) which are similarly suggestive of internal PTS(s) for *C. tropicalis* proteins]. This would establish the existence of at least two different types of targeting signals for peroxisomal proteins from *C. tropicalis*, located either at an internal location(s) within these proteins or at their C-terminus (eg. the C-terminal tripeptide PTS of HDE; Aitchison et al., 1991). Combined with the identification of a novel, cleavable PTS at the N-terminus of rat 3-ketoacyl-CoA thiolase (Swinkels et al., 1991), this would be further suggestive of the existence of multiple pathways or mechanisms of peroxisomal protein import, similar to that established for protein translocation into mitochondria and chloroplasts (section 1.4.3.1).

Since a functional system for the *in vitro* import of proteins into purified *C. tropicalis* peroxisomes could not be established, related studies concerned with the *in vivo* expression and targeting of Cat from *C. tropicalis* to the peroxisomes of *C. albicans* and *S. cerevisiae* were

initiated. For this purpose, the Cat gene, and deletion mutants thereof, were subcloned downstream of the promoter for the PGK gene of S. cerevisiae and the recombinant PGK promoter/Cat sequences further subcloned into vectors pMK22 and YEp13 for expression in C. albicans and S. cerevisiae, respectively. Unfortunately, the in vivo expression of Cat sequences could not be attained in C. albicans SGY243; it was hypothesized, therefore, that the PGK promoter was not functional in this yeast. By comparison, the PGK promoter-directed expression of full-length and truncated Cat polypeptides was achieved in S. cerevisiae DL1. Subsequent determination of the subcellular location of the full-length Cat polypeptide suggested that Cat from C. tropicalis was targeted to the peroxisomes of S. cerevisiae, reaffirming that heterologous signals for import into peroxisomes are recognized in S. cerevisiae. Interestingly, the deletion of amino acids 182 to 344 of Cat (which is specific to Cat mutant C2-11bK$_D$; section 5.4.2) abolished the association of Cat with S. cerevisiae peroxisomes previously observed upon Nycodenz density gradient centrifugation. These preliminary studies suggest that at least part of the PTS for Cat resides at an internal location within its mature amino acid sequence. Parallel studies in S. cerevisiae have, however, documented that the overexpression of certain proteins [catalase A of S. cerevisiae (Hartig et al., 1990; Binder et

al., 1991) and alcohol oxidase of H. polymorpha (Distel et al., 1987)] leads to the formation of nonmembrane-bound inclusion bodies which cannot be resolved from peroxisomes upon density gradient centrifugation (Hartig et al., 1990; Binder et al., 1991). Whether the overexpression of Cat from C. tropicalis in S. cerevisiae DL1 similarly results in the formation of inclusion bodies is not known. This must, however, be ascertained prior to concluding that the lack of an association between Cat polypeptides C2-11bK$_D$ and S. cerevisiae peroxisomes results from the deletion of amino acids 182 to 344 of Cat, which eliminates the PTS. Protease protection has traditionally been used to verify protein translocation into organelles (section 5.3), but inclusion bodies are known to be protease-resistant (Binder et al., 1991), negating this as a method of choice. Since cross-reactivity between the rabbit anti-C. tropicalis Cat serum and catalase A from S. cerevisiae is not readily apparent upon immunoblot analysis (section 5.4.2; results not shown), it may be possible to use immunofluorescence and/or electron microscopy to determine whether inclusion bodies are formed in S. cerevisiae transformants which overexpress Cat (should antibody cross-reactivity pose a problem, immunoreactive inclusion bodies and peroxisomes can be discriminated by size; Binder et al., 1991). The presence of inclusion bodies in S. cerevisiae transformants would preclude any

conclusions concerning the PTS of Cat. This problem might, however, be obviated using a single copy (CEN) plasmid (rather than the multicopy plasmid, YEp13, used herein; section 5.4.1) for Cat expression. Ideally, this would decrease the intracellular concentration of Cat polypeptides to a level less than the threshold required for inclusion body formation and allow import into S. cerevisiae peroxisomes. Alternatively, the absence of inclusion bodies from S. cerevisiae (YEp13Cat) transformants (or similar targeting results obtained subsequent to C2-11bK$_D$ expression at levels not inducing inclusion body formation) would be strongly suggestive of an internal PTS for Cat from C. tropicalis. Subsequent to defining the minimal PTS within amino acids 182 to 344 of Cat, its peroxisomal targeting ability could be further verified by creating CAT (or DHFR)/Cat fusion constructs and determining whether this minimal PTS functions to target these passenger proteins to S. cerevisiae peroxisomes. It would also be of interest to determine: (i) whether the functionality of this PTS is context specific [ie. does it retain its functionality when located elsewhere within a peroxisomal protein (eg. at the N- or C-terminus)] and (ii) if the functionality of this PTS is conserved across divergent species (eg. plants, insects and mammals), as has been determined for the C-terminal tripeptide PTS of firefly luciferase (Gould et al., 1990a).

BIBLIOGRAPHY

Aaji, C. and Borst, P. (1972) The gel electrophoresis of DNA. Biochim. Biophys. Acta 269, 192-200.

Adam, S.A., Marr, R.S. and Gerace, L. (1990) Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. J. Cell Biol. 111, 807-816.

Aitchison, J.D., Murray, W.W. and Rachubinski, R.A. (1991) 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. and Rachubinski, R.A. (1990) In vivo import of Candida tropicalis hydratase-dehydrogenase-epimerase into peroxisomes of Candida albicans. Curr. Genet. 17, 481-486.

Alexson, S.E.H., Fujiki, Y., Shio, H. and Lazarow, P.B. (1985) Partial disassembly of peroxisomes. J. Cell. Biol. 101, 294-305.

Allen, L.-A.H., Morand, O.H. and Raetz, C.R.H. (1989) Cytoplasmic requirement for peroxisome biogenesis in Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA 86, 7012-7016.

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.

Ammerer, G., Richter, K., Hartter, E. and Ruis, H. (1981) Synthesis of Saccharomyces cerevisiae catalase A in vitro. Eur. J. Biochem. 113, 327-331.

An, G. and Friesen, J.D. (1980) The nucleotide sequence of tufB and four nearby tRNA structural genes of Escherichia coli. Gene 12, 33-39.

Andersson, S.G.E. and Kurland, C.G. (1990) Codon preferences in free-living microorganisms. Microbiol. Rev. 54, 198-210.

Andrews, D. (1989) Examining protein translocation in cell-free systems and microinjected Xenopus oocytes. BioTechniques 7, 960-967.

Anthony, C. (1982) The Biochemistry of Methylotrophs. New York: Academic Press, pp. 269-295.

Antonenkov, V.D. (1989) Dehydrogenases of the pentose phosphate pathway in rat liver peroxisomes. Eur. J. Biochem. 183, 75-82.

Antonenkov, V.D. and Panchenko, L.F. (1978) Organization of urate oxidase in peroxisomal nucleoids. FEBS Lett. 88, 151-154.

Appelkvist, E.L. (1987) In vitro labeling of peroxisomal cholesterol with radioactive precursors. Bioscience Rep. 7, 853-858.

Appelkvist, E.L., Brunk, U. and Dallner, G. (1981) Isolation of peroxisomes from rat liver using sucrose and Percoll gradients. J. Biochem. Biophys. Methods 5, 203-217.

Appelkvist, E.L. and Dallner, G. (1980) Possible involvement of fatty acid binding protein in peroxisomal β-oxidation of fatty acids. Biochim. Biophys. Acta 617, 156-160.

Appelkvist, E.L. and Dallner, G. (1987) Dolichol metabolism and peroxisomes. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 53-66.

Appelkvist, E.L. and Kalén, A. (1989) Biosynthesis of dolichol by rat liver peroxisomes. Eur. J. Biochem. 185, 503-509.

Appelkvist, E.L., Reinhart, M., Fischer, R., Billheimer, J. and Dallner, G. (1990) Presence of individual enzymes of cholesterol biosynthesis in rat liver peroxisomes. Arch. Biochem. Biophys. 282, 318-325.

Arola, L., Herrera, E. and Alemany, M. (1977) A new method for deproteinization of small samples of blood plasma for amino acid determination. Anal. Biochem. 82, 236-239.

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. Current Protocols in Molecular Biology (Volume 2). New York: Greene Publishing Associates and Wiley-Interscience, 1988.

Aviv, H. and Leder, P. (1972) 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.

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.

Baudhuin, P. (1969) Liver peroxisomes, cytology and function. Ann. N.Y. Acad. Sci. 168, 214-228.

Baudhuin, P., Beaufay, J., Rahman-Li, Y., Sellinger, O.Z., Wattiaux, R., Jacques, P. and de Duve,C. (1964) Tissue fractionation studies. Intracellular distribution of monoamine oxidase, aspartate aminotransferase, alanine aminotransferase, D-amino acid oxidase and catalase in rat-liver tissue. Biochem. J. 92, 179-184.

Baudhuin, P., Müller, M., Poole, B. and de Duve, C. (1965) Nonmitochondrial oxidizing particles (microbodies) in rat liver and kidney and in Tetrahymena pyriformis. Biochem. Biophys. Res. Commun. 20, 53-59.

Baumgart, E., Völkl, A., Hashimoto, T. and Fahimi, H.D. (1989) Biogenesis of peroxisomes: Immunocytochemical investigation of peroxisomal membrane proteins in proliferating rat liver peroxisomes and catalase-negative membrane loops. J. Cell Biol. 108, 2221-2231.

Beaufay, H. (1966) La Centrifugation en Gradient de Densité, Ceuterick, Louvain.

Beaufay, H., Jacques, P., Baudhuin, P., Sellinger, O.Z., Berthet, J. and de Duve, C. (1964) Tissue fractionation

studies: Resolution of mitochondrial fractions from rat liver into three distinct populations of cytoplasmic particles by means of density equilibration in various gradients. Biochem. J. 92, 184-205.

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, 329-349.

Beevers, H. (1979) Microbodies in higher plants. Ann. Rev. Plant Physiol. 30, 159-193.

Beevers, H. (1982) Glyoxysomes in higher plants. Ann. N.Y. Acad. Sci. 386, 542-545.

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

Behrends, W., Birkhan, R. and Kindl, H. (1990) Transition form of microbodies. Overlapping of two sets of marker proteins during the rearrangement of glyoxysomes into leaf peroxisomes. Biol. Chem. Hoppe-Seyler 371, 85-94.

Behrends, W., Engeland, K. and Kindl, H. (1988) Characterization of two forms of the multifunctional protein acting in fatty acid β-oxidation. Arch. Biochem. Biophys. 263, 161-169.

Beitner, R. and Naor, Z. (1972) Isoenzymes of $NADP^+$- and NAD^+-glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in rat adipose tissue. Biochim. Biophys. Acta. 276, 572-575.

Bell, G.I., Najarian, R.C., Mullenbach, G.T. and Hallewell, R.A. (1986) cDNA sequence coding for human kidney catalase. Nucl. Acids Res. 14, 5561-5562.

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. (1982a) The primary structure of the Saccharomyces cerevisiae gene for alcohol dehydrogenase I. J. Biol. Chem. 257, 3018-3025.

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

Benz, R. (1985) Porin from bacterial and mitochondrial outer membranes. CRC Crit. Rev. Biochem. 19, 145-190.

Berge, R.K., Nilsson, A. and Husoy, A.-M. (1988) Rapid stimulation of liver palmitoyl-CoA synthetase, carnitine palmitoyltransferase and glycerophosphate acyltransferase compared to peroxisomal β-oxidation and palmitoyl-CoA hydrolase in rats fed high-fat diets. Biochim. Biophys. Acta 960, 417-426.

Bergner, U. and Tanner, W. (1981) Occurrence of several glycoproteins in glyoxysomal membranes of castor beans. FEBS Lett. 131, 68-72.

Bethards, L.A., Skadsen, R.W. and Scandalios, J.G. (1987) Isolation and characterization of a cDNA clone for the Cat2 gene in maize and its homology with other catalases. Proc. Natl. Acad. Sci. USA 84, 6830-6834.

Bieber, L.L. and Fiol, C.J. (1986) Characterization and properties of carnitine acyltransferases. Biochem. Soc. Trans. 14, 674-676.

