UNIVERSITY
of
GLASGOW

ATP-binding cassette transporters of *Paracoccidiodes brasiliensis*

By

Christopher H Gray B.Sc (Hons)

A thesis presented for the degree of Doctor of Philosophy
in
The Faculty of Science at the University of Glasgow

Institute of Biomedical and Life Sciences
Division of Infection & Immunity
University of Glasgow
Glasgow G21 8QQ

April 2001

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

The author was a recipient of an Institute of Biomedical and Life Sciences studentship. Except where stated all results described in this thesis were obtained by the author's own efforts.

20/4/1

Christopher H Gray

20th April 2001

Acknowledgements

I wish to give special thanks to Dr. Adrian R Walmsley for his supervision, support and encouragement throughout this project and for the proof reading of this thesis.

I would also like to thank the Institute of Biomedical & Life at the University of Glasgow for the provision of the funds that sustained me throughout the project.

This work would not have been possible without the extended assistance I received from members of the group. Particular thanks go to Dr. M. Ines Borges-Walmsley for knowledge, advice and assistance. Dr. Borges-Walmsley also prepared all the nucleic acid materials supplied for this study. I would also like to make special mention of Dr. Paul Blackburn who has been resolute in his support, although his poor emu spotting skills almost brought the whole endeavour to an end.

To the members of the Division of Infection and Immunity South Laboratory I extend my warmest appreciation. They are recognised for their enthusiasm and friendship....and for putting up with my exclusive brand of "wit".

Thanks also go to Caroline for listening to my moans when things were bad and for keeping my feet on the ground when things were good. I would like to apologise for all those Saturday afternoons when we "just popped into the lab for 10 minutes" and stayed for 2 hours.

This thesis is dedicated to my Mum and Dad whose selfless and unfailing support has allowed me to get to this point. I extend to them my greatest appreciation and love. Without them none of this would have been possible. I would also thank Simon, Jennifer and Connor for their encouragement and understanding.

Summary

ATP-binding cassette (ABC) transporters represent one of the largest recognised superfamilies of proteins. They are ubiquitous in nature and have been implicated in a wide range of physiological functions, including multi-drug resistance (MDR). All of these transporters are structurally related, with active transport complexes exhibiting the minimal complement of twelve transmembrane helices and two nucleotide-binding domains. Despite the structural similarity, the range of functions performed by these proteins is variable. In many cases the substrate range is equally diverse with the proteins displaying the capacity to translocate substrates which appear structurally unrelated. This intensifies the problem of MDR where a number of classes of drug are susceptible to efflux.

A number of ABC transporters have been identified in fungi including *Saccharomyces cerevisiae* and *Candida albicans*. An increasing wealth of information is emerging for a few of these transporters. The current knowledge on the various characteristics of the *S.cerevisiae* transporter PDR5 approaches that relating to P-glycoprotein (Pgp) of mammalian cells. The ABC transporter CDR1 from *Candida albicans* has been shown to confer clinical drug resistance in a number of disease cases.

Paracoccidiodes brasiliensis is a thermally dimorphic human pathogen that is exclusive to South America. Inhalation is the route of infection with disease progression originating in the lungs and progressing to other parts of the body. Paracoccidiodiomycosis is the most prevalent systemic mycosis in Latin America. Exposure to the organism is most common in rural workers, particularly those involved in deforestation. Disease progression occurs when the host is immunocompromised and tends to follow and extended period of latency which can last over ten years. Outgrowth of the organism is dependent on a morphological transition from the mycelial to the yeast form of the organism. This transition is inhibited by the female hormone oestrogen resulting in the clinical syndrome displaying a bias towards males. The drug of choice is ketoconazole, a member of the azole family of anti-fungal agents. Azoles act on an intermediate enzyme in the ergosterol biosynthesis pathway, lanosterol demethylase, resulting in the perturbation of cell membrane integrity. Resistance to azole drugs can occur in two ways. A mutation in lanosterol demethylase can alter the drug-binding site so that the target is not recognised. Alternatively, the intracellular concentration of azole can be reduced by the ABC transporter mediated active efflux of the drug.

