

**Analysis of *Saccharomyces cerevisiae* genetic background and
mitochondrial DNA polymerase variants on maintenance of the
mitochondrial genome.**

by

Matthew J. Young

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Microbiology

University of Manitoba

Winnipeg

Canada

Copyright © 2008 by Matthew J. Young

Thesis Abstract

The contribution of yeast strain background, specifically auxotrophic markers, to stability and fidelity of mtDNA replication was investigated. In summary, the *ade2*, *his3 Δ 200*, and *hap1* mutations have complex effects on mitochondrial functions, the severity of which appears to depend on other components in the genetic background of the strain. These results are important as many commonly used laboratory strains are related to the respiratory hampered S288c strain and are used for studies of orthologous human mutations associated with various mitochondrial diseases. These observations have added to our understanding of fungal mtDNA replication and have informed the mitochondrial community of problematic strains that need to be considered when using this model organism.

The function of the yeast mitochondrial DNA polymerase (Mip1p) carboxyl-terminal extension (CTE) was investigated both *in vivo* and *in vitro* by genetically engineering various truncations of the CTE. The respiratory competence of *mip1 Δ 175* and *mip1 Δ 205* cells, in which Mip1p lacks the C-terminal 175 and 205 residues respectively, are indistinguishable from that of wild-type. In contrast, strains harbouring Mip1p Δ 351, Mip1p Δ 279, Mip1p Δ 241, and Mip1p Δ 222 rapidly lose mtDNA. At a low frequency, *mip1 Δ 216* cells grow poorly on glycerol. Fluorescence microscopy and Southern blot analysis revealed lower levels of mtDNA in these cells, and rapid loss of mtDNA during fermentative growth. Therefore, only the polymerase-proximal segment of the Mip1p CTE is necessary for mitochondrial function.

To determine more precisely the defects associated with polymerase truncation variants, these proteins were overexpressed in yeast and used in a novel non-radioactive

mtDNA polymerase assay. The threonine-661 and alanine-661 variants, shown by others to be responsible for the increased mtDNA mutability of various laboratory yeast strains at increased temperature, were examined in combination with CTE-truncations. These experiments suggest that exonuclease function is not effected in the alanine-661 variant at 37°C whereas polymerase activity is, and this higher relative level of exonuclease activity could be a contributing factor to mtDNA instability in S288c-related strains. Lastly, isogenic CTE truncation variants all have less DNA polymerase activity than their parental wild-type. Based on these results, several possible roles for the function of the CTE in mtDNA replication are suggested.

Acknowledgements

Financial support for the work completed in this thesis was provided by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC), awarded to Dr. Deborah Court and a University of Manitoba Graduate Fellowship (UMGF), awarded to M. Young.

I would first like to express thanks to my supervisor and mentor Dr. Court for her support, guidance, and encouragement throughout my research and especially during my thesis writing. I recognize the value of Deb's advice and direction and value the experiences I have gained while working in her laboratory.

I would also like to convey my gratitude to my committee members: Dr. Georg Hausner, Dr. Ivan Oresnik and Dr. Erwin Huebner. I appreciate each one of you for your thought provoking discussions, invaluable mentorship and positive reinforcement during my studies.

Thank you to all the Microbiology graduate students and members of the Court lab over the years especially Denice Bay (Dr. Tron), Mingyi Li, Bill Summers, Enisa Zildzic, Georgia Lefas, Xi Wang, Margot Arntfield, Kristen Creek, Ryan Szajkowski, Stephen Szczerba, Remy-Martin Gratton, Sung Hoon Kim, Joanna Hessman, Anna Motnenko, Suhaila Selamat and Suman Lakhi. Also thank you to Dr. Elizabeth Worobec, and previous members of her lab especially Ayush and Kari Kumar.

Thank you to my family (Carolyn, Anthony, Mom (and Bob), Dad, Erin, Josh, DeAngelo, Pat, Tom, Gayle, Ed, and Grandma Jean) for all of your support. A very special thank you to my wife Carolyn (especially for her technical writing skills!) and my son Anthony whose support and patience have made completing my thesis possible. I love you both.

Table of Contents

Thesis Abstract	ii
Acknowledgements	iv
Table of Contents.....	v
List of Tables	ix
List of Figures	x
List of Copyrighted Material for which Permission was Obtained	xii
List of Abbreviations	xiii
CHAPTER 1 Literature review	1
1.1 Yeast as a model organism to study human mitochondrial disease	1
1.1.1 Examples of yeast orthologs of human mitochondrial disease	3
1.1.2 Human ortholog complements a yeast mutant with mitochondrial dysfunction	8
1.1.3 Genetic and chemical suppressors of orthologous disease genes identified in yeast.....	8
1.1.4 Yeast as model to study the pathophysiology of orthologous human mutations	10
1.1.5 Choosing the best yeast strain as a model of human mitochondrial disease	11
1.2 Seven families of DNA polymerases	12
1.2.1 Family A DNA polymerases.....	13
1.2.2 Family B DNA polymerases	14
1.2.3 Family C DNA polymerases	15
1.2.4 Family D DNA polymerases.....	15
1.2.5 Family X DNA polymerases.....	15
1.2.6 Family Y DNA polymerases.....	16
1.2.7 Family reverse transcriptase (RT) DNA polymerases	17
1.3 Mitochondrial DNA polymerases (γ -subfamily)	17
1.3.1 History of mitochondrial DNA polymerases, specifically yeast Mip1p	17
1.3.2 Sequence organization of pol- γ catalytic subunit.....	19
1.3.3 Molecular modeling of pol- γ catalytic subunit	40
1.4 Yeast mitochondrial DNA replication	50
1.4.1 Yeast mitochondrial DNA	50
1.4.2 Yeast mitochondrial DNA bidirectional mtDNA model.....	51
1.4.3 Yeast rolling circle mtDNA replication and vegetative segregation.....	54
1.5 The mitochondrial nucleoid.....	60
1.6 Thesis objectives.....	70

CHAPTER 2 Contributions of the S288c genetic background and common auxotrophic markers on mitochondrial DNA maintenance in <i>Saccharomyces</i>		72
2.1 Abstract		72
2.2 Introduction		73
2.3 Materials and Methods.....		86
2.3.1 General methods.....		86
2.3.2 Media, growth conditions, and genetic methods.....		86
2.3.3 Strain construction.....		88
2.3.4 Assessment of mtDNA maintenance, replication fidelity and respiratory competence		95
2.4 Results		97
2.4.1 YPH499 based strains and the <i>his3-Δ200</i> S150 strains show temperature-sensitive loss of respiratory competence.		97
2.4.2 Loss of respiratory competence is accompanied by loss of mtDNA in the YPH499 but not in the S150 background.....		103
2.4.3 Low levels of point mutations were detected in all S150 background strains, DACY1, and DACY3 while higher levels were detected in YPH499 and DACY2.....		103
2.5 Discussion		106
CHAPTER 3 The carboxyl-terminal extension on fungal mitochondrial DNA polymerases: Identification of a critical region of the enzyme from <i>Saccharomyces cerevisiae</i>		114
3.1 Abstract		115
3.2 Introduction		115
3.3 Materials and Methods.....		121
3.3.1 Strains and growth conditions.....		121
3.3.2 Strain Construction.....		121
3.3.3 Fluorescence microscopy of cells harbouring Mip1p-GFP fusions.....		126
3.3.4 Assessment of mtDNA maintenance and replication fidelity.....		127
3.3.5 Phylogenetic Analyses.....		128
3.4 Results and Discussion.....		131
3.4.1 Phylogenetic analysis of fungal mitochondrial DNA polymerases		131
3.4.2 Truncation variants of <i>Saccharomyces</i> mtDNA polymerase.....		136
3.4.3 Respiratory competence of Mip1p variants		137
3.4.4 Mutator phenotype of <i>mip1Δ216</i> cells		139
3.4.5 Loss of mitochondrial DNA in <i>mip1Δ216</i> cells.....		142
3.4.6 mtDNA maintenance by <i>mip1Δ216</i> cells in non-fermentable carbon sources		149

3.4.7 Possible functions of the CTE.....	150
CHAPTER 4 Deleterious effects of various Mip1p carboxyl-terminal truncations and the Mip1p alanine-661 variant on DNA polymerase activity.....	154
4.1 Abstract	154
4.2 Introduction	155
4.3 Materials and Methods.....	162
4.3.1 General methods.....	162
4.3.2 Media	163
4.3.3 Isolation of yeast genomic DNA.....	164
4.3.4 Strain and plasmid construction.....	165
4.3.5 Induction of recombinant Mip1p[S] and Mip1p[Σ] and variants.....	186
4.3.6 Isolation of mitochondria and mitochondrial membrane proteins	187
4.3.7 Characterization of the association of Mip1p with the mitochondrial membrane	190
4.3.8 Western blotting	191
4.3.9 Release of mitochondrial matrix proteins from intact mitochondria	192
4.3.10 DNA polymerase activity measurement of mitochondrial membrane protein fractions using the incorporation of Digoxigenin-11-dUTP.....	193
4.3.11 Nickel-nitrilotriacetic acid chromatography of Mip1p[S]Δ175-6xHis-HA and Mip1p[S]-6xHis-HA.....	194
4.3.11.1 Denaturing conditions.....	194
4.3.11.2 Non-denaturing conditions.....	195
4.4 Results.....	196
4.4.1 <i>GAL1p</i> induction of N-terminal variants of Mip1p-yEGFP expressed from a 2μ plasmid	196
4.4.2 Association of Mip1p with mitochondrial membranes	198
4.4.3 Expression and partial purification of Mip1p[S], Mip1p[Σ], and truncation variants	206
4.4.4 A non-radioactive assay for mitochondrial DNA polymerase.....	207
4.4.5 An attempt to purify Mip1p[S]Δ175 and Mip1p[S] by Nickel-NTA chromatography	225
4.5 Discussion	230
Chapter 5 Conclusions and Future work.....	249
5.1 Conclusions	249
5.2 Future Work	252
6. Appendix.....	258
6.1 Sequence of BS70.....	258
6.2 dsRed Sequence.....	258

6.3 Fusion of Mip1p N-terminus to the CTE	258
6.4 Deletion of the <i>HO</i> endonuclease gene	263
6.5 Alignment used for mitochondrial DNA polymerase phylogeny.....	268
6.6 Phylogeny of eukaryotic porins and porin motifs	278
6.6.1 Introduction.....	278
6.6.2 Results and discussion	279
6.6.2.1 Phylogenetic history of mitochondrial porins	279
6.6.2.2 Phylogenetic history of porin paralogs	287
6.6.2.3 Signature motifs for eukaryotic porins	288
References	295

List of Tables

Table 1.1 Yeast orthologs to known human mitochondrial disease genes ^a	4
Table 1.2 Nucleoid proteins and proteins implicated in mitochondrial DNA transactions	65
Table 2.1 Yeast strains used for mitochondrial studies	74
Table 2.2 Yeast strains used in this study	87
Table 2.3 Primers used in this study	89
Table 2.4 Percent respiratory competence for various laboratory strains	100
Table 2.5 Erythromycin-Resistance per 10 ⁸ cells	109
Table 3.1 Primers used in this study	124
Table 3.2 Sources of mitochondrial DNA polymerase primary sequences used for phylogenetic analyses.....	129
Table 4.1 Primers and oligonucleotides used in this study	166
Table 4.2 Plasmids utilized and constructed in this study	171
Table 4.3 Student's t-Test scores for variant membrane fractions at 10 minutes and 30°C compared to negative control.	220
Table 4.4 RLU production from three separate trials of the Mip1p[Σ] and sigma-variant polymerase assays.....	221
Table 4.5 RLU production from three separate trials of the Mip1p[S] and S288c-variant polymerase assays.....	222
Table 6.1 Plasmids utilized and constructed in this study	264
Table 6.2 Appendix Primers and oligonucleotides used in this study	265
Table 6.3. Distribution of mitochondrial porin motifs in the sequences in the current analysis and public protein sequence databases.	290

List of Figures

Figure 1.1 Fungal mtDNA polymerases.	21
Figure 1.2 Klenow fragment of <i>E. coli</i> DNA polymerase I.	41
Figure 1.3 Pentacovalent transition state in the mechanism of the 5' to 3' DNA polymerase reaction	43
Figure 1.4 Cartoon of the Klenow fragment of <i>E. coli</i> DNA polymerase I shuttling between editing and polymerization modes	46
Figure 1.5 Homology model of human DNA pol γ polymerase domain	48
Figure 1.6 The bidirectional model of the yeast mitochondrial DNA replication fork	52
Figure 1.7 The rolling circle replication model of yeast mitochondrial DNA	57
Figure 1.8 Conceptualization of the yeast mitochondrial nucleoid	63
Figure 2.1 Region of the <i>HIS3</i> locus of <i>S. cerevisiae</i>	81
Figure 2.2 Comparison of respiratory competence	83
Figure 2.3 Transplacement vector strategy	90
Figure 2.4 Respiratory competence of various yeast strains grown for four days at 37°C in YP10D	98
Figure 2.5 mtDNA maintenance in cells grown at 37°C for four days.....	104
Figure 2.6 Point mutations in mtDNA.....	107
Figure 3.1. Schematic diagram of the wild-type Mip1p (top line) and four truncation variants	118
Figure 3.2. Phylogenetic analysis of mtDNA polymerases	132
Figure 3.3. Point mutations in mtDNA.....	140
Figure 3.4. Time course analysis of mtDNA maintenance	143
Figure 3.5. Quantitative analysis of mtDNA loss and growth rate	147
Figure 4.1 CLUSTALW (http://align.genome.jp/) alignment of Mip1p sequences (1254 amino acid residues) obtained from S150 and YPH499 laboratory strains.	156
Figure 4.2 pBG1805 Gateway®-type yeast expression plasmid	176
Figure 4.3 Yeast recombination-based, 3'-end modification of pMIPS28 harbouring the <i>MIP1</i> gene	178
Figure 4.4 Targeting of Mip1p amino-terminal variants to the mitochondria	199

Figure 4.5 Release of mitochondrial matrix proteins from intact mitochondria by hypotonic lysis and homogenization.....	202
Figure 4.6 Characterization of the association of Mip1p[S] with the mitochondrial membrane	204
Figure 4.7 Western blot analysis of sucrose gradient membrane protein fractions from yeast cells expressing Mip1p[S] and Mip1p[Σ] truncation variants.....	208
Figure 4.8 Dot blot of DIG-11-dUTP incorporated into DNA by mitochondrial membrane protein fractions	211
Figure 4.9 Quantitation of recombinant mtDNA polymerase activity at 30 and 37°C ..	214
Figure 4.10 An attempt to purify Mip1p[S] Δ 175 and Mip1p[S] by Nickel-NTA chromatography	227
Figure 4.11 ClustalX (Thompson <i>et al.</i> , 1997) alignment of a portion of selected family A group polymerases	236
Figure 4.12 Structure of <i>Escherichia coli</i> DNA polymerase I Klenow fragment (PBD structure accession number 1KLN) bound to duplex DNA highlighting the homologous Mip1p threonine-661 region and DNA bound in the 3'-5' exonuclease domain	238
Figure 4.13 Structure of <i>Bacillus stearothermophilus</i> DNA polymerase I (PBD structure accession number 1NJW) bound to duplex DNA highlighting the homologous Mip1p threonine-661 region and DNA bound in the polymerase domain.....	241
Figure 4.14 Putative hydrogen bond formed between <i>B. stearothermophilus</i> DNA polymerase I residues E620 and N326.....	243
Figure 5.1 Hypothetical models rationalizing decreased polymerase activity in Mip1p variants.	256
Figure 6.1 Clustal W (http://align.genome.jp/) alignment of TRP1 from the plasmid BS70 and the S288c genome.	259
Figure 6.2 DNA sequence of the pVT100U:: <i>dsRed</i> <i>Discosoma</i> sp. dsRed open reading frame.	261
Figure 6.3 Mitochondrial DNA polymerase alignment used for phylogenetic analyses.	269
Figure 6.4. Evolutionary history of eukaryotic porin sequences – stramenopiles and plants.	280
Figure 6.5. Evolutionary history of eukaryotic porin sequences – fungi and animals. ..	282

List of Copyrighted Material for which Permission was Obtained

From Beese, L. S., Derbyshire, V. and Steitz, T. A., (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **260**: 352-5. Reprinted with permission from AAAS.

Adapted by permission from Macmillan Publishers Ltd: *Nat Struct Mol Biol* (Graziewicz, M. A., Longley, M. J., Bienstock, R. J., Zeviani, M. and Copeland, W. C., (2004) Structure-function defects of human mitochondrial DNA polymerase in autosomal dominant progressive external ophthalmoplegia. *Nat Struct Mol Biol* **11**: 770-6). Copyright 2004.

Joyce, C. M. and Steitz, T. A., (1995) Polymerase structures and function: variations on a theme? *J Bacteriol* **177** (22): 6321-9. Reproduced/amended with permission from American Society for Microbiology.

Young, M. J., Bay, D. C., Hausner, G. and Court, D. A., (2007) The evolutionary history of mitochondrial porins. *BMC Evol Biol* **7**: 31. Reproduced with permission.

Young, M. J., Theriault, S. S., Li, M. and Court, D. A., (2006) The carboxyl-terminal extension on fungal mitochondrial DNA polymerases: identification of a critical region of the enzyme from *Saccharomyces cerevisiae*. *Yeast* **23**: 101-16. Copyright John Wiley & Sons Limited. Reproduced with permission.

List of Abbreviations

Å	Angstrom
α-phosphate	alpha phosphate
γ-phosphate	gamma phosphate
<i>AAC2</i>	gene encoding the major ADP/ATP carrier of the yeast mitochondrial inner membrane
<i>ABF2</i>	gene encoding a high-mobility group protein involved in recombination of mtDNA
<i>ACO1</i>	gene encoding yeast mitochondrial aconitase
ad	autosomal dominant
adPEO	autosomal dominant progressive external ophthalmoplegia
<i>ADE2</i>	gene encoding Phosphoribosylaminoimidazole carboxylase, required for adenine biosynthesis
ADP	adenosine diphosphate
AMP	ampicillin
<i>ANT1</i>	gene encoding human adenine nucleotide transporter
ar	autosomal recessive
arPEO	autosomal recessive progressive external ophthalmoplegia
ATP	adenosine triphosphate
<i>B.</i>	<i>Bacillus stearothermophilus</i>
<i>stearothermophilus</i>	
BER	base excision repair
bla	beta-lactamase gene for ampicillin resistance in <i>E.coli</i>
bp	base pair
°C	degrees Celsius
C-terminus	carboxy terminus
Ca ²⁺	Calcium ion
<i>CCC1</i>	a yeast gene encoding a putative vacuolar Ca ²⁺ /Mn ²⁺ transporter
<i>CCE1</i>	a yeast gene that encodes a mitochondrial cruciform cutting endonuclease
Cdc9p	yeast mtDNA ligase
[¹⁴ C]dTTP	[¹⁴ C]deoxythymidine triphosphate
cm	centimetre
CMT	Charcot-Marie-Tooth disease
CMT-2A2	Charcot-Marie-Tooth disease-2A2
COO ⁻	Carboxylate ion
<i>CRD1</i>	yeast cardiolipin synthase gene
CTE	carboxyl-terminal extension
DAPI	4',6-diamidino-2-phenylindole
dATP	deoxyadenosine triphosphate
ddGTP	dideoxyguanosine triphosphate
ddH ₂ O	double distilled water
<i>DDP1</i>	human gene encoding deafness dystonia polypeptide 1
DIG labelled DNA	Digoxigenin-11-2'-deoxy-uridine-5'- triphosphate labelled DNA prepared using the Roche DIG DNA labelling Kit

DinB	<i>E. coli</i> DNA polymerase IV
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DP1	archaeal DNA polymerases small subunit
DP2	archaeal DNA polymerases large subunit
DRP	dynamain-related protein
dRP	5'-deoxyribose 5'-phosphate
dsRed	<i>Discosoma</i> red fluorescent protein
<i>E. coli</i>	<i>Escherichia coli</i>
<i>E. coli dnaQ</i>	MutD, <i>E. coli</i> DNA polymerase III epsilon subunit
EDTA	ethylenediamine-tetraacetic acid
EF-Tu	a mitochondrial elongation factor
Er-R	Erythromycin-resistance
Est2p	reverse transcriptase subunit of the yeast telomerase holoenzyme
Exo	exonuclease
FRDA	human Frataxin gene
FZO1	yeast gene encoding a mitochondrial outermembrane GTPase or dynamain-related protein
g	gram
<i>GAL1p</i>	yeast galactose inducible promoter
G (G needle)	gauge
G1 phase	Gap 1 growth phase of cell cycle
GFP	green fluorescence protein
GTPases	family of hydrolase enzymes that can bind and hydrolyze GTP
HA	hemagglutinin epitope tag
<i>HAP1</i>	yeast gene encoding a zinc finger transcription factor involved in the complex regulation of gene expression in response to levels of heme and oxygen
<i>HIS3</i>	yeast gene encoding imidazoleglycerol-phosphate dehydratase, catalyzes the sixth step in histidine biosynthesis
Hmi1p	a yeast mitochondrial-localized helicase
hrs	hours
IgG	immunoglobulin G
IROs	integrative recombinant oligonucleotides
kb	kilobase
kDa	kiloDalton
LB	Luria-Bertani medium
LEU	leucine
LRJ	Left recombination junction
M	molar
MELAS	mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes
<i>MFN1</i>	a human gene orthologous to yeast <i>FZO1</i>
<i>MFN2</i>	mitofusin 2, human orthologous of yeast <i>FZO1</i>
<i>MGM1</i>	yeast gene encoding Mitochondrial GTPase related to dynamain
<i>MHR1</i>	yeast gene encoding Mhr1p, involved in homologous

	recombination in mitochondria and in transcription regulation in nucleus
<i>MIP1</i>	gene encoding the yeast mitochondrial DNA polymerase subunit 1
<i>MIP1</i>[S]	S288c variant of <i>MIP1</i>
Mip1p [Σ]	Σ1278b variant of Mip1p
μl	microlitres
mg	milligram
ml	millilitres
mM	millimolar
Mmm1p	a trans-outermembrane protein required for mtDNA maintenance
Mg²⁺	Magnesium ion
MPP	mitochondrial processing peptidase
<i>MRM1</i>	yeast gene encoding a mitochondrial site-specific 21 S rRNA
	ribose methyltransferase
MRP	mitochondrial RNA processing
MT	Mohr-Tranebjaerg syndrome
mtDNA	mitochondrial deoxyribonucleic acid
MTS	matrix targeting sequences
mtSSB	mitochondrial single-stranded DNA binding protein
<i>MTF1</i>	yeast gene encoding a mitochondrial RNA polymerase transcription factor
ng	nanograms
Ni-NTA	Nickel-nitrilotriacetic acid
nt	nucleotide
N-terminus	amino terminus
OD₆₀₀	optical density 600 nanometres
OMP	orotidine-5'-phosphate
<i>OPA1</i>	human gene encoding optic atrophy, orthologous to gene <i>MGM1</i> in yeast
ORF	open reading frame
ori	origin of replication
p	probability
p55	human mitochondrial DNA polymerase processivity factor
pBS	pBluescript
PCR	polymerase chain reaction
PEO	progressive external ophthalmoplegia
<i>PET56</i>	also known as <i>MRM1</i> , a yeast gene encoding a ribose methyltransferase that modifies a functionally critical, conserved nucleotide in mitochondrial 21S rRNA
Pif1p	yeast mitochondrial helicase
PMSF	phenyl methylsulphonyl fluoride
Pol	polymerase
<i>polA</i>	gene that encodes E. coli DNA polymerase I
<i>polB</i>	gene that encodes E. coli DNA polymerase II
<i>polC</i>	gene that encodes E. coli DNA polymerase III alpha-subunit
PolD	Family D DNA polymerase

pol γ	human mtDNA polymerase catalytic core
<i>POLG</i>	gene that encodes the human polymerase catalytic core
Pol α	polymerase alpha, yeast Pol1p
Pol δ	polymerase delta, Pol3p
Pol ϵ	polymerase epsilon, Pol2p
Pol ζ	polymerase zeta, Rev3p
Pol η	polymerase eta, Rad30p
Pol θ	polymerase theta, human family A polymerase
Pol ι	polymerase iota, family Y DNA polymerase
Pol κ	polymerase kappa, family Y DNA polymerase
Pol λ	polymerase lambda, Pol4p
Pol μ	polymerase mu, eukaryotic members of family X DNA polymerases
Pol ν	polymerase nu, human family A polymerase
Pol σ	polymerase sigma, Trf4p
poly (rA):oligo(dT)	adenylate ribohomopolymer templates
Ppi	inorganic pyrophosphate
Put2p	soluble yeast mitochondrial matrix enzyme, Delta-1-pyrroline-5-carboxylate dehydrogenase
QIAGEN kit	QIAquick Gel Extraction kit
r strand	reverse strand
<i>rho</i>⁰	rho zero, cells without mtDNA
<i>rho</i>⁻	rho minus, cells harbouring deletions in mtDNA
<i>rho</i>⁺	rho plus, cell containing wild-type mtDNA
Rim1p	yeast single stranded mtDNA binding protein
RLU	relative light units
RNase P	ribonuclease P
RNase MRP	ribonuclease MRP
<i>RNR1</i>	yeast gene encoding ribonucleotide reductase gene
ROS	reactive oxygen species
rpm	revolutions per minute
Rpo41p	yeast mitochondrial RNA polymerase
RRJ	right recombination junction
RT	reverse transcriptase
S phase	synthesis phase of the cell cycle
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	sorbitol, ethylenediamine-tetraacetic acid, MOPS buffer
<i>SML1</i>	yeast gene encoding a ribonucleotide reductase inhibitor involved in regulating dNTP production
ssDNA	single stranded deoxyribonucleic acid
TAP	tandem affinity purification
<i>Taq</i>	<i>Thermus aquaticus</i>

TAZI	yeast gene encoding Lyso-phosphatidylcholine acyltransferase, required for normal phospholipid content of mitochondrial membranes
TBG	Tris-Cl, β -mercaptoethanol, glycerol
TCA	tricarboxylic acid cycle
TdT	terminal deoxynucleotidyl transferase
TF	transcription factor
TIM8	yeast gene encoding a mitochondrial intermembrane space protein which mediates import and insertion of mitochondrial inner membrane proteins
tRNA-Leu	Transfer RNA with anticodon for leucine
TRP	tryptophan
TSB	Tris buffered saline
TMS	two membrane-spanning autonomous mtDNA replisome
TOM	translocase of the mitochondrial outer membrane
TTC	2,3,5-triphenyl tetrazolium chloride
TUF1	yeast gene that encodes the mitochondrial elongation factor EF-Tu
UmuD'₂C	<i>E. coli</i> DNA polymerase V
URA3	yeast gene encoding Orotidine-5'-phosphate (OMP) decarboxylase required for uracil biosynthesis
YBB	yeast breaking buffer
YFH1	yeast Frataxin homolog, encodes Frataxin
yEGFP	yeast enhanced green fluorescent protein
YG-Er	media: yeast extract, glycerol, potassium phosphate, erythromycin
YP2D	media: yeast extract, peptone, 2% dextrose
YP10D	media: yeast extract, peptone, 10% dextrose
YPG	media: yeast extract, peptone, 3% glycerol

CHAPTER 1 Literature review

This review will discuss *Saccharomyces cerevisiae* as a model for studies of human disease, DNA polymerases with an emphasis on the catalytic subunit of mitochondrial DNA polymerases, and finally mitochondrial DNA replication and organization. These topics form a comprehensive background for the studies done in this thesis. These studies include effects of yeast auxotrophic markers and the carboxyl-terminal extension of the mitochondrial DNA polymerase on mtDNA maintenance.

1.1 Yeast as a model organism to study human mitochondrial disease

Much of our understanding of mitochondrial biology comes from studies using yeast. *Saccharomyces cerevisiae* is a commonly used model eukaryote for studying various aspects of mitochondrial processes such as mitochondrial fusion, division, protein import, and mitochondrial DNA (mtDNA) replication (Hoppins *et al.*, 2007, Neupert and Herrmann, 2007, Lecrenier and Foury, 2000, Schmitt and Clayton, 1993). These mitochondrial processes have been studied in detail due to the ease of yeast genetic manipulation. In particular, *S. cerevisiae* is easily transformed (Gietz and Woods, 2002) and collections having genome-wide coverage are available, which include ORF plasmid libraries, ORF deletion strains and gene fusion libraries using GFP and other tags (Gelperin *et al.*, 2005, Winzeler *et al.*, 1999, Huh *et al.*, 2003, Ghaemmaghami *et al.*, 2003). Additionally, in yeast, respiratory growth may be bypassed when a fermentable carbon source is provided. This makes yeast a powerful organism to study the processes of mtDNA transactions, whereby a mutation in a gene for mtDNA replication or maintenance in an obligate aerobe would be lethal. Wild-type yeast contain mtDNA

(*rho*⁺) and grow on nonfermentable carbon sources. Mutations may arise in cells which can result in large mtDNA deletions or point mutations (*rho*⁻), or complete loss of mtDNA (*rho*^o). These *rho*⁻ and *rho*^o cells are termed cytoplasmic *petites* as defined by Boris Ephrussi, while wild-type, *rho*⁺, are defined as *grande* (Ephrussi *et al.*, 1966).

Is *Saccharomyces* a good model for human studies? Humans are obligately aerobic and contain one hundred to ten thousand circular copies of a 16.6-kb mtDNA genome per cell (Stuart *et al.*, 2006). In contrast, yeast are facultative anaerobes and contain twenty to one hundred predominantly linear copies of a 85.8-kb mtDNA genome (Foury *et al.*, 1998). However, although the sizes of mtDNAs are different between yeast and man the coding capacity is similar as both genomes encode components of the electron transport chain and the mitochondrial translation machinery. Yeast are believed to have a different mechanism or mechanism(s) of mtDNA replication (see below), and contain a single mtDNA polymerase subunit (Mip1p), while the human polymerase possesses a catalytic core (pol γ) and a processivity factor (p55). Nonetheless, orthologs, or homologs that have resulted from speciation, of human disease-associated genes are found in yeast. Thirty to forty percent of the known human disease-associated genes share significant sequence identity with yeast genes (Foury, 1997). In fact more recently, yeast has been used to screen for candidate mitochondrial genes associated with human diseases (Steinmetz *et al.*, 2002). In this screen, a library of yeast deletion mutants was screened for respiratory deficiency and eight human orthologs were identified that were known to be associated with diseases for which a mitochondrial pathophysiology was plausible, but had not been proven. Additionally, this screen also identified twenty-one orthologs known to be implicated in mitochondrial disease. These genes, and six other

orthologs associated with human mitochondrial disease, are listed in Table 1.1. Finally, the power of using yeast to identify genetic and pharmacological suppressors of orthologous mitochondrial disease-associated mutations has also been suggested (see below, Schwimmer *et al.*, 2006). These genetic similarities shared between yeast and human mitochondria make yeast a powerful model to study mitochondrial disease.

1.1.1 Examples of yeast orthologs of human mitochondrial disease

Various human diseases are associated with genes encoding components of the machinery that carry out the mitochondrial processes described above (Hoppins *et al.*, 2007, Neupert and Herrmann, 2007, Copeland, 2008). A specific example of a gene encoding a component of the mitochondrial fusion and division machinery is human *MFN2* (mitofusin 2), which is orthologous to yeast *FZO1* (Hoppins *et al.*, 2007). *FZO1* encodes a mitochondrial outermembrane GTPase or dynamin-related protein (DRP), which contributes to mitochondrial fusion and fission dynamics (Hoppins *et al.*, 2007, Sesaki *et al.*, 2003). Mutations in this gene are associated with defects in long peripheral neurons in the neurodegenerative Charcot-Marie-Tooth (CMT) disease. CMT diseases are a genetically and clinically assorted group of hereditary sensory and motor neuropathies. Kijima *et al.* reported patients with either CMT-2A2, or unclassified CMT, had mutations located within or immediately upstream of the *MFN2* GTPase domain, critical for mitofusin 2 function, and within two coiled-coil domains, critical for mitochondrial targeting (Kijima *et al.*, 2005). These authors stated that formation of the mitochondrial network would be required to maintain functional peripheral nerve axons. *MFN1* is another human ortholog of *FZO1* that exists in the human genome.

Table 1.1 Yeast orthologs to known human mitochondrial disease genes^a

Yeast gene	Human gene	Pathophysiologic- al defect	Disease
<i>SDH1</i>	<i>SDHA</i>	RCC II subunit	Leigh syndrome; Deficiency of succinate dehydrogenase
<i>SDH2</i>	<i>SDHB</i>	RCC II subunit	Familial extraadrenal Pheochromocytoma
<i>BCS1</i>	<i>BCS1L</i>	RCC III assembly	Tubulopathy, encephalopathy, and liver failure due to CIII deficiency
<i>SHY1</i>	<i>SURF1</i>	RCC IV assembly	Leigh syndrome
<i>SCO1</i>	<i>SCO1</i>	RCC IV assembly	Hepatic failure, early-onset, and neurologic disorder due to cytochrome c oxidase deficiency
<i>SCO1</i>	<i>SCO2</i>	RCC IV assembly	Fatal infantile cardioencephalomyopathy due to Cox deficiency
<i>COX10</i>	<i>COX10</i>	RCC IV assembly	Deficiency of complex IV; Leigh syndrome
<i>LAT1</i>	<i>PDX1</i>	pyruvate DH	Pyruvate dehydrogenase E3-binding protein deficiency
<i>PDA1</i>	<i>PDHA1</i>	pyruvate DH	Pyruvate dehydrogenase deficiency; Leigh syndrome
<i>PDA1</i>	<i>BCKDH A</i>	AA catabolism (DH)	Maple syrup urine disease (MSUD)
<i>PDB1</i>	<i>BCKDH B</i>	AA catabolism (DH)	Maple syrup urine disease (MSUD)
<i>KGD2</i>	<i>DBT</i>	AA catabolism (DH)	Maple syrup urine disease (MSUD)
<i>LPD1</i>	<i>DLD</i>	pyruvate/AA/ α -KGDH	Dihydrolipoamide dehydrogenase deficiency; Leigh syndrome
<i>MIS1</i>	<i>MTHFD1</i>	AA metabolism	Deficiency of MTHFD1; associated with an increase in the genetically determined risk that a woman will bear a child with a neural tube defect
<i>GCV3</i>	<i>GCSH</i>	AA metabolism	Non-ketotic hyperglycinemia, type III (NKH3)
<i>YHM1</i>	<i>SLC25A1 5</i>	small-molecule transport	Deficiency of ornithine translocase; hyperornithinemia-hyperammonemia-homocitrullinuria (HHH syndrome)
<i>PET8</i>	<i>SLC25A1 5</i>	small-molecule transport	Deficiency of ornithine translocase; hyperornithinemia-hyperammonemia-homocitrullinuria (HHH syndrome)

Yeast gene	Human gene	Pathophysiologic defect	disease
<i>FUM1</i>	<i>FH</i>	TCA-cycle enzyme	Deficiency of fumarate hydratase; leiomyomatosis and renal cell cancer
<i>MIP1</i>	<i>POLG</i>	maintenance of mtDNA	Progressive external ophthalmoplegia with mitochondrial DNA deletions (PEO); Involved in male infertility (MI); Alpers syndrome; ataxia-neuropathy
<i>HEM14</i>	<i>PPOX</i>	heme biosynthesis	Porphyria variegata (VP)
<i>YTA12</i>	<i>SPG7</i>	ATP-dependent protease	Hereditary spastic paraplegia (HSP)
<i>CCC2</i>	<i>ATP7B</i>	copper-transport ATPase	Wilson disease (WD)
<i>TAZ1</i>	<i>G4.5</i>	Inner mitochondrial membrane lipid remodeling	Barth syndrome ^b
<i>MGM1</i>	<i>OPA1</i>	dynamin related GTPase, mitochondrial division and fusion	dominant optic atrophy disease ^c
<i>FZO1</i>	<i>MFN2</i>	dynamin related GTPase, mitochondrial division and fusion	neurodegenerative Charcot-Marie-Tooth disease-2A2 (CMT-2A2) ^d
<i>TIM8</i>	<i>DDP1</i>	Mitochondrial protein import	Mohr-Tranebjaerg syndrome ^e
<i>tL(UAA)Q</i>	<i>tRNA-Leu (UUR)</i>	mitochondrial translation	mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS syndrome) ^f
<i>YFH1</i>	<i>FRDA</i>	mitochondrial iron accumulation	Friedreich ataxia ^g

^aThe table is adapted from (Steinmetz *et al.*, 2002) with the exceptions of b, c, d, e, f and g, which were taken from (Ma *et al.*, 2004), (Sesaki *et al.*, 2003), (Hoppins *et al.*, 2007), (Paschen *et al.*, 2000), (Feuermann *et al.*, 2003), and (Chen and Kaplan, 2000) respectively. RCC, respiratory-chain complex; DH, dehydrogenase; AA, amino acid; KG, ketoglutarate.

In vivo single knock-out mutants of either *MFN1* or *MFN2* in mouse embryonic fibroblasts can maintain mitochondrial fusion while double mutants cannot (Hoppins *et al.*, 2007).

Another example of a DRP is the mitochondrial inner membrane *MGMI* of yeast, orthologous to human *OPA1* (Hoppins *et al.*, 2007). Mutations in *OPA1* are associated with defects in retinal ganglion neurons in dominant optic atrophy disease. Optic atrophy is characterized by an early childhood onset of visual impairment with a range of moderate to severe loss of visual acuity, temporal optic disc pallor, colour vision deficits, and centrocecal scotoma (a localized defect in the visual field bordered by an area of normal vision) of variable density (Votruba *et al.*, 1998). *OPA1* mutations are believed to affect the integrity of mitochondria resulting in the impairment of ATP generation and over time this impairment could affect the survival of retinal ganglion cells (Alexander *et al.*, 2000).

The yeast ortholog of human *DDP1*, encoding deafness dystonia polypeptide 1, is *TIM8*, a mitochondrial intermembrane space protein, which mediates import and insertion of mitochondrial inner membrane proteins (Neupert and Herrmann, 2007). In humans, loss-of-function mutations of *DDP1* are associated with Mohr-Tranebjaerg (MT) syndrome, a progressive neurodegenerative disorder associated with deafness, visual disability leading to cortical blindness, dystonia, fractures and mental deficiency (Tranebjaerg *et al.*, 1995, Neupert and Herrmann, 2007). Therefore, MT syndrome is likely caused by a defect in mitochondrial protein import.

The human mtDNA polymerase (pol γ , encoded by *POLG*) is orthologous to yeast mtDNA polymerase, Mip1p (encoded by *MIPI*). This enzyme bears the sole

responsibility of replicating mtDNA; therefore, it is probably not surprising that mutations in this gene are associated with human mitochondrial disease. Nearly ninety pathogenic mutations have been found in *POLG* and are associated with diseases such as progressive external ophthalmoplegia (PEO), Alpers syndrome, ataxia-neuropathy, and male infertility (Copeland, 2008, and the Human Polymerase Gamma Mutation Database, <http://dir-apps.niehs.nih.gov/polg>).

PEO is a disease that typically has a later onset in patients between 18 and 40 years of age. Patients exhibit mtDNA depletion and/or accumulation of mtDNA mutations and deletions (Copeland, 2008). Symptoms include ptosis (drooping of upper eyelids) and progressive weakening of external eye muscles, resulting in ophthalmoparesis (paralysis of the extraocular muscles, which are responsible for eye movement). Genetically, PEO may be sporadic, autosomal dominant (ad) or autosomal recessive (ar). With one exception, the mutations associated with adPEO are all located in the region of *POLG* encoding the polymerase domain. Parkinsonism has also been documented to co-occur with adPEO. Most mutations are arPEO forms, and often patients with PEO have two different mutant alleles of *POLG* (compound heterozygote).

Alpers syndrome is associated with more than 25 mutations in the *POLG* gene (Copeland, 2008). This autosomal recessive disease is rare, but nonetheless devastating, having an early onset, whereby patients typically die within the first decade of life. Symptoms of the disease include mtDNA depletion, deafness, progressive cerebral degeneration, spastic quadriparesis (weakness of all four limbs), blindness, and liver failure.

Ataxia-neuropathy syndrome is an autosomal recessive disease associated with nonsense and missense mutations in *POLG*. Symptoms include mtDNA depletion, damage to nerves of the peripheral nervous system, speech disorder, cognitive impairment, involuntary movements, psychiatric symptoms, involuntary twitching of muscles, and seizures (Copeland, 2008).

Alterations of a polymorphic CAG repeat of 10 codons at the 3'-end of the *POLG* gene has been reported to be associated with loss of sperm quality and contribute to 5-10% of the male infertility cases in European populations (reviewed in Copeland, 2008). However, other independent studies have not been able to reproduce this association (Copeland, 2008).

1.1.2 Human ortholog complements a yeast mutant with mitochondrial dysfunction

S. cerevisiae has been employed as a model to quickly study and characterize a number of mitochondrial disease genes *in vivo*. For example Ma and coworkers showed by complementation analysis that the human *TAZ* gene complements the mitochondrial dysfunction of a yeast mutant harbouring a deletion of *TAZI* (Ma *et al.*, 2004 and Table 1.1). Mutations in the human *TAZ* gene are associated with Barth syndrome. Symptoms of this disease include abnormal mitochondria, cardiomyopathy (weakening of the heart muscle), skeletal myopathy, and low amounts of neutrophil granulocytes (type of white blood cell) (Barth *et al.*, 1983).

1.1.3 Genetic and chemical suppressors of orthologous disease genes identified in yeast

In a study by Feuermann *et al.*, the equivalent human mtDNA tRNA-Leu mutations, associated with the neurodegenerative disease MELAS were constructed in the orthologous yeast mtDNA gene (Feuermann *et al.*, 2003). The resulting mutants

showed respiratory defects, accumulated mtDNA deletions at a high frequency, and displayed abnormal mitochondrial morphology. Interestingly, these mitochondrial dysfunctions could be genetically suppressed by overexpression of EF-Tu (encoded by *TUFI*), a mitochondrial elongation factor. MELAS is characterized by mitochondrial myopathy (muscle disease), encephalopathy (disease of the brain), lactic acidosis (build up of lactic acid in the body), and stroke-like episodes (Feuermann *et al.*, 2003).

In a similar screen, *CCCI*, encoding a putative vacuolar $\text{Ca}^{2+}/\text{Mn}^{2+}$ transporter, was found to suppress mitochondrial dysfunction in a yeast model of Friedreich's Ataxia (a *YFHI*, yeast Frataxin homolog, deletion mutant) by limiting the accumulation of iron in the mitochondria (Chen and Kaplan, 2000). *YFHI* encodes Frataxin, which regulates iron accumulation in the mitochondria. Friedreich's ataxia results from a triplet expansion in the first intron of the human Frataxin gene, *FRDA*, and this disorder is lethal affecting the nervous system and heart. The resulting mutant *FRDA* mRNAs are decreased and patients exhibit increased iron deposits in heart biopsies and in mitochondria of fibroblasts (Chen and Kaplan, 2000). In the absence of Yfh1p in yeast, iron accumulates in the mitochondria and mtDNA mutations accumulate.

Baruffini *et al.* studied two orthologous *POLG* mutations associated with adPEO and arPEO, Y955C (Y757C in yeast) and G268A (G224A in yeast) respectively (Baruffini *et al.*, 2006). The *mip1*^{Y757C} haploid strain completely lost mtDNA (*rho*^o), while 79% of the *mip1*^{G224A} haploid cells maintained respiratory competence. Similar to the phenotypes in humans, the diploid strain *MIP1/mip1*^{Y757C} showed a dominant negative phenotype in comparison to *MIP1/MIP1* (68% and 99% of cells maintained respiratory competence, respectively), while the *MIP1/mip1*^{G224A} diploid displayed a recessive

phenotype (> 95% respiratory competence). Genetic suppression of the *MIP1/mip1*^{Y757C} diploid and *mip1*^{G224A} haploid *petite* phenotypes, was demonstrated by increasing the pool of dNTPs via ribonucleotide reductase (*RNR1*) overexpression and by deletion of its inhibitor (*SML1*) from the yeast genome. Chemical suppression of these phenotypes was also demonstrated using the anti-oxidant agent dihydrolipoic acid (Baruffini *et al.*, 2006). These authors therefore proposed that patients suffering from PEO may benefit from the development of antioxidant therapy.

1.1.4 Yeast as model to study the pathophysiology of orthologous human mutations

The last examples of chemical and genetic suppression of orthologous *POLG* and tRNA-Leu mutations, validates yeast as a model to characterize the cellular consequences of human mutations of mitochondrial disease. Interestingly, mutations in other components of the human mtDNA replication machinery and a gene encoding an ATP/ADP transporter are also associated with PEO. These include the genes encoding the pol γ accessory subunit (*POLG2*), the mitochondrial DNA helicase (*TWINKLE*) and the adenine nucleotide transporter (*ANT1*) (Copeland, 2008). In a study by Fontanesi *et al.*, *S. cerevisiae* was employed to characterize mutations in *ANT1* associated with adPEO (Fontanesi *et al.*, 2004). In haploid yeast the equivalent human mutations in the yeast ortholog *AAC2*, caused growth defects on non-fermentable carbon sources, reductions in cellular respiration, mitochondrial cytochromes, and cytochrome c oxidase activity, and led to defects in ADP versus ATP transport compared to wild-type *AAC2*. This study also examined the heterozygous conditions of the wild-type *AAC2* with adPEO mutant variants of *AAC2*. The control *AAC2* strains harboured either *AAC2* or “humanized wild-type variants” on a centromeric plasmid, and were referred to as homoallelic

(homozygous), while the heteroallelic (heterozygous) *AAC2* strains harboured the equivalent *ANT1* adPEO mutations on the plasmid borne *AAC2*. Dominant traits of reduced cytochrome content and increased mtDNA instability were observed in the heteroallelic strains. Other studies of PEO- and Alpers-*POLG*-associated mutations were analyzed in yeast models by Stuart *et al.* and Baruffini *et al.* (Stuart *et al.*, 2006, Baruffini *et al.*, 2007a).

In summary, analyses in yeast are now common place to quickly assess the molecular consequences of mitochondrial disease-associated mutations *in vivo*. These observations are potentially relevant to understanding the pathologies observed in human patients afflicted with these disorders (Stuart *et al.*, 2006). However, the model organism itself must be well-characterized for these analyses and comparisons among them to be valid.

1.1.5 Choosing the best yeast strain as a model of human mitochondrial disease

Recent work from several labs has revealed critical differences in common laboratory strains. For example, Baruffini *et al.* reported that a single nucleotide polymorphism in the yeast *MIP1* changes an evolutionarily conserved threonine at position 661 (T661) to an alanine (A661) in *S. cerevisiae* strains related to S288c (W303-1B, BY4741, BY4742, and BY4743) (Baruffini *et al.*, 2007b). They showed that an alanine residue at position 661 is responsible for an increase in point mutations and deletions within the mitochondrial genome. The S288c strain has also been documented as being inappropriate for mitochondrial studies due to the mutated *HAPI* locus (Sherman, 2002). *HAPI* encodes a transcription factor that senses cellular heme status and increases expression of aerobic genes in response to oxygen (Ter Linde and

Steensma, 2002). Also, the contribution of auxotrophic marker genes to the interpretation of respiratory phenotypes is exemplified by studies utilizing strains bearing the *his3-Δ200* allele. This allele harbours a 1036-base pair deletion that removes the entire *HIS3* gene and a bidirectional promoter required for transcription of *MRM1* (Chapter 2), which encodes a mitochondrial site-specific ribose methyltransferase (Struhl, 1985, Sirum-Connolly and Mason, 1993, Rieger *et al.*, 1997). As a result, *his3-Δ200* strains display respiratory deficiencies (Zhang *et al.*, 2003, Sirum-Connolly and Mason, 1993, Rieger *et al.*, 1997). The vast majority of the examples mentioned above of complementing, genetically and chemically suppressing, and studying the pathophysiology of human orthologs in yeast employed S288c-, W303-1B- or W303-related strains (Chen and Kaplan, 2000 and Table 2.1). Due to these observations, in Chapter 2 we evaluate combinations of different auxotrophic markers in two separate genetic background strains. In addition, the two Mip1p variants harbouring either the A661 or T661 residue are also examined in Chapter 4.

1.2 Seven families of DNA polymerases

The remainder of the work carried out in this thesis focused on a novel region of the yeast Mip1p enzyme, the carboxyl-terminal extension. Therefore, a brief introduction to DNA polymerases will be presented, followed by details of mtDNA polymerases, mtDNA replication, and mtDNA organization into nucleoids.

All organisms require enzymes to replicate the genetic information stored in their genome and to pass this information on to their offspring. DNA polymerases are found in all domains of life (Eukarya, Eubacteria, and Archaea) as well as are encoded by viruses, phages, and plasmids. Currently, sixteen DNA polymerases have been identified in

humans, while five and nine have been recognized in *E. coli* and *S. cerevisiae* respectively (Bebenek and Kunkel, 2004). The evolution of the DNA replication apparatus is complex and probably involved numerous gene duplications, gene losses, and lateral gene transfers between viruses, plasmids and their hosts (Filee *et al.*, 2002). Due to this complex evolutionary history, and the wide-spread distribution of DNA polymerases, a unified classification of DNA polymerases was established in 1991 (Ito and Braithwaite, 1991). This classification system organized over forty DNA polymerases known at that time, into four families, Family A, B, C and X. Since this analysis, the classification has expanded to a total of seven families with the newly added families D, Y, and reverse transcriptase (Filee *et al.*, 2002, Kempeneers *et al.*, 2005).

1.2.1 Family A DNA polymerases

Family A DNA polymerases are DNA-dependent DNA polymerases, which share amino acid sequence homology with the first discovered DNA polymerase, *E. coli* DNA polymerase I encoded by the *polA* gene (Lehman *et al.*, 1958). These enzymes include subfamilies of bacterial, bacteriophage, eukaryotic and mitochondrial DNA polymerases (gamma (γ) subfamily, although this group was initially classified as RNA-dependent DNA polymerases (Graziewicz *et al.*, 2006)). Also grouped into this family were a few 5' to 3' exonucleases (ex. T5 and T7 exonucleases) that share sequence homology with *E. coli* DNA polymerase I (Braithwaite and Ito, 1993). Representatives of this family such as the Klenow fragment of *E. coli* DNA polymerase I, *Taq* DNA polymerase, *Bacillus stearothermophilus* DNA polymerase I, and T7 DNA polymerase have been crystallized with various DNA substrates and have contributed to our overall understanding of the DNA synthesis reaction (Beese *et al.*, 1993, Eom *et al.*, 1996, Kiefer *et al.*, 1998, Doublet

et al., 1998). Two newly identified human family A polymerases, Pol θ (theta) and Pol ν (nu), are predicted to be involved in interstrand cross-link repair based on homology to *Drosophila* mus308 (Bebenek and Kunkel, 2004). Family A polymerases are resistant to the drug aphidicolin but are very sensitive to dideoxynucleotide inhibitors.

1.2.2 Family B DNA polymerases

Family B DNA polymerases are DNA-dependent DNA polymerases, which share amino acid sequence homology with *E. coli* DNA polymerase II encoded by the *polB* gene. The family B enzymes comprise gamma proteobacterial, bacteriophage, archaeobacterial, eukaryotic cell, eukaryotic viral, and eukaryotic linear DNA plasmid-encoded DNA polymerases (Braithwaite and Ito, 1993). Two categories of polymerases can be distinguished in this family due to their DNA priming properties. Category one includes RNA/DNA-priming polymerases of Archaea, Eukaryota, Enterobacteria and various eukaryotic and prokaryotic viruses. Examples include the three eukaryotic polymerases α (alpha), δ (delta), and ϵ (epsilon), which are involved in the bulk of the nuclear DNA replication. In addition, the eukaryotic polymerase ζ (zeta), which performs translesional DNA replication/bypasses DNA lesions, is also part of this category (McCulloch and Kunkel, 2008). Category two includes protein-priming type polymerases encoded by linear plasmids of mitochondria (fungi and two plants), bacteriophages, and of the eukaryotic Adenovirus (Filee *et al.*, 2002). Most, if not all, Family B DNA polymerases are sensitive to aphidicolin and relatively resistant to dideoxynucleotide inhibitors.

1.2.3 Family C DNA polymerases

Family C DNA polymerases are bacterial DNA-dependent DNA polymerases representative of the primary bacterial-chromosome replicative DNA polymerases and *E. coli dnaQ* (MutD, *E. coli* DNA polymerase III ϵ subunit). Members of this family share amino acid sequence homology with the *E. coli* DNA polymerase III α -subunit encoded by the *polC* gene. Enzymes of this family also include cryptic phages as well as plasmid-encoded DNA polymerases (Braithwaite and Ito, 1993, Filee *et al.*, 2002). No eukaryotic or archaeal representatives have been discovered yet for this family.

1.2.4 Family D DNA polymerases

Family D DNA polymerases (PolD) are archaeal DNA-dependent DNA polymerases, which share amino acid sequence homology with *Pyrococcus furiosus* DNA polymerase II of the Euryarchaeota subdomain of the Archaea (Shen *et al.*, 2004). PolDs are composed of two subunits: a large DP2 catalytic subunit and a small DP1 subunit, which has low but significant homology to the eukaryotic DNA polymerase δ .

1.2.5 Family X DNA polymerases

Family X DNA polymerases are DNA-dependent DNA polymerases, which share amino acid sequence homology with eukaryotic DNA polymerase β (beta) (Ito and Braithwaite, 1991). Mammalian enzymes of this family are relatively small and inaccurate enzymes involved in DNA repair processes such as base excision repair (BER) and repair of double-strand breaks (Moon *et al.*, 2007). Within eukaryotes two types of family X paralogs seem to be present in a wide variety of metazoans 1) DNA polymerase β and 2) terminal deoxynucleotidyl transferase (TdT) (Filee *et al.*, 2002). TdT is a template-independent polymerase that catalyzes the addition of random

dNTPs to the 3'-hydroxyl terminus of DNA. This process occurs at the junctions of immunoglobulin genes during V(D)J recombination (N-regions), which results in increased immunological diversity. The family X polymerases have members in the Archaea, Eukaryota, and Eubacteria. More recently discovered eukaryotic members include Pol λ (lambda), Pol μ (mu), and Pol σ (sigma). Both Pol β and λ have DNA polymerase activity and 5'-deoxyribose 5'-phosphate (dRP) lyase activity required for BER. Pols λ , μ and TdT all have a BRCT domain (similar to BRCA-1 C-terminal protein-protein interaction domain) associated with nonhomologous end-joining of double-stranded DNA breaks induced by DNA damage and/or during V(D)J recombination (Moon *et al.*, 2007).

1.2.6 Family Y DNA polymerases

Family Y DNA polymerases are DNA-dependent DNA polymerases, which share amino acid sequence homology with *E. coli* DNA polymerase IV (DinB) and DNA polymerase V (UmuD'₂C). These family Y enzymes are found in the three domains, Archaea, Eukaryota, and Eubacteria as well as in some bacteriophages and plasmids (Filee *et al.*, 2002). The Y family polymerases are characterized by their ability to bypass the stalling of a replication fork during DNA replication. This ability to synthesize DNA past a DNA lesion is referred to as translesion synthesis (TLS, Jarosz *et al.*, 2007). Eukaryotic polymerases include η (eta), κ (kappa) and ι (iota). Rev1, a template-dependent deoxycytidyl transferase is also part of the Y family (Bebenek and Kunkel, 2004).

1.2.7 Family reverse transcriptase (RT) DNA polymerases

Family RT DNA polymerases are RNA-dependent DNA polymerases, which share amino acid sequence homology with retroviral reverse transcriptases or eukaryotic telomerases (Kempeneers *et al.*, 2005, Nakamura *et al.*, 1997). Because these enzymes are RNA-dependent, they are typically analyzed separately from the DNA-dependent enzymes.

In summary, the known DNA polymerases have been organized into seven families: A, B, C, D, X, Y, and RT. Nine DNA-dependent polymerase have been identified in *Saccharomyces* and are representative of family A (Mip1p (γ)), family B (Pol1p (α), Pol2p (ϵ), Pol3p (δ) and Rev3p(ζ)), family X (Pol4p (λ) and Trf4p (σ)), and family Y (Rad30p (η) and Rev1p). In addition, a family RT representative is also present in yeast, Est2p (reverse transcriptase subunit of the telomerase holoenzyme (Lingner *et al.*, 1997)).

1.3 Mitochondrial DNA polymerases (γ -subfamily)

1.3.1 History of mitochondrial DNA polymerases, specifically yeast Mip1p

MtDNA polymerases make up the γ -subfamily of family A DNA polymerases. For mammals it has been estimated that mtDNA polymerase comprises one to five percent of the total cellular DNA polymerase activity and mtDNA accounts for one percent of the total cellular DNA (Graziewicz *et al.*, 2006). The low natural occurrence of these molecules is fascinating as mtDNA maintenance is essential for the life of obligate aerobes and mutations in the *POLG* gene are associated with various human genetic diseases. The first cloned and sequenced mtDNA polymerase was isolated from *S. cerevisiae*, *MIP1* (**M**itochondrial DNA **P**olymerase, Foury, 1989). Foury noted sequence

similarities between Mip1p and eukaryotic as well as viral DNA polymerases (Foury, 1989). This study also showed that the carboxyl-terminal region of Mip1p harbours a motif, which has sequence similarity with reverse transcriptases and Mip1p-containing fractions had reverse transcriptase activity i.e. the ability to copy adenylate ribohomopolymer templates (poly (rA)·oligo(dT)) as was demonstrated for higher eukaryotic mitochondrial DNA polymerases in the 1970s (reviewed by Kaguni, 2004).

Subsequently, the coding sequences for human, mouse, chicken, *Xenopus*, *Drosophila*, *Schizosaccharomyces pombe*, and *Pychia pastoris* pol γ were determined (Graziewicz *et al.*, 2006). In addition, a database search in 2006 found twenty additional fungal homologs of *MIP1* (Young *et al.*, 2006). The animal polymerase gammas possess two distinct subunits 1) a catalytic core, containing DNA polymerase and 3'-5' proofreading exonuclease activities and 2) an accessory subunit that enhances catalytic activity and serves as a processivity factor (Fan *et al.*, 2006). Processivity is defined as the extent of DNA polymerization in a single binding event (Foury and Vanderstraeten, 1992). Carrodeguas *et al.* solved the three-dimensional structure of the mouse pol γ accessory subunit, demonstrating that the subunit crystallized as a dimer (Carrodeguas *et al.*, 2001). Accessory subunits share striking sequence and structural similarities with class IIa aminoacyl-tRNA synthetases (Fan *et al.*, 1999); however, residues critical for ATP binding and anticodon recognition in tRNA synthetases are not present in the accessory subunit (Carrodeguas *et al.*, 2001). The relatedness of accessory subunits to tRNA synthetases is apparent in surface loops involved in tRNA recognition by tRNA synthetases that also appear to be important for the interaction of the mouse accessory subunit with ssDNA (Carrodeguas *et al.*, 2001). To date no homolog of an animal

accessory subunit has been found in the yeast genome and the Mip1p catalytic subunit has been shown to function as a processive monomer (Foury and Vanderstraeten, 1992). In addition, a portion of the carboxyl-terminal extension (CTE) of Mip1p, which is conserved in ten Saccharomycetales and located proximal to the polymerase domain, is essential for mtDNA maintenance (Young *et al.*, 2006, Chapters 3 and 4, Figure 1.1).

As part of this work examined the fungal-specific CTE of Mip1p (Chapters 3 and 4) as well as the Mip1p[S] and Mip1p[Σ] variants that exist in many commonly used laboratory strains, this review will focus on the sequence organization and molecular structural model of the polymerase catalytic subunit only.

1.3.2 Sequence organization of pol- γ catalytic subunit

The Mip1p polypeptide harbours three conserved 3'-5' exonuclease motifs and three conserved DNA polymerase motifs shared by family A DNA polymerases (Figure 1.1, Ito and Braithwaite, 1990). Interestingly, Ito and Braithwaite suggested that this finding is consistent with the endosymbiotic theory. Briefly, this theory states that mitochondria evolved from endocytosis of an ancient bacterium by an ancient anaerobic cell and a symbiotic relationship was developed. Theoretically, the aerobic bacterium could handle the toxic oxygen for the anaerobic cell and the cell would protect and ingest food for the bacterium (Margulis, 1968). Support for this theory comes from the observation that yeast mitochondrial RNA and DNA polymerases are homologous to T-odd (T3 and T7) bacteriophage RNA and DNA polymerases (Ito and Braithwaite, 1990, Filee *et al.*, 2002). Recently, it has been suggested that it is possible that some of the genes encoding components of the replication machinery may have been acquired as part of protomitochondrial genome, in the form of integrated phage genes, which were then

transferred to the nucleus. Although there is no evidence that T-odd phage genes laterally transferred into mtDNA this theory is supported by the finding that some phage-like genes exist in alpha-proteobacterial genomes (reviewed in Shutt and Gray, 2006).

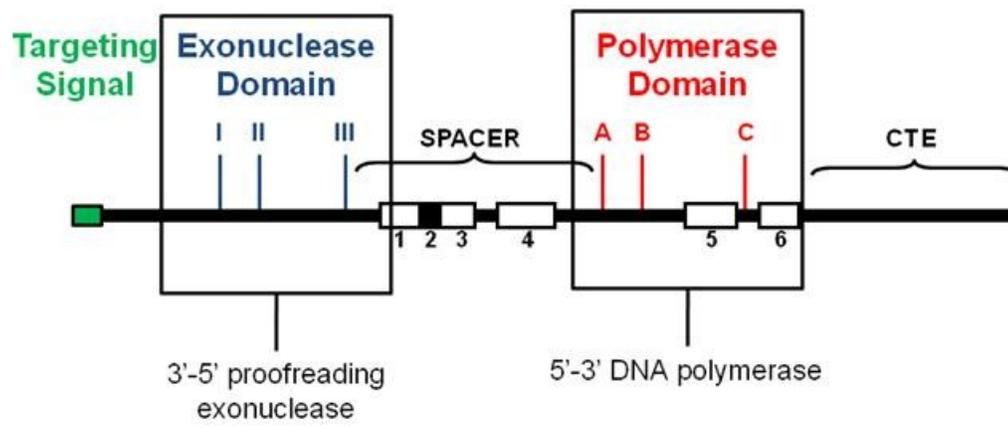
Alternatively, perhaps the phage genes were acquired by the nucleus from a replicating phage genome in an infected endosymbiotic bacterium. Whatever the case, three organisms should probably be considered for the establishment of mitochondrial endosymbiosis: host, bacterium, and phage (Shutt and Gray, 2006).

In addition to the exonuclease and polymerase motifs shared among the family A polymerases, six conserved elements ($\gamma 1$ to $\gamma 6$) among the γ -polymerases have been identified (Kaguni, 2004, Chapter 3 and Figure 1.1). Most of the gamma-specific elements ($\gamma 1$ to $\gamma 4$) are found within a γ -polymerase spacer region, which is almost twice as long as the spacer region in other family A polymerases (Graziewicz *et al.*, 2006). It has been suggested that the spacer region of mtDNA polymerases may be involved in subunit interaction (for animal polymerases), DNA binding, and/or functional coupling between activities of the polymerase and exonuclease domains. Luo and Kaguni have shown that mutations in three out of four of these elements ($\gamma 1$, $\gamma 3$ and $\gamma 4$) alter polymerase activity, processivity, and/or DNA binding affinity of purified mutant holoenzymes from *Drosophila* (Luo and Kaguni, 2005). Deletion mutants lacking $\gamma 1$ and $\gamma 4$ had weak interactions with the accessory subunit while all three deletion mutants had low DNA polymerase activity. Various site-directed missense mutations in all three elements caused reduced processivity. Additionally, human mutations in the spacer region have been linked to various diseases including PEO, Alpers, and ataxia-neuropathy-like syndromes (Luo and Kaguni, 2005).

Figure 1.1 Fungal mtDNA polymerases. **A.** Schematic diagram of Mip1p. I, II, and III indicate the location of three conserved exonuclease motifs that define the 3'-5' exonuclease domain, while A, B, and C indicate the location of the three conserved polymerase motifs that define the polymerase domain. Boxes numbered 1-6 indicate the positions of the mtDNA polymerase γ -specific sequences. The γ_2 sequence is shaded black to delineate it from γ_1 and γ_3 . "CTE" indicates the location of the carboxyl-terminal extension unique to fungal polymerases. Targeting signal refers to the amino-terminal region of the protein required for proper targeting and import into the mitochondrion. **B.** Conservation among fungal mtDNA polymerases. An alignment of full length mtDNA polymerases was generated using PRALINE (Simossis and Heringa, 2003), and conservation scores are presented. 'Consistency' refers to how consistently an amino acid is found at a given position in the final multiple alignment and in pre-processed alignment blocks (Heringa, 1999). In the 10-point consistency scale, zero (blue) indicates the least conserved alignment position, while * (red) indicates the most conserved position. Gamma specific sequences ($\gamma_1 - \gamma_6$) and polymerase motifs A, B, and C (Pol A, Pol B, and Pol C) are indicated by boxed residues on the alignment. Sc, *Saccharomyces cerevisiae* (S288c); Spa, *Saccharomyces paradoxus*; Sba, *Saccharomyces bayanus*; Cgl, *Candida glabrata*; Kwa, *Kluyveromyces waltii*; Debaryomyces_ha, *Debaryomyces hansanii*; Ncr, *Neurospora crassa*; Ego, *Eremothecium gossypii*; Klactis, *Kluyveromyces lactis*; Gze, *Gibberella zeae*; Pan, *Podospora anserina*; Mgr, *Magnaporthe grisea*; Hca, *Histoplasma capsulatum*; Calb, *Candida albicans*; Sca, *Saccharomyces castellii*; Ani, *Aspergillus nidulans*; Yarrowia_lipoly, *Yarrowia lipolytica*; Ppa, *Pichia pastoris*; aspfum, *Aspergillus fumigatus*; Sku, *Saccharomyces*

kudriavzevii. Accession numbers for all of these mtDNA polymerase sequences can be found in Table 3.2.

A.



B.

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

	10	20	30	40	50
Sc		MDYERTV	LKKRSRWGLY	VVV	
Spa					
Sba					
Cgl		MLCALR	ARAQVGVRLP	LPL	
Kwa					
Debaryomyces_ha		MI	LHKQAFSYSK	NVV	
Ncr		MLTPVR	CRTVPNATVA	TAA	
Ego			MIKC	GLT	
Klactis			MRY	SI	
Gze			MEDAN	AKFII	
Pan		MLLPYR	RVAAGRDAPV	ASFIR	
Mgr	MHGLGPD T T	SQTNI I FAFT	TCCQRQAVAF	VRMLKPLRGC	I ANGNSTAFV
Hca	MLLSSCAT	SRGVLRISSK	PLRLPVLTRT	TRHQR	L
Calb		MS	NRLLLRGIT	RSHRSFVTSL	RRR
Sca					
Ani	MLKGHCTP	YSRGVPRFYA	QKTRIRSF AQ	SQA	
Yarrowia_lipoly					
Ppa					
aspfum					
Sku					
Consistency	0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 1 1 1 0	1 1 1 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0

	60	70	80	90	100
Sc	EQRGTS M	TKLMVRSECM	LRMVRRRPLR	VQFCARWFST	KKNTAEAPRI
Spa	M	TKLMVTSHCM	LRMVRRRPLR	VQFYTRWYST	KKNNAEAPRI
Sba				VRWCRRYAT	RKSSAEAPRI
Cgl	PLLLRVP	NMAL		PV	TVRWAHVGE G
Kwa			MRLQRAT	QCARVWRKVR	CYSIQEKPRI
Debaryomyces_ha	RIWGRSL	K		SLAKT	RDQFKEAPRI
Ncr	RVLRRAN	LFSR	Y	PRQLGHLRWD	STIAQVLERK
Ego	RLKSVRG	ASV		RGY	SDAARKQARN
Klactis	RLLQKGA	GSI		RR	VRYFHRSIV
Gze	L	TLHLSRC	TRQSTTST		E I I LDRRI VQA
Pan	LARHCAGRRP	ISSRRWIS		Q	TAATQDETTE
Mgr	RLARHRCRTH	LRPGRWQSSV	AVTAGDAVAS	SKGLSRKNTD	AFSGPSSGARF
Hca	DFR WYTAEAS	ASPGGGKQLQ	TLAGMFALCF	ISMHDCLAH D	I FLVPSTARF
Calb	QQQQQQQ	Q		LLQQLQQQHP	LQQFTEEP RV
Sca					
Ani	SPLAARY	YS		T	QDGAV ERQRK
Yarrowia_lipoly				MKQ	GN
Ppa				MI	IRRFVRH
aspfum					LSTSKS S I R I
Sku	MANVM	VTSQWLVR		MVR	RRPLCVRYS T
Consistency	0 0 0 1 1 1 1 1 1 1	1 1 1 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 1	2 2 1 2 2 3 2 2 3 3	2 2 3 2 3 5 3 4 8 4

	110	120	130	140	150
Sc	NPVGIQYLGE	SLQRQVFGSC	GKDEV	EQSDKLMEL	SKKSLKDHGL
Spa	NPVGIQYLGE	SLQRQVFGNC	SGRDEA	GQSDKLMEL	SKKSLKDHGL
Sba	NPVGIQYLGE	SLHRQVFGSG	GDEGEV	HGDKLIEL	SKQSLKDHGL
Cgl	NPLGIQYLSE	SLQQQVFPGA	TGHS	EDPELVKL	AKLSLSNHDL
Kwa	NPVGIQHLSK	HLHEQVFGNQ	SEGKTEDSVS	EAEKKALVRL	SQSFLLKHHGL
Debaryomyces_ha	NPVGVQYLSE	ELHKKLFPKT	KSTDYL	KPKNPALLEL	AKEHLRYNEL
Ncr	NEIGVQQLSE	HLYKQLFPRG	NTDP	PAPELIEL	AKDHLARHDL
Ego	NPVGIQYMSH	SLHKQLFGAS	ESMRRRN	LEAARGKLVKW	SKLKLKHHGL
Klactis	NPVGIQYLSN	SLHEQIFPSK	DKYNKFF	TEEEKQELINL	SKTFLRNHDL
Gze	NEIGVQQLSS	HVFEQIFPDG	VKP	PPQELVEL	SKDHLRRHDL
Pan	NEIGVQQLSS	HIYDQLFPKG	NTSP	PSKELVDL	ARDHLRRHDL
Mgr	NEIGVQHLSL	HVFKQVFPPLG	PDP	PPKHLVKL	AEDHLARHDL
Hca	NEIGVQQLSS	HVHSQIFLRE	PTP	PNPELIAL	SKDHLRRHDL
Calb	NQLGIQYLSN	DLHKKVFPPT	STKDYL	SPQHPQLLEI	SKKHLQENEL
Sca	---	SQSLHSQIFGNK	KHSFKEDAMT	PEERTHLIDL	SKQLKSSHGL
Ani	NEIGVQQLSD	HVYSQIFFNK	PTP	PDPNLVAL	SKDHLARHDL
Yarrowia_lipoly	NPMGIQMLSS	GLYNQLFGDY	NQTNLSLT	KPEMVKLEAM	AKHHLMKHEL
Ppa	NAVGIQHLLAP	SLQNQLFGKS	SP	RLDEDLINM	AKRHLKSNL
aspfum	---	---	---	---	---
Sku	NPVGIQYLGE	SLQRQVFGSC	GSRDET	GQSDKLMEL	SKKSLKDHGL
Consistency	8468784764	5745879443	3321110000	0143349647	7734954749

	160	170	180	190	200		
Sc	WGKKTLLITDP	ISFPLPPLQ	GRS	LDEHFQ	KIGRFNSEPY	KSFCEDKFTE	
Spa	WGKKTLLITDP	ISFPLPPLQ	GRS	LDEHFQ	KIGQFSSEPY	KSFCEGKFTE	
Sba	WGKKTLLITDP	IAFPLPPLQ	GQS	LDEHFQ	KIGQFNSEPY	KSFCEGKFTE	
Cgl	LGRKTTITKP	IAFKLPELQ	GSS	LDEHFQ	RLGLYASEPY	RTMALEKFKS	
Kwa	LGKKTALISEP	ISFDVPKLQ	GRT	LDEHFQ	KLGHYAAQPF	RDLAGSKFTE	
Debaryomyces_ha	LGKKTALISEP	INIDNFPDLE	GKN	LDEHFF	RIGTEASQPY	LSIAESFLSP	
Ncr	LGKTTDKTTP	IAFQLPALV	GDT	LDEHFF	KLGVDAAEPP	LTHAKQFADA	
Ego	LGSKTSISEP	ISFELPQLQ	GTT	LDEHFQ	RLGHFASDPY	LSLCENKSTE	
Klactis	LGKPTNISDP	ISFELPKLQ	GRI	LDEHFQ	KLGHFASDPY	KSMCDNKFKN	
Gze	LGKNTDKSDP	IAFDIPLPQ	GRT	LDEHFF	KLGI DAAEPF	LTFAKQFSMA	
Pan	LGKNTDTSAP	VAFDLPLLV	GST	LDEHFF	KLGVDAAEPP	LTHAKRFALA	
Mgr	LGRTTDTSEP	ISFDLPLPH	GKT	LDEHFF	KLGI DAAEPY	LTRAKQFARA	
Hca	LGKTQEGSNP	IAFDIPLLE	GQT	MDEHFF	KLGMDCSEPY	LTKAKQYAKI	
Calb	LGKKTQITDP	ITIKNFPSLV	GKS	MDEHFN	KLGGKSSDPY	LAMATNFKK	
Sca	LGKKTALISEP	ISFELPKLQ	GNS	LDEHFQ	KLGHFASDPY	KTMVAVNKFTT	
Ani	LGKAQEHADP	VAFDLPLGLQ	GQT	LDEHFF	KLGMDSSEPY	LTKAKQYAVV	
Yarrowia_lipoly	LGKKTKKENI	IDFDLPLML	GSS	LDEHFF	RMAKVMGEPY	MDI IQGMMTK	
Ppa	IGKKSLSLNDP	IEIQLPSLQ	GNT	LDEHFF	KLGSASLEPY	NPLIDELNL	
aspfum	---	---	MPALQ	GQS	MDEHFF	KLGMDSSEPY	LTKAKSYTTV
Sku	WGKKTLLITDP	ISFPLPPLQ	GQS	LDEHFQ	KIGQFNSEPY	KSFCEDKFTE	
Consistency	6975734558	857047*496	*5709***5	88944678*8	5646454454		

	310	320	330	340	350		
Sc	A	LI PLNTL	NK EQVI I GHNV A	YDRARVLEEY	NFRDSKAFFL		
Spa	T	LI PMNTL	TK EQVVI GHNV A	YDRARVLEEY	NFRDSKAFFL		
Sba	A	LI PMNTL	RK EQVVI GHNV A	YDRARVLEEY	NFQDSKAFFL		
Cgl	D	KHEVYE	H LI PMDTL	NK EKLI I GHNV G	YDRGRVLEEY	NLKPSKAFFI	
Kwa	P	T	H LI PLDTL	NK TRLVI GHNV G	YDRARVLEEY	NCQESKAFFL	
Debaryomyces_ha	K	KYNDWE	H LMPMDCL	NN PKLI I GYNV S	YDRARVLEEY	NI KQSKGFFL	
Ncr	K	A	Q LVPMDGP	TV DRI I VGHNI G	YDRAKI LEEY	DLKQTRNFFL	
Ego	D	N	FVH	LI PLNTM	TR PQVI I GHNV G	YDRSKVLEEY	NLAPSKAFFM
Klactis	D	N	FKH	LI PMNTL	ER PKLI VGHNV G	YDRSKVLEEY	NFTPSKAYFI
Gze	D	R	Q	LI PFGDP	TK DRI I VGHNV G	FDRARVLEEY	SLKQTRNTFL
Pan	E	H	Q	LI PVGDP	TK ERVI I GHNI G	YDRARI LEEY	SLTQTRNSFL
Mgr	D	R	H	LI PLGDP	SK DRI I VGHNI G	YDRARI LEEY	SLTQTRNAFL
Hca	E	R	Q	LI PLGSP	SQ ARI I VGHNV G	YDRARTLEEY	NI KKTRNFFI
Calb	I	T	DQIDWN	H LI PMNCA	KK PKVI I GYNV S	YDRARI LEEY	NI KQTKAFYL
Sca	S	S	FKH	LI PLGENTK	V EKLI I GHNV G	YDRARVQEEY	NFNHNSKFFFL
Ani	E	I	Q	LVPLGDP	TK PRI I VGHNI G	YDRARVLEEY	DLKQNTANFFL
Yarrowia_lipoly	I	V	NHPQES	MH LI PLGPP	DE PRI VGHNI G	YDRQAVKEEY	SLKQTKNFFM
Ppa	D	H	LI PLNTF	ET EKFI VGHNV S	YDRARALEEY	NLKQTKAFWM	
aspfum	E	V	Q	LI PLGDK	KQ PRI VGHNI G	YDRARVLEEY	GMEQTRNFFL
Sku	A	A	LI PMNTL	HK EQVVI GHNV S	YDRARVLEEY	NFRDSKAFFL	
Consistency	0 0 2 1 0 0 0 1 5	* 9 * 7 6 5 4 0 4 6	5 7 7 9 9 * 9 * 9 7	9 * * 8 8 8 8 9 * *	7 6 5 5 7 7 5 7 9 8		

	360	370	380	390	400		
Sc	D	TQSLHI ASF	GLCSRQRPMF	MKNNK	KKKEA	- - EVES	- - EV HPEI
Spa	D	TQSLHI ASF	GLCSRQRPMF	MKNNK	KKKEE	- I EVES	- EV QSEI
Sba	D	TQSLHI ASF	GLCSRQRPMF	MKNNK	KKKEE	PEPAET	- EA QSEI
Cgl	D	TLSLHVATA	GLCSRQRPLF	LKNKK	LKENL	TAEPEE	EEV TEAL
Kwa	D	TMSLHI ASS	GMCSRQRQW	MQKRK	AQDKQ	VVDEVE	- - EE
Debaryomyces_ha	D	GMAHVAIS	GI CSQQRPTW	QKHKK	SKS QL	DSSKDE	SPA I ESDI DYFDKK
Ncr	D	TMSLHVAVN	GMCSQQRPTW	MKHKK	ARE LR	EKAEHE	SASV ELQEVLQGGG
Ego	D	TMSLHI SSS	GMCSRQRPFY	LKNQK	TKQE	NDGDL	DYELN EDL
Klactis	D	TMSLHVASS	GMCSRQRPTY	MKKL	KRDEE	V AQLNAE	NDDL PPI
Gze	D	TMSLHVAVN	GMCSQQRPTW	MMHK	KNREL	R EQVATEI	PSH DLATVLSTQG
Pan	D	TMSLHVAVN	GMCSQQRPTW	MKHK	KDREF	R DKIAKE	SSNA ELVELLNNGA
Mgr	D	TMSLHVAVN	GMCSRQRPTW	MKHK	KNREL	R DKIAQET	DSI ELA GLLNSNS
Hca	D	TMSLHVAVN	GMCSQQRPTW	MRHR	KNRDL	R DKLLK	ENNSI ELA ALLENK
Calb	D	GMSLHVATS	GI CSRQRPKW	AKYN	KVKKLS	QEDVDEI	EGT MEEEEEEEEGE
Sca	D	TQSLHVASS	GLCSRQRPLF	MQM	KRRKEK	EEEQL	QQETE SEEDGINGIS
Ani	D	TMSLHVAVN	GMCSQQRPTW	MRHK	KNRDL	R DKIASD	SNV ELAALI ENKM
Yarrowia_lipoly	D	TLSLHVSVN	GMCSRQRGTW	MKARK	GQKRM	KKL	LDD EDL SNEENVNEL
Ppa	D	TMALHVSVS	GMCTRQRGTW	I KH	KKKEQEN	TMSVASI	KSK VQNEELQELL
aspfum	D	TMSLHVAVN	GMCSQQRPTW	MRHK	KNKDL	R DKIANE	SNV ELA ALLESKM
Sku	D	TQSLHI ASF	GLCSRQRPMF	MKNNK	KKETE	AEAEVE	AEVQ SEIS
Consistency	* 8 7 9 * * 9 8 5 5	* 8 * 9 7 * * 8 5 6	7 7 5 5 * 4 7 6 4 3	3 4 3 3 3 7 2 3 4 4	4 3 4 3 1 2 1 1 1 1		

	410	420	430	440	450
Sc	S I E D Y D D P			W L N V S A L N S L K D	
Spa	S I E D Y D D P			W L N V S A L N S L K D	
Sba	S I E D Y D D P			W L N V S A L N S L K D	
Cgl	E L N D V A D P			W L K V S T M N A L S D	
Kwa	E L A F N N D P			W T T V S T L N S L K D	
Debaryomyces_ha	L S A V E I A N E L L D D P			W L N K G S P N S L A N	
Ncr	L T A E E A D L			W V D K S S I N S L R D	
Ego	E L G N			W Y N L S T M N S L R D	
Klactis	A D V D E D S P			W I N K S S A N S L K D	
Gze	M H D E E D L			W V E R S S V N S L R D	
Pan	F S A Q E E E L			W V E R S S I N S L R D	
Mgr	L G Q E E E E L			W V E R S S V N S L R D	
Hca	L T E E E E E L			W V G R S S V N S L R D	
Calb	M S Q E L G Q L S S E S N D V S D Y A K L L S D G F Q V A V E L E V A E D P			W L L M G A P N S L A N	
Sca	T T T L N T L E E E E D P			W L N I S T M N S L S D	
Ani	I R E E E E E L			W V G R S S V N S L R D	
Yarrowia_lipoly	D I S D D L A D A Y S E D P			W V M V S A T N S L A E	
Ppa	D S T S I E D P M L E D N P			W I N K S S L N S L E R	
aspfum	L S E E E E E L			W V G R S S V N S L R D	
Sku	I E D Y D D P			W L N V A A L N S L K D	
Consistency	0 0 0 0 0 0 3 3 5 4	4 5 6 4 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0	* 7 4 4 8 7 5 * 9 * 6 8

	460	470	480	490	500
Sc	V A K F H C K I D L D K T D R D F F A S T D K S T I I E N F Q K L V N Y C A T D V T A T S Q V F D				
Spa	V A K F H C K I N L D K A D R D F F A S T D K S T I I E N F Q K L V N Y C A T D V T A T S Q V F D				
Sba	V A K F H C K I K L D K A D R D F F A A T D K S T I I E N F Q K L V N Y C A T D V T A T S Q V F D				
Cgl	V A L F H C G I K M D K T T R D Y F A S T N K Q D V I D K F Q L L V N Y C A T D V I T T S K V F D				
Kwa	V A L L H C G I K L D K E Q R E D F G T T N K N D I I D K F N E M V N Y C A N D V E A T S K V F D				
Debaryomyces_ha	V A D F H C N I K L D K T D R D Y F G T E D P Q D V I N N F N N L M D Y C A K D V D A T Y S V T A				
Ncr	V A Q F H L N V K I D K D I R D V F A E T D R N V I L N Q L D D L L T Y C A A D V Q V T H Q V Y Q				
Ego	V A K L H C N I N L E K T L R D S F E A L D K Q E V I D K F Q Q L V D Y C S T D V E T T S K V F D				
Klactis	V A R F H C N I K M S K E A R E S F A T L D K F E I I N N F D S M V E Y C A F D V E A T S K V F D				
Gze	V A K F H L N V S I D K T I R N D F G E L D R D G V V E K L D E L L T Y C A A D V S I T H R V Y Q				
Pan	V A Q F H L G V T I D K E A R N A F G E F D R D G I N Q Q L D E L L T Y C A A D V A I T H R V Y Q				
Mgr	V A K F H L N V S I D K A I R D D F S V L D R D G I M A K L E E L L D Y C A A D V A V T H R V Y S				
Hca	V A K F H C N V T I D K A Q R D H F G D L D R S G I L E R L D E L L D Y C A A D V S I T H R V Y K				
Calb	V A R F H C N I H L D K T D R D Y F A T E D P N L V I D N F P K L M D Y C A K D V D A T F E V T K				
Sca	V A L F H C G I K M D K E P R D F F A S L D K Q V V I D N F Q K L V T Y C A T D V E A T S H V F D				
Ani	V A K F H C D V T I D K S Q R D Y F G E L E R P Q I L A K L D E L L D Y C A A D V A I T H R V Y K				
Yarrowia_lipoly	V A H L H L N G L V M S K D D R E L F S Q E S L G E L Q Q R V K Q A I D Y N I T D V E I T A Q V F K				
Ppa	V A E F Y C K I K L K K D I R N S F A T E D P D E L R G N F D E L M Q Y C A K D V F A T G K V F Q				
aspfum	V A K F H C D V T I D K S Q R D H F G E L N R E Q I L G K L E E L L D Y C A A D V A I T H R V Y K				
Sku	V A K F H C K I N L D K A D R D F F A S T D K S T V I E N F Q K L V N Y C A T D V T A T S K V F D				
Consistency	* * 5 8 9 7 5 0 9 5	8 8 * 5 3 * 7 3 * 6	5 4 7 6 4 3 8 6 5 6	6 5 6 8 7 5 * 9 9 5	* * 4 6 * 5 6 * 7 5