Bieber, L.L., Valkner, K. and Farrell, S. (1982) Carnitine acyl transferases of liver peroxisomes. Ann. N.Y. Acad. Sci. 386, 395-396.

Bieker, K.L. and Silhavy, T.J. (1990) The genetics of protein secretion in E. coli. Trends Genet. 6, 329-334.

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

Björkhem, I., Kase, B.F. and Pederson, J.I. (1984) Mechanism of peroxisomal 24-hydroxylation of 3α,7α,12α-trihydroxy-5β-cholestanoic acid in rat liver. Biochim. Biophys. Acta 796, 142-145.

Björkhem, I., Kase, B.F. and Pederson, J.I. (1985) Role of

peroxisomes in the biosynthesis of bile acids. Scand. J. Clin. Lab. Invest. 45 (Suppl. 177), 23-31.

Blum, J.J. (1973) Localization of some enzymes of β-oxidation of fatty acids in the peroxisomes of Tetrahymena. J. Protozool. 20, 688-692.

Bodnar, A.G. (1991) Doctoral Thesis, McMaster University, Hamilton, Ontario, Canada.

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. (In press).

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, 248-254.

Breathnach, R. and Chambon, P. (1981) Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. 50, 349-383.

Breeuwer, M. and Goldfarb, D.S. (1990) Facilitated nuclear transport of histone H1 and other small nucleophilic proteins. Cell 60, 999-1008.

Breidenbach, R.W. and Beevers, H. (1967) Association of the glyoxylate cycle enzymes in a novel subcellular

particle from castor bean endosperm. Biochem. Biophys. Res. Commun. 27, 462-469.

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.

Bronfman, M., Inestrosa, N.C. and Leighton, F. (1979) Fatty acid oxidation by human liver peroxisomes. Biochem. Biophys. Res. Commun. 88, 1030-1036.

Bronfman, M. and Leighton, F. (1984) Carnitine acyltransferase and acyl-coenzyme A hydrolase activities in human liver. Quantitative analysis of their subcellular localization. Biochem. J. 224, 721-730.

Bruinenberg, P.G., Blaauw, M., Kazemier, B. and Geert, A.B. (1990) Cloning and sequencing of the malate synthase gene from Hansenula polymorpha. Yeast 6, 245-254.

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

Brul, S., Westerveld, A., Strijland, A., Wanders, R.J.A., Schram, A.W., Heymans, H.S.A., Schutgens, R.B.H., Van den Bosch, H. and Tager, J.M. (1988a) Genetic heterogeneity in the cerebrohepatorenal (Zellweger) syndrome and other inherited disorders with a generalized impairment of peroxisomal functions. J. Clin. Invest. 81, 1710-1715.

Brul, S., Wiemer, E.A.C., Westerveld, A., Strijland, A., Wanders, R.J.A., Schram, A.W., Heymans, H.S.A., Schutgens, R.B.H., Van den Bosch, H. and Tager, J.M. (1988b) Kinetics of the assembly of peroxisomes after fusion of complementary cell lines from patients with the cerebro-hepato-renal (Zellweger) syndrome and related disorders. Biochem. Biophys. Res. Commun. 152, 1083-1089.

Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.-T. and Dobberstein, B. (1987) A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. Methods Enzymol. 155, 416-433.

Bulmer, M. (1987) Coevolution of codon usage and transfer RNA abundance. Nature 325, 728-730.

Burke, J.J. and Trelease, R.M. (1975) Cytochemical demonstration of malate synthase and glycolate oxidase in microbodies of cucumber cotyledons. Plant Physiol. 56, 710-717.

Burke, R.L., Tekamp-Olson, P. and Najarian, R. (1983) The isolation, characterization, and sequence of the pyruvate kinase gene of Saccharomyces cerevisiae. J. Biol. Chem. 258, 2193-2201.

Burnette, W.N. (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 Nedergaard, J. (1982) Peroxisomal β-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.

Cerdan, S., Künnecke, B., Dölle, A. and Seelig, J. (1988) In situ metabolism of 1,ω medium chain dicarboxylic acids in the liver of intact rats as detected by ^{13}C and ^{1}H NMR. J. Biol. Chem. 263, 11664-11674.

Chamberlin, M. and Ring, J. (1973) Characterization of T7-specific ribonucleic acid polymerase. J. Biol. Chem. 248, 2235-2244.

Chang, Y.F. (1978) Lysine metabolism in the rat brain: Blood-brain barrier transport, formation of pipecolic acid and human hyperpipecolatemia. J. Neurochem. 30, 355-360.

Chatterjee, B., Song, C.S., Kim, J.M. and Roy, A.K. (1988) Cloning, sequencing and regulation of rat liver carnitine octanoyltransferase: transcriptional stimulation of the enzyme during peroxisome proliferation. Biochemistry 27, 9000-9006.

Chen, C., Kalb, V.F., Turi, T.G. and Loper, J.C. (1988)

Primary structure of the cytochrome P450 lanosterol 14α-demethylase gene from Candida tropicalis. DNA Cell Biol. 7, 617-626.

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, L. and Carbon, J. (1976) A colony bank containing synthetic Col EI hybrid plasmids representative of the entire E. coli genome. Cell 9, 91-99.

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

Cohen, C. and Parry, D.A.D. (1986) α-Helical coiled coils- a widespread motif in proteins. Trends Biochem. Sci. 11, 245-248.

Cohen, G., Rapatz, W. and Ruis, H. (1988) Sequence of the Saccharomyces cerevisiae CTA1 gene and amino acid sequence of catalase A derived from it. Eur. J. Biochem. 176, 159-163.

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.

Connock, M.J. and Perry, S.R. (1983) Detection of acyl-CoA β-oxidation enzymes in peroxisomes (microperoxisomes) of mouse heart. Biochem. Int. 6, 545-551.

Cooper, T.G. and Beevers, H. (1969) β-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.

Corner, B.E. and Poulter, R.T.M. (1989) Interspecific complementation analysis by protoplast fusion of Candida tropicalis and Candida albicans adenine auxotrophs. J. Bacteriol. 171, 3586-3589.

Coudron, P.E., Frerman, F.E. and Schowalter, D.B. (1983) Chemical and catalytic properties of the peroxisomal acyl-coenzyme A oxidase from Candida tropicalis. Arch.

Biochem. Biophys. 226, 324-336.

Crane, D.I. and Masters, C.J. (1986) The effect of clofibrate on the phospholipid composition of the peroxisomal membranes in mouse liver. Biochim. Biophys. Acta 876, 256-263.

Cromartie, T.H. and Walsh, C.T. (1975) Rat kidney L-α-hydroxy acid oxidase: isolation of enzyme with one flavin coenzyme per two subunits. Biochemistry 14, 2588-2596.

Crouse, G.F., Leys, E.J., McEwan, R.N., Frayne, E.G. and Kellems, R.E. (1985) Analysis of the mouse dhfr promoter region: existence of a divergently transcribed gene. Mol. Cell. Biol. 5, 1847-1858.

Danpure, C.J., Cooper, P.J., Wise, P.J. and Jennings, P.R. (1989) An enzyme trafficking defect in two patients with primary hyperoxaluria type 1: Peroxisomal alanine:glyoxylate aminotransferase rerouted to mitochondria. J. Cell Biol. 108, 1345-1352.

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

Davis, W.L., Matthews, J.L. and Goodman, D.B.P. (1989) Glyoxylate cycle in the rat liver: effect of vitamin D_3 treatment. FASEB J. 3, 1651-1655.

Declercq, P.E., Haagsman, H.P., Van Veldhoven, P., Debeer, L.J., Van Golde, L.M.G. and Mannaerts, G.P. (1984) Rat liver dihydroxyacetone-phosphate acyltransferases and their contribution to glycerolipid synthesis. J. Biol. Chem. 259, 9064-9075.

de Duve, C. (1965a) The separation and characterization of subcellular particles. Harvey Lectures Ser. 59, 49-87.

de Duve, C. (1965b) Functions of microbodies (peroxisomes). Abstr. Vth Meeting Am. Soc. Cell. Biol. 27, 25A.

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

de Duve, C. (1989) Commentary on 'Intracellular localization of catalase and of some oxidases in rat liver'. Biochim. Biophys. Acta 1000, 317-320.

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

de Duve, C., Beaufay, H., Jacques, P., Rahman-Li, Y., Sellinger, O.Z., Wattiaux, R. and de Coninck, S. (1960) Intracellular localization of catalase and of some oxidases in rat liver. Biochim. Biophys. Acta 40, 186-187.

de Duve, C. Gianetto, R., Appelmans, F. and Wattiaux, R. (1953) Enzymic content of the mitochondria fraction. Nature 172, 1143-1144.

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 tissue. Biochem. J. 60, 604-617.

Delcuve, G., Downing, W., Lewis, H. and Dennis, P.P. (1980) Nucleotide sequence of the proximal portion of the RNA polymerase β subunit gene of Escherichia coli. Gene 11, 367-373.

della-Cioppa, G. and Kishore, G.M. (1988) Import of a precursor protein into chloroplasts is inhibited by the herbicide glyphosate. EMBO J. 7, 1299-1305.

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

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.

de Wet, J.R., Wood, K.W., DeLuca, M., Helinski, D.R. and Subramani, S. (1987) Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. 264, 13929-13937.

Diczfalusy, U. and Alexson, S.E.H. (1988) Peroxisomal chain-shortening of prostaglandin $F_{2\alpha}$. J. Lipid Res. 29, 1629-1636.

Diczfalusy, U. and Alexson, S.E.H. (1990) Identification of metabolites from peroxisomal β-oxidation of

prostaglandins. J. Lipid Res. 31, 307-314.

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

Dmochowska, A., Dignard, D., Maleszka, R. and Thomas, D.Y. (1990) Structure and transcriptional control of the Saccharomyces cerevisiae POX1 gene encoding acyl-coenzyme A oxidase. Gene 88, 247-252.

Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J., Kingsman, S.M., Perkins, R.E., Conroy, S.C., Dunbar, B. and Fothergill, L.A. (1982) Conservation of high efficiency promoter sequences in Saccharomyces cerevisiae. Nucl. Acids Res. 10, 2625-2637.

Dommes, P., Dommes, V. and Kunau, W.-H. (1983) β-oxidation in Candida tropicalis: partial purification and biological function of an inducible 2,4-dienoyl coenzyme A reductase. J. Biol. Chem. 258, 10846-10852.

Donaldson, R.P. and Beevers, H. (1977) Lipid composition of organelles from germinating castor bean endosperm. Plant Physiol. 59, 259-263.

Donaldson, R.P., Tolbert, N.E. and Schnarrenberger, C. (1972) A comparison of microbody membranes with microsomes and mitochondria from plant and animal tissue. Arch. Biochem. Biophys. 152, 199-215.

Douglas, S.A., Criddle, R.S. and Breidenbach, R.W. (1973) Characterization of deoxyribonucleic acid species from castor bean endosperm. Plant Physiol. 51, 902-906.

Douma, A.C., Vennhuis, M., deKoning, W., Evers, M. and Harder, W. (1985) Dihydroxyacetone synthase is localized in the peroxisomal matrix of methanol-grown Hansenula polymorpha. Arch. Microbiol. 143, 237-243.

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.

Dovey, H.F., Parsons, M. and Wang, C.C. (1988) Biogenesis of glycosomes of Trypanosoma brucei: An in vitro model of 3-phosphoglycerate kinase import. Proc. Natl. Acad. Sci. USA 85, 2598-2602.

Dubourdieu, M. and Fox, J.L. (1977) Amino acid sequence of Desulfovibrio vulgaris flavodoxin. J. Biol. Chem. 252, 1453-1463.

Dynan, W.S., Sazer, S., Tjian, R. and Schimke, R.T. (1986) Transcription factor Sp1 recognizes a DNA sequence in the mouse dihydrofolate reductase promoter. Nature 319, 246-248.

Dynan, W.S. and Tjian, R. (1985) Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature 316, 774-778.

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.

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

Ekström, T.J., Chojnacki, T. and Dallner, G. (1987) The α-saturation and terminal events in dolichol biosynthesis. J. Biol. Chem. 262, 4090-4097.