The remit of this thesis was to screen the genome of *P.brasiliensis* for loci encoding ABC transporters. Particular interest would be paid to proteins that may be involved with drug resistance. Initially a strategy employed by other investigators was employed. Two degenerate oligonucleotides (MDR1 and MDR2), designed on the basis of conserved motifs from the nucleotide binding domains of a number of ABC transporters were used in PCR. A single amplicon was obtained which showed homology to other sequences coding for nucleotide-binding domains. This amplicon was used to screen a lambda genomic library and a single phage was isolated. This phage clone contains part of a gene with displayed considerable identity to an ABC half-transporter from *Aspergillus fumigatus*. The complete sequence of the gene was then obtained using genome walking SSP-PCR. The gene, *pft1*, spans 2627bp and is interrupted by 2 introns of 68bp and 77bp.

A similar approach was employed in a search for additional genes. The amino acid sequences from a number of PDR5-like fungal ABC transporters were aligned. This allowed the identification of conserved motifs that were specific to these fungal proteins. Application of these primers allowed the amplification of a second amplicon. The sequence of this amplicon displayed considerable homology to the fungal transporter genes. This fragment was proposed to belong to a second *P.brasiliensis* gene, *pft2*. Screening of the lambda phage genomic library allowed the isolation of a single clone. Subcloning of fragments from this phage clone allowed the greater part of this gene sequence to be determined.

RT-PCR based transcript analysis of these genes proved that both were expressed and spliced by *P.brasiliensis*. The expression of *pft1* occurred only on stimulation with an azole drug. This result represents the first evidence linking a fungal half-transporter with drug resistance. Further analysis of the *pft1* transcript revealed that the mRNA was shorter than expected at the 5' end. 5'RACE RT-PCR suggested that a second ATG codon was the true translational start and not the ATG suggested by the genomic sequence. The acquisition of cDNA allowed progression to recombinant protein expression. All further experiments concentrate on *pft1*, as half-transporters represent a more novel and less characterised system.

A number of constructs were created in pET21b. The objective was to produce a soluble nucleotide-binding domain of Pft1. Subcellular localisation of the resultant proteins caused a number of problems. Larger constructs appeared to partition to the membrane which would complicate purification procedures. The level of expression of smaller portions of the NBD was very high and the protein persisted mainly as insoluble inclusion bodies. Despite these problems, a 28.0kDa NBD was selected for

purification using nickel affinity chromatography. This pure protein, termed Pft1NBDTW, was then applied to fluorescence based experiments. Measurement of tryptophan fluorescence and fluorescence energy transfer to ligands allowed the demonstration of binding of both ATP and ketoconazole to the NBD. Stopped flow analysis allowed the visualisation of the transient kinetics of ATP binding. Further experiments, examining the steady-state binding of ketoconazole and ATP indicated that the NBD had a high affinity for ketoconazole.