	510	520	530	540	550
Sc	E I F P V F L K K C	P H P V S F A G L K	S L S K C I L P T K	L N D W N D Y L N S	S E S L Y Q Q S K V
Spa	K I F P V F L K K C	P H P V S F A G L K	S L S K C I L P T K	L E D W N D Y L N S	S E S L Y Q Q S K V
Sba	K I F P V F L K K C	P H P V S F A G L K	S M S K C V L P T K	L N D W N K Y L N S	S E S L Y Q Q S K V
Cgl	K V F P S F M K K C	P H P V S F G G L K	A L S S C I L P T D	L O I K W Q N Y I K N	S E N K Y Q D S K K
Kwa	A V F P L F L S K C	P H P I S F G A L R	L L S T S I L T T K	D G T W K D Y I N R	S E T L Y Q K S K Q
Debaryomyces_ha	K L F S Q F R Q K V	P H P V S F A A L R	H L G T L M L P T T	- K N W D N Y I E T	A E E V Y Q K N R E
Ncr	V V F P N F L G V C	P H P V S F A A L R	H L A S V I L P V N	- K T W D T Y I E T	A E A T Y L Q M L H
Ego	K V Y P L F R E Q C	P H P V S F G A L R	F L S N C I L P V Q	P K M W K Q Y I D G	A E T L Y Q N A R Q
Klactis	E V F P K F L Q K C	P H P V S F G A L R	F L S T S I L P T R	V G K W E K Y L N S	S E T L F Q D N K A
Gze	I V F P N F L K T C	P H P V S F A A L R	H L S S V I L P V D	- K S W D A Y I K N	A E A T Y H T L S D
Pan	V V F P N F L R V C	P H P V S F A A L R	H L S N V I L P V N	- K T W D S Y I A N	A E A T Y R K L S D
Mgr	I V F P N F L E V C	P H P V S F A A L R	H L S S V I L P V N	- E S W D S Y I A N	A E A T Y Q K L S E
Hca	Q V F P N F L E V C	P H P V S F G A L Q	H L S S V I L P V N	- N T W D E Y L K N	A E E T Y Q R R L E
Calb	K L F P V F I Q R I	P H P V S F A A L R	L L G T L I L P T T	- T K W P K Y I E A	A E S L Y E S N R Q
Sca	K V L P R F L K K C	P H P V S F G A L K	T L S S F I L P T R	R A E W D D Y I N R	S E K L Y Q E S K L
Ani	K V F L N F L E T C	P H P V S F G A L R	H L S S V I L P V N	- Q T W K E Y L D N	A E S T Y N Q R L G
Yarrowia_lipoly	K I L P L F L D T C	P H P V S F T A L K	N I S S A I L P T D	- S E W K D Y L N R	C E N M Y Q K M K K
Ppa	K V Y P K F K K L I	P H P V T L A A L K	D I S S C I L P T T	- T K W E D Y I E T	S E R L Y Q E S R R
aspfum	K V F P N F L E V C	P H P V S F G A L R	H L S S V I L P V N	- E T W K E Y I T N	A E A T Y H Q R V D
Sku	K I F P V F L K K C	P H P V S F A G L K	S L S K C I L P T K	L N D W N N Y L N S	S E S L Y Q Q S K V
Consistency	5 8 8 8 4 * 7 6 5 8	* * * 9 9 9 6 7 * 8	4 9 8 6 4 9 * 9 7 5	0 4 4 * 5 5	* 8 5 5 7 * 5 6 9 7 5 4 5 3

γ_1

	560	570	580	590	600
Sc	Q I E S K I V Q I I	K D I V L L K D K P	- - - - -	D F Y L K D P W L S	Q L D W T T K P L R L T K K G
Spa	E I E S K I V Q I I	K D I V L L K D Q P	- - - - -	D I Y L K D P W L S	Q L D W T T K P L R L T K K G
Sba	E I E S K I V Q I I	K D I V L L K D Q P	- - - - -	D E Y R K D P W L S	Q L D W S A K P L R M T K K G
Cgl	V I D D K I I E I V	Q E T V K L K D Q K	G - F Q	D V I K T N P W L R	Q L D W E I K P L R L T K K G
Kwa	S I E K K I L T I V	E D I V K L K D N L	- - - - -	S A V N D D P W L R	Q L D W T I K P L R I T K K G
Debaryomyces_ha	V S T I L K E R A	N A F V N Y I D Q N	N E S L K P	D W E T D P W L S	Q L N W T I K V Q R L K K N G
Ncr	G V Q E R L F T L M	E R T L D Y K A D P	- - - - -	E K Y L S D P W L S	Q L D W S G Q E I K M A K P K
Ego	A I E E K I I T I I	E E I V K L K D K P	- - - - -	E V Y E N D P W Y K	Q L D W T I S P L K L T K K G
Klactis	K I E D K I V K I V	E D T L Q L K N Q T	- - - - -	E L I K N D P W L K	Q L D W T A V P V K Y T K S G
Gze	A V Q E R L I G L T	E S A L E M K D S P	- - - - -	G E Y E N D P W M Q	Q L D W S G Q E I K M A K G K
Pan	G V R E R L V A L T	D K A R E I K K Q P	- - - - -	E K W S T D P W M S	Q L D W S G Q E V K M T K G T
Mgr	A V Q E R L I A L T	E K A L E I K D D Q	- - - - -	D K W S A D P W L S	Q L D W S G Q E I K M V K G K
Hca	Y V Q K R L V E L C	D S A L A V K D K P	- - - - -	E V Y E N D P W L R	Q L D W S G Q E I R M V K G R
Calb	E V A T T L Q Q L A	N D L V A H I Q N K	K Q - V K	P D Y E N D P W L S	Q L N W K L K E T R L K K D G
Sca	E I E R K I I L I I	E E L V K I K D D P	V K - L	D Q I K D D P W L R	Q L D W T I K P L R L T K K G
Ani	D V Q R R L V E L C	D E A L S V K D D P	- - - - -	E K Y M N D P W L R	Q L D W S G Q E V K M V K G K
Yarrowia_lipoly	S I D E G L Q S L V	T E A V K L K D E A	D K - - -	E P W L N D P W L S	Q L D W T I K P I K M V K I P
Ppa	M I E K N L H V I C	E E T V K L K D D P	- - - - -	K P W E N D P W L S	Q L D W T I D P I R L T K K G
aspfum	D V Q R R L V E L C	N E A L K V K E N P	- - - - -	D I Y M N D P W L R	Q L D W S G Q E I K L V K G K
Sku	E I E S K I V Q I I	K D I V L L K D Q P	- - - - -	D V Y L K D P W L A	Q L D W T A K P L R V T K K G
Consistency	4 9 6 5 6 8 6 4 7 5	5 6 5 7 4 6 8 7 5 5	0 0 0 0 0 6 3 5 4 5	9 * * 8 6 * * 9 * 7	4 6 5 7 8 7 6 * 4 5

γ_1

	610	620	630	640	650
Sc	VP	AKCQKLPGFPE	WYRQLFSP	KD	T
Spa	VP	AKCQKLPGFPE	WYRQLFSP	KD	T
Sba	VP	AKCQKLPGFPE	WYRQLFSP	SD	T
Cgl	VP	VKNQKLPGYPE	WYRSLFPNKE	A	
Kwa	IP	AKGQKLPGYPE	WYRQLFPT	RL	S
Debaryomyces_ha	EP	ANQAFLLTGYPE	WYRDLF	KT	VT
Ncr	KKGD	VERPA	LNQKLPGY	PQ	WYKDLFVKVP
Ego	VP	AKRQKMPGYPA	WYRSLFTG	T	D
Klactis	EI	RKNQKLPGYPE	WYRSLFP	SKS	T
Gze	GN	EPPRPV	ARQKKPGMPK	WYKDLF	QSSD
Pan	PT	SPPREY	KNQKMPGEPN	WYRNLF	PKAK
Mgr	KKGD	PPRPA	ARQKKPGMPQ	WYKDLF	VKND
Hca	KKGD	PPRPA	ARQKMPGMPK	WYKALF	PTST
Calb	TP	YK	KTAYMTGYPE	WYRELF	KTSK
Sca	VP	V	KGQKLPGYPE	WYRSLFP	SKT
Ani	KKGD	PPRPA	ARQKKPGMPQ	WYKDLF	SSNT
Yarrowia_lipoly	AK	QGGGERLP	KNVKFPGYPE	WYKALF	KSSA
Ppa	E	I	HKNQKLPGYPN	WYKQLI	VK
aspfum	KKGD	PPRPA	ARQKKPGMPQ	WYKDLF	PSAT
Sku	IP	AKGQKLPGFPE	WYRQLFSP	KD	T
Consistency	4400001115	648868*6*6	**85*956	44	1000000000

γ_2

γ_3

	660	670	680	690	700
Sc	VEPKITIKSR	IIPILFKLSW	ENSPVI	WSKE	SGWCF
Spa	IEPKITIKSR	IIPILFKLSW	ENSPVI	WSKE	SGWCF
Sba	IEPKITIKSR	IIPILFKLSW	ENSPVI	WSKE	SGWCF
Cgl	ELPQITMRSR	IIPILFKLSW	EDSPVI	WTED	CGWCF
Kwa	EFPQVSI RNR	LIPILFKLSW	EGEPVI	WSAT	DGWCF
Debaryomyces_ha	REMNI	STRTR	VTPLLLK	KLKW	EGYPL
Ncr	EFINLTVRSR	IAPLLLKLSW	EGYPL	FWS	DQ
Ego	KKPNI	TIRSR	LIPILFK	LTSW	EGYPV
Klactis	TRPQITIRTR	QVPLFFK	LTSW	EDCPV	VY
Gze	SPI NITVRTR	VAPLLLK	MMSW	DGYPL	FWS
Pan	GPI NLTVRTR	IAPLLL	RRLSW	DGHPL	FWS
Mgr	SPI GITVRTR	IAPLLL	RRLAW	DGHPL	VWS
Hca	SAINLSVRTR	LAPILL	KLSW	LGYPV	VWS
Calb	EEMNLT	LTKTR	ITPLLL	RRLKW	EGHPL
Sca	QQPQITIKSR	IIPILFK	LTSW	EGHPV	F
Ani	ADINLTVRTR	IAPILL	KLSW	DTHPLI	WSD
Yarrowia_lipoly	GPMHITTKTR	ITPLLL	RRLKW	EGYPV	VWS
Ppa	NELKLT	TKTR	TSPLLL	KLAW	NGNPL
aspfum	ADINLTVRTR	IAPILL	KLSW	DGYPL	TWS
Sku	LEPKITIKSR	IIPILFK	LTSW	ENSPVI	WSKE
Consistency	344588787*	75*877897*	764*	868755	4**6*57645

γ_3

	710	720	730	740	750
Sc	V L A D S V S Q E E	E E I R T H N L	--- G L Q C T G	V L F K V P H P N G	P T F N C T N L L T
Spa	V L A D S V S Q E E	E E I R M Q N L	--- G S Q P T G	V L L K V P H P N G	P T S N C T N L L T
Sba	V L A D S I S Q E E	E Q I R S R D L G S	S S S S L S A S T D	V L F K V P H P N G	P A F N C T N L L T
Cgl	R L A D E A T I K N	H F E K L V E M G	Q Y K K T D I G N D	L L F K V P H P N G	P D N N C T T L L S
Kwa	M L A N K E V S A D	P	--- G Y V M F K I P H P N G	Q Q Y N C T T L V S	
Debaryomyces_ha	S Y P R A K L N D E	D L E K F L P Q L	--- R D D G N S Y	E L F K V P H P E G	P G K R C T S I L S
Ncr	T P V Q F E D P D V	D D R L	--- R M D V D H	K Y F K L P H K D G	P N A R C V N P M A
Ego	A I A N S P D F D R	N V	--- G Y T A F R I P H P S S	A E Q R T T T L F S	
Klactis	V K A V E V D D Y H	E E	--- D T T L F K I P H P N G	P E F N C S T L L S	
Gze	V E C K F D E K E	I K L R	--- D D H A H	A Y F K L P H K D G	P T A R C T N P M A
Pan	K P C D F E G E T V	L A V K	--- E D T K H	A Y F K L P H K D G	P T A R C T N P M A
Mgr	I E C D M S D E T N	E K L R	--- T D G G	I Y F K L P H K D G	P L A R C A N P M A
Hca	A P S N M A E E K N	L S L R	--- D D R E H	L Y F K L P H K D G	P D A R C S S P L A
Calb	K K Y V Q P K L T E	E E L M E T R	D A L L T K G K S Y	V L F K V P H P N G	P S L R C T H I L S
Sca	V V A D E D S I S Q	Y R E L S H S	I D D N G D S D Q	I L L K V P H P N G	P T F N C T T L L S
Ani	V A C N M T E E K N	P E L H	--- N D R K H	I Y F K L P H K D G	P T A R C V S P L A
Yarrowia_lipoly	K L A T L D T E P		H L R D D G S T	L C F R I P H K D G	G N T R T T Q L M G
Ppa	V M V S L K K L S E	D P G F E S	I R A E D L K N F	T F V K V P H P D G	P S A R V T N C M T
aspfum	V L C D M S E E K N	L E L R	--- N D R K H	V Y F K L P H K D G	P Q A R C V N P L A
Sku	I L A D S V S E E E	E Q L R M	Q N Q G S P S T D	V L F K V P H P N G	P T F N C T N L L T
Consistency	5 4 5 5 4 4 4 4 4 4	3 3 3 3 0 0 0 0 0 0	0 0 0 0 0 3 2 2 4 3	5 5 8 9 8 * * 6 6 9	7 4 4 7 8 7 6 5 7 6

γ_4

	760	770	780	790	800
Sc	K S Y N H F F E K G	V L K S E S E L A H	Q A L Q I N S S G S	Y W M S A R E R I Q	S Q F V V P S C K F
Spa	K S Y N H F F E K G	V L K S E S E L A H	Q A L Q I N S S G S	Y W M S A R E R I Q	S Q F V V P S S K F
Sba	K S Y N H F F E K G	V L K S E S E L A H	Q A L Q I N S S G S	Y W M S A R E R I Q	S Q F V V P S S K F
Cgl	K P Y V H F F E K G	I L K S E S E L A S	I A L K M N A S G S	Y W M S A R E R I M	S Q F V V P K E D F
Kwa	K Q Y I Q F F N K G	I L T S H S E V A Q	E A L E I N S S C A	Y W M S A R E R V M	S Q F V V S S K D F
Debaryomyces_ha	K S Y L R Y F E S G	I L T S E Y S F A Q	E I L S L N S T A S	Y W M G N R Q R I M	D Q F V I Y S D A
Ncr	K G Y L P Y F E K G	I L S S E Y P Y A K	E A L E M N A S C S	Y W I S A R E R I K	N Q M V V Y E D Q L
Ego	K P Y M H F M D K N	V L Q S E Y S L A R	E A L Q I N S S G S	Y W T A A R E R I K	T Q H V V S K E D F
Klactis	K P F M H Y F E K G	V L K S Q S K L A H	D A L Q I N A S G A	Y W M S A R E R I M	S Q H V V S V N D F
Gze	K G Y L S Y F E S G	M L S S E Y T Y A K	E A L E M N A S C S	Y W I S A R D R I K	S Q M V V Y D Q D L
Pan	K S Y L S Y F E K G	T L S S E Y D Y A K	E A L E M N A S C S	Y W I S S R D R I M	S Q L V V Y E S N L
Mgr	K G Y L G Y F E S G	L L S S E Y S Y A K	E A L E M N A S C S	Y W I S A R D R I M	S Q L V V Y E N E V
Hca	K G Y L Q Y F E G G	K L S S Q F A L A K	E A L E M N A S C S	Y W I S A R D R I M	S Q M V V Y H N D I
Calb	K L Y V R N F E D G	T L T S E Y D Y A T	K I L N L N N E A S	Y W L G N R N R I K	D Q F V V Y N Q Q G
Sca	K P Y I H Y F E K N	V L T S D S E L A H	Q A L N M N S S G S	Y W M A A R E R I M	S Q F V V S G K D F
Ani	K G Y L Q Y F E R G	T L S S Q Y A L A K	E A L E M N A S C S	Y W I S A R D R I M	G Q L V V Y E N D V
Yarrowia_lipoly	K S F L G H F E K G	T L S S D N P Q T K	E M F E M S V C S S	Y W V S N R S R I M	S Q H C I W D G D G
Ppa	K S C L G F F E K G	F L N S Q Y P L A K	D A L Q M A V A S S	Y W T S S R E R I M	N Q F V V F E D D
aspfum	K G Y L Q F F E R G	T L S S Q Y A L A K	E A L E M N A S C S	Y W I S A R D R I M	S Q M V V Y E D Q V
Sku	K S Y N H F F E K G	V L K S E S E L A H	Q A L Q I N S S G S	Y W M S A R E R I Q	S Q F V V P S S K V
Consistency	4 8 5 4 7 9 9 6 8	5 * 6 * 7 5 5 6 9 5	6 8 9 6 7 8 6 8 4 9	* * 6 8 7 * 7 * 9 6	7 * 5 9 9 4 5 4 5 4

γ_4

	810	820	830	840	850	
Sc	PNEFQSL	AKSSLNNE	KTNDLAIIP	KIVPMGTITR	RAVENAWLTA	
Spa	PNEFHSL	ARSNLNNG	KTNDLAIIP	KIVPMGTITR	RAVENTWLTA	
Sba	PNEFRSLA	AKPVPNTA	SADDLAIIP	KIVPMGTVTR	RAVENTWLTA	
Cgl	PEQFLPIK	VGKDLKPT	ESSQAIIIP	QIIPMGTVTR	RAVESTWLTA	
Kwa	PEQ---	RALDHKPS	SDKDASIIIP	SVIPMGTVTR	RSVENTWLTA	
Debaryomyces_ha	NHKKNK	FDTKVDSK	KHKDMGIIIP	KLCSMGTI	RATENTWLTA	
Ncr	PPS---	FVNKDADS	NTPIGGFVLP	QVIPMGTI	RAVERTWLTA	
Ego	PDQ---	AMVDGTPS	RDDQVGAIIIP	KIIPMGTVTR	RAVENTWLTA	
Klactis	KSE---	NAKTPFD	KDLDVGIIP	KLVPMGTVTR	RSVENTWLTA	
Gze	RVMRGST	RGKP	KKDKPGFIIP	QIIPMGTI	RAVENTWLTA	
Pan	PEPVRPK	DLPE	DAPTGGFIIP	QVIPMGTVTR	RAVENTWLTA	
Mgr	SAPTASASTK	KSRNTKRT	SETKSGFIIP	QIIPMGTI	RAVENTWLTA	
Hca	ERLQNK	SSK	STLKQGYLP	QIIPMGTI	RAVENTWLTA	
Calb	KNKFF---	DTIKKSK	EHSSMGIIIP	NLATMGTI	RATENTWLTA	
Sca	NKEFNPL	NKVI	GTPG	EKEELGIIP	TVIPMGTVTR	RAVENTWLTA
Ani	RPA---	SPSSKN	GDQKLYIIP	QIIPMGTI	RAVENTWLTA	
Yarrowia_lipoly	V---	DMG	RGPGGIIP	RIIPMGTI	RCVENTWLTA	
Ppa	---	---	MGYLP	QIIPMGTI	RAVENTWLTA	
aspfum	KQS---	GSTGE	KSNRLGFIIP	QIIPMGTI	RAVENTWLTA	
Sku	PNEFLSL	AKPAPSSD	KADDLAIIP	KIVPMGTITR	RAVENTWLTA	
Consistency	3331111100	0011212333	533347698*	6888**9**	*88*99***	

γ4

	860	870	880	890	900
Sc	SNAKANRI	GS ELKTQVKAPP	GYCFVGDVD	SEELWIASLV	GDSIFNVHGG
Spa	SNAKANRI	GS ELKTQVKAPP	GYCFVGDVD	SEELWIASLV	GDSIFNVHGG
Sba	SNAKSNRI	GS ELKTQVKAPP	GYCFVGDVD	SEELWIASLV	GDSIFNVHGG
Cgl	SNAKKNRI	GS ELKAQI CAPE	GYSFVGDVD	SEELWIASLF	GDSVLRIGHG
Kwa	SNAKKNRI	GS ELKAQI QAPP	GYAFVGDVD	SEELWIASLV	SDSLFINHGG
Debaryomyces_ha	SNSKKNRI	GS ELKAMI EAPE	GYTFVGDVD	SEELWIASLV	GDSMFEMHGS
Ncr	SNAKKNRV	GS ELKAMV RAPP	GYV FVGDVD	SEELWIASVV	GDATFKLHGG
Ego	SNAKANRI	GS ELKTKI VAPE	GYSFVGDVD	SEELWIASLI	GDSVFNHGG
Klactis	SNAKKNRI	GS ELKCNV EAPP	GYTFVGDVD	SEELWIASLV	GDSIFDIHGA
Gze	SNAKKNRV	GS ELKAMV KAPR	GYCFVGDVD	SEELWIASLV	GDATFKLHGG
Pan	SNAKKNRV	GS ELKAMV RAPP	GYV FVGDVD	SEELWIASVV	GDATFKLHGG
Mgr	SNAKKNRV	GS ELKAMV KAPP	GYCFVGDVD	SEELWIASVV	GDATFKLHGG
Hca	SNAKANRV	GS ELKAMV RAPP	GYV FVGDVD	SEELWIASLV	GDAQFQIGHG
Calb	SNAKKNRI	GS ELKSLV EAPK	GYCFVGDVD	SEELWIASLV	GDSMLQIGHG
Sca	SNAKANRI	GS ELKAQV KAPK	GYCFVGDVD	SEELWIASLV	GDSVFNHGG
Ani	SNAKANRV	GS ELKAMI KAPP	GYAFVGDVD	SQELWIASLI	GDAQFQLHGG
Yarrowia_lipoly	SNAKKNRI	GS ELKSMI KAPP	GYSFVGDVD	SEELWIASLY	GDSCFKNHGS
Ppa	SNAKKNRL	GS ELKSLI EAPK	GYCFVGDVD	SEELWIASLI	GDSVFKIGHG
aspfum	SNAKANRV	GS ELKAMI KAPP	GYV FVGDVD	SQELWIASLV	GDAQFQLHGG
Sku	SNAKANRI	GS ELKTQVKAPP	GYCFVGDVD	SEELWIASLV	GDSIFNVHGG
Consistency	**9*6**9**	**6696**6	*5*****9**	88	9*85857**8

Pol I

	910	920	930	940	950	
Sc	T A I G W M C L E G	T K N E G T D L H T	K T A Q I L G C S R	N E A K I F N Y G R	I Y G A G A K F A S	
Spa	T A I G W M C L E G	T K N E G T D L H T	K T A Q I L G C S R	N E A K I F N Y G R	I Y G A G A K F A S	
Sba	T A I G W M C L E G	T K N E G T D L H T	K T A Q I L G C S R	N E A K I F N Y G R	I Y G A G A K F A S	
Cgl	T P I G W M C L E G	T K A E G S D L H T	K T A Q I L S C D R	N E A K I F N Y G R	I Y G A G V K F A S	
Kwa	T P I G W M C L E G	S K N E G T D L H T	K T A E I L G C S R	N E A K I F N Y G R	I Y G A G L K F A T	
Debaryomyces_ha	T A L G W M T L E G	D K S E K T D L H S	K T A E I L G I L R	N D A K V F N Y G R	I Y G A G V K F A T	
Ncr	N A I G F M T L E G	T K S Q G T D L H S	R T A S I L G I T R	N D A K V F N Y G R	I Y G A G L K F A S	
Ego	T A I G Y M C L E G	T K D A G T D M H T	K T A S I L G C S R	N E A K I F N Y G R	I Y G A G V R F A A	
Klactis	T A I G W M C L E G	T K S L G T D L H S	K T A H I L G C S R	D E A K I F N Y G R	I Y G A G V K F A A	
Gze	N A V G F M T L E G	T K S A G T D L H S	R T A S I L G I T R	N N A K I F N Y G R	I Y G A G L K F A A	
Pan	N A I G F M T L E G	T K A A G T D L H S	R T A A I L G I T R	N D A K V F N Y G R	I Y G A G L K F A G	
Mgr	N A I G F M T L E G	T K A A G T D L H S	R T A S I L G I T R	N D A K V F N Y G R	I Y G A G V K F A G	
Hca	N A I G F M T L E G	T K S A G T D L H S	K T A N I L G I S R	N D A K V F N Y G R	I Y G A G L K F A A	
Calb	T A L G W M T L E G	E K S Q K T D L H S	K T A S I L G I S R	N D A K V F N Y G R	I Y G A G V K F A T	
Sca	T A I G W M C L E G	T K N E G T D L H T	K T A E I L G C S R	N E A K I F N Y G R	I Y G A G V K F A G	
Ani	N A I G F M T L E G	S K A A G T D M H S	R T A K I L G I S R	N D A K V F N Y G R	I Y G A G V K F A A	
Yarrowia_lipoly	T A I G W M T L E G	T K S E G T D L H S	K T A S I L G I S R	N E A K I F N Y G R	I Y G A G L K F A E	
Ppa	T A I G W M T L E G	T K N E G T D L H S	K T A K I L G I S R	N E A K I F N Y G R	I Y G A G I K F T T	
aspfum	N A I G F M T L E G	S K A A G T D M H S	R T A Q I L G I S R	N D A K V F N Y G R	I Y G A G V K F A A	
Sku	T A I G W M C L E G	T K N E G T D L H T	K T A Q I L G C S R	N E A K I F N Y G R	I Y G A G A K F A S	
Consistency	7 8 9 * 6 * 6 * * *	7 * 6 6 8 9 * 9 * 7	8 * * 5 * * * 9 6	7 * 9 7 * * 9 * * * * *	* * * * * 6	9 * 9 5

Pol II

	960	970	980	990	1000
Sc	Q L L K R F N P S L	T D E E T K K I A N	K L Y E N T K G K T	K R - S K L F K K F	- W Y G G S E S I L
Spa	Q L L K R F N P S L	T D E E T K K I A N	K L Y E N T K G K T	K R - S K L F K K F	- W Y G G S E S I L
Sba	Q L L K R F N P S L	T D D E T K K I A N	K L Y E N T K G K T	K R - S K L F K K F	- W Y G G S E S I L
Cgl	Q L L K K F N P T L	S E K E A K E T A V	K L Y E S T K G K T	K R - S K V F K K F	- W Y G G S E S I L
Kwa	Q L L K Q F N P G L	S D S E A K S T A E	K L Y N S T K G L T	K R - S H I F K K F	- W Y S G S E S I L
Debaryomyces_ha	R L L K Q C N S S L	S D V E A E L M A K	T L Y A K T K G Q T	N T - S K F L D R R	M Y H G G T E S I M
Ncr	Q L L R Q F N P S L	T E A E T T A I A T	K L Y D A T K G A K	T N R K S L Y K R P	F W R G G T E S F V
Ego	Q L L K Q F N P T L	S E E Q T R A T A E	R L Y E A T K G K T	Y H - S K I F K K F	- W F G G S E S I L
Klactis	Q L L K K F N P M L	S E D E A K K T A E	R L Y E N T K G K K	Q R - S T I F K T F	- W Y G G S E S I L
Gze	T L L R Q F N P N L	S E K Q T M E T A S	R L Y A N T K G T K	T N R K L L Y K R P	F W R G G T E S F V
Pan	Q L L R Q F N P S L	S E Q E T L D I A G	K L Y A T T K G T K	T N R K S L Y K R P	F W R G G T E S F V
Mgr	Q L L R Q F N P N L	S E Q E T V E T A S	K L Y A A T K G A K	T N R K A L Y K R P	F W R G G T E S F V
Hca	T L L R Q F N P S M	S E Q E T K R V A A	K L Y K E T K G T R	T N R K I L S D N P	F W R G G T E S F V
Calb	R L L K Q F N A N I	T D E E A D K V A R	Q L Y D S T K G R T	A V - S K Y L P L R	I Y Y G G T E S I M
Sca	T L L K R F N P S L	S D D E A K K T A I	K L Y E N T K G K T	K R - S R I F K K F	- W Y G G S E S I L
Ani	T L L R Q F N P S M	S E R E T Q E V A S	K L Y K E T K G A K	T T R R L L S D N P	F W R G G T E S F V
Yarrowia_lipoly	A L L K Q F S P N V	T D A E A S T L A G	Q L Y T S T K G K K	I Y - - - S I - - G	K Y S G G T E S I L
Ppa	T L L K K F N P A L	S D A E A K A T A N	A L Y T A T K G I S	G R - - - Y D K K S	I W Y G G S E S I I
aspfum	T L L R Q F N P S M	S E K E T Q E V A T	N L Y R E T K G T R	T T R R I L S E N P	F W R G G T E S F V
Sku	Q L L K R F N P S L	T D E E T K K I A N	K L Y E N T K G K T	K R - S K L F K K F	- W Y G G S E S I L
Consistency	6 * * 8 7 9 9 8 6 8	8 8 5 9 7 5 5 6 * 4	6 * * 5 5 * * * 5 6	4 5 1 5 3 6 5 6 5 3	1 8 5 9 * 7 * * 7 7

γ5

	1010	1020	1030	1040	1050
Sc	F N K L E S I A E Q E T P K T P V L G C G I T Y S L M K K N L R A N S F L P S R I N W A I Q S S G V				
Spa	F N K L E S I A E Q E T P K T P V L G C G I T Y S L M K K N L R A N S F L P S R I N W A I Q S S G V				
Sba	F N K L E S I A E Q E T P M T P V L G C G I T Y S L M K K N L R A N S F L P S R I N W A I Q S S G V				
Cgl	F N K L E S I A E Q D N P M T P V L G C G I T N S L M K Q N L K A N S F L P S R V N W A I Q S S G V				
Kwa	F N K L E H I A E Q D S P M T P V L G A G I T Y S L M K R N L G T N T F L P S R I N W V I Q S S G V				
Debaryomyces_ha	F N M L E S I A Y Q E D P K T P V L G A A I T D A L T I A N L N K N N Y L T S R I N W V I Q S S G V				
Ncr	F N M L E E F A E Q E R P R T P V L G A G I T E A L M S R W V S K G G F L T S R I N W A I Q S S G V				
Ego	F N K L E H I A E Q D Q P R T P V L G A G I T V S L L K R N L G K D T F L P S R I N W A I Q S S G V				
Klactis	F N K L E S I A E Q D E P K T P V L G A G I T Y S L M K K N L G K D T F L P S R I N W V I Q S S G V				
Gze	F N K L E E F A E Q D K A R T P V L G A G I T E A L M G R Y I S K G G Y L T S R I N W A I Q S S G V				
Pan	F N K L E E F A E Q S H P R T P V L G A G I T E A L T A R F L S K G G Y L T S R I N W A I Q S S G V				
Mgr	F N K L E E F A E Q E R P R T P V L G A G I T E A L M S R F M S K G G Y L T S R I N W A I Q S S G V				
Hca	F N K L E E F A E Q E R P R T P V L G A G I T E A L M R R F I N K G S F M T S R I N W A I Q S S G V				
Calb	F N A L E A I A Q Q E E P K T P V L G A S I T D A L N V K N L N T N Q Y M T S R V N W T I Q S S G V				
Sca	F N K L E S I A E Q D E P K T P V L G C G I T F S L M K R N L R A N S F L P S R I N W A I Q S S G V				
Ani	F N K L E E F A D Q E R P R T P V L G A G I T E A L M R R F I N R G S F M T S R I N W A I Q S S G V				
Yarrowia_lipoly	F N K L E E I A T Q D M P E T P V L G A V I T K A L S S Q N L K K S S F M T S R V N W T I Q S S G V				
Ppa	F N R L E A I A E M A H P K T P V L G A G I T A P L Q K A N L S T N N F L T S R I N W A I Q S S G V				
aspfum	F N K L E E F A D Q E R P R T P V L G A G I T E A L M R R F I N K G S F M T S R I N W A I Q S S G V				
Sku	F N K L E S I A E Q E I P K T P V L G C G I T Y S L M K K N L R A N S F L P S R I N W A I Q S S G V				
Consistency	* * 7 * * 6 7 * 7 9	7 4 9 6 * * * * 7	8 * * 4 7 * 6 5 6 5	8 5 5 6 6 8 8 6 * *	9 * * 7 * * * * * *

$\gamma 5$

	1060	1070	1080	1090	1100
Sc	D Y L H L L C C S M E Y I I K K Y N L E A R L C I S I H D E I R F L V S E K D K Y R A A M A L Q I S				
Spa	D Y L H L L C C S M E Y I I K K Y N L E A R L C I S I H D E I R F L V S E K D K Y R A A M A L Q I S				
Sba	D Y L H L L C C S M E Y I I K K Y S L D A R L C I S I H D E I R F L V S E K D K Y R A A M A L Q I S				
Cgl	D Y L H L L C C S M N Y L I S M Y N I E A R L I I S I H D E I R Y L V A D K D K Y K L A L A L Q I S				
Kwa	D Y L H L L C C S M N Y L I K K Y S L K A R L S I S I H D E I R Y L T A E E D K Y R T A M A L Q I S				
Debaryomyces_ha	D Y L H L L I V A M E F L L K K F R I D G R L I T V H D E L R Y M V K S E D K Y K T A L L L Q I S				
Ncr	D Y L H L L I I A M D Y L T R R F N L A C R L A I T V H D E I R Y L A E E P D K Y R V A M A L Q I A				
Ego	D Y L H L L F C S M H Y L I E K Y R L R A R V C I S I H D E L R Y L V A N E D R Y R V A L A L Q V A				
Klactis	D Y L H I L C C S M N Y L I E K Y S I D A R L M I S I H D E I R Y I V S E K D K Y R I T L A L Q I S				
Gze	D Y L H L L V V A M D Y L I R R F N I D A R V A I T V H D E I R Y L V T E K D K Y R T A L A L Q V A				
Pan	D Y L H L L I V S M D Y L T R R F N I A C R L A I T V H D E I R Y L V Q E H D K Y R A A M A L Q I S				
Mgr	D Y L H L L I V S M D Y L M R R F N I D A R L A I T V H D E I R Y L V K D H D K Y R A A M A L Q V A				
Hca	D Y L H L L I I S M D F L I R R F N L D A R L A I T V H D E I R Y L V K E H D K Y R A A M A L Q V A				
Calb	D Y L H L L I I S M E Y L I E K Y G I D A R L A I T V H D E L R Y L V K E T D K Y K A A L L L Q I S				
Sca	D Y L H L L C C S M N Y L I L K Y G L D A R L T I S I H D E I R Y L S S E K D K Y K V A M A L Q I S				
Ani	D Y L H M L I I A M D Y L I R R F N I Q A R L A I T V H D E I R Y L V K E Q D K Y R A A L A L Q V S				
Yarrowia_lipoly	D Y L H M L V V S M E Y L A A K Y D I D M R L C I T V H D E L R F L V K N E D A Y R A A M A L Q V S				
Ppa	D Y L H L L I I S M D Y L I K L F D I D A R L C I T V H D E I R Y L V K E E D K F R A A Y A L Q I S				
aspfum	D Y L H L L I I S M D Y L I R R F N I D A R L A I T V H D E I R Y L V K E E D K Y R A A M A L Q V A				
Sku	D Y L H L L C C S M E Y I I K K Y N L D A R L C I S I H D E I R F L V S E K D K Y R A A M A L Q I S				
Consistency	* * * * 9 * 5 6 8 * 6 9 9	7 6 7 8 6 8 6	7 * 9 5 * 7 9 * * *	9 * 8 9 8 5 7 5 * 9	9 8 6 9 8 8 * * 9 8

$\gamma 5$

Pol III

$\gamma 6$

	1110.	1120.	1130.	1140.	1150		
Sc	NIWTRAMF	CQ QMGI	NELPQN	CAFFSQVDI	D SVIRKEVNMD	CITPSN	KTAI
Spa	NIWTRAMF	CQ QMGI	NELPQN	CAFFSQVDI	D SVIRKEVNMD	CITPSN	KIAI
Sba	NIWTRAMF	SQ QMGI	NELPQN	CAFFSMVDI	D SVMRKEVNMD	CITPSN	KTAI
Cgl	NLWTRAMF	AE QMGI	RELDPQN	CAFFSAID	D KVLNRKEVDLD	CVTPSN	EIPI
Kwa	NLWTRAI	FCE QMGI	RDLPQN	CAFFSAVDI	D HVMRKEVDMD	CITPSN	SKPI
Debaryomyces_ha	NLWTRAMF	CE QLGI	KEVPQS	CAFFSEVDI	D NILRKEVTL	CVTPSH	SNSI
Ncr	NLWTRVMF	AQ QVGI	QDLPQS	CAFFSAVDI	D HVLRKEVDMD	CITPSN	PIPI
Ego	NLWTRAI	FCE QI GI	DDLDPQN	CAFFSAVDI	D HVLRKEVDMD	CVTPSN	QTPV
Klactis	NLWTRSYF	SE QMCI	NDLPQN	CAFFSAVDV	D RVMRKEVDMD	CITPSN	PNAI
Gze	NLWTRAMF	SQ QVGI	NDLPQS	CAYFSAVDI	D HVLRKEVDME	CITPSH	QTPI
Pan	NLWTRAMF	AQ QVGI	HDLPQS	CAFFSAVDI	D TVLRKEVDMP	CITPSN	PIPI
Mgr	NLWTRAMF	AE QVGI	RDLPQS	AAYFSAVDI	D HVLRKEVDMD	CITPSH	STPI
Hca	NLWTRAMF	SQ EMGI	DDLDPQS	CAYFSAVDI	D TVLRKEVDMD	CVTPSH	PNKI
Calb	NIWTRAMF	CE QLGI	KEVPQS	CAFFSEVDI	D HVLRKEVSMD	CVTPSN	PTPI
Sca	NIWTRAI	FCE QMGI	DDLDPQN	CAFFSAVDI	D HVMRKEVNMD	CITLNSN	KVAI
Ani	NVWTRAMF	SQ QVGI	NDLPQS	CAYFSQVDI	D HVLRKEVDMD	CVTPSH	PHKI
Yarrowia_lipoly	NIWTRGMF	CQ QSGI	DDIPQG	CAFFSAVDI	D HVLRKEVDFD	CITPTH	PDPI
Ppa	NLWTRAMF	CQ QLGI	NEVPQS	CAFFSAVDL	D FVLRKEVDLD	CVTPSN	PHPI
aspfum	NVWTRAMF	SQ QVGI	DDLDPQS	CAYFSAVDI	D HVLRKEVDMD	CVTPSH	PDPI
Sku	NIWTRAMF	VN KWE	-	-	-	-	-
Consistency	8 8 8 8 5 7	9 6 8 9 4 7 7 9 9 6	8 9 7 9 9 6 8 9 8 9	4 8 7 9 9 9 9 6 7 7	9 8 9 8 8 6 4 3 5 8		

γ6

	1160.	1170.	1180.	1190.	1200				
Sc	PHGEALDI	NQ LLDKSN	SKLKG	KPNLDI	D - - - -	SKVSQYAYN	YREPVFEEYN		
Spa	PHGEALDI	NQ LLVKPN	SKLKG	QPNLDI	D - - - -	SEVSQYTYN	YREPVFEEYN		
Sba	PHGEALDI	NQ LLTKPN	SALG	NPNVEI	D - - - -	SQVSQYPYT	FREPVFEEYN		
Cgl	PHGESIDI	VK LLERPE	CHLE	LPRETV	D - - - -	LSKFGFK	KREPVFVKYN		
Kwa	PHGESIDI	LR LLELP	ESTLK	EASEEI	D - - - -	VSGMPYQ	PREPVFSRQN		
Debaryomyces_ha	PPGESLDI	HK ILEKCK	DEKF	MIGAPRK	- - - -	KASKLSKIEYC	NRDRVISNLD		
Ncr	AHGESIDI	FQ ILEKGD	DAKL	DDSI	VPQS	QYAPRLENI	PYTPRVPMQRLR		
Ego	PHGERVTI	TD LLARP	DAALN	NPNPAVN	- - - -	VADYPVV	QRQPVLAYYN		
Klactis	PHGESLDI	LD ILKIN	GYL	EPTDAVD	- - - -	VSQFPYR	HRESVFRQYN		
Gze	PPGESIDI	NT LLEKGE	VARL	DPAI	VPDAKH	APKLDHIA	YTPRVPIMEGIR		
Pan	PPGETVDI	NQ LLSKGS	LARL	DPTITP	DPTY	APRVNHIKYV	PRKPVMEQLQ		
Mgr	PHGESLDI	KT LLEKGD	AAFL	DPAI	VPRAHA	PNFDDGIEY	TPRKSVMTLRL		
Hca	PHGESLDI	EN LLKGF	EAQL	DPSI	EPTSL	-	PTPEKYHYTPRKPVLSLQ		
Calb	PLGESLNI	GQ LLLKQ	CHGDI	LADKNT	KP - - -	LTLRTTKYS	PRQPIINQLD		
Sca	PHGENLDI	NK LLSME	GSRLE	SSDAI	DT - - -	VDVSSLPYK	YREPVFSQYN		
Ani	PHGESLNI	TQ LLDKN	DAARL	DPSI	TPI	SR -	PTPEKYTYTPRKTVMSALQ		
Yarrowia_lipoly	PHGESLDI	EG LLARK	DARLE	DATHVI	DHSI	DKKAQKFPY	TSRTPVLAALAE		
Ppa	PCGKSLDI	YQ LLQ	QEDI	KGA	DFPRT	MHLN	- - - -	DVHYR	KRTPVIMFD
aspfum	PHGESLDI	SQ LLDKQ	QEAYL	DPSV	VPVSP	-	PTPQKYTYTPRESVMATLQ		
Sku	-	-	-	-	-	-	-		
Consistency	8 5 9 8 6 7 6 9 3 4	8 9 4 6 3 4 3 4 3 3	4 4 4 3 3 3 4 1 0 0	0 2 2 4 4 4 4 3 8 4	3 9 5 6 8 5 4 3 5 5				

	1210	1220	1230	1240	1250	
Sc	KSYTP	EFLK	YFLAMQV	QSDKRDVN	LEDE YLRECTSKEY	
Spa	RSYTP	EFLK	YFLAMQV	QSDKRVN	LEEE YLRECAKEY	
Sba	KTYTE	EFLK	YFLAMQV	QSDKRVN	LEEE YLRECTSKEY	
Cgl	NNYSA	DFLK	YFLKMQI	QTTSWKVS	LEEE YIQNTEKDM	
Kwa	KAYSK	DFLH	YFLRMQI	QSSKWKVD	LEQE YLRHSNEKGH	
Debaryomyces_ha	KELNT	DMKV	AKI KLQN	SVDKTAWR	LN INNYI NIKKH IEFDKFNNEF	
Ncr	ERAEA	GDHQ	AFLRFI RAQI	TNSDEELK	RIAE TRYSDPYGAF	
Ego	STYDA	AFLN	FFLDMQI	QDSKASVA	DIKHYREFLLRDAF	
Klactis	KRYSK	DFLR	YSLAMQI	QDTKWKID	EIEKSYLRAMEAKF	
Gze	EAA	Q	ESI PFL KAQI	ASDEKEIR	DIVYA LQKERLAAMP	
Pan	EEHHA	HEDEK	SKLRFI RAQI	ANDEAEFR	EILRE TRGEAQRKY	
Mgr	ETDP	ST	ADLSFI KAQI	TADDAELK	EILAQ KRS	
Hca	SS	RD	PNYI KAQI	CNDKELR	EIKG MTMKSSQPAS	
Calb	KNLT	IA	EKI AKI ELQT	SIDKDEWK	KNLYN LAKVSRNKQN	
Sca	RAYT	KD	FWQYFLRMQV	QDSKWKVD	GLETE YIRLKTSNDF	
Ani	TINN	LAFI	KAQI	TKDDKELR	DIKDVTKLNNPVNA	
Yarrowia_lipoly	S	N	PS	KVI SSSL DQI	GGDKRAIR	KLSKL
Ppa	KAVD	DK	TRKFMVSLQI	AQDKTEFT	KWKRS	RGELIVH
aspfum	ATNN	PAFI	KAQI	TKDDKELR	EIKE	VTKTKAATSS
Sku						
Consistency	5 4 3 3 1 0 1 1 2 3	0 0 1 3 5 7 4 5 9 7	4 4 6 5 3 5 5 4 0 0	0 0 0 0 0 4 6 4 4 4	3 3 3 2 2 2 3 2 2 2	

	1260	1270	1280	1290	1300
Sc	ARDGNTAEYS	LLDYI KDVEK	GKRTKVRIMG	SNFL DGTKN	AKADQRI RLP
Spa	ARDGNTTEYS	LLDYI HDVGK	GKRTKIRIMG	SNFL DGTKN	AKIDQRI RPP
Sba	ARDGNTAEYS	LSDYLDQVKK	GKRAKIRIMG	SNFL DGTEEN	ANTSRRTEPS
Cgl	TNE DRLEYG	FLDYLSDVKK	GKRDKLTIMS	KAFI SDALSS	RGLDLDAQLN
Kwa	RKK RVEEYL	FI DYMQDTKE	GKHPKINIMS	ESPP NTAPI	VELQNEGV DV
Debaryomyces_ha	DRRNTMFEKA	PALNANARAV	PKPMKSKPMK	RKIEEYDIG	SNGDELLAKE
Ncr	SLASNGRVSG	NPHQRHA AVH	ASTKTA AAPS	KPSI ASRFDS	VSQASRI KSV
Ego	SADRAETVTM	ADYL AANGG	GKLSINAI DT	TTL	
Klactis	RLVDNEVMTM	ADYIRGNTPD	GKPKKQSI MS	DEVMYLKEND	IFDI DAELGK
Gze	KKERPMSKKE	LQSVLPYRAY	PQLM PMDE	PFI LTEAMKQ	NQNA YAHKNN
Pan	IPKVPVKI LK	KPYASAKLL	PGAARYGQRE	PVMVREALEG	VRYKSWPGGS
Mgr	ESVPPKPKA	KQAMVPYHSH	AQLM EEPL	SVTDALGARF	GTGPSFAAKG
Hca	NSSSSVTKST	TSYARSRRTK	STIPPHAEPQ	RAILMEVQPA	LLNGSTKQAA
Calb	GKNNSTE	KPQRV	IKRKPI VEYD	LQEDNNKKRE	IKYK KDPKPS
Sca	YKHGYTSEYG	LLDYLDQVKK	GKRRKKSIME	NSFLQDLIEV	ESSDHGDI LP
Ani	KCTVPRSDTR	AKATAKPVAE	PQKAILMDI G	SGLYGGGVRD	FTQGSRRPPQV
Yarrowia_lipoly					
Ppa					
aspfum	TARSRFRNGSK	STN FPAAE	PQKAILMDI D	SGLY RDFRY	LPQGSRRHPPF
Sku					
Consistency	2 3 2 2 2 3 3 2 2 2	2 2 2 2 2 2 2 3 2 2	3 4 2 2 2 2 2 3 3 3	3 2 3 3 1 2 2 2 2 2	2 2 2 2 2 2 1 2 2 2

	1310	1320	1330	1340	1350	
Sc	VNMPDYPTLH	KI AND	SAIPEKQLLE	NRRKKENR	I DDE	
Spa	VNMPDYTSLH	KI AND	SAI PKEQLLK	NKGKKENGVV	LASE	
Sba	ANMLDYSSLH	KI AND	SIIPQEQLVK	DKEREKKEV	SIDE	
Cgl	LDPSKTEI H	NI TSD	DKSRL	QAQEDFSLH	DSVSTLEHI D	TKSPTREKAT
Kwa	SGAANTTRRP	QNEI	VLSHD	SGMDQLKDLE	FQHMLATNND	QI
Debaryomyces_ha	GKPDSDNRNE	NANKAI	YKPT	IKKSTPNI KK	DNKSCPQAS	KMASQE
Ncr	AAGSDEPTIR	ATKAQG		KAMAKASGK	LAASKDT	
Ego		APPQPQ		QRKPASRRKP	RVI	
Klactis	KSKEKKNRSF	KKLAAK		HKLPSIDNVS	NDIADI VTK	AVN
Gze	KL G	LG	KYSNG			
Pan	GS GSKKWDG	GRENG		RERVQWDFTR		
Mgr	KWG	WERFNG	PRAGG		PSRAPTRLG	
Hca	NFVKS PHFSA	HRAVW		KPRPSPGP		
Calb	KIATTTTTTT	KKK		ASSNNSKTKS	NGKSDS	
Sca	ENELDTENYS	ILHA		IASERP AVVA	K	
Ani	NLHRQPWKPR	PTARA				
Yarrowia_lipoly						
Ppa						
aspfum	NINRQTWKPR	PAARP				
Sku						
Consistency	2 2 2 1 2 2 1 2 2 2	2 2 3 2 1 0 0 0 0 0	2 2 2 2 2 1 2 2 1 1	1 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	

	1360	1370	1380	1390	1400	
Sc	NKKKLTRKKN	TTPMERKYKR	VYGGRKAFEA	FY	ECANKP	
Spa	NKKKLTRKKN	TTPMERKYKR	VYGGRKAFEA	FY	ECANKP	
Sba	NKRTLTKKKK	VTPMEWKYKR	VYGGRKAFEA	FY	ECANKP	
Cgl	TISTASLEIN	TNSKKGRRRT	GRLLEKYNK	VI GGKNAFDA	FY	ASIDQP
Kwa	NKKPTTKRVL	KSTRRTVAKR	VI GGSAA YRA	FD	ECMNMP	
Debaryomyces_ha	NSPVRSFS	STNYKHHTNN	HKALGSSTNA	SGRKAPSFKP	YKQNVSVVVP	
Ncr	VLNVTIKKK	VAAPEMAAVP	STSSSES SKA	SATTSTTTT E		
Ego	NRSHRI A	IGGPAAFDAF	HASVDLAEED	A	DLR	
Klactis	DTGNGKEKLS	SSNSKIKTRK	LKQKRLDSNK	SI KSNKSLK		
Gze						
Pan						
Mgr						
Hca						
Calb	TSAMTTFDL	KTDPQRP	TYNRRKK	RL F	EKP	
Sca						
Ani						
Yarrowia_lipoly						
Ppa						
aspfum						
Sku						
Consistency	0 0 0 0 0 0 0 0 0 0	0 0 0 0 1 1 1 0 0	0 1 0 0 1 0 1 0 0 0	1 0 1 1 1 1 1 0 1 1	1 0 0 0 0 0 0 1 1 1	

	1410.	1420.	1430.	1440.	1450
Sc	LDYTLETEKQ	FFNIPID---	GVIDDVLND	KSNY- KKKPS	QARTASS--S
Spa	LDYTLETEKQ	FFNIPID---	GIIDDVLSD	KSGY- KKRPS	QAGTASP--S
Sba	LDYTLETEKQ	FFNIPID---	GVVDDVLND	KTDY- RRRPS	RARTTSS--S
Cgl	KKFVELQEQR	IIN-----	EVIHDIIEG	REDTPTKKPR	QSKPKST--T
Kwa	RAFADKRQRK	LVDAEID--	KSVVDEVVTN	KCAT- QKKT K	NNRPSST--G
Debaryomyces_ha	NKSQSRLSKL	FFESEPRRKT	DPQVQTTQAK	PNSPPSRQPS	YTRSGSYRT
Ncr	NATASPSS-	-----	SNVDA	KKTTSKTKPT	HKKETEG--E
Ego	---AAAI-	-----	---	TRNYLRDRPH	L-----
Klactis	VNTTIGSDC-	-----	---	RDNYAQRKRA	FNVAI----
Gze	-----	-----	-----	-----	-----
Pan	-----	-----	-----	-----	-----
Mgr	-----	-----	-----	-----	-----
Hca	-----	-----	-----	-----	-----
Calb	QAK-----	-----	-----	-----	-----
Sca	-----	-----	-----	-----	-----
Ani	-----	-----	-----	-----	-----
Yarrowia_lipoly	-----	-----	-----	-----	-----
Ppa	-----	-----	-----	-----	-----
aspfum	-----	-----	-----	-----	-----
Sku	-----	-----	-----	-----	-----
Consistency	0 0 0 0 0 0 1 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	1 0 0 0 0 1 1 1 1 0	0 0 0 0 0 0 0 0 0 0

	1460.	1470.	1480.	1490.	1500
Sc	PIRKTAKAV-	HSKKL PARKS	STTNRNLVEL	ERDI T I SREY	-----
Spa	PIRKTAKAV-	HSKKL PASKP	DTTKRSL LGL	ERDI T I SRGY	-----
Sba	T R K A A K T A	PPKKI PARKP	NTKEMGVLEL	ERDNTLPKGN	TQKLH-
Cgl	PVPRAKQRT-	HPF-----	---TENLVKG	STHLTK-	-----
Kwa	NTMPRQGRS-	TPETKI SFKP	ATHERRTGAN	TS-----	-----
Debaryomyces_ha	PI QQP KDP TK	KPATSYLANS	SAKRKKSADT	YTQRHLFNRA	STFI PGNGI M
Ncr	PFPSLDDPVI	AARLEAVSKT	SPGTRASVAA	KLDALASFCH	ASCC- GC----
Ego	-----	-----	-----	-----	-----
Klactis	-----	-----	-----	-----	-----
Gze	-----	-----	-----	-----	-----
Pan	-----	-----	-----	-----	-----
Mgr	-----	-----	-----	-----	-----
Hca	-----	-----	-----	-----	-----
Calb	-----	-----	-----	-----	-----
Sca	-----	-----	-----	-----	-----
Ani	-----	-----	-----	-----	-----
Yarrowia_lipoly	-----	-----	-----	-----	-----
Ppa	-----	-----	-----	-----	-----
aspfum	-----	-----	-----	-----	-----
Sku	-----	-----	-----	-----	-----
Consistency	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0

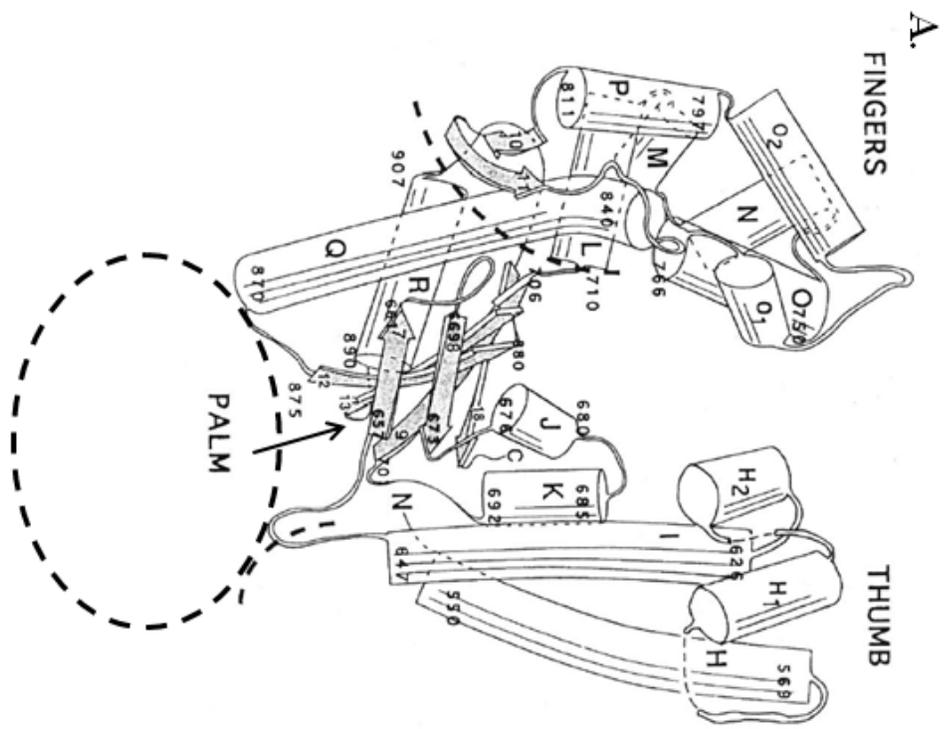
Sc	. .
Spa	- -
Sba	- -
Cgl	- -
Kwa	- -
Debaryomyces_ha	K H
Ncr	- -
Ego	- -
Klactis	- -
Gze	- -
Pan	- -
Mgr	- -
Hca	- -
Calb	- -
Sca	- -
Ani	- -
Yarrowia_lipoly	- -
Ppa	- -
aspfum	- -
Sku	- -
Consistency	0 0

1.3.3 Molecular modeling of pol- γ catalytic subunit

Currently, an X-ray crystal structure of a mtDNA polymerase catalytic subunit does not exist. However, as mentioned above, various crystal structures of family A-type polymerases bound to DNA have been solved (Doublet *et al.*, 1998, Eom *et al.*, 1996, Kiefer *et al.*, 1998, Beese *et al.*, 1993). Additionally, one study developed a homology model of the human mtDNA polymerase catalytic domain in complex with DNA (Graziewicz *et al.*, 2004). All of these studies suggest that the molecular architecture of a polymerase domain is conserved. All of the polymerase domain crystal structures, and the homology model of the human γ -polymerase domain, resemble a “right-hand” with “fingers”, “thumb” and “palm” subdomains (Figure 1.2 A. Joyce and Steitz, 1995, Kaguni, 2004, Graziewicz *et al.*, 2006). A cleft is formed between the polymerase and the exonuclease domains of the Klenow fragment and has been shown to bind duplex DNA in a DNA-Klenow editing complex (Figure 1.2 B., Beese *et al.*, 1993).

The palm subdomain is the location of the polymerase catalytic site. The polymerase active site harbours three conserved carboxylates and other polar residues at the base of the polymerase cleft. In the human γ -polymerase homology model these residues are equivalent to D1135 (Klenow D882), E1136 (Klenow E883), and D890 (Klenow D705), which are located in the Pol A motif (D890) and the Pol C motif (D1135 and E1136) and are absolutely conserved among mtDNA polymerases. A mechanism of phosphoryl transfer for catalysis of the polymerase reaction has been proposed. This reaction involves nucleophilic attack by the 3' hydroxyl of the primer on the dNTP α -phosphate, with the subsequent release of PPi (Figure 1.3, Joyce and Steitz, 1995).

Figure 1.2 Klenow fragment of *E. coli* DNA polymerase I. **A.** The Klenow fragment polymerase domain. The structure of this domain resembles a “right-hand” with “fingers”, “thumb” and “palm” subdomains. α -helices are depicted as cylinders lettered from the fragment N-terminus (starting at H) and β -strands are depicted as arrows. The exonuclease domain is not depicted to clearly delineate the polymerase domain. The approximate position of the exonuclease domain is represented by a dotted circle. **B.** Space filled model of the Klenow fragment with bound duplex DNA. Notice that the DNA is bound in a cleft between the polymerase domain and the exonuclease domain. The 3'-end of the primer strand is bound in the exonuclease active site (All panels were taken with permission from the following sources: panel A was taken from (Joyce and Steitz, 1995), panel B was taken from (Polesky *et al.*, 1990).



B.

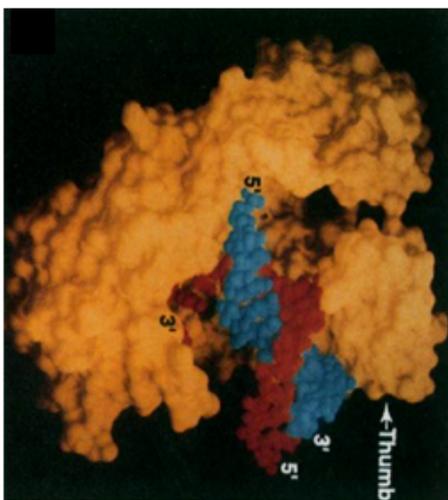
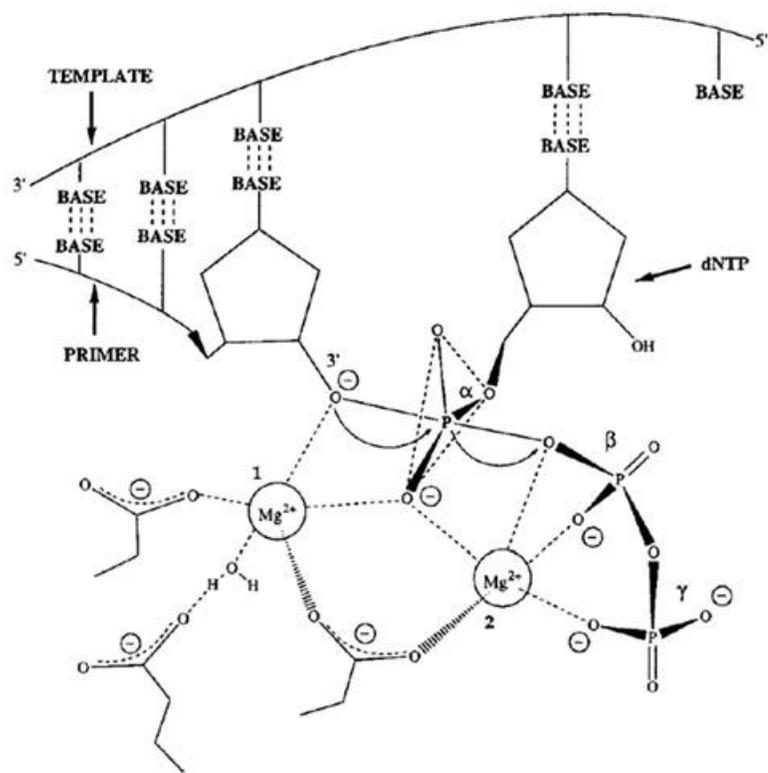


Figure 1.3 Pentacovalent transition state in the mechanism of the 5' to 3' DNA polymerase reaction. In this model catalysis is mediated by two divalent magnesium ions (Mg^{2+}). Carboxylate ligands (COO^-) are in generic positions as not to represent any particular DNA polymerase. In Klenow the side chains harbouring these conserved carboxylates are D705, D882, and E883 (Taken with permission from (Joyce and Steitz, 1995)).



In this model the three conserved carboxylates anchor two magnesium ions. One magnesium promotes deprotonation of the primer strand 3' hydroxyl while the other facilitates formation of the pentacovalent transition state at the α -phosphate of the dNTP and the release of PPi. Crystal structures of DNA polymerases suggest that as the nascent primer strand elongates via DNA polymerase activity, the 3'-end can shuttle between the polymerase active site and the exonuclease active site (Figure 1.4, Joyce and Steitz, 1995).

As was mentioned above a homology model of human pol γ catalytic domain in complex with DNA was developed based on the sequences of solved family A polymerases (*Thermus aquaticus* DNA polymerase I, *Bacillus stearothermophilus* DNA polymerase I, *E. coli* DNA polymerase I Klenow fragment, and T7 DNA polymerase) and was shown to display well-defined fingers, thumb, and palm subdomains (Figure 1.5, Graziewicz *et al.*, 2004). This model has proven to be a useful tool in predicting the potential effect of disease mutations as is the case with the two autosomal dominant mutations that cause PEO encoding the amino acid substitutions Y995C and R943H. The pol γ model predicts that the R943H mutation would disrupt the interaction between the R943 side chain and the two oxygen atoms of the γ -phosphate of the incoming dNTP thereby interfering with the nucleotidyl-transfer reaction (Figure 1.5). *In vitro* this substitution caused a decrease to 0.2% of wild-type polymerase activity and replication stalling. The Y995, E895, and Y951 side chains of pol γ form a hydrophobic pocket that harbours the incoming dNTP. The Y955C mutation is predicted to significantly disrupt this interaction and *in vitro* the Y955C enzyme exhibits the most severe biochemical defects with respect to polymerase activity, processivity, and incorrect nucleotide

Figure 1.4 Cartoon of the Klenow fragment of *E. coli* DNA polymerase I shuttling between editing and polymerization modes. “P” indicates the polymerization active site while “E” indicates the exonuclease active site (Taken with permission from (Joyce and Steitz, 1995)).

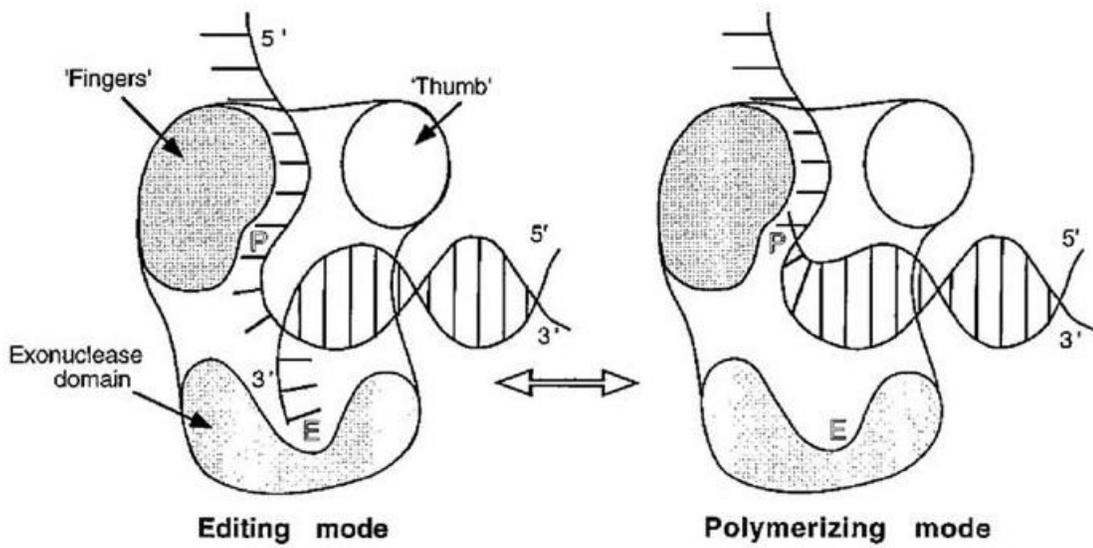
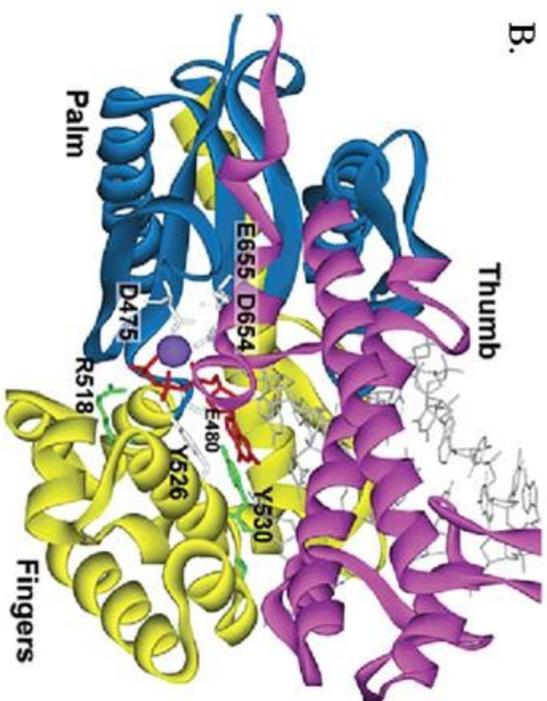
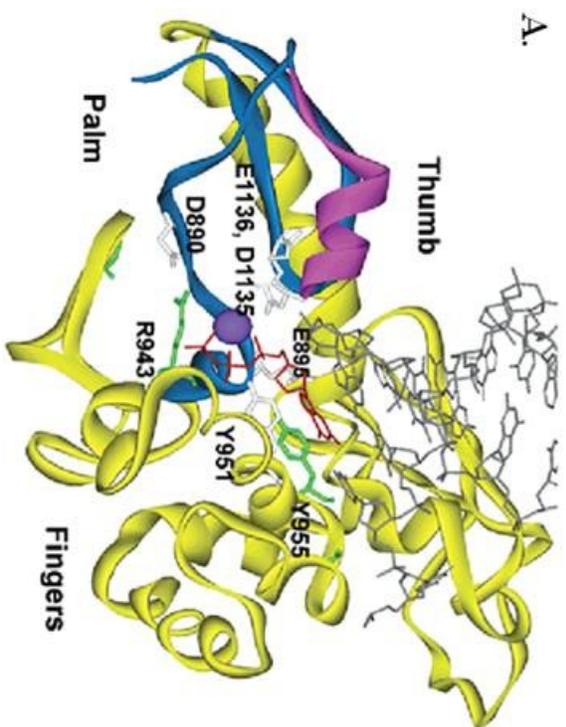


Figure 1.5 Homology model of human DNA pol γ polymerase domain. **A.** Model of residues 871 to 1145 of the human pol γ polymerase domain. Side chains of residues that are mutated in autosomal dominant PEO are coloured green. The side chains of other residues involved in catalysis are depicted in white. Residues D890, D1135, and E1136 harbour the conserved carboxylates suggested to be required for polymerase reaction (Figure 1.3). **B.** The T7 DNA polymerase ternary complex. In both panels, the thumb, fingers and palm subdomains are magenta, yellow and blue, respectively. The template and primer strands are coloured gray, the incoming ddGTP is coloured red and the magnesium ion is shown as a purple sphere.



selection (Graziewicz *et al.*, 2004). PEO patients with the Y995 mutation also exhibit the most severe clinical phenotype.

1.4 Yeast mitochondrial DNA replication

No definitive model of mtDNA replication exists for either yeast or man. In yeast, one model is that of the bidirectional origin model (Lecrenier and Foury, 2000), while other models of rolling circle replication and recombination-dependent replication via DNA priming of 3'-invading DNA have been proposed (Lecrenier and Foury, 2000, Maleszka *et al.*, 1991, Shibata and Ling, 2007). For human mtDNA replication, two models have been proposed whereby mtDNA replication is either 1) asymmetric (strand displacement model) or it is 2) symmetric, semidiscontinuous DNA replication with coupled leading and lagging strand DNA synthesis (strand-coupled model, Graziewicz *et al.*, 2006). Only the yeast mtDNA replication models will be discussed here.

1.4.1 Yeast mitochondrial DNA

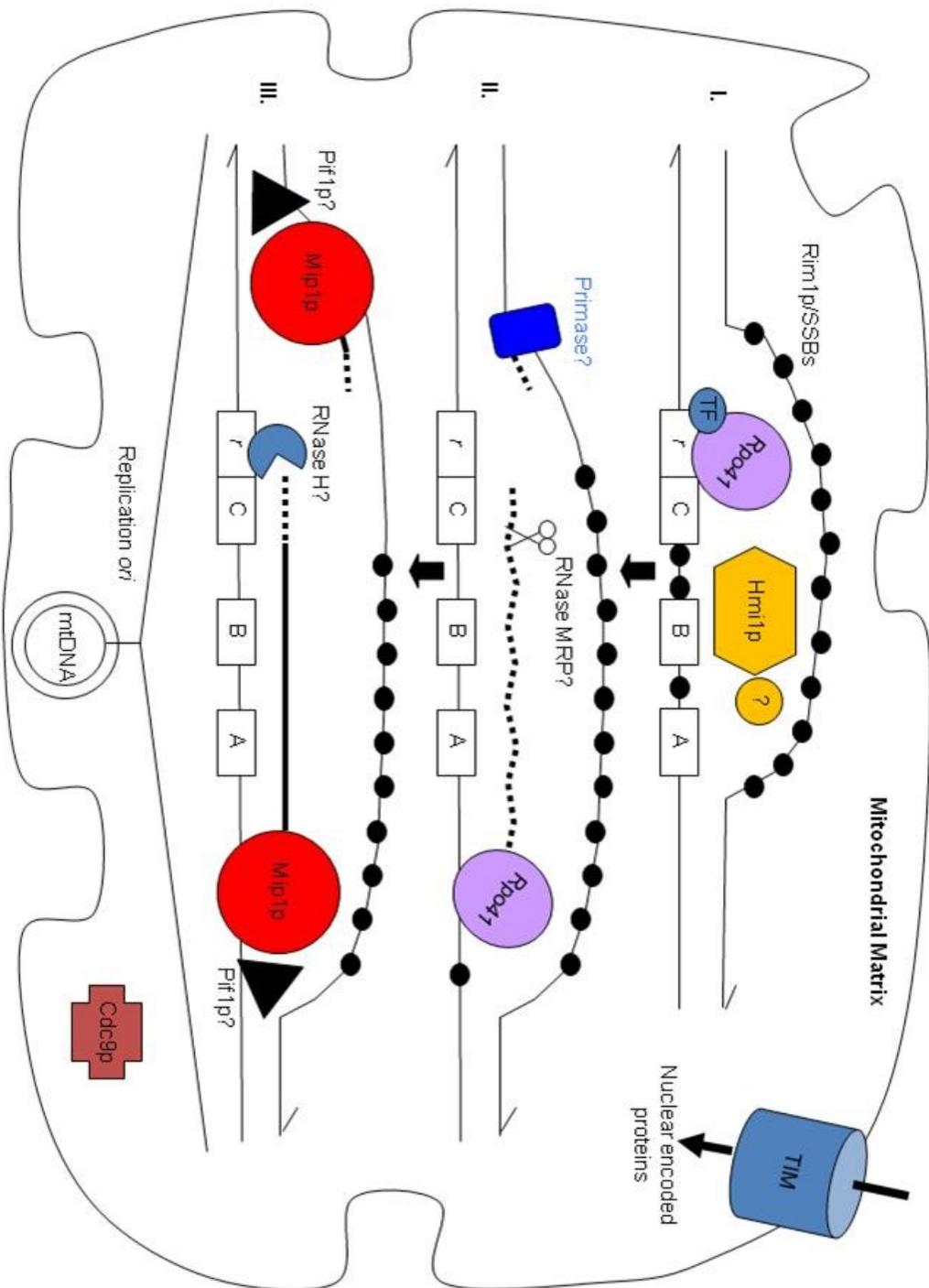
The yeast mitochondrial genome has been sequenced from the laboratory strain FY1679 (isogenic derivative of S288c) and has been determined to be 85.8-kb in size (Foury *et al.*, 1998). Unlike the intron-less human mtDNA, the sequenced yeast mtDNA contains 13 introns some of which are “optional” or present in some strains but absent in others (ex. the Omega intron, Lambowitz and Belfort, 1993). The coding capacity of yeast mtDNA is similar to that of human (the main difference being yeast mtDNA does not encode components of complex I). The yeast mitochondrial genome encodes three subunits of cytochrome c oxidase complex IV (*COX1*, *COX2*, *COX3*), three subunits of the ATP synthase complex V (*ATP6*, *ATP8*, *ATP9*), apocytochrome *b* (*COB*) subunit of complex III, a ribosomal protein (*VARI*), several intron-related ORFs, large (21S) and

small (15S) ribosomal RNAs, 24 tRNAs that can recognize all codons, and a 9S RNA component of RNase P (Foury *et al.*, 1998). Yeast mtDNAs also contain seven or eight origin-like (*ori*) elements, of which three or four are thought to be active bidirectional sites of DNA replication initiation (Lecrenier and Foury, 2000). Additionally, these *ori* elements have been proposed to function with Hsp60 in nucleoid division, which is likely dependent on membrane association of mtDNA (Kaufman *et al.*, 2003).

1.4.2 Yeast mitochondrial DNA bidirectional mtDNA model

Hmi1p is a mitochondrial-localized helicase required for mtDNA maintenance; however, its helicase activity is not essential for maintaining the mitochondrial genome (Monroe *et al.*, 2005). It has been suggested that Hmi1p may be required to bind DNA and assemble replication proteins on either a collapsed replication fork or at the origin of replication (*ori*). The mtDNA *ori* sequences are approximately 300 bp long and are formed by three GC clusters separated by AT stretches (Figure 1.6). Transcription of the *r* strand is hypothesized to be initiated from the *r* promoter located upstream of the three GC clusters, by the mitochondrial transcription apparatus (Rpo41p, RNA polymerase, and Mtf1p, transcription factor for promoter binding and specificity of Rpo41p). An RNase MRP activity would then process this transcript into a primer for initiation of non-*r* strand synthesis using the *r* stand template via Mip1p (Lecrenier and Foury, 2000, Graves *et al.*, 1998). RNase MRP is a site-specific endoribonuclease found in yeast and man. In mammals a portion of RNase MRP has been shown to localize in mitochondria, but this has not been shown in fungi. *In vitro* RNAs initiated at mtDNA *oris* from yeast and mammals were cleaved by RNase MRP and could then be extended by DNA polymerase

Figure 1.6 The bidirectional model of the yeast mitochondrial DNA replication fork. **I.** Hmi1p may bind to the replication origin (*ori*) and assemble replication proteins (circle with a question mark). A, B, and C represent the three GC clusters separated by AT stretches and *r* represents the *r* promoter. The mitochondrial transcription apparatus (Rpo41p, mitochondrial RNA polymerase and Mtf1p, TF) is shown bound to the *r* promoter. Single-stranded DNA binding proteins (SSBs, Rim1p, depicted as black dots) bind to and stabilize single-stranded DNA during mtDNA replication. **II.** The mitochondrial transcription apparatus extends a new RNA that would be cleaved by an RNase MRP-like activity to form a primer for the mitochondrial DNA polymerase, Mip1p. Likewise, a primase activity on the non-*r* mtDNA strand could form a short primer for Mip1p. **III.** Separate Mip1p enzymes extend the primers on the *r* and non-*r* strands while Pif1p may unwind the replication fork ahead of mtDNA synthesis in both situations. Before the polymerase reaches the 5'-end of a primer a 5'-3' RNase H would have to degrade the RNA primer and Mip1p would fill in the remaining DNA. Mitochondrial Cdc9p ligase would then seal the nick completing replication (modified from Lecrenier and Foury, 2000).



activity (reviewed in Lecrenier and Foury, 2000). These observations suggest an involvement of RNase MRP in mitochondrial primer generation.

The Pif1p mtDNA helicase has been shown to associate with mtDNA and either protects mtDNA from double-stranded breaks or participates in the repair of double-stranded breaks. Pif1p has 5'-3' DNA helicase activity and may recognize certain DNA secondary structures in the regions that form during mtDNA replication (Cheng *et al.*, 2007). Therefore, Pif1p may be the replicative helicase that unwinds mtDNA ahead of the replication fork although this function has not been proven.

Primer synthesis on the non-*r* strand has been hypothesized to be generated by a primase activity upstream of the *r* promoter (Lecrenier and Foury, 2000). Subsequently, Mip1p would use this primer to replicate a new *r* strand in the opposite direction as non-*r* strand synthesis (Figure 1.6). Before Mip1p reaches the 5'-end of the RNA primer a 5'-3' RNase H would have to degrade the RNA primer. Mip1p could then replicate mtDNA up to the first deoxyribonucleotide incorporated in the nascent chain followed by ligation of the first and last deoxyribonucleotides in this chain by the bona fide mtDNA ligase, Cdc9p (Donahue *et al.*, 2001).

1.4.3 Yeast rolling circle mtDNA replication and vegetative segregation

Due to the location of mtDNA near the electron transport chain, which generates reactive oxygen species, mtDNA accumulates oxidative damage at a higher rate than nuclear DNA. One might expect that mtDNA would be heterogeneous (or heteroplasmic) due to the high copy number and mutation rate. However, within a single cell mtDNA copies are typically identical, a state referred to as homoplasmy. In yeast, homoplasmy is achieved by a heteroplasmic cell during mitotic cell division. After ten to

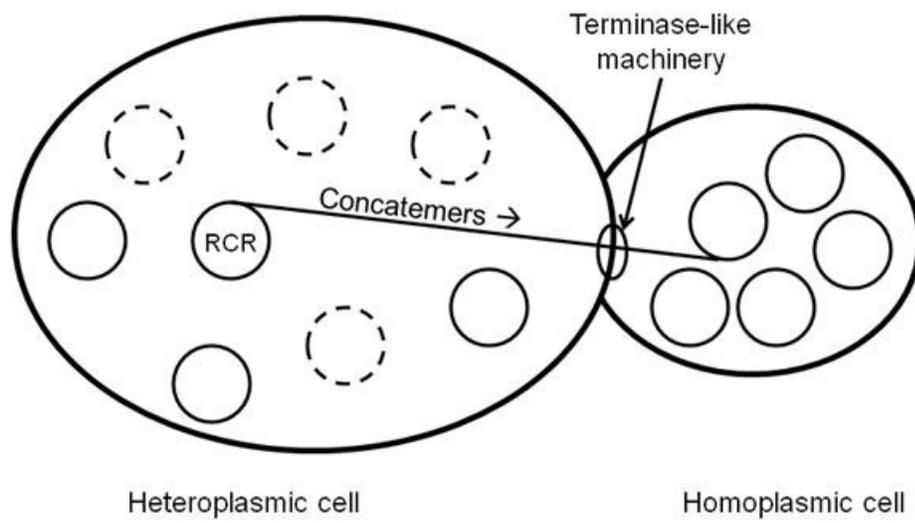
twenty cell divisions most progeny of the heteroplasmic cell are homoplasmic. This process of segregating homoplasmic cells (buds) from a heteroplasmic mother cell during mitosis is defined as vegetative segregation. Recently Shibata and Ling have proposed a model whereby vegetative segregation in *S. cerevisiae* occurs by partitioning of monomeric mtDNA into growing daughter buds via rolling circle mtDNA replication from the mother cell (Shibata and Ling, 2007). Rolling circle replication of mtDNA in various yeasts has also been suggested based on pulse field gel electrophoresis analysis of linear tandem repeats of mtDNA concatamers as well as electron microscopic analysis of mtDNA in a lariat form (Maleszka *et al.*, 1991). Similarly, Bendich observed that mtDNAs isolated from *S. cerevisiae* were linear molecules ranging in size from 60-kb to several hundreds of kilobases (Bendich, 1996).

Analyses of *rho*⁺ and *rho*⁻ mtDNA from mother cells indicated that the major mtDNA species are concatamers whereas monomers were minor species. In contrast, the majority of mtDNA in daughter buds were found to be circular monomers (Ling and Shibata, 2002). Evidence supporting a rolling circle-type replication model, came from a pulse-chase experiment of [¹⁴C]deoxythymidine ([¹⁴C]dTTP) labelled mtDNA during mitosis. Because mtDNA is replicated continuously throughout the mitotic cycle in yeast (Sena *et al.*, 1975), and nuclear DNA is replicated only during the S (synthesis) phase of the cell cycle, mtDNA can be selectively labelled when yeast are arrested at the G1 phase of growth with the mating pheromone α -factor. In cells arrested at G1 [¹⁴C]dTTP was incorporated into mtDNA concatamers at a much higher density than into mtDNA circular monomers. After removal of [¹⁴C]dTTP and addition of excess deoxythymidine, cells were grown for another hour prior to release from G1 arrest and synchronous

growth for 1.5 hours. Then mother cells and buds were separated. The daughter buds of dividing cells contained monomers as the major species at a similar relative density of [¹⁴C]thymidine as the concatamers in the mother cells. This result indicated that concatamers in mother cells are precursors of monomers in buds and supports a rolling circle-type replication model (Figure 1.7 Ling and Shibata, 2004). In this model mtDNA is transmitted into budding daughter cells via concatamers resulting in a number of clonal mtDNA copies within a budding daughter cell. Concatamers are selectively partitioned to daughter cell buds by a putative terminase and are processed into circular unit sized mtDNAs. Selective partitioning of concatamers formed on a limited number of mtDNAs could explain the quick segregation of homoplasmic cells observed in yeast.