Ellis, J. (1987) Proteins as molecular chaperones. Nature 328, 378-379.

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) PAS1, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. Cell 64, 499-510.

Erickson, A.H. and Blobel, G. (1983) Cell-free translation of messenger RNA in a wheat germ system. Meth. Enzymol. 96, 38-50.

Ettinger, W.F. and Harada, J.J. (1990) Translational or post-translational processes affect differentially the accumulation of isocitrate lyase and malate synthase proteins and enzyme activities in embryos and seedlings of Brassica napus. Arch. Biochem. Biophys. 281, 139-143.

Fahimi, H.D. and Sies, H. Peroxisomes in Biology and Medicine. Berlin: Springer-Verlag, 1987.

Farabaugh, P.J. (1978) Sequence of the lacI gene. Nature 274, 765-769.

Farrell, S.O., Fiol, C.J., Reddy, J.K. and Bieber, L.L. (1984) Properties of purified carnitine acyltransferases of mouse liver peroxisomes. J. Biol. Chem. 259, 13089-13095.

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) Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137, 266-267.

Finlay, B.J. and Fenchel, T. (1989) FEMS Microbiol. Lett. 65, 311-314.

Fita, I. and Rossman, M.G. (1985) The NADPH binding site on beef liver catalase. Proc. Natl. Acad. Sci. USA 82, 1604-1608.

Flatmark, T., Nilsson, A., Kvannes, J., Eikhom, T.S., Fukami, M.H., Kryvi, H. and Christiansen, E.N. (1988) On the mechanism of induction of enzyme systems for peroxisomal β-oxidation of fatty acids in rat liver by diets rich in partially hydrogenated fish oil. Biochim. Biophys. Acta 962, 122-130.

Flügge, U.-I. (1990) Import of proteins into chloroplasts. J. Cell Sci. 96, 351-354.

Frerman, F.E. (1988) Acyl-CoA dehydrogenases, electron transfer flavoprotein and electron transfer flavoprotein dehydrogenase. Biochem. Soc. Trans. 16, 416-418.

Frevert, J. and Kindl, H. (1978) Plant microbody proteins. Purification and glycoprotein nature of glyoxysomal isocitrate lyase from cucumber cotyledons. Eur. J. Biochem. 92, 35-43.

Frevert, J. and Kindl, H. (1980) A bifunctional enzyme from glyoxysomes. Purification of a protein possessing enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase

activities. Eur. J. Biochem. 107, 79-86.

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

Frischauf, A.M., Murray, N. and Lehrach, H. (1987) phage vectors - EMBL series. Meth. Enzymol. 153, 103-115.

Fröhlich, K.-U., Entian, K.-D. and Mecke, D. (1985) The primary structure of the yeast hexokinase PII gene (HXK2) which is responsible for glucose repression. Gene 36, 105-111.

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

Fujiki, Y., Hubbard, A.L., Fowler, S. and Lazarow, P.B. (1982b) Isolation of intracellular membranes by means of sodium carbonate treatment. Application to endoplasmic reticulum. J. Cell. Biol. 93, 97-102.

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., 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.

Fujiwara, S., Ohashi, H. and Noguchi, T. (1987) Comparison of intraperoxisomal localization form and properties of amphibian (Rana catesbeiana) uricase with those of other animal uricases. Comp. Biochem. Physiol. 86B, 23-26.

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

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.

Gänsler, H. and Rouiller, C. (1956) Modifications physiologiques et pathologiques du chondriome. Schweiz Z. Allgem. Pathol. Bacteriol. 19, 217-243.

Garland, P.B., Yates, D.W. and Haddock, B.A. (1970) Spectrophotometric studies of acyl-coenzyme A synthetases of rat liver mitochondria. Biochem. J. 119, 553-564.

Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120, 97-120.

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

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

Gaylor, J.L. (1981) In: Biosynthesis of Isoprenoid Compounds, Vol. I (Porter, J.W. and Spurgeon, S.L., eds.), New York, John Wiley and Sons, pp. 481-544.

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

Gödecke, A., Veenhuis, M., Roggenkamp, R., Janowicz, Z.A. and Hollenberg, C.P. (1989) Biosynthesis of the peroxisomal dihydroxyacetone synthase from Hansenula polymorpha in Saccharomyces cerevisiae induces growth but not proliferation of peroxisomes. Curr. Genet. 16, 13-20.

Goldenberg, H. (1977) Organization of purine degradation in the liver of a teleost (carp; Cyprinus carpio L.). A study of its subcellular distribution. Mol. Cell. Biochem. 16, 17-21.

Goldfarb, D.S., Gariepy, J., Schoolnik, G. and Kornberg, R.D. (1986) Synthetic peptides as nuclear localization signals. Nature 322, 641-644.

Goldfischer, S. (1969) Further observations of the

peroxidatic activities of microbodies (peroxisomes). J. Histochem. Cytochem. 17, 681-685.

Goldfischer, S., Collins, J., Rapin, I., Neumann, P., Neglia, W., Spiro, A.J., Ishii, T., Roels, F., Vamecq, J. and Van Hoof, F. (1986) Pseudo-Zellweger syndrome: Deficiencies in several peroxisomal oxidative activities. J. Pediatr. 108, 25-32.

Goldfischer, S. and Reddy, J.K. (1984) Peroxisomes (microbodies) in cell pathology. Int. Rev. Exp. Path. 26, 45-84.

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.

Goodman, D.B., 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. (1985) Dihydroxyacetone synthase is an abundant constituent of the methanol-induced peroxisome of Candida boidinii. J. Biol. Chem. 260, 7108-7113.

Gorgas, K. (1982) Serial section analysis of peroxisomal shape and membrane relationships in the mouse preputial gland. Ann. N.Y. Acad. Sci. 386, 519-522.

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

Gorgas, K. (1985) Serial section analysis of mouse hepatic peroxisomes. Anat. Embryol. 172, 21-32.

Gorgas, K. (1987) Morphogenesis of peroxisomes in lipid-synthesizing epithelia. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 3-17.

Goudie, R.B., Horne, C.H.W. and Wilkinson, P.C. (1966) A simple method for producing antibody specific to a single selected diffusible antigen. Lancet ii, 1224-1226.

Gould, S.J., Keller, G.-A., Schneider, M., Howell, S.H.,

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

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 peroxisomal targeting signals located at the carboxy terminus of four peroxisomal proteins. J. Cell Biol. 107, 897-905.

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

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

Gouy, M. and Gautier, C. (1982) Codon usage in bacteria: correlation with gene expressivity. Nucl. Acids Res. 10, 7055-7074.

Graham, I.A., Smith, L.M., Brown, J.W.S., Leaver, C.L. and Smith, S.M. (1989) The malate synthase gene of cucumber. Plant Mol. Biol. 13, 673-684.

Grasso, P. (1985) Peroxisome proliferation and hepatotoxicity in rodents. Biochem. Soc. Trans. 13, 861-862.

Graves, L.B. and Becker, W.M. (1974) Beta-oxidation in glyoxysomes from Euglena. J. Protozool. 21, 771-774.

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.

Guarente, L. (1987) Regulatory proteins in yeast. Annu. Rev. Genet. 21, 425-452.

Guarente, L. (1988) UASs and enhancers: common mechanism of transcriptional activation in yeast and mammals. Cell

52, 303-305.

Gubler, U. and Hoffman, B.J. (1983) A simple and very efficient method for generating cDNA libraries. Gene 25, 263-269.

Haeuptle, M.-T., Frank, R. and Dobberstein, B. (1986) Translation arrest by oligodeoxynucleotides complementary to mRNA coding sequences yields polypeptides of predetermined length. Nucl. Acids Res. 14, 1427-1448.

Hagey, L.R. and Krisans, S.K. (1982) Degradation of cholesterol to propionic acid by rat liver peroxisomes. Biochem. Biophys. Res. Commun. 107, 834-841.

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.

Hajra, A.K., Burke, C.L. and Jones, C.L. (1979) Subcellular localization of acyl-coenzyme A: dihydroxyacetone phosphate acyltransferase in rat liver peroxisomes (microbodies). J. Biol. Chem. 254, 10896-10900.

Hallermayer, G., Zimmerman, R. and Neupert, W. (1977) Kinetic studies on the transport of cytoplasmically synthesized proteins into the mitochondria in intact cells of Neurospora crassa. Eur. J. Biochem. 81, 523-532.

Hamilton, G.A., Bukthal, D.J., Mortensen, R.M. and Zerby, K.W. (1979) Reactions of cysteamine and other amine metabolites with glyoxylate and oxygen catalyzed by mammalian D-amino acid oxidase. Proc. Natl. Acad. Sci. USA 76, 2625-2629.

Hanahan, D. (1983) Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol. 166, 557-580.

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.

Hardeman, D., Versantvoort, C., van den Brink, J.M. and van den Bosch, H. (1990) Studies on peroxisomal membranes. Biochim. Biophys. Acta 1027, 149-154.

Harson, M.M., Conder, M.J. and Lord, J.M. (1983) Endoplasmic reticulum and glyoxysomal membranes from castor-bean endosperm: interaction between membrane glycoproteins and organelle matrix proteins. Planta 157, 143-149.

Hart, D.T., Baudhuin, P., Opperdoes, F.R. and de Duve, C. (1987) Biogenesis of the glycosome in Trypanosoma brucei: the synthesis, translocation and turnover of glycosomal polypeptides. EMBO J. 6, 1403-1411.

Hart, D.T., Misset, O., Edwards, S.W. and Opperdoes, F.R. (1984) A comparison of the glycosomes (microbodies) isolated from Trypanosoma brucei bloodstream form and cultured procyclic trypomastigotes. Mol. Biochem. Parasitol. 12, 25-36.

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.

Hartig, A. and Ruis, H. (1986) Nucleotide sequence of the Saccharomyces cerevisiae CTT1 gene and deduced amino acid sequence of yeast catalase T. Eur. J. Biochem. 160, 487-490.

Hartl, F.-U. and Neupert, W. (1990) Protein sorting to mitochondria: evolutionary conservations of folding and assembly. Science 247, 930-938.

Hartl, F.-U., Ostermann, J., Guiard, B. and Neupert, W. (1987) Successive translocation into and out of the mitochondrial matrix: Targeting of proteins to the intermembrane space by a bipartite signal peptide. Cell 51, 1027-1037.

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

Hase, J., Müller, U., Riezman, H. and Schatz, G. (1984) A 70-kd protein of the yeast mitochondrial outer membrane

is targeted and anchored via its extreme amino terminus. EMBO J. 3, 3157-3164.

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

Hashimoto, T. (1987) Comparison of enzymes of lipid β-oxidation in peroxisomes and mitochondria. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 94-104.

Hay, R., Böhni, P. and Gasser, S. (1984) How mitochondria import proteins. Biochim. Biophys. Acta 779, 65-87.

Hemming, F.W. (1983) In: Biosynthesis of Isoprenoid Compounds, Vol. II (Porter, J.W. and Spurgeon, S.L., eds.), New York, John Wiley and Sons, pp. 305-354.

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) Sequences responsible for transcription termination on a gene segment in Saccharomyces cerevisiae. Mol. Cell. Biol. 4, 1515-1520.

Heymans, H.S.A., Oorthuys, J.W.E., Nelck, G., Wanders, R.J.A. and Schutgens, R.B.H. (1985) Rhizomelic chondrodysplasia punctata: another peroxisomal disorder. N. Engl. J. Med. 313, 187-188.

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.

Heymans, H.S.A., van den Bosch, H., Schutgens, R.B.H., Tegelaers, W.H.H., Walter, J.U., Müller-Höcker, J. and Borst, P. (1984) Deficiency of plasmalogens in the cerebro-hepato-renal (Zellweger) syndrome. Eur. J. Pediatr. 142, 10-15.

Higashi, T. and Peters, T.J. (1963) Studies on rat liver catalase. II. Incorporation of ^{14}C-leucine into catalase of liver cell fractions in vivo. J. Biol. Chem. 238, 3952-3954.