Table of Contents

Chapter1 General Introduction

1.1 *Paracoccidiodes brasiliensis*

1.1.1 Morphology of *P.brasiliensis* 2

1.1.2 Paracoccidiodiomycosis; Disease and Infection 7

1.1.3 Molecular Studies of *P.brasiliensis* 9

1.1.4 Gp43 and the Immunobiology of Paracoccidiodiomycosis 13

1.2 ATP-Binding Cassette Transporters

1.2.1 The structural organisation of ABC transport proteins 19

1.2.2 The Nucleotide-Binding Domain 22

1.2.3 The Membrane Spanning Domain 32

1.2.4 The Fungal ATP-binding Cassette Transporters 37

1.2.5 Fungal Drug Resistance and ABC Multidrug Transporters 39

1.2.6 The Regulation of ABC Transport Mediated Fungal Drug Resistance 45

1.2.7 Conclusions 48

Chapter 2 Materials and Methods

2.1 Reagents and Materials

2.1.1 Origin of materials and reagents 49

2.1.2 Preparation of chemical and media solutions 49

2.1.3 Sterilisation of reagents and materials 49

2.1.4 Centrifugation 49

2.2 Microorganisms; Growth and Storage

2.2.1 Growth of Bacteria 50

2.2.2 Growth of *Saccharomyces cerevisiae* 50

2.2.3 Long term storage of bacteria 52

2.2.4 Antibiotics 52

2.2.5 Growth and preparation of *E.coli* LE392 for λ phage infection 53

2.2.6 Propagation of λ phage on solid media 53

2.2.7 Propagation of λ phage in liquid media 53

2.3 Isolation and Analysis of Nucleic Acids

2.3.1 Agarose gel electrophoresis of DNA 54
2.3.2 Small scale preparation of λ DNA 55
2.3.3 Alcohol precipitation of DNA 56
2.3.4 Small scale preparation of plasmid DNA 56
2.3.5 Purification of *P.brasiliensis* DNA from LMP agarose gels 57
2.3.6 Extraction of DNA from agarose gels 57
2.3.7 Purification of DNA from buffered solutions 58
2.3.8 Alternative procedure for gel extraction/purification of DNA 59
2.3.9 Purification of mRNA from total nucleic acid 60

2.4 Use of the Polymerase Chain Reaction (PCR)

2.4.1 Oligonucleotide primers for PCR 61
2.4.2 Routine PCR 63
2.4.3 Single Specific Primer PCR (SSP-PCR) 65
2.4.4 PCR based DNA sequencing 65
2.4.5 Reverse transcriptase PCR (RT-PCR) 65
2.4.6 Nested 5' SMART™ RACE RT-PCR 66

2.5 Universal GenomeWalker™ SSP-PCR

2.5.1 Preparation of DNA for library construction 68
2.5.2 Ligation of GenomeWalker™ adapters to the blunt termini of restriction... 69
2.5.3 PCR genome walking using the GenomeWalker™ libraries 69

2.6 Enzyme Mediated Manipulations of DNA

2.6.1 Restriction enzyme digestion 71
2.6.2 Fluorescein labelling of probes for hybridisation 72
2.6.3 A-tailing of proof-read RT-PCR amplicons to allow A-T cloning 72
2.6.4 Dephosphorylation of vector ends 73
2.6.5 Ligation of DNA to plasmid vectors 73

2.7 Transformation of *E.coli* and *S.cerevisiae* with plasmid DNA

2.7.1 Transformation of chemically competent *E.coli* 75
2.7.2 Transformation of bacteria by electroporation 75
2.7.3 Transformation of *S.cerevisiae* with plasmid DNA 75
2.7.4 Selective screening of recombinant bacterial clones 76