Interestingly, mtDNA maintenance seems to be dependent on homologous recombination as double *MHR1 CCE1* deletion mutants are completely devoid of mtDNA (Ling and Shibata, 2002). *MHR1* encodes a mitochondrial matrix protein that pairs single-stranded DNA and homologous superhelical closed circular double-stranded DNA to form heteroduplex joints in the absence of ATP *in vitro*. Homologous pairing and formation of the heteroduplex joint is essential to homologous DNA recombination. *CCE1* encodes a mitochondrial cruciform cutting endonuclease that cleaves Holliday junctions formed during recombination of mtDNA (Kleff *et al.*, 1992). Cells harbouring only a mutant *mhr1* allele show a temperature-sensitive phenotype with respect to maintenance of mtDNA. When cultured at a nonpermissive temperature some cells become *rho*^o but some are *rho*⁻, indicating *rho*⁻ mtDNA can be maintained in the absence of Mhr1p, albeit at reduced levels (Shibata and Ling, 2007). Cells that harbour only a deletion for the active *CCE1*, exhibit little, if any, increase in the generation of *rho*⁻ cells.

Figure 1.7 The rolling circle replication model of yeast mitochondrial DNA. A heteroplasmic mother cell is represented as having two different forms of mtDNA, indicated by solid and dashed circles. A few mtDNAs are randomly selected as templates for rolling circle replication to form concatamers (solid circle). Concatamers are transmitted to the daughter bud by putative terminase-like machinery and monomeric mtDNA circles are formed during transmission resulting in a homoplasmic daughter cell (modified from Shibata and Ling, 2007).



It has been suggested that Mhr1p functions in the initiation of rolling circle mtDNA replication (Shibata and Ling, 2007). Support for this assumption is based on the observations that 1) *mhr1* mutants display a decrease in intracellular concatameric mtDNA while 2) overexpression of Mhr1p increased the intracellular concatamers (Ling and Shibata, 2004). In this model Mhr1p would catalyze the heteroduplex joint formation between a single-stranded region of one mtDNA (a 3' single-stranded tail derived either from a double-stranded breakage or a single-stranded gap) and another intact closed circular double-stranded mtDNA to form a joint molecule. Mitochondrial Pif1p helicase might participate in the stabilization or resolution of such mtDNA recombination intermediates (Cheng *et al.*, 2007). The paired 3' terminus could then act as a primer for mtDNA synthesis of the double-stranded mtDNA. mtDNA synthesis would continue and replace the entire strand, which would eventually be released as a single-strand DNA.

Rolling circle replication could then initiate and continue while the lagging strand could be synthesized as Okazaki fragments (Shibata and Ling, 2007). This is an interesting proposition to explain the quick vegetative segregation of mtDNA in yeast; however, future experimentation will have to be carried out to add support to this model. Future work could address if other proteins, in addition to Mhr1p, are required to initiate rolling circle replication. If there is an initiation complex perhaps chemical crosslinking followed by co-immunoprecipitation with Mhr1p could identify other rolling circle replication initiation factors. Another question is do yeast have a phage-like terminase enzyme, which uses concatamers as substrates to package copies of mtDNA into buds? Perhaps a G1-arrest and release type of experiment could be employed followed by formaldehyde cross-linking to trap terminase-like subunits on mtDNA concatamers,

similar to the experiments done to identify protein constituents of the mitochondrial nucleoid (Chen *et al.*, 2005).

Rolling circle replication is an exciting area of mitochondrial research and may have relevance to understanding human mtDNA replication, which is also not fully understood. MtDNA recombination in human cells has been controversial; however, recent examples of this process have been reported (D'Aurelio *et al.*, 2004, Kajander *et al.*, 2001). Vegetative segregation occurs in mammals during oogenesis, whereby healthy progeny of a heteroplasmic mother are homoplasmic (reviewed in Shibata and Ling, 2007). Therefore, an interesting question arises, do human mtDNAs form concatamers by rolling circle replication? Recently, reconstitution of a minimal human mtDNA replisome *in vitro*, consisting of polymerase γ (catalytic and accessory subunits), *TWINKLE* helicase, and mtSSB (mitochondrial single-stranded DNA binding protein), on double-stranded minicircles consisting of a 90-nt oligonucleotide annealed to a 70-nt ssDNA minicircle, permitted rolling circle DNA synthesis of single-stranded DNA products greater than 15,000 nts long (Korhonen *et al.*, 2004). T-odd bacteriophages generate genome concatamers during DNA replication. Perhaps rolling circle replication to generate concatamers of mtDNA is reminiscent of the T-odd bacteriophage origin of mtDNA replication machinery like the homologous yeast and human mtDNA polymerases and mtRNA polymerases (Shutt and Gray, 2006).

1.5 The mitochondrial nucleoid

Mitochondrial nucleoids are mtDNA-protein complexes that exist in the mitochondrial matrix and are the heritable units of mtDNA (Kaufman *et al.*, 2003). It has been estimated that an average of three to four mtDNA molecules (Miyakawa *et al.*,

1987) and twenty-two different proteins (Chen *et al.*, 2005) are present per mitochondrial nucleoid. Various observations have led to the hypothesis that an inner membrane-bound mtDNA segregation apparatus exists (Meeusen and Nunnari, 2003, Azpiroz and Butow, 1993). First, within the mitochondrial reticulum of the yeast cell cortex, nucleoids are distributed regularly, as visualized by fluorescence microscopy of DAPI-stained punctate mtDNA suggesting discrete anchorage points (Nunnari *et al.*, 1997, Meeusen and Nunnari, 2003). Second, after fusion of two haploid yeasts during mating, punctate mtDNA remained localized to one lobe of the zygotic cell (Nunnari *et al.*, 1997). In contrast, when mitochondrial proteins were labelled with different fluorophores prior to mating, the mitochondrial proteins rapidly diffused throughout the mitochondrial reticulum of the zygote. Therefore, during mating, mitochondria fuse and mitochondrial proteins intermix, whereas mtDNA appear to have a regular distribution within the mitochondrial reticulum that ensures distribution to daughter cells (Nunnari *et al.*, 1997).

Although a bona fide inner membrane component of the mtDNA segregation apparatus has not yet been discovered, an outer membrane component has, Mmm1p. Meeusen and Nunnari have defined this segregation machine as the two membrane-spanning autonomous mtDNA replisome (TMS, Meeusen and Nunnari, 2003). The TMS comprises 1) Mgm101p, a component of the mitochondrial nucleoid (Chen *et al.*, 2005) that is essential for mtDNA maintenance and is required for repair of oxidatively damaged mtDNA (Meeusen *et al.*, 1999) 2) Mmm1p, a trans-outermembrane protein required for mtDNA maintenance that coimmunoprecipitates with Mgm101p and has a role in mitochondrial morphology, 3) potentially Mip1p, as shown by colocalization of Mip1p-GFP with a Mmm1p-dsRed fusion protein and 4) replicating mtDNA, as shown

by association of Mmm1p-GFP and Mgm101p-GFP foci with BrdU foci (Figure 1.8 Meeusen and Nunnari, 2003). Extraordinarily, the TMS components (Mip1p-GFP, Mgm101p-GFP, and Mmm1p-dsRed) persist and remain colocalized in the absence of mtDNA (*rho*^o cells) and TMSs arise through self-replication of existing TMSs, which are then faithfully inherited into daughter buds (Meeusen and Nunnari, 2003). Self-replication was shown by time-lapse analysis of fusion proteins with associated fluorophores in *rho*^o and *rho*⁺ cells. These analyses showed that TMSs arise through self-replication of existing structures which are then faithfully inherited into daughter buds regardless of the presence or absence of mtDNA (Meeusen and Nunnari, 2003). Presumably membrane-bound nucleoids harbour proteins needed for mtDNA transactions such as DNA replication, repair, and recombination. Indeed Butow's group discovered examples of proteins involved in these processes that bind yeast mtDNA (Figure 1.8, Table 1.2, Kaufman *et al.*, 2000, Chen *et al.*, 2005). For example Rim1p, the single-stranded mtDNA binding protein (Van Dyck *et al.*, 1992), Mgm101p, mentioned above for its role in repair of oxidative damage to mtDNA, and Abf2, a high-mobility group protein involved in recombination of mtDNA (Zelenaya-Troitskaya *et al.*, 1998) were all found to be constituents of the nucleoid (Kaufman *et al.*, 2000, Chen *et al.*, 2005). Perhaps the most interesting finding from Butow's group was that most of the protein constituents of the nucleoid are not obviously related to mtDNA transactions. For example Chen *et al.* defined three functional categories of nucleoid proteins with functions other than mtDNA transactions: a) protein import and mitochondrial biogenesis b) citric acid cycle and upstream glycolytic steps and c) amino acid metabolism (Table 1.2 and Figure 1.8, Chen *et al.*, 2005).

Figure 1.8 Conceptualization of the yeast mitochondrial nucleoid. For simplicity only one circular mitochondrial DNA (mtDNA) is shown with associated proteins. The two-membrane spanning (TMS) complex is depicted interacting with Mgm101p as indicated by a thin solid arrow and potentially with Mip1p indicated by a thin dashed arrow. Purple shapes represent proteins that are implicated in mtDNA repair, replication, recombination and maintenance. Green shapes represent proteins that are required for mtDNA replication, recombination, repair and maintenance and have been experimentally shown to bind mtDNA. Orange shapes represent proteins that are implicated in protein import and have been experimentally shown to bind mtDNA. Blue shapes represent enzymes of the tricarboxylic acid (TCA) cycle and upstream glycolytic steps experimentally shown to bind mtDNA. Yellow shapes represent enzymes for amino acid metabolism that have been experimentally shown to bind to mtDNA with the exception of Arg5,6p. Mnp1p is a putative mitochondrial ribosomal subunit and is coloured brown. See Table 1.2 for details of each nucleoid protein and references.

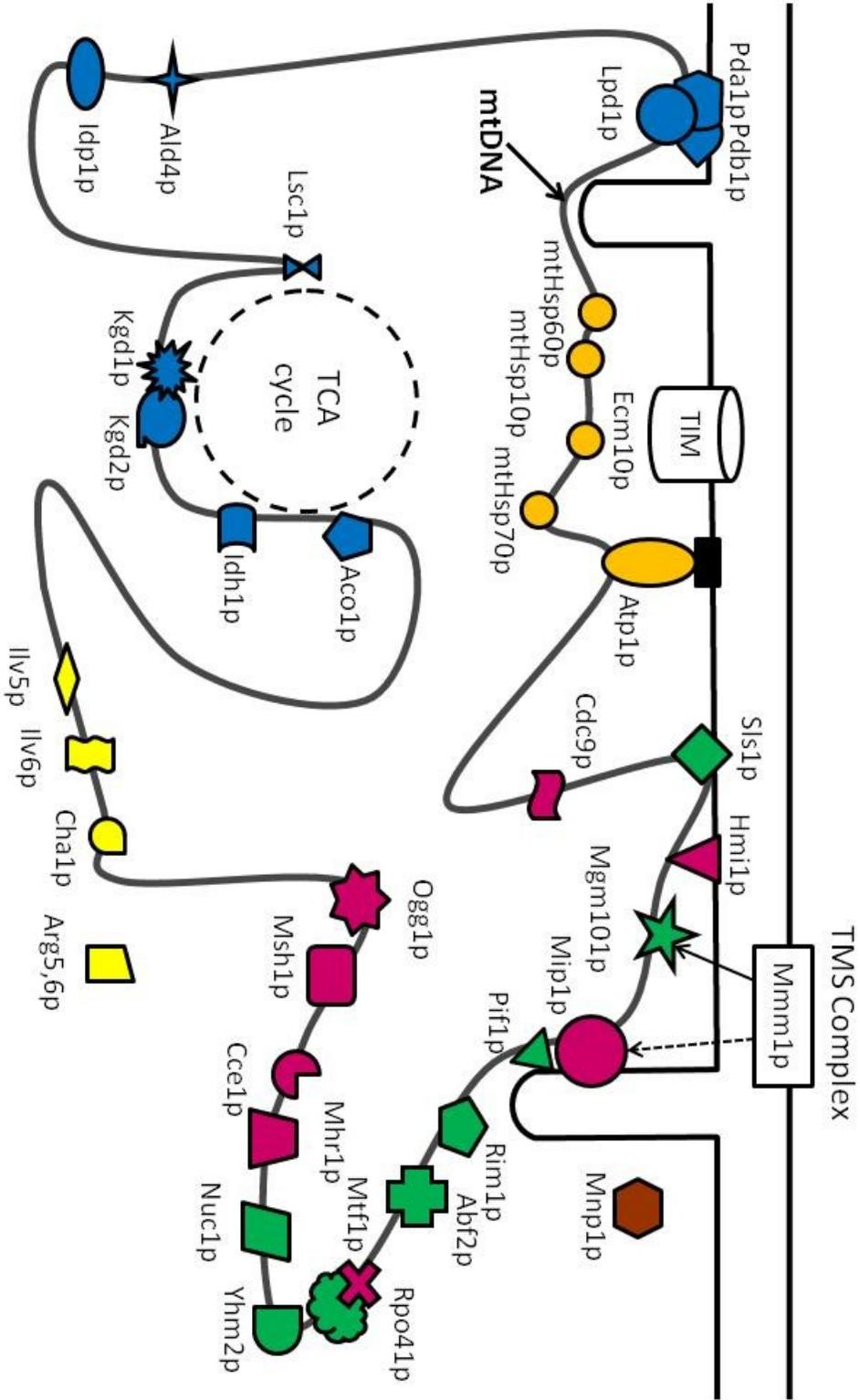


Table 1.2 Nucleoid proteins and proteins implicated in mitochondrial DNA transactions

Gene	Protein	Functions
<i>Category I: mtDNA Replication, recombination, repair and maintenance</i>		
<i>ABF2*</i>	Abf2p	mtDNA packaging, <i>rho</i> + genome maintenance, high mobility group protein
<i>MGM101*</i>	Mgm101p	Maintenance of <i>rho</i> + and ori-less mtDNA, mtDNA repair, Putative replication initiator
<i>RIM1*</i>	Single-stranded DNA binding protein	mtDNA replication
<i>RPO41*</i>	DNA-directed RNA polymerase	Mitochondrial transcription, <i>rho</i> + mtDNA maintenance
<i>MTF1**</i>	Transcription factor	Structural similarity to S-adenosylmethionine-dependent methyltransferases and functional similarity to bacterial sigma-factors, interacts with mitochondrial core polymerase Rpo41p
<i>SLS1*</i>	Sls1p	<i>rho</i> + mtDNA maintenance, Inner membrane protein, implicated in mitochondrial translation
<i>YHM2*</i>	Mitochondrial carrier protein, suppressor of abf2	Mitochondrial DNA-binding protein, involved in mtDNA replication and segregation of mitochondrial genomes; member of the mitochondrial carrier protein family
<i>NUC1*</i>	Major mitochondrial nuclease	RNase and DNA endo- and exonucleolytic activities; has roles in mitochondrial recombination, apoptosis and maintenance of ploidy
<i>MIP1**</i>	DNA polymerase γ	Catalytic subunit of the mitochondrial DNA polymerase; conserved C-terminal segment is required for the maintenance of mitochondrial genome
<i>CCE1**</i>	Cruciform cutting endonuclease	cleaves Holliday junctions formed during recombination of mitochondrial DNA
<i>CDC9**</i>	DNA ligase	found in the nucleus and mitochondria, an essential enzyme that joins Okazaki fragments during DNA replication; also acts in nucleotide excision repair, base excision repair, and recombination

Gene	Protein	Functions
<i>HMI1</i> **	DNA helicase	Required for the maintenance of the mitochondrial genome; not required for mitochondrial transcription; has homology to <i>E. coli</i> helicase uvrD
<i>MHR1</i> **	Mhr1p, mitochondrial homologous recombination	Implicated in recombination of mtDNA and suggested to initiate rolling circle replication
<i>MSH1</i> **	MutS homolog	DNA-binding protein of the mitochondria involved in repair of mitochondrial DNA, has ATPase activity and binds to DNA mismatches
<i>OGG1</i> **	8-oxoguanine glycosylase/lyase	Excises 7,8-dihydro-8-oxoguanine residues located opposite cytosine or thymine residues in DNA, repairs oxidative damage to mitochondrial DNA
<i>PIF1</i> *	DNA helicase	involved in telomere formation and elongation; acts as a catalytic inhibitor of telomerase; also plays a role in repair and recombination of mitochondrial DNA
<i>Category II: Protein import & mitochondrial biogenesis</i>		
<i>HSP60</i> *	mtHsp60p	Mitochondrial chaperonin, heat shock protein 60
<i>HSP10</i> *	mtHsp10p	Mitochondrial chaperonin, heat shock protein 10
<i>SSC1</i> *	mtHsp70p	Protein import, heat shock protein 70, subunit of Endo.Sccl
<i>ECM10</i> *	mtHsp70p	Heat shock protein of the Hsp70 family, plays a role in protein translocation, interacts with Mge1p in an ATP-dependent manner; overexpression induces extensive mitochondrial DNA aggregations
<i>ATP1</i> *	α -Subunit of F1-ATPase	ATP synthesis, protein import
<i>Category III: Citric acid cycle and upstream glycolytic steps</i>		
<i>ACO1</i> *	Mitochondrial aconitase	Citric acid cycle
<i>ALD4</i> *	Aldehyde dehydrogenase	Ethanol metabolism

Gene	Protein	Functions
<i>IDH1</i> *	NAD ⁺ -dependent isocitrate dehydrogenase, subunit 1	Citric acid cycle
<i>IDP1</i> *	NADP ⁺ -dependent isocitrate dehydrogenase	Oxidative decarboxylation of isocitrate
<i>KGD1</i> *	2-oxoglutarate dehydrogenase, E1 component of KGDC	Citric acid cycle
<i>KGD2</i> *	2-oxoglutarate dehydrogenase, E2 component of KGDC	Citric acid cycle
<i>LPD1</i> *	Dihydrolipoamide dehydrogenase, E3 component of PDHC, KGDC, and BCADC	Citric acid cycle, catabolism of branched-chain amino acids
<i>LSC1</i> *	Succinate-CoA ligase, α subunit	Citric acid cycle
<i>PDA1</i> *	Pyruvate dehydrogenase, E1 α -subunit of PDHC	Oxidation of pyruvate
<i>PDB1</i> *	Pyruvate dehydrogenase, E1 β -subunit of PDHC	Oxidation of pyruvate
<i>Category IV: Amino Acid metabolism</i>		
<i>ILV5</i> *	Acetohydroxyacid reductoisomerase	Biosynthesis of Val, Ile, and Leu
<i>ILV6</i> *	Acetolactate synthase regulatory subunit	Biosynthesis of Val, Ile, and Leu
<i>CHAI</i> *	L-serine/L-threonine deaminase	Catabolism of hydroxyl amino acids, Serine and threonine ammonia-lyase
<i>ARG5,6</i> *	acetylglutamate kinase and N-acetyl-gamma-glutamyl-phosphate reductase	Protein that is processed in the mitochondrion to yield acetylglutamate kinase and N-acetyl-gamma-glutamyl-phosphate reductase, which catalyze the 2nd and 3rd steps in arginine biosynthesis

Gene	Protein	Functions
<i>Category V: Mitochondrial translation</i>		
<i>MNPI</i> *	Putative mitochondrial ribosomal subunit (similar to <i>E. coli</i> L7/L12)	Required for normal respiratory growth

* - Bona fide nucleoid components – Experimentally demonstrated by Cho *et al.*, 1998, Kaufman *et al.*, 2000, Miyakawa *et al.*, 2000, Sato *et al.*, 2002, Umezaki and Miyakawa, 2002, Sakasegawa *et al.*, 2003, Sato and Miyakawa, 2004, Chen *et al.*, 2005, Cheng *et al.*, 2007, Hall *et al.*, 2004

** - Other mtDNA interacting proteins (reviewed in Nosek *et al.*, 2006)

Out of seventeen proteins, Butow's group choose to study the role of aconitase (*ACO1*) as a bifunctional nucleoid protein (Chen *et al.*, 2005). MtDNA is very unstable in *aco1Δ* cells. MtDNA stability was investigated in yeast null mutants of citrate synthase (*cit1Δ*, *cit2Δ*, and *cit3Δ*) to block the metabolic flux through the tricarboxylic acid (TCA) cycle upstream of aconitase, and in three *aco1* mutants harbouring separate mutations to alter each of the three cysteine residues necessary for coordination of the iron-sulfur center required for enzyme catalysis. These experiments proved that Aco1p contributes to mtDNA stability independently of the rate of flux through the TCA cycle and of aconitase activity (Chen *et al.*, 2005). When Aco1p and the three cysteine mutants thereof were constitutively expressed separately in a mtDNA unstable *abf2Δ* mutant, they all could rescue the mtDNA unstable phenotype. Therefore, Aco1p is an example of a mitochondrial TCA cycle enzyme that moonlights in mtDNA maintenance function. Recently, it was suggested that perhaps Aco1p provides mtDNA stability during conditions of oxidative stress (Shadel, 2005) and that Aco1p may be involved in mitochondrial nucleoid remodelling required due to changes in cellular metabolism (Chen *et al.*, 2005). For example during respiratory growth reactive oxygen species generated by the electron transport chain complexes I and III are known to damage mitochondrial lipids, proteins and mutagenize the mtDNA (Wallace, 2007). Perhaps oxidation of aconitase results in reallocation from the TCA cycle to the nucleoid to allow stabilization of mtDNA via a type of protective remodelling (Shadel, 2005). In this way aconitase could suffer the brunt of the ROS thereby protecting the valuable genes encoded by the mitochondrial genome. Clearly, the intricate details of the mitochondrial nucleoid and mtDNA replication are just beginning to be understood.

1.6 Thesis objectives

As was discussed at the beginning of this review, yeast is a commonly used model organism for studies of human mitochondrial disease as many orthologous mitochondrial disease-associated genes exist between these two species. Because of the wide-spread use of yeast as a model organism the first major objective of this thesis was to determine the contributions of several logical genes shown to contribute to loss of mtDNA maintenance and/or stability. The genes examined in chapter 2 are *MIP1*, *HAP1*, *ADE2*, and *HIS3/MRM1*. The effect of various genes on mitochondrial function was assessed in two different parental background strains (either *MIP1*[S] or *MIP1*[Σ]) harbouring all combinations of either wild-type or mutant forms of *HAP1*, *ADE2*, and *HIS3*. We monitored respiratory competence of yeast cultures at 30 and 37°C. Fidelity of mtDNA replication was monitored by the erythromycin-resistance assay and stability of mtDNA was monitored by fluorescence microscopy of 4',6-diamidino-2-phenylindole-stained yeast cells.

One difference between yeast and man that was the focus of this research is that of fungal specific CTE of Mip1p. Because yeast is a commonly used model for mitochondrial studies it is necessary to fully comprehend the details of mtDNA replication including the function of the CTE in this process. To this end, with the abundance of genomic sequence data available on the world-wide-web we searched for phylogenetic relationships among fungal mtDNA polymerases. The main goals here were 1) to determine if other fungal mtDNA polymerases harbour CTEs, 2) to determine if they do have CTEs is there any conservation within the amino acid sequence and potentially any putative structural regions, and 3) to determine the function of the CTE *in*

vivo. To specifically address goals 1) and 2) phylogenetic analyses were carried out. To address goal 3) various truncation mutants were constructed and analyzed for the ability of cells harbouring them to respire and maintain mtDNA.

Based on the observations from chapters 2 and 3, polymerase activities of Mip1p variants overexpressed and partially purified from yeast mitochondria were characterized. The main goals here were 1) to determine if the CTE contributes to DNA polymerase activity and 2) to determine if the alanine-661 substitution in Mip1p[S] detrimentally effects DNA polymerase activity compared to the threonine-661 version of Mip1p[Σ].

CHAPTER 2 Contributions of the S288c genetic background and common auxotrophic markers on mitochondrial DNA maintenance in *Saccharomyces*

2.1 Abstract

Saccharomyces cerevisiae is an excellent model system for the study of many eukaryotic processes, including mitochondrial DNA (mtDNA) maintenance. Work from several laboratories has shown that the sequenced strain S288c carries alleles of several genes that have a negative impact on mitochondrial DNA maintenance. The first of the genes considered in the current study is *MIP1* that encodes the mtDNA polymerase. *MIP1*[S], found in strain S288c, is associated with increased mtDNA mutation and deletion (Baruffini *et al.*, 2007b) and encodes a polymerase in which a conserved threonine at position 661 is replaced by an alanine. In contrast *MIP1*[Σ], found in strain Σ 1278b and harbouring the conserved Mip1p T661 residue, is associated with stable maintenance of the mitochondrial genome (Baruffini *et al.*, 2007b). Contributions of wild-type and mutant alleles *HIS3*, *ADE2*, and *HAP1* were then investigated in strains expressing Mip1p[S] or Mip1p[Σ]. Even in the *MIP1*[Σ] background, the presence of the *his3 Δ 200* allele, harbouring a deletion of the promoter for *MRM1*, is associated with loss of respiratory competence, but not of mtDNA, at 37°C. In contrast, mtDNA loss, and error-prone synthesis, measured as the acquisition of erythromycin-resistance, is highest in the *MIP1*[S] background, and is exacerbated by the *ade2* mutation. Thus, a complex set of genetic factors in commonly used laboratory strains make differential contributions to mtDNA maintenance, making strain background an important consideration in establishing and comparing yeast model systems.

2.2 Introduction

Saccharomyces cerevisiae is one of the most powerful model eukaryotes to study many fundamental biological processes, through biochemical and genetic analyses. *S. cerevisiae* has also emerged as an excellent model system for the study of processes in higher eukaryotes, including those associated with human mitochondrial disease (for reviews see (Barrientos, 2003, Foury and Kucej, 2002, Schwimmer *et al.*, 2006, Lecrenier and Foury, 2000, Shadel, 1999). Recent studies have shown that over twenty known human mitochondrial disease genes have yeast orthologs that, when deleted from the sequenced S288c yeast genome, cause a respiratory-deficient phenotype (Steinmetz *et al.*, 2002, Winzeler *et al.*, 1999, Table 1.1).

Part of the power of using yeast is the ease of genetic manipulation, which is in part due to the array of auxotrophic markers that have been engineered into a plethora of strains (see http://wiki.yeastgenome.org/index.php/Commonly_used_strains and Table 2.1). One of the most common backgrounds is that of S288c, the strain that was used for determining the first yeast genome sequence (Goffeau *et al.*, 1996). S288c-related strains have been used subsequently for making ORF libraries (Gelperin *et al.*, 2005), whole-genome deletion studies (http://sequence-www.stanford.edu/group/yeast_deletion_project/deletions3.html) (Winzeler *et al.*, 1999), GFP-tagging the proteome (Huh *et al.*, 2003), TAP-tagging the proteome (Ghaemmaghami *et al.*, 2003) as well as studying human mitochondrial disease (Steinmetz *et al.*, 2002, Ma *et al.*, 2004, Stuart *et al.*, 2006, Iida *et al.*, 1998). Other strain backgrounds include W303 (Baruffini *et al.*, 2006, Chen and Kaplan, 2000, Barrientos *et al.*, 2002) which has been stated to be closely related to S288c (Baruffini *et al.*, 2007b,

Table 2.1 Yeast strains used for mitochondrial studies

Strain Name	Genotype	Background Relation	Relevant Yeast Gene Studied	Human Ortholog	Experiment	Reference
aGH1	<i>MATa gal1delta152 leu2-3 leu2-112 trp1-289</i>		HAP1/YLR256W, Zinc finger transcription factor			Becerra <i>et al.</i> , 2002
aW303deltaCOX10	<i>MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 cox10::HIS3</i>	W303/S288c-like	COX10/YPL172C, Heme A: farnesyltransferase	>gnl UG Hs#S1732347 <i>Homo sapiens COX10</i> (yeast) homolog, yeast class III ^a	functional complementation of a yeast cox10 mutant	Glerum and Tzagoloff, 1994
BJ3345	<i>MATalpha ade2-101 his3-delta200 leu2-delta1 lys2-801 trp1-delta101 ura3-52</i>	S288c				Derived from S288c, John Woolford personal communication
BWG1-7a	<i>MATa ade1-200 his4-519 leu2-3,112 ura3-52</i>		HAP1/YLR256W, Zinc finger transcription factor			Mattoon <i>et al.</i> , 1990
BY4741	<i>MATa his3delta1 leu2delta0 met15delta0 MIP1[S]^b ura3delta0</i>	S288c				Brachmann <i>et al.</i> , 1998, Baruffini <i>et al.</i> , 2007b
BY4742	<i>MATalpha his3delta1 leu2delta0 lys2delta0 MIP1[S] ura3delta0</i>	S288c				Brachmann <i>et al.</i> , 1998, Baruffini <i>et al.</i> , 2007b
BY4743	<i>MATa/MATalpha his3delta1/his3delta1 leu2delta0/leu2delta0 lys2delta0/+ met15delta0/+ MIP1[S]/ MIP1[S] ura3delta0/ura3delta0</i>	S288c				ATCC 201390, Baruffini <i>et al.</i> , 2007b

Strain Name	Genotype	Background Relation	Relevant Yeast Gene Studied	Human Ortholog	Experiment	Reference
CG379	<i>MATalpha ade5 his7-2 leu2-3 leu2-112 trp1-289amber ura3-52 [Kil-0]</i>	S288c				CG379 is closely related to S288c, Craig N. Giroux personal communication
COP161 U7	<i>MATalpha ade lys ura3 [rho+]</i>					Sirum-Connolly and Mason, 1993
COP161 U7 F11 (or just F11)	<i>MATalpha ade lys ura3 [rho-]</i>					Sirum-Connolly and Mason, 1993
D273-10B/A1	<i>MATalpha mal [rho+] MIP1[sigma]</i>					Sherman, 1964, Baruffini <i>et al.</i> , 2007b
DL1	<i>MATalpha his3-11,15 leu2-3,112 ura3-251,328,372</i>	isogenic or congeneric to D273-10B	CRD1/YDL142C, Cardiolipin synthase	hypothetical protein (LOC54675), mRNA, yeast class III	Studied interactions of <i>his3delta-200</i> and <i>CRD1</i> with respect to temperature sensitive phenotypes in yeast	Dunstan <i>et al.</i> , 1997, Zhang <i>et al.</i> , 2003
DY150	<i>MATa ura3-52 leu2-3 leu2-112 trp1-1 ade2-1 his3-11 can1-100</i>	W303/S288c-like				
E134	<i>MATalpha lys2-InsE-A14 ade5-1 his7-2 leu2-3,112 trp1-289 ura3-52</i>	CG379 (S288c)	MIP1/YOR330C, Catalytic subunit of the mitochondrial DNA polymerase	>gnl UG Hs#S1727412 <i>Homo sapiens</i> polymerase (DNA directed), gamma (POLG), nuclear gene encoding mitochondrial protein, mRNA, yeast class III	Studied progressive external ophthalmoplegia-associated mutations in yeast	isogenic with CG379 Jin <i>et al.</i> , 2003, CG379 is closely related to S288c, Craig N. Giroux personal communication, Stuart <i>et al.</i> , 2006
EM93	<i>MATa/MATalpha SUC2/SUC2 GAL2/gal2 MAL/MAL mel/mel CUP1/cup1 FLO1/flo1</i> , heterozygous for <i>MIP1</i> A and G at position 1981					Parent of S288c contributes ~ 90% of the S288C genome, Mortimer and Johnston, 1986, Baruffini <i>et al.</i> , 2007b
FGY2	<i>MATalpha ura3-52 lys2-52 lys2-801 ade2-101 trp1-delta1 his3-delta200 leu2-delta1 crd1::URA3</i>	isogenic with FGY3, CG990/YPH2 74 (S288c)	TAZ1/YPR140W, Lyso-phosphatidylcholine acyltransferase	human G4.5 encoding a putative acyltransferase	The human <i>TAZ</i> gene complements mitochondrial dysfunction in the yeast <i>taz1Delta</i> mutant. Implications for Barth syndrome Ma <i>et al.</i> , 2004	Ma <i>et al.</i> , 2004

Strain Name	Genotype	Background Relation	Relevant Yeast Gene Studied	Human Ortholog	Experiment	Reference
FGY3	<i>MATalpha ura3-52 lys2-52 lys2-801 ade2-101 trp1-delta1 his3-delta200 leu2-delta1</i>	CG990/YPH274 (S288c)	TAZ1/YPR140W, Lyso-phosphatidylcholine acyltransferase CRD1/YDL142C, Cardiolipin synthase	human G4.5 encoding a putative acyltransferase	The human TAZ gene complements mitochondrial dysfunction in the yeast <i>taz1</i> Delta mutant. Implications for Barth syndrome Ma <i>et al.</i> , 2004	Ma <i>et al.</i> , 2004, FGY3 is a spore derived from CG990 Jiang <i>et al.</i> , 1997, CG990 is Phil Hieter's YPH274, personal communication from Craig N. Giroux
FGY3	<i>MATalpha ura3-52 lys2-52 lys2-801 ade2-101 trp1-delta1 his3-delta200 leu2-delta1</i>	CG990/YPH274 (S288c)	CRD1/YDL142C, Cardiolipin synthase Zhong <i>et al.</i> , 2004	<i>Homo sapiens</i> hypothetical protein (LOC54675), mRNA, yeast class III Steinmetz <i>et al.</i> , 2002	Absence of cardiolipin results in temperature sensitivity, respiratory defects, and mitochondrial DNA instability indepent of <i>pet56</i> (<i>mrm1</i>)	Zhong <i>et al.</i> , 2004, FGY3 is a spore derived from CG990 Jiang <i>et al.</i> , 1997, CG990 is Phil Hieter's YPH274, personal communication from Craig N. Giroux
FL100	<i>MATa MIP1[sigma]</i>					http://wiki.yeastgenome.org/index.php/Commonly_used_strains , Baruffini <i>et al.</i> , 2007b
HMD22	<i>MATa ura3-52 leu2-3,112 lys2 his3deltaHindIII arg8::hisG cox2::ARG8m</i>	D273-10B	CRD1/YDL142C, Cardiolipin synthase Zhang <i>et al.</i> , 2003	<i>Homo sapiens</i> hypothetical protein (LOC54675), mRNA, yeast class III Steinmetz <i>et al.</i> , 2002	Cardiolipin is not required to maintain mitochondrial DNA stability or cell viability for <i>Saccharomyces cerevisiae</i> grown at elevated temperatures	Zhang <i>et al.</i> , 2003, Dunstan <i>et al.</i> , 1997
HSD2	<i>MATa ura3-52 leu2-3,112 lys2 his3deltaHindIII arg8::hisG cox2::ARG8m crd1delta::HIS3</i>	D273-10B	CRD1/YDL142C, Cardiolipin synthase Zhang <i>et al.</i> , 2003	<i>Homo sapiens</i> hypothetical protein (LOC54675), mRNA, yeast class III Steinmetz <i>et al.</i> , 2002	Cardiolipin is not required to maintain mitochondrial DNA stability or cell viability for <i>Saccharomyces cerevisiae</i> grown at elevated temperatures	Zhang <i>et al.</i> , 2003
JWY1466	<i>MATa can1 his3-delta200 leu2-delta1 lys2-801 trp1-delta101 ura3-52</i>	S288c				Derived from S288c John Woolford personal communication
KSY24 F11	<i>MATalpha ade lys pet56::URA3 [rho-]</i>	COP161 U7 F11 [rho-]				Sirum-Connolly and Mason, 1993
KSY35	<i>HIS3 strain, MATa ade2-101 leu2-D1 lys2-801 trp1-D101 ura3-52 [rho+]</i>	MM1403				Sirum-Connolly and Mason, 1993

Strain Name	Genotype	Background Relation	Relevant Yeast Gene Studied	Human Ortholog	Experiment	Reference
MB2	<i>MATa/alpha ADE2/ade2-101ochre his3/his3-delta200 leu2/leu2-delta1 lys2-801amber/lys2-801amber trp1-289/TRP1 ura3-52/ura3-52</i>	crossing MB1 and YP102 Maarse <i>et al.</i> , 1992	Tim8/YJR135W-A , Mitochondrial intermembrane space protein mediating import and insertion of polytopic inner membrane proteins	DDP1 (deafness/dystonia peptide 1) gene, associated with Mohr-Tranebjaerg syndrome	It is demonstrated that import of human Tim23 into yeast and rat mitochondria is dependent on a high membrane potential. A mechanism to explain the pathology of MTS is discussed	Maarse <i>et al.</i> , 1992, Paschen <i>et al.</i> , 2000
MCC123	<i>MATa ura3-52 ade2 kar1-1</i>		tL(UAA)Q, Mitochondrial tRNA-Leu	Human tRNA-Leu (UUR)	The yeast counterparts of human 'MELAS' mutations cause mitochondrial dysfunction that can be rescued by overexpression of the mitochondrial translation factor EF-Tu	Feuermann <i>et al.</i> , 2003
MM1402	<i>Mata/Matalpha ade2-101/+ canr1/+ his3-delta200/his3-delta200 leu2-delta1/leu2-delta1 lys2-801/lys2-801 trp1-delta101/trp1-delta101 ura3-52/ura3-52 rpl16a::TRP1/+ rpl16b::LEU2/+</i>	BJ3345 and JWY1466 (S288c)				Genotype, Thomas L. Mason personal communication, MM1402 was derived from the parent strains BJ3345 and JWY1466, John Woolford personal communication
MM1403	<i>MATa ade2-101 his3-D200 leu2-D1 lys2-801 trp1-D101 ura3-52 [rho+]</i>	MM1402 (S288c)	PET56/YOR201C, Ribose methyltransferase	>gnl UG Hs#S3438785 <i>Homo sapiens</i> hypothetical protein FLJ22578 (FLJ22578), mRNA, yeast class III Steinmetz <i>et al.</i> , 2002	Cells reduced in <i>MRM1</i> activity (his3-delta200) were deficient in formation of functional large subunits of the mitochondrial ribosome	Haploid derived from MM1402 Thomas L. Mason personal communication Sirum-Connolly and Mason, 1993
RM11-1a	<i>MIP1[S]-like allele but has Thr661</i>					Baruffini <i>et al.</i> , 2007b
Sigma1278b	<i>MIP1[sigma]</i>	directly related to S288c				Winzeler <i>et al.</i> , 2003, Baruffini <i>et al.</i> , 2007b
S288c	<i>MATalpha SUC2 gal2 mal mel flo1 flo8-1 hap1 ho bio1 bio6</i> , probably <i>MIP1[S]</i>					http://wiki.yeastgenome.org/index.php/Commonly_used_strains , Baruffini <i>et al.</i> , 2007b
W303 isogenic strains		W303/S288c-like	YFH1/YDL120W, Frataxin, regulates mitochondrial iron accumulation	>gnl UG Hs#S1726774 <i>Homo sapiens</i> Friedreich ataxia (FRDA), mRNA, not class III	CCC1 Suppresses Mitochondrial Damage in the Yeast Model of Friedreich's Ataxia by Limiting Mitochondrial Iron Accumulation	Chen and Kaplan, 2000

Strain Name	Genotype	Background Relation	Relevant Yeast Gene Studied	Human Ortholog	Experiment	Reference
W303	<i>MATa/MATalpha</i> (<i>leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15</i>) [<i>phi</i> ⁺]	directly related to S288c	SHY1/YGR112W, Mitochondrial inner membrane protein required for assembly of cytochrome c oxidase	>gnl UG Hs#S1727967 Homo sapiens surfeit 1 (SURF1), mRNA, yeast class III	Shy1p is necessary for full expression of mitochondrial <i>COX1</i> in the yeast model of Leigh's syndrome	Winzeler <i>et al.</i> , 2003, http://wiki.yeastgenome.org/index.php/Commonly_used_strains , Barrientos <i>et al.</i> , 2002
W303-Dtaz	<i>MATa/MATalpha</i> (<i>leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15</i>) [<i>phi</i> ⁺] <i>taz1delta</i>	W303/S288c-like	TAZ1/YPR140W, Lyso-phosphatidylcholine acyltransferase	human G4.5 encodes a putative acyltransferase involved in phospholipid biosynthesis	The human <i>TAZ</i> gene complements mitochondrial dysfunction in the yeast <i>taz1Delta</i> mutant. Implications for Barth syndrome	Vaz <i>et al.</i> , 2003, Ma <i>et al.</i> , 2004
W303-1A	<i>MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15</i>	W303/S288c-like	MNP1/YGL068W, Protein associated with the mitochondrial nucleoid; putative mitochondrial ribosomal protein	>gnl UG Hs#S1727713 Homo sapiens mitochondrial ribosomal protein L12 (MRPL12), mRNA, not class III	null mutation of <i>MNP1</i> led to a respiratory-deficient phenotype	http://wiki.yeastgenome.org/index.php/Commonly_used_strains , Sato and Miyakawa, 2004
W303-1B	<i>MATalpha leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 MIP1[S]</i>	W303/S288c-like	MIP1/YOR330C, Catalytic subunit of the mitochondrial DNA polymerase	>gnl UG Hs#S1727412 Homo sapiens polymerase (DNA directed), gamma (POLG), nuclear gene encoding mitochondrial protein, mRNA, yeast class III	Genetic and chemical rescue of the <i>Saccharomyces cerevisiae</i> phenotype induced by mitochondrial DNA polymerase mutations associated with progressive external ophthalmoplegia in humans	http://wiki.yeastgenome.org/index.php/Commonly_used_strains , Baruffini <i>et al.</i> , 2006, Baruffini <i>et al.</i> , 2007b
YGM128	<i>MATalpha ura1 ura2 leu2::KanR, CanR</i>		tL(UAA)Q, Mitochondrial tRNA-Leu	Human tRNA-Leu (UUR)	The yeast counterparts of human 'MELAS' mutations cause mitochondrial dysfunction that can be rescued by overexpression of the mitochondrial translation factor EF-Tu	Feuermann <i>et al.</i> , 2003
YH747	<i>MATa lys2-A14 ade5-1 his7-2 leu2-3,112 trp1-289 ura3-52</i>	isogenic to E134 (S288c)				Stuart <i>et al.</i> , 2006
YJM789	<i>MIP1[S]</i> -like allele but has Thr661					Baruffini <i>et al.</i> , 2007b
YPH274	<i>MATa/alpha</i> (<i>ura3-52 lys2-801 amber ade2-101 ochre trp1-delta1 his3-delta200 leu2-delta1</i>)	Derived from YNN216 (congenic with S288c)				Sikorski and Hieter, 1989
YPH499	<i>MATa ura3-52 lys2-801 amber ade2-101 ochre trp1-delta63 his3-delta200 leu2-delta1 MIP1[S]</i>	Derived from YNN216 (congenic with S288c)	CCC2/YDR270W, Cu(+2)-transporting P-type ATPase	>gnl UG Hs#S1726319 Homo sapiens ATPase, Cu ⁺⁺ transporting, beta polypeptide (Wilson disease) (ATP7B), mRNA, yeast class III	Analysis of functional domains of Wilson disease protein (ATP7B) in <i>Saccharomyces cerevisiae</i>	Sikorski and Hieter, 1989, Iida <i>et al.</i> , 1998

Strain Name	Genotype	Background Relation	Relevant Yeast Gene Studied	Human Ortholog	Experiment	Reference
YPH499	<i>MATa ura3-52 lys2-801amber ade2-101ochre trp1-delta63 his3-delta200 leu2-delta1 MIP1[S]</i>	Derived from YNN216 (congenic with S288c)	APT20/YPR020W, Subunit g of the mitochondrial F1F0 ATP synthase		Profiling Phosphoproteins of Yeast Mitochondria Revealed a Role of Phosphorylation in Assembly of the ATP Synthase, 80 phosphorylation sites in 48 different proteins were identified. These mitochondrial phosphoproteins are involved in critical mitochondrial functions, including energy metabolism, protein biogenesis, fatty acid metabolism, metabolite transport, and redox regulation. By combining yeast genetics and in vitro biochemical analysis, it was found that phosphorylation of a serine residue in subunit g (Atp20) regulates dimerization of the mitochondrial ATP synthase.	Sikorski and Hieter, 1989, Reinders <i>et al.</i> , 2007
YPH500	<i>MATalpha ura3-52 lys2-801amber ade2-101ochre trp1-delta63 his3-delta200 leu2-delta1</i>	Derived from YNN216 (congenic with S288c)	CRD1/YDL142C, Cardiolipin synthase	Homo sapiens hypothetical protein (LOC54675), mRNA, yeast class III	Cardiolipin is not required to maintain mitochondrial DNA stability or cell viability for <i>Saccharomyces cerevisiae</i> grown at elevated temperatures	Sikorski and Hieter, 1989, Zhang <i>et al.</i> , 2003
YZD2	<i>MATalpha ura3-52 lys2-801amber ade2-101ochre trp1-delta63 his3-delta200 leu2-delta1 crd1delta::HIS3</i>	YPH500, Derived from YNN216 (congenic with S288c)	CRD1/YDL142C, Cardiolipin synthase	Homo sapiens hypothetical protein (LOC54675), mRNA, yeast class III	Cardiolipin is not required to maintain mitochondrial DNA stability or cell viability for <i>Saccharomyces cerevisiae</i> grown at elevated temperatures	Zhang <i>et al.</i> , 2003
YZD5	<i>MATalpha his3-11,15 leu2-3,112 ura3-251,328,372 crd1delta::HIS3</i>	DL1 (isogenic or congenic to D273-10B)	CRD1/YDL142C, Cardiolipin synthase	Homo sapiens hypothetical protein (LOC54675), mRNA, yeast class III	Cardiolipin is not required to maintain mitochondrial DNA stability or cell viability for <i>Saccharomyces cerevisiae</i> grown at elevated temperatures	Dunstan <i>et al.</i> , 1997, Zhang <i>et al.</i> , 2003
ZGY1	<i>MATalpha ura3-52 lys2-52 lys2-801 ade2-101 trp1-delta1 his3-delta200 leu2-delta1 taz1::KanMX4</i>	isogenic with FGY3, CG990/YPH274 (S288c)	TAZ1/YPR140W, Lyso-phosphatidylcholine acyltransferase	human G4.5 encoding a putative acyltransferase	The human <i>TAZ</i> gene complements mitochondrial dysfunction in the yeast <i>taz1Delta</i> mutant. Implications for Barth syndrome	Ma <i>et al.</i> , 2004

^aYeast class III are genes that when deleted cause defects in the fitness of *Saccharomyces cerevisiae* when grown on nonfermentable carbon sources Steinmetz *et al.*, 2002.

^b*MIP1* alleles are either S288c-related, *MIP1*[S], or Sigma1278b-related, *MIP1*[sigma], and have been sequenced and defined by Baruffini *et al.*, 2007b or by us in the case of YPH499.

Winzler *et al.*, 2003). In many cases only the relevant genotype, but not the background of the strain is known or reported (Table 2.1).

Two issues arise with the use of different strain backgrounds. First, the selectable markers therein may influence some of the traits under study. Secondly, the classical yeast genetics used for early strain development involved crossing and backcrossing to transfer markers from one strain into another. This process has introduced “genetic baggage” that may not be obvious until a particular trait is studied.

The contribution of auxotrophic marker genes to the interpretation of phenotypes is exemplified by studies utilizing strains bearing the *his3-Δ200* allele. This allele harbours a 1036-base pair deletion that removes the entire *HIS3* gene and a bidirectional promoter required for transcription of *MRM1* (*PET56*, Figure 2.1), which encodes a mitochondrial site-specific ribose methyltransferase (Struhl, 1985, Sirum-Connolly and Mason, 1993, Rieger *et al.*, 1997). As a result, *his3-Δ200* strains display respiratory deficiencies (Zhang *et al.*, 2003, Zhong *et al.*, 2004, Figure 2.2). Zhang *et al.* (2003) studied the deletion of the cardiolipin synthase gene (*CRD1*) in D273-10B (*MRM1*) and YPH500 (*mrm1/his3Δ200*) backgrounds. Ultimately, the authors concluded that mitochondrial DNA (mtDNA) instability, loss of cellular viability, and respiratory defects exhibited at high temperatures were due to the presence of the *mrm1* allele and the synergistic effect of *crd1Δ* and *mrm*, but not the *crd1Δ* alone. In contrast, Zhong *et al.* (Zhong *et al.*, 2004) reported that deletion of *CRD1* caused temperature sensitivity, respiratory defects, and mtDNA instability, independently of *mrm1*. Interestingly, the *crd1Δ* strains that maintain the best respiratory competence at the end of both of these studies are both prototrophic for *ADE2* (a DL1/D273-10B derived strain

Figure 2.1 Region of the *HIS3* locus of *S. cerevisiae*. *his3Δ1* is an allele containing an internal 187-bp deletion of the *HIS3* gene (663 bp) whereas *his3Δ-200* contains a 1040-bp deletion of the entire *HIS3* gene as well as the poly (dA-dT) upstream promoter element (17 bp), box with arrowheads. This promoter acts bidirectionally to activate transcription of both genes; the direction of transcription is indicated by the arrows (Sirum-Connolly and Mason, 1993).

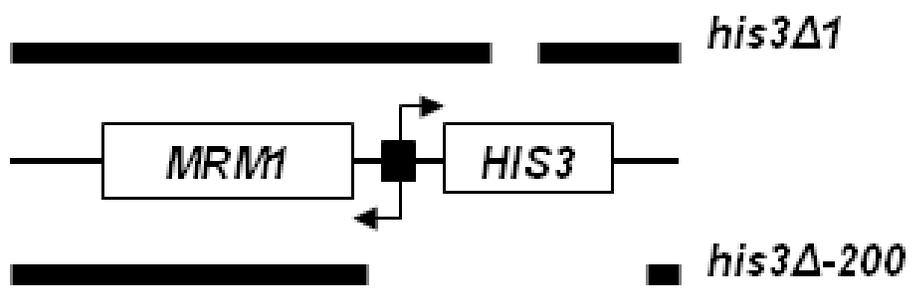
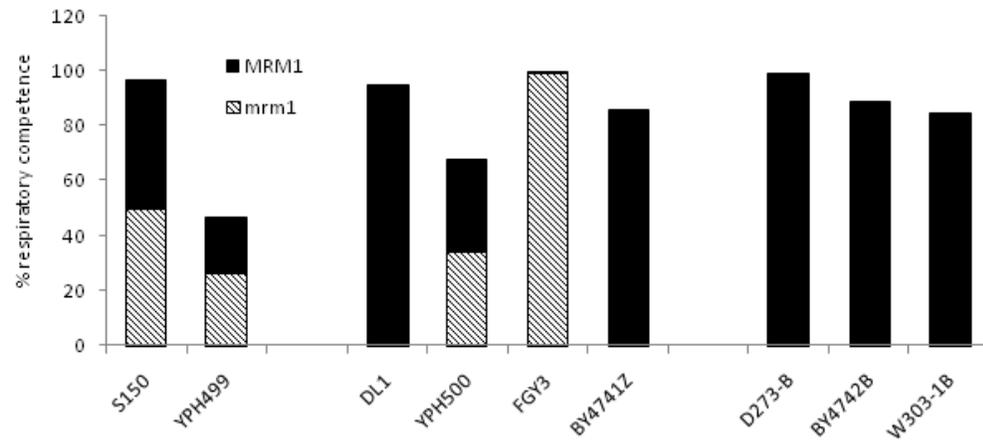


Figure 2.2 Comparison of respiratory competence. Data from this work (37°C, S150 and YPH499) are compared with those from DL1, YPH500 (37°C, Zhang *et al.*, 2003), BY4741 (37°C), FGY3 (39°C, Zhong *et al.*, 2004) and D273-10B/A1, BY4742 and W303-1B (36°C, Baruffini *et al.*, 2007b). Data are grouped according to the experimental methods, as described in the text. Respiratory competence in *mrm1* (*his3Δ200*) variants are indicated by hatched bars, and the competence of the isogenic *MRM1* variants are indicated by the total height of the solid and hatched bars. Relevant genotypes are indicated below each strain. Question marks indicate that the indicated allele has been inferred from whether or not the strain is known to be derived from S288c (*hap1* *MIP1*[*S*]) or Σ1278b (in the case of S150, *HAP1* and *MIP1*[*Σ*]).



<i>MIP1</i> allele	Σ	S	$\Sigma?$	S?	S?	S	Σ	S	S
<i>HAP1</i>	+	-	+	-?	-?	-?	+	-?	-?
<i>ADE2</i>	+	-	+	-	-	+	+	+	-

(Zhang *et al.*, 2003) and a BY4741 strain (Zhong *et al.*, 2004). In contrast the strains that did not maintain respiratory competence well (YPH500 (Zhang *et al.*, 2003) and FGY3/YPH274 (Zhong *et al.*, 2004)) are both *ade2-101ochre* auxotrophs (Figure 2.2). Thus, synergistic effects may result from the use of multiple markers.

The issue of “genetic baggage” has also been exemplified in studies of mitochondrial function. The S288c strain has been documented as being inappropriate for mitochondrial studies due to the mutated *HAPI* locus (Sherman, 2002). *HAPI* encodes a transcription factor that senses cellular heme status and increases expression of aerobic genes in response to oxygen (Ter Linde and Steensma, 2002). The effect of *hap1* mutations varies from a slight increase in mtDNA instability (*rho*⁻) (Mattoon *et al.*, 1990) in a BWG1-7a background, to a *petite* phenotype and diminished respiratory capacity in strain aGH1 (Becerra *et al.*, 2002, Table 2.1).

One possible explanation for this apparent discrepancy could come from a recent example provided by Baruffini *et al.* (Baruffini *et al.*, 2007b). It was shown that a single nucleotide polymorphism in the yeast mtDNA polymerase gene (*MIP1*) changes an evolutionarily conserved threonine at position 661 in *MIP1*[Σ] to an alanine in *S. cerevisiae* strains related to S288c, *MIP1*[S] (W303-1B, BY4741, BY4742, and BY4743). They showed that an alanine residue at position 661 is responsible for an increase in point mutations and deletions within the mitochondrial genome.

Due to the discrepancies in the literature noted above concerning respiratory competence and the wide spread use of S288c, we re-evaluated some of the issues pertinent to choice of yeast background strain on respiration and stability of the mitochondrial genome. The auxotrophic marker *ade2-101ochre* is common to the YPH

series of strains (Sikorski and Hieter, 1989) and is used for colorimetric determination of respiratory competence (Shadel, 1999). In this current study, this marker was examined alone and in combination with the *his3-Δ200* (*mrm1*) mutation discussed above, in backgrounds containing wild-type and mutant *HAPI* alleles. Finally, the role of the *MIP1* allele was considered by examining these marker alleles in the S288c-related strain YPH499 (*MIP1*[S]), and in S150 (*MIP1*[Σ]), Table 2.2 (Sherman *et al.*, 1986). The results herein demonstrate that multiple factors contribute to the maintenance of mitochondrial function, and that strain background is an important consideration in both designing experiments and comparing results obtained by different research groups.

2.3 Materials and Methods

2.3.1 General methods

Restriction digestions were carried out according to manufacturers' procedures. General molecular biological procedures were carried out according to Sambrook and Russell, 2001 and yeast transformations were carried out according to Gietz and Woods, 2002. Linear DNA fragments obtained from restriction digestion of plasmids or PCR amplification and used for yeast transformation or cloning were gel extracted/purified using the QIAquick Gel Extraction Kit (QIAGEN Inc., Mississauga, ON). When necessary yeast cells were screened to ensure they were haploid using the rapid assessment of *S. cerevisiae* mating type by PCR (Huxley *et al.*, 1990).

2.3.2 Media, growth conditions, and genetic methods.

Yeast were grown on various media including: YPG (1% yeast extract, 2% peptone, 3% glycerol); YG-Er (2% yeast extract, 3% glycerol, 50 mM potassium phosphate, pH 6.5, 4 g/L erythromycin); YP2D (1% yeast extract, 2% peptone, 2%

Table 2.2 Yeast strains used in this study

Strain	Genotype	Source or reference
YPH499	<i>MATa ura3-52 lys2-801amber ade2-101ochre trp1-Δ63 his3-Δ200 leu2-Δ1 hap1^a MIP1[S]^b</i> , Derived from the diploid strain YNN216, which is congenic with S288c	Sikorski and Hieter, 1989
YNN216	<i>MATα/α ura3-52/ura3-52 lys2-801amber/lys2-801amber ade2-101ochre/ade2-101ochre</i> , Congenic to S288c	Sikorski and Hieter, 1989
S288c	<i>MATα SUC2 gal2 mal mel flo1 flo8-1 hap1 ho bio1 bio6</i>	Mortimer and Johnston, 1986
S150	<i>MATa leu2-3,112 his3-Δ1 trp1-289 ura3-52 MIP1[Σ]^c HAP1</i>	Steger <i>et al.</i> , 1990
DACS1	[S150] <i>URA3</i>	This study
DACS2	[S150] <i>his3-Δ200</i>	This study
DACS3	[S150] <i>ade2-101ochre URA3</i>	This study
DACS4	[S150] <i>ade2-101ochre</i>	This study
DACS5	[S150] <i>his3-Δ200 ade2-101ochre</i>	This study
DACS6	[S150] <i>his3-Δ200 URA3</i>	This study
DACS7	[S150] <i>his3-Δ200 ade2-101ochre URA3</i>	This study
DACS8	[S150] <i>his3-Δ200 ade2-101ochre hap1::caURA3^b</i>	This study
DACS9	[S150] <i>hap1::caURA3^b</i>	This study
DACY1	[YPH499] <i>ADE2</i>	This study
DACY2	[YPH499] <i>HIS3</i>	This study
DACY3	[YPH499] <i>HIS3 ADE2</i>	This study

^a*HAP1* or *hap1* alleles were determined in YPH499 and S150 by either the presence or absence of a PCR product generated with CloneHAP1F/CloneHAP1R primers

^b*MIP1* alleles were identified by sequencing and are from either S288c, *MIP1*[S], or Σ1278b, *MIP1*[Σ], background strain as defined by Baruffini *et al.*, 2007b

^c*caURA3* is the *Candida albicans* ortholog of *S. cerevisiae URA3* Sheff and Thorn, 2004

dextrose); YP10D (1% yeast extract, 2% peptone, 10% dextrose); SC-URA (0.67% Difco™ nitrogen base without amino acids, 4% dextrose, all amino acids except uracil, pH 5.6, amino acid concentrations used were as described by (Sherman, 1991); SC-ADE (0.67% Difco™ nitrogen base without amino acids, 4% dextrose, all amino acids except adenine, pH 5.6); SC-HIS (0.67% Difco™ nitrogen base without amino acids, 4% dextrose, all amino acids except histidine, pH 5.6); 5-fluoroorotic acid (SC+FOA) (0.7% Difco™ nitrogen base without amino acids, 2% dextrose, 0.1% FOA, 0.007% uracil, and all required amino acids). All components were obtained from Fisher Scientific (Ottawa, Ontario) except for erythromycin and FOA which were obtained from Research Products International Corp. (Mt. Prospect, Illinois) and ZYMO RESEARCH (Orange, USA) respectively. Percentages indicate weight per volume and all solid media contained 2% Bacto™ agar.

2.3.3 Strain construction

S. cerevisiae strains used in this study are listed in Table 2.2. Primers used to construct and confirm all strains are listed in Table 2.3.

The DACS2 strain, used to construct DACS5 – 8, was constructed using a transplacement vector and a method modified from (Scherer and Davis, 1979, Figure 2.3A). pBS::*URA3* was constructed by Heather McIntosh (University of Manitoba). Briefly, the *Bgl*III fragment from pFL34 was inserted into the *Bam*HI site of pBluescript to make pBS::*URA3*. An ~1.1 kb region of the *his3-Δ200* allele was PCR amplified from YPH499 genomic DNA using the HUS/HDS primer set (Table 2.3), restriction digested, gel purified and inserted directionally into the *Eco*RI and *Kpn*I sites of pBS::*URA3*

Table 2.3 Primers used in this study

Primer Name	Primer Sequence (5' to 3')*	T _m (°C)	Gene Target
ADE2DS	<u>AAAGGTACCATGAACTGTATCGAAAC</u> G	50	<i>ADE2</i>
ade2DSE	<u>AAAGAATTCATGAACTGTATCGAAAC</u> G	50	<i>ADE2</i>
ADE2US	<u>AAAGAATTCTATTACAGCTATGCTGA</u> C	50	<i>ADE2</i>
CaURA3F	CATCAATTGACGTTGATAACC	56	<i>caURA3</i>
CaURA3R	TAATCCAACCTCCAGGTGTC	56	<i>caURA3</i>
CloneHAP1F	<u>AAAGGATCCTTAGCAAACCACAGTAC</u> GAGC	62	<i>HAP1</i>
CloneHAP1R	<u>AAAGGATCCCAACTTCTTCAACCTCC</u> AATTAT	62	<i>HAP1</i>
DHAPF	TTAGATGTCGAATACCCCTTATAATTC ATCTGTGCCTTCCATTGCATCCATGAC C TGA GTTTAGCTTGCCTCGTCC	56	<i>caURA3</i> of pKT209
DHAPR	ATGTCCCTTTCTATTTATACATAAACA AACCGCATCACTCCCATATTGGAAAAT C TCGATGAATTCGAGCTCG	54	<i>caURA3</i> of pKT209
DScrMIPF	GAGGACATGCTGTATCGG	56	<i>MIP1</i>
DScrMIPR	AGCTAACGAGAAGATGCAG	56	<i>MIP1</i>
GenomicMIP_Se q1R	GCACATTCATAGAACGCTTC	58	<i>MIP1</i>
HDS	<u>AAAGAATTCTTCAGTTCATCGACTAT</u> C	50	<i>HIS3</i>
HUS	<u>AAAGGTACCTCTGTGATTGGTTCAGG</u>	50	<i>HIS3</i>
MIPDSofERev	GAAGATGTAGATAATCCACTC	58	<i>MIP1</i>
MIPSeq2R	CTGGACTTGCATCGCAAG	56	<i>MIP1</i>
MIPSeq3bR	ACACCTGTGCATTGCAATCC	60	<i>MIP1</i>
MIPSeq3R	CTTTTAAGTAACTGACTCGC	56	<i>MIP1</i>
MIPSeq4bR	GTTGATTTATCTGTGGAAGC	56	<i>MIP1</i>
MIPSeq4R	CTCCTTCGACCATATTACGG	60	<i>MIP1</i>
MIPSeq5R	GTTGTATTCTTCGAGAACCC	58	<i>MIP1</i>
URA3_INT_FO R	AAACGAAGATAAATCATGTCTG	56	<i>URA3</i>
URA3_R	TTGTGAGTTTAGTATACATGC	56	<i>URA3</i>
XMA	TCTGTGCTTGTCCACC	50	<i>HIS3</i>
XMB	TTGCACAGGTGTTGGC	50	<i>HIS3</i>

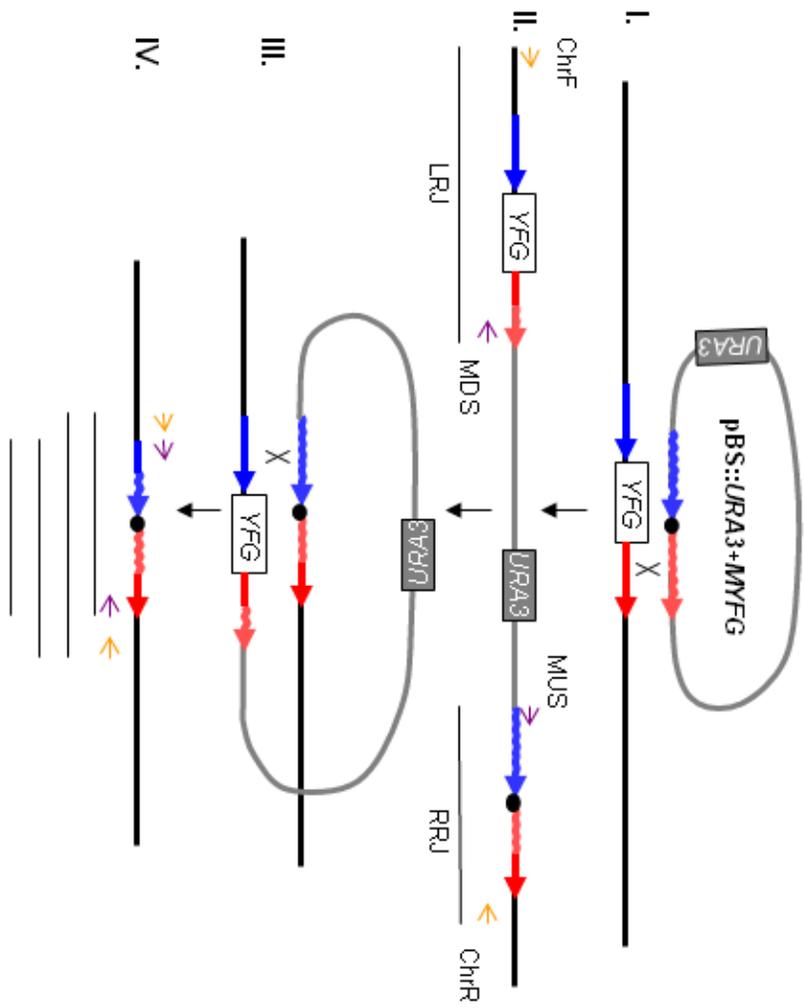
*Primer sequences that anneal to the target gene are in bold. Restriction endonuclease recognition sites engineered into the 5'-end of primers are underlined while normal font only represents regions engineered into the 5'-end of primers for recombination into yeast. *TGA*, represents a stop codon engineered into DHAPF.

Figure 2.3 A. Transplacement vector strategy. The mutant allele of your favorite gene (MYFG) was PCR amplified from an appropriate genome using the MUS and MDS primer set (purple) then ligated into the pBS::URA3 vector. This insert is drawn as two wavy arrows connected by a dot representing either a point mutation or deletion. **I)** The recombinant pBS::URA3+MYFG plasmid was transformed into *Saccharomyces cerevisiae*, harbouring YFG (white box) and the single crossover was selected on SC-URA plates. **II)** The single recombination event between the plasmid, right side wavy arrow, and the correct YFG chromosomal locus, right side solid arrow, was screened by PCR. The left recombination junction (LRJ) was identified by ChrF (top strand) and MDS (bottom strand) primers and the right recombination junction (RRJ) was confirmed by and MUS (top strand) and ChrR (bottom strand) primers. **III)** Strains harbouring the integrated plasmid at the YFG locus were plated onto 5-fluoroorotic acid plates to select for Ura⁻ colonies, which survive by crossing out the URA3 gene for uracil biosynthesis (grey box) via a second recombination event upstream of the first, on the left-hand-side arrow of both the plasmid and the chromosome. **IV)** In the resulting strain the presence of the MYFG point mutation or deletion at the YFG locus was identified by PCR using the following primer sets: ChrF and MDS; ChrF and ChrR; MUS and ChrR; MUS and MDS.

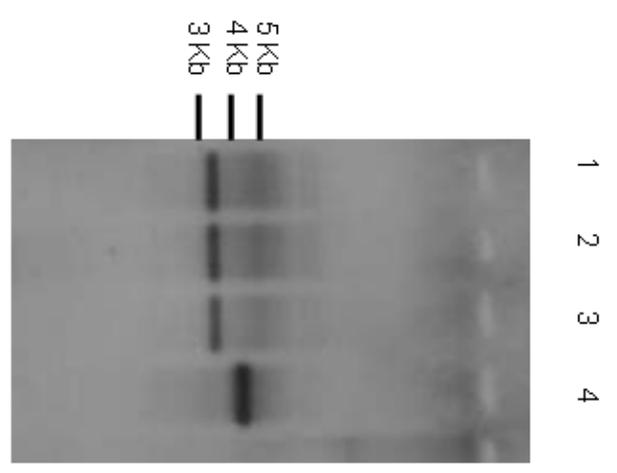
B. *his3-Δ200* strains, harbouring the 1036-bp deletion, were confirmed by Southern blotting. Yeast genomic DNA was digested with *EcoRI* and *XbaI* to identify the genomic 3439-bp band in three separate isolates harbouring *his3-Δ200* (lanes 1 – 3). As a control the isogenic parent S150 was included to compare *his3-Δ200* to *his3-Δ1* (4288-bp fragment, lane 4). Three, four, and five-kb are indicated on the left-hand-side of the blot. For *ade2-101ochre* mutants, the nonfunctional *ade2* gene was confirmed by red pigment

production on YPD as well as adenine auxotrophy or the inability to grow on SC-ADE plates.

A.



B.



to make the transplacement vector pBS::*URA3*::*his3-Δ200*. This plasmid was confirmed by restriction mapping and sequencing and then transformed into S150 using the standard lithium acetate transformation protocol (Gietz and Woods, 2002) and plated onto SC-URA. Integration of the plasmid at the *HIS3* locus in the uracil prototrophs was confirmed by PCR across the sites of chromosomal plasmid insertion using the HUS/XMB (1.2 kb amplicon) and XMA/HDS (2 kb amplicon) primer sets (Figure 2.3A and Table 2.3). The confirmed strains were then grown for two days in SC-URA broth to an OD₆₀₀ of ~1. To screen for colonies in which the cassette had been removed by recombination, 1 ml of the SC-URA cultures was centrifuged at 6000 rpm for 2 minutes and then 900 μl of supernatant was discarded followed by plating of the cells in the remaining 100 μl onto a SC+FOA plate and incubated at 30°C. After four days, FOA-resistant colonies were picked and streaked onto fresh SC+FOA plates, as well as onto YPG plates to test for the ability of the colonies to respire. FOA-resistant colonies were then screened with XMA/XMB and HUS/HDS primer sets, which should yield either a 1.2-kb or a 1.1-kb amplicon, respectively, if the *his3-Δ1* allele is replaced with *his3-Δ200*. Southern blot analysis using a *HIS3*-specific probe was carried out to confirm the colony PCR results (Figure 2.3B). This strain was designated DACS2.

To create *ade2* versions of S150 an ~2.2 kb region containing the *ade2-101ochre* locus was PCR amplified from YPH499 genomic DNA using the AUS/ADSE primer set (Table 2.3), and restriction digested, gel purified and cloned into the *EcoRI* site of pBluescript creating pBS::*ade2-101ochre*. The pBS::*ade2-101ochre* plasmid was then cut with *EcoRI* and the *ade2-101ochre* fragment was gel purified and subcloned into the *EcoRI* site in the multiple cloning site of pBS::*URA3* such that the *URA3* ORF was not

disrupted, creating the transplacement vector pBS::*URA3+ade2-101ochre*. The transplacement vector was then transformed into S150, DACS1, and DACS2 and transformants were plated onto SC-URA. Approximately twenty uracil prototrophic colonies were inoculated into FOA broth and allowed to grow for 48 hours at 30°C with shaking to an OD₆₀₀ of 2.5 followed by plating for isolated colonies on SC+FOA plates. Red colonies, indicative of successful recombination to remove the *ADE2* allele (Figure 2.3A), appeared at a low frequency after four days within the FOA-resistant population and these were picked to YPD plates and streaked out for isolated colonies. Single isolate adenine auxotrophs from YPD plates were confirmed by the inability to grow on SC-ADE drop out media.

URA3 prototrophs were constructed by transformation of S150, DACS2, DACS4, and DACS5 with a 0.86-kb gel purified PCR product amplified from pBS::*URA3* using the primers URA3_R/URA3_INT_FOR (Table 2.3). *URA3* transformants were selected on SC-URA dropout media.

The *HAPI* gene was disrupted by a PCR-mediated gene disruption. The *Candida albicans URA3* (*caURA3*) gene was PCR amplified from the pKT209 plasmid (Sheff and Thorn, 2004) using the DHAPF/DHAPR primers (Table 2.3) followed by gel purification. Engineered into the 5'-ends of the primers were 55 nucleotides of homology for recombination at the *HAPI* locus to remove most of the ORF. Within primer DHAPF a stop codon was introduced just after amino acid 17 of the *HAPI* ORF to avoid engineering a primer within a large AT-rich region just upstream of the start codon. The purified PCR product was transformed into the *HAPI* wild-type strains S150 and DACS5 and transformants were selected on SC-URA. The double crossover recombination event

was screened using the CLONEHAP1F/CAURA3R and CAURA3F/CLONEHAP1R primer sets (Table 2.3) which cover the upstream and downstream region for both crossovers.

The YPH499 background strain was transformed to *ADE2* prototrophy by replacing the *ade2-101ochre* locus with the wild-type *ADE2* locus. *ADE2* was PCR amplified from S150 genomic DNA using ADE2US/ADE2DS primers (Table 2.3) and gel purified. The PCR product was then used for yeast transformation and YPH499 *ADE2* prototrophs were selected on SC-ADE plates. The *his3-Δ200* allele was replaced with the wild-type *HIS3* gene including the bidirectional promoter from the plasmid pBluescript::*HIS3* (BS50, a gift from W. Neupert, University of Munich). The 1.768-kb *Bam*HI genomic fragment harbouring the *HIS3* gene was cut out of the BS50 plasmid and gel purified as described above, followed by transformation into YPH499 and DACY1 and *HIS*⁺ colonies were selected for on SC-HIS plates to create DACY2 and DACY3 respectively. Due to the position of the deletion at the *his3-Δ200* allele (-205 to 835, relative to the first ATG of the *HIS3 ORF*, <http://www.yeastgenome.org/alleletable.shtml>), recombination of the *HIS3* fragment would have to restore the bidirectional promoter as well as the *HIS3 ORF*.

2.3.4 Assessment of mtDNA maintenance, replication fidelity and respiratory competence

Slides for fluorescence microscopy of 4',6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich, St. Louis, MO, USA) stained cultures were prepared to monitor the presence of mtDNA. Cultures grown for 96 hours at 30°C or 37°C as described for respiratory competence were diluted to 0.7 OD₆₀₀ in 5 ml of YP10D broth and DAPI was added to a final concentration of 2.5 µg/ml. Cultures were incubated for 40 minutes at

30°C with shaking, followed by washing of 1.6 ml of cells with 1 ml ddH₂O and resuspension in 150 µl ddH₂O. 40 µl of the DAPI-stained cells were combined with an equal volume of 0.7% low gelling agarose (Mandel, Guelph, ON) made in ddH₂O, transferred to a slide for observation by Zeiss Epi-fluorescent microscopy.

Erythromycin-resistance was monitored according to (Young *et al.*, 2006, Hu *et al.*, 1995). Briefly, 5×10^8 – 1×10^9 cells were spread onto each of two YG-Er plates from an overnight culture grown in 30 ml of a 50:50 mixture of YPG:YP2D. Appropriate dilutions were also spread onto YPG plates to determine the actual number of respiratory-competent cells spread on the YG-Er plates.

Respiratory competence was monitored over 96 hours. Yeast were first grown in YPG broth for 48 hours to ensure a 100% respiring culture. Cultures were then diluted to 0.01 OD₆₀₀ in YP10D broth for growth with shaking at 30°C or 0.16 OD₆₀₀ for growth with shaking at 37°C. Cultures were diluted down to 0.01 OD₆₀₀ (for 30°C) or 0.16 OD₆₀₀ (for 37°C) every 24 hours and glucose levels were continually monitored (Young and Court, 2004) to ensure sufficient (> 1%) glucose to avoid selection of respiratory-competent cells. At 24, 72, and 96 hours of growth, cultures were diluted down and then spread plate to give significant plate counts on YP2D plates. Respiratory competence was determined by diagnosing respiration deficiency based on Raut's observation that white colonies following growth on tetrazolium plates are unable to respire whereas red colonies are respiratory competent (Ogur *et al.*, 1957). YP2D plates grown at 30°C for two days were overlaid with 2,3,5-triphenyl tetrazolium chloride (TTC) top agar (1% Bacto agar, 0.067 M phosphate buffer, pH 7.0, 0.1% TTC) as described by (Ogur *et al.*, 1957). Plates were incubated in the dark for at least one hour and percent respiratory

competence was calculated as the number of red colonies divided by the total number of colonies multiplied by 100.

2.4 Results

2.4.1 YPH499 based strains and the *his3-Δ200* S150 strains show temperature-sensitive loss of respiratory competence.

To investigate the contributions of *ADE2*, *MRM1* promoter/*his3-Δ200*, *HAP1*, and *MIPI*, cells harbouring different combinations of these alleles (Table 2.2) were grown in YPD for four days. Cultures were transferred to fresh media each day to ensure sufficient glucose, thereby preventing selection for respiratory competent cells. Due to the variability in the data (Table 2.4), median values are presented to show trends (Figure 2.4). Respiratory competence of YPH499 ranged from 52.2 to 92.9% by the fourth day of growth at 30°C (Table 2.4). All other strains examined in the study maintained at least 89% respiratory competence over four days growth at 30°C (Table 2.4). As a more sensitive assay, we also examined differences in respiratory competence of the various strains at 37°C. The S150 background strain and two mutants made thereof, DACS4 (*ade2*) and DACS9 (*hap1*), maintained greater than ninety percent respiratory competence over the four day period at 37°C (Figure 2.4 and Table 2.4). In contrast, the presence of the *his3-Δ200* allele in the same background strain (DACCS2) caused a slight decrease in respiratory competence by day three and a more dramatic decrease was observed by day four where the median value was 50%. The loss of *ADE2* function in DACCS5 (*ade2*) in a S150 *his3-Δ200* background was not associated with any further loss of respiration compared to DACCS2 (*ADE2*, *his3-Δ200*). Interestingly, when the *HAP1*

Figure 2.4 Respiratory competence of various yeast strains grown for four days at 37°C in YP10D. Respiratory competence was determined as described in Materials and Methods, day 1 (D1), day 3 (D3), and day 4 (D4). The relevant genotypes for each are shown below strain names/abbreviations. Y1 (DAC Y1), Y2 (DAC Y2), Y3 (DAC Y3), S2 (DAC S2), S4 (DAC S4), S5 (DAC S5), S8 (DAC S8), and S9 (DAC S9). ^aA representative experiment is shown for S150, DAC S4 (S4) and DAC S9 (S9) as all strains maintained median respiratory competence greater than 90% over four days growth in YP10D (Table 2.4).

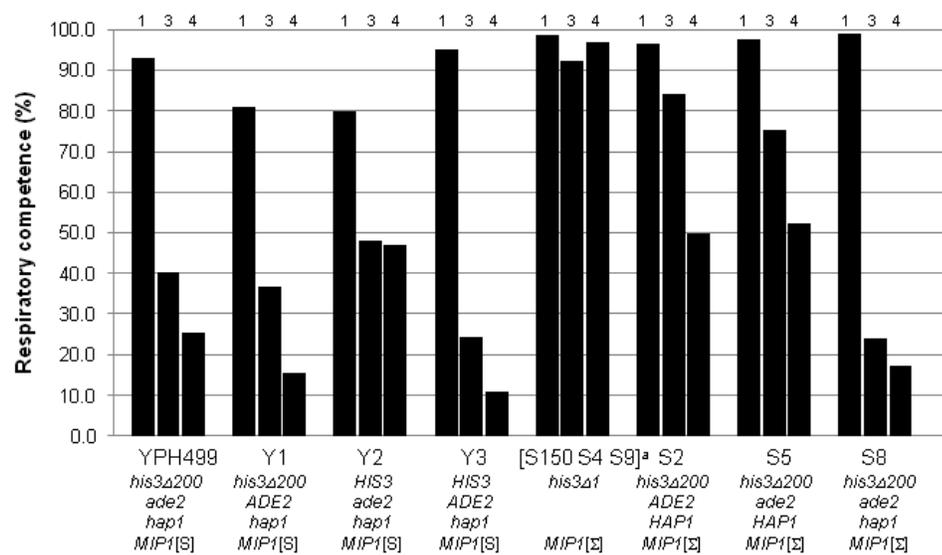


Table 2.4 Percent respiratory competence for various laboratory strains

Strain	Temp. (number of replicates)	Incubatio n (Days)	Respiratory Competence (%)			
			Median	Average	Standard Deviation	Range
YPH499	30°C (3)	1	98.7	98.3	2.0	96.1 - 100
		3	86.3	76.7	19.5	89.5 - 54.3
		4	81.9	75.5	21.4	52.2 - 92.9
YPH499	37°C (3)	1	93.1	87.9	11.3	74.9 - 95.7
		3	40.2	35.4	17.9	15.7 - 50.4
		4	25.6	29.7	20.8	11.3 - 52.2
DACY1	30°C (3)	1	98.7	99.0	0.6	98.6 - 99.7
		3	97.9	97.3	1.8	98.6 - 95.2
		4	96.5	94.8	3.6	97.3 - 90.7
DACY1	37°C (3)	1	81.0	79.6	13.4	65.5 - 92.2
		3	36.6	30.1	26.0	1.5 - 52.2
		4	15.5	26.9	29.8	4.5 - 60.7
DACY2	30°C (3)	1	96.1	95.6	2.8	92.6 - 98.2
		3	92.0	92.0	1.2	90.8 - 93.1
		4	89.6	90.4	1.5	89.4 - 92.1
DACY2	37°C (3)	1	79.9	79.9	4.9	75.0 - 84.9
		3	47.9	47.8	8.7	39.0 - 56.4
		4	46.9	38.5	18.4	17.4 - 51.2
DACY3	30°C (3)	1	99.6	99.5	0.7	98.7 - 100
		3	99.0	99.0	0.5	98.0 - 99.5
		4	98.8	98.4	0.9	97.2 - 98.9
DACY3	37°C (3)	1	95.2	95.4	3.3	92.3 - 98.9
		3	24.3	27.8	26.8	3.0 - 56.3
		4	11.0	15.3	10.4	7.7-27.2

Strain	Temp. (number of replicates)	Incubatio n (Days)	Respiratory Competence (%)			
			Median	Average	Standard Deviation	Range
S150	30°C (3)	1	99.4	99.5	0.5	98.7 - 100
		3	99.2	98.6	1.0	97.4 - 99.2
		4	98.9	98.6	1.4	97.0 - 99.8
S150	37°C (4)	1	98.7	99.0	0.7	98.7 - 100
		3	92.3	91.7	8.2	82.5 - 99.7
		4	97.0	93.9	7.4	82.9 - 98.9
DACs1 ^a	30°C (1)	1		99.5		
		3		99.5		
		4		98.9		
DACs1	37°C (2)	1		99.4		98.7 - 100
		3		98.6		97.5 - 99.7
		4		98.2		97.4 - 98.9
DACs2	30°C (3)	1	99.6	99.4	0.6	98.7 - 99.8
		3	98.0	97.3	0.9	96.5 - 98.7
		4	98.0	96.6	0.9	92.5 - 99.3
DACs2	37°C (3)	1	96.4	96.4	3.1	93.3 - 99.6
		3	83.4	84.1	6.1	87.4 - 90.5
		4	50.0	48.9	8.3	40.1 - 56.5
DACs3	30°C (2)	1		96.3		94.7 - 97.9
		3		93.9		89.2 - 96.5
		4		93.8		88.7 - 98.9
DACs3	37°C (2)	1		96.3		94.7 - 97.9
		3		92.9		89.2 - 96.5
		4		93.8		88.7 - 98.9
DACs4	30°C (3)	1	2 replicates	99.0	1.0	98.3-99.6
		3		97.7	1.2	96.7 - 99.2
		4		99.5	0.0	97.9 - 99.6
DACs4	37°C (3)	1	99.6	98.2	2.6	95.3 - 99.8
		3	97.4	91.7	11.4	78.5 - 99.1
		4	98.1	89.6	15.0	72.3 - 98.4

Strain	Temp. (number of replicates)	Incubatio n (Days)	Respiratory Competence (%)			
			Median	Average	Standard Deviation	Range
DAC55	30°C (3)	1	99.6	99.6	0.0	99.1 - 100
		3	99.4	99.3	0.6	98.6 - 99.7
		4	96.9	95.6	5.2	89.9 - 100
DAC55	37°C (3)	1	97.7	97.4	1.1	97.7 - 98.3
		3	75.2	66.4	17.1	46.7 - 77.3
		4	52.2	46.7	21.0	23.5 - 64.4
DAC57	30°C (1)	1		95.2		
		3		100.0		
		4		99.3		
DAC57	37°C (1)	1		98.2		
		3		85.6		
		4		75.8		
DAC58	30°C (3)	1	99.8	99.8	0.1	99.7 - 100
		3	99.3	99.0	0.6	98.4 - 99.5
		4	98.7	98.7	1.3	97.3 - 100
DAC58	37°C (3)	1	98.9	98.8	0.7	98.1 - 99.4
		3	24.0	30.4	16.9	17.8 - 49.6
		4	17.2	12.9	11.4	0 - 21.7
DAC59	30°C (3)	1	100.0	99.7	0.5	99.7 - 100
		3	98.2	98.0	0.4	98.3 - 97.7
		4	97.4	96.9	1.2	95.6 - 97.8
DAC59	37°C (3)	1	99.8	99.3	1.1	98.0 - 100
		3	95.5	96.4	1.5	95.5 - 98.0
		4	94.8	95.1	0.7	94.6 - 95.9

^aOnly one or two replicates were performed for strains highlighted

gene was disrupted in the S150 *his3-Δ200 ade2 hap1* strain DACS8, respiratory competence was reduced by three-fold on days three and four in comparison to DACS2 and DACS5.