Hijikata, M., Ishii, N., Kagamiyana, 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: occurrence of two closely related but differentially regulated genes. 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.

Hitzeman, R.A., Hagie, F.E., Hayflick, J.S., Chen, C.Y., Seeburg, P.H. and Derynck, R. (1982) The primary structure of the Saccharomyces cerevisiae gene for 3-phosphoglycerate kinase. Nucl. Acids Res. 10, 7791-7808.

Hoffman, K.E. and Gilmore, R. (1988) Guanosine triphosphate promotes the post-translational integration of opsin into the endoplasmic reticulum membrane. J. Biol. Chem. 263, 4381-4385.

Hogg, J.F. (1969) Peroxisomes in Tetrahymena and their relation to gluconeogenesis. Ann. NY Acad. Sci. 168, 209-381.

Holland, J.P. and Holland, M.J. (1979) The primary structure of a glyceraldehyde-3-phosphate dehydrogenase gene from Saccharomyces cerevisiae. J. Biol. Chem. 254, 9839-9845.

Holland, J.P. and Holland, M.J. (1980) Structural comparison of two nontandemly repeated yeast glyceraldehyde-3-phosphate dehydrogenase genes. J. Biol. Chem. 255, 2596-2605.

Holland, M.J., Holland, J.P., Thill, G.P. and Jackson, K.A. (1981) The primary structures of two yeast enolase genes. J. Biol. Chem. 256, 1385-1395.

Holm, L. (1986) Codon usage and gene expression. Nucl. Acids Res. 14, 3075-3087.

Holmes, D.S. and Quigley, M. (1981) A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114, 193-197.

Hölttä, E. (1977) Oxidation of spermidine and spermine in

rat liver: Purification and properties of polyamine oxidase. Biochemistry 16, 91-100.

Holtzmann, E. (1982) Peroxisomes in nervous tissue. Ann. N.Y. Acad. Sci. 386, 523-525.

Honinberg, B.M., Volkmann, D., Entzeroth, R. and Scholtyseck, E. (1984) A freeze-fracture electron microscope study of Trichomonas vaginalis donné and Trichomonas foetus (Reidmuller). J. Protozool. 31, 116-131.

Hopp, T.P. and Woods, K.R. (1981) Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78, 3824-3828.

Hori, S.H. and Yonezawa, S. (1972) The in vitro inter-conversion by mercaptans of type I glucose-6-phosphate dehydrogenase isozymes of rat. J. Histochem. Cytochem. 20, 804-810.

Hörtner, H., Ammerer, G., Hartter, E., Hamilton, B., Rytka, J., Bilinski, T. and Ruis, H. (1982) Regulation and synthesis of catalases and iso-1-cytochrome C in Saccharomyces cerevisiae by glucose, oxygen and heme. Eur. J. Biochem. 128, 179-184.

Hovic, R. and Osmundsen, H. (1987) Peroxisomal β-oxidation of long-chain fatty acids possessing different extents of unsaturation. Biochem. J. 247, 531-535.

Hruban, Z. and Rechcigl Jr., M. Microbodies and Related Particles. New York: Academic Press, 1969.

Hruban, Z. and Swift, H. (1964) Uricase: localization in hepatic microbodies. Science 146, 1316-1318.

Hryb, D.J. and Hogg, J.F. (1976) A peroxisomal fatty acyl-coenzyme A oxidase in Tetrahymena pyriformis. Fed. Proc. 35, 1501.

Huang, A.H.C. and Beevers, H. (1973) Localization of enzymes within microbodies. J. Cell. Biol. 58, 379-389.

Huang, S., Elliott, R.C., Liu, P.-S., Koduri, R.K., Weickmann, J.L., Lee, J.-H., Blair, L.C., Ghosh-Dastidar, P., Bradshaw, R.A., Bryan, K.M., Einarson, B., Kendall, R.L., Kolacz, K.H. and Saito, K. (1987) Specificity of cotranslational amino-terminal processing of proteins in yeast. Biochemistry 26,

8242-8246.

Hutzler, J. and Dancis, J. (1975) Lysine-ketoglutarate reductase in human tissues. Biochim. Biophys. Acta 377, 42-51.

Ikemura, T. (1985) Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. 2, 13-34.

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.

Ishii, N., Hijikata, M., Osumi, T. and Hashimoto, T. (1987) Structural organization of the gene for rat enoyl-CoA hydratase: 3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme. J. Biol. Chem. 262, 8144-8150.

Ishii, S., Xu, Y.H., Stratton, R.H., Roe, B.A., Merlino, G.T. and Pastan, I. (1985) Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. Proc. Natl. Acad. Sci. USA 82, 4920-4924.

Itakura, K., Rossi, J.J. and Wallace, R.B. (1984) Synthesis and use of synthetic oligonucleotides. Annu. Rev. Biochem. 53, 323-356.

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

Jacobs, P., Brockly, F., Massaer, M., Loriau, R., Guillaume, J.P., Ciccarelli, E., Heinderyckx, M., Cravador, A., Biemans, R., van Elsen, A., Herzog, A. and Bollen, A. (1987) Porcine D-amino acid oxidase: determination of the mRNA nucleotide sequence by the characterization of genomic and cDNA clones. Gene 59, 55-61.

Janowicz, Z.A., Eckart, M.R., Drewke, C., Roggenkamp, R.O. 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.

Johnson, P.J., D'Oliveira, C.E., Gorrell, T.E. and Müller, M. (1990) Molecular analysis of the hydrogenosomal

ferredoxin of the anaerobic protist Trichomonas vaginalis. Proc. Natl. Acad. Sci. USA 87, 6097-6101.

Jones, C.T. (1980) Is there a glyoxylate cycle in the liver of the fetal guinea pig? Biochem. Biophys. Res. Commun. 95, 849-856.

Kafatos, F.C., Jones, C.W. and Efstratiadis, A. (1979) Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucl. Acids Res. 7, 1541-1552.

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., 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. 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., Sakasegawa, Y. and Tan, H. (1989) Expression and transport of Candida tropicalis peroxisomal acyl-coenzyme A oxidase in the yeast Candida maltosa. Agric. Biol. Chem. 53, 179-186.

Kase, B.F. and Björkhem, I. (1989) Peroxisomal bile acid CoA: amino acid N-acyltransferase in rat liver. J. Biol. Chem. 264, 9220-9223.

Kase, B.F., Björkhem, I. and Pederson, J.I. (1983) Formation of cholic acid from 3α,7α,12α-trihydroxy-5β-cholestanoic acid by rat liver peroxisomes. J. Lipid Res. 24, 1560-1567.

Kawamoto, S., Nozaki, C., Tanaka, A. and Fukui, S. (1978) Fatty acid β-oxidation system 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.

Keilin, J. (1959) The biological significance of uric acid and guanine exretion. Biol. Rev. 34, 265-275.

Keith, T.P., Riley, M.A., Kreitman, M., Lewontin, R.C., Curtis, D. and Chambers, G. (1987) Sequence of the structural gene for xanthine dehydrogenase (rosy locus) in Drosophila melanogaster. Genetics 116, 67-73.

Keller, G.-A., Barton, M.C., Shapiro, D.J. and Singer, S.J. (1985) 3-Hydroxy-3-methylglutaryl-coenzyme A reductase is present in peroxisomes in normal rat liver cells. Proc. Natl. Acad. Sci. USA 82, 770-774.

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, 1-5.

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

Keller, G.-A., Pazirandeh, M. and Krisans, S. (1986) 3-Hydroxy-3-methylglutaryl coenzyme A reductase localization in rat liver peroxisomes and microsomes of control and cholestyramine-treated animals: Quantitative biochemical and immunoelectron microscopical analyses. J. Cell Biol. 103, 875-886.

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.

Kelly, R., Miller, S.M., Kurtz, M.B. and Kirsch, D.R. (1987) Directed mutagenesis in Candida albicans: one-step gene disruption to isolate ura 3 mutants. Mol. Cell. Biol. 7, 198-208.

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

Kindl, H. (1987) In: The Biochemistry of Plants (Stumpf, P.K. and Conn, E.E., eds.), London, Academic Press, pp. 31-52.

Kindl, H. and Lazarow, P.B. Peroxisomes and Glyoxysomes. New York: The New York Academy of Sciences, 1982.

Kirkman, H.N. and Gaetani, G.F. (1984) Catalase: A tetrameric enzyme with four tightly bound molecules of NADPH. Proc. Natl. Acad. Sci. USA 81, 4343-4347.

Kirwin, P.M., Meadows, J.W., Shackleton, J.B., Musgrove, J.E., Elderfield, P.D., Mould, R., Hay, N.A. and Robinson, C. (1989) ATP-dependent import of a luminal protein by isolated thylakoid vesicles. EMBO J. 8, 2251-2255.

Kleene, R., Pfanner, N., Pfaller, R., Link, T.A., Sebald, W., Neupert, W. and Tropschug, M. (1987) Mitochondrial porin of Neurospora crassa: cDNA cloning, in vitro expression and import into mitochondria. EMBO J. 6, 2627-2633.

Köller, W. and Kindl, H. (1978) Studies supporting the concept of glyoxyperoxisomes as intermediary organelles in transformation of glyoxysomes into peroxisomes. Z. Naturforsch. 33c, 962-968.

Kolvraa, S. and Gregerson, N. (1986) In vitro studies on the oxidation of medium-chain dicarboxylic acids in rat liver. Biochim. Biophys. Acta 876, 515-525.

Kondrup, J. and Lazarow, P.B. (1982) Peroxisomal beta oxidation in intact rat hepatocytes: quantitation of its flux. Ann. N.Y. Acad. Sci. 386, 404-405.

Kopetzki, E., Entian, K.-D. and Mecke, D. (1985) Complete nucleotide sequence of the hexokinase PI gene (HXK1) of Saccharomyces cerevisiae. Gene 39, 95-102.

Kornberg, A. and Pricer Jr., W.E. (1953) Enzymatic synthesis of the coenzyme A derivatives of long chain fatty acids. J. Biol. Chem. 204, 329-343.

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. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucl. Acids Res. 15, 8125-8148.

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

Kozak, M. (1989b) Evaluation of the fidelity of initiation of translation in reticulocyte lysates from commercial

sources. Nucl. Acids Res. 18, 2828.

Kozak, M. (1989c) Context effects and inefficient initation at non-AUG codons in eucaryotic cell-free translation systems. Mol. Cell. Biol. 9, 5073-5080.

Krieg, P.A. and Melton, D.A. (1987) In vitro RNA synthesis with SP6 RNA polymerase. Methods Enzymol. 155, 397-415.

Krisans, S.K., Thompson, S.L., Pena, L.A., Kok, E. and Javitt, N.B. (1985) Bile acid synthesis in rat liver peroxisomes: metabolism of 26-hydroxycholesterol to 3β-hydroxy-5-cholenoic acid. J. Lipid Res. 26, 1324-1332.

Kruse, C. and Kindl, H. (1983) Oligomerization of malate synthase during glyoxysome biosynthesis. Arch. Biochem. Biophys. 233, 629-638.

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

Kunau, W.-H., Kionka, C., Ledebur, A., Mateblowski, M., Moreno de la Garza, M., Schultz-Borchard, U., Thieringer, R. and Veenhuis, M. (1987) β-oxidation systems in eukaryotic microorganisms. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 128-140.

Kurihara, T., Ueda, M. and Tanaka, A. (1988) Occurrence and possible roles of acetoacetyl-CoA thiolase and 3-ketoacyl-CoA thiolase in peroxisomes of an n-alkane-grown yeast, Candida tropicalis. FEBS Lett. 229, 215-218.

Kurihara, T., Ueda, M. and Tanaka, A. (1989) Peroxisomal acetoacetyl-CoA thiolase and 3-ketoacyl-CoA thiolase from an 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., Lai, M. and Kirsch, D.R. (1987) Development of autonomously replicating plasmids for Candida albicans. Mol. Cell.

Biol. 7, 209-217.

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

Labarca, P., Wolff, D., Sato, U., Necochea, C. and Leighton, F. (1986) Large cation-selection pores from rat liver peroxisomal membranes incorporated into planar lipid bilayers. J. Membr. Biol. 94, 285-291.