2.8 Southern transfer and hybridisation of DNA

2.8.1 Southern transfer of DNA from agarose gels 77

2.8.2 Transfer of DNA to nitro-cellulose membranes from λ plaque lawns 78

2.8.3 DNA:DNA hybridisation 78

2.8.4 Washing and detection of bound probe 79

2.8.5 Analysis of results derived from agarose gels 79

2.8.6 Analysis of results from λ plaque lifts 80

2.9 Heterologous expression of recombinant proteins

2.9.1 Expression of recombinant proteins from pET21b in *Epicurian coli*.. 80

2.9.2 Preparation of total cell protein (TCP) samples for analysis... 81

2.9.3 Expression of recombinant proteins from pYES2 in *S.cerevisiae* 81

2.9.4 Electrophoresis of protein under denaturing conditions... 82

2.9.5 Assembly of electrophoresis apparatus 83

2.9.6 Preparation of samples for SDS-PAGE 83

2.9.7 Electrophoresis protocol 84

2.9.8 Analysis of SDS-PAGE gels 84

2.10 Analysis of recombinant expression by western blotting

2.10.1 Western transfer of protein onto nitro-cellulose membrane 85

2.10.2 Processing of the western blot 86

2.11 Purification of His_6 tagged proteins

2.11.1 Purification of His_6 tagged proteins under denaturing conditions 86

2.11.2 Larger scalePurification of His_6 tagged proteins under denaturing conditions 87

2.11.3 Purification of His_6 tagged proteins under native conditions 88

2.11.4 Dialysis of purified protein

2.12 Ligand binding analysis

2.12.1 Stopped-flow analysis of nucleotide/protein interactions 89

2.12.2 Steady-state analysis of protein/drug interactions 91

Chapter 3 Isolation and sequencing of *pft1* an MDR-like ABC half transporter from *Paracoccidiodes brasiliensis*

3.1 PCR amplification of an MDR-like gene fragment using degenerate PCR 92

3.2 Screening of lambda libraries for the *pft1* gene 93

3.3 Subcloning of *P.brasiliensis* DNA from λ*pft1* 100
3.4 Genome walking to complete the *pft1* sequence 104
3.5 The sequence of *pft1* 106
3.6 Discussion 115

Chapter 4 Identification and sequencing of *pft2*, a full-sized ABC transporter gene from *Paracoccidiodes brasiliensis*

4.1 Multiple alignments and the design of degenerate oligonucleotides 120
4.2 PCR amplification of a gene fragment using degenerate PCR 120
4.3 Screening of lambda libraries for the *pft2* gene 123
4.4 Subcloning of *P.brasiliensis* DNA from λ*pft2* 126
4.5 The sequence of *pft2* 131
4.6 Discussion 133

Chapter 5 Expression analysis of *pft1* and *pft2* transcripts at mRNA level

5.1 Introns and splicing in *pft1* 144
5.2 RT-PCR expression analysis of *pft1* and *pft2* on drug treatment 146
5.3 SMART™ RACE RT-PCR analysis of the 5' terminus of *pft1* 148
5.4 Full length cDNA for *pft1* 151
5.5 Discussion 153

Chapter 6 Over-expression of Pft1 domains and the purification of the Pft1 nucleotide-binding domain

6.1 Initial construction of a pET21b plasmid for over expression.... 160
6.2 Over-expression of Pft1NBD 162
6.3 Construction of expression systems for Pft1 NBD truncates 163
6.4 Denaturing purification of Pft1NBDTW 169
6.5 Native purification of Pft1NBDTW 170
2.11 Discussion

Chapter 7 Ligand Binding Studies for Pft1NBDTW
7.1 Fluorescence based analysis of ATP binding to Pft1NBDTW 176

7.2 Steady state binding studies for azole binding to Pft1NBDTW 179
7.3 ATP alters the binding of ketoconazole on Pft1NBDTW 183
7.4 Heterologous complementation of *S.cerevisiae* 185
7.5 Discussion 185

Chapter 8 Final Discussion

References

Appendices

Appendix I
Appendix II

List of Tables and Figures

Tables

Chapter 2

2.1	Microorganisms and strains thereof, used in this study	51
2.2	Oligonucleotide primers used for PCR	61

Figures

Chapter1

1.1	The thermal dimorphism of *P.brasiliensis*	4
1.2	The general modular structure of ABC transporters	20
1.3	The ATP hydrolysing domains of ArsA and HisP	29

Chapter 3

3.1a	Degenerate PCR screening for MDR-like genes	94
3.1b	pGEM-T™ Easy clones containing the MDR1/MDR2 fragment	95
3.2	BLAST output on screening with the MDR1/MDR2 fragment	96
3.3	MDR1/MDR2 screening of phage DNA library DNAs	98
3.4a/b	Southern hybridisation screens of λ plaque lawns	99
3.5a	Restriction digests of λ*pft1* DNA	101
3.5b	Southern blot of gel in Fig3.5a	102
3.6	Cloning of the λ*pft1* derived *Xba*I fragment in pZERO™	103
3.7a/b	SSP-PCR reactions	105
3.8a/b	GenomeWalker™ PCRs	107
3.9	Contiguous DNA assembly of the *pft1* sequence	108
3.10	The nucleotide sequence and deduced amino acid sequence of *pft1*	109
3.11	Comparative alignment of afuMDR2 and Pft1	113