After four days at 37°C YPH499, DACY1, and DACY3 only maintained median levels of 26, 16, and 11% respiratory competence while ~50% of DACY2 cells were respiratory competent (Figure 2.4). While it may appear that DACY2 has an increased ability to maintain respiratory competence at 37°C it is important to note that all YPH499 backgrounds ranged from ~5 to 61% respiratory competence, suggesting reduced levels for all YPH499 derivatives (Table 2.4).

2.4.2 Loss of respiratory competence is accompanied by loss of mtDNA in the YPH499 but not in the S150 background.

The loss of respiration ability in the YPH499 strain is correlated with a loss of mtDNA (Figure 2.5 median of 0 and average of 0.6 punctate mitochondrial nucleoids per cell) as indicated by DAPI staining of the cultures grown over four days at 37°C. In contrast, the loss of respiratory competence in the S150 background strains (DACs5 and DACs8) was not accompanied by the loss of mtDNA (Figure 2.5) suggesting different mechanisms may be responsible for the loss of respiratory ability in the two genetic backgrounds.

2.4.3 Low levels of point mutations were detected in all S150 background strains, DACY1, and DACY3 while higher levels were detected in YPH499 and DACY2.

Erythromycin-resistance (Er-R) was used as an indication of the level of point mutations in the mitochondrial genome. Due to the loss of mtDNA at 37°C in the YPH499 background, erythromycin-resistance could only be tested in cells grown at 30°C. All strains constructed in the S150 background produced on average 2.2

Figure 2.5 mtDNA maintenance in cells grown at 37°C for four days. **A.** DAPI stained puncta were counted in thirty DAPI -stained cells of each strain, and the median for each strain is shown as a line. **B.** DAPI staining. After four days growth in YP10D at 37°C cultures were stained with DAPI for 40 minutes as described previously (Young *et al.*, 2006) M, punctuate mitochondrial staining; N, nuclear staining.

erythromycin-resistance cells/ 1×10^8 cells (Figure 2.6 and Table 2.5). YPH499 produced the highest level of erythromycin-resistant cells, $15/1 \times 10^8$ on average (Figure 2.6). In this background, restoration of the *ADE2* gene in DACY1 and DACY3 (*HIS3 ADE2*) caused a decrease ($0.5/1 \times 10^8$ and $2.8/1 \times 10^8$ respectively) in erythromycin-resistant cells compared to YPH499. Restoration of only the *MRMI* promoter in DACY2 did not have an effect on the frequency of point mutations when compared to the YPH499 background, indicating that the *ADE2* locus is the contributing factor in decreased erythromycin-resistance in DACY3.

2.5 Discussion

Recent descriptions of respiratory defects associated with S288c-related strains, specifically those harbouring *MIP1*[S] (Baruffini *et al.*, 2007b) and *hap1* (Sherman, 2002) alleles, impelled us to evaluate the effects of these alleles in combination with potentially influential auxotrophic markers, namely *ade2-101ochre* and *his3Δ-200* (*mrm1*).

We observed a temperature-sensitive phenotype at 37°C for the S288c-related (YPH499) *MIP1*[S] strains used in our study. Although different experimental approaches were used (see below), less than 40% respiratory competence was detected in YPH500, presumably *MIP1*[S] (Zhang *et al.*, 2003). In comparison, the S150 *MIP1*[Σ] strain maintained a high level of respiratory competence (>90%) throughout the experiment (Figure 2.2). Similarly, greater than 90% of DL1 cells (isogenic or congenic to D273-10B, Zhang *et al.*, 2003), presumably *MIP1*[Σ], and D273-10B/A1 (*MIP1*[Σ], Baruffini *et al.*, 2007b) maintained mitochondrial function after prolonged treatment at 37°C. These data reinforce the conclusion of Baruffini *et al.*, 2007b regarding the

Figure 2.6 Point mutations in mtDNA. Relative fidelity of mtDNA replication was assessed as described previously (Young *et al.*, 2006). Each of three independent transformants for each strain was replicated three times. Averages are shown as open bars, with the standard deviation for each, and the median for each strain is displayed as solid black bars. ‘All S150’ refers to S150 and DACS1 – DACS9 (Table 2.2).

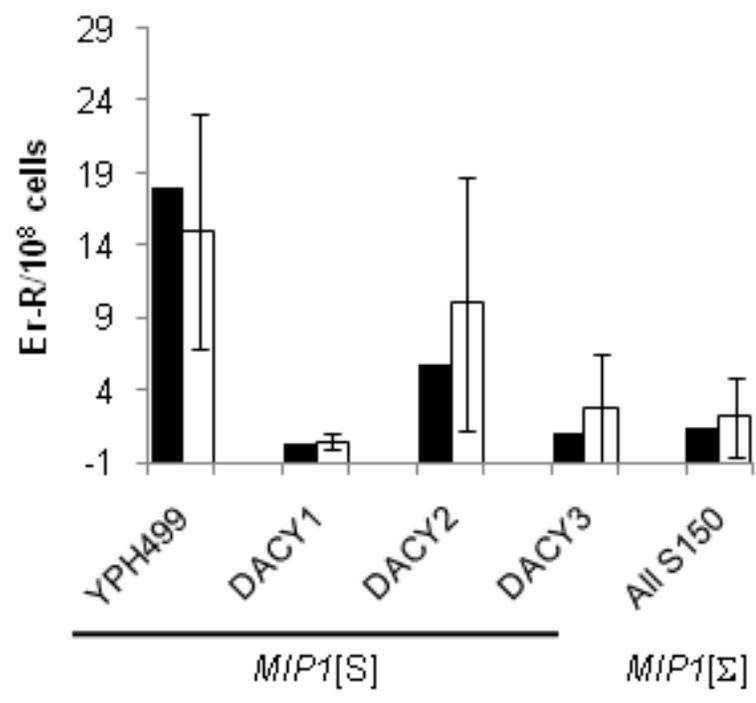


Table 2.5 Erythromycin-Resistance per 10⁸ cells

Strain	Number of Replicates	Median	Average	Standard Deviation	Range
YPH499	9	18.0	15.0	8.1	3.9 - 25.2
DACY1	13	0.2	0.5	0.6	0.1 - 2.3
DACY2	9	5.8	10.0	8.8	1.8 - 24.5
DACY3	10	1.0	2.8	3.8	0.1 - 10.4
S150	8	2.2	2.6	2.7	0.2 - 8.6
DACS1	7	2.1	2.2	2.3	0.2 - 6.6
DACS2	9	1	3	3.2	0.4 - 8.2
DACS3	7	3.1	3.3	1.8	0.1 - 5.6
DACS4	6	1.4	2.3	2.6	0 - 7.1
DACS5	4	0.6	1	1	0.24 - 2.4
DACS6	6	1.7	2.4	2.3	0.34 - 6.2
DACS7	8	1.7	2.1	1.4	0.5 - 3.7
DACS8	3	1.1	1.1	0.4	0.7 - 1.6
DACS9	3	0.2	0.2	0.2	0 - 0.2
All S150	60	1.3	2.2	2.2	0 - 8.6

contribution of the *MIP1* allele to the maintenance of mitochondrial function.

However, BY4741, BY4742, and W303-1B, all harbouring *MIP1*[S], displayed greater than 85% respiratory competence (Figure 2.2) suggesting that the *MRM1* allele drastically reduced the effect of the *MIP1*[S] allele in these strains. This is supported by the YPH499 data and the S150 (*MIP1*[Σ]) strains harbouring the *mrm1* allele, that showed decreased mitochondrial function after four days at 37°C. However, replacement of the *mrm1* with *MRM1* in YPH499-derived strains did not restore respiratory competence to levels observed in S150. Zhang *et al.* also complemented YPH500 cells with *MRM1* on a plasmid (Zhang *et al.*, 2003), which resulted in an increase in respiratory competence from about 34% to only 68% suggesting a role for other factors (see below). A similar approach was taken by Zhong *et al.* (Zhong *et al.*, 2004), in the S288c-related strain FGY3 (parent strain YPH274, Table 2.1). In this case, almost all of the cells retained respiratory function after incubation for 96 and 194 hours at 37°C in both the *mrm1* and *MRM1*-complemented strains. The reason for this discrepancy is not clear.

The *HAP1* gene, which encodes a heme-dependant transcription factor, has obvious links to mitochondrial status, and it has been suggested by Sherman that *HAP1* strains such as D273-10B be used for mitochondrial studies (Sherman, 2002). Nonetheless, strains with the S288c background (*hap1*) are widely used for this work (Table 2.1). The effect of the *HAP1* allele may be enhanced by the *MIP1*[S] allele that is present in the S288c-derived strains (Baruffini *et al.*, 2007b). In the *MIP1*[Σ] strain used in the current study, respiratory incompetence was not observed in DACS9 (*hap1* Δ).

However, there was a three-fold drop in respiratory competence in DACS8, with *MIP1*[Σ] in combination with *mrm1*, *hap1* Δ and *ade2-101ochre*, in comparison to DACS2 (*mrm1*) and DACS5 (*mrm1*, *ade2-101ochre*). This drop likely reflects negative synergism between *hap1* Δ and other genes required for respiration (*his3- Δ 200/mrm1*).

The other important consideration is the method used to estimate levels of respiratory competence. In the current work, and that of Baruffini *et al.*, 2007b) cells were maintained in a replicative state of growth, wherein they were repeatedly subcultured (Young and Court, 2004) or replica-plated (Baruffini *et al.*, 2007b). In contrast, the methods of Zhong *et al.*, 2004) and Zhang *et al.*, 2003) involve chronologically aging the cultures for at least four days in YPD or YPG before plating to determine the number of *petites*. Growing these strains into the post-diauxic and stationary phases of growth would select for respiration, due to the fact that cells survive by consuming nonfermentable by-products of glucose fermentation (Werner-Washburne *et al.*, 1996). Respiring cells may also have a selective advantage when in or entering stationary phase. Given the similar levels of respiratory competence between BY4741 (Zhong *et al.*, 2004) and BY4742 (Baruffini *et al.*, 2007b), these methodological difference cannot explain the FGY3 results (Figure 2.2, Zhong *et al.*, 2004). The FGY3 respiratory competence levels were expected to resemble those of YHP499/500, because all three of these strains are derived, through multiple crosses, from YNN216, which is congenic to S288c (discussed by Kumar *et al.*, 2003). Thus, it would be expected that FGY3 is also *MIP1*[S] and *hap1*. However, it has been shown by Kumar (Kumar *et al.*, 2003), that the parents of YNN216, namely YNN214 and YNN215 are not isogenic, and are polymorphic for the *ECM38* gene that encodes the γ -glutamyl transpeptidase

enzyme. Therefore, it is possible that other polymorphisms have been differentially propagated in the lineages leading to YPH499/YPH500 and FGY3 (YPH274), and that some of these differences may contribute to respiratory competence at high temperature.

The lack of respiratory competence could reflect a continuum of mutations, from the accumulation of point mutations to the complete loss of mtDNA. All strains constructed in the S150 background maintained approximately the same level of point mutations in the mitochondrial genome, as indicated by the acquisition of erythromycin-resistance (Figure 2.6). YPH499 and DACY2 (*HIS3 ade2-101ochre*) had high levels of erythromycin-resistance compared to the S150 strains, and to YPH499 strains in which the *ADE2* allele had been restored, DACY1 (*ADE2*, $p < 0.05$) and DACY3 (*HIS3 ADE2*, $p < 0.05$). Thus, at 30°C, the *ade2-101ochre* allele contributes to the mutational load on the mitochondrial genome in the YPH499 background strain. This defect is therefore linked to the *MIP1*[S] allele because there was no obvious difference in erythromycin-resistance in any of the S150 *ade2-101ochre MIP1*[Σ] strains. Presumably the less effective polymerase is more sensitive to reduced levels of dATP in mitochondria, thereby contributing to the moderate increase in point mutations.

In addition, the allele at the *MIP1* locus contributes to the degree of mtDNA loss associated with the loss of respiratory competence. In the *MIP1*[Σ] S150 strain, 17% of cells are *rho*^o after four days at 37°C; this value is increased to around 30% when the *mrm1* allele is introduced, singly, or in combination with *ade2* and/or *hap1* Δ (Figure 2.5). In contrast, we did observe a dramatic decrease in DAPI-stained mtDNA in YPH499 background strain (Figure 2.5). Only 23% of the DAPI-stained YPH499 cells grown at

37°C on day four maintained mtDNA, which is at the level of respiratory competence generally observed on the same day (Figures 2.4 and 2.5). In the W303 (*MIP1*[S], *MRM1*) background ~40% of cells were *rho*⁻ and 24% were *rho*^o after five days at 38.5°C (Baruffini *et al.*, 2007b). These observations again suggest synergistic or additive effects of *mrm1* with *MIP1*[S].

An interesting observation was made regarding red/white screening using the *ade2* allele. When six DACS8 white colonies were picked, from YPD plates from the last day of the respiratory competence experiment, to YPG plates, these colonies were able to grow. In combination with the DAPI-results, this suggests that visible red colour development in *ade2* strains depends on a threshold level of mitochondrial function.

In conclusion, the maintenance of mitochondrial function in two different *S. cerevisiae* strains, S150 and the S288c-related strain YPH499, is affected, to different degrees, by several genetic factors including *mrm1*, *hap1*, *MIP1*[S], and *ade2-101ochre*. With the accumulation of more strain-specific DNA sequence data, such as that from the yeast genome resequencing project (<http://www.sanger.ac.uk/Teams/Team71/durbin/sgrp/index.shtml>), it is likely that other strain-specific alleles will be identified that may contribute to mtDNA maintenance. It is already clear that this process requires a “plethora” of genes (Contamine and Picard, 2000). With such information strains can be chosen on the basis of robust maintenance of mitochondrial function, or strains with *mrm1/his3Δ200*, *hap1*, *MIP1*[S], and *ade2* alleles could be used for studies of more subtle effects. Increased knowledge of the genomes of individual strains will also enhance comparisons of experimental results obtained by different research groups. All yeast strains were not engineered equally!

CHAPTER 3 The carboxyl-terminal extension on fungal mitochondrial DNA polymerases: Identification of a critical region of the enzyme from *Saccharomyces cerevisiae*

Contributions of authors

Most of the work in this Chapter has been published in

Young, M.J., Theriault, S.S., Li, M., and Court, D.A. (2006) *Yeast* **23**:101-116.

Steven Theriault and Mingyi Li began this project, and were responsible for designing the original truncation mutants (*mip1Δ175*, *mipΔ216*, *mipΔ279*, *mipΔ351*) and generating the mutant version of the *MIP1* gene. Initial trials of expressing these variants in yeast were made by these students, but issues with strain construction were not resolved. This work is described in their M.Sc. theses.

All of the yeast strains expressing the truncated polymerases, and the subsequent laboratory and bioinformatic analyses were carried out as part of the Ph.D. work described in this thesis. The last unpublished mutants were analyzed with the assistance of Suman Lakhi (*mip1Δ205*, *mipΔ222*, *mipΔ241*).

3.1 Abstract

Fungal mitochondrial DNA (mtDNA) polymerases, in comparison to their metazoan counterparts, harbour unique carboxyl-terminal extensions (CTEs) of varying lengths and unknown function. To determine the essential regions of the 279-residue CTE of the yeast enzyme (Mip1p), several CTE-truncation variants were expressed in *Saccharomyces cerevisiae*. Respiratory competence of *mip1* Δ 175 and *mip1* Δ 205 cells, in which Mip1p lacks the C-terminal 175 and 205 residues respectively, were indistinguishable from that of wild-type. In contrast, strains harbouring Mip1p Δ 351 and Mip1p Δ 279 rapidly lose mtDNA. Likewise *mip1* Δ 241 and *mip1* Δ 222 transformants lacked the ability to respire. Approximately one in six *mip1* Δ 216 transformants grew on glycerol, albeit poorly. Fluorescence microscopy and Southern blot analysis revealed lower levels of mtDNA in *mip1* Δ 216 cells, and the rapid loss of mtDNA during fermentative, but not respiratory growth. Therefore, only the polymerase-proximal segment of the Mip1p CTE is necessary for mitochondrial function. Comparison of this essential segment with the sequences of other fungal mtDNA polymerases revealed novel features shared among the mtDNA polymerases of the *Saccharomycetales*.

3.2 Introduction

Mitochondria are the sites of oxidative phosphorylation in eukaryotic cells. These organelles harbour mitochondrial genomes or mitochondrial DNAs (mtDNAs) that encode subunits of the respiratory chain, rRNAs, tRNAs, and in some cases, ribosomal proteins. In obligately-aerobic organisms, such as humans, base-substitution, deletion, and duplication mutations in mtDNA result in a variety of degenerative diseases. These

mutations most likely result from errors in mtDNA replication (reviewed by Kaguni, 2004).

The ascomycetous fungus *Saccharomyces cerevisiae* has been used extensively for studies of mtDNA replication, due in part to its ability to bypass the need for respiration in the presence of fermentable carbon sources. Haploid yeast cells typically contain 40-50 copies (Grimes *et al.*, 1974) of an 86-kb mtDNA with a circular genetic map (Foury *et al.*, 1998). Studies of mtDNA replication in this organism have revealed a complex process that likely includes bidirectional replication from origins (reviewed in Lecrenier and Foury, 2000) and/or rolling-circle type replication (Maleszka *et al.*, 1991), which may be involved in creating concatamers that are transmitted to developing buds during asexual reproduction (Ling and Shibata, 2002; Ling and Shibata, 2004).

Mitochondrial DNAs in yeast are found in nucleoids, membrane-associated, DNA-protein complexes (Williamson and Fennell, 1979; Miyakawa *et al.*, 1987), that have been implicated in mitochondrial segregation (Azpiroz and Butow, 1993; Nunnari *et al.*, 1997) and are the sites of mtDNA replication (Meeusen and Nunnari, 2003). At least 20 different proteins can be crosslinked to mtDNA (Kaufman *et al.*, 2000), including the soluble DNA binding proteins, Rim1p (single-stranded DNA binding protein) Abf2p, a high-mobility group protein involved in maintenance, segregation, recombination and copy number control of mtDNA (Diffley and Stillman, 1991; Zelenaya-Troitskaya *et al.*, 1998; MacAlpine *et al.*, 1998) and Mgm101p, which may be involved in repair of oxidative damage to DNA (Meeusen *et al.*, 1999). An intriguing subset of mtDNA associated proteins are bifunctional (Kaufman *et al.*, 2000), including Ilv5p, an acetohydroxyacid reductoisomerase, that links amino acid starvation to nucleoid number

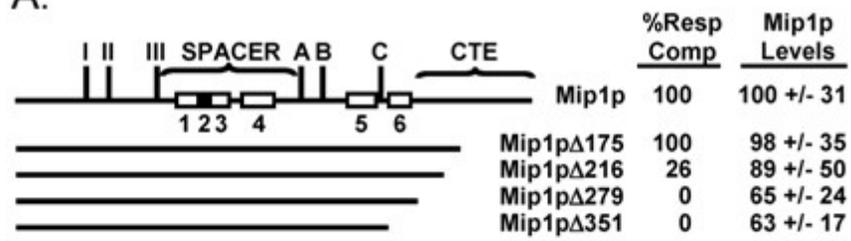
(Zelenaya-Troitskaya *et al.*, 1995; MacAlpine *et al.*, 1998) and mitochondrial aconitase, a protein that may be a key link between mitochondrial metabolism and mtDNA maintenance, in addition to its enzymatic role in the Krebs's cycle (Chen *et al.*, 2005).

Nucleoids are physically associated with the mitochondrial inner membrane, but a nucleoid specific, inner membrane component has yet to be identified. A subset of nucleoids are associated with the integral outer membrane protein Mmm1p (Hobbs *et al.*, 2001) suggesting a “two membrane spanning complex” (TMS, Meeusen and Nunnari, 2003). A mtDNA binding protein, Mgm101p, (Meeusen *et al.*, 1999) and Mip1p, the mitochondrial DNA polymerase, are also associated with the TMS (Meeusen and Nunnari, 2003). Recently, a second outer membrane protein, Mmm2p, was identified that is also associated with nucleoids, but appears to be part of a complex distinct from that harbouring Mmm1p (Youngman *et al.*, 2004).

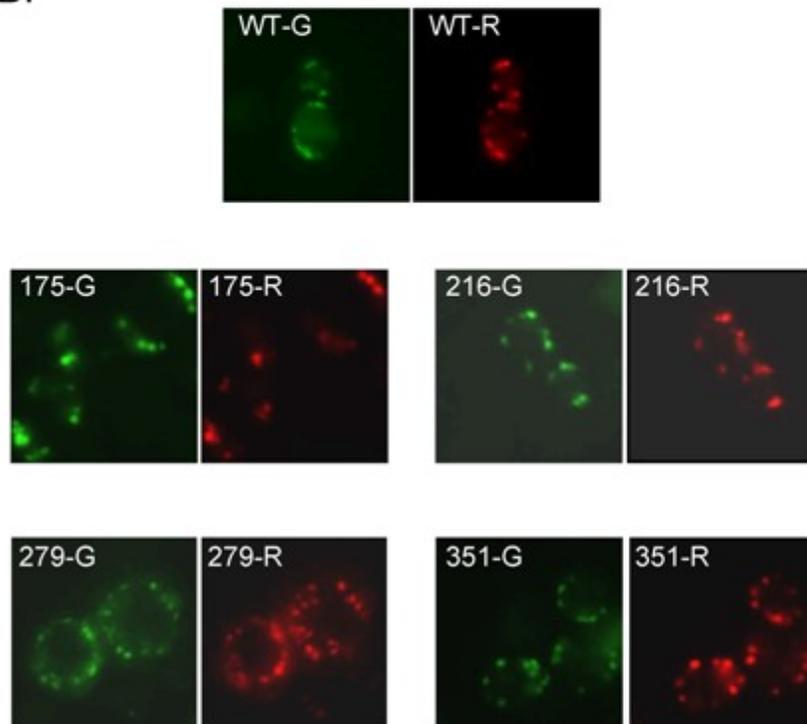
Mip1p is a member of the subclass of gamma (γ) or mtDNA polymerases within the family A group of DNA polymerases (Ito and Braithwaite, 1990; Blanco *et al.*, 1991). Within the polymerase domain of these enzymes exist three conserved sequence motifs: Pol A, Pol B, and Pol C (Ito and Braithwaite, 1990; reviewed in Kaguni, 2004; Figure 3.1A) and a proofreading 3'-5' exonuclease domain characterized by three conserved sequence motifs, namely Exo I, Exo II, and Exo III (Bernad *et al.*, 1989), all of which are present in the mitochondrial enzymes (reviewed in Kaguni, 2004, Figure 3.1A). MtDNA polymerases can be differentiated from other members of family A based on four conserved signature sequences (γ 1 – γ 4) that are located within a “spacer region” between the exonuclease and polymerase domains, and two regions (γ 5 and γ 6) that flank the carboxyl-terminal polymerase motif (Pol C, Figure 3.1A). Conserved sequence

Figure 3.1. Schematic diagram of the wild-type Mip1p (top line) and four truncation variants. I, II, and III indicate the location of the three conserved exonuclease motifs that define the 3'-5' exonuclease domain, while A, B, and C indicate the location of the three conserved polymerase motifs that define the polymerase domain (Kaguni, 2004). Boxes numbered 1 through 6 in the wild-type Mip1p indicate the positions of the mtDNA polymerase γ -specific sequences. The locations of γ 1- γ 4 are taken from Luo et al., 2005; γ 5 and γ 6 comprise residues 805-876 and 905-959, respectively, of the Mip1p preprotein (L.S. Kaguni, personal communication). The γ 2 sequence is shaded black to clearly delineate it from γ 1 and γ 3. "CTE" indicates the location of the carboxyl-terminal extension unique to fungal polymerases. Variants with truncations in this region are indicated below. The "% Resp Comp" column indicates the percentage of between 36 and 67 independent transformants of each strain that were able to respire after transfer from the initial SC-HIS plate to YPG plates. The "Mip1p level" column indicates the relative levels of the Mip1p variants, determined as described in Materials and Methods. The value for wild-type Mip1p was arbitrarily set to one. **B)** Fluorescence of Mip1p-GFP variants. The left- and right-hand panels show Mip1p-GFP (G) and dsRED (R) fluorescence, respectively, from the same field of view. Truncation variants are indicated by the number of amino acid residues removed from the C-terminus; WT indicates wild-type cells. Images obtained from 0.2 μ m slices are presented.

A.



B.



elements γ_1 , γ_3 , and γ_4 may function as part of the polymerase catalytic domain; the roles of γ_5 and γ_6 have not been investigated (Luo and Kaguni, 2005) .

In comparison to the other γ polymerases, there are two remarkable features of the yeast enzyme: the apparent lack of an accessory subunit (for example, Vanderstraeten *et al.*, 1998), and the presence of a long extension to the C-terminal side of the γ_6 region (Foury, 1989; Figure 3.1). γ polymerases from *Drosophila* (Olson *et al.*, 1995), *Xenopus* (Insdorf and Bogenhagen, 1989), mouse (Johnson *et al.*, 2000) and humans (Gray and Wong, 1992) consist of a catalytic core and an accessory subunit, which stimulates catalysis and increases the processivity of the catalytic core (reviewed in Kaguni, 2004). The carboxyl-terminal regions of the accessory subunits resemble class IIa aminoacyl-tRNA synthetases, suggesting a role for RNA primer binding during initiation of mtDNA synthesis (Carrodeguas and Bogenhagen, 2000; Carrodeguas *et al.*, 1999; Carrodeguas *et al.*, 2001; Fan *et al.*, 1999; Fan and Kaguni, 2001; Iyengar *et al.*, 2002).

The C-terminal extension (CTE) on the fungal enzymes has remained enigmatic. The only clue to its function was the analysis of a yeast mtDNA polymerase harbouring a single amino acid substitution at residue 1001 of the preprotein (Hu *et al.*, 1995). Cells harbouring this mutant, form respiratory incompetent *petites* at increased levels and have a 50-60 fold increase in the frequency of erythromycin-resistant cells, indicative of mtDNA point mutations in the 21S rRNA gene (Hu *et al.*, 1995).

Using the genome sequence data now available, we have carried out a more comprehensive search for CTEs in mtDNA polymerases, analyzed the distribution of these sequences and examined them for conserved features. Based on the resulting information, C-terminal truncation variants of Mip1p were generated and the ability of

cells harbouring these mutant polymerases to replicate and maintain mtDNA was analysed.

3.3 Materials and Methods

3.3.1 Strains and growth conditions

S150 (*leu2-3,112*, *his3-Δ1*, *trp 1-289*, *ura 3-52*, Steger *et al.*, 1990) was the wild-type strain used for expression of Mip1p truncation variants in this study. Yeast were grown on various media including: YPG (1% yeast extract, 2% peptone, 3% glycerol); YG-Er (2% yeast extract, 3% glycerol, 50 mM potassium phosphate, pH 6.5, 4 g/L erythromycin); YP2D (1% yeast extract, 2% peptone, 2% dextrose); YP10D (1% yeast extract, 2% peptone, 10% dextrose); YPL (1% yeast extract, 2% peptone, 2% lactic acid, pH 5.2); SC-HIS (0.67% Difco™ nitrogen base without amino acids, 4% dextrose, all amino acids except histidine, pH 5.6).

3.3.2 Strain Construction

Mutant versions of the mtDNA polymerase gene were generated by site-directed mutagenesis (Kunkel *et al.*, 1987) using the Muta-Gene kit from Bio-Rad (Mississauga, Canada). Base pairs and amino acid residue numbers refer to the 1254-residue protein that would be encoded from the start codon downstream of the transcriptional start point identified by Foury (1989). The template DNA was obtained from plasmid pBluescript::*MIP1*-F3, which contains basepairs (bp) 2538 to 3762 of the coding sequence for the Mip1p preprotein and extends 31 bp past the stop codon. Primers were used to introduce a stop codon [**BOLD**] followed by a *Bgl*II site [underlined] for insertion of the selectable marker *HIS3* into the *MIP1* gene. The stop codons were introduced at positions 3114, 2925, and 2709 of the *MIP1* coding sequence to create the genes *mip1Δ216*, *mip1Δ279*, and *mip1Δ351*, respectively. The primers used for the site-directed

mutagenesis were synthesized by Canadian Life Technologies Inc., (Burlington, Canada) and are as follows:

for *mip1*Δ351: 5'GCGAAAAGGACAAATAGATCTCAGAGCTGCTATGG;

for *mip1*Δ279: 5'ATCAATCAACTGCTATAGATCTGACAAATCCAATAG and

for *mip1*Δ216: 5'GGCTAGAAGATGAGTAGATCTGCGGGAGTGTAC.

The sequences of the regions to be expressed as part of the mutated polymerases were confirmed by DNA sequencing. After mutagenesis, the *Bgl*III site was used for the insertion of a *Bam*HI fragment containing the *HIS3* gene. The mutated sequences were released from pBluescript SK⁻ by double digestion with *Xba*I and *Pvu*II, followed by gel purification using the QIAquick® Gel Extraction Kit (QIAGEN Inc, Mississauga, ON). Mutated fragments were then transformed into S150 yeast cells using the lithium acetate method (Gietz *et al.*, 1997). Insertion of the truncation cassette at the correct position was first confirmed by colony PCR according to Sambrook and Russell, 2001) except that the colony was first resuspended in water, and then diluted into a standard PCR reaction mixture containing 3.1 mM MgSO₄. The primers used correspond to base pairs 2585-2605 of the *MIP1* open reading frame for the upstream primer and 11-31 base pairs downstream of the wild-type stop codon for the downstream primer. Southern blot analysis using a *MIP1* specific probe was carried out to confirm correct isolates identified by colony PCR.

To create the coding sequence for *mip1*Δ175, a *HIS3* marker was ligated into the *Bam*HI site (position 3232 bp of the *MIP1* gene) in *MIP1*-F3 cloned in pAS1 (Gietz *et al.*, 1997). The resulting disruption construct was transformed into S150 yeast and confirmed as described above. The *mip1*Δ175 construct encodes an additional ten C-

terminal amino acids (AARSCSLACT) due to the extension of the reading frame into the *HIS3* cassette.

The *mip1* Δ 175 and *mip1* Δ 216 yeast mutations were reconstructed and additional mutations, *mip1* Δ 241, *mip1* Δ 222, and *mip1* Δ 205, were introduced into yeast using a one step gene replacement strategy. D175_209F, D216_209F, DA1_209F, DA2_209F, and D205_209F primers (Table 3.1) were engineered to amplify a *caURA3* gene from the plasmid pKT209 (EUROSCARF, Sheff and Thorn, 2004) when each was used separately in a PCR reaction with the DA11_209R primer. In addition, the D175_209F, D216_209F, DA1_209F, DA2_209F, and D205_209F primers contain stop codons to truncate the ORF encoding Mip1p at 175, 216, 241, 222 or 205 C-terminal residues respectively. The 1566-bp PCR products were transformed into S150 according to (Gietz and Woods, 2002) and transformants were selected on SC-URA and SGlycerol-URA. MIPDSofE/CaURA3R and DeltaScrMIPRev/CaURA3F primer sets were used to confirm upstream and downstream recombination junctions respectively via colony PCR of potential transformants as described.

GFP (green fluorescent protein) – Mip1p fusions were created as described by (Sheff and Thorn, 2004). The template for PCR was the pKT128 vector containing the yEGFP open reading frame and the *SpHIS5* marker, obtained from EUROSCARF. The following *MIP1*-specific sequences were included in the primers; [F5] and [R3] refer to the template specific segments of the primers (Sheff and Thorn, 2004).

Mip1p Δ 351 5'CCATGATGAGATTAGATTTTTGGTGAGCGAAAAGGACAAA[F5]

Mip1p Δ 279 5'CATTCCCTCATGGGGAGGCGCTTGATATCAATCAACTGCTA[F5]

Mip1p Δ 216 5'CCAGTCAGATAAGCGCGATGTGAATCGGCTAGAAGATGAG[F5]

Table 3.1 Primers used in this study				
Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
D216_209F	TCTTGCGATGCAAGTCCAGTCAG ATAAGCGCGATGTGAATCGGCTA GAAGATGAGTAGGTTTAGCTTGCC TCGTCC	56	pKT209	Bold = sequence for recombination at the <i>MIP1</i> locus; <i>italics</i> = engineered stop codon to create <i>MIP1D216</i> ; normal = primer for amplification of <i>caURA3</i> gene from pKT209
DA1_209F	ATATGCCTATAACTACAGAGAAC CTGTATTTGAAGAATATAATAAAT CTTATACTTAGGTTTAGCTTGCCTC GTCC	56	pKT209	Bold = sequence for recombination at the <i>MIP1</i> locus; <i>italics</i> = engineered stop codon to create <i>MIP1D241</i> ; normal = primer for amplification of <i>caURA3</i> gene from pKT209
DA2_209F	AGAGTTCTTAAAATATTTTCTTGC GATGCAAGTCCAGTCAGATAAGC GCGATGTGTAGGTTTAGCTTGCCTC GTCC	56	pKT209	Bold = sequence for recombination at the <i>MIP1</i> locus; <i>italics</i> = engineered stop codon to create <i>MIP1D222</i> ; normal = primer for amplification of <i>caURA3</i> gene from pKT209
D175_209F	CTATATAAAGGATGTCGAGAAGG GCAAAGGACTAAAGTACGTATT ATGGGATCCTAGGTTTAGCTTGCCT CGTCC	56	pKT209	Bold = sequence for recombination at the <i>MIP1</i> locus; <i>italics</i> = engineered stop codon to create <i>MIP1D175</i> ; normal = primer for amplification of <i>caURA3</i> gene from pKT209
D205_209F	TGTGAATCGGCTAGAAGATGAGT ATCTGCGGGAGTGTACATCCAAA GAATACGCTTAGGTTTAGCTTGCCT CGTCC	56	pKT209	Bold = sequence for recombination at the <i>MIP1</i> locus; <i>italics</i> = engineered stop codon to create <i>MIP1D205</i> ; normal = primer for amplification of <i>caURA3</i> gene from pKT209

Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
DAII_209R	TATTATAATGTGCTGTATATATAA ATACAAATGCGAAAGCTAATGCA GATTTTGCTCGATGAATTCGAGCT CG	54	pKT209	Bold = sequence for recombination at the <i>MIP1</i> locus; normal = R3 primer (Sheff and Thorn, 2004) for amplification of <i>caURA3</i> gene from pKT209
CaURA3R	TAATCCAACCTCCAGGTGTC	56	<i>caURA3</i>	use with MIPDSofE to screen upstream recombination junctions at <i>MIP1</i> locus
CaURA3F	CATCAATTGACGTTGATACC	56	<i>caURA3</i>	use with DeltaScrMIPRev to screen downstream recombination junctions at <i>MIP1</i> locus
DeltaScrMIP Rev	AGCTAACGAGAAGATGCAG	56	<i>MIP1</i>	Use with CaURA3F to screen downstream recombination junctions at <i>MIP1</i> locus
MIPDSofE	GAGTGGATTATCTACATCTTC	58	<i>MIP1</i>	Use with CaURA3R to screen upstream recombination junctions at <i>MIP1</i> locus

Mip1p Δ 175 5'CGAGAAGGGCAAAAGGACTAAAGTACGTATTATGGGATCC[F5]

Mip1p 5'GGTTGAGCTGGAAAGGGACATTACTATTTCTAGAGAGTAC[F5]

Mip1p Reverse (used for all constructions)

5'TAATGTGCTGTATATATAAATACAAATGCGAAAGCTAATG[R3].

The resulting PCR products were transformed into wild-type S150 cells according to (Gietz *et al.*, 1997). Colony PCR was used to confirm the upstream and downstream recombination junctions and the absence of the wild-type *MIP1* allele. Maintenance of respiratory function was indistinguishable in the wild-type cells and those expressing Mip1p-GFP, as observed by others (Meeusen and Nunnari, 2003). To visualize mitochondria, the resulting cells were transformed with VT100, which encodes a mitochondrially-targeted dsRED fluorescent protein (Meeusen and Nunnari, 2003).

3.3.3 Fluorescence microscopy of cells harbouring Mip1p-GFP fusions

A Zeiss Axio Imager Z1 was used for fluorescence microscopy and images were collected with an AxioCam Mrm camera. For each field, z-stacks of four 0.2 μ m slices were obtained. The fluorescence of the Mip1p-GFP variants was normalized to the emission from the mitochondrial dsRED, which is imported into mitochondria in the absence of mtDNA (Wong *et al.*, 2000). To avoid bias toward mitochondria with GFP fluorescence, 15-20 mitochondrial puncta, in at least three different cells, were chosen based on dsRED fluorescence only. Fluorescence was then measured in the red and green channels separately, corrected for exposure time, and summed for the entire stack. These values were used to calculate a ratio of Mip1-GFP (green) to mitochondrial (dsRED) fluorescence. The ratio for the wild-type cells was set to 1.

3.3.4 Assessment of mtDNA maintenance and replication fidelity

Respiratory competence was monitored over 96 hours. Yeast were first grown in YPL broth for 48 hours to ensure a 100% respiring culture. Cultures were then diluted to 0.01 OD₆₀₀ in YP10D broth for growth with shaking at 30°C. Subsequently, cultures were diluted to 0.01 OD₆₀₀ every 24 hours and glucose levels were continually monitored (Young and Court, 2004) to ensure sufficient (> 1%) glucose to avoid selection of respiratory-competent cells.

Slides for fluorescence microscopy of 4',6-diamidino-2-phenylindole (DAPI) stained cultures were prepared to monitor the presence of mtDNA during growth. Cultures were diluted to 0.7 OD₆₀₀ in 5 ml of YP10D broth and DAPI was added to a final concentration of 2.5 µg/ml. Cultures were incubated for 40 minutes with shaking at 30°C followed by washing with ddH₂O and resuspension in a 1/5 dilution of YP10D. The DAPI-stained culture was combined with an equal volume of 0.7% low gelling agarose (FMC, sold by Mandel, Guelph, ON) made in ddH₂O, followed by visualization using Zeiss Epi-fluorescent microscopy. At the indicated time points, cells were removed, diluted appropriately and spread onto YP2D plates to determine the number of viable cells. The resulting colonies were overlaid with 2,3,5-triphenyl tetrazolium chloride agar according to (Ogur *et al.*, 1957). Respiratory competence was calculated as the number of red, respiring colonies divided by the total number of colonies multiplied by 100.

Southern blot analysis was used to estimate relative content of intact mtDNA. Cells were harvested by centrifugation from the same cultures used for DAPI-staining and frozen at -60°C until used for genomic DNA preparation (Philippsen *et al.*, 1991). A 980-bp region of the mitochondrially-encoded *COX2* locus (bp 52 of the *COX2* ORF to

275 bp past the stop codon) was cloned into pBluescript SK⁻, and used as a template for generation of a digoxigenin-labelled DNA probe (Roche, Laval, PQ). A 1.5-kb *Hind*III genomic fragment harbouring the nuclear *RIM1* gene (a generous gift of F. Foury, Université Catholique de Louvain, Belgium), was gel purified and used to generate a nuclear DNA probe. Hybridization was carried out according to the DIG DNA Labeling Kit manual (Roche, Laval, PQ).

Erythromycin-resistance was monitored according to (Hu *et al.*, 1995) with the following modifications. 3.3 – 13.3 OD₆₀₀ units (~1 – 4 x10⁸ cells) were spread onto YG-Er plates from an overnight culture grown in 20 ml of a 50:50 mixture of YPG:YP2D. Appropriate dilutions were also spread onto YPG plates to determine the actual number of respiratory-competent cells on the YG-Er plates. This step was necessary due to the generation of a high number of spontaneous *petites* by *mip1Δ216* strains.

3.3.5 Phylogenetic Analyses

A total of 25 sequences (Table 3.2) were aligned with ClustalX (Thompson *et al.*, 1997) and alignments were manually edited with GeneDoc v 2.5.010 (Nicholas and Nicholas, 1997). The amino-terminal sequences (*S. cerevisiae* residues 1-37), containing the putative mitochondrial targeting signals, and most of the CTEs (residues 1007 to 1254) aligned poorly and were removed. In addition, short gaps were introduced by ClustalX to several sequences to maintain the overall alignment; the resulting non-informative positions were also removed. These positions correspond to residues 61 – 67, 204, 388, 476, 531 – 537, 548, 551, 596 – 638, and 987– 988 of the *S. cerevisiae* mtDNA polymerase. Parsimony, maximum likelihood (ML), and neighbor-joining (NJ) phylogenetic analyses of the mtDNA polymerase sequences were carried out using the

Table 3.2 Sources of mitochondrial DNA polymerase primary sequences used for phylogenetic analyses.

Organism	Accession Number^a
Metazoans	
<i>Drosophila melanogaster</i>	<u>AAF53338</u>
<i>Homo sapiens</i>	<u>AAH50559</u>
Basidiomycetes	
<i>Cryptococcus neoformans</i> var <i>grubii</i>	<u>AACO01000141</u> translation of nucleotides 185165 – 183051, 183048 – 182818 and 182815 – 181124 ^b
<i>Phanerochaete chrysosporium</i>	proteinId 23153 <u>Phanerochaete chrysosporium:gw.121.2.1</u>
Ascomycetes	
<i>Aspergillus fumigatus</i>	contig41 at <u>http://www.genedb.org/genedb/asp/665938 – 662894</u>
<i>Aspergillus nidulans</i>	<u>EAA65359</u>
<i>Candida albicans</i>	<u>AACO01000065</u> translation of nucleotides 77331 – 81017
<i>Candida glabrata</i>	<u>CAG58839</u>
<i>Debaryomyces hansenii</i>	<u>CAG89264</u>
<i>Eremothecium gossypii</i>	<u>AAS52277</u>
<i>Gibberella zeae</i>	<u>EAA74646</u>
<i>Histoplasma capsulatum</i>	HISTO_LF.Contig359 at <u>www.genome.wustl.edu/blast/histo_client.cgi</u> translation of nucleotides 555043-558543
<i>Kluyveromyces lactis</i>	<u>CAG97904</u>
<i>Kluyveromyces waltii</i>	<u>AADM01000080</u> translation of nucleotides 54465 – 58145
<i>Magnaporthe grisea</i>	<u>EAA46570</u>

Table 3.2 continued

<i>Neurospora crassa</i>	<u>AAD21034.1</u>
<i>Pichia pastoris</i>	<u>AAB17118.1</u>
<i>Podospora anserina</i>	contig_824 at <u>http://podospora.igmors.u-psud.fr</u> ; translation of nucleotides 3994 – 7461
<i>Saccharomyces bayanus</i>	<u>AACA01000477</u> (Partial ^c) translation of nucleotides 3 – 3749
<i>Saccharomyces castellii</i>	<u>AACF01000124</u> (Partial) translation of nucleotides 2 – 3316
<i>Saccharomyces cerevisiae</i>	<u>CAA99652.1</u>
<i>Saccharomyces kluyveri</i>	<u>AACE01000598</u> (Partial) translation of nucleotides 1380 – 2267
<i>Saccharomyces kudriavzevii</i>	<u>AACI01000043</u> (Partial ^d) translation of nucleotides 2811 – 34
<i>Saccharomyces paradoxus</i>	<u>AABY01000003</u> translation of nucleotides 136136 – 132363
<i>Schizosaccharomyces pombe</i>	<u>CAA88012</u>
<i>Yarrowia lipolytica</i>	<u>CAG78428</u>

^a accession numbers refer to predicted amino-acid sequences, with the exception of those that were generated by conceptual translation of the indicated nucleic acid sequence and are therefore indicated as “translation of nucleotides”. The positions given include the putative stop codon.

^b the indicated contig, when translated, contained three segments detected by TBLASTN as being similar to known mitochondrial DNA polymerases. Each of the three segments was in a different reading frame; hence a putative mtDNA polymerase reading frame was estimated by introducing two 2-bp deletions as indicated.

^c partial indicates that a reading frame encoding only part of the expected sequence of a mtDNA polymerase was identified. Partial sequences were only used in the analysis in Fig. 2A if the entire putative 3’-5’ exonuclease and polymerase domains could be identified

^d the *S. kluyveri* sequence data was used only for the alignment in Fig. 2B as only the 3’ sequence including the polymerase domain and the CTE were available

appropriate programs contained within PHYLIP (Felsenstein, 2002). PROTDIST was used in combination with NEIGHBOR to generate the tree presented in Figure 3.2. The data were then processed through the bootstrap procedure (Felsenstein, 1985, SEQBOOT, 1000 replicates). The resulting bootstrapped replicate-based phylogenetic estimates were evaluated using PROTDIST, NEIGHBOR and finally CONSENSE to obtain the level of support for the nodes in the NJ tree. Likewise, the bootstrapped dataset was analyzed with PROTPARS and PROML to obtain support for parsimony and ML phylogenetic trees.

The PRALINE online server (<http://ibivu.cs.vu.nl/programs/pralinewww/>; Simossis and Heringa, 2003) was employed for the multiple sequence alignment and secondary structural predictions outlined in Figure 3.2.

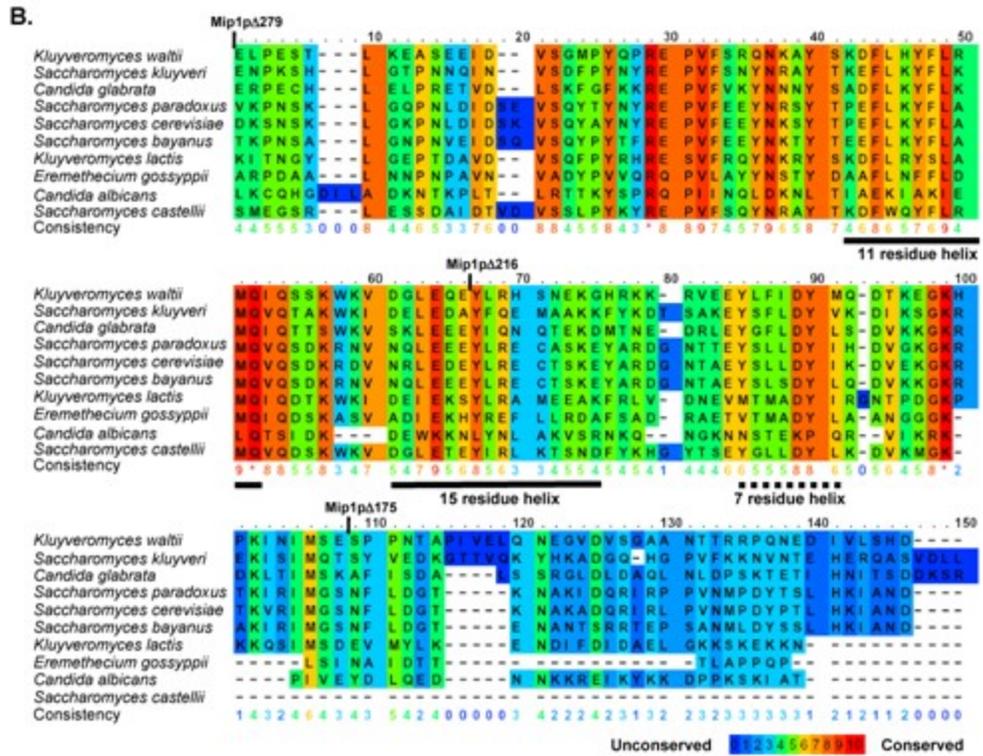
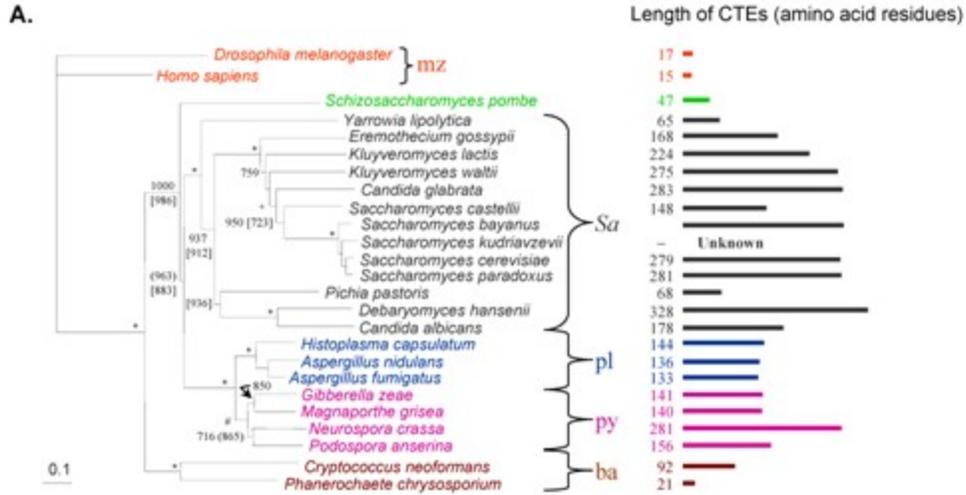
3.4 Results and Discussion

3.4.1 Phylogenetic analysis of fungal mitochondrial DNA polymerases

Carboxyl-terminal extensions were first noted on the mtDNA polymerases of *Saccharomyces* (Hu *et al.*, 1995) and *Neurospora* (Ko, T. and Bertrand, H., personal communication, see Table 3.2). With the wealth of new genome sequence data, it was possible to both investigate the phylogenetic relationships among mtDNA polymerases and determine the presence of CTEs among the taxonomic groups. The focus of this study was the Ascomycota, including members of the pyrenomycetes, the plectomycetes, and yeast-like representatives. Mitochondrial DNA polymerase sequences from the club fungi (the Basidiomycota) and the metazoans (*Drosophila melanogaster* and *Homo sapiens*) were used as outgroups. The latter two enzymes do not possess C-terminal extensions (Figure 3.2A); therefore, the core regions of the polymerase sequence,

Figure 3.2. Phylogenetic analysis of mtDNA polymerases. **(A)** Phylogenetic analysis of the “core” polymerase region including the exonuclease and polymerase domains. Analysis was carried out as described in Materials and Methods. Taxa are colour-coded: metazoans (mz, red), *Schizosaccharomyces pombe* (green), *Saccharomycetales* (*Sa*, grey), plectomycete fungi (pl, blue), pyrenomycete fungi (py, purple) and basidiomycetes (b, brown). Nodes that received support from bootstrap analysis, when combined with NJ, ML and Parsimony analyses above 95% are indicated by an asterisk (*), above 85% by a number sign (#) and above 75% by a plus sign (+). For lower levels of support, NJ values are indicated; ML and parsimony values are presented in square brackets and parentheses, respectively. A schematic map to the right of the phylogenetic tree shows the lengths of the CTEs for the various polymerases. The CTE has been designated as the sequence 26 residues past the end of the predicted polymerase domain (Blanco *et al.*, 1991); this encompasses residues 975 to 1254 of the *S. cerevisiae* preprotein. The contig containing the *S. kudruavzevii* *MIP1* gene contains the entire sequence used for the alignment, but ends before the coding sequence for the CTE; hence the length of the CTE is listed as unknown. **(B)** Comparative analysis of predicted CTE sequences. An alignment of the CTEs was generated using PRALINE (Simossis and Heringa, 2003), and conservation scores and predicted secondary structures for various yeast CTEs are presented. Consistency refers to how consistently an amino acid is found at a given position in the final multiple alignment and in pre-processed alignment blocks (Heringa, 1999). In the ten-point consistency scale, zero (purple) indicates the least conserved alignment position, while * (red) indicates the most conserved positions. Only the first 150 positions of the 328-residue alignment are shown, because the overall similarity

among the C-terminal segments is very low. The solid and dashed lines indicate long and short structural features that are predicted in at least 8 of 10 sequences. The positions of the various truncations are indicated at the top of the alignment.



encompassing the 3'-5' exonuclease and polymerase domains were used for phylogenetic analyses (see Materials and Methods).

ClustalX alignment of unambiguous positions, followed by parsimony, neighbour-joining and most-likelihood analyses, revealed evolutionary relationships among the mtDNA polymerases (Figure 3.2A) that mirrored those shown by similar analyses of both large (Dujon *et al.*, 2004) and small (Cavalier-Smith, 2001) rDNA sequences. Notably, the metazoan enzymes are most distant from those of the fungi and those of the Basidiomycetes group separately from the Ascomycetes.

The CTEs of the various polymerases were defined as all residues to the C-terminal side of the core enzyme (residues 975 to 1254 of the *S. cerevisiae* Mip1p preprotein). The N-terminus of CTE therefore is 26 amino acids from the end of the predicted polymerase domain (Blanco *et al.*, 1991) that terminates at residue 949 and includes the pol C motif (Figure 3.1). The CTE also extends 16 residues past the γ 6 region described by Kaguni (2004). The lengths of resulting CTEs are plotted alongside the phylogenetic tree (Figure 3.2A). Remarkably, the lengths of the CTEs do not correlate strictly with the phylogenetic positions of the organisms. For example, the CTE on the mtDNA polymerase of *Neurospora* is almost twice as long as that of the other pyrenomycete fungi. Similarly, within the tight grouping of *Pichia*, *Debaromyces*, and *Candida*, the lengths of the CTEs vary from 68 to 328 residues. Overall, several size classes are apparent: “long” CTEs of about 280 residues found in *Neurospora* and some yeast-like fungi; “medium length” CTEs of about 150 residues in filamentous fungi, *C. albicans*, and *Eremothecium*; and “short” CTEs of less than 93 residues in the ascomycetes *Yarrowia*, *Schizosaccharomyces* and *Pichia*, as well as the basidiomycete

Cryptococcus. CTEs in the polymerases from human, *Drosophila* and the white-rot fungus *Phanerochaete* are less than 22 residues in length.

BLAST searches of fungal CTEs do not reveal significant similarities with any known proteins (data not shown). Therefore, in an attempt to identify important features of CTEs, a multiple alignment of all CTEs was performed. No obvious conservation was observed across all polymerases. However, when CTEs from ten representatives of the *Saccharomycetales* were aligned with PRALINE (Simossis and Heringa, 2003), a gradient of sequence relatedness was observed, with most similarity near the polymerase domain (residues 10 – 69 of the alignment in Figure 3.2B). Some conservation is observed between residues 70 and 100, and the remaining sequences are unrelated (residues 101-150 in Figure 3.2B and data not shown). Only three residues are absolutely conserved, namely arginine at position 29 of the alignment (R1001 of the Mip1p preprotein), glutamine at position 52 (Q1024), and lysine at position 99 (K1070). Once the function of the CTE is known in more detail, the precise role(s) of these residues can be addressed through the analysis of single-residue variants.

3.4.2 Truncation variants of *Saccharomyces* mtDNA polymerase

As an initial approach to determine the importance of the CTE of *S. cerevisiae* Mip1p, a series of truncation variants was generated (Figure 3.1). The largest deletion, in Mip1p Δ 351, leaves only five amino acids to the C-terminal side of the conserved HDEIRFLV motif in the Pol C domain (Kaguni, 2004). The segment removed includes 71 well-conserved positions, including the mtDNA polymerase-specific region, γ 6, and therefore is likely to disrupt part of the polymerase domain. The truncation removes all of the CTE in Mip1p Δ 279, and its terminus corresponds to the last residues shared between

the fungal and metazoan mtDNA polymerases. Mip1p Δ 216 retains most of the highly conserved region in fungal CTEs, while Mip1p Δ 175 lacks only the highly variable C-terminal 63% of the CTE (see Figure 3.2B). Mip1p Δ 205 contains all of the 15-residue helix while it is absent in Mip1p Δ 222. Finally, Mip1p Δ 241 lacks all the predicted helices in the conserved part of the CTE (Figure 4.1A). The genes encoding the truncation variants were transformed into haploid S150 yeast to replace the wild-type *MIP1*.

3.4.3 Respiratory competence of Mip1p variants

Following selection for insertion of the *mip1* truncation mutations on SC-HIS plates, single colonies were tested for respiration. *mip1 Δ 351* cells rapidly lost respiratory competence; all of the 14 colonies taken directly from the transformation plates from two separate experiments did not grow on glycerol-containing medium (Figure 3.1). Furthermore, mtDNA was not detected in these cells by DAPI-staining (data not shown). These results confirm that the entire polymerase domain, including γ 6, is required for mtDNA polymerase function. The initial cells used for the transformation could respire, as indicated by the respiratory ability of transformants in which the *HIS3* marker in the *mip1* truncation plasmids had successfully recombined into the *his3 Δ 1* locus (data not shown). In general, these *HIS*⁺ cells produced large colonies on the transformation plates, while recombination into the *MIP1* locus gave rise to small colonies. Similarly, all of *mip1 Δ 279*, *mip1 Δ 241*, and *mip1 Δ 222* colonies obtained from the transformations were respiratory incompetent (Figure 3.1 and data not shown) and mtDNA staining was not detected for *mip1 Δ 279* (data not shown). The Mip1p Δ 279 variant retains all of the residues conserved between fungal and metazoan polymerases, indicating that all or part of the CTE is required for mtDNA maintenance in yeast.

To confirm that Mip1p Δ 279 and Mip1p Δ 351 variant polymerases were imported into and maintained in mitochondria, fluorescence microscopy of cells harbouring fusions of GFP to the truncated C-terminal ends of these polymerases was performed. These transformants lacked the ability to respire, as did their *mip1 Δ 279* and *mip1 Δ 351* counterparts. In both Mip1p Δ 279-GFP and Mip1p Δ 351-GFP strains, green fluorescence co-localized with red fluorescence from mitochondrially-targeted dsRED (Figure 3.1B). Assuming that any reduction in mitochondrial import in *petite* cells would affect the import of both proteins equally, the ratios of GFP and dsRED fluorescence were used to estimate the mitochondrial levels of the Mip1p variants relative to that of the wild-type protein (Figure 3.1A). The average levels of both truncation variants were lower than that of wild-type protein, but given the standard deviations in each dataset, these differences are not statistically significant. Thus, the defects in *mip1 Δ 279* and *mip1 Δ 351* cells are not likely the result of severely reduced levels of mtDNA polymerase; a stabilizing effect of the GFP moiety cannot be ruled out.

mip1 Δ 216 transformants were of two types: those unable to respire and those that showed weak growth on glycerol-containing plates; the polymerase levels were again similar to those in wild-type cells (Figure 3.1A). Therefore, the inclusion of residues 975 to 1038 only partially compensates for the lack of function in Mip1p Δ 279, and the C-terminal GFP fusion does not affect the phenotype. To further analyze the poorly respiring *mip1 Δ 216* cultures, respiratory competence was assessed after growth in glucose, which is non-selective for mtDNA function. Under these conditions, *mip1 Δ 216* cells rapidly lost respiratory competence; after 24 hours growth, approximately 2% of the

colonies could respire. In contrast, about 99% of the wild-type cells were respiratory competent (data not shown).

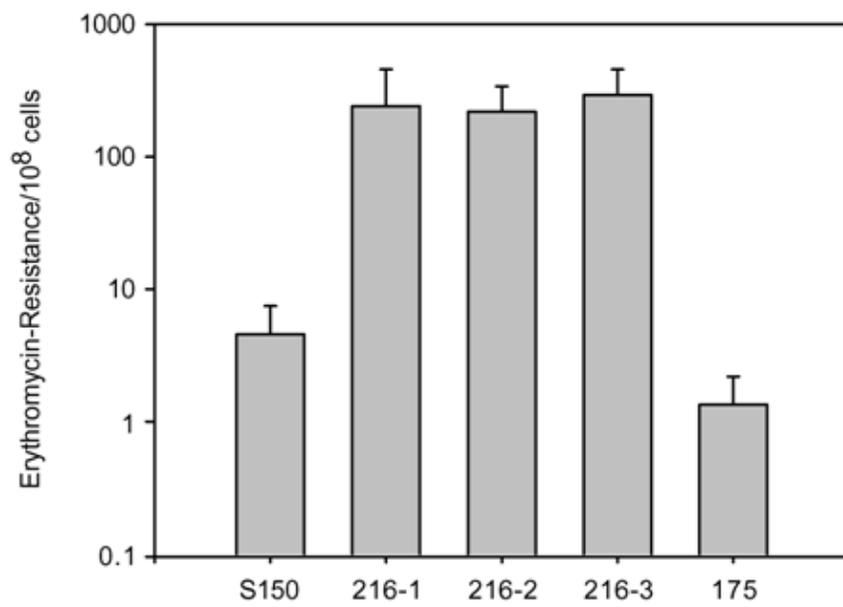
mip1Δ175 and *mip1Δ205* transformants produced large, respiratory-competent colonies. After four days growth at 30°C in glucose, almost all cells retained respiratory competence (98.2% and 95.9% respectively). For *mip1Δ175* cells mtDNA staining (100% of 77 cells), and levels of Mip1pΔ175-GFP were similar to those of Mip1p-GFP in wild-type cells (Figure 3.1A). Therefore, replacement of the terminal part of the CTE with another protein does not alter the wild-type phenotype. As *mip1Δ175* and *mip1Δ205* cells were indistinguishable from wild-type, the entire CTE is not required for mtDNA maintenance, and further truncations in the C-terminal region of the CTE were not made.

3.4.4 Mutator phenotype of *mip1Δ216* cells

The *mip1Δ216* mutant was ideal for further analysis because it was capable of long-term respiratory growth on media containing non-fermentable carbon sources, indicating that it maintained functional mitochondrial genomes. Therefore, it was possible to assess the defects associated with this truncated polymerase. There are several possible reasons for the loss of respiratory competence in *mip1Δ216* cells: complete loss of mtDNA, major deletions and rearrangements of the mitochondrial genome, or multiple point mutations.

Because *mip1Δ216* cells can maintain mtDNA under selective pressure, fidelity of mtDNA replication can be investigated using the erythromycin-resistance assay (Hu *et al.*, 1995). Cultures of wild-type S150 cells produced between 2 and 8 erythromycin-resistant colonies per 10⁸ cells (Figure 3.3), in the same order of magnitude as observed by Hu *et al.* (1995) for W303-1B. Resistant *mip1Δ175* cells appeared at the same

Figure 3.3. Point mutations in mtDNA. The relative fidelity of mtDNA replication in wild-type (S150) and *mip1*Δ175 (175) cells was assessed as the frequency of erythromycin-resistant colony formation per 10⁸ cells, as described by Hu *et al.*, 1995. Each strain was analyzed in triplicate and averages and standard deviations are presented. Data obtained for three independent isolates of *mip*Δ216 (216-1, 216-2, 216-3) are presented.

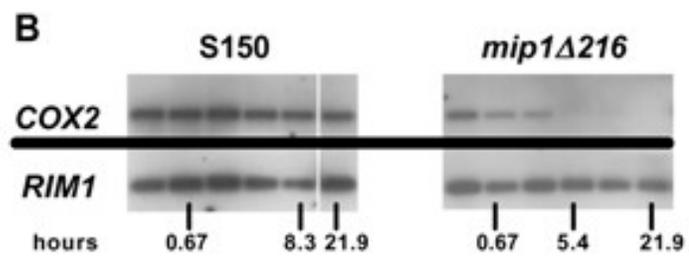
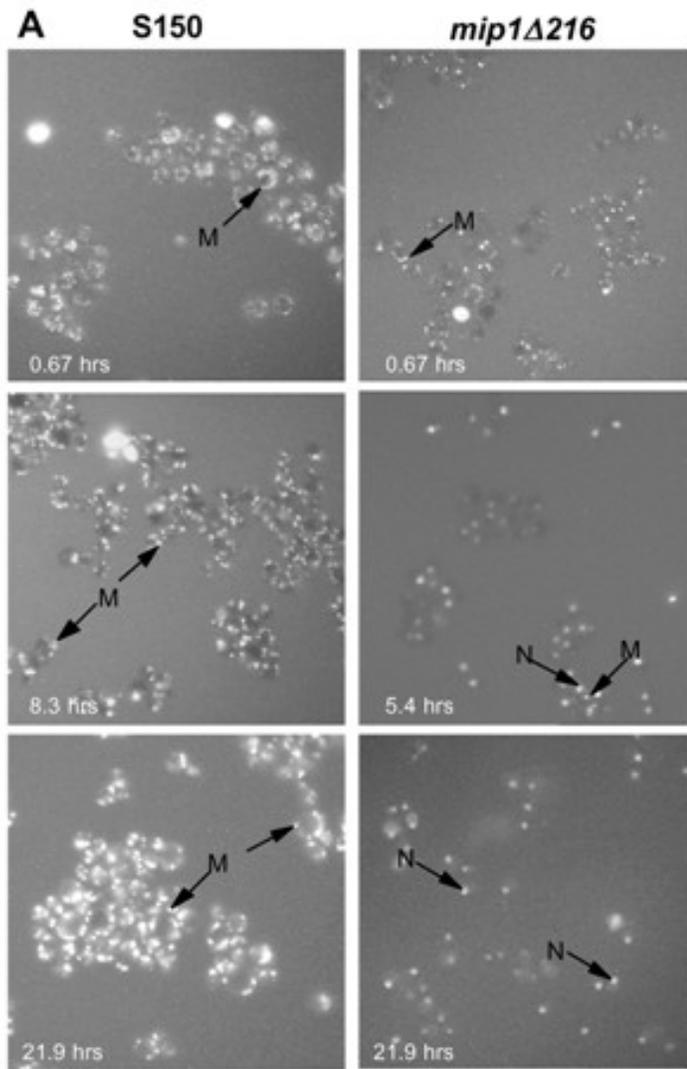


frequency, indicating that the 175 C-terminal residues are not involved in replication fidelity. Erythromycin-resistance of the *mip1Δ205* mutant was on average in the same order of magnitude as S150 and *mip1Δ175*. When performed with three separate isolates, the erythromycin-resistant colonies per 10^8 was 2.9, 6.9, and 3.8 (average 4.5, standard deviation 2.1). However, each of three independent *mip1Δ216* transformants produced highly variable levels of resistant cells in replicate experiments; overall there were between 10 and 100 times more resistant cells than seen for the wild-type control (Figure 3.3). These observations suggest that the Mip1pΔ216 polymerase is more error-prone than the wild-type enzyme. The level of erythromycin-resistance acquired by *mip1Δ216* cells was similar to that observed for mtDNA polymerases with point mutations in the proofreading 3'→5' exonuclease domain (Hu *et al.*, 1995). For the latter mutants high variation between experiments was also observed. In the case of *mip1Δ216* cells, this variation likely reflects the reduced mtDNA copy number (see below). Assuming most mutations occur in the YPD culture of each isolate prior to plating on erythromycin, cells with a low number of mtDNA molecules would have an increased probability of transmitting an erythromycin-resistant genome to their progeny cells. Therefore, the number of resistant cells in the culture would be greatly influenced by when the mutation occurred – early or late. The variation in total mtDNA levels among the isolates (see below) would compound this effect.

3.4.5 Loss of mitochondrial DNA in *mip1Δ216* cells

Mitochondrial DNA maintenance by *mip1Δ216* cells was investigated during non-selective growth on glucose, using both DAPI staining of whole cells and Southern blot analysis (Figure 3.4). Cultures were first grown in lactate to ensure respiratory activity at

Figure 3.4. Time course analysis of mtDNA maintenance. Each experiment was performed with three independent *mip1* Δ 216 or S150 isolates and a representative experiment is shown. **(A)** DAPI staining. S150 and *mip1* Δ 216 cells were grown under non-selective conditions (YP10D). The presence of mtDNA was assessed by DAPI-staining of cultures at the indicated time points, which include the 40 min of the staining procedure. M, punctate mitochondrial staining; N, nuclear staining. Nuclear staining is relatively weak in cells with strong mitochondrial signals. **(B)** Southern blot analysis. Whole cell DNA was isolated from S150 (right) and *mip1* Δ 216 cells (left) immediately after transfer to YP10D (0 hrs) and following 0.67, 2.7, 5.4, 8.3, and 21.9 hrs growth in YP10D. DNA was then digested with *Hind*III and *Sac*II and analyzed by Southern blot using probes for mtDNA (*COX2*) and nuclear DNA (*RIM1*).



the start of the experiment. DAPI staining is most effective in glucose-grown cells (data not shown); therefore all cells were exposed to glucose for 40 min during staining. After staining, almost all wild-type and *mip1Δ216* cells taken directly from the YPL culture showed strong, punctate mtDNA staining in the periphery of the cell (Figure 3.4A). However, the *mip1Δ216* cells usually contained fewer nucleoids (one to three per focal plane) than did the wild-type (at least five). For the wild-type cells, the staining pattern remained virtually unchanged through almost 22 hours of growth in glucose (Figure 3.4A). However, the number of *mip1Δ216* cells with DAPI mtDNA staining was drastically reduced (to 29.3%) after only 5.4 hours of growth and none of the cells showed staining after 22 hours (Figs. 3.4A and 3.5A). This trend was observed in the three strains, each derived from an independent *mip1Δ216* transformant. In similar experiments, wild-type and *mip1Δ175* cells maintained high levels of mtDNA staining during 96 hours of growth in glucose (data not shown).

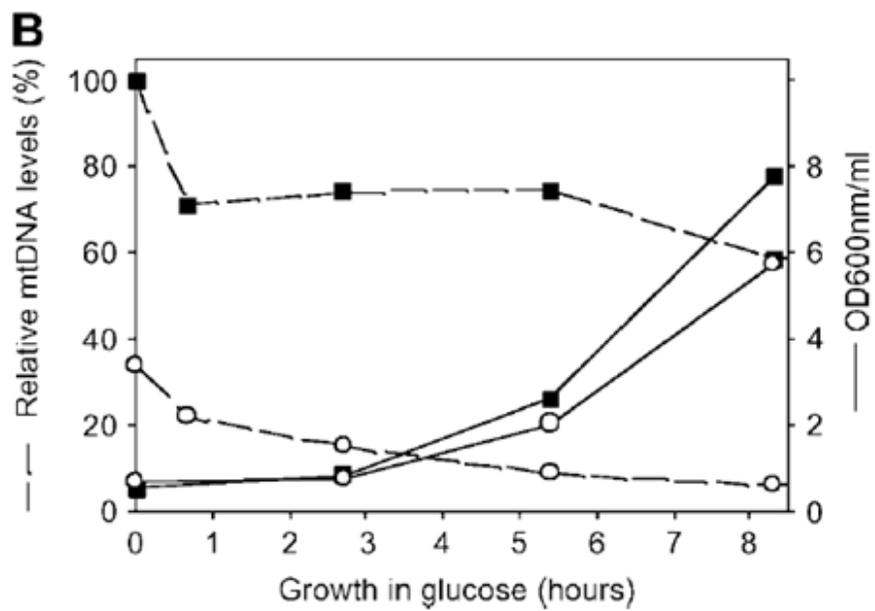
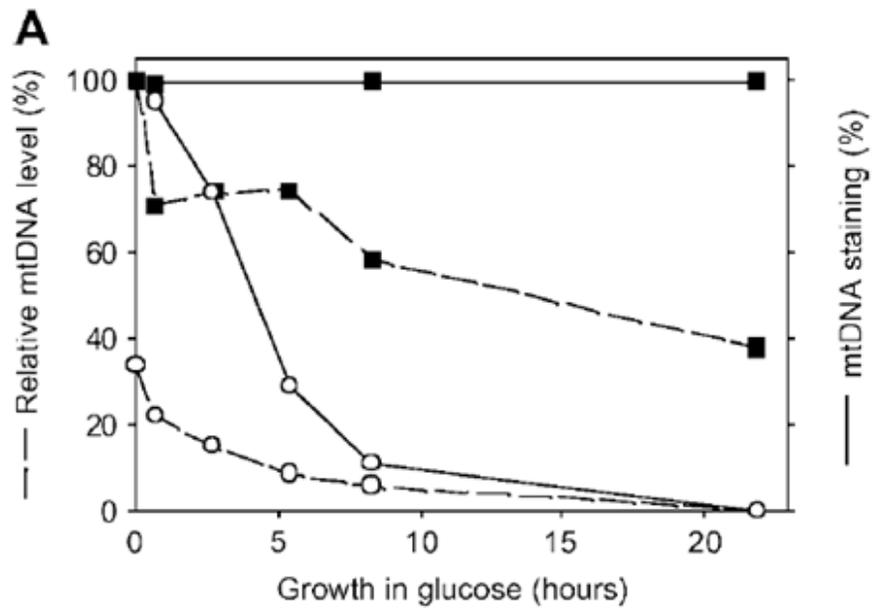
The loss of intact mtDNA was confirmed by Southern blot analysis (Figure 3.4B) of wild-type and *mip1Δ216* cells. Following initial growth in YPL, whole cell DNA from both strains contained a single *Sac*II fragment detected by a probe for the mitochondrial gene *COX2*. However, compared to the signal for the nuclear gene *RIM1*, the relative amount of intact mtDNA in the *mip1Δ216* cells was about one-third of that of the wild-type (Figs. 4B and 5A), in agreement with the DAPI-staining. During non-selective growth on glucose, the three independent *mip1Δ216* strains rapidly lost *COX2* hybridization (Figure 3.4B shows a representative). These results also indicate that the *mip1Δ216* cells had not acquired suppressor mutations during their initial growth in YPL.

Together, the DAPI-staining and Southern blot data revealed mtDNA loss that paralleled the growth of the *mip1Δ216* culture (Figure 3.5B). After 0.67 hours of growth in glucose, the cells were in early lag phase. At this point, mitochondrial staining was observed in about 95% of the *mip1Δ216* cells, but the *COX2* hybridization signal was reduced from 34% of wild-type levels to only 22% (Figs. 3.4B and 3.5B). Similarly, wild-type cells maintained 100% DAPI staining, and a similar reduction of about 30% in the relative level of mtDNA was observed. Thus, both cell types appear to undergo a reduction in relative mtDNA levels upon transfer to a fermentable carbon source. The cells were in stationary phase following growth in YPL, and upon transfer to fresh media, nuclear DNA replication during lag phase would result in a reduction in the ratio of mtDNA to nuclear DNA, if the rate of nuclear DNA replication exceeds that of mtDNA.

After 2.7 hours of growth in glucose, the cells were just beginning to emerge from lag phase; mtDNA staining was observed in only 74% of the *mip1Δ216* cells, in contrast to its presence in virtually all wild-type cells (Figure 3.5A). In comparison to the 0.67 hr timepoint, the relative *COX2* hybridization signal was reduced from 22% to 15% of wild-type levels in the mutant cells, and remained unchanged at about 70% of the initial level in the S150 cells (Figs. 3.4B and 3.5B). Smearing of the *COX2* hybridization band was not observed (Figure 3.4B and data not shown), suggesting that large scale DNA rearrangements do not accompany mtDNA loss, but the *COX2* signal is weak in the *mip1Δ216* DNA, and minor alternative forms may not have been detectable.

In the exponential growth phase, the loss of mtDNA continued in parallel to the loss of DAPI staining in *mip1Δ216* cells (Figure 3.5B), suggesting a defect in replication rather than in mtDNA distribution. In the latter case, one would expect the maintenance

Figure 3.5. Quantitative analysis of mtDNA loss and growth rate. (A) The percentage of cells with mtDNA DAPI staining (solid line) and the relative levels of mtDNA (dashed lines) are presented for the S150 and *mip1Δ216* isolates shown in Fig. 3.4. Comparative mtDNA levels were estimated from the ratio of the *COX2* to the *RIM1* signals (see panel B of Fig. 3.4). The ratio for the S150 cells grown overnight in YPL was set to 1 and used to standardize relative mtDNA levels for both S150 (filled squares) and *mip1Δ216* (open circles) cells during the time course. (B) mtDNA loss during cell division. The solid lines indicate the growth of the cultures used for Fig. 3.4, and dashed lines the relative mtDNA levels. Growth was assessed spectrophotometrically at 600 nm.



of a few cells with multiple nucleoids resembling those observed at the 0.67 hour timepoint. However, a reduction in the number of nucleoids per cell is obvious by 5.4 hours (Figure 3.4A). Throughout the remainder of the experiment, cells were maintained in log phase (see Materials and Methods) and a dilution effect was observed, in that the mtDNA levels in *mip1Δ216* cells decreased by about half with each cell division (Figure 3.5B). For example, in the representative experiment shown in Figure 3.4A, after 5.4 hours, the *mip1Δ216* cells had doubled, and the levels of mtDNA dropped from 15% (2.7 hrs) of wild-type levels to 8.7% (Figure 3.5B). Likewise, the fraction of cells with DAPI-stained mitochondria fell from 74% (2.7 hrs) to 29% (Figure 3.5A). These results suggest an almost complete failure of mtDNA replication in *mip1Δ216* cells. In contrast, throughout the experiment (21.9 hrs), multiple nucleoids were visible in all of the wild-type cells, and mtDNA was maintained, albeit at reduced levels (70-40%; Figs. 3.4B and 3.5A). Similarly reduced levels of mtDNA in glucose-grown *S. cerevisiae* cells has been described previously (Dujon, 1981).

3.4.6 mtDNA maintenance by *mip1Δ216* cells in non-fermentable carbon sources

The *mip1Δ216* cells have very different mtDNA maintenance profiles when grown in fermentable vs non-fermentable carbon sources: rapid mtDNA loss vs maintenance of low levels of mtDNA in all cells. This effect is not related to glucose repression, as similar rates of mtDNA loss are observed upon growth in the non-repressing, fermentable carbon source raffinose (data not shown). One explanation for this observation is derived from the growth rates of the wild-type and *mip1Δ216* mutant cells. In liquid culture, *mip1Δ216* cells grow more slowly when respiring (doubling time of 5.3 +/- 0.6 hours in YPG) than do the wild-type cells (3.6 +/- 0.1 hr). In glucose, both

cell types grow rapidly, with doubling times of 2.2 +/- 0.3 and 2.4 +/- 0.3 hours, respectively. The reduced respiratory growth rate of *mip1Δ216* cells reflects lower respiration and energy production due to the intrinsically lower levels of mtDNA. Interestingly, most *mip1Δ216* cells are *rho*⁺ during growth on glycerol (Figure 3.4A, 0.67 hr timepoint), suggesting efficient transmission of mtDNA to daughter cells. Perhaps the slower growth rate during respiration allows for sufficient replication by the compromised polymerase to ensure high enough mtDNA levels for transmission to daughter cells. In contrast, the rapid growth rate in glucose may be insufficient for replication and partitioning of limited amounts of mtDNA. Investigation of this question using *in organello* measurements of mtDNA replication rates is not practical for two reasons. First, the mutant cells appear to harbour near wild-type levels of Mip1pΔ216 (Figure 3.1B), but have lower levels of mtDNA (Figure 3.4 and 3.5); thus the ratios of mtDNA to polymerase are different than in wild-type cells. Second, the relative levels of mtDNA are variable in different YPL cultures derived from a single isolate. For example, three cultures derived from a single *mip1Δ216* isolate contained 34-75% of wild-type levels of mtDNA, as determined by Southern blotting (data not shown). As a result, the polymerase to template ratio is not constant in the mutant cells and valid comparisons with the wild-type can not be made. Therefore, future approaches to this question will involve biochemical characterization of over-expressed, purified Mip1p variants on standardized templates.

3.4.7 Possible functions of the CTE

The data presented herein demonstrate that only the γ 6-proximal 74 residues of the CTE (residues 975 to 1049) are sufficient for wild-type polymerase function. This

region, maintained in Mip1p Δ 205 but not in Mip1p Δ 279, is the most highly conserved CTE segment among mtDNA polymerases of *Saccharomycetales* (Figure 3.2B and Figure 4.1A), but it is not well-conserved among more distantly-related organisms. The sequence of this region does not contain any known motif or predicted secondary structure that reveals its potential function. Given the lack of predicted sequence similarity between the CTE and the accessory subunits of metazoan mtDNA polymerases, it is unlikely that the CTE functions in polymerase loading at primer-template junctions in the same manner as the metazoan proteins.

Examination of the fungal CTE sequences did reveal several salient features. All CTEs are very polar (47-71%) and their isoelectric points are 9 or higher, with exception of the CTEs from *Schizosaccharomyces* (4.4), *S. castellii* (5.0) and *Eremothecium* (4.5), suggesting that these extensions could be involved in non-specific mtDNA binding. An arginine residue equivalent to R1001 is also present in all ascomycete CTEs examined. Hu *et al.* (1995) determined that cells harbouring a point mutation converting this arginine to isoleucine produced *rho*⁻ cells at increased frequency (20% at 28°C). This polymerase also displayed a moderate mutator phenotype (50 erythromycin-resistant colonies/10⁸ cells). Mip1p Δ 216 includes this arginine, but produces a more severe phenotype, indicating that residues C-terminal to 1038 also contribute to mtDNA maintenance and fidelity of replication.

Secondary structure predictions for the conserved segments of the ascomycete CTEs (Figure 3.2B) reveal three potential α -helical regions, one of which contains the conserved glutamine residue, Q1024. This putative helix is very hydrophobic, while the 15-residue helix (Figure 3.2B) contains 13 polar residues and the 7-residue helix is

amphipathic. In Mip1p Δ 216 the putative 15-residue helix is interrupted and the smaller C-terminal one is completely removed. Mip1p Δ 205 truncates following the putative 15-residue helix. These cells are indistinguishable from wild-type, indicating the importance of this region.

In addition to DNA binding, there are several other possible functions of the CTE. First, it may be required for folding or stability of the mtDNA polymerase. It was difficult to use temperature sensitivity of the Mip1p Δ 216 polymerase to assess stability because mtDNA loss occurred so rapidly at 30°C, and was complete after 24 hours at either 30°C (Figure 3.4A) or 37°C (data not shown). In contrast, *mip1 Δ 175* and *mip1 Δ 205* cells grew and maintained respiratory competence indistinguishably from the wild-type at 37°C, indicating that the C-terminal segment of the CTE is not required for stability of the polymerase at 37°C. An argument against a possible role in folding or stability is that all of the Mip1p-GFP fusion variants co-localized with mitochondrial dsRED (Figure 3.1B), indicating that these proteins are stably maintained in mitochondria. However, stabilizing effects of the GFP moiety cannot be ruled out.

The CTE may be involved in protein-protein interactions with the mtDNA replication machinery, the nucleoid structure or the two-membrane spanning complex. Several yeast two-hybrid screens using Mip1p as bait (Uetz *et al.*, 2000; Ito *et al.*, 2000; Li and Court (unpublished results)), have not revealed any putative mitochondrial interaction partners. An interaction with Bur2p was detected, but it is a cytosolic RNA polymerase II binding protein, so the nature of the interaction is not clear (Hazbun *et al.*, 2003). A bacterial two-hybrid screen revealed a more relevant interaction, with Sed1p, which was subsequently shown to be required for maintenance of normal levels of Mip1p

and mitochondrial genome integrity (Phadnis and Ayres Sia, 2004). Mip1p has been shown to colocalize with the TMS protein Mmm1p by fluorescence microscopy (Meeusen and Nunnari, 2003), but a direct interaction between the two proteins was not demonstrated. Future studies in our lab will include suppressor screens to identify putative interactions between the CTE of Mip1p and other components of the mitochondrial DNA replication machinery.

CHAPTER 4 Deleterious effects of various Mip1p carboxyl-terminal truncations and the Mip1p alanine-661 variant on DNA polymerase activity

4.1 Abstract

In chapter 2 it was demonstrated that various genetic factors in commonly used laboratory strains make differential contributions to mitochondrial DNA (mtDNA) maintenance. Similar to Baruffini *et al.* (2007b) we observed temperature-sensitive mtDNA loss and error-prone mtDNA synthesis in strains related to S288c, when compared to strains related to Σ 1278b. S288c-related strains encode a variant form of Mip1p wherein a conserved threonine, present in Σ 1278b-related strains (*MIP1*[Σ]), at position 661, is replaced by an alanine (*MIP1*[S]). In chapter 3 it was demonstrated that truncations of the yeast mtDNA polymerase (Mip1p) carboxyl-terminal extension (CTE) are associated with mtDNA loss and in one case (*MIP1 Δ 216*) mtDNA instability *in vivo*. To investigate whether or not mtDNA instability and loss in these strains results directly from altered DNA polymerase activity, Mip1p variants were overexpressed and partially purified from yeast mitochondrial membrane fractions. A non-radioactive DNA polymerase assay was used to detect the activity of these enzymes. These results suggest that threonine 661 contributes to more robust DNA polymerase activity of Mip1p[Σ] in comparison to Mip1p[S] at 30 and 37°C. In addition, these results may also suggest that exonuclease function is not affected by the alanine-661 variant at 37°C whereas polymerase activity is, and this higher ratio of exonuclease activity could be a contributing factor to mtDNA instability in S288c-related strains. Lastly, isogenic CTE truncation variants all have less DNA polymerase activity compared to their parental wild-type polymerases. Based on these results several possible roles for the function of the CTE in mtDNA replication are suggested.