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

Lake, B.G. and Gray, T.J.B. (1985) Species differences in hepatic peroxisome proliferation. Biochem. Soc. Trans. 13, 859-861.

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

Lazarow, P.B. (1984) The peroxisomal membrane. In: Membrane Structure and Function. Vol. 5 (Bittar, E.E., ed.), New York, Wiley Interscience, pp. 1-31.

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

Lazarow, P.B. (1987b) The role of peroxisomes in mammalian cell metabolism. J. Inher. Metab. Dis. 10 (Suppl. 1), 11-22.

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. and Fujiki, Y. (1985) Biogenesis of peroxisomes. Annu. Rev. Cell Biol. 1, 489-530.

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, 285-298.

Lazarow, P.B., Shio, H. and Robbi, M. (1980) Biogenesis of peroxisomes and the peroxisome reticulum hypothesis. In: 31st Mosbach Collaq. Biol. Chem. Organelle Formation (Bucher, T., Sebald, W. and Weiss, H., eds.), New York, Springer-Verlag, pp. 187-206.

Lazdunski, C.J. and Benedetti, H. (1990) Insertion and translocation of proteins into and through membranes. FEBS Lett. 268, 408-414.

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

Lee, C. and Beckwith, J. (1986) Cotranslational and posttranslation protein translocation in prokaryotic systems. Annu. Rev. Cell Biol. 2, 315-336.

Legg, P.G. and Wood, R.L. (1970) New observations on microbodies. A cytochemical study on CPIB-treated rat liver. 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. Improved isolation procedures, automated analysis, biochemical and morphological properties of fractions. J. Cell. Biol. 37, 482-513.

Lemmens, M., Verheyden, K., Van Veldhoven, P., Vereecke, J., Mannaerts, G.P. and Carmeliet, E. (1989) Single-channel analysis of a large conductance channel in peroxisomes from rat liver. Biochim. Biophys. Acta 984, 351-359.

Lettir, M. and Dubach, U.C. (1981) The activity pattern of two peroxisomal oxidases in the rat nephron. FEBS Lett. 127, 250-252.

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

Lightfoot, D.A. (1988) Magnesium-dependance of in vitro translation programmed by gene-specific mRNAs. Nucl. Acids Res. 16, 4164.

Lipman, D.J. and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches. Science 227, 1435-1441.

Liscum, L., Cummings, R.D., Anderson, R.G.W., De Martino, G.M., Goldstein, J.L. and Brown, M.S. (1983) 3-Hydroxy-3-methylglutaryl coenzyme A reductase: a transmembrane glycoprotein of the endoplasmic reticulum with N-linked "high-mannose" oligosaccharides. Proc. Natl. Acad. Sci. USA 80, 7165-7169.

Loewen, P.C. and Switala, J. (1987) Purification and characterization of catalase-1 from Bacillus subtilis. Biochem. Cell Biol. 65, 939-947.

Lord, J.M. and Roberts, L.M. (1983) Formation of glyoxysomes. Int. Rev. Cytol. (Suppl.) 15, 115-156.

Maccecchini, M.-L., Rudin, Y., Blobel, G. and Schatz, G. (1979) Import of proteins into mitrochondria: precursor forms of the extramitochondrially made F_1-ATPase subunits in yeast. Proc. Natl. Acad. Sci. USA 76, 343-347.

Mandel, M. and Higa, A. (1970) Calcium dependent bacteriophage DNA infection. J. Mol. Biol. 53, 159-162.

Maniatis, T., Fritsch, E.F. and Sambrook, J. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 1982.

Mannaerts, G.P., Debeer, L.J., Thomas, J. and DeShepper, P.J. (1979) Mitochondrial and peroxisomal fatty acid oxidation in liver. Homogenates and isolated hepatocytes from control and clofibrate-treated rats. J. Biol. Chem. 54, 4585-4595.

Mannaerts, G.P. and Van Veldhoven, P.P. (1990) The peroxisome: functional properties in health and disease. Biochem. Soc. Trans. 18, 87-89.

Mannaerts, G.P., Van Veldhoven, P., Van Broekhoven, A., Vanderbroeck, G. and Debeer, L.J. (1982) Evidence that peroxisomal acyl-CoA synthetase is located at the cytoplasmic side of the peroxisomal membrane. Biochem. J. 204, 17-23.

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.

Masters, C.J. and Crane, D. (1984) The role of peroxisomes in lipid matabolism. Trends Biochem. Sci. 9, 314-317.

Masters, C.J. and Crane, D. (1985) On the role of peroxisomes in the metabolism of lipids - evidence from studies on mammalian tissues in vivo. Mol. Cell. Biochem. 65, 23-35.

Masters, C.J. and Holmes, R.S. (1977) Peroxisomes: new aspects of cell physiology and biochemistry. Physiol. Rev. 57, 816-882.

Max, E.E. (1984) New twist to DNA methylation. Nature 310, 100.

Maxam, A.M. and Gilbert, W. (1977) A new method of sequencing DNA. Proc. Natl. Acad. Sci. USA 74, 560-564.

McCammon, M.T., Dowds, C.A., Orth, K., Moomaw, C.R., Slaughter, C.A. and Goodman, J.M. (1990b) Sorting of peroxisomal membrane protein PMP47 from Candida boidinii into peroxisomal membranes of Saccharomyces cerevisiae. J. Biol. Chem. 265, 20098-20105.

McCammon, M.T., Veenhuis, M., Trapp, S.B. and Goodman, J.M. (1990a) Association of glyoxylate and beta-oxidation enzymes with peroxisomes of Saccharomyces cerevisiae. J. Bacteriol. 172, 5816-5827.

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.

Mead, D.A., Skorupa, E.S. and Kempter, B. (1985) Single stranded DNA SP6 promoter plasmids for engineering mutant RNAs and proteins: synthesis of a 'stretched' preproparathyroid hormone. Nucl. Acids Res. 13, 1103-1118.

Mellor, R.B., Bowden, L. and Lord, J.M. (1978) Glycoproteins of the glyoxysomal matrix. FEBS Lett. 90, 275-278.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acids Res. 12, 7035-7056.

Messing, J., Crea, R. and Seeburg, P.H. (1981) A system for

shotgun DNA sequencing. Nucl. Acids Res. 9, 309-321.

Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P.H. (1977) Filamentous coliphage M13 as a cloning vehicle: Insertion of a Hind II fragment of the lac regulatory region in the M13 replicative form in vivo. Proc. Natl. Acad. Sci. USA 74, 3642-3646.

Messing, J. and Vieira, J. (1982) A new pair of M13 vectors for selecting either strand of a double-digest restriction fragment. Gene 19, 269-276.

Michels, P.A.M., Poliszcak, A., Osinga, K.A., Misset, O., Van Beeuman, J., Wieringa, R.K., 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.

Middlekoop, E., Strijland, A. and Tager, J.M. (1991) Does aminotriazole inhibit import of catalase into peroxisomes by retarding unfolding? FEBS Lett. 279, 79-82.

Middleton, B. (1973) The oxoacyl-coenzyme A thiolases of animal tissues. Biochem. J. 132, 717-730.

Mierendorf, R. and Jendrisak, J. (1985) Optimizing transcription performance of Riboprobe SP6 and T7 RNA polymerases. Promega Notes 1, 2-6.

Mihalik, S.J. and Rhead, W. (1989) L-pipecolic acid oxidation in the rabbit and cynomolgus monkey. Evidence for differing organellar locations and cofactor requirements in each species. J. Biol. Chem. 264, 2509-2517.

Misset, O., Van Beeumen, J., Lambeir, A.M., Van der Meer, R. and Opperdoes, F.R. (1987) Glyceraldehyde-phosphate dehydrogenase from Trypanosoma brucei. Comparison of the glycosomal and cytosolic isoenzymes. Eur. J. Biochem. 162, 501-507.

Miyazawa, S., Furuta, S., Osumi, T., Hashimoto, T. and Ui, 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., Osaza, H., Osumi, T. and Hashimoto, T. (1983) Purification and properties of carnitine octanoyltransferase and carnitine palmitoyltransferase from rat liver. J. Biochem. 94, 529-542.

Miyazawa, S., Osumi, T. and Hashimoto, T. (1980) The presence of a new 3-oxoacyl-CoA thiolase in rat liver peroxisomes. Eur. J. Biochem. 103, 589-596.

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.

Montgomery, D.L., Leung, D.W., Smith, M., Shalit, P., Faye, G. and Hall, B.D. (1980) Isolation and sequence of the gene for iso-2-cytochrome c in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA 77, 541-545.

Moreno de la Garza, M., Schultz-Borchard, U., Crabb, J.W. and Kunua, W.-H. (1985) Peroxisomal β-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.

Morgan, A.R., Lee, J.S., Pulleyblank, D.E., Murray, N.L. and Evans, D.H. (1979) Review: Ethidium fluorescence assays, Part 1. Physicochemical studies. Nucl. Acids Res. 7, 547-569.

Morris, H.R., Larsen, B.S. and Billheimer, J.T. (1988) A mass spectrometric study of the structure of sterol carrier protein SCP_2 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, 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.

Movva, N.R., Nakamura, K. and Inouye, M. (1980) Gene structure of the ompA protein, a major surface protein of Escherichia coli required for cell-cell interaction. J. Mol. Biol. 143, 317-328.

Müller, M. (1975) Biochemistry of protozoan microbodies: peroxisomes, α-glycerophosphate oxidase bodies, hydrogenosomes. Annu. Rev. Microbiol. 29, 467-483.

Müller, M. (1988) Energy metabolism of protozoa without mitochondria. Annu. Rev. Microbiol. 42, 465-488.

Murthy, M.R.N., Reid, III, T.J., Sicignano, A., Tanaka, N. and Rossman, M.G. (1981) Structure of beef liver catalase. J. Mol. Biol. 152, 465-499.

Nakamura, K., Pirtle, R.M., Pirtle, I.L., Takeishi, K. and Inouye, M. (1980) Messenger ribonucleic acid of the lipoprotein of the Escherichia coli outer membrane. J. Biol. Chem. 255, 210-216.

Nakashima, H., Yaman[illegible] M., Goto, K., Osumi, T., Hashimoto, T. and Endo, H. (1989) Isolation and characterization of the rat catalase-encoding gene. Gene 79, 270-288.

Nedergaard, J., Alexson, S. and Cannon, B. (1980) Cold adaption in the rat: increased brown fat peroxisomal β-oxidation relative to maximal mitochondrial oxidative capacity. Amer. J. Physiol. 239, C208-C216.

Neupert, W., Hartl, F.-U., Craig, E.A. and Pfanner, N. (1990) How do polypeptides cross the mitochondrial membranes? Cell 63, 447-450.

Newmeyer, D.D. and Forbes, D.J. (1988) Nuclear protein import can be separated into distinct steps in vitro: nuclear pore binding and translocation. Cell 52, 641-653.

Newmeyer, D.D. and Forbes, D.J. (1990) An N-ethylmaleimide-sensitive cytosolic factor necessary for nuclear protein import: requirement in signal-mediated binding to the nuclear pore. J. Cell Biol. 110, 547-557.

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.

Ni, W., Turley, R.B. and Trelease, R.N. (1990) Characterization of a cDNA encoding cottonseed catalase. Biochim. Biophys. Acta 1049, 219-222.

Nicolay, K., Veenhuis, M., Sulter, G.J. and Harder, W. (1987) A ^{31}P NMR study of the internal pH of yeast peroxisomes. Arch. Microbiol. 147, 37-41.

Nilsson, A., Thomassen, M.S. and Christiansen, E. (1984) Long-chain acyl-CoA levels in liver from rats fed high-fat diets: Is it of significance for an increased peroxisomal β-oxidation. Lipids 19, 187-194.

Noguchi, T. (1987) Amino acid metabolism in animal peroxisomes. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 234-243.

Noguchi, T., Fujiwara, S. and Hayashi, S. (1986) Evolution of allantoinase and allantoicase involved in urate degradation in liver peroxisomes. J. Biol. Chem. 261, 4221-4223.

Noguchi, T. and Takada, Y. (1978) Peroxisomal localization of serine:pyruvate aminotransferase in human liver. J. Biol. Chem. 253, 7598-7600.