Chapter 4

4.1	The design of PCR primers targeting PDR subfamily genes	121
4.2	PDR1/PDR2 degenerate PCR analysis of genomic DNA	122
4.3	BLAST output on screening with the PDR1/PDR2 fragment	124
4.4	Southern hybridisation screens of λ plaque lawns	125
4.5a	Restriction digests of λ*pft2* DNA	127
4.5b	Southern blot of gel in Fig4.5a	128
4.6	Restriction enzyme digestion of plasmids	130
4.7	Contiguous DNA assembly of the *pft2* sequence	132

4.8 The partial nucleotide sequence ..... of *pft2* 134
4.9 Comparative alignments of AtrA and Pft2 138

Chapter 5
5.1 RT-PCR analysis of *pft1* transcripts 145
5.2 Azole dependant induction of *pft1* expression 147
5.3 induction of *pft2* expression 149
5.4a/b 5' SMART™ RT-PCR 150
5.5 A comparison of ATG start sites of *pft1* 152
5.6 RT-PCR generating full length *pft1* cDNA amplicons 154

Chapter 6
6.1 Strategy for the construction of expression plasmids 159
6.2a Kyte-Doolittle hydropathy plot of the amino acid sequence of Pft1 161
6.2b Restriction enzyme analysis of pET21NBD clones 161
6.3 Western blot of Pft1NBD 164
6.4 Restriction enzyme analysis of pET21bNBDT/TW clones 166
6.5 Expression of Pft1NBDT and Pft1NBDTW 167
6.6 Subcellular localisation of Pft1NBDT and Pft1NBDTW 168
6.7 Purification of Pft1NBDTW under denaturing conditions 171
6.8 Purification of Pft1NBDTW under native conditions 173

Chapter 7
7.1 Analysis of Pft1NBDTW on binding of MANT-ATP 178
7.2 Analysis of MANT-ATP on binding to Pft1NBDTW 180
7.3 FRET transfer of Pft1NBDTW to ketoconazole 182
7.4 Titration of ketoconazole against Pft1NBDTW 184
7.5 Titration of ketoconazole against Pft1NBDTW with 10mM ATP 186

List of Abbreviations

ABC	ATP-Binding Cassette
ADP	Adenosine diphosphate
ALV	Avian Leukemia Virus
APC	Antigen Presenting Cell
APS	Ammonium Persulphate
Atm	Atmosphere
ATP	Adenosine 5' triphosphate
BLAST	Basic Logical Alignment Search Tool
cAMP	cyclic adenosine monophosphate
CIAP	Calf Intestinal Alkaline Phosphatase
CIC	Circulating Immune Complex
DAPI	4,6,-diamidino-2-phenylindole
DEPC	Diethyl Pyrocarbonate
DMSO	Dimethyl Sulphoxide
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetracetate
GM-CSF	Granulocyte Macrophage-Colony Stimulating Factor
GSP	Gene Specific Primer
GST	Glutathione-S-Transferase
HIV	Human Immunodeficiency Virus
HRP	Horse Radish Peroxidase
IAAP	Iodoazidoprazosin
IFN-γ	Interferon-γ
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITS	Internal Transcribed Spacer
kb	Kilobase(s)
KDa	Kilo Dalton(s)
LB	Luria Bertani