4.2 Introduction

Carboxyl-terminal extensions (CTEs) of fungal mitochondrial DNA (mtDNA) polymerases were first noted in *Saccharomyces* (Hu *et al.*, 1995) and *Neurospora* (Ko T, Bertrand H, personal communication); however, the function of the CTEs was unknown. We have previously shown that mtDNA polymerases from ascomycetous fungi, including yeast-like fungi, pyrenomycetes, and plectomycetes, contain variable length CTEs (Young *et al.*, 2006, Chapter 3). We showed that the 279-residue CTE of the mitochondrial DNA polymerase (Mip1p) from *Saccharomyces cerevisiae* is essential for respiratory growth and maintenance of mtDNA *in vivo*. An interesting mutant with 216 C-terminal residues deleted from Mip1p (encoded by *MIPIΔ216*, Figure 4.1 A) maintained mtDNA at lower levels when compared to the wild-type, during respiratory growth, but could not maintain mtDNA during fermentative growth. Therefore, this mutant defined a region essential for mtDNA maintenance between the truncations Mip1pΔ279 and Mip1pΔ216 (residues 975 to 1038 of Mip1p, Chapter 3, Young *et al.*, 2006 and Figure 4.1 A). Likewise, the *MIPIΔ216* mutant in combination with another truncation mutant, *MIPIΔ205*, encoding a polymerase lacking 205 C-terminal residues, also defined a region that contributes to mtDNA maintenance and fidelity, as *MIPIΔ205* cells maintained wild-type levels of respiratory competence and mtDNA point mutations (residues 1038 to 1049 of Mip1p, Chapter 3, Young *et al.*, 2006 and Figure 4.1 A).

A PRALINE alignment of CTEs from ten representatives of the Saccharomycetales predicted three conserved alpha-helical regions within the amino-

Figure 4.1 A. CLUSTALW (<http://align.genome.jp/>) alignment of Mip1p sequences (1254 amino acid residues) obtained from S150 and YPH499 laboratory strains. The black triangle at the amino-termini represents the MitoProtII predicted cleavage site, between residues 41 and 42, of the mitochondrial targeting sequence (probability of export to mitochondria 0.9927 and 0.9946 for S150 and YPH499 respectively, Claros and Vincens, 1996). Bold residues in the alignment represent polymorphisms identified by sequencing that have been previously defined by Baruffini *et al.* (2007b). Sequencing revealed that S150 has the *MIP1*[Σ] version of the gene while YPH499 has the *MIP1*[S]. The strictly conserved threonine 661 residue is indicated by a bold number. Exonuclease (Exo) and polymerase (Pol) motifs (white boxes) are as defined by Ito and Braithwaite (Ito and Braithwaite, 1990) and gamma specific sequences (γ 1 – γ 6, black line under both sequences) are as defined by Luo and Kaguni (2005, and personal communication from Laurie S. Kaguni). The conserved residues in γ 4, SYW, that when substituted with AAA in *Drosophila* polymerase gamma caused an increase in exonuclease activity, are highlighted by a black bar above this sequence (Luo and Kaguni, 2005). Within the carboxyl-terminal extension (CTE, carboxyl-terminal 279 amino acid residues) black bars between the two sequences are predicted alpha-helical regions defined by Young *et al.* (2006). Various truncations are represented as Mip1p and the number of carboxyl-terminal amino acid residues deleted from the CTE. The two unlabeled arrows above and below the sequences delimit the region encoded by the PCR product generated from S150 genomic DNA using the primers MIP4_F_PolyFix and RevPolyFix. This PCR fragment was used to replace the region between *AvrII* and *PstI* restriction endonuclease sites within the pMIPS28 plasmid; the equivalent region of these restriction sites on the Mip1p

sequence are shown (see text for details). **B.** Conceptual translation of the two possible Mip1p N-termini and that of the site-directed mutant Mip1p[LongM→I]-yEGFP. The five amino-terminal methionines of the 1280-residue conceptual translation of *MIP1* are highlighted in bold. The four isoleucines that replaced methionines in the site-directed mutant are highlighted in bold italics. pMIP1[Long] represents the N-terminus of the Mip1p variant encoded by pMIP1[Long]-yEGFP, pMIP1[Short] represents that of pMIP1[Short]-yEGFP, and pMIP1[LongM→I] represents Mip1p[LongM→I]-yEGFP.

A.

41-42
▼

S150_Mip1p MTKLMVRFECMLRMVRRRPLRVQFCARWFSTKKNAAEAPRINPVGIQYLGESLQRQVFGSCGGKDEVEQSDKLMELSKKSLKDHLGWGKKTLLITDPI SFPLPPLQGRSLDEHFQKIGRFN
YPH499_Mip1p MTKLMVRS⁸ECMLRMVRRRPLRVQFCARWFSTKKN³⁵TAEAPRINPVGIQYLGESLQRQVFGSCGGKDEVEQSDKLMELSKKSLKDHLGWGKKTLLITDPI SFPLPPLQGRSLDEHFQKIGRFN

Exo I Exo II

S150_Mip1p SEPYKSFCEDKFTEMVARPAEWLKPGWVKYVPGMAPVEVAYPDEELVVFDVETLYNVSDYPTLATALSSTAWYLWCSPFICGGDDPAALIPLNTLNKEQV²²²VIGHNVAYDRARVLEEYNF
YPH499_Mip1p SEPYKSFCEDKFTEMVARPAEWLKPGWVKYVPGMAPVEVAYPDEELVVFDVETLYNVSDYPTLATALSSTAWYLWCSPFICGGDDPAALIPLNTLNKEQV²²²IIGHNVAYDRARVLEEYNF

Exo III

S150_Mip1p RDSKAFFLDTQSLHIASFGLCSRQRPFMFNKNNKKEAEVESEVHPEISIEDYDDPWLNV³⁵⁷SLKDVAKFHCKIDLKTD³⁵⁷DRDFFASTDKSTIIENFQKLVNYCATDVTATSQVDFKIFP
YPH499_Mip1p RDSKAFFLDTQSLHIASFGLCSRQRPFMFNKNNKKEAEVESEVHPEISIEDYDDPWLNV³⁵⁷SLKDVAKFHCKIDLKTD³⁵⁷DRDFFASTDKSTIIENFQKLVNYCATDVTATSQVDFEIFP

γ 1 γ 2

S150_Mip1p VFLKKCPHPVSVFAGLKSLSKCILPTKLN⁵⁴¹DNDYLNSS⁵⁴¹ESLYQQSKVQIESKIVQIIKDIVLLKDKPDFY⁵⁴¹LKDPWLSQLDWTTKPLFLTKKGVPAK⁵⁴¹CQKLPGFPEWYRQLFPSKDTVEPKI
YPH499_Mip1p VFLKKCPHPVSVFAGLKSLSKCILPTKLN⁵⁴¹DNDYLNSS⁵⁴¹ESLYQQSKVQIESKIVQIIKDIVLLKDKPDFY⁵⁴¹LKDPWLSQLDWTTKPLFLTKKGVPAK⁵⁴¹CQKLPGFPEWYRQLFPSKDTVEPKI

γ 3 → AvrII γ 4

S150_Mip1p TIKSRIIPILFKLSWENSPVIWSKESGWC⁶⁶¹FNVPHEQVEIYKAKNYVLADSVSQEEEEIR⁶⁶¹MN⁶⁶¹NLGLQCTGVIFKVPHPNGPTFNCTNLLTKSYNHFFEKVGLKSESELAHQALQINSSGSY
YPH499_Mip1p TIKSRIIPILFKLSWENSPVIWSKESGWC⁶⁶¹FNVPHEQVEIYKAKNYVLADSVSQEEEEIR⁶⁶¹TH⁶⁶¹NLGLQCTGVIFKVPHPNGPTFNCTNLLTKSYNHFFEKVGLKSESELAHQALQINSSGSY

540-541

Pol I

S150_Mip1p WMSARERIQSQFVVP⁶¹⁶NCKFPNEFQSLSAKSSLNNEKTNDLAIIPKIVPMGTITRRAVEN⁶⁶¹TWLTASNAKANRIGSELKTQVKAPPGYCFVGADVDSEELWIASLVGDSIFNVHGGTAIGW
YPH499_Mip1p WMSARERIQSQFVVP⁶¹⁶SCKFPNEFQSLSAKSSLNNEKTNDLAIIPKIVPMGTITRRAVEN⁶⁶¹AWLTASNAKANRIGSELKTQVKAPPGYCFVGADVDSEELWIASLVGDSIFNVHGGTAIGW

616 661

terminal 100 residues of the CTEs (Young *et al.*, 2006, Figure 4.1 A). All three of these helices are removed in the respiratory-deficient mutant lacking 279 C-terminal residues, *MIP1Δ279* (Figure 4.1 A). Interestingly, within the ORF of *MIP1Δ216* only the predicted eleven-residue-helix is maintained, while nearly two-thirds of the largest predicted 15-residue helix is truncated and the smallest seven-residue helix is completely removed (Figure 4.1 A). Further genetic evidence for the requirement of the 15-residue helix was gained from 1) a *MIP1Δ222* mutant expressing a truncation of 222 C-terminal residues, Mip1pΔ222, that deletes the entire 15-residue helix and was unable to grow on glycerol containing media and 2) *MIP1Δ205* and *MIP1Δ175* in which Mip1p retains that helix in both truncations, and both mutants maintained wild-type levels of respiratory competence and low levels of point mutations in the mtDNA (Chapter 3 and Figure 4.1 A).

To further investigate the mechanisms for the maintenance of wild-type levels of mtDNA and fidelity in Mip1pΔ175 and Mip1pΔ205 and also the loss of mtDNA in mutants lacking either all three of the predicted helices (Mip1pΔ279 and Mip1pΔ241), two of the helices (Mip1pΔ222) or a truncation within a helix (Mip1pΔ216), we overexpressed the wild-type and truncation polymerase variants in yeast. There were two additional points to consider. First, during the course of this work a study by Baruffini *et al.* (2007b) identified a single nucleotide polymorphism in the *MIP1* gene of strains related to S288c that changes a strictly conserved threonine 661 in *MIP1*[Σ] to alanine (*MIP1*[S], see Chapter 2). They showed that this change causes an increase in mtDNA deletions and point mutations. As all of our plasmid variants encoding various truncation mutants were initially constructed with a *MIP1* gene cloned from BY4700 (Gelperin *et*

al., 2005), an S288c-related strain (*MIP1*[S], Baruffini *et al.*, 2007b), each variant was changed to the wild-type Σ 1278b *MIP1* form (*MIP1*[Σ]).

The second consideration was that yeast nuclear-encoded mitochondrial proteins, like Mip1p, are translated in the cytoplasm. Proteins destined for the mitochondrial matrix typically contain an N-terminal signal sequences in the peptide, which is initially recognized by a translocase of the mitochondrial outer membrane (TOM) for import through the mitochondrial outer membrane (Neupert and Herrmann, 2007). N-terminal signal sequences specific for targeting proteins to the mitochondrial matrix are also called matrix targeting sequences (MTSs). These MTSs in most situations are proteolytically cleaved by a mitochondrial processing peptidase (MPP) within the mitochondrial matrix. The *MIP1* ORF contains five methionine codons located within the first 40 codons encoding the N-terminus of the predicted 1280 amino acid Mip1p (Figure 4.1 B). Foury demonstrated that transcription of *MIP1* occurs after the first putative translational start codon of the ORF *in vivo*, resulting in a predicted 1254-amino acid residue Mip1p preprotein, pre-Mip1p (Foury, 1989). Similarly, the MitoProt II computational method (<http://ihg.gsf.de/ihg/mitoprot.html>, Claros and Vincens, 1996) predicts that when Mip1p is translated from the second ATG of the ORF there is a probability of 0.9946 for targeting of the enzyme to mitochondria and a cleavage site is predicted between residues 41 and 42 (Figure 4.1 B). In contrast, when translation begins from the first ATG of the *MIP1* ORF a probability of 0.2294 for targeting is calculated and no MPP cleavage site was predicted by the program. To confirm these predictions, short (1254 amino acids) and long (1280 amino acids) versions of Mip1p were created and analyzed for import into mitochondria.

Here we examine overexpressed mitochondrial DNA polymerase variants for the following: 1) targeting to mitochondria, 2) association with a mitochondrial membrane fraction, and 3) DNA polymerase activity of mitochondrial membrane fractions obtained from yeast over-expressing wild-type and truncation variants of both *MIP1*[S] and *MIP1*[Σ] forms.

4.3 Materials and Methods

4.3.1 General methods

Restriction digestions were carried out according to manufacturers' procedures. General molecular biological procedures were carried out according to (Sambrook and Russell, 2001) and yeast transformations were carried out according to (Gietz and Woods, 2002). Linear DNA fragments obtained from restriction digestion of plasmids or PCR amplification, and used for yeast transformation or cloning, were gel purified using the QIAquick Gel Extraction Kit (QIAGEN Inc., Mississauga, ON). Plasmid and PCR product sequencing was carried out according to the manufacturers' procedures in the BigDye® Terminator v1.1 Cycle Sequencing Kit Protocol manual (Applied Biosystems, Foster City, CA). Unless otherwise stated, all centrifugation steps were carried out in a SORVALL® pico microfuge (MANDEL, Guelph, Ontario). SDS-PAGE gels were made and set up according to the Mini-PROTEAN® II Electrophoresis Cell Instruction Manual and Western blotting was carried out according to the Mini-PROTEAN® 3 Cell Instruction Manual, both of which were obtained from BIO-RAD (Hercules, CA).

4.3.2 Media

Escherichia coli cells were grown in Luria-Bertani medium (LB, 1% Bacto-tryptone, 0.5% Bacto-yeast extract, 1% NaCl and 1.5% Bacto-agar when LB plates were made). When transformed with betalactamase-encoding plasmids ampicillin was added to LB broth or plates at a final concentration of 100 µg/ml.

Yeast were grown on various media including: YPG (1% yeast extract, 2% peptone, 3% glycerol); YG-Er (2% yeast extract, 3% glycerol, 50 mM potassium phosphate, pH 6.5, 4 g/L erythromycin); YP2D (1% yeast extract, 2% peptone, 2% dextrose); YP10D (1% yeast extract, 2% peptone, 10% dextrose); various synthetic complete media including, SC-TRP, SC-TRP-URA, SC-TRP-LEU, SC-LEU (0.67% Difco™ nitrogen base pH 5.6 without amino acids, 4% dextrose, all amino acids except the ones indicated after “SC” where -TRP, -URA, and -LEU indicate the absence of tryptophan, uracil, and leucine respectively, and the amino acid concentrations used were as described by (Sherman, 1991). SRaff-LEU (0.67% Difco™ nitrogen base pH5.6 without amino acids, 2% raffinose contains all amino acids except leucine; raffinose was made as a filter-sterilized 20% stock solution and was added to media after autoclaving to prevent conversion to glucose); 3xYPGal (3% yeast extract, 6% peptone, and 6% galactose; galactose was made as a filter-sterilized 20% stock solution and was added to the solution after autoclaving). SRaff-TRP-LEU/2% galactose (0.67% Difco™ nitrogen base without amino acids, 2% raffinose, all amino acids except tryptophan and leucine, pH 5.6; raffinose and galactose were made, filter sterilized, and added to media as described above). All components were obtained from Fisher Scientific (Ottawa, Ontario) except for erythromycin which was obtained from Research Products International Corp.

(Mt. Prospect, Illinois). Percentages indicate weight per volume and all solid media contained 2% Bacto™ agar.

4.3.3 Isolation of yeast genomic DNA

Yeast cultures were grown in 5 ml YP10D for 19 to 20 hours at 30°C with shaking to approximately 15 OD₆₀₀. 1.4 ml was placed into each of two microfuge tubes and centrifuged at 5,000 rpm for 2 min followed by removal of the supernatant. Each pellet was washed with 1 ml 1.2 M sorbitol/50 mM EDTA pH 8.5, vortexed briefly and centrifuged at 5,000 rpm. The supernatant was removed and the two pellets were combined in 1 ml 1.2 M sorbitol/50 mM EDTA pH 8.5/0.113 mg/ml lyticase (276 units/mg, Sigma-Aldrich, St. Louis, MO), followed by incubation at 30°C for one hour with gentle shaking. Next, cells were spun for 2 minutes at 5,000 rpm, the supernatant was removed and cells were resuspended in 1 ml of 50 mM EDTA, pH 8.5, 0.2% SDS, followed by heating at 65°C for 15 minutes. 100 µl of 5 M potassium acetate was added and samples were mixed and cooled on ice for 10 minutes. Tubes were centrifuged at 13,000 rpm at 4°C for 10 minutes and then approximately 500 µl of supernatant was transferred to new microfuge tubes. 1 ml of 95% ethanol and 75 µl 3 M sodium acetate, pH 7.0 was then added to each 500 µl sample and tubes were mixed and incubated at -20°C for 1 hour or overnight. Microfuge tubes were centrifuged at 13,000 rpm at 4°C for 15 minutes, and the supernatant was discarded. Pellets were washed with 70% ethanol and then dried. Each pellet was dissolved in 30 µl TE (10 mM Tris-Cl, pH 8.0/1 mM EDTA) + RNaseA (50 µg/ml).

4.3.4 Strain and plasmid construction

Primers used to construct and confirm strains and plasmids are listed in Table 4.1.

Plasmids are listed in Table 4.2.

The *TRP1+dsRed* gene-disruption cassette was cloned in two steps. First, pUC19::*TRP1* was created. The *TRP1* gene from BS70 (a gift from Walter Neupert, Universität München), was cut out with *SstI* (*SacI*) and *SalI* as an 897-bp fragment. The *TRP1* fragment was then ligated in to pUC19 cut with the same enzymes. Note, *TRP1* in BS70 is polymorphic compared to the S288c genomic sequence available at the *Saccharomyces* Genome Database (SGD, appendix Figure 6.1). To create the plasmid pUC19::*TRP1+dsRed*, harbouring the disruption cassette, the plasmid pVT100U::*dsRed* (a gift from Jodi M. Nunnari, University of California, Davis) was cut with *SphI* to release the 1695-bp *dsRed* gene flanked by the *ADHI* promoter and 3'-*ADHI* terminator. pUC19::*TRP1* was also cut with *SphI*, which does not disrupt the *TRP1* gene, and gel purified. The 1695-bp *dsRed* fragment was ligated to pUC19::*TRP1* and a clone with both genes oriented in the same direction (5'-3') was selected as the template for PCR-mediated gene disruption in yeast.

To create a strain in which the *MIP1* gene is replaced by *dsRED*, encoding a mitochondrial targeted red fluorescent protein, the gene-disruption cassette was PCR amplified from pUC19::*TRP1+dsRed* using the US_*SstI*/DS_*SphI* primer set, followed by transformation into the laboratory strain S150 (*MATa leu2-3,112 his3-Δ1 trp1-289 ura3-52 MIP1[Σ] HAP1*). Transformants were selected on SC-TRP plates and colonies were screened for disruption of *MIP1* by: 1) screening recombination junctions using the DeltaScrMIPFor/DeltaScrREDRev (1213-bp PCR product, flanking the 5'-end recombination junction) and DsRedFor/DeltaScrMIPRev (1099-bp, flanking the 3'-end

Table 4.1 Primers and oligonucleotides used in this study				
Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
US_SstI	GGGTGACCTCGGGCGCACTTTTAGCC GATACTCAGTAAACAACAATAGGT_GC CAGTGAATTTCGAGCTC	56	pUC19::TRP1+dsRed	Bold = sequence for recombination at the <i>MIP1</i> locus; normal = primer for amplification of <i>TRP1+dsRed</i>
DS_SphI	TAATGTGCTGTATATATAAATACAAAT GCGAAAGCTAATGCAGATTTTGC_AC AGCTATGACCATGATTAC	56	pUC19::TRP1+dsRed	Bold = sequence for recombination at the <i>MIP1</i> locus, normal = primer for amplification of <i>TRP1+dsRed</i>
DeltaScrM IPFor	GAGGACATGCTGTATCGG	56	<i>MIP1</i>	Screen <i>mip1D::TRP1+dsRed</i> locus
DeltaScrM IPRev	AGCTAACGAGAAGATGCAG	56	<i>MIP1</i>	Screen <i>mip1D::TRP1+dsRed</i> locus
DeltaScrR EDRev	TGTCATTGCGGATATGGTG	56	<i>dsRed</i>	Screen <i>mip1D::TRP1+dsRed</i> locus
DsRedFor	CAGTTCCAGTACGGCTCC	58	<i>dsRed</i>	Screen <i>mip1D::TRP1+dsRed</i> locus
NtermMIP FixUS	GTCAAGGAGAAGGAATTATCAACAAGT TTGTACAAAAAAGCAGGCTACAAACTC GAGATGACGAAATTGATGGTTAGATC		pBG1805:: <i>MIP1</i>	IRO for modification of 5'-end of the <i>MIP1</i> gene relative to transcriptional ATG; italics = <i>XhoI</i> site engineered into IRO
NtermMIP FixDS	CATCGAGCACAAAACGACACGCAGC GGCCGCCGCCGCACCATTCGCAGCATGC ATTCAGATCTAACCATCAATTCGT		pBG1805:: <i>MIP1</i>	IRO for modification of 5'-end of the <i>MIP1</i> gene relative to transcriptional ATG
BG1805L EU2F	AATTTTTTTGATTTCGGTAATCTCCGAA CAGAAGGAAGAACGAAGGAAGGAGC ACAGACTT_CCTCGAGGAGAACTTCTA G	58	<i>LEU2</i> of YEp13	Bold = pNTBG1805:: <i>MIP1</i> sequence for recombination to remove the <i>URA3</i> gene and replace it with <i>LEU2</i> ; normal = primer for amplification of <i>LEU2</i> from YEp13

Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
BG1805L EU2R	TACAATTCCTGATGCGGTATTTCTC CTTACGCATCTGTGCGGTATTTACA CCGCATA_GTCGACTACGTCGTTAAGG	58	<i>LEU2</i> of YEp13	Bold = pNTBG1805:: <i>MIP1</i> sequence for recombination to remove the <i>URA3</i> gene and replace it with <i>LEU2</i> ; normal = primer for amplification of <i>LEU2</i> from YEp13
MIPDSofE Rev	GAAGATGTAGATAATCCACTC	58	<i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>
MipSeq1R	AGCATAATCAGGAACATCG	54	HA-epitope tag in pBG1805:: <i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>
MIPSeq2R	CTGGACTTGCATCGCAAG	56	<i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>
MIPSeq3b R	ACACCTGTGCATTGCAATCC	60	<i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>
MIPSeq3R	CTTTTAAGTAACTGACTCGC	56	<i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>
MIPSeq4b R	GTTGATTTATCTGTGGAAGC	56	<i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>
MIPSeq4R	CTCCTTCGACCATATTACGG	60	<i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>
MIPSeq5R	GTTGTATTCTTCGAGAACCC	58	<i>MIP1</i>	Sequencing plasmids derived from pBG1805:: <i>MIP1</i>

Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
ADH1_For	CATCATCATCACCATCATGGTAGAAT CTTTTATCCATACGATGTTTCTGATTA TGCTTGA_CGCCACTTCTAAATAAGCG	56	ADH1 terminator from pKT128, (Sheff and Thorn, 2004)	Bold = pBG1805::MIP1-derived sequence for recombination to remove the Protease 3C site and the Protein A ZZ domain from the 3'-end of the gene (Gelperin <i>et al.</i> , 2005); <i>italics</i> = stop codon engineered into IRO; normal = primer for amplification and addition of ADH1 terminator
GFP_Rev	CGGCCAGTGAGCGCGCGTAATACGAC TCACTATAGGGCGAATTGGGTACCC GGCGCGCC_GAGGCAAGCTAAACAGAT C	56	ADH1 terminator from pKT128, (Sheff and Thorn, 2004)	Bold = pBG1805::MIP1-derived sequence for recombination; normal = primer for amplification and addition of ADH1 terminator
MIP_TAG_For	TTGGTTGAGCTGGAAAGGG	58	<i>MIP1</i>	Sequencing plasmids derived from pBG1805::MIP1, to confirm modifications made to the 3'-end of the gene
GFP_For	AGCACTACAAATAGAAATTTGGTTGA GCTGGAAAGGGACATTACTATTTCTA GAGAGTAC_GGTGACGGTGCTGGTTTA	56	pKT128	Bold = pBG1805::MIP1-derived sequence for recombination; normal = primer for in-frame fusion to yEGFP and addition of ADH1 terminator
OLIdltM_for	GTCGAGATGGGGATTATATGTAGTTGTT GAGCAACGAGGGACAAGTATTACGAAA TTGATAGTTAGATCTGAATGCATCC		pBG1805::MIP1-yEGFP+LEU2	IRO for changing the 2nd, 3rd, and 4th methionines to isoleucines; bold italics = mutations introduced into IRO
OLIdltM_rev	GGAGAACCATCGAGCACAAAACACTGCAC ACGCAGCGGCCCGCCGCGCACAAATTCG CAGGATGCATTCAGATCTAACTATCA		pBG1805::MIP1-yEGFP+LEU2	IRO for changing the 3rd, 4th, and 5th methionines to isoleucines; bold italics = mutations introduced into IRO

Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
Truncation_Scrn_R	Truncation_Scrn_R	56	ADH1 terminator	Sequencing primer that binds in ADH1 terminator of pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> variants, to confirm various truncation mutants
BG_delta175_F	GTACAGCCTCCTAGACTATATAAAGGATGTCGAGAAGGGCAAAGGACTAAAGTACGTATTATGGGATCCAACCCAGCTT		pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D175</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta175_R	AAAAGATTCTACCATGATGGTGATGATGATGTCTAGACACATCAACCACTTTGTACAAGAAAGCTGGGTTGGATCCCATA		pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D175</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta279_F	GGACTGCATAACCCCTCGAACAAAACCGCCATTCCTCATGGGGAGGCGCTTGATATCAATCAACTGCTAAACCCAGCTT		pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D279</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta279_R	AAAAGATTCTACCATGATGGTGATGATGATGTCTAGACACATCAACCACTTTGTACAAGAAAGCTGGGTTTAGCAGTTGA		pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D279</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta216_F	GTTCTTAAAATATTTTCTTGCGATGCAAGTCCAGTCAGATAAGCGCGATGTGAATCGGCTAGAAGATGAGAACCCAGCTT		pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D216</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta216_R	AAAAGATTCTACCATGATGGTGATGATGATGTCTAGACACATCAACCACTTTGTACAAGAAAGCTGGGTTCTCATCTTCT		pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D216</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta_alpha1_F	CAGCAAAGTATCACAATATGCCTATAACTACAGAGAACCTGTATTTGAAGAATATAATAAATCTTATACTAACCCAGCTT		pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D241</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain

Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
BG_delta_alpha1_R	AAAAGATTCTACCATGATGGTGATGATG ATGTCTAGACACATCAACCACTTTGTAC AAGAAAGCTGGGTTAGTATAAGAT		pNTBG1805:: <i>MIP1</i> - ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D241</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta_alpha2_F	TAAATCTTATACTCCAGAGTTCTTAAAA TATTTTCTTGCGATGCAAGTCCAGTCAG ATAAGCGCGATGTGAACCCAGCTT		pNTBG1805:: <i>MIP1</i> - ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D222</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
BG_delta_alpha2_R	AAAAGATTCTACCATGATGGTGATGATG ATGTCTAGACACATCAACCACTTTGTAC AAGAAAGCTGGGTTACATCGCGC		pNTBG1805:: <i>MIP1</i> - ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D222</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
D205F	GTCAGATAAGCGCGATGTGAATCGGCT AGAAGATGAGTATCTGCGGGAGTGTAC ATCCAAAGAATACGCTAACCCAGCTT		pNTBG1805:: <i>MIP1</i> - ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D205</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
D205R	AAAAGATTCTACCATGATGGTGATGATG ATGTCTAGACACATCAACCACTTTGTAC AAGAAAGCTGGGTTAGCGTATTCT		pNTBG1805:: <i>MIP1</i> - ADH1Term+ <i>LEU2</i>	IRO for <i>MIP1D205</i> truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain
MIP4_F_PolyFix	CGTAATATGGTCGAAGGAG	56	<i>MIP1</i>	for PCR amplification of the polymorphic region including the conserved threonine 661 codon of <i>MIP1</i> [sigma] from S150
RevPolyFix	TTCGTTGATCAGCTTTGGC	56	<i>MIP1</i>	for PCR amplification of the polymorphic region including the conserved threonine 661 codon of <i>MIP1</i> [sigma] from S150

Name	Description	Plasmid cut with which R.E. to create	Relevant IROs or Primers	Reference
BS70	pBluescript harbouring <i>TRP1</i>	pBluescript cut with <i>Bam</i> HI and a <i>Bg</i> II <i>TRP1</i> fragment was cloned into this site	N/A	a gift from W. Neupert, University of Munich
pUC19:: <i>TRP1</i>	<i>TRP1</i> subcloned from BS70 as a <i>Sst</i> I (<i>Sac</i> I)- <i>Sal</i> I 897-bp fragment	puC19 cut with <i>Sst</i> I (<i>Sac</i> I) and <i>Sal</i> I	N/A	This work
pVT100U:: <i>dsRed</i>	<i>dsRed</i> fluorescent protein gene can be cut out as a cassette with <i>Sph</i> I		N/A	a gift from Jodi M. Nunnari, University of California, Davis
pUC19:: <i>TRP1+dsRed</i>	<i>dsRed</i> subcloned from pVT100U:: <i>dsRed</i> ; template for one step gene replacement	pUC19:: <i>TRP1</i> cut with <i>Sph</i> I	N/A	This work
pBG1805:: <i>MIP1</i>	Open Biosystems Mip1p expression plasmid	N/A	N/A	Open Biosystems (Huntsville, AL)
pNTBG1805:: <i>MIP1</i>	N-terminus adjusted relative to first ATG of transcription	pBG1805:: <i>MIP1</i> with <i>Sph</i> I	NtermMIPFixUS/NtermMIPFixDS	This work
pNTBG1805:: <i>MIP1+LEU2</i>	Replacement of <i>URA3</i> with <i>LEU2</i>	pNTBG1805:: <i>MIP1</i> cut with <i>Stu</i> I	BG1805LEU2F/BG1805LEU2R	This work
pBG1805:: <i>MIP1+LEU2</i>	Replacement of <i>URA3</i> with <i>LEU2</i>	pBG1805:: <i>MIP1</i> cut with <i>Stu</i> I	BG1805LEU2F/BG1805LEU2R	This work

Name	Description	Plasmid cut with which R.E. to create	Relevant IROs or Primers	Reference
pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> (or pMIPS28)	Addition of ADH1 terminator and removal of the Protease 3C site and the Protein A ZZ domain from the 3'-end of the gene	pNTBG1805:: <i>MIP1</i> + <i>LEU2</i> cut with <i>AscI</i>	ADH1_For /GFP_Rev	This work
pBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i>	Addition of ADH1 terminator and removal of the Protease 3C site and the Protein A ZZ domain from the 3'-end of the gene	pBG1805:: <i>MIP1</i> + <i>LEU2</i> cut with <i>AscI</i>	ADH1_For /GFP_Rev	This work
pMIP1[Short]-yEGFP	In-frame carboxyl-terminal fusion of <i>MIP1</i> to yEGFP and addition of ADH1 terminator	pNTBG1805:: <i>MIP1</i> + <i>LEU2</i> cut with <i>AscI</i>	GFP_For/GFP_Rev	This work (Remy-Martin's project)
pMIP1[Long]-yEGFP	In-frame carboxyl-terminal fusion of <i>MIP1</i> to yEGFP and addition of ADH1 terminator	pBG1805:: <i>MIP1</i> + <i>LEU2</i> cut with <i>AscI</i>	GFP_For/GFP_Rev	This work (Remy-Martin's project)
pMIP1[LongM-->I]-yEGFP	Site directed mutation of the 2nd, 3rd, 4th, and 5th methionines to isoleucine	pBG1805:: <i>MIP1</i> -yEGFP+ <i>LEU2</i> cut with <i>SphI</i>	OLIdltM_for/OLIdltM_rev	This work (Remy-Martin's project)
pD175S28	<i>MIP1</i> Δ175 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> cut with <i>BclI</i>	BG_delta175_F/BG_delta175_R	This work
pD205S28	<i>MIP1</i> Δ205 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> cut with <i>PstI</i>	D205F/D205R	This work

Name	Description	Plasmid cut with which R.E. to create	Relevant IROs or Primers	Reference
pD216S28	<i>MIP1</i> Δ216 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> cut with <i>PstI</i>	BG_delta216_F/BG_delta216_R	This work
pD222S28	<i>MIP1</i> Δ222 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> cut with <i>PstI</i>	BG_delta_alpha2_F/BG_delta_alpha2_R	This work
pD241S28	<i>MIP1</i> Δ241 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> cut with <i>PstI</i>	BG_delta_alpha1_F/BG_delta_alpha1_R	This work
pD279S28	<i>MIP1</i> Δ279 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> cut with <i>PstI</i>	BG_delta279_F/BG_delta279_R	This work
pMIPSig	Replacement of the S288c alanine 661 polymorphic region with threonine from S150	pNTBG1805:: <i>MIP1</i> -ADH1Term+ <i>LEU2</i> cut with <i>AvrII</i> and <i>PstI</i> to remove 1540-bp fragment	MIP4_F_PolyFix/RevPolyFix PCR product from S150 genomic DNA	This work
pD175Sig	Replacement of the S288c alanine 661 polymorphic region with threonine from S150 and <i>MIP1</i> Δ175 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pMIPSig cut with <i>BclI</i>	BG_delta175_F/BG_delta175_R	This work
pD205Sig	Replacement of the S288c alanine 661 polymorphic region with threonine from S150 and <i>MIP1</i> Δ205 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pMIPSig cut with <i>PstI</i>	D205F/D205R	This work

Name	Description	Plasmid cut with which R.E. to create	Relevant IROs or Primers	Reference
pD216Sig	Replacement of the S288c alanine 661 polymorphic region with threonine from S150 and <i>MIP1</i> Δ175 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pMIPSig cut with <i>Pst</i> I	BG_delta216_F/BG_delta216_R	This work
pD222Sig	Replacement of the S288c alanine 661 polymorphic region with threonine from S150 and <i>MIP1</i> Δ222 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pMIPSig cut with <i>Pst</i> I	BG_delta_alpha2_F/BG_delta_alpha2_R	This work
pD241Sig	Replacement of the S288c alanine 661 polymorphic region with threonine from S150 and <i>MIP1</i> Δ241 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pMIPSig cut with <i>Pst</i> I	BG_delta_alpha1_F/BG_delta_alpha1_R	This work
pD279Sig	Replacement of the S288c alanine 661 polymorphic region with threonine from S150 and <i>MIP1</i> Δ279 truncation of 3'-end of the <i>MIP1</i> gene; attB2, HIS-6, and HA-epitope remain	pMIPSig cut with <i>Pst</i> I	BG_delta279_F/BG_delta279_R	This work

recombination junction) primer sets 2) screening for the absence of the *MIP1* gene via DeltaScrMIPFor/ DeltaScrMIPFor primers, which generates a 3086-bp PCR product when *MIP1* is replaced with *TRP1+dsRed* and 4263-bp product from wild-type *MIP1* 3) the inability of the PCR-confirmed *mip1Δ* strains to grow on YPG and 4) red punctuate mitochondrial fluorescence via by Zeiss Epi-fluorescent microscopy, indicating dsRed expression.

The *MIP1* plasmid that was modified in this study, pBG1805::*MIP1* (Figure 4.2), was purchased from Open Biosystems (Huntsville, AL). The ORF was confirmed by sequencing using MIPDSofERev and the MipSeq series of primers shown in Table 4.1. Variants of pBG1805::*MIP1* were constructed using a modification of the elegant *delitto perfetto* method developed by Francesca Storici (Figure 4.3, Storici and Resnick, 2006).

pNTBG1805::*MIP1*

Foury (1989) has previously shown by S1 nuclease mapping of the 5'-termini of *MIP1* transcripts that transcription initiates after the first putative translational start codon of the ORF (Figure 4.1 A and B) and therefore, results in an enzyme that is 1254 amino acid residues long. However, the pBG1805::*MIP1* was engineered based on conceptual translation starting at the upstream methionine to yield a 1280 amino acid preprotein. Therefore, we modified the 5'-end of the *MIP1* ORF in pBG1805::*MIP1* linearized with *SphI* (Figure 4.2) to create pNTBG1805::*MIP1*. This plasmid has 78-bp of the 5'-region, including the first ATG, removed and introduces an *XhoI* site just upstream of the new start codon (the second in-frame methionine codon see Figure 4.1 B). Integrative recombinant oligonucleotides (IROs), NtermMIPFixUS and

Figure 4.2 pBG1805 Gateway®-type yeast expression plasmid. The plasmid used in this study was pBG1805::*MIP1* purchased from Open Biosystems, where *MIP1* is the open reading frame (ORF) flanked by *attB* sites. *attB1* and *attB2*, Gateway-modified *att* site that site-specifically recombine with only *attP1* and *attP2* respectively; *URA3*, encodes orotidine-5'-phosphate (OMP) decarboxylase required for uracil prototrophy in yeast; *bla*, beta-lactamase gene for ampicillin resistance in *E. coli*; *GALI*, yeast galactose inducible promoter. The C-terminal tag consists of the following: 6xHis, hexahistidine-tag for protein purification by nickel-nitrilotriacetic acid chromatography; HA, Hemagglutinin epitope tag; Protease 3C, protease 3C cleavage site; Protein A (ZZ domain), the immunoglobulin G (IgG) binding ZZ domain of protein A from *Staphylococcus aureus*. Taken from yeast_orf_manual.pdf from Open Biosystems with permission from Dr. Elizabeth Grayhack who created the plasmid.

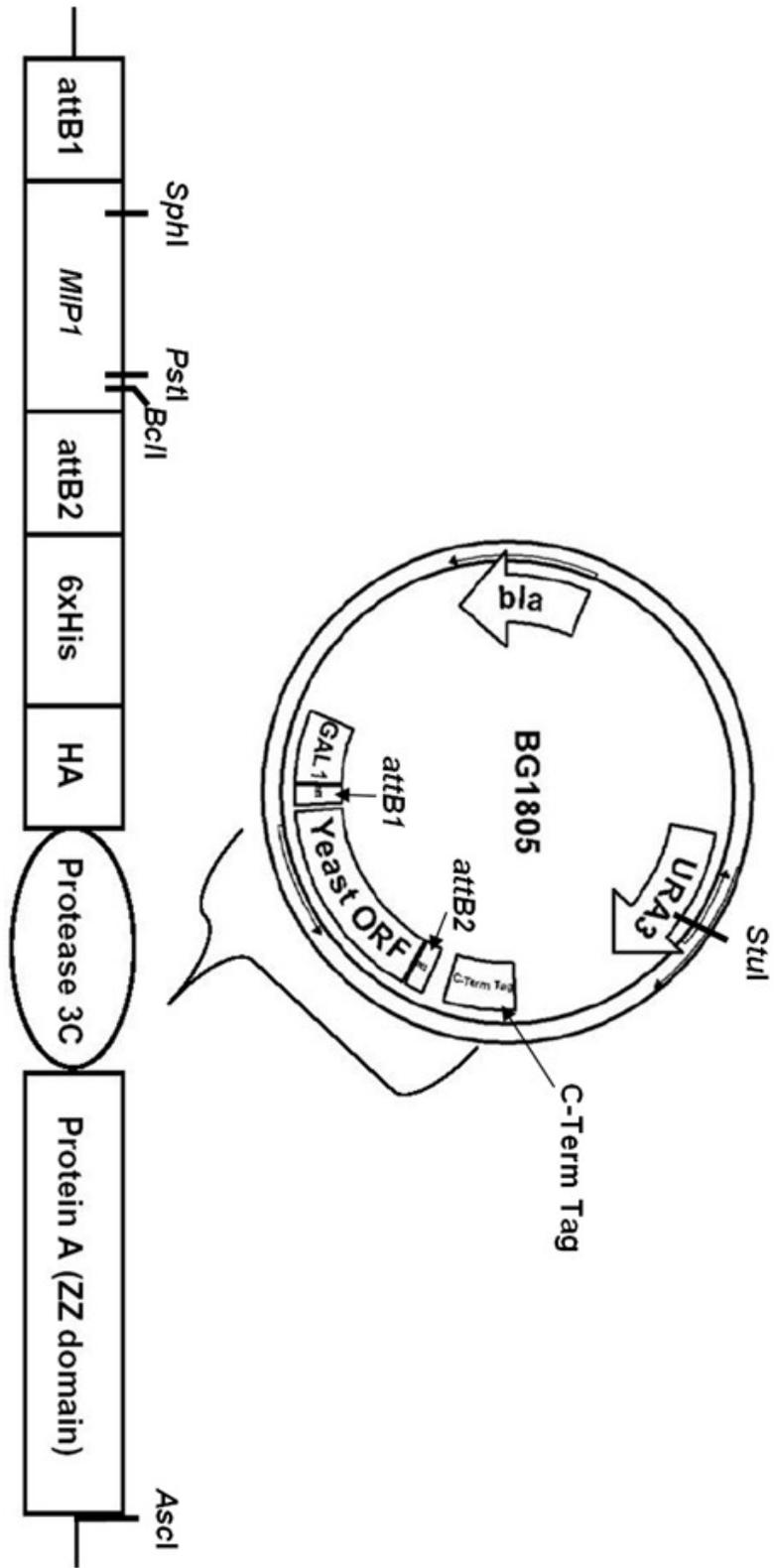
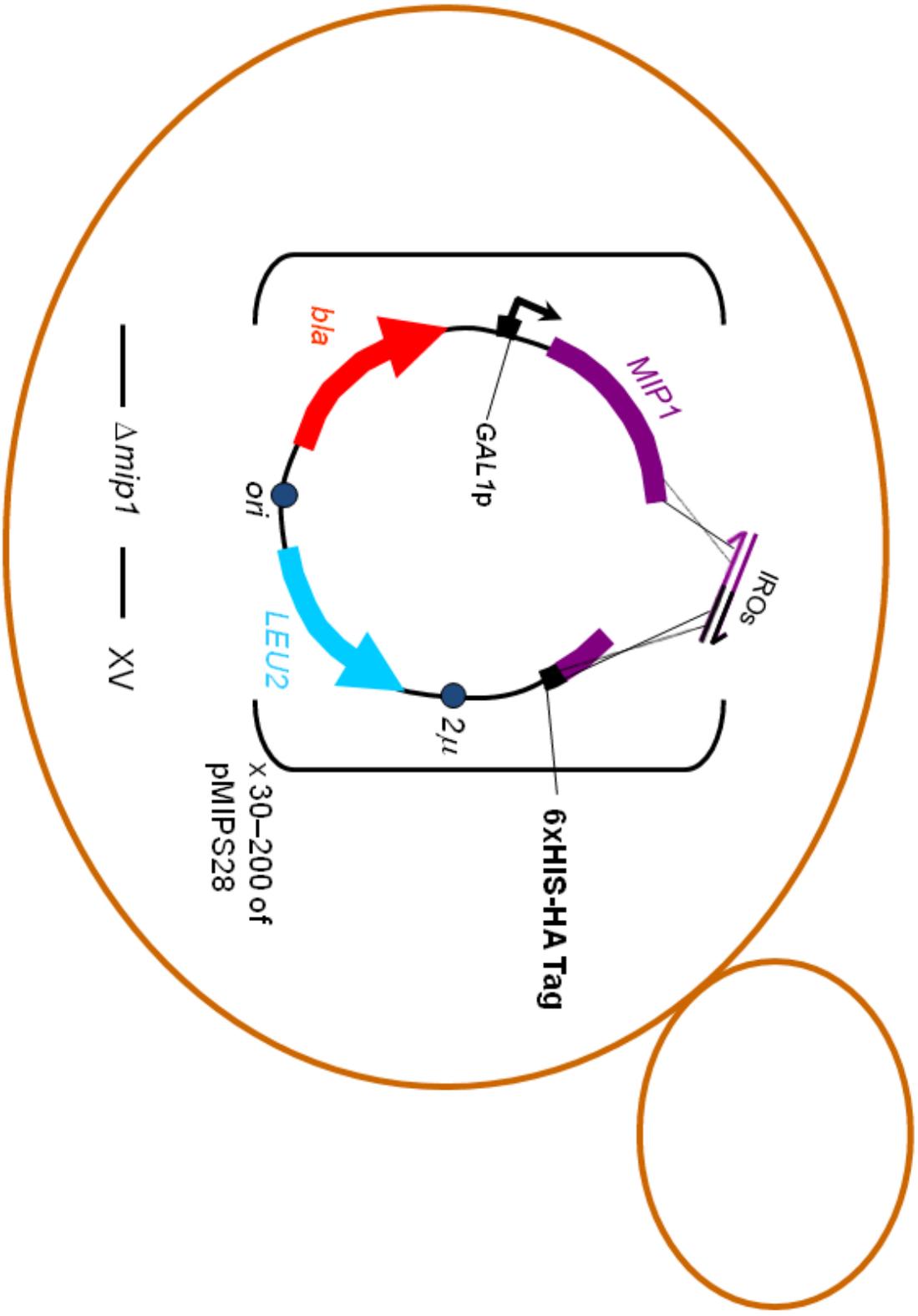


Figure 4.3 Yeast recombination-based, 3'-end modification of pMIPS28 harbouring the *MIP1* gene. The restriction-digested, linearized plasmid was electroporated into S150 *mip1* Δ (*leu2* auxotroph, *MIP1* gene on chromosome XV removed, represented by a thin line) with integrative recombinant oligonucleotides (IROs). The yeast homologous recombination machinery recombines the IROs with pMIPS28 (thin lines between IROs and pMIPS28) regenerating a circular plasmid and allowing it to replicate via the two micron origin of replication (2 μ) and to complement the S150 *mip1* Δ *leu2* auxotrophy. GAL1p, yeast galactose inducible promoter (arrow shows direction of transcription); *bla*, beta-lactamase gene for ampicillin resistance in *E. coli*; *ori*, *E. coli* plasmid origin of replication.



NtermMIPFixDS, were annealed and extended first by heating to 94°C for 1 minute followed by incubation at 68°C for 3 minutes in the following reaction mixture: 10 µl of each IRO (0.1 nmol/µl), 2 µl recombinant Taq polymerase provided by Jack Switala (University of Manitoba), 10 µl 10x Invitrogen PCR buffer, 4 µl 10 µM dNTPs, with one of 2, 4, or 6 µl 100 mM MgSO₄, and ddH₂O to bring the final volume to 100 µl. After extension all three tubes (2, 4, 6 µl MgSO₄) were combined together and precipitated at -20°C overnight with two volumes of 95% ethanol and a final concentration of 0.3 M sodium acetate pH 5.2. Extended IROs were then centrifuged at 13,000 rpm for 15 minutes, washed with 70% ethanol, dried, and resuspended in 60 µl ddH₂O. 100 ng of Antarctic phosphatase (NEB, Ipswich, MA) treated *SphI*-cut pBG1805::*MIP1* was then added to 11 µl of extended IROs and brought up to 100 µl, followed by ethanol precipitation and washing as described above. The plasmid and IROs mix was then resuspended in 9 µl ddH₂O and electroporated into S150 *mip1Δ*.

Electroporation conditions were modified from (Thompson *et al.*, 1998) and were as follows: S150 *mip1Δ* was grown overnight in 20 ml YP10D. Cells were diluted down to 0.35 OD₆₀₀ the next morning and grown to 1 – 1.5 OD₆₀₀ in 50 ml YP10D. Next, cells were centrifuged at 5,000 rpm in a Sorvall SS-34 rotor for 5 minutes and the pellet was resuspended in 16 ml 0.1 M lithium acetate, 10 mM dithiothreitol, 10 mM Tris-Cl, pH 7.5, 1 mM EDTA and incubated at room temperature for 1 hour. Cells were collected by centrifugation at 5,000 rpm in a Sorvall SS-34 rotor for 5 minutes and resuspended in 16 ml ice-cold ddH₂O. This wash step was repeated and then cells were suspended in 6.4 ml ice-cold 1 M sorbitol and collected by centrifugation. The pellet was resuspended in 54 µl of ice-cold 1 M sorbitol (to give a total volume of approximately 160 µl of cell

suspension) and 44 μ l of the suspension was combined with 9 μ l of the plasmid and IROs mix and incubated on ice for 5 minutes. The cell/DNA mix was then transferred to a 0.2 cm electroporation cuvette and pulsed using the “Sc2” program with the BioRad Micropulser. 1 ml of ice-cold 1 M sorbitol was then added, the solution was mixed, followed by transfer to a 1.5 ml microfuge tube. The mixture was centrifuged at 6,000 rpm for 2 minutes and 800 μ l of the supernatant was discarded leaving ~200 μ l and the cell pellet. The cells were resuspended in the remaining 200 μ l and 100 μ l was plated in duplicate on SC-TRP-URA. After four days of incubation at 30°C, DNA was collected from the colonies on the plate by a “smash-and-grab” crude DNA preparation as described in (Hoppins, 2006). This DNA was resuspended in 30 μ l TE and 6 μ l was pulsed with 50 μ l electrocompetent DH5 α *Escherichia coli* using the “Ec1” program of the BioRad Micropulser. 1 ml of ice-cold LB was added to the *E. coli* in the cuvette and gently mixed, followed by transfer into a microfuge tube and incubation at 37°C with gentle mixing for three hours. Cells were centrifuged at 13,000 rpm for 1 minute and ~900 μ l of the supernatant was removed while leaving ~100 μ l and the *E. coli* pellet. This remaining 100 μ l was then mixed with a pipette tip and spread onto an LB/ampicillin (AMP) plate followed by incubating overnight at 37°C.

Small *E. coli* colonies were inoculated at the end of the next day into 5 ml LB/AMP and incubated with shaking overnight at 37°C. The slightly turbid cultures were diluted early in the morning 1 in 4 into 20 ml LB/AMP and allowed to grow for another four hours. 15 ml of the culture was centrifuged at 12,000 rpm in a Sorvall SS-34 rotor and plasmid DNA was prepared using the QIAprep® miniprep kit (Mississauga, Ontario) with the following modifications. For 15 ml of culture, the pellet was resuspended in 250

μl of P1 buffer (step 1), the QIAprep® spin column was washed with buffer PB (step 7), buffer PE was allowed to sit on the membrane for 5 minutes (step 8), and 35 μl prewarmed (65°C) ddH₂O was added, instead of buffer EB, and let stand at room temperature for 8 minutes before centrifugation (step 10).

The potential pNTBG1805::*MIP1* plasmids were initially screened by linearization of the plasmid with *XhoI* to confirm introduction of this site via the IRO NtermMIPFixUS. Candidate were then sequenced with MIPSeq5R to confirm correct recombination at the 5'-end of the *MIP1* gene.

pNTBG1805::*MIP1+LEU2* and pBG1805::*MIP1+LEU2*

The second modification to pBG1805::*MIP1* constructs was replacement of the *URA3* gene with *LEU2* in pNTBG1805::*MIP1* and pBG1805::*MIP1* to create pNTBG1805::*MIP1+LEU2* and pBG1805::*MIP1+LEU2*. *LEU2* was PCR-amplified from YEp13 (Broach *et al.*, 1979) using BG1805LEU2F/BG1805LEU2R primers and recombination was carried out as described above with pNTBG1805::*MIP1* and pBG1805::*MIP1* linearized with *StuI*, which cuts in *URA3* (Figure 4.2). Approximately 200 ng of purified PCR product was precipitated with 50 ng of purified dephosphorylated vector and this mixture was used for electroporation of S150 *mip1Δ*. Transformants harboring recombinant plasmids were selected on SC-TRP-LEU.

pMIP1[Short]-yEGFP, pMIP1[Long]-yEGFP and pMIP1[LongM→I]-yEGFP

pMIP1[Short]-yEGFP and pMIP1[Long]-yEGFP plasmids were constructed to see if both the short and long forms of Mip1p, with respect to the amino-terminus, are targeted to mitochondria. In all Mip1p-yEGFP constructs created here, the primer GFP_For was engineered to fuse the last codon of the *MIP1* gene, encoding a tyrosine

residue, in-frame with the linker region just upstream of yEGFP described in Sheff and Thorn (2004). These constructs therefore lack the *attB2* site as well as the 3'-end encoding the carboxyl-terminal 6xHis-HA tag (Figure 4.2). pNTBG1805::*MIP1+LEU2* and pBG1805::*MIP1+LEU2* were linearized with *AscI*, located immediately after the stop codon in the plasmids, and a yEGFP-containing PCR product was amplified from pKT128 (Sheff and Thorn, 2004) with GFP_For/GFP_Rev for recombination in S150 *mip1Δ*. The correct recombination at the 3'-end of the *MIP1* gene in the plasmid was confirmed by sequencing with MIP_TAG_For.

pMIP1[LongM→I]-yEGFP was created by cutting pBG1805::*MIP1-yEGFP+LEU2* with *SphI* (position 111 of *MIP1* ORF) to linearize it at the 5'-end of the gene, followed by recombination in yeast with the extended IROs OLIdltM_for/OLIdltM_rev to create site-directed mutations of the 2nd, 3rd, 4th, and 5th methionine codons to isoleucine codons (Figure 4.1 B). This was done to ensure that translation starts from the first ATG. The correct recombination at the 5'-end of the *MIP1* gene in the plasmid was confirmed by sequencing with MIPSeq5R.

pMIP1[Short]-yEGFP, pMIP1[Long]-yEGFP and pMIP1[LongM→I]-yEGFP were all electroporated into S150 *mip1Δ* as described above and selected for on SC-TRP-LEU. Transformants were grown overnight in 20 ml SC-TRP-LEU and the next morning were centrifuged down at 6,000 rpm in a Sorvall SS-34 rotor and the supernatant was discarded. The cells were resuspended in 10 ml SRaff-TRP-LEU/2% galactose at 0.3 OD₆₀₀ and grown for 3-4 hours. 1.12 OD₆₀₀ of cells were washed twice with ddH₂O followed by resuspension in 150 μl ddH₂O (7.5 OD₆₀₀). 40 μl of this resuspension was mixed with 40 μl of 0.7% low-gelling agarose (made in ddH₂O) on a glass slide and a 22

x 60 mm cover-glass was quickly added and gently pressed on to get an even layer of cells. Slides were made in duplicate and sealed with clear nail polish. Fluorescence was detected in the green (GFP, Mip1p variants tagged with yEGFP) and red (rhodamine, mitochondrial targeted dsRed) channels using a Zeiss Axio Imager Z1 epi-fluorescent microscope. This work was performed with the assistance from an honors project student, Remy-Martin Gratton.

pNTBG1805::*MIP1*-ADH1Term+*LEU2* (pMIPS28) and pBG1805::*MIP1*-ADH1Term+ *LEU2*

To create pMIPS28 and pBG1805::*MIP1*-ADH1Term+ *LEU2*, the next modification made was the one step addition of a 5'-end *ADH1* terminator PCR product (amplified from pKT128, Sheff and Thorn, 2004) and removal of DNA encoding the Protease 3C cleavage site and the Protein A ZZ domain from the 3'-end of the *MIP1* gene from both plasmids (Figure 4.2). Both plasmids were separately linearized with *AscI* (located immediately downstream of the stop codon) and purified, followed by precipitation with the *ADH1* PCR product amplified with the ADH1_For /GFP_Rev primer set, and electroporated into S150 *mip1Δ*. The correct recombination at the 3'-end of the *MIP1* gene was confirmed by sequencing with the sequencing primer MIP_TAG_For. The final plasmid, pMIPS28, is the wild-type S288c derived *MIP1* gene used for *in vitro* analyses and used to construct all other S288c-*MIP1*-truncation plasmids and the Mip1p[Σ] wild-type plasmid (see below).

pD175S28, pD205S28, pD216S28, pD222S28, pD241S28, and pD279S28

Plasmids were named by deletion in and structure of the *MIP1* gene i.e. pD205S28 encodes the Mip1p from S288c (Mip1p[S]) with a carboxyl-terminal deletion (D) of 205 amino acids. pD205S28, pD216S28, pD222S28, pD241S28, and pD279S28 were all constructed in a similar manner. pMIPS28 was cut with *PstI* (located in the coding region of the CTE, Figure 4.2) and IROs engineered for the coding sequence in *MIP1Δ205*, *MIP1Δ216*, *MIP1Δ222*, *MIP1Δ241*, and *MIP1Δ279* (Table 4.1) were annealed, extended, mixed with vector, precipitated, and electroporated into yeast. The resulting plasmids were collected from yeast, electroporated into *E. coli*, and DNA was prepared as described above. All constructs were confirmed by *AscI/SphI* restriction endonuclease mapping and sequencing using the Truncation_Scrn_R primer to confirm the 3'-end recombination junction.

pD175S28 was made by cutting pMIPS28, passed through the *dam*⁻ *E. coli* strain GM2163, with *BclI* (blocked by *dam* methylation and this site is also located in the CTE coding region but further downstream than *PstI*) followed by the same method described above but using the *MIP1Δ175* specific IROs.

pMIPSig, pD175Sig, pD205Sig, pD216Sig, pD222Sig, pD241Sig, and pD279Sig

pMIPSig containing the *MIP1[Σ]* allele (Baruffini *et al.*, 2007b) was made by cutting pMIPS28 with *AvrII* and *PstI* to remove the 1540-bp fragment (see Figure 4.1 A) followed by recombination in S150 *mip1Δ* with the MIP4_F_PolyFix/RevPolyFix PCR product generated from S150 (*MIP1[Σ]*) genomic DNA. The entire *MIP1* gene in this plasmid was sequenced. pD175Sig, pD205Sig, pD216Sig, pD222Sig, pD241Sig, and

pD279Sig were created and confirmed exactly as their S288c-counterparts described above, except the pMIPSig was the starting plasmid instead of pMIPS28.

4.3.5 Induction of recombinant Mip1p[S] and Mip1p[Σ] and variants

Wild-type S150 was chosen over S150 *mip1Δ* as the background strain for expression of all Mip1p variants as it gave the best expression and mitochondrial membrane protein fractions isolated from untransformed S150 cells gave very little background polymerase activity (see results).

S150 was electroporated with each of the plasmids encoding Mip1p[S], Mip1p[Σ] and truncation variants described above. Transformants were selected on SC-LEU plates after three or four days growth at 30°C. Cells were collected by pipetting 4 ml of SC-LEU onto the first plate, followed by collecting colonies with a sterile glass slide to obtain a homogenous mixture, and then pipetting this mix to the second plate and repeating the process. The mixture of cells was then aliquoted to two microfuge tubes and spun at 6,000 rpm. To ensure the same concentration of cells in each aliquot, approximately 0.8 ml of supernatant was removed from each replicate tube and then the remaining 0.7 ml plus the pellet was mixed and combined together into one microfuge tube. This tube was vortexed for 10 seconds to mix cells and then ~0.75 ml was aliquoted to each of two screw capped microfuge tubes containing 0.2 ml sterile glycerol. These tubes were vortexed and 200 µl aliquots were frozen down at -60°C.

Induction of transcription of the plasmid borne *MIP1*-variants from the *GAL1p* was carried out according to Gelperin *et al.* (2005) with the following modifications. One aliquot of cells from -60°C was allowed to thaw at room temperature and then was inoculated into 50 ml SC-LEU and allowed to grow approximately 14 hrs at 30°C with

shaking at 220 – 240 rpm. An OD₆₀₀ was taken and approximately 50 OD₆₀₀ of cells was centrifuged at 6,000 rpm in a Sorvall SS-34 rotor at room temperature for 10 minutes, followed by resuspension of the pellet in 50 ml SRaff-LEU. These cells were grown for nine hours at 30°C and then subcultured into 500 ml SRaff-LEU to give an OD₆₀₀ of 0.1. The cultures were grown for another twelve hours to an OD₆₀₀ of 1 to 2. Induction of recombinant polymerase was initiated by adding 250 ml of 3x YPGal in the morning and allowing cultures to grow a final five hours. After induction, the culture was centrifuged in two 500 ml centrifuge tubes in a Sorvall GSA rotor at 6,000 rpm for 10 minutes at 4°C and the supernatant was decanted. Both pellets were resuspended together in 2 ml of 50 mM Tris-Cl, pH 8.0, 30% glycerol (TG) and then cells were aliquoted to microfuge tubes on ice. Cells were centrifuged for 2 minutes at 6,000 rpm in a SORVALL® pico microfuge, supernatant was removed and the wet weight of cells in each tube was determined. Next, 0.5 ml of TG was added to each pellet and mixed with a 1 ml Pipetman™ followed by a 10 second vortex. Cells in 0.5 ml TG were immediately frozen down at -60°C for preparing mitochondrial membrane proteins.

4.3.6 Isolation of mitochondria and mitochondrial membrane proteins

Procedures for isolation of mitochondria and mitochondrial membrane proteins were derived from those of Mueller *et al.* (2004) and Lewandowska *et al.* (2006). 1 to 2 g wet weight of induced frozen cell pellets were thawed at room temperature and then immediately put on ice and all subsequent steps were carried out at 4°C or on ice. Cells were centrifuged at 6,000 rpm for 2 minutes and TG was aspirated off without disturbing the cell pellet. The cell pellets were resuspended in one volume of cold yeast breaking buffer (YBB, 0.65 M sorbitol, 0.1 M Tris-Cl, 5 mM EDTA, 0.2% BSA, pH 8.0, 1 mM

PMSF) and pooled together in a 15-ml conical tube. Two volumes of YBB was then added to the cells in the conical tube (total of three volumes of the initial cell pellet) and vortexed for 5 seconds. Next, three volumes of glass beads (0.45 – 0.53 mm) were added and the tube was vortexed for 1 minute followed by incubating on ice for 1 minute. This step was repeated five more times. Conical tubes were centrifuged at 3,000 rpm for 5 minutes to remove cell debris and the supernatant was transferred into microfuge tubes and kept on ice (Sup 1). The cell pellet was extracted a second time by resuspending in 2.2 volumes of fresh YBB and vortexing for 1 minute followed by incubation on ice for 1 minute. Vortexing was repeated twice more followed by centrifugation at 3,000 rpm for 6 minutes and the resulting pellet was discarded and the supernatant was transferred into fresh microfuge tubes and kept on ice (Sup 2). Microfuge tubes containing Sup 1 and Sup 2 were next centrifuged at 12,000 rpm for 8 minutes to collect mitochondrial pellets. The supernatant obtained from the Sup 2-tubes was discarded, while the more concentrated one from the Sup 1-tubes was aliquoted to new microfuge tubes (Sup 3). Each mitochondria-containing pellet was washed with 500 μ l SEM (250 mM sucrose, 1 mM EDTA, 10 mM MOPS, pH 7.5 with KOH) followed by an additional centrifugation at 3,500 rpm for 4 minutes and the pellets were discarded while the mitochondrial supernatant was transferred to fresh microfuge tubes on ice. Sup 3 tubes were centrifuged again at 12,000 rpm. These mitochondrial pellets were also washed with SEM and spun at low rpm as described. All of the mitochondrial supernatants obtained from the 3,500 rpm spins were then pelleted at 12,000 rpm for 8 minutes and the SEM was gently aspirated off. For every 1.2 g of wet weight cells, mitochondria were resuspended in 0.6 ml of hypotonic lysis buffer (20 mM MOPS, pH 8.5, 3 mM β -mercaptoethanol, 1 mM

PMSF, 1 µg/ml leupeptine, final concentration was typically 2.3 – 2.5 mg mitochondrial protein per 350 µl hypotonic lysis buffer). The volume of mitochondria was estimated by measuring the total volume of the resuspension with a 1 ml Pipetman™. A 20 µl aliquot of mitochondria was mixed with an equal volume of 2x SDS-PAGE sample buffer (0.0625 M Tris-Cl, pH 6.8, 10% glycerol, 2% SDS, 5% β-mercaptoethanol, and 0.05% bromophenol blue) for determining the release of mitochondrial matrix proteins (see below) and a 3 µl aliquot was frozen at -20°C for determining the protein concentration of mitochondria via a Bradford assay (Sigma-Aldrich, St. Louis, MO).

The remaining mitochondria were subjected to homogenization with a 27 ½ G needle through eleven cycles of drawing up and expelling the sample. The material was vortexed for 30 seconds followed by incubation on ice for 30 seconds and then vortexed and incubated on ice again. The homogenization and vortexing was repeated exactly as just described. One tenth of the volume of 20 mM MOPS, pH 8.5, 4 M NaCl, 3 mM β-mercaptoethanol, 1 mM PMSF, and 1 µg/ml leupeptine was added to the homogenized mitochondria to bring the salt concentration to 0.4 M. The mitochondria were subjected to homogenization with the same 27 ½ G needle through eleven more cycles and then vortexed followed by incubation on ice. The sample was vortexed and incubated on ice again. The mitochondria were homogenized through a final eleven cycles, this time without vortexing. The final volume was measured with a 1 ml Pipetman™ and 3 µl was taken for a Bradford assay. A 50 µl-aliquot was taken and centrifuged at 13,000 rpm for 15 minutes to obtain a supernatant and a pellet fraction to measure the release of mitochondrial matrix proteins (see below). The supernatant was mixed with an equal volume of 2x SDS-PAGE sample buffer and an equal volume of 2x SDS-PAGE sample

buffer was added to the pellet. Both were heated at 95°C for 3 minutes and then stored at -60°C.

The remaining homogenized mitochondria were loaded onto a sucrose gradient (0.25 ml 60%, 0.25 ml 55%, 2.25 ml 18% (wt/wt) sucrose in 20 mM Tris-Cl pH 7.4, 0.5 mM EDTA, 5 mM β -mercaptoethanol) and centrifuged in a Beckman TLA 100.3 rotor at 60,000 rpm for 1 hour at 4°C. 75 μ l fractions containing mitochondrial membrane proteins from the interface between 55% and 60% sucrose were collected from the bottom portion of the tube (Lewandowska *et al.*, 2006). For each fraction 3 μ l and 12 μ l aliquots were taken for Bradford assay and SDS-PAGE respectively. The remaining ~60 μ l was mixed with 50 μ l sterile glycerol and 11 μ l aliquots were frozen down at -60°C.

4.3.7 Characterization of the association of Mip1p with the mitochondrial membrane

Mitochondria were isolated from 0.5 g wet weight of S150 cells transformed with pMIPS28 as described above. Approximately 0.1 g of cells was aliquoted into each of four tubes. After washing mitochondrial pellets in 500 μ l SEM and pelleting mitochondria at 12,000 rpm for 8 minutes, the mitochondria were resuspended in 20 μ l SEM each (x 10 pellets) and all tubes were combined. 3 μ l was aliquoted out for a Bradford assay. For each wash condition 20 μ l of 9.9 μ g/ μ l of mitochondria was used. Each 20 μ l aliquot was centrifuged at 13,000 rpm for 10 minutes and the volume of supernatant was measured (15 μ l) and discarded. Each pellet was resuspended in 1 ml of either 6 M urea, 1 M NaCl, 1% TritonX-100, 0.1 M Na₂CO₃, or SEM and incubated on ice for 30 minutes. Samples were then centrifuged at 13,000 rpm for 30 minutes and supernatants were aspirated into new microfuge tubes. 15 μ l of 2x SDS-PAGE sample

buffer was added to each pellet followed by heating at 95°C for 3 minutes and storing samples at -60°C. Soluble proteins in the supernatants were precipitated with 5% TCA (trichloroacetic acid) for NaCl and TritonX-100 treated samples or 10% TCA for urea and Na₂CO₃ treatments. Tubes were then incubated on ice for 12 minutes followed by a 13,000 rpm centrifugation for 15 minutes. Supernatants were discarded and pellets were washed with 1 ml ice cold acetone followed by centrifugation at 13,000 rpm for 10 minutes. Acetone was aspirated off and pellets were dried in at 56°C for 5 minutes. 15 µl of 2x SDS-PAGE sample buffer was added to each pellet and pellets were then heated at 95°C for 3 minutes. 5 µl of each sample was loaded onto two 12% SDS-PAGE gels and run at 100 mV for 1 hour and 46 minutes followed by Western blotting.

4.3.8 Western blotting

12 % SDS-PAGE gels were transferred to nitrocellulose membranes overnight at 10 mV/30 mA. For the blot shown in Figure 4.6, the membrane was stained with Ponceau S (0.2% Ponceau S/3.15% TCA), scanned and then cut into two parts between the 30 and 40 kDa molecular weight marker of the BenchMark™ Protein Ladder (Invitrogen, Carlsbad, California). Next the two pieces of the blot were washed with shaking for 5 minutes in 1x TBS (10 mM Tris-Cl, pH 7.5, 150 mM NaCl) and then blocked in 1xTBS/6.5% skim milk for 30 minutes with shaking. The top part of the blot, with molecular markers above 120 kDa, was incubated with a 1 in 1000 dilution of Anti-HA Tag mouse monoclonal antibody (Applied Biological Material Inc., Richmond, BC) in 1x TBS/6.5% skim milk/0.2% Tween-20 for 1 hour. The bottom part of the blot was incubated for the same amount of time with 1 in 1000 dilution of anti-Tim17p rabbit polyclonal antibody (a gift from K. Hell, Universität München) in 1x TBS/6.5% skim

milk/0.05% Tween-20. Next, both pieces were washed twice with 1x TBS/0.3% TritonX-100 for 25 minutes followed by washing twice with 1x TBS for 8 minutes. The top part of the blot was next incubated in a 1 in 6,000 dilution of anti-mouse IgG (whole molecule)–alkaline phosphatase antibody produced in goat (Sigma-Aldrich, St. Louis, MO, A-3562) in 1x TBS/6.5% skim milk and the bottom part of the blot was incubated in a 1 in 2,000 dilution goat anti-rabbit IgG (H + L) alkaline phosphatase conjugate (ZYMED, San Francisco, CA) in 1x TBS/6.5% skim milk for 1 hour with shaking. The wash steps were repeated. Membranes were equilibrated for 2 minutes in detection buffer (0.1 M Tris-Cl, 0.1 M NaCl, pH 9.5) and then 1 ml of a 1 in 100 dilution of CDP-Star (Roche Laval, Quebec, Cat. No. 11 685 627 001), made in detection buffer, was rinsed over the blot by pipetting five times followed by sandwiching the pieces of the blot in a page protector. The blot was detected on Kodak Scientific Imaging Film X-OMAT™ LS (Fisher Scientific, Ottawa, Ontario) or using the Alpha Innotech FluorChem 8900 according to manufacturers' procedures.

For the Western blot shown in Figure 4.7, the membranes were processed exactly the same as described above except the Ponceau S stained membrane was cut into three pieces, between 30 and 40 kDa and between 80 and 90 kDa of the BenchMark™ Protein Ladder, and the middle piece (40 to 80 kDa) was probed with a 1 in 1000 dilution of anti-Put2p rabbit polyclonal antibody in 1x TBS/6.5% skim milk (a gift from Dr. Debkumar Pain, New Jersey Medical School, Newark, NJ).

4.3.9 Release of mitochondrial matrix proteins from intact mitochondria

Disruption of mitochondrial membranes was determined by Western blot analysis of mitochondrial fractions using the mitochondrial inner membrane specific anti-Tim17p

antibody and the anti-Put2p antibody, which reacts with the soluble mitochondrial matrix protein Put2p. 1) Mitochondria resuspended in hypotonic lysis before homogenizing and centrifugation, and equivalent amounts of 2) mitochondrial pellet obtained after homogenization and centrifugation and 3) mitochondrial supernatant after homogenizing and centrifugation were loaded onto a 12% SDS-PAGE gel and analyzed by Western blotting as described above.

4.3.10 DNA polymerase activity measurement of mitochondrial membrane protein fractions using the incorporation of Digoxigenin-11-dUTP

Digoxigenin-11-2'-deoxy-uridine-5'- triphosphate, alkali-stable (DIG-11-dUTP) was purchased from Roche (Laval, Quebec Cat. No. 11 093 088 910). A 10x nucleotide mixture was made as follows: 1 mM each of dATP, dGTP, dCTP, 0.9125 mM dTTP, 0.0875 mM DIG-11-dUTP, 10 mM Tris-Cl pH 7.5. An approximately 4-kb plasmid (MIPF3RI/SallpBSΔBam, also called pBluescript::MIP1-F3, Young *et al.*, 2006) was linearized with *EcoRI* (Invitrogen, Carlsbad, California) and was used as the template DNA. This DNA was boiled for 10 minutes at 100°C prior to addition to the polymerase reaction. Polymerase reactions were set up in 10 µl volumes as follows: 7 µg of mitochondrial membrane protein fraction, 675 ng (determined before boiling) linearized, boiled template, 1 µl 10x hexanucleotide mix (Roche, Laval, Quebec Cat. No. 11 277 081 001), 1 µl 10x nucleotide mix, 0.5 µl 1 mg/ml aphidicolin (a 1 in 10 dilution made in ddH₂O of 10 mg/ml aphidicolin in DMSO). After 5, 10, 15 and 30 minutes, at 30°C or 37°C, 2 µl aliquots of each polymerase reaction were mixed with 24.25 mM EDTA/10 mM Tris pH 8.0 to stop DNA polymerase and samples were stored at -20°C. For each time point of the experiment 1 µl samples were spotted in triplicate onto Hybond-N nylon

(Fisher Scientific, Ottawa, Ontario) and the dot blot was UV crosslinked for 3 minutes. Following crosslinking the blot was equilibrated in malic acid buffer (0.1 M malic acid, 0.15 M NaCl, pH 7.5) for 2 minutes and then incubated in 40 ml blocking buffer (malic acid buffer/1x blocking reagent, Roche Laval, Quebec, Cat. No. 1096176)/0.015% SDS with shaking for 30 minutes. Anti-Digoxigenin-AP conjugate (Roche Laval, Quebec, Cat. No. 1093274) was then added to the dot blot in blocking buffer/0.015% SDS at a 1 in 10,000 dilution and it was incubated with shaking for 1 hour. The blot was rinsed three times with ddH₂O and then twice for 15 minutes each with malic acid buffer/0.3% Tween 20/0.015% SDS with shaking (rinsed with ddH₂O between washes). It was equilibrated for 2 minutes in detection buffer (0.1 M Tris-Cl, 0.1 M NaCl, pH 9.5) and then 1 ml of a 1 in 100 dilution of CDP-Star (Roche Laval, Quebec, Cat. No. 11 685 627 001), made in detection buffer, was rinsed over the blot by pipetting five times followed by sandwiching the blot in a page protector. Quantitation and detection were carried out using the Alpha Innotech FluorChem 8900 according to manufacturers' procedures.

4.3.11 Nickel-nitrilotriacetic acid chromatography of Mip1p[S] Δ 175-6xHis-HA and Mip1p[S]-6xHis-HA

4.3.11.1 Denaturing conditions

Mitochondria were isolated from 0.19 g wet weight of cells expressing Mip1p[S] Δ 175-6xHis-HA as described above with the following modification. After the SEM wash and centrifugation at 12,000 rpm for 8 minutes, mitochondrial pellets were suspended in 57.5 μ l SEM, to generate a final volume of 67.5 μ l. 400 μ l of buffer B (100 mM NaH₂PO₄, 10 mM Tris-Cl, 8 M urea, pH 8.0) was added to the mitochondria followed by vortexing twice briefly to mix. Mitochondria were then homogenized with a

1 cc, 25G 5/8 needle through six cycles of drawing up and expelling the sample. 200 μ l of a slurry of Nickel-nitrilotriacetic acid (Ni-NTA) resin was washed with 1 ml of buffer B and centrifuged at 13,000 rpm for 20 seconds followed by aspiration of the supernatant. The mitochondria in buffer B were added to the Ni-NTA and mixed at room temperature for one hour. Ni-NTA resin was pelleted at 13,000 rpm for 20 seconds and the supernatant was removed but saved to run on an SDS-PAGE gel. The Ni-NTA was washed three times for 5 minutes each with 1 ml of buffer C (100 mM NaH_2PO_4 , 10 mM Tris-Cl, 8 M urea, pH 6.3) followed by centrifugation at 13,000 rpm for 20 seconds and discarding the supernatant. To elute protein 30 μ l of buffer C/100 mM EDTA was added to the resin and the Ni-NTA was mixed for 10 minutes followed by centrifugation at 13,000 rpm for 20 seconds. Two elutions were obtained; each was mixed with an equal volume of 2x SDS-PAGE sample buffer, heated to 95°C for three minutes and loaded onto a 9% SDS-PAGE gel followed by Western blotting with a 1 in 6000 dilution of Anti-HA Tag mouse monoclonal antibody in 1x TBS/6.5% skim milk as described above.

4.3.11.2 Non-denaturing conditions

Mitochondria were prepared from 0.6 g wet weight of cells expressing Mip1p[S]-6xHis-HA as describe above with the following modification. After the final homogenization with a with a 27 ½ G needle, the approximately 250 μ l of mitochondria were centrifuged at 13,000 rpm for 20 minutes and the supernatant (~200 μ l) was transferred into a separate microfuge tube from the pellet. Both the supernatant and pellet tubes were kept on ice. 200 μ l of 20 mM MOPS/pH 8.5/0.4 M NaCl/3 mM β -mercaptoethanol/1 mM PMSF/1 μ g/ml leupeptine (an equal supernatant volume) was

added to the pellet and the pellet was resuspended. Two microfuge tubes of 0.5 ml of Ni-NTA slurry were washed twice with 1 ml of TBG (50 mM Tris-Cl, pH 8.0, 2 mM β -mercaptoethanol, and 30% glycerol) followed by centrifugation at 13,000 rpm for 20 seconds and discarding the supernatant. The approximately 200 μ l of each of mitochondrial supernatant and pellet were added to separate Ni-NTA tubes and mixed for one hour at room temperature. The samples were centrifuged at 13,000 rpm for 20 seconds and the supernatant was discarded. Next the Ni-NTA was washed three times with 1 ml TBG for five minutes with mixing followed by centrifugation at 13,000 rpm for 20 seconds and discarding the supernatant. Protein was eluted off of the Ni-NTA by mixing with 50 μ l TBG/100 mM imidazole for five minutes followed by centrifugation at 13,000 rpm for 20 seconds and collecting 45 μ l which was then mixed with 45 μ l 2x SDS-PAGE sample buffer and heated at 95°C for three minutes. This process was repeated with both TBG/250 mM imidazole and buffer C/100 mM EDTA respectively and then samples were loaded onto a 9% SDS-PAGE gel followed by Western blotting with a 1 in 5,000 dilution of Anti-HA Tag mouse monoclonal antibody in 1x TBS/6.5% skim milk as described above. This protocol was also carried out with homogenized mitochondria prepared from 0.6 g wet weight of cells that were not fractionated into pellet and supernatant fractions.

4.4 Results

4.4.1 *GAL1p* induction of N-terminal variants of Mip1p-yEGFP expressed from a 2 μ plasmid

Before a suitable overexpression system could be developed it was important to test possible MTSs, as five methionine codons are located early in the 5'-region of the

MIP1 ORF. To determine whether or not Mip1p could be overexpressed and targeted to the mitochondria *in vivo*, genes encoding Mip1p precursors with different potential targeting signals were constructed (Materials and Methods). pMIP1[Long]-yEGFP encodes the 1280-amino acid Mip1p preprotein (pre-Mip1p) linked at the CTE to yeast enhanced green fluorescent protein (yEGFP), while pMIP1[Short]-yEGFP contains the *MIP1* ORF starting from the second ATG (1254 amino acid pre-Mip1p as defined by Foury, 1989, and Figure 4.1 B) linked to yEGFP. The pMIP1[LongM→I]-yEGFP plasmid encodes a yEGFP-linked, 1280-amino acid residue Mip1p, with the four N-terminal internal methionine codons mutated to isoleucine codons. This construct was made to ensure translation started from the first ATG only. These plasmids were created by a yeast recombination-based cloning strategy conceptually adapted from Storici and Resnick (Figure 4.3, Storici and Resnick, 2006).

The pMIP1[Long]-yEGFP, pMIP1[Short]-yEGFP, and pMIP1[LongM→I]-yEGFP plasmids were separately electroporated into S150 *mip1Δ*. S150 *mip1Δ* lacks the ability to respire, because the *MIP1* gene is disrupted with the *dsRED* gene, encoding a mitochondrially-targeted red fluorescent protein, dsRed. This protein can be used for fluorescence co-localization studies of proteins imported into mitochondria (Young *et al.*, 2006, Meeusen and Nunnari, 2003). S150 *mip1Δ* transformants, each harbouring separate plasmids (pMIP1[Long]-yEGFP, pMIP1[Short]-yEGFP, or pMIP1[LongM→I]-yEGFP), were subjected to galactose induction, followed by fluorescence microscopic analysis of potential green and red fluorophores (Figure 4.4). In contrast to our prediction, all three Mip1p-yEGFP variants, Mip1p[Long]-yEGFP, Mip1p[Short]-yEGFP, and Mip1p[LongM→I]-yEGFP, formed green puncta that co-localized with mitochondrially-

targeted dsRed, indicating they were all targeted to mitochondria. However, the Mip1p[LongM→I]-yEGFP transformant appears to have increased cytosolic staining relative to pMIP1[Long]-yEGFP and pMIP1[Short]-yEGFP (Figure 4.4) suggesting that not all of the overexpressed Mip1p[LongM→I]-yEGFP is correctly targeted.

The finding that the long form of Mip1p-yEGFP with the four N-terminal internal methionines mutated to isoleucines is targeted to mitochondria was interesting as there is no MTS predicted by MitoProt II. However, it was decided to engineer the expression plasmids to start transcription from the second start codon, as defined by Foury (1989), by deleting 78-bp upstream of the second ATG in all of our constructs (see Materials and Methods).

4.4.2 Association of Mip1p with mitochondrial membranes

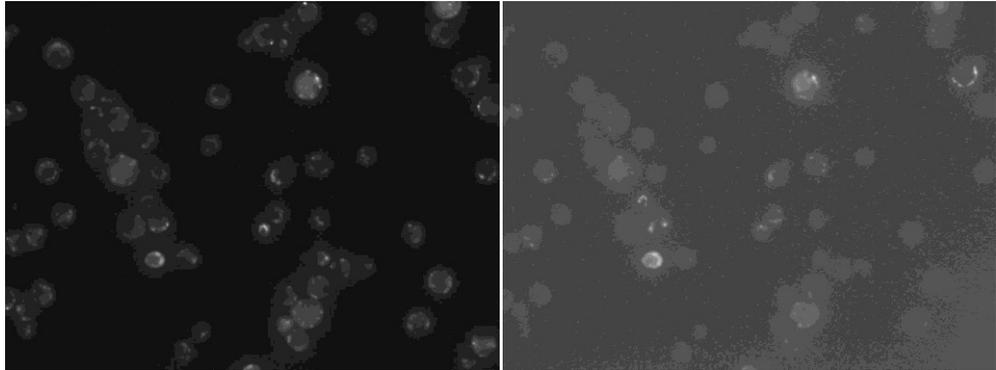
Meeusen and Nunnari have proposed a two membrane-spanning (TMS), autonomous mtDNA replisome model (Meeusen and Nunnari, 2003). The TMS comprises 1) Mgm101p, a component of the mitochondrial nucleoid (Chen *et al.*, 2005) essential for mtDNA maintenance and possibly required for repair of oxidative mtDNA damage 2) Mmm1p, a trans-outermembrane protein required for mtDNA maintenance that coimmunoprecipitates with Mgm101p and has a role in mitochondrial morphology, 3) potentially Mip1p, as shown by colocalization with a Mmm1p-dsRed fusion protein and 4) replicating mtDNA, as shown by association of Mmm1p-GFP and Mgm101p-GFP foci with BrdU foci (Meeusen and Nunnari, 2003). Based on their observation that Mip1p-GFP colocalizes with Mmm1p-dsRed of the TMS, and is therefore likely linked to the mitochondrial inner membrane, the association of Mip1p with a mitochondrial membrane pellet fraction was initially investigated. The efficiency of releasing

Figure 4.4 Targeting of Mip1p amino-terminal variants to the mitochondria. Left- and right-hand panels show green (Mip1p-yEGFP, Mip1p variants tagged with yEGFP) and rhodamine (Mito-dsRed, mitochondrial-targeted dsRed) fluorescence, respectively, from the same field of view for S150 *mip1Δ* harbouring either **A.** pMIP1[Short]-yEGFP, **B.** pMIP1[Long]-yEGFP or **C.** pMIP1[LongM→I]-yEGFP plasmids. S150 *mip1Δ* expresses a chromosomal copy of mitochondrial targeted dsRed due to disruption of *MIP1* with the *TRP1+dsRed* cassette (see Materials and Methods). Note that because dsRed was expressed as a single copy from the genome, red fluorescence was much less intense than that of Mip1p-yEGFP variants. Images obtained from 0.2 μm slices using a 100x objective are presented.