Noguchi, T., Takada, Y. and Fujiwara, S. (1979) Degradation of uric acid to urea and glyoxyate in peroxisomes. J. Biol. Chem. 254, 5272-5275.

Noland, B.J., Arebalo, R.E., Hansbury, E. and Scallen, T.J. (1980) Purification and properties of sterol carrier protein$_2$. J. Biol. Chem. 255, 4282-4289.

Norrander, J., Kempe, T. and Messing, J. (1983) Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26, 101-106.

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, A.B., Podber, E., Ryan, J. and Noe, E. (1953) Biochemical heterogeneity of the cytoplasmic particles isolated from rat liver homogenate. J. Histochem. Cytochem. 1, 27-46.

Novikoff, A.B. and Shin, W.Y. (1964) The endoplasmic reticulum in the Golgi zone and its relation to microbodies: Golgi apparatus and autophagic vacuoles in rat liver cells. J. Microsc. (Oxf.) 3, 187-206.

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.

Nunberg, J.H., Kaufman, R.J., Chang, A.C.Y., Cohen, S.N. and Schimke, R.T. (1980) Structure and genomic organization of the mouse dihydrofolate reductase gene. Cell 19, 355-364.

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.

Ogden, J.E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Efficient expression of the Saccharomyces cerevisiae PGK gene depends on an upstream activation sequence but does not require TATA sequences. Mol. Cell. Biol. 6, 4335-4343.

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: sequence analysis and comparison with peroxisomal and cytosolic catalases from other sources. Eur. J. Biochem. 170, 105-110.

Okayama, H. and Berg, P. (1982) High-efficiency cloning of full-length cDNA. Mol. Cell. Biol. 2, 161-170.

Okazaki, K., 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. (1982) The glycosome. Ann. N.Y. Acad. Sci. 386, 542-545.

Opperdoes, F.R. (1987a) Compartmentation of carbohydrate metabolism in trypanosomes. Annu. Rev. Microbiol. 41, 127-151.

Opperdoes, F.R. (1987b) Topogenesis of glycolytic enzymes in Trypanosoma brucei. Biochem. Soc. Symp. 53, 123-129.

Opperdoes, F.R. (1988) Glycosomes may provide clues to the import of peroxisomal proteins. Trends Biochem. Sci. 13, 255-260.

Opperdoes, F.R., Baudhuin, P., Coppens, I., De Roe, C., Edwards, S.W., Weijers, P.J. and Misset, O. (1984) Purification, morphometric analysis, and characterization of the glycosomes (microbodies) of the protozoan hemoflagellate Trypanosoma brucei. J. Cell Biol. 98, 1178-1184.

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 kinase of Trypanosoma brucei. EMBO J. 4, 3811-3817.

Osmundsen, H. (1982) Peroxisomal β-oxidation of long fatty acids: effects of high fat diets. Ann. N.Y. Acad. Sci. 386, 13-27.

Osmundsen, H. and Hovik, R. (1988) β-Oxidation of poly-unsaturated fatty acids. Biochem. Soc. Trans. 16, 420-422.

Osmundsen, H., Thomassen, M.S., Hiltunen, J.K. and Berge, R.K. (1987) Physiological role of peroxisomal beta-oxidation. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 152-159.

Ostermann, J., Horwich, A.L., Neupert, W. and Hartl, F.-U. (1989) Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. Nature 341, 125-130.

Osumi, M. (1976) Possible existence of DNA in yeast microbody. J. Electron. Microsc. 25, 43-47.

Osumi, M., 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, M., Kazama, H. and Sato, S. (1978) Microbody-associated DNA in Candida tropicalis pK233 cells. FEBS Lett. 90, 309-312.

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, M., Sato, M., Sakai, T. and Suzuki, M. (1979) Fine structure of crystalloid in the yeast Kloeckera microbodies. J. Electron Microsc. 28, 295-300.

Osumi, T. and Hashimoto, T. (1984) The inducible fatty acid oxidation system in mammalian peroxisomes. Trends Biochem. Sci. 9, 317-319.

Osumi, T., Ishii, N., Hijikata, M., Kamijo, K., Ozasa, H., Furuta, 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.

Palosaari, P.M. and Hiltunen, J.K. (1990) Peroxisomal bifunctional protein from rat liver is a trifunctional enzyme possessing 2-enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase and Δ^3, Δ^2-enoyl-CoA isomerase

activities. J. Biol. Chem. 265, 2446-2449.

Panchenko, L.F. and Antonenkov, V.D. (1984) Glucose-6-phosphate dehydrogenase of rat liver peroxisomes. Experientia 40, 467-468.

Parish, R.W. (1975) The isolation and characterization of peroxisomes (microbodies) from baker's yeast, Saccharomyces cerevisiae. Arch. Microbiol. 105, 187-192.

Parker, J.M.R., Guo, D. and Hodges, R.S. (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites. Biochemistry 25, 5425-5432.

Pederson, 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.

Pederson, J.I., Kase, B.F., Prydz, K. and Björkhem, I. (1987) Liver peroxisomes and bile acid formation. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 67-77.

Pelham, H.R.B. and Jackson, R.J. (1976) An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. 67, 247-256.

Pfanner, N., Hartl, F.-U. and Neupert, W. (1988) Import of proteins into mitochondria: a multi-step process. Eur. J. Biochem. 175, 205-212.

Pfanner, N. and Neupert, W. (1990) The mitochondrial protein import apparatus. Annu. Rev. Biochem. 59, 331-353.

Pfanner, N., Rassow, J., Weinhues, U., Hergersberg, C., Söllner, T., Becker, K. and Neupert, W. (1990) Contact sites between inner and outer membranes: structure and role in protein translocation into the mitochondria. Biochim. Biophys. Acta 1018, 239-242.

Pfeifer, K., Arcangioli, B. and Guarente, L. (1987) Yeast HAP1 activator competes with the factor RC2 for binding to the upstream activation site UAS1 of the CYC1 gene. Cell 49, 9-18.

Porter, T.D. and Kasper, C.B. (1985) Coding nucleotide sequence of rat NADPH-cytochrome P-450 oxidoreductase cDNA and identification of flavin-binding domains. Proc. Natl. Acad. Sci. USA 82, 973-977.

Poulos, A., Sharp, P. and Whiting, M. (1984) Infantile Refsum's disease (phytanic acid storage disease): a variant of Zellweger's syndrome? Clin. Genet. 26, 579-586.

Prydz, K., Kase, B.F., Björkhem, I. and Pederson, J.I. (1986) Formation of chenodeoxycholic acid from 3α,7α-dihydroxy-5β-cholestanoic acid by rat liver peroxisomes. J. Lipid Res. 27, 622-628.

Pugsley, A.P. Protein Targeting. San Diego, California: Academic Press, Inc., 1989.

Purdue, P.E., Takada, Y. and Danpure, C.J. (1990) Identification of mutations associated with peroxisome-to-mitochondrion mistargeting of alanine/glyoxylate aminotransferase in primary hyperoxaluria type 1. J. Cell Biol. 111, 2341-2351.

Quan, F., Korneluk, R.G., Tropak, M.B. and Gravel, R.A. (1986) Isolation and characterization of the human catalase gene. Nucl. Acids Res. 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. and Lazarow, P.B. (1985) Cloning of a cDNA coding for peroxisomal acyl-CoA oxidase from the yeast Candida tropicalis pK233. Proc. Natl. Acad. Sci. USA 82, 3973-3977.

Rachubinski, R.A., Fujiki, 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 of closed circular duplex DNA: The closed circular DNA in HeLa cells. Proc. Natl. Acad. Sci. USA 57, 1514-1521.

Rawn, J.D. Biochemistry. Burlington, North Carolina: Neil

Patterson Publishers, 1989.

Reddy, J.K., Warren, J.R., Reddy, M.K. and Lalwani, N.D. (1982) Hepatic and renal effects of peroxisomal proliferators: biological implications. Ann. N.Y. Acad. Sci. 386, 81-110.

Reddy, M.K., Usuda, N., Reddy, M.N., Kuczmarski, E.R., Rao, M.S. and Reddy, J.K. (1987) Purification, properties and immunocytochemical localization of human liver peroxisomal enoyl-CoA hydratase/3-hydroxyacylCoA dehydrogenase. Proc. Natl. Acad. Sci. USA 84, 3214-3218.

Redinbaugh, M.G., Wadsworth, G.J. and Scandalios, J.G. (1988) Characterization of catalase transcripts and their differential expression in maize. Biochim. Biophys. Acta 951, 104-116.

Reubsaet, F.A.G., Veerkamp, J.H., Brückwilder, M.L.P., Trijbels, J.M.F. and Monnens, L.A.H. (1990) The involvement of fatty acid binding protein in peroxisomal fatty acid oxidation. FEBS Lett. 267, 229-230.

Rhodin, J. (1954) Doctoral Thesis, Karolinska Institute, Stockholm, Sweden.

Richter, K., Ammerer, G., Hartter, E. and Ruis, H. (1980) The effect of δ-aminolevulinate on catalase T-messenger RNA levels in δ-aminolevulinate synthase-defective mutants of Saccharomyces cerevisiae. J. Biol. Chem. 255, 8019-8022.

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.

Rigatuso, J.L., Legg, P.G. and Wood, R.L. (1970) Microbody formation in regenerating rat liver. J. Histochem. Cytochem. 18, 893-900.

Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113, 237-251.

Roberts, B. (1989) Nuclear location signal-mediated protein

transport. Biochim. Biophys. Acta 1008, 263-280.

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.

Rodricks, J.V. and Turnball, D. (1987) Interspecies differences in peroxisomes and peroxisome proliferation. Toxicol. Indust. Health 3, 197-212.

Roggenkamp, R., Didion, T. and Kowallik, K.V. (1989) Formation of irregular gaint peroxisomes by overproduction of the crystalloid core protein methanol oxidase in the methylotrophic yeast Hansenula polymorpha. Mol. Cell. Biol. 9, 988-994.

Roitsch, T. and Lehle, L. (1989) Requirements for efficient in vitro transcription and translation: a study using invertase as a probe. Biochim. Biophys. Acta 1009, 19-26.

Rosenkrantz, M., Alam, T., Kim. K.-S., Clark, B.J., Srere, P.A. and Guarente, L.P. (1986) Mitochrondrial and nonmitochrondrial citrate synthases in Saccharomyces cerevisiae are encoded by distinct homologous genes. Mol. Cell. Biol. 6, 4509-4515.

Rouiller, C. and Bernhard, W. (1956) "Microbodies" and the problem of mitochondrial regeneration in liver cells. J. Biophys. Biochem. Cytol. Suppl. 2, 355-359.

Russell, D.W., Smith, M., Williamson, V.M. and Young, E.T. (1983) Nucleotide sequence of the yeast alcohol dehydrogenase II gene. J. Biol. Chem. 258, 2674-2682.

Russo, J.J. and Black, V.H. (1982) Hormone-dependent changes in peroxisomal enzyme activity in guinea pig adrenal. J. Biol. Chem. 257, 3883-3889.

Saggerson, E. and Zammit, V. (1986) Carnitine metabolism and function. Biochem. Soc. Trans. 14, 671-683.

Saier Jr., M.H., Werner, P.K. and Müller, M. (1989) Insertion of proteins into bacterial membranes: Mechanism, characteristics, and comparisons with the eucaryotic process. Microbiol. Rev. 53, 333-366.

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.

Sambrook, J., Fritsch, E.F. and Maniatis, T. Molecular Cloning. A Laboratory Manual (Second Edition). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 1989.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Sanglard, D. and Loper, J.C. (1989) Characterization of the alkane-inducible cytochrome P450 (P450alk) gene from the yeast Candida tropicalis: identification of a new P450 gene family. Gene 76, 121-136.

Sansone, G. and Hamilton, J.G. (1969) Glyceryl ether, wax ester and triglyceride composition of the mouse preputial gland. Lipids 4, 435-440.

Santos, M.J., Imanaka, T., Shio, H., Small, G.M. and Lazarow, P.B. (1988) Peroxisomal membrane ghosts in Zellweger syndrome - Aberrant organelle assembly. Science 239, 1536-1538.