LMP	Low Melting Point
LMW	Low Molecular Weight
MANT	2'-(or 3')-O-(N-methylanthramiloyl) adenosine 5' triphosphate
MAP	Mitogen Activated Protein
Mb	Megabase(s)
MBP	Maltose-Binding Protein
MCS	Multiple Cloning Site
MDR	Multi-Drug Resistance
MMLV	Moloney Murine Leukemia Virus
mRNA	Messenger Ribonucleic Acid
MSD	Membrane Spanning Domain
MTS	Methanethiosulphonate
NBD	Nucleotide-Binding Domain
NCBI	National Centre for Biotechnology Information
PBS	Phosphate Buffered Saline
PCM	Paracoccidiodiomycosis
PCR	Polymerase Chain Reaction
PDR	Pleiotropic Drug Resistance
PDRE	Pleiotropic Drug Responsive Element
PEG	Polyethylene Glycol
PFU	Plaque Forming Units
PMN	Polymorphonuclear Leukocytes
PMT	Photomultiplier Tube
RAPD	Random Amplified Polymorphic DNA
RNA	Ribonucleic Acid
RPM	Revolutions Per Minute
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
SA-PMP	Streptavidin-Paramagnetic Particle
SDS	Sodium Dodacyl Sulphate
SDS-PAGE	Sodium Dodacyl Sulphate-Polyacrylamide Gel Electrophoresis
SMART	Small Molecules Attaching to the 3' End of RNA Transcripts
STRE	Stress Responsive Element
dT	Deoxythymidylate

TCP	Total Cell Protein
TEMED	N,N,N'N'-tetramethylethylenediamine
TM	Transmembrane
TNP	2'-(or 3')-O-(trinitrophenyl)
UPM	Universal Primer Mix
X-Gal	5-Bromo-4-chloro-3-indolyl-β-D-galactoside

Chapter 1

General Introduction

Paracoccidiodes brasiliensis is a pathogenic fungus and the etiological agent of paracoccidiodiomycosis (PCM). PCM is the most prevelant systemic mycosis in Latin America. Disease is observed in discreet pockets within an area that is restricted to the regions between Mexico and northern Argentina. Even within these limits not every country reports disease. Approximately 80% of cases are recorded in Brazil, where PCM is a significant clinical problem. The majority of cases are clustered in areas of tropical and humid forest, the suspected environmental niche of *P.brasiliensis*. A number of patients are individuals involved in the deforestation industry. The disease manifests initially as a respiratory infection causing pulmonary lesions. This typically follows an extended period of latency (up to 15 years) where the organism remains dormant in the lower lung. Outgrowth often occurs when the host becomes immuno-compromised and, if untreated, can spread to other organs causing lesions in the skin, adrenal glands, mucous membranes and lymph nodes (Brummer *et al.*, 1993).

As with so many pathogens restricted to the developing world, the characterisation of *P.brasiliensis* is poor in comparison with pathogens causing disease in the developed world. This is particularly true with respect to studies involving current applications of molecular biology, biochemistry and immunobiology. However, the increasing profile of the South American scientific community, coupled with an increased interest in tropical disease in centres of the West has led to an elevation in the frequency of papers relating to this organism in recent years. Sequences derived from *P.brasiliensis* have been submitted to sequence databases. The total number of distinct genes represented in these data is 13, and the number will continue to increase (Source; NCBI Genbank database (www.ncbi.nih.nlm.gov)).

Within the field of medical mycology, an area which receives intense interest is the phenomenon of multi-drug resistance. Indeed resistance to the popular azole drugs has already been demonstrated in the important pathogen, *Candida albicans*. (Hitchcock, 1993; Miyazaki *et al.*, 1998; Prasad *et al.*, 1995). Such drug resistance is often mediated by the active efflux of the cytotoxic drug from the microbial cell by membrane transport proteins. These can belong to either the major facilitator superfamily or the ATP-binding cassette transporter (ABC transporter) superfamily. ABC tranporters are ATP hydrolysing, energy dependant influx or efflux systems. They are universal and are described in organisms from bacteria to man (Higgins, 1992). In *Saccharomyces*

cerevisiae there are 31 genes coding for proteins of this type (Decottignies & Goffeau, 1999). The biological function of ABC transporters is wide and varied but popular themes of research often focus on processes such as antigen presentation (Karttunen *et al.*, 1999), solute and nutrient uptake (Hung *et al.*, 1998) and multidrug resistance (Higgins, 1995; Tobin *et al.*, 1997; Van Veen *et al.*, 1996). It is accepted that ABC transporters perform a wide range of additional and often essential tasks contributing to cellular homeostasis. Prior to the studies reported in the following chapters of this thesis, no ABC transporters had been described in *P.brasiliensis.*