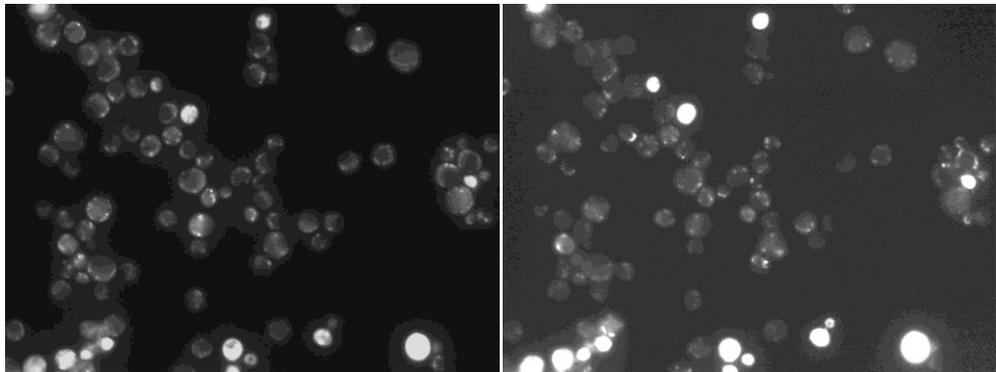
Mip1p-yEGFP

Mito-dsRed

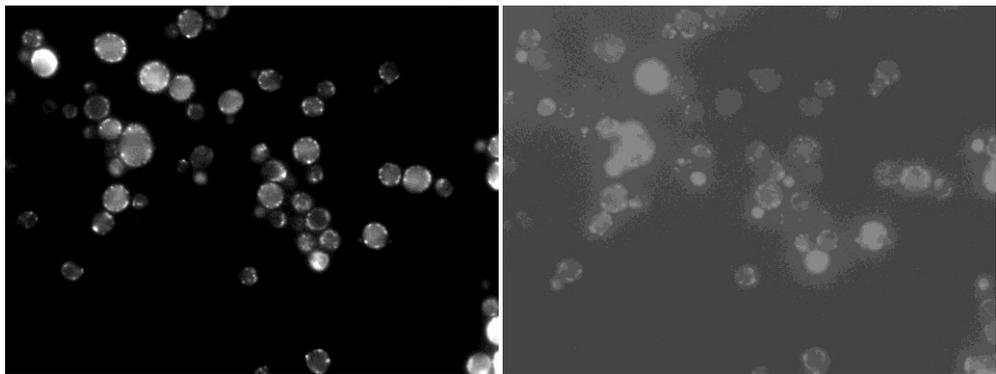
A.



B.



C.



mitochondrial matrix proteins during the preparation of mitochondria was determined based on Western blot analysis of the following mitochondrial fractions: 1) mitochondria resuspended in hypotonic lysis buffer before homogenization and centrifugation, 2) mitochondrial pellets and 3) supernatants obtained after homogenization and centrifugation (Figure 4.5). Proteins present in the supernatant fraction (S) include Put2p and it was estimated that 80% and 60% of Put2p was released from total mitochondria (M) for the Mip1p[S] and Mip1p[Σ] mitochondria shown in Figure 4.5. Therefore, using this method, at least 60% of the soluble mitochondrial proteins were released by this method. Nearly all of Tim17p, translocase of the inner mitochondrial membrane, was located in the pellet fraction. Likewise all of the polymerase detected by immunoblotting was localized only in the pellet fraction (Figure 4.5).

In an attempt to elucidate the nature of the association of Mip1p with the mitochondrial membrane fraction, mitochondria isolated from yeast expressing the Mip1p[S] wild-type version were subjected to various washing treatments (Figure 4.6). Peripheral membrane-associated proteins can typically be released from membranes due to their sensitivity to salt washes, while treatment of membranes with urea and Na₂CO₃ disrupt protein-protein interactions (Rodriguez-Cousino *et al.*, 1995, Horazdovsky and Emr, 1993). Detergents such as TritonX-100 have been used to disrupt direct associations of proteins with cellular membranes (Horazdovsky and Emr, 1993). The majority of the Mip1p[S] and Tim17p following Na₂CO₃ treatment was localized in the pellet fraction, while only a very faint band of each was visible in the supernatant fraction (Figure 4.6). In contrast, the mitochondrial matrix protein Hep1p, was found only in the supernatant fraction

Figure 4.5 Release of mitochondrial matrix proteins from intact mitochondria by hypotonic lysis and homogenization. Mitochondria resuspended in hypotonic lysis buffer before homogenization and centrifugation (M), and corresponding amounts of mitochondrial pellet obtained after homogenization and centrifugation (P), and supernatant after homogenization and centrifugation (S) were run on a 12% SDS-PAGE gel. Mitochondrial membrane fractions (Me) isolated via sucrose gradient centrifugation were also loaded onto the gel. Molecular weight markers are indicated on the left-hand side of the blot. The blot was cut into three pieces and each piece was separately subjected to immunodetection by: 1) anti-HA Tag mouse monoclonal antibody to detect the HA-tagged Mip1p[S] and Mip1p[Σ] (top of blot), 2) the mitochondrial matrix specific anti-Put2p rabbit polyclonal antibody (middle panel) and 3) the mitochondrial inner membrane specific anti-Tim17p rabbit polyclonal antibody (lower panel). The ratio of both Put2p to Tim17p (Put/Tim) and Mip1p to Tim17p (Mip/Tim) are presented below the blot. The amount of Put1p and Mip1p for each of Mip1p[S] and Mip1p[Σ] in the M fraction was set to 1.

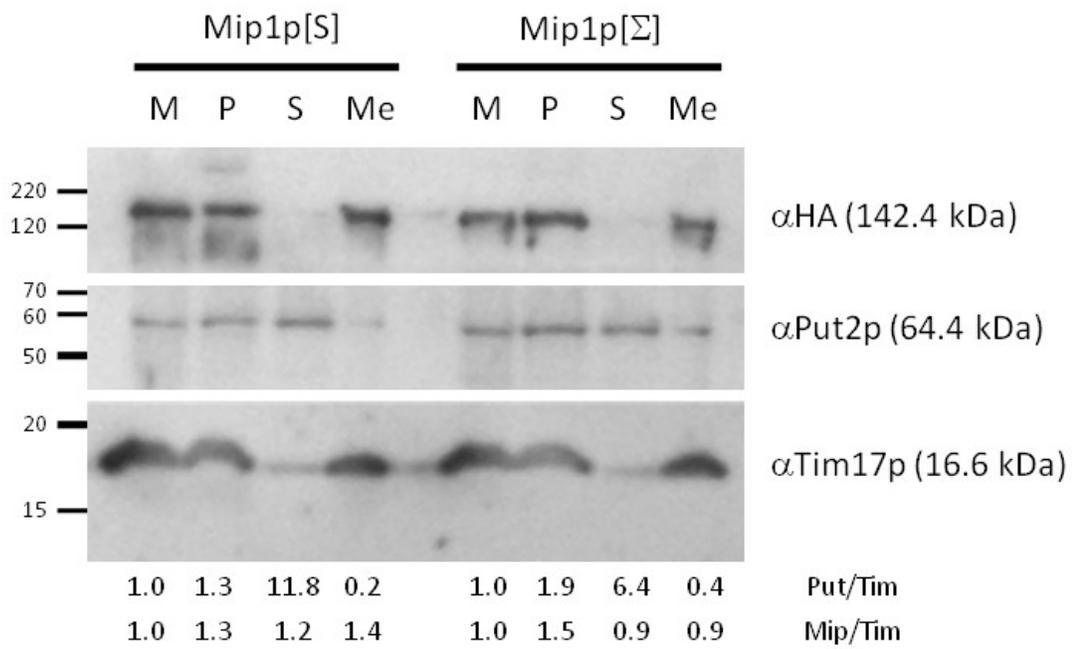
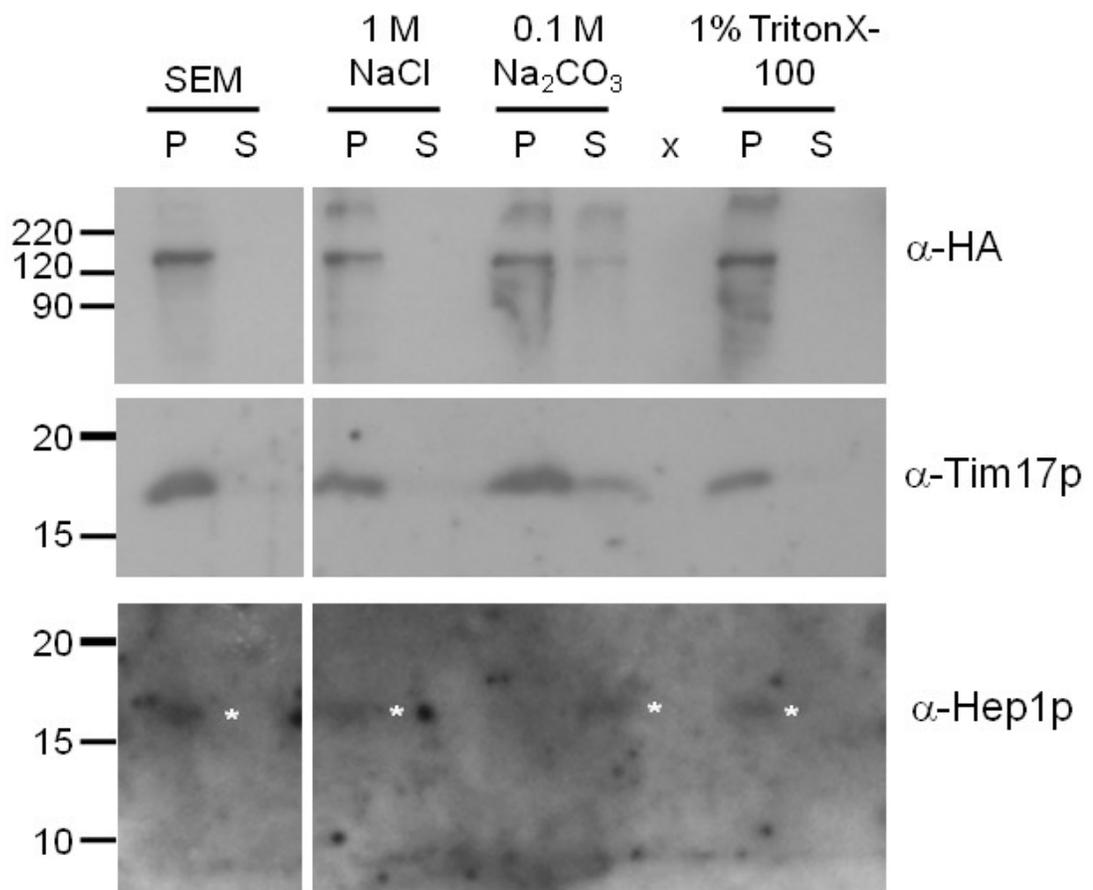


Figure 4.6 Characterization of the association of Mip1p[S] with the mitochondrial membrane. Equal aliquots of mitochondria were adjusted to 1 M NaCl, 0.1 M Na₂CO₃, 1% TritonX-100 or left untreated (SEM), incubated and insoluble material collected by centrifugation. Aliquots of resulting pellets (P) and supernatants (S) were analyzed by SDS-PAGE and Western blot. The top half of the blot was probed for Mip1p[S]-6xHis-HA with anti-HA Tag mouse monoclonal antibody (α -HA), predicted molecular weight of 142.4 kDa after cleavage of predicted signal sequence, and the bottom half with anti-Tim17p rabbit polyclonal antibody (α -Tim17p, 16.6 kDa). A duplicate gel was loaded and run at the same time as the first and the bottom half of the blot obtained from this gel was probed with anti-Hep1p rabbit polyclonal antibody (α -Hep1p, 19.9 kDa, a gift from K. Hell, University of Munich). Due to the low signal obtained from this antibody, white asterisks are used to indicate Hep1p signals. The molecular weights in kDa obtained from the BenchMark™ Protein Ladder are indicated on the left-hand side of the blot.



of the Na₂CO₃ treated mitochondria, indicating sufficient mitochondrial membrane disruption to release proteins from the matrix. For the remaining treatments, all submitochondrial profiles were similar to the control (re-suspending mitochondria in isotonic SEM); Mip1p[S], Tim17p, and Hep1p were consistently localized to the pellet fractions, suggesting that 1% TritonX-100 and 1 M NaCl were ineffective at disrupting mitochondria.

4.4.3 Expression and partial purification of Mip1p[S], Mip1p[Σ], and truncation variants

The S150 (*MIP1*[Σ]) laboratory strain was employed as the yeast strain for polymerase expression as it gave the best expression and mitochondrial membrane protein fractions isolated from untransformed cells gave minimal background polymerase activity (Figure 4.8). Recombinant forms of Mip1p were constructed and electroporated into yeast as outlined in Materials and Methods. Separate cultures harbouring plasmids encoding each form of Mip1p: Mip1p[S], Mip1p[S]Δ175, Mip1p[S]Δ205, Mip1p[S]Δ216, Mip1p[S]Δ222, Mip1p[S]Δ241, Mip1p[S]Δ279, Mip1p[Σ], Mip1p[Σ]Δ175, Mip1p[Σ]Δ205, Mip1p[Σ]Δ216, Mip1p[Σ]Δ222, Mip1p[Σ]Δ241, and Mip1p[Σ]Δ279, were initially grown from a -60°C aliquot overnight under glucose-repressing conditions to repress transcription of the polymerases. The next day glucose repression was lifted by growing the cells in raffinose-containing media for 21 hours followed by 5 hours of galactose induction of *MIP1* transcripts from *GALI*p.

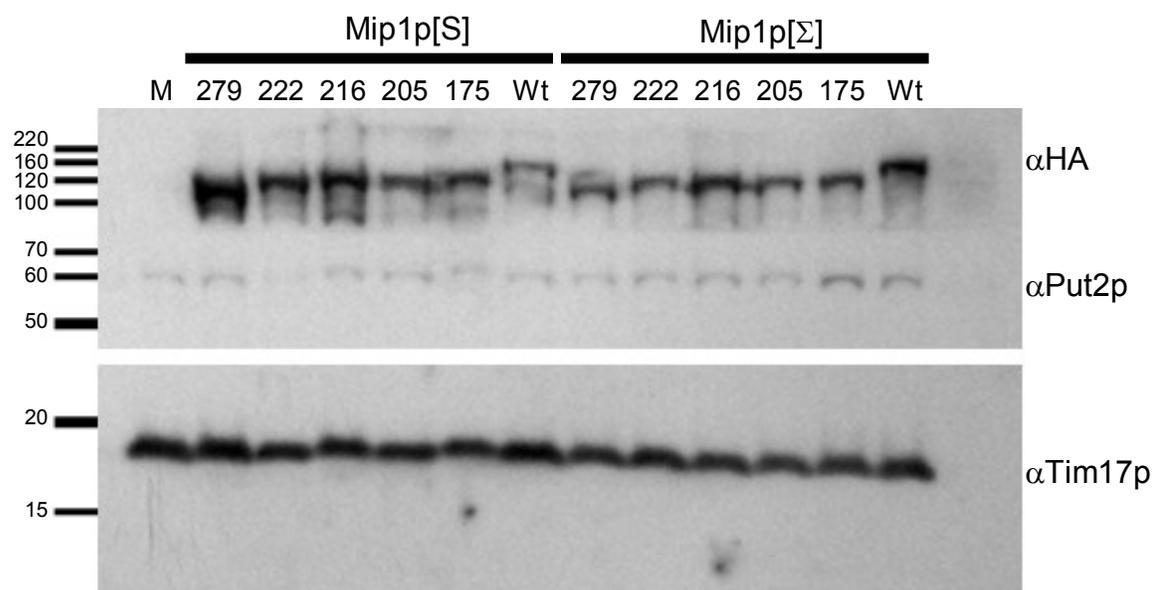
Initially, Mip1p expression was confirmed by Western blot analysis of a crude protein extract from yeast (Kushnirov, 2000) and in crude mitochondrial preparations. Mitochondria were isolated by a method derived from (Mueller *et al.*, 2004) and were

subjected to a hypotonic lysis according to (Lewandowska *et al.*, 2006). Following hypotonic treatment mitochondria were homogenized and a mitochondrial membrane protein fraction was obtained by sucrose gradient ultracentrifugation (Lewandowska *et al.*, 2006). The expression of recombinant polymerases was confirmed by Western blot analysis for the C-terminal HA epitope tag present on each variant (Figure 4.7, predicted molecular weights of N-terminally-processed mature polymerases are: 109.8, 116.6, 117.3, 118.7, 122.1, and 142.4 kDa for Mip1p Δ 279, Mip1p Δ 222, Mip1p Δ 216, Mip1p Δ 205, Mip1p Δ 175 and wild-type versions respectively). A strong signal for the mitochondrial inner membrane protein Tim17p was detected in the mitochondrial membrane protein fractions, while the signal for the mitochondrial matrix protein Put2p was much weaker (on average the Tim17p signal was 5.1-times greater than Put2p) than the Put2p signal observed in total mitochondria (on average 1.4-times more Tim17p in Lanes labelled M, Figure 4.5). Most of the polymerase variants underwent very little N-terminal degradation as indicated by little to no lower molecular weight bands below the full-length mtDNA polymerase. The amount of C-terminal degradation could not be assessed due to the C-terminal location of the HA epitope.

4.4.4 A non-radioactive assay for mitochondrial DNA polymerase

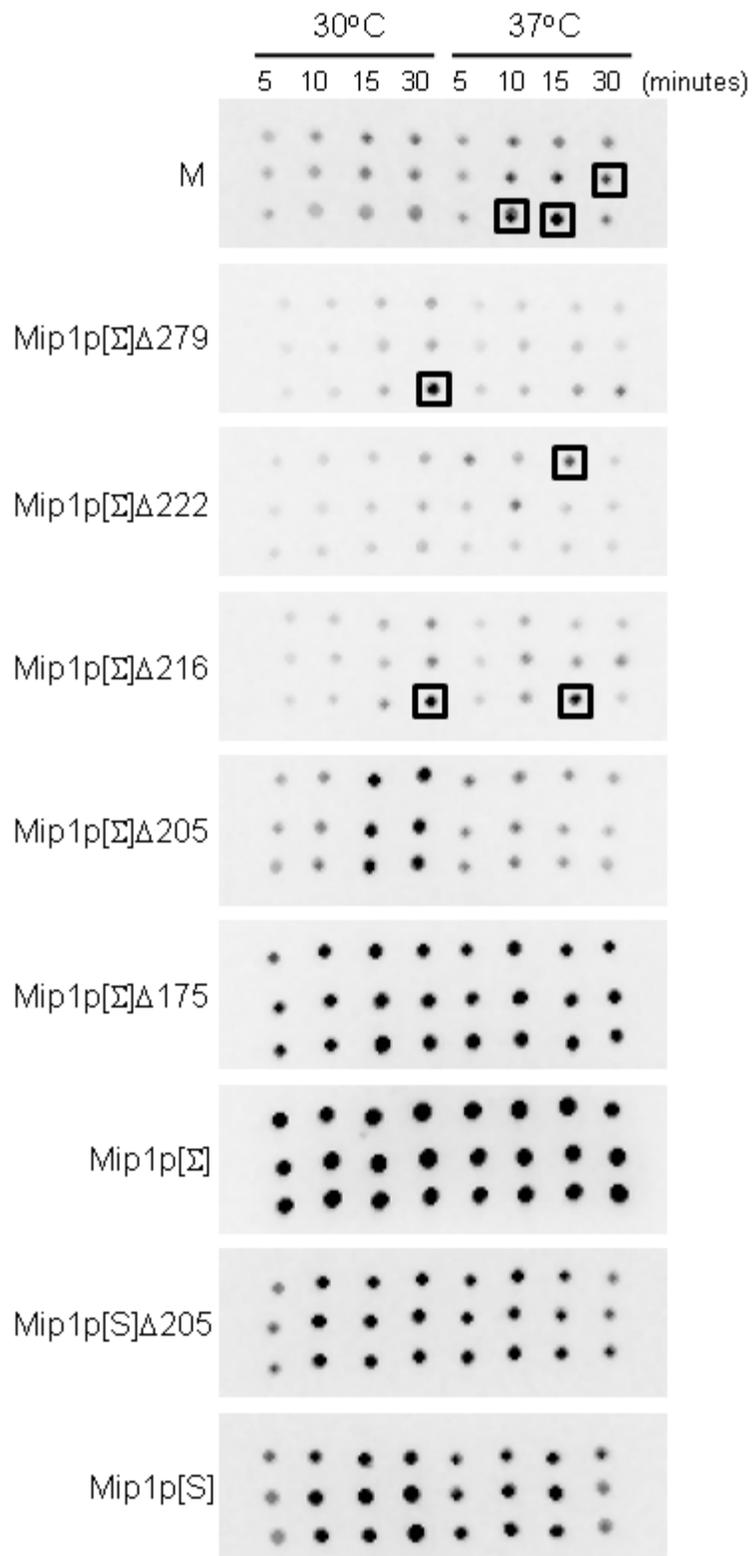
Procedures involving radioactive nucleotide for determining Mip1p activity from mitochondrial membrane protein fractions and from soluble mitochondrial proteins have been described (Lewandowska *et al.*, 2006, Foury, 1989). We developed a non-radioactive assay that is based on the Roche DIG DNA Labeling Kit (Laval, Quebec, Materials and Methods). Isolated mitochondrial membrane protein fractions were combined with DIG-11-dUTP nucleotide mix, template plasmid DNA linearized by

Figure 4.7 Western blot analysis of sucrose gradient membrane protein fractions from yeast cells expressing Mip1p[S] and Mip1p[Σ] truncation variants. M, 279, 222, 216, 205, 175, and Wt are untransformed S150 mitochondrial membranes, and those from cells expressing Mip1p Δ 279, Mip1p Δ 222, Mip1p Δ 216, Mip1p Δ 205, Mip1p Δ 175 and wild-type Mip1p respectively. The blot was cut into three strips and each strip was probed with a separate antibody (Materials and Methods). The top strip of the blot was probed with anti-HA Tag antibody to detect the HA-tagged Mip1p[S] and Mip1p[Σ] variants. The middle strip of the blot was probed with anti-Put2p rabbit polyclonal antibody (predicted molecular weight 64.4 kDa) and the bottom strip was probed with anti-Tim17p rabbit polyclonal antibody (predicted molecular weight 16.6 kDa). Molecular weight markers on the left-hand side of the blot from top to bottom are 220, 160, 120, 100, 70, 60, 50, 20, and 15 kDa. 14 μ g of sucrose gradient purified mitochondrial membranes of each fraction was loaded.



restriction endonuclease digestion, and reaction buffer/hexanucleotide mix from the Roche kit in the presence of aphidicolin, known to inhibit four nuclear DNA polymerases (Pol1p, Pol2p, Pol3p, and Pol5p, Budd *et al.*, 1989, Shimizu *et al.*, 2002). Following the polymerase reactions, samples were spotted and DIG-labelled DNA was detected with anti-DIG antibody and CDP-Star. A minimal amount of background DNA synthesis was obtained with control mitochondrial membrane proteins isolated from the untransformed expression strain S150 (panel M, Figure 4.8). From 5 to 30 minutes this minimal amount of signal increased in amount; at 30 minutes the background signal approximately doubled at 30°C relative to the initial signal. At 37°C, synthesis increased until the 10 or 15 minute time point, after which the signal remained constant, or decreased by less than 20%. The synthesis at each time point was subtracted as background from the corresponding time points of experiments using membrane fractions containing overexpressed polymerases. DNA synthesis was barely detectable in fractions obtained from yeast cultures expressing Mip1p[Σ]Δ279, Mip1p[Σ]Δ222, and Mip1p[Σ]Δ216 (Figure 4.8). Transformants expressing Mip1p[Σ]Δ205, Mip1p[Σ]Δ175, Mip1p[Σ], Mip1p[S]Δ205, and Mip1p[S] contained mitochondrial membrane protein fractions that incorporated DIG-11-dUTP into DNA and the resulting chemiluminescence signals were significantly different from those of the control mitochondrial membrane proteins ($p < 0.03$ from a Student's t-Test for the ten minute time point at 30/37°C for all five transformants expressing the aforementioned polymerases compared to the control membrane proteins, except for Mip1p[S] at 30°C, for which $p = 0.056$, Figure 4.8, Figure 4.9 and Table 4.3). When template DNA was not added to a reaction containing the Mip1p[Σ] fraction, no detectable signal was observed, indicating that contaminating

Figure 4.8 Dot blot of DIG-11-dUTP incorporated into DNA by mitochondrial membrane protein fractions. Cells expressing Mip1p[S] and Mip1p[Σ] wild-type and truncation variants were examined for the ability of their mitochondrial membrane protein fractions to synthesize DNA. DNA polymerase activity was monitored over 30 minutes at 30 and 37°C for each strain (indicated on the left side of the blot). At 5, 10, 15, and 30 minutes of the experiment an aliquot was removed and the reaction was stopped. 1 μl of the reaction was spotted onto the nylon membrane in triplicate. M, mitochondrial membrane proteins from untransformed S150. Aberrant spots are boxed and were not used for calculating relative light units.

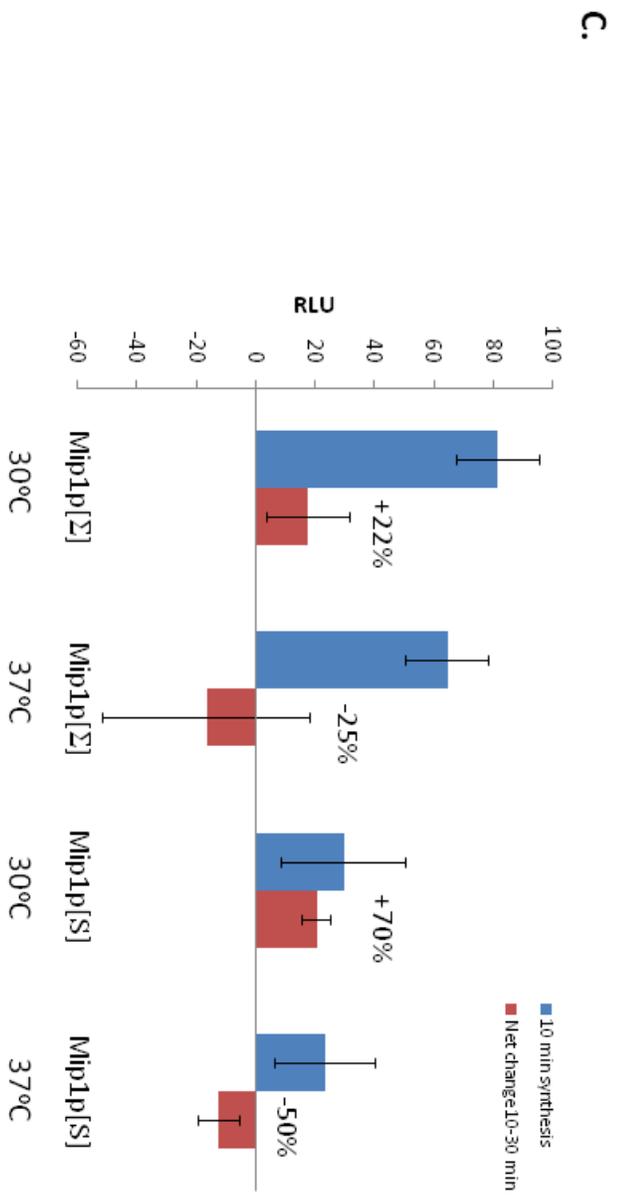
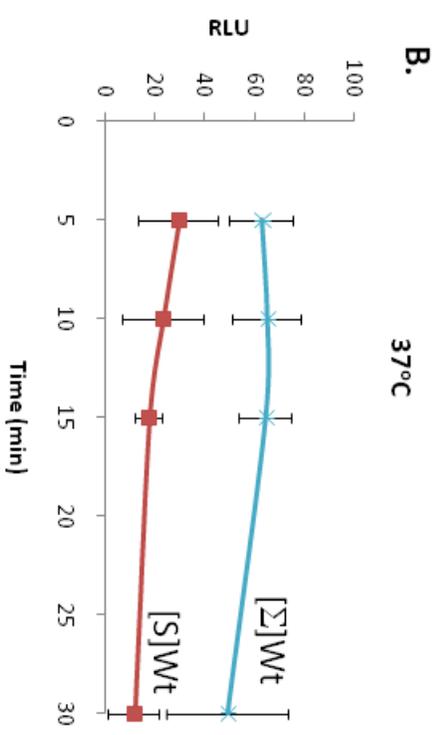
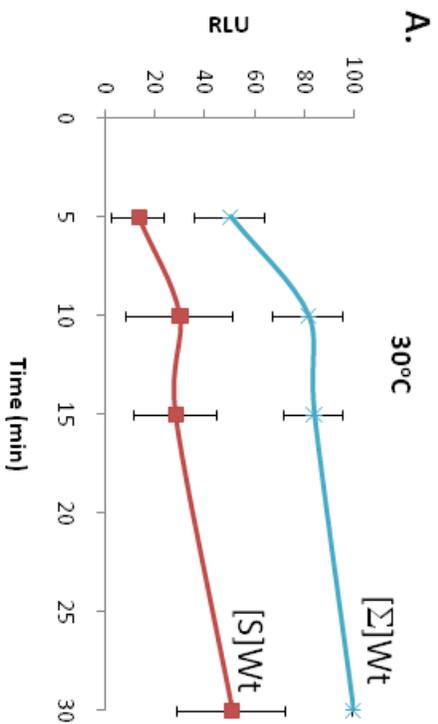


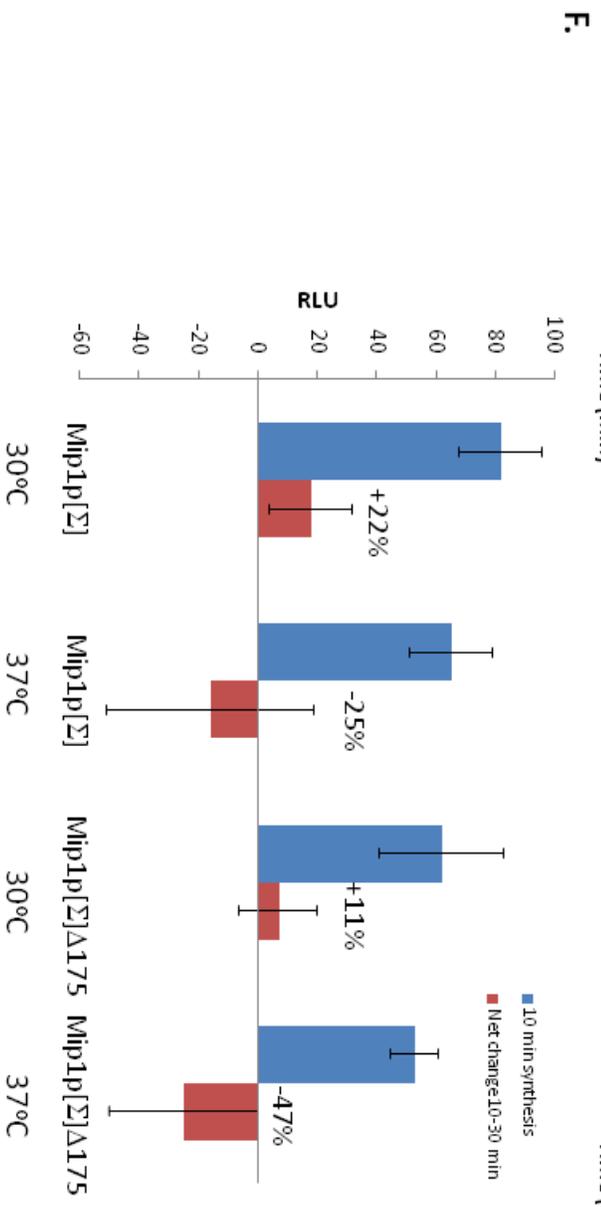
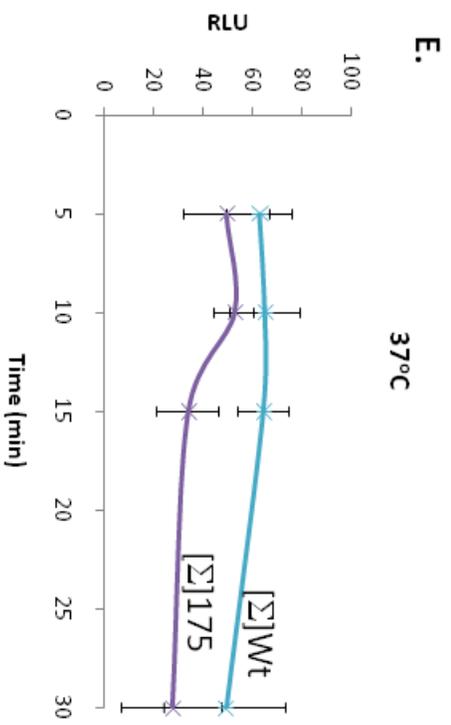
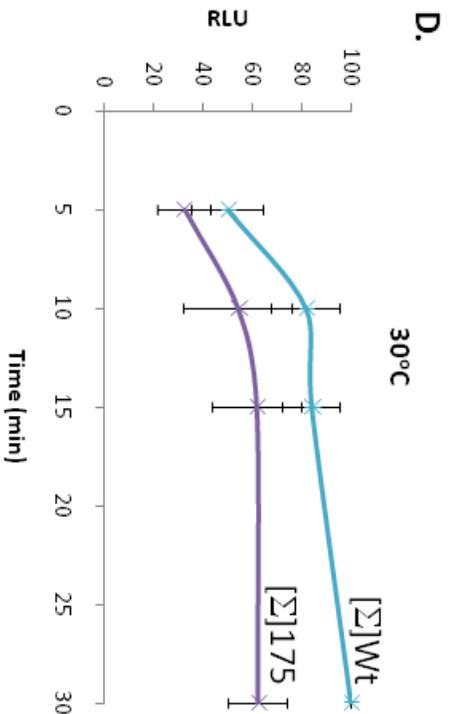
mtDNA is insignificant (data not shown). Likewise, if the Klenow fragment of *E. coli* DNA polymerase I or the Mip1p[Σ]-containing fraction was not added to the reaction, DIG-11-dUTP-labelled DNA was not detected on the dot blot (data not shown).

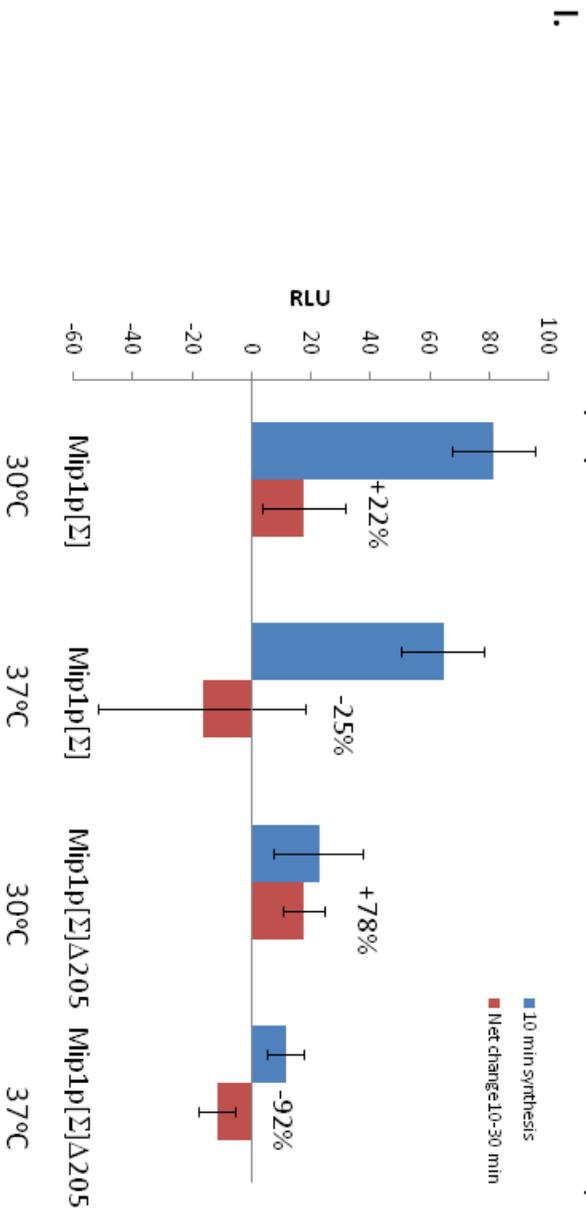
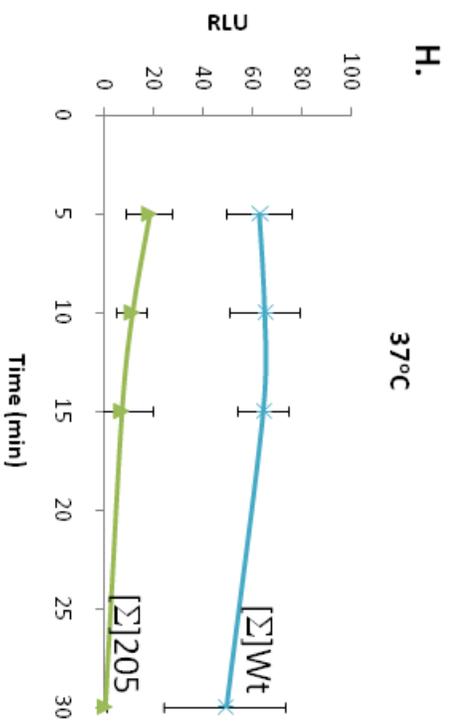
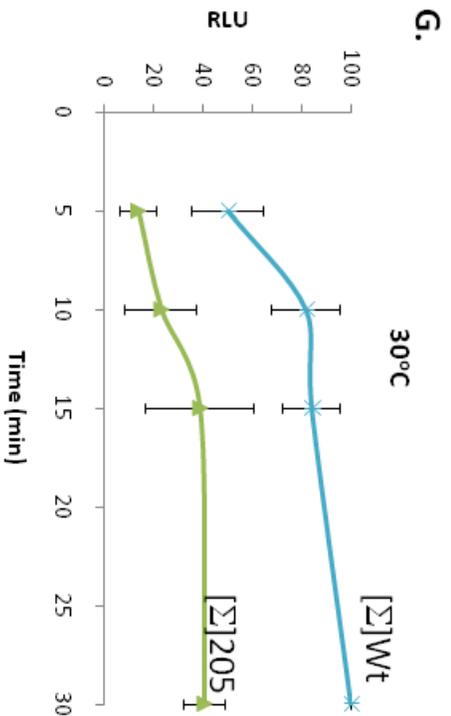
The total amount of DIG-labelled DNA obtained for each isolated mitochondrial membrane fraction in three different experiments (trials 1, 2, and 3) at 5, 10, 15 and 30 minutes, and at 30 and 37°C, is presented as relative light units (RLU) in five separate panels in Figure 4.9. The reaction containing the Mip1p[Σ] membrane fraction at 30 minutes and at 30°C, produced the most DIG-labelled DNA, as the most intense chemiluminescent signal from the anti-DIG antibody was detected in all trials at this temperature and time point. Therefore, for each experiment this signal was set to 100 RLU. All other time points for all polymerase variants are presented relative to this signal (Figure 4.9). The relative RLU data for individual trials are presented in Tables 4.4 and Table 4.5.

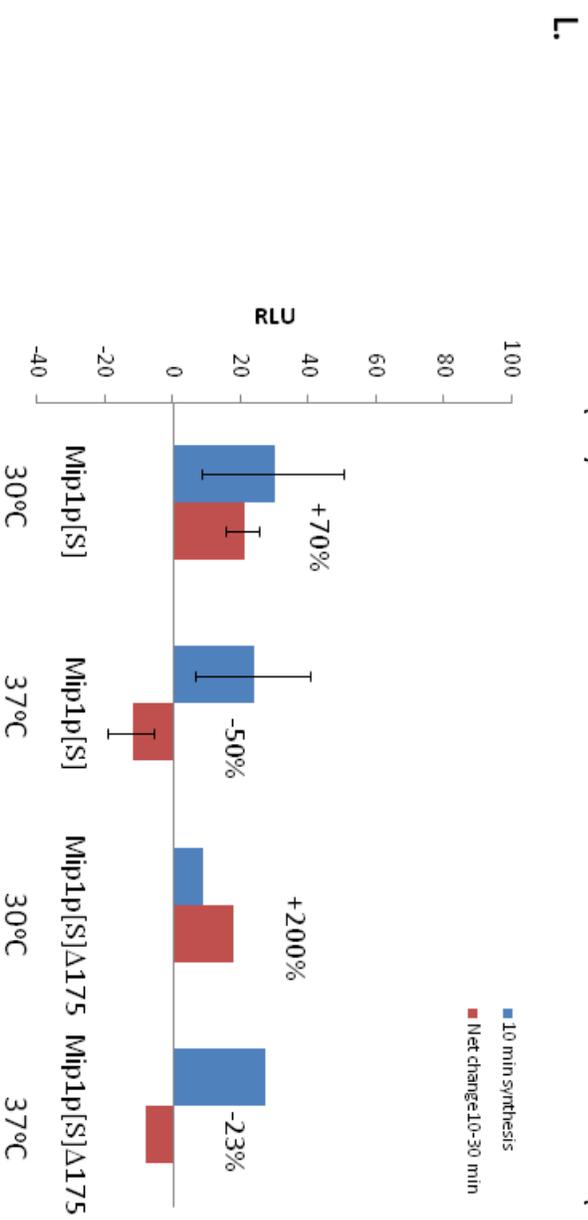
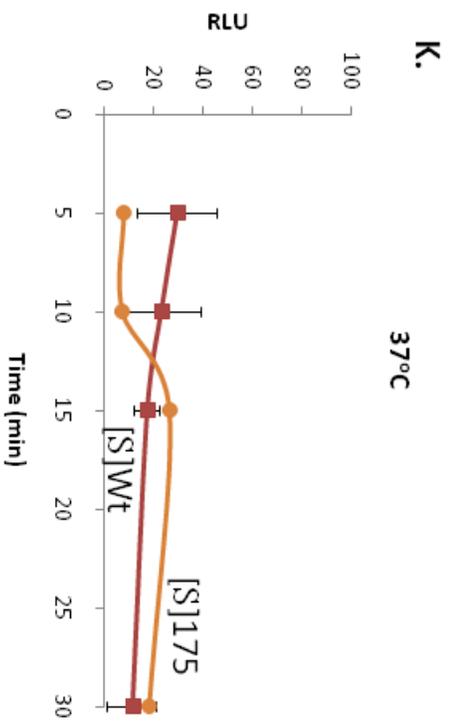
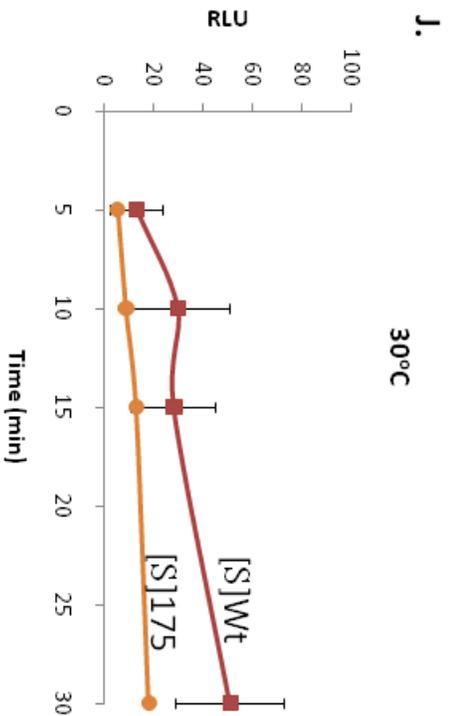
At 30°C the Mip1p[Σ] and Mip1p[S] polymerase fractions produced steady increases in RLUs over 30 minutes (Figure 4.9A. and C., Table 4.4, and Table 4.5). The net change in RLUs from 10 to 30 minutes was 22% and 70% for Mip1p[Σ] and Mip1p[S] respectively, indicating continuous DNA synthesis during the course of the experiment for both of these fractions (Figure 4.9 C). However, on average the incorporation rate for Mip1p[S] was approximately one third of that for Mip1p[Σ] at the beginning of the experiment (5 to 15 minutes) and about half by the end of the experiment (Figure 4.9A and B).

Figure 4.9 Quantitation of recombinant mtDNA polymerase activity at 30 and 37°C. Chemiluminescence obtained from DIG-11-dUTP-DNA polymerase dot blots was quantitated with an Alpha Innotech FluorChem 8900. [S]205, [S]175, [S]Wt, [Σ]205, [Σ]175, and [Σ]Wt are fractions containing Mip1p[S]Δ205, Mip1p[S]Δ175, Mip1p[S], Mip1p[Σ]Δ205, Mip1p[Σ]Δ175, and Mip1p[Σ] respectively. DNA synthesis for transformants expressing Mip1p[Σ]Δ205, Mip1p[Σ]Δ175, Mip1p[S]Δ205, Mip1p[S]Δ175, and Mip1p[S] was quantitated relative to that from the transformant expressing Mip1p[Σ], which incorporated the largest amount of DIG-11-dUTP into DNA in all trials. Mip1p[Σ] DNA synthesized at 30 minutes and 30°C was set to 100 relative light units (RLU). The amount of DNA synthesized by each membrane fraction was also normalized to the level of Mip1p[Σ], because different levels of expression were observed by Western blot analysis (Figure 4.7). The experiment was repeated three times, with the exception of Mip1p[S]Δ175, and averages are presented as bars and standard deviations as thin black lines. Aberrant spots, such as those boxed in Figure 4.8 were not used in the analysis. Net change (red bars) is defined as the difference in RLUs from 30 minutes to 10 minutes and the percent net change value is indicated above the bar. Percent change is defined as the average loss or gain of RLUs after 30 minutes divided by the average RLUs for the 10 minute time point multiplied by 100. Positive values indicate DNA synthesis while negative values indicate DNA degradation.









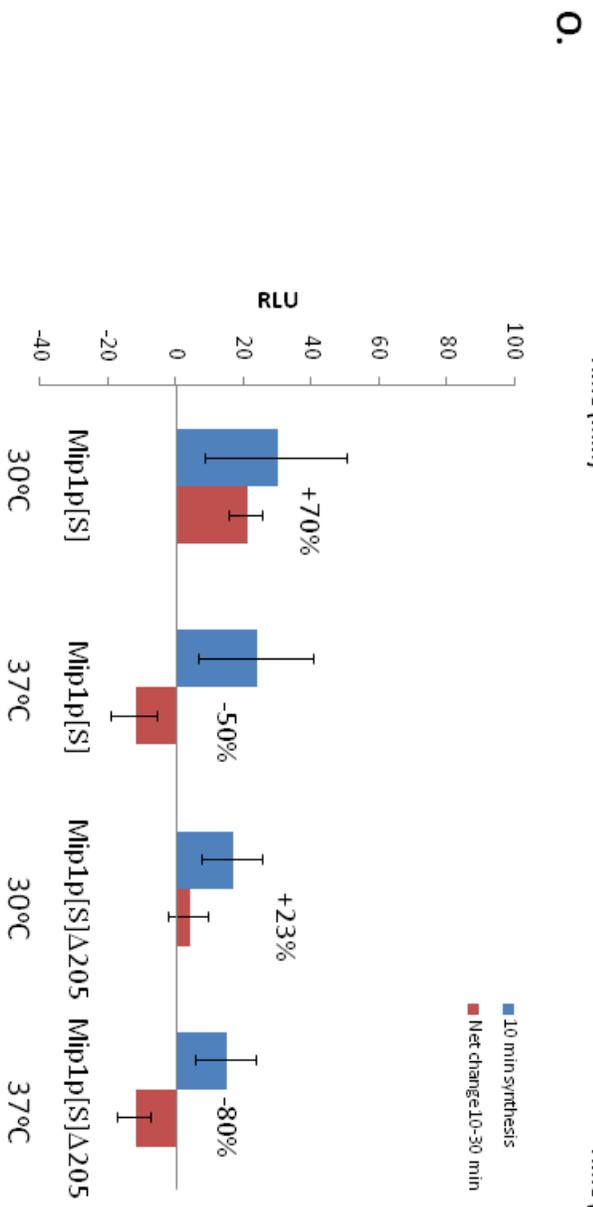
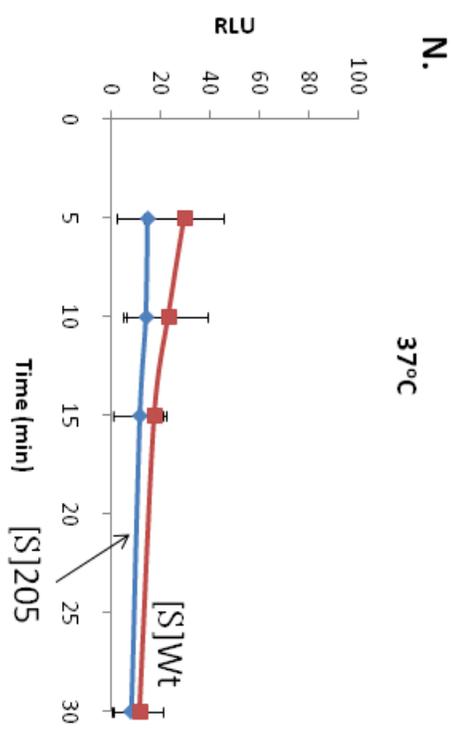
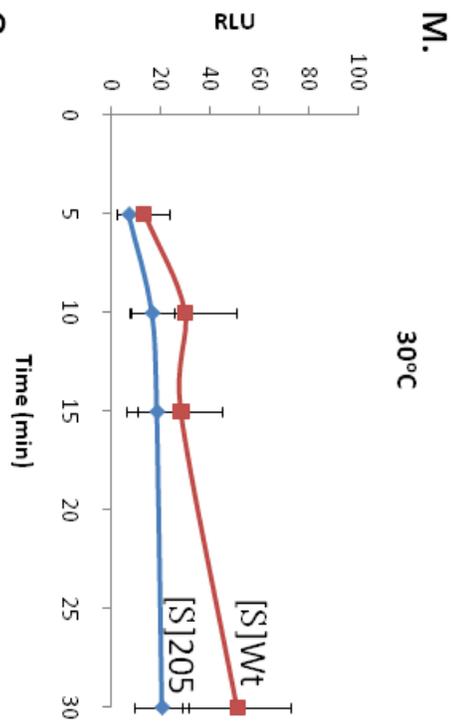


Table 4.3 Student's t-Test scores for variant membrane fractions at 10 minutes and 30°C compared to negative control.

Mip1p Variant	Average^a	Variance	P-value^b
Membranes	142	2	n/a ^c
Mip1p[Σ]Δ279	112	1	3 x 10 ^{-5d}
Mip1p[S]Δ279	148	112	0.4
Mip1p[Σ]Δ222	115	1	5 x 10 ^{-5d}
Mip1p[S]Δ222	120	1	1 x 10 ^{-6d}
Mip1p[Σ]Δ216	121	4	0 ^d
Mip1p[S]Δ216	n/d ^e		
Mip1p[Σ]Δ205	170	79	0.03
Mip1p[S]Δ205	230	109	0.004
Mip1p[Σ]Δ175	257	28	0
Mip1p[S]Δ175	n/d		
Mip1p[Σ]	526	6475	0.01
Mip1p[S]	240	1764	0.06

^a average integrative density values (/1000) measured for two or three individual spots from the same sample

^b calculated using the t-Test function of Excel 2007, comparing the data from the negative control membrane only fraction with that of each variant membrane fraction

^c n/a not applicable

^d t-Test score was less than 0.03, but the signal was weaker than that of the membranes alone

^e n/d not determined

Table 4.4 RLU production from three separate trials of the Mip1p[Σ] and sigma-variant polymerase assays. RLUs presented are relative to the Mip1p[Σ]-containing fraction at 30 minutes and 30°C, which was set to 100%.

Time (min)	[Σ]Wt		[Σ]175		[Σ]205	
	30°C	37°C	30°C	37°C	30°C	37°C
5	35, 52, 64	51, 77, 62	21, 41, 37	30, 62, 58	6, 19, 19	10, 29, 18
10	66, 86, 94	58, 57, 82	32, 58, 75	44, 58, 58	10, 39, 22	5, 14, 17
15	73, 84, 96	58, 77, 60	45, 61, 81	23, 48, 33	29, 64, 24	0, 0, 23
30	100, 100, 100	77, 42, 30	49, 71, 67	37, 43, 4.9	32, 49, 43	0, 0, 2

Table 4.5 RLU production from three separate trials of the Mip1p[S] and S288c-variant polymerase assays. RLUs presented are relative to the Mip1p[Σ]-containing fraction at 30 minutes and 30°C, which was set to 100%.

Time (min)	[S]Wt		[S]175*		[S]205	
	30°C	37°C	30°C	37°C	30°C	37°C
5	12, 25, 4	18, 48, 24	6	8	5, 13, 5	10, 29, 7
10	31, 51, 8	21, 41, 9	9	8	17, 26, 8	13, 24, 7
15	25, 47, 14	15, 24, 15	13	27	13, 33, 12	4, 9, 23
30	47, 75, 32	7, 24, 6	18	19	14, 34, 16	2, 7, 16

*The Mip1p[S]Δ175 was performed once

Activities of truncated polymerases were then compared to those of the corresponding full length enzyme. At 30°C Mip1p[Σ] Δ 175-containing chemiluminescent spots produced fewer RLUs on average compared to Mip1p[Σ]; however, based on the overlapping of standard deviations during the initial 15 minutes of the experiment, the difference between these two enzymes is not statistically significant. At later time points the net difference in DNA production was greater in Mip1p[Σ] *i.e.* Mip1p[Σ] Δ 175 DNA synthesis levelled off around 15 minutes (Figure 4.9 D). On average the Mip1p[Σ] Δ 175 fraction at 30°C synthesized at least 1.5 times more RLUs compared to Mip1p[Σ] Δ 205 at all time points (Table 4.4). Similarly, synthesis by the Mip1p[Σ] Δ 205-containing fraction also levelled off by 15 minutes and at least 2.2-fold less DIG-labelled DNA was synthesized in comparison to Mip1p[Σ] at all time points (Figure 4.9 G).

Similar to the results obtained from the Σ 1278b truncation variant polymerases, on average the S288c Mip1p Δ 175 and Mip1p Δ 205 truncation variants at 30°C synthesized fewer RLUs in comparison to the Mip1p[S] wild-type (Figure 4.9 J – O). However, because Mip1p[S] Δ 175 was only examined in one experiment and the differences seen between Mip1p[S] and Mip1p[S] Δ 205 are not statistically significant, this observation is not definitive. On average all Σ 1278b and S288c truncation variant polymerase fractions produced net positive changes in RLUs from 10 to 30 minutes at 30°C, indicating that DNA was maintained and/or synthesized at this temperature during the course of these experiments (Figure 4.9 F, I, L, and O).

Next, we examined the membrane fractions containing wild-type Mip1p[Σ] and wild-type Mip1p[S] at 37°C to determine if Mip1p[S] polymerase activity is related to

mtDNA loss that has been observed in *MIP1[S]* cells (Chapter 2 and Baruffini *et al.*, 2007b) and truncation variants *in vivo* (Chapter 3).

At 5 minutes Mip1p[Σ] initially synthesized more DIG-labelled DNA at 37°C in comparison to at 30°C; however, after 5 minutes DNA synthesis reached a plateau and by 30 minutes decreased slightly, with the exception of trial 1, in which there was a continuous increase in RLU (Figure 4.9 B and Table 4.4). The average net change in RLUs from 10 to 30 minutes was -25% of that synthesized in the first 10 minutes indicating DNA loss at 37°C (Figure 4.9 C).

The Mip1p[S] version is half as active as the Mip1p[Σ] after 5 minutes at 37°C on average, but appears almost twice as active as at 30°C possibly due to more rapid heating of the sample (Figure 4.9 A and B). In both Mip1p[Σ] and Mip1p[S], DNA synthesis appeared to stop after 5 minutes at 37°C but a more dramatic loss of RLUs was observed between 5 and 30 minutes with Mip1p[S] (Figure 4.9 B and C). By the end of the experiment approximately four-times fewer RLUs were observed for Mip1p[S] in comparison to Mip1p[Σ]. While the RLU loss was similar for both reactions the average net loss of RLUs from 10 to 30 minutes of the Mip1p[S] reaction was higher at -50% of that synthesized in the first 10 minutes (Figure 4.9 C).

Reactions containing the two CTE-truncation variants, Mip1p[Σ] Δ 175 and Mip1p[Σ] Δ 205, produced fewer RLUs on average relative to that containing Mip1p[Σ] at all time points at 37°C (Figure 4.9 E and H). However, during the initial time points of the experiment the difference between Mip1p[Σ] Δ 175 and Mip1p[Σ] was not statistically significant (Figure 4.9 E). Nonetheless, it was interesting that at 30 minutes the fraction of DNA that was lost was nearly twice as high in the Mip1p[Σ] Δ 175 reaction compared

to the wild-type (Figure 4.9 F). Mip1p[Σ] Δ 205-containing fractions produced only approximately 30% of the maximal RLUs synthesized by Mip1p[Σ] at 37°C during the first 5 minutes (Figure 4.9 H). Following this time point, the signal decreased to zero or almost zero, indicating an almost complete loss of DIG-labelled DNA (-92%) by the end of the experiment (Figure 4.9 I).

Truncated polymerases of S288c-origin at 37°C produced very weak signals that were not statistically different from those of Mip1p[S] wild-type.

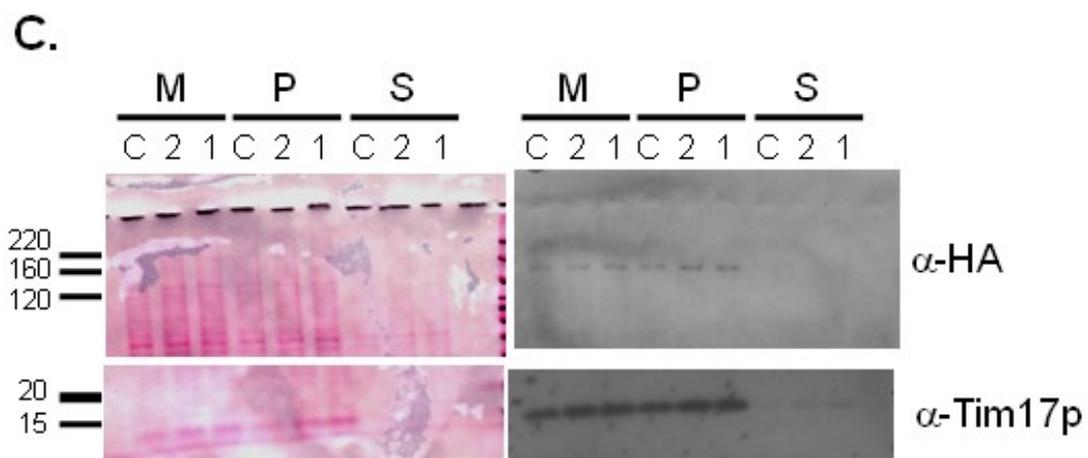
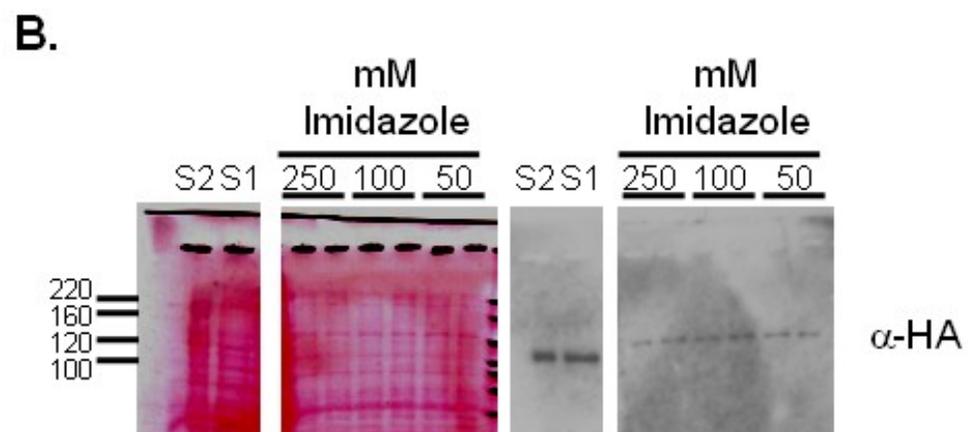
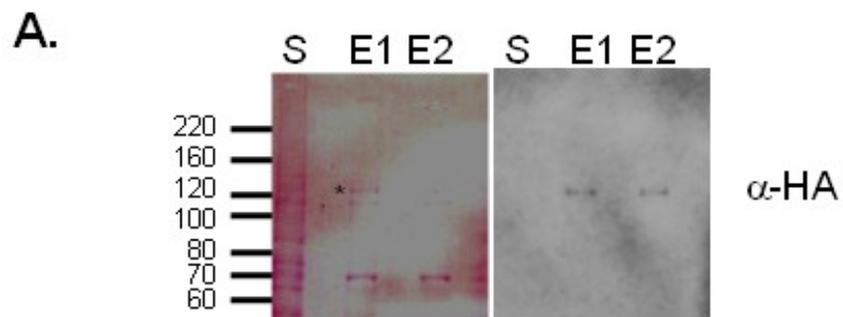
4.4.5 An attempt to purify Mip1p[S] Δ 175 and Mip1p[S] by Nickel-NTA chromatography

More precise assessments of polymerase and exonuclease activities could be carried out with a purified enzyme. Therefore, mitochondria harbouring Mip1p[S] Δ 175 were treated under denaturing and nondenaturing conditions in an attempt to purify the polymerase via the C-terminal 6xHis Tag (Figure 4.10, A and B respectively). Following Nickel-NTA (Ni-NTA) purification under denaturing conditions, a faint band approximately the correct molecular weight of Mip1p[S] Δ 175, was visible in the lane containing elution 1 on the Ponceau S stained membrane (Figure 4.10 A). This band was in fact Mip1p[S] Δ 175 as indicated by Western blotting. Mip1p[S] Δ 175 was also detected in elution 2 of the Western blot even though it was not visible after Ponceau S staining (Figure 4.10 A). The eluted Mip1p[S] Δ 175 was not pure, as indicated by the presence of a clear 70-kDa band on the Ponceau S-stained blot. The 70-kDa protein may represent a Mip1p binding protein which remained tightly associated in the presence of 8 M urea, or a histidine-rich protein that co-purified on the Ni-NTA resin.

In an attempt to obtain purified catalytically active polymerase, Mip1p[S] and Mip1p[S] Δ 175 mitochondria were subjected to non-denaturing conditions followed by Ni-NTA chromatography. On the blot of mitochondria treated under nondenaturing conditions following Ni-NTA purification, a faint polymerase band was detected in elutions with different imidazole concentrations (Figure 4.10 B right panel). However, much more protein was visible on the Ponceau S stained blot (left panel) of the eluted samples. The pattern of protein banding was similar following all elutions, indicating that Mip1p[S] Δ 175 was not enriched under these conditions. A significant amount of Mip1p[S] Δ 175 was visualized in the supernatant fractions indicating a lack of binding of Mip1p[S] Δ 175 to Ni-NTA (S1 and S2, Figure 4.10 B). Perhaps intact mitochondria were preventing access of Ni-NTA to either soluble or membrane associated tagged Mip1p. Therefore, in a final attempt to purify Mip1p[S] under nondenaturing conditions, homogenized mitochondria were fractionated into pellet (P) and supernatant (S) fractions. These fractions as well as total homogenized mitochondria (M) were then subjected to Ni-NTA chromatography (Figure 4.10 C). No successful enrichment of polymerase was obtained from any of the non-denaturing fractions. Contaminating proteins in total homogenized mitochondria and pellet purifications included Tim17p, as indicated by immunodetection (Figure 4.10 C), suggesting that mitochondrial membranes were trapped in the slurry of nickel-NTA resin.

Figure 4.10 An attempt to purify Mip1p[S] Δ 175 and Mip1p[S] by Nickel-NTA chromatography. For each panel 12 μ l of sample was loaded into each lane of a 9% SDS-PAGE gel followed by Western blot analysis. Ponceau S stained blots are shown to the left of the Western blot in each panel. **A.** Results of Ni-NTA chromatography of Mip1p[S] Δ 175 under denaturing conditions. Following homogenization mitochondria were incubated with Nickel-NTA Sepharose resin. After collection of the resin by centrifugation, the supernatant (S) was saved for analysis and bound proteins removed by two consecutive elutions with buffer C/100 mM EDTA (E1 and E2). All three fractions were analyzed by SDS-PAGE and immunoblotting. The blot was immunodetected with anti-HA antibody. Images of X-ray film following detection with CDP-Star are shown. A faint band on the Ponceau S stained blot of approximately 122 kDa, the size expected for Mip1p[S] Δ 175, is marked by an asterisk. **B.** Results of Ni-NTA chromatography of Mip1p[S] Δ 175 under non-denaturing conditions. Immunodetection was carried out exactly as in part A. After the homogenization step the entire sample of mitochondria was processed as described in Materials and Methods. S1, is supernatant that was saved following the one hour incubation and removal of resin by centrifugation. S2, is supernatant obtained after the first TBG wash. Elutions performed in duplicate on the same tube, with either 250, 100, or 50 mM imidazole are indicated on the right-hand-side of the blot. **C.** Results of Ni-NTA chromatography of Mip1p[S] under non-denaturing conditions. Homogenized mitochondria (M), a pellet fraction (P) and a supernatant fraction (S) of homogenized mitochondria were subjected to Ni-NTA chromatography. Samples were eluted in the following order: (1) TBG/100 mM imidazole, (2) TBG/250 mM imidazole, and (C) buffer C/100 mM imidazole. The blot was cut into two pieces

between 40 and 50 kDa of the BenchMark ladderTM. Immunodetection of the top piece was carried out with anti-HA antibody and anti-Tim17p antibody for the bottom section. Molecular weight markers in kDa are indicated on the left-hand-side of all blots.



4.5 Discussion

The first goal of this work was to determine the shortest targeting signal to allow import of Mip1p into mitochondria. It was not surprising that Mip1p[Short]-yEGFP colocalized with mitochondrial dsRed as Foury has shown that transcription occurs downstream of the first ATG of the *MIP1* gene (Figure 4.4, Foury, 1989). Interestingly, Mip1p[Long]-yEGFP also colocalized with dsRed. However, it could not be ruled out that transcription from *GAL1p* results in a heterogeneous population of transcriptional start sites, which could lead to expression of a short-form *MIP1* transcript encoding the 1254-amino acid “short” Mip1p. To prevent this possibility, a polymerase, Mip1p[LongM→I]-yEGFP, was engineered that has site-directed mutations of the 2nd, 3rd, 4th, and 5th amino-terminal methionine codons to ones for isoleucine. The finding that Mip1p[LongM→I]-yEGFP was targeted to the mitochondria was also unexpected as MitoProt II calculated a low probability of 0.2294 for targeting of both this form and the 1280-residue form potentially encoded by *MIP1*. In addition, the program does not predict a mitochondrial processing peptidase cleavage site for either of these long pre-proteins (Claros and Vincens, 1996). However, it has been observed that approximately one quarter of randomly-generated sequences, encoding nine amino-acid residues, inserted at the 5'-end of the yeast *COX4* gene lacking its own mitochondrial targeting signal, correctly target Cox4p to the mitochondria (Lemire *et al.*, 1989). Due to the importance of positive charge in the signal sequence and because long side-chains of lysine and arginine residues may project from the helical surface at unusual angles, sequences predicted to be non-amphiphilic, and therefore not adhering to the “amphiphilic helix model” of a targeting signal, may in fact be amphiphilic (Lemire *et al.*, 1989). This could explain why the Mip1p[LongM→I]-yEGFP polymerase is targeted

to the mitochondria. This polymerase harbours five lysine and twelve arginine residues within the first 67 N-terminal residues (Figure 4.1 B).

In the experiment shown in Figure 4.4, cytosolic staining appears to be increased in the Mip1p[LongM→I]-yEGFP transformant (Figure 4.4). Therefore, the efficiency of mitochondrial import of this variant may be reduced. Future experiments could examine the import of Mip1p[LongM→I]-yEGFP expressed from the endogenous *MIP1* promoter followed by more precise calculations of relative mitochondrial and cytosolic fluorescence. Additionally, overexpressed GFP-tagged versions could be analyzed by Western blotting using anti-GFP antibody, to quantitate the relative amounts of Mip1p[LongM→I]-yEGFP in mitochondria isolated by subcellular fractionation. Finally, the ability of the three N-terminal variants to complement deletions of wild-type *MIP1* should be considered.

The second goal was to obtain partially-purified overexpressed enzyme from yeast to assess polymerase activity. Although the finding that the long form of Mip1p was imported into mitochondria was interesting, we engineered our expression plasmids to start transcription at the second start codon, as defined by Foury (Foury, 1989). The sub-mitochondrial localization of Mip1p was then investigated (Figure 4.5). In sucrose gradient membrane fractions we estimated 60% and 80% lysis of Mip1p[Σ] and Mip1p[S] mitochondria subjected to hypotonic treatment and homogenization, based on the release of Put2p from the mitochondrial matrix (Figure 4.5). Overexpressed polymerases of these forms were always localized in the Tim17p-containing pellet fractions and were either barely detectable or absent in the Put2p-containing supernatant fractions (Figure 4.5). Likewise when mitochondria from yeast overexpressing

polymerase variants were subjected to hypotonic treatment, homogenization, and sucrose gradient ultracentrifugation, a strong association of the polymerases was still observed with mitochondrial membranes (Figure 4.7).

Several additional treatments were used to attempt to release Mip1p from the membrane fraction. Mitochondria from yeast expressing Mip1p[S] were resuspended in sodium carbonate, which resulted in more complete mitochondrial lysis, as indicated by the presence of Hep1p only in the supernatant fraction (Figure 4.6). Under these conditions, Mip1p[S] was found mostly in the pellet fraction, which also contained the majority of Tim17p. The low level of Mip1p[S] observed in the supernatant fraction of the sodium carbonate treatment suggests that a possible protein-protein interaction exists between a subset of Mip1p molecules and a yet unidentified partner. This low level of Mip1p could represent a fraction of polymerase that was released from the mitochondrial protein import machinery during or just after translocation into the mitochondria. Due to the low level of Tim17p detected in the supernatant fraction, the second scenario is the most likely explanation. Other future experiments will test the efficiency of mitochondrial lysis via sodium carbonate using the stronger Put2p antibody, as the Hep1p antibody gives a very weak chemiluminescent signal and high amounts of background when tested at different dilutions using different stringency washes (data not shown). The Put2p antibody would allow for quantitation of the matrix-associated signal. In addition, rather than TritonX-100, other non-ionic/non-denaturing detergents such as Nonidet P-40, N-octylglucoside, and Tween-20 could be tested for the ability to disrupt mitochondria and/or solubilize Mip1p (Longley *et al.*, 1998).

The consistent localization of Mip1p, and variants thereof, in mitochondrial membrane-containing pellet fractions, agrees with a model of a TMS replisome where Mip1p is associated with the mitochondrial inner membrane (Meeusen *et al.*, 2004). Genetic and cytological evidence suggest mtDNA inheritance is non-random and nucleoids are distributed somewhat regularly within the mitochondria, presumably by attachment to the inner membrane (Meeusen *et al.*, 2004). If Mip1p is not directly or indirectly associated with the membrane, but is associated with the nucleoid, then perhaps removing the mtDNA with DNaseI could release Mip1p into a more soluble fraction. Future studies will investigate this possibility.

Mip1p-containing membrane protein fractions were capable of DNA synthesis *in vitro* as detected with the non-radioactive assay developed in this work. In fact, this is the first evidence that a mitochondrial DNA polymerase, Mip1p, can utilize DIG-11-dUTP as a substrate and incorporate it into DNA. Initially at 5 minutes all polymerases synthesized more DNA on average at 37°C than at 30°C (Figure 4.9, Table 4.4, and Table 4.5). This initial burst of DNA synthesis likely reflects that 37°C samples achieved an optimal synthesis temperature faster than at 30°C. However, after this point, there was a steady decrease in the total amount of DIG-labelled DNA, with the exception of Mip1p[S] Δ 175; however, this polymerase was only examined once (Figure 4.9 J, K and L). This assay should be repeated with pre-warmed reaction mixtures to confirm this prediction.

With this assay, it was shown that the Mip1p[S]-containing membrane fractions do not replicate DNA as well as the Mip1p[Σ]-containing fractions at 30 or 37°C (Figure 4.9). This observation agrees with results in the literature that yeast strains expressing Mip1p[S] generate a high frequency of *petites* and mtDNA deletions in comparison to

Mip1p[Σ] strains at 37°C (Baruffini *et al.*, 2007b). Interestingly, greater than half of the maximal RLUs produced by Mip1p[S] were lost at 37°C, suggesting that the DIG-labelled DNA was degraded. There was no significant decrease in chemiluminescent signal detected after 30 minutes from mitochondrial membrane protein fractions obtained from untransformed cells at 37°C (Figure 4.8). This observation suggests that the average decrease observed for all polymerase fractions at 37°C may be due to the exonuclease activity of these enzymes and not due to a contaminating nuclease. This hypothesis is further supported by the fact that all of the enzymes at 30°C do not show a decrease in RLUs; they continue to synthesize DNA or a plateau is observed (Figure 4.9). Therefore, the deletions in mtDNA observed by Baruffini *et al.* (2007b) and in the YPH499 strain (Chapter 2) may result from excess Mip1p exonuclease activity.

This difference observed at 37°C between Mip1p[S] and Mip1p[Σ] could represent relative thermal instability of Mip1p[S] whereby the enzyme can initially elongate the primer-template complex. With time, the polymerase domain becomes disordered and only the proofreading domain retains activity and degrades the nascent DNA. A related result has been observed in a study by Luo and Kaguni (Luo and Kaguni, 2005), where a variant of *Drosophila* polymerase-gamma had three residues, SYW, replaced by AAA, in the γ 4 region of the polymerase (Figure 4.1). This polymerase had barely detectable levels of DNA polymerase activity (<1%) and had a lower DNA binding activity (approximately one quarter of wild-type), but it exhibited a three-fold increase in proofreading activity compared to the wild-type (Luo and Kaguni, 2005). This type of scenario could explain the increase in *rho*^o cells harbouring the alanine 661 Mip1p[S] version seen in YPH499 (Chapter 2) and by Baruffini *et al.*, 2007b. In this

scenario, the Mip1p[S] polymerase domain would eventually become disordered at 37°C, but the enzyme could still bind to the template and the exonuclease domain could degrade mtDNA. Alternatively, proofreading activity may function at an increased rate relative to DNA polymerization, therefore, causing a net decrease in labelled DNA. Either way, the presence of the threonine 661 residue in Mip1p[Σ] may be the contributing factor to stability of enzyme folding, stability of polymerase-domain interactions with DNA or a combination of these two processes.

Interestingly, the yeast threonine 661 residue present in the Mip1p[Σ] version, is conserved in fungal as well as human, chicken, mouse, frog and fly mtDNA polymerases (Figure 4.11). Upon closer examination, the twenty-three fungi including ascomycetes and basidiomycetes shown in Figure 3.2, also contain this threonine residue (data not shown). The bacterial DNA polymerases shown in Figure 4.11 all have an acidic residue at this position. An aspartate is present in the Klenow fragment of *E. coli* DNA polymerase I and in *Taq* DNA polymerase, while *Bacillus stearotherophilus* DNA polymerase I contains a glutamate residue. All of these bacterial enzymes shown in this alignment have been crystallized with double-stranded DNA and in fact the *Bacillus* DNA polymerase I crystal is catalytically active (Beese *et al.*, 1993, Eom *et al.*, 1996, Kiefer *et al.*, 1998). Within the structure of Klenow bound to duplex DNA this aspartate residue (D673) is located at the end of a β-strand within a cleft formed between the polymerase and exonuclease domains, and is 5 Å from the phosphodiester backbone chain B of the duplex DNA (See Figure 4.12). The DNA binding site is exclusive of the

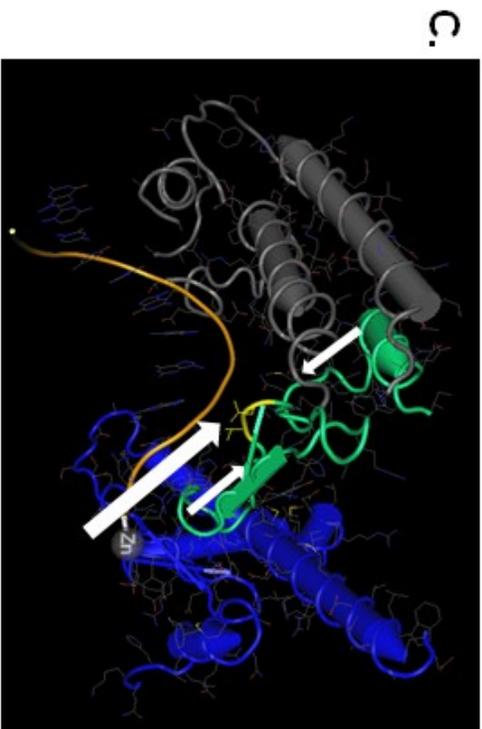
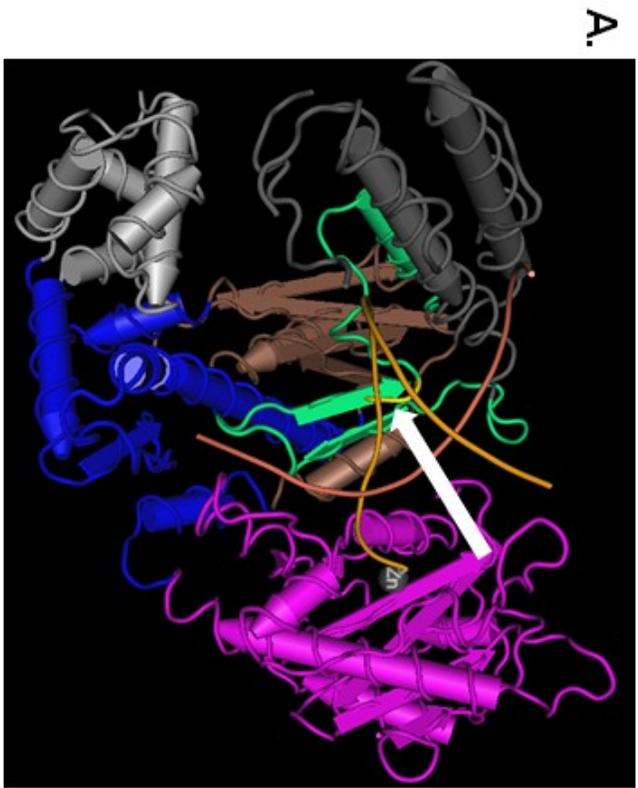
Figure 4.11 ClustalX (Thompson *et al.*, 1997) alignment of a portion of selected family A group polymerases. Five metazoan and four *Saccharomycetales* gamma subfamily mitochondrial DNA polymerases are shown as well as three bacterial polymerases. The alignment shown includes residues 640 to 754 of the *Saccharomyces cerevisiae* Mip1p[Σ] preprotein (1254 amino acids long). The metazoans include: *Mus musculus* (AAF82772, Mm), *Homo sapiens* (NP_001119603, Hs), *Gallus gallus* (Q92076, Ggal), *Xenopus laevis* (NP_001081464, Xl), and *Drosophila melanogaster* (Q27607, Dmel); the *Saccharomycetales* include: *S. cerevisiae* (conceptual translation of sequence obtained in this work, Sc), *Saccharomyces paradoxus* (AABY01000003, Spa), *Eremothecium gossypii* (AAS52277, Egos), and *Candida albicans* (XP_716738, Cal); the bacterial polymerases include: *Thermus aquaticus* DNA polymerase (P19821, Taq), *Bacillus stearothermophilus* DNA polymerase I (1NJW_A, Bste), and *Escherichia coli* DNA polymerase I Klenow fragment (1D8Y_A, Klenow). All NCBI accession numbers shown in parenthesis refer to proteins except for *S. paradoxus*, which is a nucleotide accession number. The arrow above the alignment indicates the *S. cerevisiae* 661 threonine residue, which is absolutely conserved in the mitochondrial DNA polymerases. Pol I indicates the DNA polymerase I domain defined by Ito and Braithwaite (Ito and Braithwaite, 1990). The alignment was visualized in GeneDoc V2.5.010 (Nicholas and Nicholas, 1997).

Poll

	940	*	960	*	980	*	1000	
Mm_AAF8277	YGATILPQVV	TAGTITRR	RAVEPTWLTASNAR	PPDRV	GSEILKAMVQAP	PG-YV	LVGADVDSQETWIAAVL	
Hs_NP_0011	YGATILPQVV	TAGTITRR	RAVEPTWLTASNAR	PPDRV	GSEILKAMVQAP	PG-YT	LVGADVDSQETWIAAVL	
Ggal_Q9207	YGATILPQVV	TAGTITRR	RAVEPTWLTASNAR	ADRV	GSEILKAMVQV	PG-YSL	VGADVDSQETWIAAVL	
X1_NP_0010	YGATILAQVV	SAGTITRR	RAVEPTWLTASNAR	ADRV	GSEILKAMVQV	PG-YHL	IGADVDSQETWIAAVL	
Dmel_Q2760	YGATICPQVV	ACGTLTR	RAMEPTWLTASNAR	SPDR	GSEILKSMVQAP	PG-YR	LVGADVDSQETWIASVL	
Sc_1254_12	LAIIPKIV	EMGTLTR	RAVENTWLTASNAR	KANR	GSEILKTQVKA	PG-YC	FVADVDSQETWIASLV	
Spa_AABY01	LAIIPKIV	EMGTLTR	RAVENTWLTASNAR	KANR	GSEILKTQVKA	PG-YC	FVADVDSQETWIASLV	
Egos_AAS52	VGAIIPKII	PMGTVTR	RAVENTWLTASNAR	KANR	GSEILKTKIVA	HEG-YSE	VGADVDSQETWIASLV	
Cal_XP_716	MGIIIPN	LATMGTITR	RATENTWLTASNAR	KNR	GSEILKSLVEA	KEG-YC	FVADVDSQETWIASLV	
Bste_INJW	TKRVHTI	FNCAITQ	GLR	LSSTEPNLQNIPIR	LEEGR	KIRCAFV	PESESDWLLIF	ADYSQIEIRVLAHI
Taq_P19821	TGRHTR	FNQTA	TALGRL	SSSDENLQNI	PVRTPLQ	QRIRRA	FAEEG-WLL	VALDYSQIEIRVLAHI
klenow_1D8	TGRVHTS	YHCAV	TALGRL	SSSTDPNLONI	PVRNEE	GRRIRCA	FAIED-YV	VSDYSQIEIRVLAHI

	*	1020	*	1040	*																																						
Mm_AAF8277	GDAHFA	GMHGCTAF	FGWMTLQGR	KRSRGTDL	HSKTAATVGI	REHAKIIFNYG	931																																				
Hs_NP_0011	GDAHFA	GMHGCTAF	FGWMTLQGR	KRSRGTDL	HSKTAATVGI	REHAKIIFNYG	952																																				
Ggal_Q9207	GEAHEA	GMHGCTAF	FGWMTLQGR	KSDGTD	DLHSKTAATVGI	REHAKIIFNYG	364																																				
X1_NP_0010	GEAHEA	GIHGCTAF	FGWMTLQGR	KSSGTD	DLHSKTAATVGI	REHAKIIFNYG	922																																				
Dmel_Q2760	GDAYAC	GEHGA	PLG	WMTLSGSKS	NGSMDH	SITAKAVGISRDHAKVINYA	870																																				
Sc_1254_12	GDSIEN	VHGGTA	IGMM	CLEGT	KNEGTDL	HTKTAQILGCSR	NEAKIIFNYG	754																																			
Spa_AABY01	GDSIEN	VHGGTA	IGMM	CLEGT	KNEGTDL	HTKTAQILGCSR	NEAKIIFNYG	755																																			
Egos_AAS52	GDSVEN	IHG	GTAIGYM	CLEGT	KDAGTDM	HTKTASILGCSR	NEAKIIFNYG	742																																			
Cal_XP_716	GDSMLQ	IHG	GTAIGW	M	TLEGEKSQ	KTDLH	SKTASILGISRN	DAKVENYG	828																																		
Bste_INJW	AE-----	DDN	IMEA	FR	RLD	DHTKT	AMDI	FQVSE	DEVTPN	MRQ	QAKAVN	HEG	415																														
Taq_P19821	SG-----	DEN	L	IR	V	FQ	GR	D	H	T	E	T	A	S	W	M	F	G	P	R	E	A	V	D	P	I	M	R	R	A	K	T	I	N	G	668							
klenow_1D8	SR-----	DK	G	L	L	T	A	F	A	E	G	K	D	I	H	R	A	T	A	A	E	V	F	G	L	P	L	E	T	V	T	S	E	O	R	S	A	K	A	I	N	G	440

Figure 4.12 Structure of *Escherichia coli* DNA polymerase I Klenow fragment (PDB structure accession number 1KLN) bound to duplex DNA highlighting the homologous Mip1p threonine-661 region and DNA bound in the 3'-5' exonuclease domain. α -helices are shown as cylinders and β -stands are shown as arrows. In each of panel the threonine (T672) and aspartate (D673, same position as Mip1p threonine-661 in the alignment shown in Figure 4.11) residues are highlighted in yellow and indicated by a white arrow. Two smaller white arrows indicate the S658 and M647 residues found in β 7 and α I respectively. **A.** Half-open “right hand” of DNA polymerase I. The polymerase domain consists of “fingers”, “thumb”, and “palm” subdomains. The fingers subdomain is comprised of predominantly the light gray and blue α -helices, the thumb subdomain consists of mostly two dark grey α -helices and the palm subdomain consists predominantly of brown β -strands and a brown α -helix, while the exonuclease domain is represented as purple α -helices and β -strands. The 3'-end of the orange DNA strand is located within the exonuclease domain near a zinc ion. The duplex DNA is bound in a cleft (containing T672 and D673) formed between the polymerase and exonuclease domains. The second DNA strand is coloured red. **B.** Rotation of DNA polymerase I shown in **A.**, 90° towards the reader. **C.** Visualization of side chains. The same angle is shown as in **B.** and the exonuclease domain, fingers subdomain, and the red DNA strand are removed to more simply highlight T672 and D673 side chains and their proximity to DNA. Images in Figures 4.13 and 4.14 were generated with Cn3D version 4.1 (NCBI, Bethesda, MO and Beese *et al.*, 1993).



polymerase active site and the proofreading active site (in the exonuclease domain, Figure 4.12).