Scandalios, J.G., Liu, E.H. and Campeau, M.A. (1972) The effects of intragenic and intergenic complementation on catalase structure and function in maize: a molecular approach to heterosis. Arch. Biochem. Biophys. 153, 695-705.

Schaffner, W. and Weissman, C. (1973) A rapid, sensitive, and specific method for the determination of protein in dilute solution. Anal. Biochem. 56, 502-514.

Schena, M. (1989) The evolutionary conservation of eukaryotic gene transcription. Experientia 45, 972-983.

Schepers, L., Casteels, M., Vamecq, J., Parmentier, G., Van Veldhoven, P.P. and Mannaerts, G.P. (1988) β-Oxidation of the carboxyl side chain of prostaglandin E_2 in rat peroxisomes and mitochondria. J. Biol. Chem. 263, 2724-2731.

Schleif, R.F. and Wensink, P.C. Practical Methods in Molecular Biology. New York: Springer-Verlag New York, Inc., 1981.

Schnarrenberger, C. and Oeser, A. (1984) Two isoenzymes of glucosephosphate isomerase from spinach leaves and their intracellular compartmentation. Eur. J. Biochem. 45, 77-82.

Schopfer, D., Bajracharya, D., Bergfeld, R. and Falk, H. (1976) Phytochrome-mediated transformation of glyoxysomes into peroxisomes in the cotyledons of mustard (Sinapis alba L.) seedlings. Planta 133, 73-80.

Schrakamp, G., Roosenboom, C.F.P., Schutgens, R.B.H., Wanders, R.J.A., Heymans, H.S.A., Tager, J.A. and van den Bosch, H. (1985a) Alkyl dihydroxyacetone phosphate synthase in human fibroblasts and its deficiency in Zellweger syndrome. J. Lipid Res. 26, 867-873.

Schrakamp, G., Schalkwijk, C.G., Schutgens, R.B.H., Wanders, R.J.A., Tager, J.M. and van den Bosch, H. (1988) Plasmalogen biosynthesis in peroxisomal disorders: fatty alcohol versus alkylglycerol precursors. J. Lipid Res. 29, 325-334.

Schrakamp, G., Schutgens, R.B.H., Wanders, R.J.A., Heymans, H.S.A., Tager, J.A. and van den Bosch, H. (1985b) The cerebro-hepato-renal (Zellweger) syndrome. Impaired de novo biosynthesis of plasmalogens in cultured skin fibroblasts. Biochim. Biophys. Acta 833, 170-174.

Schroeder, W.A., Shelton, J.R., Shelton, J.B., Robberson, 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., Romeijn, G.J., Wanders, R.J.A., van den Bosch, H., Schrakamp, G. and Heymans, H.S.A. (1984) Deficiency of acyl CoA:dihydroxyacetone phosphate acyltransferase in patients with Zellweger (cerebro-hepato-renal) syndrome. Biochem. Biophys. Res. Commun. 120, 179-184.

Scotto, J.M., Hadchouel, M., Odievre, M., Laudat, M.-H., Saudubray, J.-M., Dulac, M., Beucler, I. and Blaune, P. (1982) Infantile phytanic acid storage disease, a possible variant of Refsum's disease: three cases, including ultrastructural studies of the liver. J. Inher. Metab. Dis. 5, 83-90.

Sharp, P.A., Sugden, B. and Sambrook, J. (1973) Detection of two restriction endonuclease activities in Haemophilus parainfluenzae using analytical agarose-ethidium bromide electrophoresis. Biochemistry 12, 3055-3063.

Sharp, P.M. and Bulmer, M. (1988) Selective differences among translation termination codons. Gene 63, 141-145.

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.

Shaw, W.V. (1975) Chloramphenicol acetyltransferase from chloramphenicol resistant bacteria. Methods Enzymol. 43, 737-755.

Shimizu, S., Yasui, K., Tani, Y. and Yamada, H. (1979) Acyl-CoA oxidase from Candida tropicalis. Biochem. Biophys. Res. Commun. 91, 108-113.

Silver, P.A. (1991) How proteins enter the nucleus. Cell 64, 489-497.

Silver, P. and Goodson, H. (1989) Nuclear protein transport. Crit. Rev. Biochem. Mol. Biol. 24, 419-435.

Singh, H., Derwas, N. and Poulos, A. (1987) Very long chain fatty acid β-oxidation by rat liver mitochondria and peroxisomes. Arch. Biochem. Biophys. 259, 382-390.

Singh, H. and Poulos, A. (1988) Distinct long chain and very long chain fatty acyl CoA synthetases in rat liver peroxisomes and microsomes. Arch. Biochem. Biophys. 266, 486-495.

Singh, I., Moser, A.E., Goldfischer, J. 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.

Skerjanc, I. (1990) Mitochondrial import: properties of precursor proteins. Biochem. Cell Biol. 68, 9-16.

Skjeldal, O.H. and Stokke, O. (1987) The subcellular localization of phytanic acid oxidase in rat liver. Biochim. Biophys. Acta 921, 38-42.

Small, G.M., Brolly, D. and Connock, M.J. (1980) Palmitoyl-CoA oxidase: Detection in several guinea pig tissues and peroxisomal localisation in mucosa of small intestine. Life Sci. 27, 1743-1751.

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. 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., 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.

Smeekens, S., Weisbeek, P. and Robinson, C. (1990) Protein transport into and within chloroplasts. Trends Biochem. Sci. 15, 73-76.

Smith, M., Leung, D.W., Gillam, S., Astell, C.R., Montgomery, D.L. and Hall, B.D. (1979) Sequence of the gene for iso-1-cytochrome c in Saccharomyces cerevisiae. Cell 16, 753-761.

Söllner, T., Griffiths, G., Pfaller, R., Pfanner, N. and 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 mitochrondrial import receptor for the ADP/ATP carrier. Cell 62, 107-115.

Sommer, J.M., Thissen, J.A., Parsons, M. and Wang, C.C. (1990) Characterization of an in vitro assay for import of 3-phosphoglycerate kinase into the glycosomes of Trypanosoma brucei. Mol. Cell. Biol. 10, 4545-4554.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Spevak, W., Hartig, A., Meindl, P. and Ruis, H. (1986) Heme control region of the catalase T gene of the yeast Saccharomyces cerevisiae. Mol. Gen. Genet. 203, 73-78.

Staack, H., Binstock, J.F. and Schulz, H. (1978)

Purification and properties of a pig heart thiolase with broad chain length specificity and comparison of thiolases from pig heart and Escherichia coli. J. Biol. Chem. 253, 1827-1831.

Stenderup, A. and Bak, A.L. (1968) Deoxyribonucleic acid base composition of some species within the genus Candida. J. Gen. Microbiol. 52, 231-236.

Stokke, O., Skrede, S., Ek, J. and Björkhem, I. (1984) Refsum's disease, adrenoleucodystrophy, and the Zellweger syndrome. Scand. J. Clin. Lab. Invest. 44, 463-464.

Stone, D. and Phillips, A.W. (1977) The amino acid sequence of dihydrofolate reductase from L1210 cells. FEBS Lett. 74, 85-87.

Struhl, K. (1987) Promoters, activator proteins, and the mechanism of transcription initiation in yeast. Cell 49, 295-297.

Struhl, K. (1989) Molecular mechanisms of transcription initiation in yeast. Annu. Rev. Biochem. 58, 1051-1077.

Stuart, R.A., Neupert, W. and Tropschug, M. (1987) Deficiency in mRNA splicing in a cytochrome c mutant of Neurospora crassa: importance of carboxy terminus for import of apocytochrome c into mitochondria. EMBO J. 6, 2131-2137.

Stuart, R.A., Nicholson, D.W., Harmey, M.A. and Neupert, W. (1988) Biogenesis of mitochondrial proteins. Biochem. Soc. Trans. 16, 701-702.

Suga, T., Horie, S. and Ishii, H. (1982) Changes in peroxisomal enzymes of rat liver under various physiological and pathological conditions. Ann. N.Y. Acad. Sci. 386, 430-432.

Sugita, Y., Tobe, T., Sakamoto, T. and Higashi, T. (1982) Immature precursor catalase in subcellular fractions of rat liver. J. Biochem. 92, 509-515.

Sutcliffe, J.G. (1978) pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. 5, 2721-2728.

Sutter, T.R., Sangard, D. and Loper, J.C. (1990) Isolation and characterization of the alkane-inducible NADPH-

cytochrome P-450 oxidoreductase gene from Candida tropicalis. J. Biol. Chem. 265, 16428-16436.

Suzuki, H., Yamada, J., Watanabe, T. and Suga, T. (1989) Compartmentation of dicarboxylic acid β-oxidation in rat liver: importance of peroxisomes in the metabolism of dicarboxylic acids. Biochim. Biophys. Acta 990, 25-30.

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., 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 POX18, 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.

Takada, Y., Kaneko, N., Esumi, H., Purdue, P.E. and Danpure, C.J. (1990) Human peroxisomal L-alanine:glyoxylate aminotransferase. Evolutionary loss of a mitochondrial targeting signal by point mutation of the initiation codon. Biochem. J. 268, 517-520.

Takada, Y., Mori, T. and Noguchi, T. (1984) The effect of vitamin B_6 deficiency on alanine:glyoxylate aminotransferase isoenzymes in rat liver. Arch. Biochem. Biophys. 229, 1-6.

Takada, Y. and Noguchi, T. (1986) Ureidoglycollate lyase, a new metalloenzyme of peroxisomal urate degradation in marine fish liver. Biochem. J. 235, 391-397.

Talwar, D. and Swaiman, K.F. (1987) Peroxisomal disorders: a review of a recently recognized group of clinical entities. Clin. Pediatr. 26, 497-504.

Tan, H., Okazaki, K., Kubota, I., Kamiryo, T. and Utiyama,

H. (1990) A novel peroxisomal nonspecific lipid-transfer protein from Candida tropicalis. Eur. J. Biochem. 190, 107-112.

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. N.Y. Acad. Sci. 386, 183-199.

Thangada, S., Alvares, K., Mangino, M., Usman, M.I., Rao, M.S. and Reddy, J.K. (1989) An in vitro demonstration of peroxisome proliferation and increase in peroxisomal β-oxidation system mRNAs in cultured rat hepatocytes treated with ciprofibrate. FEBS Lett. 250, 205-210.

Theg, S.M., Bauerle, C., Olsen, L.J., Selman, B.R. and Keegstra, K. (1989) Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplast membranes. J. Biol. Chem. 264, 6730-6736.

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 peroxisomes. Mol. Cell. Biol. 11, 510-522.

Thompson, S.L., Burrows, R., Laub, R.J. and Krisans, S.K. (1987) Cholesterol synthesis in rat liver peroxisomes. Conversion of mevalonic acid to cholesterol. J. Biol. Chem. 262, 17420-17425.

Thompson, S.L. and Krisans, S.K. (1985) Evidence for peroxisomal hydroxylase activity in rat liver. Biochem. Biophys. Res. Commun. 130, 708-716.

Thompson, S.L. and Krisans, S.K. (1990) Rat liver peroxisomes catalyze the initial step in cholesterol synthesis. The condensation of acetyl-CoA units into acetoacetyl-CoA. J. Biol. Chem. 265, 5731-5735.

Titus, D.E. and Becker, W.M. (1985) Investigation of the glyoxysome-peroxisome transition in germinating cucumber cotyledons using double-label immunoelectron microscopy. J. Cell Biol. 101, 1288-1299.

Tolbert, N.E. (1971) Microbodies - peroxisomes and

glyoxysomes. Annu. Rev. Plant Physiol. 22, 45-74.

Tolbert, N.E. (1981) Metabolic pathways in peroxisomes and glyoxysomes. Annu. Rev. Biochem. 50, 133-157.

Tolbert, N.E. (1982) Leaf peroxisomes. Ann. N.Y. Acad. Sci. 386, 254-268.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA 76, 4350-4354.