Although the direct role of the Mip1p threonine-661 residue in DNA polymerase and 3'-5' exonuclease activities is unclear due to the lack of conservation between mitochondrial DNA polymerases and bacterial polymerases, it is tempting to suggest T661 may play a role in stabilizing interactions with the DNA phosphodiester backbone. This interaction could be similar to that of the *Taq* DNA polymerase S577 and Klenow T672, which interact with duplex DNA through either indirect or direct hydrogen bonds (Figure 4.12, Eom *et al.*, 1996). In the case of the *Taq* DNA polymerase-DNA complex, S577 has an indirect hydrogen bond with a phosphate group of the DNA backbone, while in that of the Klenow fragment complex, T672 has a direct hydrogen bond with a phosphate. Based on these scenarios, the T661 to A661 substitution in Mip1p could potentially disrupt the polymerase domains' interaction with DNA (Figure 4.12).

Alternatively, T661 could be involved in direct or indirect interactions with residues in the protein. D673, which aligns with T661 in Mip1p, is within 5 Å of S658, in β 7 in the cleft, and M647, in the loop between β 7 and helix I (see Figure 1.2). Similarly, in the crystal structure of *B. stearothermophilus* DNA polymerase I bound to duplex DNA, the location of the T619 and E620 residues are identical to those of the orthologous Klenow residues T672 and D673 (Figure 4.11 to Figure 4.13). In fact the E620 residue of *B. stearothermophilus* DNA polymerase I, which aligns with T661 in Mip1p, is located in the polymerase domain and is within 3.6 Å of N326 of the exonuclease domain, which would allow for a potential hydrogen bond between the side chains of these two residues (Figure 4.14). Perhaps a similar type of interaction occurs between orthologous Mip1p

Figure 4.13 Structure of *Bacillus stearothermophilus* DNA polymerase I (PBD structure accession number 1NJW) bound to duplex DNA highlighting the homologous Mip1p threonine-661 region and DNA bound in the polymerase domain. α -helices are shown as cylinders and β -stands are shown as arrows. In each panel the threonine (T619) and glutamate (E620, same position as Mip1p threonine-661 in the alignment shown in Figure 4.11) residues are highlighted in yellow and indicated by a white arrow. **A.** Half-open right hand of DNA polymerase I, which has a similar overall structure to Klenow fragment. The fingers subdomain comprises the light brown and gray α -helices, the thumb subdomain comprises the dark brown α -helices and the palm subdomain consists predominantly of green β -strands and a several green α -helices while the exonuclease domain is represented as purple α -helices and β -strands. The template DNA strand is shown in green and the primer strand in red. The duplex DNA is bound in a cleft (containing T619 and E620) formed between the polymerase and exonuclease domains. The 3'-end of the primer strand is located in the active site of the polymerase domain near a magnesium ion. **B.** Rotation of DNA polymerase I shown in **A.**, 90° towards the reader. **C.** Visualization of side chains. The same angle is show as in **B.** and the exonuclease domain, fingers subdomain, and the primer DNA strand are removed to more simply highlight T619 and E620 side chains. Other ligands in the structure include: 1. two sucrose molecules shown as grey rings with red oxygens, one located in the thumb subdomain and the other located between the exonuclease and polymerase domain, 2. Three sulphuric acids shown as yellow tetrahedral molecules with red oxygens and 3. A single magnesium ion located in the active site of the palm subdomain of the polymerase domain (Johnson *et al.*, 2003).

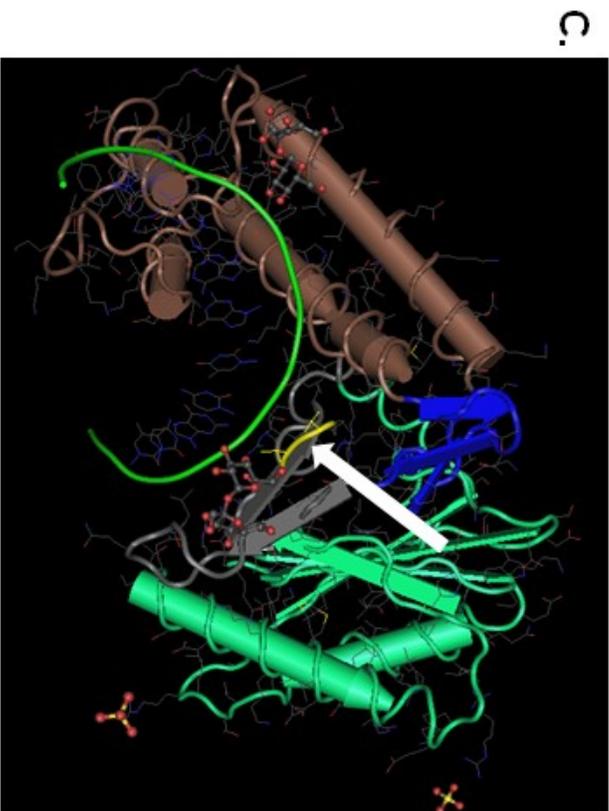
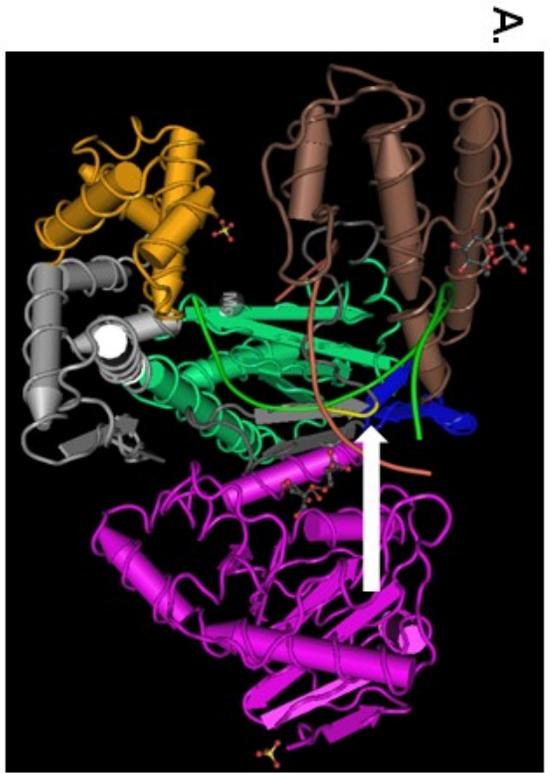
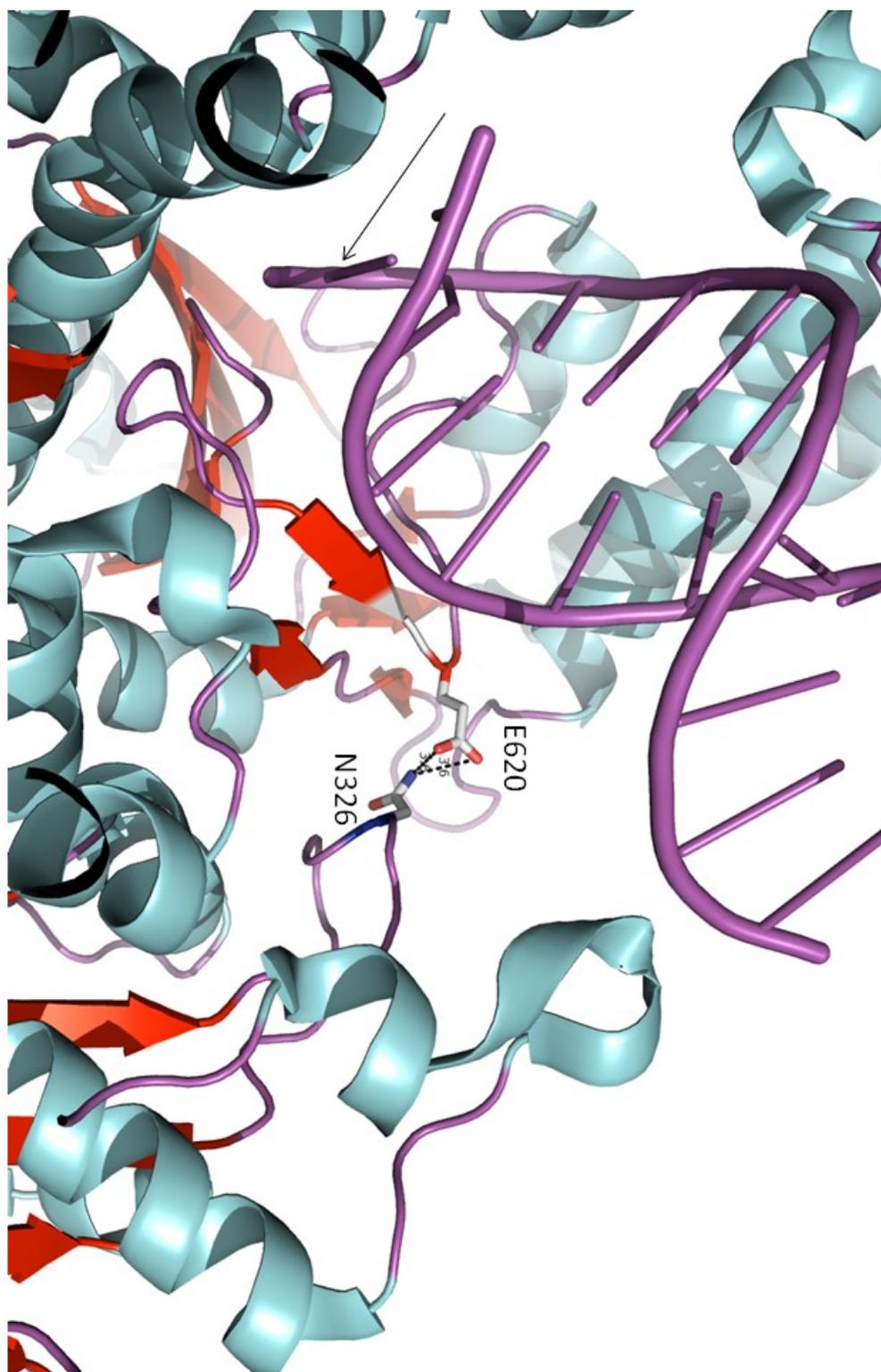


Figure 4.14 Putative hydrogen bond formed between *B. stearothermophilus* DNA polymerase I residues E620 and N326. E620 aligns with Mip1p T661 and is located in the polymerase domain of the crystal structure. N326 is located within the exonuclease domain and its side chain is 3.6 Å from the side chain of E620 (indicated by black dotted lines). An arrow represents the location of the carboxyl-terminus of the enzyme located behind the Duplex DNA (purple). Alpha helices are coloured blue while β -strands are red. The crystal structure was visualized with PyMOL. Protein Data Bank ID 1NJW.



residues to bring the polymerase and exonuclease domains into close proximity.

Uncoupling of this region in Mip1p[S], harbouring the A661 residue, could explain the decrease polymerization activity observed for this enzyme (Figure 4.9 A, B and C).

It should also be stressed that partial degradation of the enzymes at 37°C could be another explanation for the loss of polymerase activity at this temperature. In Figure 4.7 it appears that the Mip1p[S] variants may indeed be less stable than the Mip1p[Σ] variants, as indicated by some lower molecular weight bands, suggestive of protein degradation, below the intact Mip1p bands. Any carboxyl-terminal degradation would not be visible due to the location of the 6xHis-HA tags. Up to the time of loading the gel shown in Figure 4.7 all of these samples (or cultures expressing the variant polymerases) were processed on ice or at 30°C. It is therefore necessary to examine polymerase fractions by Western blot analysis after incubation at 37°C to ensure intactness of these enzymes at this temperature. Similarly, cells could be grown at 37°C overexpressing the enzymes, followed by PAGE gel analysis of the polymerases.

Future work with purified Mip1p variants could test these hypotheses of altered DNA binding ability combined with reduced polymerase activity. Differences in DNA binding could be examined by gel-mobility shift assays (Luo and Kaguni, 2005) and the secondary structure of purified Mip1p[S] and Mip1p[Σ] enzymes at 30 and 37°C could be investigated by circular dichroism to look for large scale changes in folding of these enzymes (Chan *et al.*, 2006). In addition, differences in proofreading activities would be examined (Luo and Kaguni, 2005) and ultimately crystallization of the polymerase should be attempted with the goal of determining the structure of Mip1p.

In addition to the role of the conserved T661, the function of the CTE was examined. All of the Σ 1278b-related truncation variants generated fewer RLUs on average compared to the wild-type polymerase at both temperatures (Figure 4.9 D – I). At 37°C the Mip1p[Σ] Δ 205 variant had the largest loss of RLUs by 15 minutes and almost no signal could be detected after 30 minutes (Figure 4.9 G, H, and I and Table 4.4). However, in the S288c-background, the Mip1p[S] Δ 205 variant had less net DNA loss compared to Mip1p[Σ] Δ 205 (-80% Figure 4.9 O vs. -92% Figure 4.9 I) and maintained approximately ten-fold more RLUs on average than Mip1p[Σ] Δ 205 by the end of the 30 minutes (Table 4.4 and Table 4.5). This may represent a difference in stability of the two different types of polymerases, Mip1p[S] vs. Mip1p[Σ]. Alternatively, if the Mip1p[Σ] version can bind to template DNA more tightly, this may explain why Mip1p[Σ] Δ 205 can degrade more DNA in comparison to Mip1p[S] Δ 205 and Mip1p[S]. Therefore, it is tempting to suggest that the CTE potentially contributes to the correct folding of the polymerase domain, due to the decreased DNA polymerase activity in membrane fractions harbouring overexpressed Mip1p[Σ]- and Mip1p[S]-truncation variants at 30 and 37°C (Figure 4.9 D to O). Reasons for this decrease could also be related to processivity and catalytic activity of the enzyme, where the CTE may function similarly to the accessory subunit found in the holoenzymes of animal mitochondrial DNA polymerases (Fan *et al.*, 2006). In yeast an accessory subunit has not yet been discovered. The human accessory subunit (pol γ - β or p55) shares a high level of structural similarity to class IIa aminoacyl tRNA synthetases (Fan *et al.*, 2006) and enhances DNA binding and processivity of the polymerase γ catalytic core (Lim *et al.*, 1999). The CTE does not

seem to resemble this protein based on its primary amino acid sequence or predicted secondary structure.

The polymerases harbouring the C-terminal 6xHis-HA tag was partially purified under denaturing conditions (Figure 4.10 A). Interestingly, an unidentified 70-kDa protein was also obtained in the elutions as indicated by Ponceau S staining of the blot. It is possible that this protein was bound to Mip1p during the purification procedure. This purification method was not pursued further as the goal was to obtain folded and catalytically active enzyme.

In contrast, under non-denaturing conditions the polymerases could not be successfully purified by Ni-NTA chromatography (Figure 4.10 B and C). Perhaps under non-denaturing conditions the polymerase folds in such a way that the 6xHis tag, internal to the C-terminal HA epitope, is inaccessible to Ni-NTA. Another possible explanation is that the CTE-region of Mip1p is bound to some other component(s) of the TMS and this interaction is blocking access to the 6xHis tag. Future work to obtain pure catalytically active enzyme could involve genetically removing the C-terminal HA epitope and extending the 6xHis to 12xHis. Alternatively, the enzymes could be purified using an anti-HA affinity matrix available from Roche. Other mitochondrial disruption techniques such as sonication or disruption through a French press may yield a more soluble form of Mip1p.

The potential for a mitochondrial DNA polymerase to incorporate DIG-11-dUTP into DNA is an exciting finding and has relevance for studying other members of the subfamily of gamma-type polymerases. This is interesting as differences in enzyme kinetics of purified disease-associated variant polymerases have been described in the

literature (Chan *et al.*, 2006). Additionally, gamma-type polymerases possess a unique reverse-transcriptase activity and it would be interesting to see if DIG-11-dUTP could also be used to detect this activity.

In summary, it appears that the alanine 661 substitution in Mip1p[S] and removal of 205 C-terminal residues of the CTE from the threonine 661-containing polymerase, Mip1p[Σ], may cause destabilization of the polymerase domain and/or decrease DNA binding to this part of the enzyme at 37°C. These observations will form the basis for future studies of Mip1p in our lab. Given the phylogenetic analysis done in chapter 3, the function of the CTE in mtDNA replication is likely a common mechanism among various fungi. This function could be specific to fungi as no accessory subunit appears to be required for yeast mtDNA replication.

Chapter 5 Conclusions and Future work

5.1 Conclusions

As a single cell facultative anaerobe, *Saccharomyces* is an invaluable tool for studies related to mitochondrial disease. The two distinct life cycles of haploid and diploid yeast give *Saccharomyces* the ability to reproduce in either the homozygous or heterozygous form. These forms allow the investigation of both recessive and dominant forms of orthologous human disease-associated mutations *in vivo*. A large part of the power of using this organism is the fact that it was sequenced over ten years ago (Goffeau *et al.*, 1996) and a brilliant database maintains this information, the *Saccharomyces* Genome Database (<http://www.yeastgenome.org/>). The consequence of having a well-annotated database is that most researchers have employed S288c-related strains for mitochondrial studies. We have shown that the maintenance of mitochondrial function is affected to different degrees by several genetic loci: *MRM1*, *HAPI*, *MIP1*[S], and *ade2-101ochre*.

While the effects of some genes, namely *MRM1* and *MIP1*[S], on mitochondrial function have been investigated in the work of others, the effects of these genes as well as others, namely *ADE2* and *HAPI*, have not been studied in a comprehensive and uniform way, which was done here for the first time. Also, the results of different combinations of mutations have not yet been presented and it is suggested here that different combinations of mutations likely have negative synergistic effects on mitochondrial function (namely *hap1* Δ and *his3- Δ 200/mrm1*, chapter 2). This is important given the heterogeneity of markers that appear in many standard laboratory strains, and those designed for use with different plasmids. In conclusion, researchers should choose parent strains having robust

mitochondrial functions to obtain comprehensible data from the study of orthologous human mutations.

To better understand mtDNA replication in yeast we choose to phylogenetically analyze fungal mitochondrial DNA polymerase CTEs and to characterize the CTE of Mip1p by truncation mutagenesis both *in vivo* and *in vitro*. In addition a non-radioactive, *in vitro* DNA polymerase assay was used to demonstrate reduced polymerase activity and increased temperature sensitivity associated with the Mip1p[S] protein. Various conclusions may be drawn from this work. First, mtDNA polymerase encoded by fungal representatives of the *Saccharomycetales*, plectomycete, and pyrenomycete as well as a basidiomycete (*Cryptococcus neoformans*) and a taphrinomycete (*Schizosaccharomyces pombe*) all contain putative CTEs. The length of CTEs among these fungal groups are quite variable and no obvious conservation was observed across all polymerases; however, the *Saccharomycetales* Mip1ps harbour around one-hundred residues of conservation in the N-terminal region of the CTEs. This region harbours three putative alpha helices (Figure 4.1, Young *et al.*, 2006).

In vivo analyses indicated that removal of the 279 amino acid residue CTE from the *MIP1* gene in *S. cerevisiae* was associated with rapid mtDNA loss. This suggests that this region may be required for either interacting with the mtDNA replication machinery (components of the TMS) or stability of the enzyme. A region of 205 carboxyl-terminal amino acid residues (Mip1p Δ 205) is dispensable for respiratory growth. This region is only 11 residues C-terminal to the truncation encoded by *MIP1 Δ 216*, which is associated with mtDNA instability during fermentative growth. Therefore, this 11-residue region of the CTE contributes to mtDNA maintenance and fidelity *in vivo*. In addition, the

MIP1Δ222 mutant also showed an inability to maintain mtDNA suggesting that a region of the CTE six residues long (between Mip1p Δ 222 and Mip1p Δ 216) is essential for mtDNA maintenance, either alone, or as part of a larger component of the enzyme.

A yeast expression system was developed to obtain partially purified catalytically active Mip1p from mitochondrial membrane fractions. Additionally, we developed a novel non-radioactive DNA polymerase assay and proved that Mip1p, from the gamma subfamily of family A DNA polymerases, can incorporate Digoxigenin-11-2'-deoxy-uridine-5'- triphosphate into DNA. Results from this study suggest that a conserved threonine at position 661 of Mip1p contributes to more robust DNA polymerase activity of Mip1p[Σ] in comparison to Mip1p[S]. Likewise, wild-type Mip1p enzymes (Mip1p[S] and Mip1p[Σ]) were catalytically more active than their isogenic Mip1p variants harbouring CTE truncations. Although preliminary, an exciting finding was made with respect to the exonuclease domain in the variant polymerases. It appears that the relative amount of mtDNA polymerase exonuclease activity may be a cause for DNA depletion over time at 37°C. This observation opens many possibilities for future experimentation.

Apart from analysis of one CTE point mutation (Hu *et al.*, 1995), no one has looked at CTE function. Given the phylogenetic analysis done in Chapter 3, the function of the CTE in mtDNA replication is likely a common mechanism among various fungi. This function could be fungal specific as no accessory subunit seems to be required for yeast mtDNA replication. Interestingly, the T7 phage also uses an accessory subunit (*E. coli* thioredoxin); therefore, mtDNA polymerases from the *Saccharomycetales* may represent a unique subset of accessory subunit-independent gamma-DNA polymerases.

In summary, the work done in this thesis will hopefully benefit the yeast and mitochondrial research communities at large by emphasizing the influence of yeast strain pedigree and genetic factors therein on mitochondrial function. In addition, CTEs are present on mtDNA polymerases from various fungi and a region of about 100 residues is conserved among the *Saccharomycetales*. These observations have added to our overall understanding of yeast mtDNA replication in general.

5.2 Future Work

One future project could be to further genetically dissect the function of the conserved region of the Mip1p CTE. The 11-amino acid residue region (between Mip1p Δ 216 and Mip1p Δ 205), shown to be essential for mtDNA maintenance and consisting mainly of a putative 15-residue alpha helix, could be targeted by truncation mutagenesis to narrow down the minimal essential region. The one step gene replacement strategy used in Chapter 3 could easily be employed to quickly generate 11 truncation mutants to determine the essential region. Alternatively, internal deletions or point mutations could be constructed using the *delitto perfetto* method (Storici and Resnick, 2006) to target conserved residues within this region. These mutants could then be characterized by analysis of basic mitochondrial functions such as respiratory competence, erythromycin-resistance, GFP-tagging and DAPI staining for mtDNA.

Another exciting project would be to test if the expression of the CTE in *trans* could rescue the fermentative loss of mtDNA in *MIP1 Δ 216*. The plasmids and oligonucleotides required for this analysis have already been constructed (Appendix Table 6.2). Another interesting question regarding *MIP1 Δ 216* is whether the Mip1p Δ 216 polymerase competes with Mip1p at the replication fork to inhibit wild-type mtDNA

replication during fermentative growth *i.e.* is *MIP1Δ216* a dominant or recessive mutation? This question could be addressed by mating the *MIP1Δ216* strain to a strain of the opposite mating type harbouring a wild-type *MIP1* gene.

What is the nature of the predicted association of Mip1p with the mitochondrial inner membrane? Probably the first step is to ensure that Mip1p has been properly targeted to the mitochondrial matrix. Intact mitochondria could be treated with proteinase K to ensure aggregates are not sticking to the organelles during their preparation. Following this step mitochondria should be more efficiently lysed (see suggestions below). If Mip1p remains tightly associated with the membrane, perhaps other conserved yet functionally unknown regions of the polymerase could be investigated for their requirement for membrane or TMS interactions; candidates might be the gamma-specific sequences.

Another critical aspect of future work is to obtain purified Mip1p from yeast mitochondria. Mitochondria were not easily lysed with hypotonic treatment plus mechanical shearing through a needle and attempts to bind Mip1p-6xHis-HA to Ni-NTA under non-denaturing conditions failed. Success purifying soluble catalytically active Mip1p from isolated yeast mitochondria has been documented when using mechanical extraction with an ultra turrax homogenizer (Vanderstraeten *et al.*, 1998). Human pol γ was successfully purified from recombinant baculovirus-infected insect cells treated with the non-ionic detergent Nonidet P-40 (Longley *et al.*, 1998). Copeland's group has also successfully lysed isolated mitochondria in the presence of Triton X-100. Logically the next step should be to try extracting Mip1p from mitochondria using one or a combination of these methods. As an alternative to the ultra turrax homogenizer to obtain

soluble Mip1p, one could try either disruption via sonication or French press to release the polymerase from the matrix and/or inner membrane. Another possibility is that the polymerase is entangled in mtDNA; therefore, treatment of lysed mitochondria with DNase I may help to release a more soluble form of Mip1p. Further genetic alterations could be made to the pMIPSig plasmid to either extend the coding region for the 6xHis tag to 12xHis and/or to remove the region encoding the HA epitope. Alternatively, the HA epitope itself could be used to purify the enzyme using an HA affinity matrix.

The preliminary results suggest the intriguing possibility that the Mip1p[Σ] Δ 205 CTE-mutant and the Mip1p[S] variant have a relative over abundance of exonuclease activity at high temperature compared to the wild-type, Mip1p[Σ] (Chapter 4 and Figure 5.1). Freemont *et al.* (1988) stated that Klenow excises approximately 10% of all correctly incorporated nucleotides (Freemont *et al.*, 1988). This raises three questions with respect to Mip1p variants studied in Chapter 4: 1) is the proofreading catalytic rate increased at 37°C? 2) is the polymerization domain destabilized at 37°C? and 3) is the DIG-11-dUTP causing increased melting of the 3'-end of the primer strand at the nonpermissive temperature (Figure 5.1)? To investigate the first question, Mip1p[Σ] Δ 205 membrane fractions could be incubated with pre DIG-labelled DNA to examine loss of RLUs or DNA degradation. To address the second question changes in secondary structure could be investigated using protease sensitivity assays, or by analyzing purified protein via circular dichroism. Lastly, the issue of DIG-11-dUTP causing increased melting could be addressed by comparing DIG assays with a radioactive nucleotide assay. In this type of experiment the radiolabelled dNTP should not sterically hinder the 3'-end of the primer strand as would be predicted for DIG-11-dUTP. Alternatively, DIG-

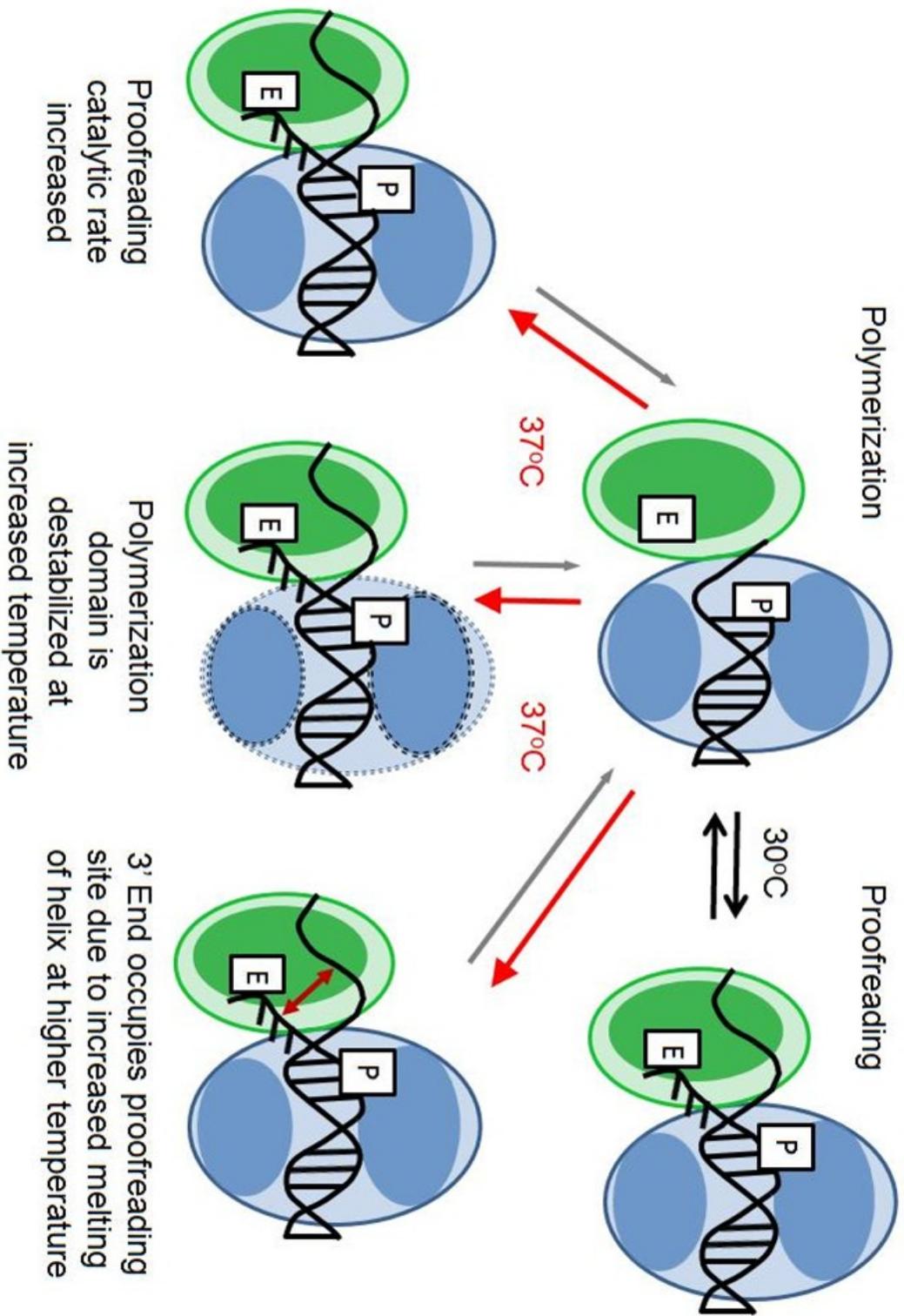
11-dUTP does not necessarily have to be incorporated into the DNA, the DIG label could be engineered onto the 5'-end of a primer.

Exonuclease deficient mutants could definitively prove that mtDNA depletion at 37°C is linked to a higher relative Mip1p exonuclease activity in CTE variants.

Exonuclease-deficient mutants have been successfully constructed in yeast and human polymerase γ and give the following targets in Mip1p for mutagenesis: exonuclease motif I (D171 and E173, Longley *et al.*, 1998) or exonuclease motifs I and II (D171 and D230 respectively, Vanderstraeten *et al.*, 1998).

With a purified enzyme in hand a plethora of experimental opportunities await. Some obvious experiments include: quantitation of enzymatic activities associated with Mip1p (polymerase and processivity, reverse transcriptase, exonuclease), investigation of DNA-binding through gel shift analyses, assessment of Mip1p-TMS interactions, and protein crystallization. One very interesting experiment would be to test Mip1p for the ability to perform rolling circle replication on oligonucleotides annealed to minicircles as was carried out with the human pol γ (Korhonen *et al.*, 2004). In fact this may be something worth examining even with partially purified polymerases because potential exists for rolling circle mechanisms in yeast (Shibata and Ling, 2007). Ideally reconstitution of the yeast replication fork with purified mtDNA, Mip1p, components of the TMS, and other protein targets implicated in mtDNA replication and recombination, could ultimately lead to a more universal and comprehensible model of mtDNA replication in yeast – or potentially all organisms. This penultimate goal would give an insight into the molecular workings of mtDNA replication which could benefit researchers trying to cure and understand human mitochondrial diseases.

Figure 5.1 Hypothetical models rationalizing decreased polymerase activity in Mip1p variants. Models are based on those of Freemont et al. (1988). The exonuclease domain (small oval) contains the catalytic site (E). The polymerization domain (large oval) contains the catalytic site (P) and the “fingers” and “thumb” domains in the “right hand” model of DNA polymerases. DNA (black) is held by the groove, and the 3' end of the primer can reside in either the P or E site. Proofreading requires denaturation of the end of the primer-template complex.



6. Appendix

6.1 Sequence of BS70

BS70 harbours a polymorphic *TRP1* gene that was cloned into pUC19 to create pUC19::*TRP1*. Due to polymorphic restriction endonuclease mapping of BS70 compared to the S288c *TRP1* gene cloned into the pBluescript *Bam*HI site, I sequenced BS70 to resolve this issue (Figure 6.1).

6.2 dsRed Sequence

To perform florescence colocalization experiments with yEGFP-tagged Mip1p variants I asked Dr. Jodi M. Nunnari if she would share her yeast pVT100U::*dsRed* plasmid with us. Kari Naylor sent us the plasmid as well as some hard-copy documents about it. The sequence found in Figure 6.2 was typed from those documents. dsRed is a red fluorescent protein cloned from a *Discosoma* species (coral anemones).

6.3 Fusion of Mip1p N-terminus to the CTE

Due to the carbon source-mediated loss (glucose or raffinose) of mtDNA in the *mip1Δ216* strain two questions arise. The first question is would expression of the CTE in *trans* rescue the loss of mtDNA when *mip1Δ216* cells are grown in these fermentable carbon sources? This rescue would be expected if the CTE has an analogous accessory subunit or clamp-like function. The second question is, alternatively, would expression of the CTE in *trans* exacerbate the carbon source-mediated loss of mtDNA? This effect would be predicted to occur if the CTE was required for the polymerase to interact with other components of the mtDNA replication machinery (i.e. TMS complex).

To address these questions we devised a genetic approach to overexpress a fusion protein consisting of the Mip1p amino-terminal matrix targeting signal and the CTE from

Figure 6.1 Clustal W (<http://align.genome.jp/>) alignment of TRP1 from the plasmid BS70 and the S288c genome. The *TRP1* start and stop codons are indicated by a line between the two sequences. A polymorphic *HindIII* site is indicated by italics in the *TRP1* gene obtained from the *Saccharomyces* Genome Database (SGD_TRP1, S288c). The fusion of *BglII* sites of the *TRP1* insert to the pBluescript *BamHI* site creating BS70 (seqbs70.seq) is indicated by boxes.

Figure 6.2 DNA sequence of the pVT100U::*dsRed* *Discosoma* sp. dsRed open reading frame.

>Ben Glick dsRed1-T1 from Jodi M. Nunnari and Kari Naylor

ATGGCCTCCACTCGTGTCCCTCGCCTCTCGCCTGGCCTCCCGGATGGCTGCTTC
CGCCAAGGTTGCCCGCCCTGCTGTCCGCGTTGCTCAGGTCAGCAAGCGCACC
ATCCAGACTGGCTCCCCCTCCAGACCCTCAAGCGCACCCAGATGACCTCCA
TCGTCAACGCCACCACCCGCCAGGCTTTCAGAAGCGCGCCTACTCTTCCGG
ATCCCCGGCGCGGGCCCGGGATCCACTAGTCGCCACCATGGCCTCCTCCGAG
GACGTCATCAAGGAGTTCATGCGCTTCAAGGTGCGCATGGAGGGCTCCGTGA
ACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGG
GCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCCCTTCGC
CTGGGACATCCTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAG
CACCCCGCCGACATCCCCGACTACAAGAAGCTGTCCTTCCCCGAGGGCTTCA
AGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCA
GGACTCCTCCCTGCAGGACGGCTCCTTCATCTACAAGGTGAAGTTCATCGGC
GTGAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACTATGGGCTGGG
AGGCCTCCACCGAGCGCCTGTACCCCCGCGACGGCGTGCTGAAGGGCGAGAT
CCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAG
TCCATCTACATGGCCAAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGG
ACTCCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAGCA
GTACGAGCGCGCCGAGGGCCGCCACCACCTGTTCTGTAG

the yeast genome. Two plasmids were engineered to encode a protein where the amino-terminal 47 residues of the 1254 Mip1p pre-protein were fused to the carboxyl-terminal 295 residues starting at K960. One plasmid encodes the 32 amino acid 6xHis-HA tag (pNCADH1) on the C-terminal end and the other does not (pNC, Table 6.1 and 6.2). These constructs were confirmed by sequencing.

The next step was to place the ORF encoding these fusion proteins under the control of a galactose inducible promoter at the *URA3* locus of the yeast genome using a PCR product amplified from either pNCADH1 or pNC. The Nterm_CTE_deltaURA3_F/R primer set was used and selection for transformants was on 5-FOA plates. I was not able to obtain transformants on 5-FOA plates.

6.4 Deletion of the *HO* endonuclease gene

During growth of S150 transformants in SC-LEU broth to induce overexpression of recombinant Mip1p, cells appeared to form “shmoos” (pear-shaped cells) indicative of yeast cell mating. It has been documented that strains carrying the *HO* gene are unstable with respect to mating type and can change from *MATa* to *MAT α* or *MAT α* to *MATa* nearly every cell division (Herskowitz, 1991 #277). Therefore, to prevent mating type switching in future yeast cell constructs, we engineered an S150 background strain harbouring a deletion of the *HO* gene (S150 Δ *ho*, Table 6.2). The wild-type *HO* gene was replaced by a simple one-step gene replacement using a PCR product amplified from pKT209. (*CaURA3*) and its disruption was confirmed by colony PCR of upstream and downstream recombination junctions using HOUSRJ_F/CaURA3R and HODSRJ_R/CaURA3F primer sets respectively (Table 6.2).

Table 6.1 Plasmids utilized and constructed in this study					
Name in book	Name	Description	Plasmid cut with which R.E. to create	Relevant IROs or Primers	Reference
NC#2	pNCADH1	Fusion of N-terminus of Mip1p to the CTE plus the 6-His-HA-tag	pNTBG1805:: <i>MIP1-ADH1</i> Term+ <i>LEU2</i> cut with <i>MscI</i>	N_Term_CTE_Fuse_F/N_Term_CTE_Fuse_R	This work
TBNA1a	pNC	Fusion of N-terminus of Mip1p to the CTE and the 6-His-HA-tag is removed	pNTBG1805:: <i>MIP1+LEU2</i> cut with <i>MscI</i> and <i>AscI</i>	N_Term_CTE_Fuse_F/N_Term_CTE_Fuse_R and Fuse_Tag_Remove_For/GFP_Rev PCR product	This work

Table 6.2 Appendix Primers and oligonucleotides used in this study

Primer Name	Primer Sequence 5'-3'	T _m (°C)	Gene/Target	Description
MIPUSdelta175_URA3_T7	CTATATAAAGGATGTCGAGA AGGGCAAAGGACTAAAGTA CGTATTATGGGATCCTGA _GT AATACGACTCACTATAGGGC	64	T7 of BS50	Bold = sequence homologous to <i>MIP1</i> for recombination to create <i>MIP1Δ175::URA3</i> by one-step gene replacement; <i>italics</i> = engineered stop codon; normal = T7 primer to amplify <i>URA3</i> from BS50
MIPDSdelta175_URA3_T3	AATAATATTATAATGTGCTGT ATATATAAATACAAATGCGA AAGCTAATGCAGATTTTGC _A ATTAACCCTCACTAAAGGG	56	T3 of BS50	Bold = sequence homologous to <i>MIP1</i> for recombination to create <i>MIP1Δ175::URA3</i> by one-step gene replacement; normal = T7 primer to amplify <i>URA3</i> from BS50
URA3_Int_Rev	CAACTAACTCCAGTAATTCC	56	<i>URA3</i>	To screen <i>MIP1Δ175::URA3</i> upstream recombination junction; used with MIPDSofE (1117 bp product)
URA3_Int_For	AAACGAAGATAAATCATGTGCG	56	<i>URA3</i>	To screen <i>MIP1Δ175::URA3</i> downstream recombination junction; used with DeltaScrMIPRev (1180 bp product)
N_Term_CTE_Fuse_F	TTGTGCTCGATGGTTCTCCACA AAGAAGAATACCGCAGAAGCA CCCAGGATTAATCCTGTGGGG ATACAGAAAACCGCCA		pNTBG1805:: <i>MIP1</i> - ADH1Term+ <i>LEU2</i> or pNTBG1805:: <i>MIP1</i> + <i>LEU2</i>	IRO to create the fusion of N-terminus of Mip1p to the CTE
N_Term_CTE_Fuse_R	CCAATTTACTATTGGATTTGTC TAGCAGTTGATTGATATCAAG CGCCTCCCATGAGGAATGGC GGTTTTCTGTATCCCC		pNTBG1805:: <i>MIP1</i> - ADH1Term+ <i>LEU2</i> or pNTBG1805:: <i>MIP1</i> + <i>LEU2</i>	IRO to create the fusion of N-terminus of Mip1p to the CTE

Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
Fuse_Tag_Remove_For	CACTACAAATAGAAATTTGG TTGAGCTGGAAAGGGACATT ACTATTTCTAGAGAGTACTGA _CGCCACTTCTAAATAAGCG	56	ADH1 terminator from pKT128, ({Sheff, 2004 #141})	Bold = pBG1805:: <i>MIP1</i> -derived sequence for recombination; <i>italics</i> = engineered stop codon; normal = primer for amplification and addition of ADH1 terminator; Used in combination with GFP_Rev to generate a PCR product for creation of pNC with N_Term_CTE_Fuse_F/N_Term_CTE_Fuse_R IROs
Nterm_CTE_deltaURA3_F	ATTGCCCAGTATTCTTAACC CAACTGCACAGAACAAAAAC CTGCAGGAAACGAAGATAAA TC_AAAAGCTGGAGCTCCACC	56	pNCADH1 or pNC	Bold = upstream <i>URA3</i> sequence; normal = pNCADH1 or pNC specific; to insert the galactose inducible fusion of the N-terminus of Mip1p to the CTE plus or minus the 6-His-HA-tag at the <i>URA3</i> locus by one step gene replacement
Nterm_CTE_deltaURA3_R	ATTAAATTGAAGCTCTAATTT GTGAGTTTAGTATACATGCA TTTACTTATAATACAGTTTT GAGGCAAGCTAAACAGATC	56	pNCADH1 or pNC	Bold = upstream <i>URA3</i> sequence; normal = pNCADH1 or pNC specific; to insert the galactose inducible fusion of the N-terminus of Mip1p to the CTE plus or minus the 6-His-HA-tag at the <i>URA3</i> locus by one step gene replacement
URA3_USRJ_F	CATCATCTCATGGATCTGC	56	Yeast <i>URA3</i> upstream area	to be used with URA3IntRev to screen the US recombination junction (641 bp fragment)
URA3_DSRJ_R	TCATTACGACCGAGATTCC	56	Yeast <i>URA3</i> downstream area	to be used with URA3IntFor to screen the DS recombination Junction (916 bp fragment)
Delta_HO_F	TTAGCTCTAAATCCATATCCT CATAAGCAGCAATCAATTCT ATCTATACTTTAAAGTTTAC TTGCCTCGTCC	56	pKT209	Bold = sequence for recombination at the <i>HO</i> locus; normal = primer for amplification of <i>caURA3</i> gene from pKT209

Primer Name	Primer Sequence 5'-3'	Tm (°C)	Gene/Target	Description
Delta_HO_R	AAAATATTAAATTTTACTTTT ATTACATACAAC TTTTAAAC TAATATACACATTTCGATGAA TTCGAGCTCG	54	pKT209	Bold = sequence for recombination at the <i>HO</i> locus; normal = R3 primer (Sheff, 2004 #141) for amplification of caURA3 gene from pKT209
HOUSRJ_F	CTAGTACTACCATTGGTACC	58	<i>HO</i> upstream region	use with CaURA3R to screen upstream recombination junctions at <i>HO</i> locus
HODSRJ_R	GAGGAAAGTTGATCAAGACC	58	<i>HO</i> downstream region	use with CaURA3F to screen downstream recombination junctions at <i>HO</i> locus

6.5 Alignment used for mitochondrial DNA polymerase phylogeny

The alignment that was used to generate the fungal phylogenetic tree in Chapter 3 Figure 3.2 is shown in Figure 6.3. The methods used to create the alignment are described in Chapter 3.

Figure 6.3 Mitochondrial DNA polymerase alignment used for phylogenetic analyses. Sequences were aligned and edited as outlined in section 3.3.5 Phylogenetic analyses. The alignment is presented as it appears in GeneDoc v 2.5.010 conservation shading mode (Nicholas, 1997 #127). Four different shading levels of conservation are presented; black, dark grey, light grey, and no shading are 100, ≥ 80 , ≥ 60 , and < 60 percent conservation in a column respectively. The top two levels are distinguished by upper (100%) and lower case ($\geq 80\%$) characters on a consensus line below the alignment.

Sc, *Saccharomyces cerevisiae* (S288c); Spa, *Saccharomyces paradoxus*; Sba, *Saccharomyces bayanus*; Cgl, *Candida glabrata*; Kwa, *Kluyveromyces waltii*; Debaryomyces_ha, *Debaryomyces hansenii*; Ncr, *Neurospora crassa*; Ego, *Eremothecium gossypi*; Klactis, *Kluyveromyces lactis*; Gze, *Gibberella zeae*; Pan, *Podospora anserina*; Mgr, *Magnaporthe grisea*; Hca, *Histoplasma capsulatum*; Calb, *Candida albicans*; Sca, *Saccharomyces castellii*; Ani, *Aspergillus nidulans*; Yarrowia_lipoly, *Yarrowia lipolytica*; Ppa, *Pichia pastoris*; aspfum, *Aspergillus fumigatus*; Sku, *Saccharomyces kudriavzevii*; Spo, *Schizosaccharomyces pombe*; cne, *Cryptococcus neoformans* var. *grubii*; Pcr, *Phanerochaete chrysosporium*; Hs, *Homo sapiens*; Dmel, *Drosophila melanogaster*. Accession numbers for all of these mtDNA polymerase sequences can be found in Table 3.1.

```

      *           20           *           40           *           60
Sc      : APRINPVGIOYLGESLQROVFGS-EQSDK---IMELSKKSLKDHGLWCKKTLITDPISFPIPPLOG : 62
Spa     : APRINPVGIOYLGESLQROVFGN-GQSDK---IMELSKKSLKDHGLWCKKTLITEPISFPIPPLOG : 62
Sku     : AARINPVGIOYLGESLQROVFGS-GQSDK---IMELSKQSLKDHGLWCKKTLIAEPISEFPIPPLOG : 62
Sba     : APRINPVGIOYLGESLHROVFGS-GHGDK---IIELSKQSLKDHGLWCKKTLITEPIAFPIPPLOG : 62
Sca     : -----SQSLHSOLFENKMTPEERTHIDLSKQLLKSGLLCKKTAISEPISFELPKLOG : 54
Cgl     : QPRTNPLGICQYLSESLQQOVFPG-GHSED-PEIVKLAKLSISNEDLLGRKTTITKPIAFKIPPELOG : 64
Kwa     : KPRINPVGIOYHLSKHLHEOVFQGNQVSEAEKKAIVRLSQSFLKKGHLGKKTATISEPISFDFVFKLOG : 66
Klactis : KPRINPVGIOYLSNSLHEOIFPS-FTEEEKQELINLSKTFILRNHDLGKPTNISDPISFELPKLOG : 65
Ego     : EPRINPVGIOYMSHSLHKQLRGASLDEAARGKIVKWSKLYLKLHGLGSKTSISEPISFELPQLOG : 66
Debaryomyc : APRINPVGVOYLSSEELHKKLFPK-YLKPKNPALELAKEDHRYNELGKKTIVEPFINIDFPDLEG : 65
Calb    : EPRVNLGICQYLSNDLHKKVFPYPT-YLSPQHPQLLEISKKHQENELGKKTQITEPITIKFESLVG : 65
Ppa     : --RINAVGICQYHAPSLQNLDFGK-----LDEDLINMAKRHLKSNLIGKSSSLNDPTEIQIFPSLOG : 59
Yarrowia : GRRINPMGICQMLSSGLYNQLFGD-LTKPMPVKLEAMAKHHDMKHELCKKTKKENIIDFDLPPMLG : 65
Ani     : VARFNEIGVQQLSDHVYSQIFDFNK-PTPPDPNIVALSKDHLAREDLGKAQEHADVAFDLPPLOG : 65
aspfum  : -----MPALOG : 6
Hca     : TARFNEIGVQQLSSHVHSCIFLRE-PTPPNPETIALSKDHLRREDDLCKTQEGSNPIAFDIPPELG : 65
Gze     : EARHNEIGVQQLSSHVFEQIFPDG-VKPPPQELVELSKDHLRREDDLCKNTDKSDPIAFDIPPELG : 65
Mgr     : GARFNEIGVQHLSDHVFKQVPLG-PDPPPKHVLAEDHDLAREDDLGRITDTSEPISEFDFPIPHG : 65
Ncr     : TARHNEIGVQQLSEHLYKQLFPRG-TDPPAPELLELAKDHLAREDDLCKTTDKTPIAFQIPALVG : 65
Pan     : TARYNEIGVQQLSSHIYDOLFPGK-TSPPSKELVDLARDHLRREHDLCKNTDTSAPVAFDIPPLVG : 65
Spo     : PPRINPVGVOYLSPALQNOVFPQQ-NTQSLHLDLAKFHAKHQLLNKETIKLPSFNELRPELOG : 65
cne     : GPRRNPGVQMLSSSLHSOLFPGQPLPKPPQALDISKRHLKDNDFPEGAAVLPETISFNUPSLRG : 66
Pcr     : -AKRNEVGVQLSRRRLHQIFRNVSFPNNPETFRIAREHDKMGLDPSQGSVLPDIFGTLPSLOG : 65
Hs      : QLRHNPLDIQMLSRGLHEQIFGQGMFGEAAVRRS---VEHLQKHGLWQOPAVPLPDVELRPEPLVG : 63
Dmel    : EYAENLVKVOIMISRNLAHQIFPQASISEQQVASAKVYKDELRRHGVDISSAPVSDVQLKIPALRG : 66
      r n g q l           q f           l           l h l g           P l G

```

```

      *           80           *           100           *           120           *
Sc      : RSLDEHFQKIGRFNSEPYKSFCEDKFT-EMVARAEWLRKPGWVKYVPGMAPVEVAYPDEELVVFD : 127
Spa     : RSLDEHFQKIGQFSSEPYKSFCEGKFT-EMVARAEWLRKPGWVKYVPGAAPVEVAYPDEELVVFD : 127
Sku     : QSLDEHFQKIGQFNSEPYKSFCLHKFT-EMVARAEWLRKPGWVKYVPGAAPVEVAYPDEELVVFD : 127
Sba     : RSLDEHFQKIGQFNSEPYKSFCEGKFT-QMVPREAWLRPGRWVKYVPGAAPVEVAYPDEELVVFD : 127
Sca     : NSLDEHFQKLGHFASEPYKTMVAVNKFT-TVLPKENKWI RTAGWMMKYE PGKAPVAVPPEENTLVFD : 119
Cgl     : SSLDEHFQRLGLYASEPYRTMALEKFK-SIPEKKNWIRKPGWYRYEKGKEPTRVDEPSEDTLVFD : 129
Kwa     : RTLDEHFQKLGHYAAQEFERDLGSKFT-EIPHRPSRDLRLSGWVRYGAGLEPERVEFPEDTIVFD : 131
Klactis : RILDEHFQKLGHFASEPYKSMCDNKFK-NIVPREKKWLRERGWYRYEPGKEPEKVPYPLEDTLVFD : 130
Ego     : TTLDEHFQRLGHFASPYLSLACENKST-EILLPRQWLRKAGWYRYAPGREPEAVASPLEDIYVFD : 131
Debaryomyc : KTLDEHFYRIGTEASQPYLSIAESFLSLPTKPKSEWLFKSGWVRYAEGEAPKEVPHLEDELVFD : 131
Calb    : KSMDEHFNKLQKSSDPYLAMATNFKLKPCKPPKSQWKFQSGWTRYTDGNPPEQVYPLEDTLVFD : 131
Ppa     : NTLDEHFYKLGASLEPYNPLIDELLNLIIEPQFEFAFQSGWTRYAPGEPSEKVPYPLEDSYFVD : 125
Yarrowia : SSLDEHFYRMAKVMGEPYMDIIQGMFTFAKPRERDVKTIISGMRYAKGKKPERVDYPKENGLVFD : 131
Ani     : QTLDEHFYKLGMDSSDPYLTYSKEYAVVNSEPELPRKTVRRSGWTKYNSDGSWEAVDAPNESMITFD : 131
aspfum  : QSMDEHFYKLGMDASEPYLTYAKSYTTVNSPQKPREWVKRSGWTKYNSDGSWEVVDAPNEPMLTFD : 72
Hca     : QTMDEHFYKLGMDCSBPYLTAQYAKIDLPPIEKKWIRRSWTKYANGTSEAVDAPNEEMITFD : 131
Gze     : RTLDEHFHKLGIIDAAEPFLTFKQFSMANPPPKERQWVHQS GWTKYNDGTTEKVEAPEEEMLSFD : 131
Mgr     : KTLDEHFHKLGIIDAAEPYLTRAKQFARASTPPKERKQVVRQSGWTKYYPDGKTEAVDAPDETSLTFD : 131
Ncr     : DTLDEHFHKLGVDAABEFLTHAKQFADAHLPKPETSWVRRSGWTKYNRDGTTEVDLPPQGNMCFD : 131
Pan     : STLDEHFHKLGVDAABEFLTHAKRFALANPPPKERKQVRRSGWTKYADGRTEQVDAPEGEAMTFD : 131
Spo     : KTISEHFYNIQLEFADPHLSKAIKFSKIDTPVQEKTKRQPGWTKYAKDGSISCVPEPDSDCMVF : 131
cne     : NNIRDHFHTLQYTAEPYASMAAREFAATKLPKAPDRWEMRPGWTKY-YSDGRAVDLGDFTLVSFD : 131
Pcr     : RTLDEHFYSIQSVAAQPWLTAKDLGSTELPPKEGDWLLQPGWTKYQSDGTPMHDSQPQMLVFD : 131
Hs      : DNLQHFRLLAQKQSLPYLEAAN-LLLQLPKPATAWAEGWTRYGE-GEA--VAIPEERALVFD : 124
Dmel    : ANIEEHFHNIKAEQVQPYEELLLPLVQ-QLPKRKRWAFHTGWTADEDEGTA--VDHELEKGLVFD : 129
      deHF g P           p w GW Y           v p e FD

```

```

      140          *          160          *          180          *          2
Sc      : VETLYNVSDYPTLATALSSTAWYLVCSPFICGDDP--AALIPLNTLNKEQVVIIGHNVAYDRARVLE : 191
Spa     : VETLYNIADYPTLATALSSTAWYLVCSPFICGDDP--TALIPMNTLTKEQVVIIGHNVAYDRARVLE : 191
Sku     : VETLYNISHYATLATALSSTAWYLVCSPFICGDDP--AALIPMNTLHKEQVVIIGHNVSYDRARVLE : 191
Sba     : VETLYNISHYPTLATALSSTAWYLVCSPFICGDDP--AALIPMNTLRKEQVVIIGHNVAYDRARVLE : 191
Sca     : VETLYKISPYPTLATALSDKAWYLVCSPFICDNNSSFKHLIPLGENTKEKLIIGHNVGYDRARVLE : 185
Cgl     : VETLYKISPYPTLATAVSDKAWYLVCSPFYICDKHEVYEHLIEMDTLNKEKLIIGHNVGYDRARVLE : 195
Kwa     : VETLYKIAPYPTLATAVSDKAWYLVCSPFICGESDVPTHLIPLDTLNKTRLVIGHNVGYDRARVLE : 197
Klactis : VEVLYKISHFPTLAVALSDDKAWYLVCSPFMCDGSDNFKHLIPLMNTLERPKLIVGHNWGYDRSKVLE : 196
Ego     : VECLYKISDYPTLAVALSDDKAWYLVCSPFYICG-EDNFVHLIPLNMTMRPQVVIIGHNVGYDRSKVLE : 196
Debaryomyc : VEVIYKKSPLYAALATCVSSKAWYLVCSPFYITN-YNDWEHLMFMDCLNNPKLIIGHNVSYDRARVLD : 196
Calb    : VEVMYKISNYEVMATCASNEAWYSVSEALLDDQIDWNHLIEMNCAKPKVVIIGHNVSYDRARILE : 197
Ppa     : VETMYKVSQYPMVAVALSDKAWYLVCSPFYITE-----HLIPLNTTFETEKFIVGHNWGYDRARALE : 185
Yarrowia : VEVLKVSQYPMVASAVSTEAWYLVCSPFVLEHQPESMHLIPLGPPDEPRVIVGHNWGYDRARQAVKE : 197
Ani     : TEVMYKEHPFAVMACAVSPTAWYAWLSPWLLRESENEIQLVPLGDPTKPRIIVGHNIGYDRARVLE : 197
aspfum  : TEVMYKEHSPFAVMACAVSPTAWYAWLSPWLLGESENEVQLIPLGDKKQPRIVVGHNIGYDRARVLE : 138
Hca     : TEVMWKESSFAVMACAASPSAWYAWLSPWLLGETENERQLIPLGSPSQARIIVGHNIGYDRARILE : 197
Gze     : TEVMWKISPFFAVMACASSPTAWYAWLSPWLLGETENDRQLIPLFGDPTKDRIIVGHNIGYDRARVLE : 197
Mgr     : TEVLWKESSFAVMACAVGPNAWYAWLSPWLLGESESDRHLIPLGDPSKDRIIVGHNIGYDRARILE : 197
Ncr     : TEVMYKDNPHYAVMACAGTPDAWYAWLSPWLLGETENKAQLVPMGDPTVDRIIVGHNIGYDRARILE : 197
Pan     : TEVMWKESSPYAVMACAATADAWYAWLSPWLLGETENEHQLIPLVGDPTKERVVIIGHNIGYDRARILE : 197
Spo     : VEVLKVSQYPMVAVALSDKAWYLVCSPFYITE-----HLIPLNTTFETEKFIVGHNWGYDRARALE : 185
cne     : VEVLKVSQYPMVAVALSDKAWYLVCSPFYITE-----HLIPLNTTFETEKFIVGHNWGYDRARALE : 185
Pcr     : VEVLKVSQYPMVAVALSDKAWYLVCSPFYITE-----HLIPLNTTFETEKFIVGHNWGYDRARALE : 185
Hs      : VEVLKVSQYPMVAVALSDKAWYLVCSPFYITE-----HLIPLNTTFETEKFIVGHNWGYDRARALE : 185
Dmel    : VEVCVSEGQAEVPLATAVSTKRWYSVSSKLTG--LSVEKLEPLDVTDSPLVVGHNWGYDRARALE : 193
      E          A a s aWY W Sp          Lip          GhN yDRar e

```

```

      00          *          220          *          240          *          260
Sc      : EYNFRDSKAEFLDTQSLHIIASFGLCSRQRFMFKNKKKK---EAEVESEVHPEISIEDYDDPWLN : 253
Spa     : EYNFRDSKAEFLDTQSLHIIASFGLCSRQRFMFKNKKKK---EIEVESEVQSEISIEDYDDPWLN : 254
Sku     : EYNFRDSKAEFLDTQSLHIIASFGLCSRQRFMFKNKKRKEEAEAEVQSEISIEDYDDPWLN : 257
Sba     : EYNFQDSKAEFLDTQSLHIIASFGLCSRQRFMFKNKKKKESE-PEPAETEAEQSEISIEDYDDPWLN : 256
Sca     : EYNFNHNSKFFELDTQSLHIVASSGLCSRQRFMFQMKRRKEKEEQLQOETESEEDLEEEEDPWLN : 251
Cgl     : EYNLKPDKAEFLDTLHIVATAGLCSRQRFVFLKNNKLNKENLTAPEPEIEEVTEALNDVADPWLK : 261
Kwa     : EYNCQESKAEFLDTMSLHIIASSGMCQRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Klactis : EYNFTPSKAYFLDTMSLHIVASSGMCQRQRFYMKKLRDEEVAQLNAENDDLPIIA-VDEDSPWLN : 261
Ego     : EYNLAPSKAEFLDTMSLHIIASSGMCQRQRFYMLKQKTKQEQENDGDLDELNEDE-----NWN : 256
Debaryomyc : EYNIKQSKGFEFLDGMALHVAISGICQQRFTWQRHKKSK-SQLDSSKDESPAIES-ELLDD-PWLN : 259
Calb    : EYNIKQTKAFYLDGMSLHVATSGICSRQRFKWKARYNKVKKDEIEGTMEEEEEEEGEEVAED-PWLL : 262
Ppa     : EYNLQKTKAFYLDGMSLHVAVSGMCTRQRTWIKHKKKEQENTMSVASIKSKVQN-PMLEDNFWLN : 250
Yarrowia : EYSLKQTKNFEFLDTMSLHIVSVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Ani     : EYDLKQTKAFYLDGMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
aspfum  : EYGMQTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Hca     : EYNIKQTKAFYLDGMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Gze     : EYSLKQTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Mgr     : EYSLTQTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Ncr     : EYDLKQTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Pan     : EYSLTQTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Spo     : EYNIKSSRNVEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
cne     : EYSIERTQTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Pcr     : EYSTDGTQTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Hs      : QYLIQGSRRMFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
Dmel    : QYLTEDTGTRNFEFLDTMSLHVAVNGMCSRQRFQWQKRAQDKQVVEVEVEE-----LAFNNDPWT : 258
      eY          f Dt slH a G cs Qrp          k          W

```

```

      *           280           *           300           *           320           *
Sc      : VSALNSLKDVAKEFHCKIDLDKTRDRDFEASTDKSTIIENFQKLVNYCATDVTATSQVDFEIEFVELK : 319
Spa     : VSALNSLKDVAKEFHCKINLDRADRDFFASTDKSTIIENFQKLVNYCATDVTATSQVDFKIEFVELK : 320
Sku     : VAALNSLKDVAKEFHCKINLDRADRDFFASTDKSTVIENFQKLVNYCATDVTATSQVDFKIEFVELK : 323
Sba     : VSALNSLKDVAKEFHCKIKLDRADRDFFAATDKSTIIENFQKLVNYCATDVTATSQVDFKIEFVELK : 322
Sca     : ISTMNSLSDVALFHCGIKMDKEPRDFEASLDKQVVIDNFQKLVNYCATDVEATSHVFDKVLRFELK : 317
Cgl     : VSTMNSLSDVALFHCGIKMDKTRDRDFEASTNKQDVIDKFQQLVNYCATDVTITTSKVFDFKVFSEFMK : 327
Kwa     : VSTLNSLKDVALFHCGIKLDRQREDFGTLTKNDIIDKFNEMVNYCANDVEATSKVFDVAFELFELS : 324
Klactis : KSSANSLKDVARFHCNIKMSKEARESFATLDKFEIINNFDMSVVEYCAFDVEATSKVFDVEVPEKELQ : 327
Ego     : LSTMNSLSDVALFHCKINLDRADRDFFEALDKQEVIDKFQQLVNYCATDVTETTSKVFDFKVFLEFRE : 322
Debaryomyc : KGSFNSLANVADFHCNIKLDKTRDRDFEGTDPQDVINNFNLMDCAKDVAATYSVTAKLESQERQ : 325
Calb    : MGAFNSLANVAREHCNIHLDRADRDFFATEDPNLVIDNFPKLMDCYCAKDVAATFEVTKKLEVFETIQ : 328
Ppa     : KSSLNSLSEVAFEFYCKIKLKDIRNDFATEDPEDELGRNFDLIMQYCAKDVAATGKVFQKVYKPEKK : 316
Yarrowia : VSATNSLAEVAHLHLNLMVMSKDDRELFESQESLGELOQRVQQAIDYNIITDVEITATQVFKKILELELD : 329
Ani     : RSSVNSLSDVAKEFHCDVTIDKQRDFEGELERPOLAKLDELDDYCAADVAITHRVYKVFLENELE : 329
aspfum  : RSSVNSLSDVAKEFHCDVTIDKQRDFEGELNREQILGKLEELDDYCAADVAITHRVYKVFLENELE : 270
Hca     : RSSVNSLSDVAKEFHCDVTIDKQRDFEGDLDRSGILERLEDELDDYCAADVAITHRVYKVFLENELE : 329
Gze     : RSSVNSLSDVAKEHNLVSDIKTIRNDFEGELDRDGVVEKLEDELDDYCAADVAITHRVYQIVFPELKL : 328
Mgr     : RSSVNSLSDVAKEHNLVSDIKAIRDDFESVLDLDRDGTMAKLEELDDYCAADVAITHRVYSIVFPELE : 329
Ncr     : KSSINSLSDVAKEHNLVSDIKDIRDVAETDRNVILNQLDDELDDYCAADVAITHRVYQVVFPELE : 329
Pan     : RSSINSLSDVAKEHNLVSDIKAIRNDFEGELDRDGINQQLDELDDYCAADVAITHRVYQVVFPELE : 329
Spo     : HSSVNSLSDVAKEFHCDVTIDKQRDFEASLEKEPILQKLNELDDYCAADVAITHRVYQVVFPELE : 327
cne     : VTMNSLAEVAALFHCGYVVDKSVRDRFEGDSDIKQIHSELHQLLSYCAADVAITHRVYQVVFPELE : 329
Pcr     : ITSANSLADVAKEHNLVSDIKSVRDRFEGDSDIKQIHSELHQLLSYCAADVAITHRVYQVVFPELE : 324
Hs      : ISSVNSLAEVHRLYVGPPELEKPRELDFVKGTMKDIENFQDLMQYCAADVAITHRVYQVVFPELE : 320
Dmel    : QSSLNSLSEVHRLYCGDTLSKEPRNDFEGVTELEQVRQSFQSLTNYCASDVEATHRILRVLYLELAE : 306
      s NsL Va h K R F l Yca Dv T v p f

```

```

      340           *           360           *           380           *
Sc      : KCPHPVSEAGLKSLSKCIETKLDWMDYLNSSSESLYQQSKVQIESKIVQIIKDIVLLKDKPDPFYLK : 385
Spa     : KCPHPVSEAGLKSLSKCIETKLDWMDYLNSSSESLYQQSKVEIESKIVQIIKDIVLLKDKPDIYLK : 386
Sku     : KCPHPVSEAGLKSLSKCIETKLDWMDYLNSSSESLYQQSKVEIESKIVQIIKDIVLLKDKPDPVYLK : 389
Sba     : KCPHPVSEAGLKSMSKCVLTKLDWMDYLNSSSESLYQQSKVEIESKIVQIIKDIVLLKDKPDEYRK : 388
Sca     : KCPHPVSEAGLKTLSSEFILTTRREWDYDYNRSEKLYQESKLEIERKILLIIEELVKIKDDPDQIKD : 383
Cgl     : KCPHPVSEAGLKSLSKCIETDQKWQNYTKNSSENKYQDSKKVIDDKIEIVQETVKLKDQKDVIKT : 393
Kwa     : KCPHPVSEAGLRLLSTSIETKLDWMDYLNSSSESLYQKSKQSEKILTIIVEDIVLKDNLDSAVND : 390
Klactis : KCPHPVSEAGLRFLSTSIETRVKWEKYLNSSETFLDQNKAKIEDKIVKIVEDTLQLKNQTELKLN : 393
Ego     : QCPHPVSEAGLRFLSNCIETVQPMKQYIDGAEATLYQNRQAIEEKIITIIIEIVLKDKEPEVYEN : 388
Debaryomyc : KVPHPVSEAAIRHLGTLMLPTTKNWDNYIETAEVYQKNREEVSTILKERANAFVNYIDNEPDWET : 391
Calb    : RHPHPVSEAAIRLLGTLMLPTTTKWPKEYTEAAESLYESNRQEVATTLQQLANDLVAHIQKPKDYEN : 394
Ppa     : LHPHPVTLAALKDISSCIEPTTTKWDYIETSERLYQESRRMIEKNLHVICETVVKLD-DP-WEN : 380
Yarrowia : TPHPVSEATLKNISSAILPTDSEWKDYLNRCENMYQKMKKSIDEGQLSLVTEAVKLDKDKPEPWLN : 395
Ani     : TPHPVSEAGLRHLSSVILPVNQTWKEYLDNAESTYNQRLGDVQRRIVELCDEALSVKDDPEKYMN : 395
aspfum  : VCPHPVSEAGLRHLSSVILPVNETWKEYITNABEATYHQRVDDVQRRIVELCNEALKVKENPDIYMN : 336
Hca     : VCPHPVSEAAIRHLSSVILPVDKSDAYTKNABEATYHTLSDAVQERLIGLTESALEMKDSPGEYEN : 395
Gze     : VCPHPVSEAAIRHLSSVILPVNESWDSYIATANABEATYQKLESEAVQERLIALTEKALEIKDDQDKWSA : 395
Mgr     : VCPHPVSEAAIRHLASVILPVNKIWDYIETABEATYQLMLHGQVQERLFTLMERTLDYKADPEKYLK : 395
Ncr     : VCPHPVSEAAIRHLNSVILPVNKIWDYIETABEATYQLMLHGQVQERLFTLMERTLDYKADPEKYLK : 395
Pan     : VCPHPVSEAAIRHLNSVILPVNKIWDYIETABEATYQLMLHGQVQERLFTLMERTLDYKADPEKYLK : 395
Spo     : VCPHPVSEAAIRHLASVILPVNHSWTRYINGVEEQYQMIQLVDQKLSQYAEKAKDLIN--DTVLK : 391
cne     : SCPHPVTLGILSMGSSFLPTDQSWKEYLRNABEATYREMDVAVKALRLLAEKLR--AE--EPKKG : 391
Pcr     : SCPSPVSEAGILTMGSALLTVNHEWERYIENARTYKELEDKVKNRLLDLARQAKDMMD--EKWKE : 388
Hs      : RCPHPVTLGMLMGSVYLVNQNWERYLAEAQGTYEELQREMKSIMDLANDACQLLS--ERYKE : 384
Dmel    : RHPHPVTLGMLMGSAYLVNSNHWERYIENARTYKELEDKVKNRLLDLARQAKDMMD--DQYRQ : 370
      cPhPvsf l Lp W Y e y

```

```

      400          *          420          *          440          *          460
Sc      : DPWLSQLDWTTKPLRLTKKGVPAKQKLPGFPEWYRQLFPSKDTEPKITIKSRIIPILFKLSWENS : 451
Spa     : DPWLSQLDWTTKPLRLTKKGVPAKQKLPGFPEWYRQLFPSKDTEPKITIKSRIIPILFKLSWENF : 452
Sku     : DPWIAQLDWTAKPLRVTKKGIIPAKGQKLPGFPEWYRQLFPSKDTEPKITIKSRIIPILFKLSWENS : 455
Sba     : DPWLSQLDWSAKPLRMTKKGVPAKQKLPGFPEWYRQLFPTSDETEPKITIKSRIIPILFKLSWENS : 454
Sca     : DPWLRQLDWTIKPLRLTKKGVPAKQKLPGFPEWYRQLFPSKTSQPQITIKSRIIPILFKLSWEGH : 449
Cgl     : NPWLRQLDWEIKPLRLTKKGVPAKQKLPGFPEWYRQLFSPKEALPQITMRSRIIPILFKLSWEDS : 459
Kwa     : DPWLRQLDWTIKPLRLTKKGIIPAKGQKLPGFPEWYRQLFPTRLSFPQVSIIRNRLIIPILFKLSWEGE : 456
Klactis : DPWIKQLDWTAVPVKYTKSGEIRKNQKLPGFPEWYRQLFPSKSTRPQITIRTRQVPLFFKLSWEDC : 459
Ego     : DPWYKQLDWTISPLKLTKKGVPAKQKMPGYPAWRSIFGTGTD-KPNITIRSRILIPILFKLSWEGY : 453
Debaryomyc : DPWLSQINWTTIKVQRLKNGEPAANQALTCGYPPEWYRDLFKTVTIEMNISTRTRVTPILLKLSWEGY : 457
Calb    : DPWLSQINWKLKTRKLDGTEYKKTAMTYGPEWYRELFKTS--EMNLTLLKTRITPILLRLKSWEGH : 458
Ppa     : DPWLSQLDWTIDPIRLTKKGEIHKNOKLPGYPNWYKQLIVKN--ELKLTTKTRTSPILLKLAWNGN : 444
Yarrowia : DPWLSQLDWTIKPIKMVKIGERLPKNVFGYPPEWYKALEKSSAGPMHITTKTRITPILLRLKSWEGY : 461
Ani     : DPWLRQLDWSGQEVKMKVGGPRPAARQKKPGMPQWYKDLFSSNTA-INLTVRTRIAPIILLKLSWDTH : 460
aspfum  : DPWLRQLDWSGQEIKLKVGPRPAARQKKPGMPQWYKDLFSSATA-INLTVRTRIAPIILLKLSWDGY : 401
Hca     : DPWLRQLDWSGQEIIRMVKGGPRPAARQKKPGMPQWYKALEFPTSTS-INLSVRTRIAPIILLKLSWLGY : 460
Gze     : DPWMLQLDWSGQEIKMAKGGPRPVARQKKPGMPQWYKDLFQSSDS-INITVTRTRVAPILLKMSWDGY : 459
Mgr     : DPWLSQLDWSGQEIKMVKGPRPAARQKKPGMPQWYKDLFVKNSD-IGITVTRTRIAPIILLRLAWDGH : 460
Ncr     : DPWLSQLDWSGQEIKMAKPERPALNQLKPGYPQWYKDLFVKVPKFINTVRSRIAPILLKLSWEGY : 461
Pan     : DPWMSQLDWSGQEVKMTKGPREYKNQKMPGEPNWRNLFPKARG-INLTVRTRIAPIILLRLSWDGH : 460
Spo     : DPWLRQLDWT--PCNLRYK--LKKATQVIVVPPKWKYCKTEK-AVITAKSRLAPIILLRLKWKKH : 452
cne     : DPWASQLDWSPKNARWSDERESAQPRKSASSPAWLTQISSNHSV-LKSNMSQRLLEPLVLRMSFKGH : 456
Pcr     : DPWLSQLDWTPKEDPWLKEKALKPKARSCTGTQONPRKIPLRX-LDITVNRTRIAPIILLRLSWQGW : 453
Hs      : DPWLWDLWDLQEFKQKAGDEMDQEDLPGHFGWYRKLCPFR---GPSLLSLQMVTPKLMALTWDFG : 447
Dmel   : NTLWLDWDLWSVQELKQLPLTVELKDLPGYPLWYRKLCKRPPAGASEISTGMIAPEKLSLCSWEGY : 436
dpWl qldW          k          pg p Wy l          r P l w

```

```

      *          480          *          500          *          520
Sc      : PVIWSKESGWCEINVPHEQVETYKAKNYVLADSI RTHNLGLQC-GVFKVPHPNGETFNCTNLLTKSY : 516
Spa     : PVIWSKESGWCEINVPHEQVETYKAKNYVLADSI RMQNLGSQP-GVLKVPHPNGETSNCNTNLLTKSY : 517
Sku     : PVIWSKESGWCEIKAPQDQEDAFKAKNYILADSLR MQNQGSPS-DVFKVPHPNGETFNCTNLLTKSY : 520
Sba     : PVIWSKESGWCEIKVSDQAEAFKAKNYVLADSI RSRDLGSSSSDVFKVPHPNGETFNCTNLLTKSY : 520
Sca     : PVFFTKESGWCEPIHKDEFEMFQKKNYVVDSSHSI DDNGDS-QILKVPHPNGETFNCTNLLSKPY : 514
Cgl     : PVIWTEGCGWCEIKPKSQEKQYKDKNYRLADTLVEMGQYKKTDDLFKVPHPNGETFNCTNLLSKPY : 525
Kwa     : PVIWSATDGCWCEAAREERMSELKGNMYLAN-EVSADPG----YVFKIHPHNGQQYNTTTLVSKQY : 517
Klactis : PVVYTNSQGWCEIRCDNSRVEEILAKNYVKAV-YHEEDT-----FKIHPHNGPEFNGSTLTKSPF : 518
Ego     : PVVWNQEHGWCEVVPDANAGTFLKKNYAIAN-DFDRNVG----YTFRIHPHSSAEQRTTTLFSKPY : 514
Debaryomyc : PLLWTDSSGWCEKVPFNEET-IKSMDDKSYPKKFLPQLRDDGNYEFKVPHPPEGGRKCTSLTKSY : 522
Calb    : PLFWVDSQGWCEKVPFDDNEIARLTKKYYKTRDALLTKGKYVFKVPHPNGETFNCTNLLSKLY : 524
Ppa     : PLYWIQTQGWCEKVPKNKTEEYELNFMVMSKGFESIRAEDLKFTVKVPHPDGE SARVTNCTMKSC : 510
Yarrowia : PVVWSDKGGWCEKRVVSEQAHFQAEAKKYKLAT-DTEPHLRDDG-TLFRIPHKDGETARCTNPLAKGY : 525
Ani     : PLIWSDKHGWCEKVPKDKAHQYENQPVVACNEEKNPELHNRK-IFKIPHKDGETARCVSPLAKGY : 525
aspfum  : PLTWSDKHGWCEKVPKDKVKKFENQPVVLDCEEKNELELRNRK-VFKIPHKDGEQARCVNPLAKGY : 466
Hca     : PLVWSDVHGWCEQVPKGEIKNFENQAVAPSNEEKNLSLRDDRE-LFKIPHKDGEFARCVSPLAKGY : 525
Gze     : PLFWSNQYGWCEKRVPRAEAAKYAKLAVECKDEKIKLRDDHA-AFKIPHKDGETARCTNPLAKGY : 524
Mgr     : PLVWSDKYGWCEKRAERNREKYINRQMI ECDDETNEKLRDGG-IFKIPHKDGEFARCVNPLAKGY : 525
Ncr     : PLFWSDDQGWCEKQVPREKAETFIQRQMTFVQDDVDRRLRMDVD-KFKIPHKDGEFARCVNPLAKGY : 526
Pan     : PLFWSDKHGWCEKRVPLADMKKYLDQMKPCDGETVLAVKEDTK-AFKIPHKDGETARCTNPLAKGY : 525
Spo     : PLAWSDTYGVVPSVEKDEIEMLLDQGLVPCS--REEDTKLDYNYIFKVPHKDGEFARCVSPLSKSY : 516
cne     : PVAYLSEHGWCEKMPVGDYFDTHGSPHMLSAS-RLEKLEESYSHVFRIPHKDGEKGNVGNPLSKGF : 521
Pcr     : PLFYSERHGWCEKRVVD-GSDFQTRATP--LTFA-ADEALQES--YSYKIPHKDGEKANVGSPLSKTF : 513
Hs      : PLHYSERHGWGYLVPDNLAKLPTGTTLESAGLYRKHCLQGGKWCFKIPHKDGENSCNVGSPFAKDF : 513
Dmel   : PLHYERQGWCEKLVDP--SEGVDRLPMEQLLCL SASKAESDMACFLKIPHKNGE SFRVGNPLSKDF : 500
P          GW f          fk PH g          K

```

```

      *           540           *           560           *           580           *
Sc      : NHFFKGVILKSESELAHQALQINSSGSYWMASARERIQSOFVVPSSCKFPNEFQSLSAKSSLNNDLA : 582
Spa     : NHFFKGVILKSESELAHQALQINSSGSYWMASARERIQSOFVVPSSCKFPNEFHSLSARSNLNNGDLA : 583
Sku     : NHFFKGVILKSESELAHQALQINSSGSYWMASARERIQSOFVVPSSKVPNEFLSLSAKPAPSSDDLA : 586
Sba     : NHFFKGVILKSESELAHQALQINSSGSYWMASARERIQSOFVVPSSCKFPNEFRSLAAKVPNTADLA : 586
Sca     : IHYFFKNVLTSDSELAHQALNMMSSGSYWMAARERIMSOFVVSQKDFNKEFNPLSNKVIGTPGELG : 580
Cgl     : VHFEEKGILKSESELASTALKMNASGSYWMASARERIMSOFVVPKEDFPEQFLPIKVGKDLKPT-QA : 590
Kwa     : IQFENKGITLTSHEVAQDALEINSSCAYWMSARERVMSOFVVSCKDFPEQFRALDHK----PSDAS : 579
Klactis : MHYFFKGVILKSQKLAHDALQINASGAYWMSARERIMSOFVVSVDNDFKSEFN-----AKTPFDDVG : 579
Ego     : MHFMKDNVILQSEYSLAREALKINSSGSYWTAAARERIKTCHVVSKEDEFDPQFAMVDG----TPSQVG : 576
Debaryomyc : LRYFFSGITLSEYSFAQELTSLNSTASYWGNRQRIMDOFVLYSDANHHKKNKFFDTKV--DSKDMG : 586
Calb    : VRNFEDGITLSEYDYATKILNLNNEASYLGNRRNRKDOFVVYNQO--GKNKFFDTIK--KSKSMG : 586
Ppa     : LGFFKGVILNSQYPLAKDALQMAVASSYWTSSRERIMNOFVVFEDD-----MG : 558
Yarrowia : LGHFFKGVILSSDNPQTKEMFEMSVCSYVWVSNRSRIMSQHCIWDGD-----GVDMGGFG : 579
Ani     : LQYFFERGTLSSQYALAKEALEMNASCYWIISARDRIMGQLVVYENDVVRPA-SPSSKNGDQKLG--- : 587
aspfum  : LQYFFERGTLSSQYALAKEALEMNASCYWIISARDRIMSOMVVYEDQVKQS-GSTGEKSN-RLG--- : 527
Hca     : LQYFFGGKILSSQFALAKEALEMNASCYWIISARDRIMSOMVVYHNDIERLQNKSSSKSTLKQG--- : 588
Gze     : LSYFFSGVILSSEYTYAKALEMNASCYWIISARDRIMSOMVVYDQDLRVMRGSTRGKPKK--KPG : 588
Mgr     : LGYFFSGLLSSEYSYAKALEMNASCYWIISARDRIMSOLVVYENEVSAPTASASTKKSRTKKS : 591
Ncr     : LPYFFKGVILSSEYPYAKALEMNASCYWIISARDRIMSOMVVYEDQLPPSQRFVNKDADSNTPIG : 592
Pan     : LSYFFKGVILSSEYDYAKALEMNASCYWIISARDRIMSOLVVYENLPEPVR--PKDLPEDAPTPG : 589
Spo     : HAYFFEGILQSDYEVAKKALEMSASCYWSSARDRIRSMVVWDKDAELGVPSSVDG-----FG : 575
cne     : VKFTESGELASAAEAADATNMNAFCYWIISRERIMDOMVVYRD-----EFG : 569
Pcr     : IKYAQDGTITSPGDEAKDALDMNAQCSYWIISARDRILKQMVVWDDKLMGFAPGAG-----KWG : 573
Hs      : LPKMDGTLQAG-PGGPRALEINKMISFWRNAHKRISSEMVVWLPRALPRAVIRHPDYDEEG-YG : 577
Dmel    : LNKTAENVLSSGDPSAARVIDIARMMSYWRNRNDRIMGQMVVWLDLSDQLPNEFTG--EKCPPI-YG : 563
      fe g L s      a a l n      syW      r Ri      Q vv