Trelease, R.N., Becker, W.M., Gruber, P.J. and Newcomb, E.H. (1971) Microbodies (glyoxysomes and peroxisomes) in cucumber cotyledons. Correlative biochemical and ultrastructural study in light- and dark-grown seedlings. Plant Physiol. 48, 461-475.

Trindade, H., Karmali, A. and Pais, M.S. (1988) One-step purification and properties of catalase from leaves of Zantedeschia aethiopica. Biochimie 70, 1759-1763.

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.

Tzagoloff, A. and Myers, A.M. (1986) Genetics of mitochondrial biogenesis. Annu. Rev. Biochem. 55, 249-285.

Ueda, M., Okada, H., Hishida, T., Teranishi, Y. and Tanaka, A. (1987) Isolation of several cDNAs encoding yeast peroxisomal enzymes. FEBS Lett. 220, 31-35.

Ueda, M., Tanaka, A. and Fukui, S. (1984) Characterization of peroxisomal and mitochondrial carnitine acetyltransferases purified from alkane-grown Candida tropicalis. Eur. J. Biochem. 138, 445-449.

Ueda, M., Yamanoi, K., Morikawa, T., Okada, H. and Tanaka, A. (1985) Peroxisomal localization of enzymes related to fatty acid β-oxidation in an n-alkane grown yeast, Candida tropicalis. Agric. Biol. Chem. 49, 1821-1828.

Usuda, N., Reddy, M.K., Hashimoto, T., Rao, M.S. and Reddy, J.K. (1988) Tissue specificity and species differences in the distribution of urate oxidase in peroxisomes. Lab. Invest. 58, 100-111.

Vamecq, J., de Hoffman, E. and Van Hoof, F. (1985) Mitochondrial and peroxisomal metabolism of glutaryl-CoA. Eur. J. Biochem. 146, 663-669.

Van Broekhoven, A., Peeters, M.C., Debeer, L.J. and Mannaerts, G.P. (1981) Subcellular distribution of coenzyme A: evidence for a separate coenzyme A pool in peroxisomes. Biochem. Biophys. Res. Commun. 100, 305-312.

Van der Krift, T.P., Leunissen, J., Teerlink, T., van Heusden, G.P.H., Verkleij, A.J. and Wirtz, K.W.A. (1985) Ultrastructural localization of a peroxisomal protein in rat liver using the specific antibody against the non-specific lipid transfer protein (sterol carrier protein 2). Biochem. Biophys. Acta 812, 387-392.

Van Veldhoven, P.P., Just, W.W. and Mannaerts, G.P. (1987) Permeability of the peroxisomal membrane to cofactors of β-oxidation: Evidence for the presence of a pore-forming protein. J. Biol. Chem. 262, 4310-4318.

Vaughn, K.C. (1989) Subperoxisomal localization of glycolate oxidase. Histochemistry 91, 99-105.

Veenhuis, M. and Harder, W. (1987) Metabolic significance and biogenesis of microbodies in yeast. In: Peroxisomes in Biology and Medicine (Fahimi, H.D. and Sies, H., eds.), Berlin, Springer-Verlag, pp. 436-458.

Veenhuis, M., Harder, W., Van Dijken, J.P. and Mayer, F. (1981) Substructure of crystalline peroxisomes in methanol-grown Hansenula polymorpha: evidence for an in vivo crystal of alcohol oxidase. Mol. Cell. Biol. 1, 949-957.

Veenhuis, M., Mateblowski, M., Kunau, W.-H. and Harder, W. (1987) Proliferation of microbodies in Saccharomyces cerevisiae. Yeast 3, 77-84.

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

82.

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.

Verner, K. and Lemire, B.D. (1989) Tight folding of a passenger protein can interfere with the targeting function of a mitochondrial presequence. EMBO J. 8, 1491-1495.

Verner, K. and Schatz, G. (1988) Protein translocation across membranes. Science 241, 1307-1313.

Verner, K. and Weber, M. (1989) Protein import into mitochondria in a homologous yeast in vitro system. J. Biol. Chem. 264, 3877-3879.

Vestweber, D. and Schatz, G. (1988) Point mutations destabilizing a precursor protein enhance its post-translational import into mitochondria. EMBO J. 7, 1147-1151.

Vieira, J. and Messing, J. (1987) Production of single-stranded plasmid DNA. Meth. Enzymol. 153, 3-11.

Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S.P., Chick, W.L. and Gilbert, W. (1978) A bacterial clone synthesizing proinsulin. Proc. Natl. Acad. Sci. USA 75, 3727-3731.

Völkl, A., Baumgart, E. and Fahimi, H.D. (1988) Localization of urate oxidase in the crystalline cores of rat liver peroxisomes by immunocytochemistry and immunoblotting. J. Histochem. Cytochem. 36, 329-336.

Völkl, A. and Lazarow, P.B. (1982) Affinity chromatography of peroxisomal proteins on lectin-sepharose columns. Ann. N.Y. Acad. Sci. 386, 504-506.

Volokita, M. and Somerville, 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. (1988) Transcending the impenetrable: how proteins come to terms with membranes. Biochim. Biophys. Acta 947, 307-333.

von Heijne, G. (1990) The signal peptide. J. Membrane Biol. 115, 195-201.

von Heijne, G. and Nishikawa, K. (1991) Chloroplast transit peptides. 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 import competence. FEBS Lett. 261, 89-92.

Wagner, P., Kunz, J., Koller, A. and Hall, M.N. (1990) Active transport of proteins into the nucleus. FEBS Lett. 275, 1-5.

Walter, P. and Blobel, G. (1983) Preparation of microsomal membranes for cotranslational protein translocation. Meth. Enzymol. 96, 84-93.

Wanders, R.J.A., Romeyn, G.J., Schutgens, R.B.H. and Tager, J.M. (1989) L-pipecolate oxidase: A distinct peroxisomal enzyme in man. Biochem. Biophys. Res. Commun. 164, 550-555.

Wanders, R.J.A., van Roermund, C.W.T., van Wijland, M.J.A., Schutgens, R.B.H., van den Bosch, H., Schram, A.W. and Tager, J.M. (1988) Direct demonstration that the deficient oxidation of very long chain fatty acids in X-linked adrenoleukodystrophy is due to an impaired ability of peroxisomes to activate very long chain fatty acids. Biochem. Biophys. Res. Commun. 153, 618-624.

Watenpaugh, K.D., Sieker, L.C. and Jensen, L.H. (1973) The binding of riboflavin-5'-phosphate in a flavoprotein: Flavodoxin at 2.0-Å resolution. Proc. Natl. Acad. Sci. USA 70, 3857-3860.

Watkins, P.A., Mihalik, S.J. and Skjeldal, O.H. (1990) Mito-chondrial oxidation of phytanic acid in human and monkey liver: Implication that Refsum's disease is not a peroxisomal disorder. Biochem. Biophys. Res. Commun. 167, 580-586.

Wattiaux, R., Wibo, M. and Baudhuin, P. Ciba Foundation

Symposium on Lysosomes. London: J. & A. Churchill, Ltd., 1963.

White, A., Handler, P., Smith, E.L., Hill, R.L. and Lehman, I.R. Principles of Biochemistry. New York: McGraw-Hill Book Company, 1978.

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.

Williams, H.E. and Wandzilak, T.R. (1989) Oxalate synthesis, transport and the hyperoxaluric syndromes. J. Urol. 141, 742-747.

Wolvetang, E.J., Wanders, R.J.A., Schutgens, R.B.H., Berden, J.A. and Tager, J.M. (1990) Properties of the ATPase activity associated with peroxisome-enriched fractions from rat liver: Comparison with mitochondrial F_1F_0-ATPase. Biochim. Biophys. Acta 1035, 6-11.

Yamada, J., Ogawa, S., Horie, S., Watanabe, F. and Suga, T. (1987) Participation of peroxisomes in the metabolism of xenobiotic acyl compounds: comparison between peroxisomal and mitochondrial β-oxidation of ω-phenyl fatty acids in rat liver. Biochim. Biophys. Acta 921, 292-301.

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.

Yamamoto, K. and Fahimi, H.D. (1987) Three-dimensional reconstruction of a peroxisomal reticulum in regenerating rat liver: evidence of interconnections between heterologous segments. J. Cell Biol. 105, 713-722.

Yamamoto, R., Kallen, C.B., Babalola, G.O., Rennert, H., Billheimer, J.T. and Straus III, J.F. (1991) Cloning and expression of a cDNA encoding human sterol carrier protein 2. Proc. Natl. Acad. Sci. USA 88, 463-467.

Yang, S.-Y. and Schulz, H. (1983) The large subunit of the fatty acid oxidation complex from Escherichia coli

is a multifunctional polypeptide. J. Biol. Chem. 258, 9780-9785.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. Gene 33, 103-119.

Yokota, S. and Fahimi, H.D. (1982) Immunoelectron microscopic localization of acyl-CoA oxidase in rat liver and kidney. Ann. N.Y. Acad. Sci. 386, 491-494.

Yokota, T., Sugisaki, H., Takanami, M. and Kaziro, Y. (1980) The nucleotide sequence of the cloned tufA gene of Escherichia coli. Gene 12, 25-31.

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.

Zaar, K., Völkl, A. and Fahimi, H.D. (1989) D-aspartate oxidase in rat, bovine and sheep kidney cortex is localized in peroxisomes. Biochem. J. 261, 233-238.

Zaheer, N. and McLean, P. (1972) Evidence for the existence and functional activity of the pentose phosphate pathway in the large particle fraction isolated from rat tissues. Biochem. Biophys. Res. Commun. 46, 167-173.

Zammit, V.A. (1986) Carnitine acyltransferases in the physiological setting: the liver. Biochem. Soc. Trans. 14, 676-679.

Zaret, K.S. and Sherman, F. (1982) DNA sequence required for efficient transcription termination in yeast. Cell 28, 563-573.

Zellweger, H. (1987) The cerebro-hepato-renal (Zellweger) syndrome and other peroxisomal disorders. Dev. Med. Child Neurol. 29, 821-829.

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.

Zoeller, R.A., Morand, O.H. and Raetz, C.R.H. (1988) A possible role for plasmalogens in protecting animal

cells against photosensitized killing. J. Biol. Chem. 263, 11590-11596.

Zubay, G. Biochemistry (Second edition). New York: Macmillan Publishing Company, 1988.

LIST OF PUBLICATIONS

Papers

Murray, W.W. and Rachubinski, R.A. (1987) 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. (1987) 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.

Gupta, R.S., Murray, W.W. and Gupta, R. (1988) Cross resistance pattern towards anticancer drugs of a human carcinoma multidrug-resistant cell line. Br. J. Cancer 58, 441-447.

Murray, W.W. and Rachubinski, R.A. (1989) Nucleotide sequence of peroxisomal catalase from the yeast Candida tropicalis pK233: Identification of an upstream BamHI site polymorphism. Nucl. Acids Res., 17, 3600.

Fernandez-Rachubinski, F., Eng, B., Murray, W.W., Blajchman, M.A. and Rachubinski, R.A. (1990) Incorporation of 7-deazadGTP during the amplification step in the polymerase chain reaction procedure improves subsequent DNA sequencing. DNA Sequence 1, 137-140.

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

Abstracts

Murray, W.W. and Gupta, R.S. (1984) Mutants of HeLa cells resistant to the adenosine analogs toyocamycin, tubercidin and 6-methylmercaptopurineriboside: A potential genetic marker for quantitative mutagenesis studies in human cells. Genetics Society of America-Genetics Society of Canada Annual Meeting.

Murray, W.W. and Rachubinski, R.A. (1986) Primary structure of a third peroxisomal fatty acyl-CoA oxidase

of the yeast Candida tropicalis. J. Cell Biol. 103, 524a (Abstract no. 1961).

Murray, W.W. and Rachubinski, R.A. (1988) The nucleotide sequence of cDNA encoding peroxisomal catalase from the yeast Candida tropicalis pK233. Advances in Gene Technology: Protein Engineering and Production, ICSU Short Reports 8, 43.

Murray, W.W., Glover, J.R. and Rachubinski, R.A. (1990) In vivo targeting of catalase from C. tropicalis to the peroxisomes of S. cerevisiae. J. Cell Biol. 111, 195a (Abstract no. 1083).