```

```

      600           *           620           *           640           *           660
Sc      : IILPKIVEMGTITRRAVENAWLTASNAKANRIGSELKTQVKAPFGYCFVVGADVSEELWIASLVGD : 648
Spa     : IILPKIVEMGTITRRAVENTWLTASNAKANRIGSELKTQVKAPFGYCFVVGADVSEELWIASLVGD : 649
Sku     : IILPKIVEMGTITRRAVENTWLTASNAKANRIGSELKTQVKAPFGYCFVVGADVSEELWIASLVGD : 652
Sba     : IILPKIVEMGTITRRAVENTWLTASNAKANRIGSELKTQVKAPFGYCFVVGADVSEELWIASLVGD : 652
Sca     : IILPTVIEPMGTITRRAVENTWLTASNAKANRIGSELKAQVKAPFGYCFVVGADVSEELWIASLVGD : 646
Cgl     : IILPQIIEPMGTITRRAVESTWLTASNAKANRIGSELKAQICAPFGYSFVVGADVSEELWIASLVGD : 656
Kwa     : IILPSVIEPMGTITRRASVENTWLTASNAKANRIGSELKAQIQAPFGYAFVVGADVSEELWIASLVSD : 645
Klactis : IILPKLVEPMGTITRRASVENTWLTASNAKANRIGSELKCNVEAPFGYTFVVGADVSEELWIASLVGD : 645
Ego     : AILPKIIEPMGTITRRAVENTWLTASNAKANRIGSELKTKIVAPFGYSFVVGADVSEELWIASLVGD : 642
Debaryomyc : IILPKLCSMGTITRRATENTWLTASNAKANRIGSELKAMIEAPFGYTFVVGADVSEELWIASLVGD : 652
Calb    : IILPFLATMGTITRRATENTWLTASNAKANRIGSELKSLVEAPFGYCFVVGADVSEELWIASLVGD : 652
Ppa     : YILPQIIEPMGTITRRAVENTWLTASNAKANRIGSELKSLIEAPFGYCFVVGADVSEELWIASLVGD : 624
Yarrowia : IILPRIIEPMGTITRRACVENTWLTASNAKANRIGSELKSMIKAPFGYSFVVGADVSEELWIASLVGD : 645
Ani     : YILPQIIEPMGTITRRAVENTWLTASNAKANRIGSELKAMIKAPFGYAFVVGADVSDQLWIASLVGD : 653
aspfum  : FILLPQIIEPMGTITRRAVENTWLTASNAKANRIGSELKAMIKAPFGYVTFVVGADVSDQLWIASLVGD : 593
Hca     : YILPQIIEPMGTITRRAVENTWLTASNAKANRIGSELKAMVRAPFGYVTFVVGADVSEELWIASLVGD : 654
Gze     : FILLPQIIEPMGTITRRAVENTWLTASNAKANRIGSELKAMVKAPFGYCFVVGADVSEELWIASLVGD : 654
Mgr     : FILLPQIIEPMGTITRRAVENTWLTASNAKANRIGSELKAMVKAPFGYCFVVGADVSEELWIASLVGD : 657
Ncr     : FVILPQVIEPMGTITTRAVERTWLTASNAKANRIGSELKAMVRAPFGYVTFVVGADVSEELWIASLVGD : 658
Pan     : FILLPQVIEPMGTITRRAVENTWLTASNAKANRIGSELKAMVRAPFGYVTFVVGADVSEELWIASLVGD : 655
Spo     : IILPFCIIEPMGTITRRAVENTWLTASNAKANRIGSELKAMIRAPFGYTFVVGADVSEELWIASLVGD : 641
cne     : MLLPQVITMGTITRRAVEATWLTASNAKANRIGSELKAMVRAPFGYSFVVGADVSEELWIASLVGD : 635
Pcr     : IILPQVITMGTITRRATEKTWLTASNAKANRIGSELKAMVRAPFGYAFVVGADVSEELWIASLVGD : 639
Hs      : AILPQVVTAGTITTRAVEPTWLTASNAKANRIGSELKAMVQAPFGYTFVVGADVSDQLWIASLVGD : 643
Dmel    : AILPQVVACGTTTRRAVEPTWLTASNAKANRIGSELKAMVQAPFGYRIFVVGADVSDQLWIASLVGD : 629
      i P      mGT TRRAvE tWLTASnak nR GSELk      AP GY fVVGADVSEELWias gd

```

```

*           680           *           700           *           720
Sc       : SIENVHGGTATGWMCLEGTRKNEGTDLHHTKTAQILGCSRNEAKIFENYGRITYGAGAKFASQLLKRFPN : 714
Spa      : SIENVHGGTATGWMCLEGTRKNEGTDLHHTKTAQILGCSRNEAKIFENYGRITYGAGAKFASQLLKRFPN : 715
Sku      : SIENVHGGTATGWMCLEGTRKNEGTDLHHTKTAQILGCSRNEAKIFENYGRITYGAGAKFASQLLKRFPN : 718
Sba      : SIENVHGGTATGWMCLEGTRKNEGTDLHHTKTAQILGCSRNEAKIFENYGRITYGAGAKFASQLLKRFPN : 718
Sca      : SVENIHGGTATGWMCLEGTRKNEGTDLHHTKTAEILGCSRNEAKIFENYGRITYGAGVKFAGTLLKRFPN : 712
Cgl      : SVLRIHGGTPTGWMCLEGTRKAEGSDLHHTKTAQILSCDRNEAKIFENYGRITYGAGVKFASQLLKRFPN : 722
Kwa      : SLQNIHGGTPTGWMCLEGTRKNEGTDLHHTKTAEILGCSRNEAKIFENYGRITYGAGLKFATQLLKRFPN : 711
Klactis  : SIQDIHGATAIGWMCLEGTRKSLGTDLHHTKTAHILGCSRDEAKIFENYGRITYGAGVKFAQLLKRFPN : 711
Ego      : SVENIHGGTATGYMCLGTRKADAGTDMHTKTASILGCSRNEAKIFENYGRITYGAGVRFPAQLLKRFPN : 708
Debaryomyc : SMEMHGSTALGMMTLEGTRKSEKTDLHHTKTAEILGILRNDKAVENYGRITYGAGVKFATRLKQCN : 718
Calb     : SMQIHGGTATGMMTLEGTRKSEKTDLHHTKTAEILGILRNDKAVENYGRITYGAGVKFATRLKQCN : 718
Ppa      : SVQKIHGTTATGMMTLEGTRKNEGTDLHHTKTAAILGISRNEAKIFENYGRITYGAGIKFTTLLKRFPN : 690
Yarrowia : SCQKNHGSTALGMMTLEGTRKSEKTDLHHTKTAAILGISRNEAKIFENYGRITYGAGLKFVFAQLLKRFPN : 711
Ani      : AQLHGGNATGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGVKFAATLLRQFNP : 719
aspfum   : AQLHGGNATGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGVKFAATLLRQFNP : 659
Hca      : AQLHGGNATGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 720
Gze      : ATBKLHGNAIGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 720
Mgr       : ATBKLHGNAIGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 723
Ncr       : ATBKLHGNAIGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 724
Pan       : ATBKLHGNAIGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 721
Spo       : SCQRLHGATAIGMMTLEGTRKSEKTDLHHTKTAAILGVSRDSKAVENYGRITYGAGLKFVFAATLLRQFNP : 707
cne      : SCQGMHGATAIGMMTLEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 701
Pcr       : AQLHGGNATGFMILEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGMRHVVALLLQANF : 705
Hs       : AHSAMHGCTATGMMTLEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 709
Dmel     : AYACEHGATPLGMMTLEGTRKAAAGTDMHSRTAKILGISRNDKAVENYGRITYGAGLKFVFAATLLRQFNP : 695
          f HG a G M LeG K gtdLH TA ilg R AK fNYgRiYGAG kfa LL fnp

```

```

*           740           *           760           *           780           *
Sc       : SLTDEETKKIANKLYENTKGTTRKSKLFLK-KFWYGGSESLFNKLESLAEQETPKTPVLCGGITYS : 779
Spa      : SLTDEETKKIANKLYENTKGTTRKSKLFLK-KFWYGGSESLFNKLESLAEQETPKTPVLCGGITYS : 780
Sku      : SLTDEETKKIANKLYENTKGTTRKSKLFLK-KFWYGGSESLFNKLESLAEQELPKTPVLCGGITYS : 783
Sba      : SLTDEETKKIANKLYENTKGTTRKSKLFLK-KFWYGGSESLFNKLESLAEQETPKTPVLCGGITYS : 783
Sca      : SLSDDEAKKTAIKLYENTKGTTRKSRIFK-KFWYGGSESLFNKLESLAEQDEPKTPVLCGGITYS : 777
Cgl      : TLSEKTAKETAVKLYESTKGTTRKSKVFK-KFWYGGSESLFNKLESLAEQDNPMTPVLCGGITNS : 787
Kwa      : GLSDSEAKSTAELYNSTKGLTKRSHIFK-KFWYGGSESLFNKLESLAEQDQSPMTPVLCGGITYS : 776
Klactis  : MLSEDEAKKTAERLYENTKGTTRKSRIFK-KFWYGGSESLFNKLESLAEQDEPKTPVLCGGITYS : 776
Ego      : TLSEEQTRATAERLYEATKGTTYHSKIFK-KFWYGGSESLFNKLESLAEQDQPRTPVLCGGITVS : 773
Debaryomyc : SLSDVFAELMAKTLYAKTKGQNTSKFLDRMYHGGTESIMFNKLESLAEQEDPKTPVLCGGITDA : 784
Calb     : NITDEEADKVARQLYDSTKGRVAVSKYLLPLRIYGGTESIMFNKLESLAEQDEPKTPVLCGGITDA : 784
Ppa      : ALSDAFAKATANALYATKGG--ISGRYDCKSIYGGSESLFNKLESLAEQDQPRTPVLCGGITAP : 754
Yarrowia : NVTDAFASTLAGQLYTSKGGK---KIYSIGKYSGGTESILFNKLESLAEQDMPETPVLCGGITKA : 773
Ani      : SMSERETQEVASKLYKETKGAKTTRRLSDPFWRGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 785
aspfum   : SMSEKETQEVATNLYRETKGTTRTRRLISEPFWRGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 725
Hca      : SMSEQETKRVAAKLYKETKGTTRTRNKLSDPFWRGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 786
Gze      : NLSEKQTMETASRLYANTKGTNTNRKLLYKPFWRGGTESFVFNKLESLAEQDKARTPVLCGGITTEA : 786
Mgr       : NLSEQETVETASKLYAATKGAKTNRKSLYKPFWRGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 789
Ncr       : SLTEAETTAIAATKLYDATKGAKTNRKSLYKPFWRGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 790
Pan       : SLSEQETLDIAGKLYATKGTNTNRKSLYKPFWRGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 787
Spo       : TLKTAFAKELAKKLYASTKGVKMSKRLQETFWSCGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 773
cne      : KLTKETAGKLANLYKSTKGAQVARNLPVSLWHGGSESYLFNLESLAEQDQPRTPVLCGGITTEA : 767
Pcr       : GMLPEFAQKLANLYASTKGNHRADFFGR-FWGGTESFVFNKLESLAEQDQPRTPVLCGGITTEA : 770
Hs       : RLTTQQAEEKAQQMYAATKGLRYRLSDEGERAKGGTESSEMFNLESLAEQDQPRTPVLCGGITTEA : 775
Dmel     : TFSASFAKAKAMKMFSLTKGKRYRLREEFFPNWGGTESSEMFNLESLAEQDQPRTPVLCGGITTEA : 761
          e A ly TKG w gG ES FN LE A p TPvLg git

```

```

      800          *          820          *          840          *          8
Sc      : LMKKNLRANSFLPSRINWAIQSSGVDYLHLHCCSMFYIIKKYNLEARLCISTHDEIRFLVSEKDKY : 845
Spa     : LMKKNLRANSFLPSRINWAIQSSGVDYLHLHCCSMFYIIKKYNLEARLCISTHDEIRFLVSEKDKY : 846
Sku     : LMKKNLRANSFLPSRINWAIQSSGVDYLHLHCCSMFYIIKKYNLEARLCISTHDEIRFLVSEKDKY : 849
Sba     : LMKKNLRANSFLPSRINWAIQSSGVDYLHLHCCSMFYIIKKYSLDARLCISTHDEIRFLVSEKDKY : 849
Sca     : LMKKNLRANSFLPSRINWAIQSSGVDYLHLHCCSMNYLILKYGLDARLCISTHDEIRFLVSEKDKY : 843
Cgl     : LMKQNLKANSFLPSRVNWAIQSSGVDYLHLHCCSMNYLISMYNIEARLCISTHDEIRFLVADKDKY : 853
Kwa     : LMKRNLTNTFLPSRINWAIQSSGVDYLHLHCCSMNYLILKYSLKARLCISTHDEIRFLVTAEDKY : 842
Klactis : LMKKNLGKDTFLPSRINWAIQSSGVDYLHLHCCSMNYLILKYSLDARLCISTHDEIRFLVSEKDKY : 842
Ego     : LLKRNLGKDTFLPSRINWAIQSSGVDYLHLHFCSMHYLIEKYRLRARCISTHDEIRFLVANEDRY : 839
Debaryomyc : LTIANLNKNYLTSRINWAIQSSGVDYLHLHIVAMEFLKFKFRIDGRLLITVHDEIRFLVVKSEDKY : 850
Calb    : LNVKNLNTNQYMTSRVNWAIQSSGVDYLHLHIIISMEYLIEKYGIDARLCITVHDEIRFLVKEEDKY : 850
Ppa     : LQKANLSTNNFLPSRINWAIQSSGVDYLHLHIIISMDYLIKLFIDARLCITVHDEIRFLVKEEDKY : 820
Yarrowia : LSSQNLKKSSEMTSRVNWAIQSSGVDYLHLHIVVSMYLAAYDIDMRLCITVHDEIRFLVKNEDAY : 839
Ani     : LMRRFINRGSEMTSRINWAIQSSGVDYLHLHIIISMDYLIIRRFNIQARLCITVHDEIRFLVKEEDKY : 851
aspfum  : LMRRFINKGSEMTSRINWAIQSSGVDYLHLHIIISMDYLIIRRFNIDARLCITVHDEIRFLVKEEDKY : 791
Hca     : LMRRFINKGSEMTSRINWAIQSSGVDYLHLHIIISMDYLIIRRFNIDARLCITVHDEIRFLVKEEDKY : 852
Gze     : LMGRYISKGGYLTSRINWAIQSSGVDYLHLHIVVAMDYLIIRRFNIDARLCITVHDEIRFLVTEKDKY : 852
Mgr     : LMSRFMSKGGYLTSRINWAIQSSGVDYLHLHIVSMDYLIIRRFNIDARLCITVHDEIRFLVVDKDKY : 855
Ncr     : LMSRWVSKGGYLTSRINWAIQSSGVDYLHLHIIISMDYLIIRRFNIDARLCITVHDEIRFLVAEEDKY : 856
Pan     : LTKARFLSKGGYLTSRINWAIQSSGVDYLHLHIVSMDYLIIRRFNIACRLCITVHDEIRFLVQEHDKY : 853
Spo     : LSSKNLKNSEMTSRVNWAIQSSGVDYLHLHLLVSMNHLIKKYLEARLCITVHDEIRFLVSSDKDKY : 839
cne     : LRKSYL-NASYLPSRVNWVQSSGVDYLHLHIVSMEYLIIKKYNIQARYLISVHDEIRFLVLAKEEDRY : 832
Pcr     : LSKEYLPGSDYMTSRINWVQSSGVDYLHLHIVSMEYLIIKKYDIKARYLISVHDEIRFLVLAKEEDRY : 836
Hs      : LEPS---QEEFMTSRVNWVQSSGVDYLHLHMLVAMKWLFEFAIDGRFCISTHDEIRFLVREEDRY : 838
Dmel    : LEADTGPQRFLPTRINWVQSSGVDYLHLHMLVSMRWLMSG---HVRFCLSFHDEIRFLVKEELSP : 824
      L          sR NW iQsgVdYLHLH sM l          R i HDE Rylv dky

```

```

      60          *          880          *          900          *          920
Sc      : RAAMALQISNLWTRAMFCQQMGINELPONCAFFSQVDIDSVIRKEVNMDCITPSNKTATPHGEALD : 911
Spa     : RAAMALQISNLWTRAMFCQQMGINELPONCAFFSQVDIDSVIRKEVNMDCITPSNKTATPHGEALD : 912
Sku     : RAAMALQISNLWTRAMFVNKWE----- : 871
Sba     : RAAMALQISNLWTRAMFCQQMGINELPONCAFFSMVDIDSVIRKEVNMDCITPSNKTATPHGEALD : 915
Sca     : KVAMALQISNLWTRAMFCQQMGIDDLPONCAFFSAVDIDHVMRKEVNMDCITLPSNKTATPHGENLD : 909
Cgl     : KLALALQISNLWTRAMFAEQMGIRELPONCAFFSAIDIDKVLKEVDLDCVTPSNEIPIPHGESID : 919
Kwa     : RTAMALQISNLWTRAMFCQQMGIRDLPONCAFFSAVDIDHVMRKEVNMDCITPSNKPPIPHGESID : 908
Klactis : RITLALQISNLWTRSYSEQMCINDLPONCAFFSAVDIDHVMRKEVNMDCITPSNPNAIPIPHGESIG : 908
Ego     : RVALALQVANLWTRAMFCQQMGIDDLPONCAFFSAVDIDHVLKEVMDDCVTPSNQTPVPHGERVT : 905
Debaryomyc : KTALLQISNLWTRAMFCQQMGIRLQSCAFFSEVDIDNLRKEVTLDCVTPSHSNIPIPHGESLD : 916
Calb    : KAALLQISNLWTRAMFCQQMGIRLQSCAFFSEVDIDHVLKEVMDDCVTPSNPTPIPIPHGESLN : 916
Ppa     : RAAYALQISNLWTRAMFCQQMGINEVPOSCAFFSAVDIDHVLKEVMDDCVTPSNPDIPIPHCKSLD : 886
Yarrowia : RAAMALQVSNLWTRAMFCQQSGIDDIPOGCAFFSAVDIDHVLKEVMDDCITPTHDPPIPHGESLD : 905
Ani     : RAALALQVSNLWTRAMFCQQVGINDLPOSCAYFSQVDIDHVLKEVMDDCVTPSHPKIPIPHGESLN : 917
aspfum  : RAAMALQVANLWTRAMFCQQVGINDLPOSCAYFSQVDIDHVLKEVMDDCVTPSHPKIPIPHGESLD : 857
Hca     : RAAMALQVANLWTRAMFCQQVGINDLPOSCAYFSQVDIDHVLKEVMDDCVTPSHPKIPIPHGESLD : 918
Gze     : RTALALQVANLWTRAMFCQQVGINDLPOSCAYFSQVDIDHVLKEVMDDCITPSHQTPPIPHGESID : 918
Mgr     : RAAMALQVANLWTRAMFAEQVGINDLPOSCAYFSQVDIDHVLKEVMDDCITPSHSTPIPHGESLD : 921
Ncr     : RVAMALQIANLWTRAMFAQQVGIQDLPOSCAFFSAVDIDHVLKEVMDDCITPSNPIPIPHGESID : 922
Pan     : RAAMALQISNLWTRAMFAQQVGIHDLPOSCAFFSAVDIDHVLKEVMDDCITPSNPIPIPHGETVD : 919
Spo     : RVAFALQVANLWTRAMFCQRLGINELPOSAVFFSSVDIDHVLKRDVKMDDCVTPSNKVPPIPHGEELT : 905
cne     : RTALALQIANLWTRAMFCQNLGIDDMPOGITFFSAVDIDHVLKEVMDDCITPSHPKVIPIPHGESLD : 898
Pcr     : RAALALQIANLWTRAMFAKLGIDDLPOGVGFFSAVDIDHVLKEVMDDCVTPSQPEPIPHGESLD : 902
Hs      : RAALALQITNLWTRAMFAKLGIDDLPOSAVFFSAVDIDHVLKEVMDDCITPSNPTGIPIPHGESLD : 904
Dmel    : KAALAMHITNLWTRAMFVSRIGLQDLPMVAFFSSVEVDHVLKREKTMDCITPSNPHGIQIPQSLD : 890
      r a alq N WTR f gi pq a fs vdid v rkev dc tps ip ge

```

```

          *           940           *
Sc       : INQLLDKSNKLGKPNIDSKVSYANYREFVFE : 945
Spa     : INQLLVKPNKLGQPNIDSEVSQYTYNYREFVFE : 946
Sku     : ----- : -
Sba     : INQLLTKPNSALGNPNIDSQVSQYPTTFREFVFE : 949
Sca     : INKLLSMEGSRLESSDDTVDVSSLPYKYREFVFS : 943
Cgl     : IVKLLERPECHLELP-ETVDLSKFGFKKREFVFV : 952
Kwa     : ILRLLELPESTLKEASEEIDVSGMPYQPREVFS : 942
Klactis : ILDLIKITNGYLGEPTDAVDVSQFPYRHRESVFR : 942
Ego     : ITDLLARPDAALNNPNPAVNVADYPVVQRQPVLA : 939
Debaryomyc : IHKILEKCKDEMIGAPKASKLSKIEYCNRDRVIS : 950
Calb    : IGQLLLKQCQHGLADKNKPLTLRTTKYSPRQPIIN : 950
Ppa     : IYQLLQQED--IKGADRTMHLNDVHYRKRTFVIE : 918
Yarrowia : IEGLIARKDARLEDATIDHSIDKFPYTSRTPVLA : 939
Ani     : ITQLLDKNDAAARLDPSPTS--RKYTYTPRKTVMS : 949
aspfum  : ISQLLDKDQEAYLDPSVPS--PKYTYTPRESVMA : 889
Hca     : IENLLKKGFEAQLDPSPTS--LKYHYTPRKPVLS : 950
Gze     : INTLLEKGEVARLDPAPDAKHAHIAYTPRVIME : 952
Mgr     : IKTLLEKGDAAFLDPAPRA-HAGIEYTPRKSVME : 954
Ncr     : IFQILEKGDDAKLDDSPQSQYANIPYTPRVVMQ : 956
Pan     : INQLLSKGLARLDPTPDPTYAHIKYVPRKPVME : 953
Spo     : IESVLEKLEQS-----GQSLEPLEQIQC : 928
cne     : IISLLEKIPRGDLGTPDDLQPP-TDIKPPVALFP : 931
Pcr     : IYKVLEKTNGGSLHPDPMIEEP-PR----- : 926
Hs      : IYQIIELTKG-----SLEKTSQPGP : 924
Dmel    : VAEAIEKAGGN-----DVSQWDWIKK : 911
          i   l                               r

```

6.6 Phylogeny of eukaryotic porins and porin motifs

All of the work in section 6.6 has been published in Young, M. J., Bay, D. C., Hausner, G. and Court, D. A., (2007) *BMC Evol Biol* 7: 31. MJY and DCB contributed equally to this work. MJY and DCB collected the sequence data and created the conceptual translations where necessary. MJY and GH carried out the phylogenetic analysis and DCB and DAC performed and analyzed the secondary structure predictions. DCB performed the intron/exon analysis. All authors participated in the design of the study, data analysis and the drafting of the manuscript and read and approved the final manuscript. All referencing may be found in this paper. The following sections are copied from the paper, with permission of the authors, as required by BMC Evolutionary Biology.

6.6.1 Introduction

Mitochondrial porins, or voltage-dependent anion-selective channels (VDAC) allow the passage of small molecules across the mitochondrial outer membrane, and are involved in complex interactions regulating organellar and cellular metabolism. More recently porin has been implicated in apoptosis initiation. Porin proteins are predicted to traverse the mitochondrial outer membrane as a series of β -strands that form a β -barrel. Numerous organisms possess multiple porin isoforms, and initial studies indicated an intriguing evolutionary history for these proteins and the genes that encode them. Specialization of porin function is suggested by developmental regulation or tissue-specific differential expression of human, mouse and plant isoforms.

In this work, the wealth of recent sequence information was used to perform a comprehensive analysis of the evolutionary history of mitochondrial porins. Fungal porin

sequences were well represented, and newly-released sequences from stramenopiles, alveolates, and seed and flowering plants were analyzed. In total 244 VDAC protein sequences were analyzed. A combination of Neighbour-Joining (NJ) and Bayesian methods was used to determine phylogenetic relationships among the proteins. The aligned sequences were also used to reassess the validity of previously described eukaryotic porin motifs and to search for signature sequences characteristic of VDACs from plants, animals and fungi.

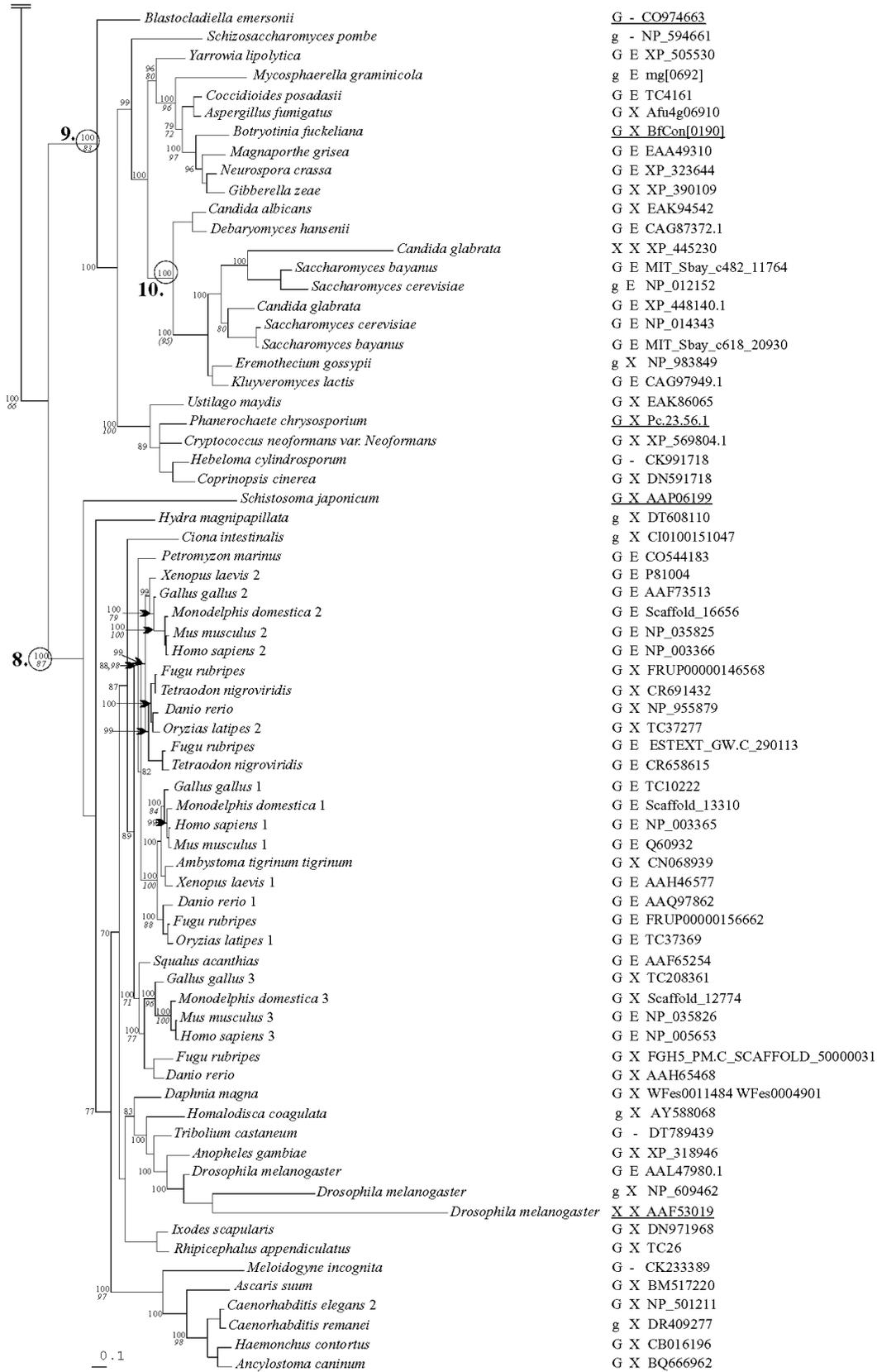
6.6.2 Results and discussion

6.6.2.1 Phylogenetic history of mitochondrial porins

A phylogenetic analysis of the aligned data set consisting of 244 VDAC proteins was carried out by NJ analysis and can be found in [Additional File 2] of Young *et al.* 2007. A subset of 141 porins from the 244 data set was analyzed by both the NJ and Bayesian algorithms (Figures 6.4 and 6.5). The VDAC amino acid sequences from members of the three crown groups (plants, animals, and fungi) formed monophyletic groupings and the branching patterns suggest that the animal and fungal porins are derived from a common ancestor. Essentially the evolution of the VDAC sequences follows the expected pattern for a highly conserved sequence, as the positions of the crown group taxa within the tree correspond in part to the expected phyletic positions based on rDNA sequences, which suggests that fungi and the metazoans share a common ancestor. The Stramenopile *Phytophthora sojae* porin sequences were used as outgroups

Figure 6.4. Evolutionary history of eukaryotic porin sequences – stramenopiles and plants. The phylogenetic tree, continued in Figure 3, is based on NJ and Bayesian analysis of 141 VDAC sequences. Stramenopile sequences were used as the out group. Levels of confidence of the nodes are only provided if support is above 66 %. The numbers are based on posterior probability values generated by Bayesian analysis and on bootstrap analysis in combination with NJ analysis (*italics*). The presence of the GLK (G) and Eukaryotic porin motifs (E) are indicated towards the right of the phylogenetic tree. Note among the plants the GLK domain appears as the STK (S) motif; an X indicates the absence of the motif. Lower case g designates a G-any-K or G-any-R, where “any” refers to any other amino acid. The lower case s indicates S-any-K or S-any-R. The minus (-) indicates that the sequence was incomplete and thus the GLK and eukaryotic porin motifs could not be identified. Nodes designated by a number (1-10) are discussed within the text. Underlined accession numbers are those of VDACs that do not contain the signature motif identified in this study.

Figure 6.5. Evolutionary history of eukaryotic porin sequences – fungi and animals. The portion of the phylogenetic tree, described in Figure 1, containing the fungal and animal data is presented. Symbols and notations are as described for Figure 1.



in our analysis as the emergence of the Stramenopiles are viewed as a basal event to the evolution of plants, metazoans and fungi.

Plant mitochondrial porins: Angiosperms, Eudicots

Among the Eudicots, mitochondrial porins appear to have a rather complex evolutionary history. Based on the available data, there are potentially up to five paralogs for some members of this group of flowering plants, including *Glycine max* and *Lotus japonicus*. In cases where whole genome data are unavailable, it is not clear whether these represent allelic variants or different loci.

Many eudicot VDACs can be derived from one node (Figure 6.4, node 1); thereafter the phylogenetic tree suggests that several gene duplications have occurred, giving rise to some of the recent eudicot VDAC paralogs. The current annotation for plant VDACs is very confusing and needs to be addressed in future efforts. Knowing the evolutionary history of the various VDACs, combined with the completion of more plant genome projects in the near future should facilitate this goal.

Plant mitochondrial porins: Angiosperms, Monocots

Within the monocots grouped by node 2 (Figure 6.4), three clades can be recognized that suggest that at least three VDAC paralogs have evolved. The three monocot VDAC paralogs can be derived from a node that received 100% Bayesian posterior probability support and 89% bootstrap support in NJ analysis, suggesting these three paralogs have a common ancestor.

Plant mitochondrial porins: Gymnosperms

Basal to node 3 (Figure 6.4) that joins the monocot and eudicot VDACs described above, a branch emerges that unites VDACs from the Gymnosperms (Seed Plants, node 4, Figure 6.4). Although there were only a limited number of sequences available, there is an indication that more than one form of mitochondrial porin is present within some members (*Welwitschia mirabilis*, and *Pinus conifera*), suggesting that VDAC paralogs also evolved within this group of plants.

A clade composed of both seed and flowering plant porins

Attached to a node (Figure 6.4, node 5) that is “basal” to the Angiosperm and Gymnosperm VDACs described above is a clade that includes mitochondrial porins from both monocots and eudicots, and the one available example of a seed plant VDAC. This clade could represent the most ancient forms of VDAC that were present in the common ancestor that gave rise to the seed and flowering plants. Again within this clade there appears to be evidence of paralogs, as more than one form of VDAC from *Arabidopsis thaliana*, *Lotus sp.*, *Glycine max*, *Medicago sp.*, and *Lycopersion sp.* are grouped within this cluster. This clade includes two groupings (nodes 6 and 7) that accommodate the latter paralogs. Also potentially allied to each of these two groupings are monocot sequences. The *Pinus conifera* porin sequence appears to be situated between the two groupings.

Overall within the land plants examined, VDAC paralogs appear to have evolved numerous times independently within different plant lineages. But ultimately one node appears to unite all the land plant porin sequences. Green algal sequences (Chlorophyta), and a Desmid sequence (Charophyta) are basal to the land plant node.

Metazoan mitochondrial porins

The evolution of metazoan VDAC genes (see node 8, Figure 6.5) has already been addressed by others. Essentially within the Chordates, the porin gene family consists of three genes (VDAC 1, 2, and 3) and these appear to be a classic example of a set of paralogous genes. Isoforms in the VDAC1 and VDAC2 clades are more closely related to each other than to those in the VDAC3 group. For the basal branching members of the Phylum Chordata: tunicates (*Molgula tectiformis*) and the Urochordata (*Ciona intestinalis*), only one porin gene could be identified [see Additional File 1 and 2 Young *et al* 2007]. Thus, the Chordate porin paralogs probably evolved within the vertebrate lineage. In general, the vertebrate VDAC3 clade appears to contain the earliest branching porin sequences. For *Petromyzon marinus* (parasitic marine lamprey) one of the oldest known taxa of living vertebrates (Hyperoartia), evidence for more than one porin gene was not recovered, although the entire genome sequence is not yet available.

Within the Phylum Arthropoda, *Drosophila melanogaster* appears to have at least four porin genes (data not shown), of which one set might represent co-orthologs, the result of a lineage specific duplication event. However, a porin gene from *D. pseudoobscura* is derived from the same node that includes a potential set of co-orthologs (data not shown) and these *D. pseudoobscura* VDACs could represent inparalogs. The two species of *Drosophila* (*melanogaster* and *pseudoobscura*) appear to have VDACs with the phylogenetic pattern expected for true paralogs; one paralog was then tandemly duplicated twice in *D. melanogaster* and once in *D. pseudoobscura*.

Fungal mitochondrial porins

Within the mycota the porin sequence phylogeny (node 9, Figure 6.5) follows the expected pattern, with the Chytrid sequence being the basal member of this group, the Ascomycota and Basidiomycota sequences forming monophyletic groupings and the ascomycetous yeast branching early within the ascomycete lineage. With respect to gene duplication events, it appears that among the Saccharomycetales a lineage is present that displays the presence of paralogs (node 10, Figure 6.5). *Candida glabrata*, and the *Saccharomyces spp.* appear to have at least two forms of the VDAC gene. The putative VDAC2 of *C. glabrata* is highly degenerate and is not present in the cluster of *Saccharomyces* VDAC2 sequences. It is of interest that we could not detect paralogs for the other yeast-like fungi, including *Eremothecium gossypii*, for which the entire genome has been sequenced. Recent genomic work on *S. cerevisiae* and allied species suggest that these organisms experienced genome duplication events during their evolution. It is interesting to speculate that the appearance of VDAC paralogs coincides with this genome duplication event, and that for some unknown reason both copies of the porin gene were maintained.

6.6.2.2 Phylogenetic history of porin paralogs

One question we tried to address in this analysis was the origin of VDAC paralogs. Did VDAC gene duplications arise early in the evolution of the eukaryotes or did gene duplications occur independently in different evolutionary lineages? If the first scenario applies it would suggest that VDAC duplication events in part paralleled the requirement for more specialized forms of VDACS as eukaryotes evolved into more complex multicellular or multi-tissue forms. Examples of such “ancient paralogs” are the genes for elongation factors involved in translation, and to a lesser extent the globin gene

paralogs of the metazoans. Alternatively, VDAC paralogs may have evolved independently in the different eukaryotic lineages.

Paralogs can arise through gene duplication, which in turn can be result of genome duplication events (polyploidy), segmental chromosome duplication as a result of unequal crossover events, or gene duplication events as a result of "retrotransposition" events. Recent comparative genomic analysis suggest that all of the above have occurred during the evolution of the eukaryotic crown groups, and several examples are revealed by the analysis of VDAC genes. For example, the presence of two versions of VDAC in *Saccharomyces*, but not other *Saccharomycetales* fungi such as *Kluyveromyces lactis* (Figure 6.5), is in agreement with the genome duplication that is postulated for *Saccharomyces*. Segmental duplication is seen in *D. melanogaster* and *D. pseudoobscura*, and appears to have led to highly divergent forms of VDAC. Evidence for retrotransposition has not been seen; in the limited data set available, paralogs both lacking and containing introns have not been identified.

Our analysis shows that paralogs have appeared several times during the evolution of VDACS from plants, metazoans, and fungi. This observation suggests that there are no "ancient" paralogs within this gene family.

6.6.2.3 Signature motifs for eukaryotic porins

The Prosite database carries one "signature" motif for eukaryotic porins (PS00558); this motif was derived from 30 animal, plant and fungal porins, including members of all major VDAC classes in these groups. To determine the universality of this motif, the current collection of porin sequences were analyzed. Of the 236 plant, animal and fungal sequences in the present analysis, 28 lack the complete C-terminal

segment where this motif is found. The signature motif is not found in 92 of the remaining 208 sequences, indicating that it is not universal (Figures 6.4 and 6.5), as noted by others. In particular, the motif is not detected in several animal sequences, such as zebrafish (DrerAAH654), nor in the limited number of stramenopile sequences available (Figure 6.4).

Analysis of the aligned regions encompassing the eukaryotic signature motif did not reveal a motif common to all eukaryotic porins; this finding is not surprising given the low degree of sequence identity across the spectrum of sequences analyzed herein. Use of the PRATT algorithms suggested that the N-terminal regions of the proteins contain sequences suitable for pattern analysis. Manual comparisons of the amino-terminal regions of the protein sequences visualized in GeneDoc did reveal motifs common to most members of the largest phylogenetic groups: plants, fungi and animals (Table 6.3). The limited number of stramenopile, charophyte and rhodophyte sequences did not allow identification of sequence patterns unique to this subgroup.

The fungi represent one of the smaller groups in terms of numbers of representative sequences [38], but the diversity in the group made it the most difficult in terms of identifying a consistent motif. The final sequence motif (Table 6.3) contains conserved residues in the N-terminal α -helix, and is anchored with a short conserved sequence around 50 residues away, which is part of the fourth putative β -strand. The only fungal VDAC sequence that did not contain this motif was that of *Botryotinia fuckeliana*. Updated sequence information may resolve this issue, as the *B. fuckeliana* VDAC sequence was derived from EST data. This fungal VDAC motif was somewhat more

Table 6.3. Distribution of mitochondrial porin motifs in the sequences in the current analysis and public protein sequence databases.

Motif	Total Sequences Analyzed (this work)	Hits	Misses	Insufficient sequence data	Hits in Swiss-Prot TrEMBL and PDB ^a	Expected random matches ^b
Animal Motif ^c	93 animal	76	4	13	81 (including 7 splice variants)	2.9 ^d (84)
Plant Motif ^e	105 plant	90	11	4	43 (splice variants not identified)	2.1e-14
Fungal Motif ^f	38 fungal	33	1	4	18 (splice variants not identified)	4.4 ^d (26)
Totals for this work	236	199	16	21	142	
PS00558 ^g	93 animal	42	41	10	61 animal sequences	6.3e-02
PS00558	105 plant	59	35	11	30 plant sequences	as above
PS00558	38 fungal	20	15	3	11 fungal sequences	as above
PS00558	5 stramenopiles	0	1	4	0 stramenopile sequences	as above
Totals for PS00558	241	121	92	28	102 (94 without splice variants)	

Motif	Total Sequences Analyzed (this work)	Hits	Misses	Insufficient sequence data	Hits in Swiss-Prot TrEMBL and PDB ^a	Expected random matches ^b
GLK	93 animal	82	11 (2 GXX/R ^h)	0	28 ⁱ	17478
GLK	105 plant	74	30 (11 GXX/R) 15 SXX	1	14	as above
GLK	38 fungal	30	8 (6 GXX/R)	0	2	as above
GLK	5 stramenopiles	5	0	0	0	as above
Totals for GLK	241	192	47 (19 GXX/R)	1	44 (of 50 VDAC sequences) ^j	

^a sequences in UniProtKB/Swiss-Prot (release 50.6), UniProtKB/TrEMBL (release 33.6), PDB (14-Sep-2006) databases

^b approximate number of expected random matches in Swiss-Prot release 41 (122564 sequences)

^c animal motif: <-X(0,30)-[PTSA]-X(1,2)-[YF]-X-[DE]-[ILVF]-[AG]-[KR]-X-[AST]-[KR]-[DE]-[ILV]-[FYST].; was not found in any plant or fungal VDAC sequences

^d expected random matches could not be calculated for constrained patterns; value was obtained using the unconstrained motif; in parentheses is the number of hits using the unconstrained motif

^e plant motif derived for Viridiplantae: [YF]-X-[DE]-[ILV]-G-[KR]-[KR]-[APST]-[KR]-D-[IL]-L-X-[KR]-D-[FHY]-X(4)-K-[FL]-[CNST]-X(4)-[ANST]-X(2)-G-X(2)-[FILV]-X-[ASTV]-[AST]-[AGS]-X(3)-[ADGNS].; was not found in any animal or fungal VDAC sequences

^f fungal motif: <-X(0,10)-[STIMLQ]-X(1,3)-[PL]-X(1,3)-[WYF]-X-[DEAG]-[ILV]-X-[RK]-X(3)-[DG]-X(0,1)-[ILV]-X(48,53)-[ILV]-X(2)-[ST]-Q-Achouak, *et al.*, -[WL].; was not found in any plant or animal VDAC sequences

^g PS00558 (EUKARYOTIC_PORIN) [PROSITE (release 19.29)] motif: [YH]-x(2)-D-[SPCAD]-x-[STA]-x(3)-[TAG]-[KR]-[LIVMF]-[DNSTA]-[DNS]-x(4)-[GSTAN]-[LIVMA]-x-[LIVMY]

^h where X represents any amino acid

ⁱ G XK/GXR included in total

^j the description (DE) filter VDAC was used due to the common occurrence of short sequences (G-L-K); this filter does not work on PDB sequence data, so only the Swiss-Prot and TrEMBL databases were searched

effective at identifying putative VDAC sequences in the public databases than the eukaryotic signature motif (Table 6.3).

The N-terminal animal VDAC motif derived in this work (Table 6.3) occurs once in each of 76 of the 80 animal VDAC sequences that had a complete N-terminal sequence in our database. Four sequences could not be accommodated without greatly reducing the specificity of the pattern, that of *Schistosoma japonicum*, and *Drosophila melanogaster* CG31722-PA and CG31722-PB and the ortholog from *D. pseudoobscura* (Dp-22). These *Drosophila* sequences diverge significantly from those of all porins in this study [see Additional File 2] and possess long N-terminal extensions. The animal VDAC motif detects 81 sequences in the public databases, including seven spliceoforms. The position from the N-terminus of the protein was allowed to vary between 0 and 30 residues, to allow detection of human VDAC2 (P45880), in which the motif begins 24 residues from the predicted N-terminus.

The plant VDAC motif derived from this study (Table 6.3) was identified in all but eleven of the plant sequences used. All of the latter VDACs have related sequences in the N-terminal region, but their inclusion led to a consensus sequence with significantly reduced specificity. Both VDAC sequences from the green algae *Chlamydomonas reinhardtii* lack the motif described above. Within the land plants, it is absent from a small group of sequences located within a cluster of sequences united by node 2 (Figure 6.4). Two closely related sequences within the cluster unified by node 5 (Figure 6.4) also lack the motif. Three additional sequences shown in [Additional File 2] also lack the plant-derived VDAC motif, namely those derived from ESTs from a club moss *Selaginella moellendorffii*, and from *Pinus taeda* (CO49625878), and from genomic

DNA of *Sorghum bicolor* (TC94332). The plant motif is not contained within the VDAC sequences from organisms of the deeper branches of the tree: the red algae (*Galdieria sulphuraria* and *Cyanidioschyzon merolae*), and the Desmid *Chlosterium peracerosum-strigosum-littorale*. The plant VDAC motif detected 43 sequences in the public databases, and anchoring to the N-terminus was not necessary to maintain specificity.

Another highly conserved sequence in VDAC is the glycine-leucine-lysine (GLK) motif, which was initially suggested to be an ATP binding site; replacement of the lysine residue with glutamate leads to porins that retain ATP binding, but are cation-selective, rather than anion selective in artificial bilayers.

The GLK sequence is present in the majority of sequences in the current survey (Table 6.3; Figures 6.1 and 6.2). Variations containing chemically similar residues are also found (G-any-R, G-any-K, where “any” is any amino acid). In the fungi and animals, the VDAC sequences lacking this motif are scattered throughout the phylogenetic tree, and there is no apparent link between the presence of this motif and the eukaryotic signature sequence. In the plants there are two clusters of VDAC sequences in which GLK is replaced by a version of STK (Figure 6.4, node 5). Also of note, all of the organisms that contain an STK-bearing isoform also have one with a GLK motif, suggesting perhaps a separation of function in the two types of isoforms.

In conclusion, sequence motifs characteristic of the members of the crown groups of organisms were identified.

References

- Alexander, C., *et al.*, (2000) OPA1, encoding a dynamin-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28. *Nat Genet* **26**: 211-5.
- Azpiroz, R. and Butow, R. A., (1993) Patterns of mitochondrial sorting in yeast zygotes. *Mol Biol Cell* **4**: 21-36.
- Barrientos, A., (2003) Yeast models of human mitochondrial diseases. *IUBMB Life* **55**: 83-95.
- Barrientos, A., Korr, D. and Tzagoloff, A., (2002) Shy1p is necessary for full expression of mitochondrial COX1 in the yeast model of Leigh's syndrome. *Embo J* **21**: 43-52.
- Barth, P. G., Scholte, H. R., Berden, J. A., Van der Klei-Van Moorsel, J. M., Luyt-Houwen, I. E., Van 't Veer-Korthof, E. T., Van der Harten, J. J. and Sobotka-Plojhar, M. A., (1983) An X-linked mitochondrial disease affecting cardiac muscle, skeletal muscle and neutrophil leucocytes. *J Neurol Sci* **62**: 327-55.
- Baruffini, E., Ferrero, I. and Foury, F., (2007a) Mitochondrial DNA defects in *Saccharomyces cerevisiae* caused by functional interactions between DNA polymerase gamma mutations associated with disease in human. *Biochim Biophys Acta* **1772**: 1225-35.
- Baruffini, E., Lodi, T., Dallabona, C. and Foury, F., (2007b) A single nucleotide polymorphism in the DNA polymerase gamma gene of *Saccharomyces cerevisiae* laboratory strains is responsible for increased mitochondrial DNA mutability. *Genetics* **177**: 1227-31.
- Baruffini, E., Lodi, T., Dallabona, C., Puglisi, A., Zeviani, M. and Ferrero, I., (2006) Genetic and chemical rescue of the *Saccharomyces cerevisiae* phenotype induced by mitochondrial DNA polymerase mutations associated with progressive external ophthalmoplegia in humans. *Hum Mol Genet* **15**: 2846-55.
- Bebenek, K. and Kunkel, T. A., (2004) Functions of DNA polymerases. *Adv Protein Chem* **69**: 137-65.
- Becerra, M., Lombardia-Ferreira, L. J., Hauser, N. C., Hoheisel, J. D., Tizon, B. and Cerdan, M. E., (2002) The yeast transcriptome in aerobic and hypoxic conditions: effects of hap1, rox1, rox3 and srb10 deletions. *Mol Microbiol* **43**: 545-55.
- Beese, L. S., Derbyshire, V. and Steitz, T. A., (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **260**: 352-5.

- Bendich, A. J., (1996) Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. *J Mol Biol* **255**: 564-88.
- Bernad, A., Blanco, L., Lazaro, J. M., Martin, G. and Salas, M., (1989) A conserved 3'-5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell* **59**: 219-28.
- Blanco, L., Bernad, A. and Salas, M., (1991) MIP1 DNA polymerase of *S. cerevisiae*: structural similarity with the *E. coli* DNA polymerase I-type enzymes. *Nucleic Acids Res* **19**: 955.
- Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P. and Boeke, J. D., (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**: 115-32.
- Braithwaite, D. K. and Ito, J., (1993) Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res* **21**: 787-802.
- 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**: 121-33.
- Budd, M. E., Sitney, K. C. and Campbell, J. L., (1989) Purification of DNA polymerase II, a distinct DNA polymerase, from *Saccharomyces cerevisiae*. *J Biol Chem* **264**: 6557-65.
- Carrodeguas, J. A. and Bogenhagen, D. F., (2000) Protein sequences conserved in prokaryotic aminoacyl-tRNA synthetases are important for the activity of the processivity factor of human mitochondrial DNA polymerase. *Nucleic Acids Res* **28**: 1237-44.
- Carrodeguas, J. A., Kobayashi, R., Lim, S. E., Copeland, W. C. and Bogenhagen, D. F., (1999) The accessory subunit of *Xenopus laevis* mitochondrial DNA polymerase gamma increases processivity of the catalytic subunit of human DNA polymerase gamma and is related to class II aminoacyl-tRNA synthetases. *Mol Cell Biol* **19**: 4039-46.
- Carrodeguas, J. A., Theis, K., Bogenhagen, D. F. and Kisker, C., (2001) Crystal structure and deletion analysis show that the accessory subunit of mammalian DNA polymerase gamma, Pol gamma B, functions as a homodimer. *Mol Cell* **7**: 43-54.
- Cavalier-Smith, T., (2001) What are fungi?
- Chan, S. S., Longley, M. J. and Copeland, W. C., (2006) Modulation of the W748S mutation in DNA polymerase gamma by the E1143G polymorphism in mitochondrial disorders. *Hum Mol Genet* **15**: 3473-83.

- Chen, O. S. and Kaplan, J., (2000) CCC1 suppresses mitochondrial damage in the yeast model of Friedreich's ataxia by limiting mitochondrial iron accumulation. *J Biol Chem* **275**: 7626-32.
- Chen, X. J., Wang, X., Kaufman, B. A. and Butow, R. A., (2005) Aconitase couples metabolic regulation to mitochondrial DNA maintenance. *Science* **307**: 714-7.
- Cheng, X., Dunaway, S. and Ivessa, A. S., (2007) The role of Pif1p, a DNA helicase in *Saccharomyces cerevisiae*, in maintaining mitochondrial DNA. *Mitochondrion* **7**: 211-22.
- Cho, J. H., Ha, S. J., Kao, L. R., Megraw, T. L. and Chae, C. B., (1998) A novel DNA-binding protein bound to the mitochondrial inner membrane restores the null mutation of mitochondrial histone Abf2p in *Saccharomyces cerevisiae*. *Mol Cell Biol* **18**: 5712-23.
- Claros, M. G. and Vincens, P., (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* **241**: 779-86.
- Contamine, V. and Picard, M., (2000) Maintenance and integrity of the mitochondrial genome: a plethora of nuclear genes in the budding yeast. *Microbiol Mol Biol Rev* **64**: 281-315.
- Copeland, W. C., (2008) Inherited mitochondrial diseases of DNA replication. *Annu Rev Med* **59**: 131-46.
- D'Aurelio, M., Gajewski, C. D., Lin, M. T., Mauck, W. M., Shao, L. Z., Lenaz, G., Moraes, C. T. and Manfredi, G., (2004) Heterologous mitochondrial DNA recombination in human cells. *Hum Mol Genet* **13**: 3171-9.
- Diffley, J. F. and Stillman, B., (1991) A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc Natl Acad Sci U S A* **88**: 7864-8.
- Donahue, S. L., Corner, B. E., Bordone, L. and Campbell, C., (2001) Mitochondrial DNA ligase function in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **29**: 1582-9.
- Doublie, S., Tabor, S., Long, A. M., Richardson, C. C. and Ellenberger, T., (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**: 251-8.
- Dujon, B., (1981) Mitochondrial genetics and functions.
- Dujon, B., *et al.*, (2004) Genome evolution in yeasts. *Nature* **430**: 35-44.
- Dunstan, H. M., Green-Willms, N. S. and Fox, T. D., (1997) In vivo analysis of *Saccharomyces cerevisiae* COX2 mRNA 5'-untranslated leader functions in mitochondrial translation initiation and translational activation. *Genetics* **147**: 87-100.

Eom, S. H., Wang, J. and Steitz, T. A., (1996) Structure of Taq polymerase with DNA at the polymerase active site. *Nature* **382**: 278-81.

Ephrussi, B., Jakob, H. and Grandchamp, S., (1966) Etudes Sur La Suppressivité Des Mutants à Déficience Respiratoire De La Levure. II. Etapes De La Mutation Grande En Petite Provoquée Par Le Facteur Suppressif. *Genetics* **54**: 1-29.

Fan, L. and Kaguni, L. S., (2001) Multiple regions of subunit interaction in *Drosophila* mitochondrial DNA polymerase: three functional domains in the accessory subunit. *Biochemistry* **40**: 4780-91.

Fan, L., Kim, S., Farr, C. L., Schaefer, K. T., Randolph, K. M., Tainer, J. A. and Kaguni, L. S., (2006) A novel processive mechanism for DNA synthesis revealed by structure, modeling and mutagenesis of the accessory subunit of human mitochondrial DNA polymerase. *J Mol Biol* **358**: 1229-43.

Fan, L., Sanschagrin, P. C., Kaguni, L. S. and Kuhn, L. A., (1999) The accessory subunit of mtDNA polymerase shares structural homology with aminoacyl-tRNA synthetases: implications for a dual role as a primer recognition factor and processivity clamp. *Proc Natl Acad Sci U S A* **96**: 9527-32.

Felsenstein, J., (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.

Felsenstein, J., (2002) PHYLIP (Phylogeny Inference Package) version 3.6a. Distributed by the author.

Feuermann, M., Francisci, S., Rinaldi, T., De Luca, C., Rohou, H., Frontali, L. and Bolotin-Fukuhara, M., (2003) The yeast counterparts of human 'MELAS' mutations cause mitochondrial dysfunction that can be rescued by overexpression of the mitochondrial translation factor EF-Tu. *EMBO Rep* **4**: 53-8.

Filee, J., Forterre, P., Sen-Lin, T. and Laurent, J., (2002) Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. *J Mol Evol* **54**: 763-73.

Fontanesi, F., Palmieri, L., Scarcia, P., Lodi, T., Donnini, C., Limongelli, A., Tiranti, V., Zeviani, M., Ferrero, I. and Viola, A. M., (2004) Mutations in AAC2, equivalent to human adPEO-associated ANT1 mutations, lead to defective oxidative phosphorylation in *Saccharomyces cerevisiae* and affect mitochondrial DNA stability. *Hum Mol Genet* **13**: 923-34.

Foury, F., (1989) Cloning and sequencing of the nuclear gene MIP1 encoding the catalytic subunit of the yeast mitochondrial DNA polymerase. *J Biol Chem* **264**: 20552-60.

- Foury, F., (1997) Human genetic diseases: a cross-talk between man and yeast. *Gene* **195**: 1-10.
- Foury, F. and Kucej, M., (2002) Yeast mitochondrial biogenesis: a model system for humans? *Curr Opin Chem Biol* **6**: 106-11.
- Foury, F., Roganti, T., Lecrenier, N. and Purnelle, B., (1998) The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. *FEBS Lett* **440**: 325-31.
- Foury, F. and Vanderstraeten, S., (1992) Yeast mitochondrial DNA mutators with deficient proofreading exonucleolytic activity. *Embo J* **11**: 2717-26.
- Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R. and Steitz, T. A., (1988) Cocystal structure of an editing complex of Klenow fragment with DNA. *Proc Natl Acad Sci U S A* **85**: 8924-8.
- Gelperin, D. M., *et al.*, (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev* **19**: 2816-26.
- Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K. and Weissman, J. S., (2003) Global analysis of protein expression in yeast. *Nature* **425**: 737-41.
- Gietz, R. D., Triggs-Raine, B., Robbins, A., Graham, K. C. and Woods, R. A., (1997) Identification of proteins that interact with a protein of interest: applications of the yeast two-hybrid system. *Mol Cell Biochem* **172**: 67-79.
- Gietz, R. D. and Woods, R. A., (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* **350**: 87-96.
- Glerum, D. M. and Tzagoloff, A., (1994) Isolation of a human cDNA for heme A:farnesyltransferase by functional complementation of a yeast *cox10* mutant. *Proc Natl Acad Sci U S A* **91**: 8452-6.
- Goffeau, A., *et al.*, (1996) Life with 6000 genes. *Science* **274**: 546, 563-7.
- Graves, T., Dante, M., Eisenhour, L. and Christianson, T. W., (1998) Precise mapping and characterization of the RNA primers of DNA replication for a yeast hypersuppressive petite by in vitro capping with guanylyltransferase. *Nucleic Acids Res* **26**: 1309-16.
- Gray, H. and Wong, T. W., (1992) Purification and identification of subunit structure of the human mitochondrial DNA polymerase. *J Biol Chem* **267**: 5835-41.

- Graziewicz, M. A., Longley, M. J., Bienstock, R. J., Zeviani, M. and Copeland, W. C., (2004) Structure-function defects of human mitochondrial DNA polymerase in autosomal dominant progressive external ophthalmoplegia. *Nat Struct Mol Biol* **11**: 770-6.
- Graziewicz, M. A., Longley, M. J. and Copeland, W. C., (2006) DNA polymerase gamma in mitochondrial DNA replication and repair. *Chem Rev* **106**: 383-405.
- Grimes, G. W., Mahler, H. R. and Perlman, P. S., (1974) Letter: Mitochondrial morphology. *Science* **185**: 630-1.
- Hall, D. A., Zhu, H., Zhu, X., Royce, T., Gerstein, M. and Snyder, M., (2004) Regulation of gene expression by a metabolic enzyme. *Science* **306**: 482-4.
- Hazbun, T. R., *et al.*, (2003) Assigning function to yeast proteins by integration of technologies. *Mol Cell* **12**: 1353-65.
- Heringa, J., (1999) Two strategies for sequence comparison: profile-preprocessed and secondary structure-induced multiple alignment. *Comput Chem* **23**: 341-64.
- Hobbs, A. E., Srinivasan, M., McCaffery, J. M. and Jensen, R. E., (2001) Mmm1p, a mitochondrial outer membrane protein, is connected to mitochondrial DNA (mtDNA) nucleoids and required for mtDNA stability. *J Cell Biol* **152**: 401-10.
- Hoppins, S., (2006) Mitochondrial Protein Import in *Neurospora crassa*
- Hoppins, S., Lackner, L. and Nunnari, J., (2007) The machines that divide and fuse mitochondria. *Annu Rev Biochem* **76**: 751-80.
- Horazdovsky, B. F. and Emr, S. D., (1993) The VPS16 gene product associates with a sedimentable protein complex and is essential for vacuolar protein sorting in yeast. *J Biol Chem* **268**: 4953-62.
- Hu, J. P., Vanderstraeten, S. and Foury, F., (1995) Isolation and characterization of ten mutator alleles of the mitochondrial DNA polymerase-encoding *MIP1* gene from *Saccharomyces cerevisiae*. *Gene* **160**: 105-10.
- Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S. and O'Shea, E. K., (2003) Global analysis of protein localization in budding yeast. *Nature* **425**: 686-91.
- Huxley, C., Green, E. D. and Dunham, I., (1990) Rapid assessment of *S. cerevisiae* mating type by PCR. *Trends Genet* **6**: 236.
- Iida, M., Terada, K., Sambongi, Y., Wakabayashi, T., Miura, N., Koyama, K., Futai, M. and Sugiyama, T., (1998) Analysis of functional domains of Wilson disease protein (ATP7B) in *Saccharomyces cerevisiae*. *FEBS Lett* **428**: 281-5.

- Insdorf, N. F. and Bogenhagen, D. F., (1989) DNA polymerase gamma from *Xenopus laevis*. I. The identification of a high molecular weight catalytic subunit by a novel DNA polymerase photolabeling procedure. *J Biol Chem* **264**: 21491-7.
- Ito, J. and Braithwaite, D. K., (1990) Yeast mitochondrial DNA polymerase is related to the family A DNA polymerases. *Nucleic Acids Res* **18**: 6716.
- Ito, J. and Braithwaite, D. K., (1991) Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res* **19**: 4045-57.
- Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S. and Sakaki, Y., (2000) Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc Natl Acad Sci U S A* **97**: 1143-7.
- Iyengar, B., Luo, N., Farr, C. L., Kaguni, L. S. and Campos, A. R., (2002) The accessory subunit of DNA polymerase gamma is essential for mitochondrial DNA maintenance and development in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* **99**: 4483-8.
- Jarosz, D. F., Beuning, P. J., Cohen, S. E. and Walker, G. C., (2007) Y-family DNA polymerases in *Escherichia coli*. *Trends Microbiol* **15**: 70-7.
- Jiang, F., Rizavi, H. S. and Greenberg, M. L., (1997) Cardiolipin is not essential for the growth of *Saccharomyces cerevisiae* on fermentable or non-fermentable carbon sources. *Mol Microbiol* **26**: 481-91.
- Jin, Y. H., Clark, A. B., Slebos, R. J., Al-Refai, H., Taylor, J. A., Kunkel, T. A., Resnick, M. A. and Gordenin, D. A., (2003) Cadmium is a mutagen that acts by inhibiting mismatch repair. *Nat Genet* **34**: 326-9.
- Johnson, A. A., Tsai, Y., Graves, S. W. and Johnson, K. A., (2000) Human mitochondrial DNA polymerase holoenzyme: reconstitution and characterization. *Biochemistry* **39**: 1702-8.
- Johnson, S. J., Taylor, J. S. and Beese, L. S., (2003) Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc Natl Acad Sci U S A* **100**: 3895-900.
- Joyce, C. M. and Steitz, T. A., (1995) Polymerase structures and function: variations on a theme? *J Bacteriol* **177**: 6321-9.
- Kaguni, L. S., (2004) DNA polymerase gamma, the mitochondrial replicase. *Annu Rev Biochem* **73**: 293-320.

- Kajander, O. A., Karhunen, P. J., Holt, I. J. and Jacobs, H. T., (2001) Prominent mitochondrial DNA recombination intermediates in human heart muscle. *EMBO Rep* **2**: 1007-12.
- Kaufman, B. A., Kolesar, J. E., Perlman, P. S. and Butow, R. A., (2003) A function for the mitochondrial chaperonin Hsp60 in the structure and transmission of mitochondrial DNA nucleoids in *Saccharomyces cerevisiae*. *J Cell Biol* **163**: 457-61.
- Kaufman, B. A., Newman, S. M., Hallberg, R. L., Slaughter, C. A., Perlman, P. S. and Butow, R. A., (2000) In organello formaldehyde crosslinking of proteins to mtDNA: identification of bifunctional proteins. *Proc Natl Acad Sci U S A* **97**: 7772-7.
- Kempeneers, V., Renders, M., Froeyen, M. and Herdewijn, P., (2005) Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization. *Nucleic Acids Res* **33**: 3828-36.
- Kiefer, J. R., Mao, C., Braman, J. C. and Beese, L. S., (1998) Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* **391**: 304-7.
- Kijima, K., *et al.*, (2005) Mitochondrial GTPase mitofusin 2 mutation in Charcot-Marie-Tooth neuropathy type 2A. *Hum Genet* **116**: 23-7.
- Kleff, S., Kemper, B. and Sternglanz, R., (1992) Identification and characterization of yeast mutants and the gene for a cruciform cutting endonuclease. *Embo J* **11**: 699-704.
- Korhonen, J. A., Pham, X. H., Pellegrini, M. and Falkenberg, M., (2004) Reconstitution of a minimal mtDNA replisome in vitro. *Embo J* **23**: 2423-9.
- Kumar, C., Sharma, R. and Bachhawat, A. K., (2003) Investigations into the polymorphisms at the ECM38 locus of two widely used *Saccharomyces cerevisiae* S288C strains, YPH499 and BY4742. *Yeast* **20**: 857-63.
- Kunkel, T. A., Roberts, J. D. and Zakour, R. A., (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol* **154**: 367-82.
- Kushnirov, V. V., (2000) Rapid and reliable protein extraction from yeast. *Yeast* **16**: 857-60.
- Lambowitz, A. M. and Belfort, M., (1993) Introns as mobile genetic elements. *Annu Rev Biochem* **62**: 587-622.
- Lecrenier, N. and Foury, F., (2000) New features of mitochondrial DNA replication system in yeast and man. *Gene* **246**: 37-48.

- Lehman, I. R., Bessman, M. J., Simms, E. S. and Kornberg, A., (1958) Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *J Biol Chem* **233**: 163-70.
- Lemire, B. D., Fankhauser, C., Baker, A. and Schatz, G., (1989) The mitochondrial targeting function of randomly generated peptide sequences correlates with predicted helical amphiphilicity. *J Biol Chem* **264**: 20206-15.
- Lewandowska, A., Gierszewska, M., Marszalek, J. and Liberek, K., (2006) Hsp78 chaperone functions in restoration of mitochondrial network following heat stress. *Biochim Biophys Acta* **1763**: 141-51.
- Lim, S. E., Longley, M. J. and Copeland, W. C., (1999) The mitochondrial p55 accessory subunit of human DNA polymerase gamma enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. *J Biol Chem* **274**: 38197-203.
- Ling, F. and Shibata, T., (2002) Recombination-dependent mtDNA partitioning: in vivo role of Mhr1p to promote pairing of homologous DNA. *EMBO J* **21**: 4730-40.
- Ling, F. and Shibata, T., (2004) Mhr1p-dependent concatemeric mitochondrial DNA formation for generating yeast mitochondrial homoplasmic cells. *Mol Biol Cell* **15**: 310-22.
- Lingner, J., Cech, T. R., Hughes, T. R. and Lundblad, V., (1997) Three Ever Shorter Telomere (EST) genes are dispensable for in vitro yeast telomerase activity. *Proc Natl Acad Sci U S A* **94**: 11190-5.
- Longley, M. J., Ropp, P. A., Lim, S. E. and Copeland, W. C., (1998) Characterization of the native and recombinant catalytic subunit of human DNA polymerase gamma: identification of residues critical for exonuclease activity and dideoxynucleotide sensitivity. *Biochemistry* **37**: 10529-39.
- Luo, N. and Kaguni, L. S., (2005) Mutations in the spacer region of *Drosophila* mitochondrial DNA polymerase affect DNA binding, processivity, and the balance between Pol and Exo function. *J Biol Chem* **280**: 2491-7.
- Ma, L., Vaz, F. M., Gu, Z., Wanders, R. J. and Greenberg, M. L., (2004) The human TAZ gene complements mitochondrial dysfunction in the yeast taz1Delta mutant. Implications for Barth syndrome. *J Biol Chem* **279**: 44394-9.
- Maarse, A. C., Blom, J., Grivell, L. A. and Meijer, M., (1992) MPI1, an essential gene encoding a mitochondrial membrane protein, is possibly involved in protein import into yeast mitochondria. *Embo J* **11**: 3619-28.

- MacAlpine, D. M., Perlman, P. S. and Butow, R. A., (1998) The high mobility group protein Abf2p influences the level of yeast mitochondrial DNA recombination intermediates in vivo. *Proc Natl Acad Sci U S A* **95**: 6739-43.
- Maleszka, R., Skelly, P. J. and Clark-Walker, G. D., (1991) Rolling circle replication of DNA in yeast mitochondria. *EMBO J* **10**: 3923-9.
- Margulis, L., (1968) Evolutionary Criteria in Thallophytes: A Radical Alternative. *Science* **161**: 1020-1022.
- Mattoon, J. R., Caravajal, E. and Guthrie, D., (1990) Effects of hap mutations on heme and cytochrome formation in yeast. *Curr Genet* **17**: 179-83.
- McCulloch, S. D. and Kunkel, T. A., (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* **18**: 148-61.
- Meeusen, S., McCaffery, J. M. and Nunnari, J., (2004) Mitochondrial fusion intermediates revealed in vitro. *Science* **305**: 1747-52.
- Meeusen, S. and Nunnari, J., (2003) Evidence for a two membrane-spanning autonomous mitochondrial DNA replisome. *J Cell Biol* **163**: 503-10.
- Meeusen, S., Tieu, Q., Wong, E., Weiss, E., Schieltz, D., Yates, J. R. and Nunnari, J., (1999) Mgm101p is a novel component of the mitochondrial nucleoid that binds DNA and is required for the repair of oxidatively damaged mitochondrial DNA. *J Cell Biol* **145**: 291-304.
- Miyakawa, I., Kitamura, Y., Jyozaki, T., Sato, H. and Umezaki, T., (2000) Simple detection of a yeast mitochondrial DNA-binding protein, Abf2p, on SDS-DNA gels. *J Gen Appl Microbiol* **46**: 311-316.
- Miyakawa, I., Sando, N., Kawano, S., Nakamura, S. and Kuroiwa, T., (1987) Isolation of morphologically intact mitochondrial nucleoids from the yeast, *Saccharomyces cerevisiae*. *J Cell Sci* **88 (Pt 4)**: 431-9.
- Monroe, D. S., Jr., Leitzel, A. K., Klein, H. L. and Matson, S. W., (2005) Biochemical and genetic characterization of Hmi1p, a yeast DNA helicase involved in the maintenance of mitochondrial DNA. *Yeast* **22**: 1269-86.
- Moon, A. F., Garcia-Diaz, M., Batra, V. K., Beard, W. A., Bebenek, K., Kunkel, T. A., Wilson, S. H. and Pedersen, L. C., (2007) The X family portrait: structural insights into biological functions of X family polymerases. *DNA Repair (Amst)* **6**: 1709-25.
- Mortimer, R. K. and Johnston, J. R., (1986) Genealogy of principal strains of the yeast genetic stock center. *Genetics* **113**: 35-43.

- Mueller, D. M., Puri, N., Kabaleeswaran, V., Terry, C., Leslie, A. G. and Walker, J. E., (2004) Ni-chelate-affinity purification and crystallization of the yeast mitochondrial F1-ATPase. *Protein Expr Purif* **37**: 479-85.
- Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., Harley, C. B. and Cech, T. R., (1997) Telomerase catalytic subunit homologs from fission yeast and human. *Science* **277**: 955-9.
- Neupert, W. and Herrmann, J. M., (2007) Translocation of proteins into mitochondria. *Annu Rev Biochem* **76**: 723-49.
- Nicholas, K. B. and Nicholas, H. B. J., (1997) GeneDoc: A tool for editing and annotating multiple sequence alignments.
- Nosek, J., Tomaska, L., Bolotin-Fukuhara, M. and Miyakawa, I., (2006) Mitochondrial chromosome structure: an insight from analysis of complete yeast genomes. *FEMS Yeast Res* **6**: 356-70.
- Nunnari, J., Marshall, W. F., Straight, A., Murray, A., Sedat, J. W. and Walter, P., (1997) Mitochondrial transmission during mating in *Saccharomyces cerevisiae* is determined by mitochondrial fusion and fission and the intramitochondrial segregation of mitochondrial DNA. *Mol Biol Cell* **8**: 1233-42.
- Ogur, M., St. John, R. and Nagai, S., (1957) Tetrazolium overlay technique for population studies of respiration deficiency in yeast. *Science* **125**: 928-9.
- Olson, M. W., Wang, Y., Elder, R. H. and Kaguni, L. S., (1995) Subunit structure of mitochondrial DNA polymerase from *Drosophila* embryos. Physical and immunological studies. *J Biol Chem* **270**: 28932-7.
- Paschen, S. A., Rothbauer, U., Kaldi, K., Bauer, M. F., Neupert, W. and Brunner, M., (2000) The role of the TIM8-13 complex in the import of Tim23 into mitochondria. *Embo J* **19**: 6392-400.
- Phadnis, N. and Ayres Sia, E., (2004) Role of the putative structural protein Sed1p in mitochondrial genome maintenance. *J Mol Biol* **342**: 1115-29.
- Philippsen, P., Stotz, A. and Scherf, C., (1991) DNA of *Saccharomyces cerevisiae*. *Methods Enzymol* **194**: 169-182.
- Polesky, A. H., Steitz, T. A., Grindley, N. D. and Joyce, C. M., (1990) Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*. *J Biol Chem* **265**: 14579-91.
- Reinders, J., Wagner, K., Zahedi, R. P., Stojanovski, D., Eyrich, B., van der Laan, M., Rehling, P., Sickmann, A., Pfanner, N. and Meisinger, C., (2007) Profiling

phosphoproteins of yeast mitochondria reveals a role of phosphorylation in assembly of the ATP synthase. *Mol Cell Proteomics* **6**: 1896-906.

Rieger, K. J., Kaniak, A., Coppee, J. Y., Aljinovic, G., Baudin-Baillieu, A., Orłowska, G., Gromadka, R., Groudinsky, O., Di Rago, J. P. and Slonimski, P. P., (1997) Large-scale phenotypic analysis--the pilot project on yeast chromosome III. *Yeast* **13**: 1547-62.

Rodriguez-Cousino, N., Lill, R., Neupert, W. and Court, D. A., (1995) Identification and initial characterization of the cytosolic protein Ycr77p. *Yeast* **11**: 581-5.

Sakasegawa, Y., Hachiya, N. S., Tsukita, S. and Kaneko, K., (2003) Ecm10p localizes in yeast mitochondrial nucleoids and its overexpression induces extensive mitochondrial DNA aggregations. *Biochem Biophys Res Commun* **309**: 217-21.

Sambrook, J. and Russell, D., (2001) *Molecular Cloning - A Laboratory Manual*

Sato, H. and Miyakawa, I., (2004) A 22 kDa protein specific for yeast mitochondrial nucleoids is an unidentified putative ribosomal protein encoded in open reading frame YGL068W. *Protoplasma* **223**: 175-82.

Sato, H., Tachifuji, A., Tamura, M. and Miyakawa, I., (2002) Identification of the YMN-1 antigen protein and biochemical analyses of protein components in the mitochondrial nucleoid fraction of the yeast *Saccharomyces cerevisiae*. *Protoplasma* **219**: 51-8.

Scherer, S. and Davis, R. W., (1979) Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc Natl Acad Sci U S A* **76**: 4951-5.

Schmitt, M. E. and Clayton, D. A., (1993) Conserved features of yeast and mammalian mitochondrial DNA replication. *Curr Opin Genet Dev* **3**: 769-74.

Schwimmer, C., Rak, M., Lefebvre-Legendre, L., Duvezin-Caubet, S., Plane, G. and di Rago, J. P., (2006) Yeast models of human mitochondrial diseases: from molecular mechanisms to drug screening. *Biotechnol J* **1**: 270-81.

Sena, E. P., Welch, J. W., Halvorson, H. O. and Fogel, S., (1975) Nuclear and mitochondrial deoxyribonucleic acid replication during mitosis in *Saccharomyces cerevisiae*. *J Bacteriol* **123**: 497-504.

Sesaki, H., Southard, S. M., Yaffe, M. P. and Jensen, R. E., (2003) Mgm1p, a dynamin-related GTPase, is essential for fusion of the mitochondrial outer membrane. *Mol Biol Cell* **14**: 2342-56.

Shadel, G. S., (1999) Yeast as a model for human mtDNA replication. *Am J Hum Genet* **65**: 1230-7.

- Shadel, G. S., (2005) Mitochondrial DNA, aconitase 'wraps' it up. *Trends Biochem Sci* **30**: 294-6.
- Sheff, M. A. and Thorn, K. S., (2004) Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast* **21**: 661-70.
- Shen, Y., Tang, X. F., Matsui, E. and Matsui, I., (2004) Subunit interaction and regulation of activity through terminal domains of the family D DNA polymerase from *Pyrococcus horikoshii*. *Biochem Soc Trans* **32**: 245-9.
- Sherman, F., (1964) Mutants of Yeast Deficient in Cytochrome C. *Genetics* **49**: 39-48.
- Sherman, F., (1991) Getting started with yeast. *Methods Enzymol* **194**: 3-21.
- Sherman, F., (2002) Getting started with yeast. *Methods Enzymol* **350**: 3-41.
- Sherman, F., Fink, G. R. and Hicks, J. B., (1986) Laboratory course manual for methods in yeast genetics / by Fred Sherman, Gerald R. Fink, and James B. Hicks.
- Shibata, T. and Ling, F., (2007) DNA recombination protein-dependent mechanism of homoplasmy and its proposed functions. *Mitochondrion* **7**: 17-23.
- Shimizu, K., Kawasaki, Y., Hiraga, S., Tawaramoto, M., Nakashima, N. and Sugino, A., (2002) The fifth essential DNA polymerase phi in *Saccharomyces cerevisiae* is localized to the nucleolus and plays an important role in synthesis of rRNA. *Proc Natl Acad Sci U S A* **99**: 9133-8.
- Shutt, T. E. and Gray, M. W., (2006) Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet* **22**: 90-5.
- Sikorski, R. S. and Hieter, P., (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19-27.
- Simossis, V. A. and Heringa, J., (2003) The PRALINE online server: optimising progressive multiple alignment on the web. *Comput Biol Chem* **27**: 511-9.
- Sirum-Connolly, K. and Mason, T. L., (1993) Functional requirement of a site-specific ribose methylation in ribosomal RNA. *Science* **262**: 1886-9.
- Steger, H. F., Sollner, T., Kiebler, M., Dietmeier, K. A., Pfaller, R., Trulzsch, K. S., Tropschug, M., Neupert, W. and Pfanner, N., (1990) Import of ADP/ATP carrier into mitochondria: two receptors act in parallel. *J Cell Biol* **111**: 2353-63.
- Steinmetz, L. M., *et al.*, (2002) Systematic screen for human disease genes in yeast. *Nat Genet* **31**: 400-4.

- Storici, F. and Resnick, M. A., (2006) The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. *Methods Enzymol* **409**: 329-45.
- Struhl, K., (1985) Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc Natl Acad Sci U S A* **82**: 8419-23.
- Stuart, G. R., Santos, J. H., Strand, M. K., Van Houten, B. and Copeland, W. C., (2006) Mitochondrial and nuclear DNA defects in *Saccharomyces cerevisiae* with mutations in DNA polymerase gamma associated with progressive external ophthalmoplegia. *Hum Mol Genet* **15**: 363-74.
- Ter Linde, J. J. and Steensma, H. Y., (2002) A microarray-assisted screen for potential Hap1 and Rox1 target genes in *Saccharomyces cerevisiae*. *Yeast* **19**: 825-40.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G., (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876-82.
- Thompson, J. R., Register, E., Curotto, J., Kurtz, M. and Kelly, R., (1998) An improved protocol for the preparation of yeast cells for transformation by electroporation. *Yeast* **14**: 565-71.
- Tranebjaerg, L., *et al.*, (1995) A new X linked recessive deafness syndrome with blindness, dystonia, fractures, and mental deficiency is linked to Xq22. *J Med Genet* **32**: 257-63.
- Uetz, P., *et al.*, (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**: 623-7.
- Umezaki, T. and Miyakawa, I., (2002) Use of SDS-DNA PAGE for detection of mitochondrial Abf2p-like proteins and mitochondrial nuclease in *Saccharomyces* yeasts and *Arxiozyma telluris*. *Cytologia* **67**: 423-428.
- Van Dyck, E., Foury, F., Stillman, B. and Brill, S. J., (1992) A single-stranded DNA binding protein required for mitochondrial DNA replication in *S. cerevisiae* is homologous to *E. coli* SSB. *Embo J* **11**: 3421-30.
- Vanderstraeten, S., Van den Brule, S., Hu, J. and Foury, F., (1998) The role of 3'-5' exonucleolytic proofreading and mismatch repair in yeast mitochondrial DNA error avoidance. *J Biol Chem* **273**: 23690-7.
- Vaz, F. M., Houtkooper, R. H., Valianpour, F., Barth, P. G. and Wanders, R. J., (2003) Only one splice variant of the human TAZ gene encodes a functional protein with a role in cardiolipin metabolism. *J Biol Chem* **278**: 43089-94.

Votruba, M., Fitzke, F. W., Holder, G. E., Carter, A., Bhattacharya, S. S. and Moore, A. T., (1998) Clinical features in affected individuals from 21 pedigrees with dominant optic atrophy. *Arch Ophthalmol* **116**: 351-8.

Wallace, D. C., (2007) Why do we still have a maternally inherited mitochondrial DNA? Insights from evolutionary medicine. *Annu Rev Biochem* **76**: 781-821.

Werner-Washburne, M., Braun, E. L., Crawford, M. E. and Peck, V. M., (1996) Stationary phase in *Saccharomyces cerevisiae*. *Mol Microbiol* **19**: 1159-66.

Williamson, D. H. and Fennell, D. J., (1979) Visualization of yeast mitochondrial DNA with the fluorescent stain "DAPI". *Methods Enzymol* **56**: 728-33.

Winzeler, E. A., Castillo-Davis, C. I., Oshiro, G., Liang, D., Richards, D. R., Zhou, Y. and Hartl, D. L., (2003) Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays. *Genetics* **163**: 79-89.

Winzeler, E. A., *et al.*, (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**: 901-6.

Wong, E. D., Wagner, J. A., Gorsich, S. W., McCaffery, J. M., Shaw, J. M. and Nunnari, J., (2000) The dynamin-related GTPase, Mgm1p, is an intermembrane space protein required for maintenance of fusion competent mitochondria. *J Cell Biol* **151**: 341-52.

Young, M. and Court, D., (2004) Quick measurement of glucose concentration in *Saccharomyces cerevisiae* cultures. *Bull. Genet. Soc. Can.* **35**: 109-110.

Young, M. J., Theriault, S. S., Li, M. and Court, D. A., (2006) The carboxyl-terminal extension on fungal mitochondrial DNA polymerases: identification of a critical region of the enzyme from *Saccharomyces cerevisiae*. *Yeast* **23**: 101-16.

Youngman, M. J., Hobbs, A. E., Burgess, S. M., Srinivasan, M. and Jensen, R. E., (2004) Mmm2p, a mitochondrial outer membrane protein required for yeast mitochondrial shape and maintenance of mtDNA nucleoids. *J Cell Biol* **164**: 677-88.

Zelenaya-Troitskaya, O., Newman, S. M., Okamoto, K., Perlman, P. S. and Butow, R. A., (1998) Functions of the high mobility group protein, Abf2p, in mitochondrial DNA segregation, recombination and copy number in *Saccharomyces cerevisiae*. *Genetics* **148**: 1763-76.

Zelenaya-Troitskaya, O., Perlman, P. S. and Butow, R. A., (1995) An enzyme in yeast mitochondria that catalyzes a step in branched-chain amino acid biosynthesis also functions in mitochondrial DNA stability. *EMBO J* **14**: 3268-76.

Zhang, M., Su, X., Mileykovskaya, E., Amoscato, A. A. and Dowhan, W., (2003) Cardiolipin is not required to maintain mitochondrial DNA stability or cell viability for *Saccharomyces cerevisiae* grown at elevated temperatures. *J Biol Chem* **278**: 35204-10.

Zhong, Q., Gohil, V. M., Ma, L. and Greenberg, M. L., (2004) Absence of cardiolipin results in temperature sensitivity, respiratory defects, and mitochondrial DNA instability independent of *pet56*. *J Biol Chem* **279**: 32294-300.