1.1 *Paracoccidiodes brasiliensis*

1.1.1 Morphology of *P.brasiliensis*

P.brasiliensis is a thermally dimorphic saprobic fungus which exists in a mycelial form at environmental temperatures (26^{0}C) and in a yeast form at temperatures consistent with internalisation in the mammalian host (37^{0}C) (Braibant *et al.*, 1996; Brummer *et al.*, 1993). (This is an inverted arrangement to the dimorphism exhibited by *C.albicans* where mycelial growth is associated with the infectious form (Balan *et al.*, 1997)). The organism was first cultured by Lutz in 1908 and reported in the Brazilian scientific literature (Lutz, 1908). Disease in man is invariably associated with the yeast form of the organism. Very often the clinical diagnosis of PCM involves the microscopic observation of the distinct "pilots wheel" morphology of this yeast form, where a mother cell is surrounded by a number of peripheral daughter cells (Hadad *et al.*, 1992). Under conditions of nutritional stress mycelia will transform and release propagules as either conidia or chlamydospores (Brummer *et al.*, 1993). It is suspected that contraction of the disease involves inhalation of propagules produced by the organism when exposed to sub-optimal conditions in the environment. The process of morphological transition and the biochemical distinctiveness of each form is the subject of ongoing research. Targeting a certain morphology or by inhibiting the morphological switching from non-pathogenic to pathogenic forms could prove a useful point of intervention for chemotherapeutics. This will only be achieved following a comprehensive understanding the molecular biology and biochemistry of this microbe.

Scientific investigations relating to the nature of *P.brasiliensis* began in earnest in the mid-1970s, with South American investigators performing rudimentary microbiological studies on this novel fungus. For example, the simple analysis of growth curves of

mycelial and yeast forms of the organism in trypticase soy broth proved worthy of publication, indicating the level of knowledge pertaining to this organism at that time (Arango & Restrepo, 1976). This is not to say that this rudimentary type of investigation desisted as knowledge progressed. For the last 25 years investigators have continued to perform very simple techniques to try and unravel some of the fundamental questions which persist about *P.brasiliensis*. For example, the outgrowth of the mycelial form on incubation in sterile distilled water demonstrated that intracellular storage compounds, in particular lipid vacuoles and glycogen deposits, allowed the continued development of the mycelia for over 6 months (De Bedout *et al.*, 1986). Such elementary observations are important, conveying the resilience of the organism and confirming it's ability to survive and increase in numbers in a hostile environmental niche. Later elementary studies have attempted to refine the laboratory handling of the organism with simple reports of differing culture dynamics on various media (Borba *et al.*, 1999).

There are considered to be three main morphologies of *P.brasiliensis*. The amount of information relating to each of these forms is variable but studies focus on either the mycelial form, the propagule form or the yeast form. Each of these is intricately associated with the ecology of the organism in its survival in the humid forest environment through to the lungs of an infected individual. The mycelial form of the pathogen is poorly described in comparison to the yeast or propagules. This is may be due to the fact that this phase is not thought to be directly associated with disease in any way. Only on sporulation will the organism be inhaled. In addition to the aforementioned studies relating to mycelial resilience there have been reports concerned with metabolism in this form. For example, the activity of cytosolic neutral proteinases, is much higher in mycelia than yeast. Inhibitor experiments implicated serine-, cysteine- and metallo-proteases in particular (San-Blas *et al.*, 1998).

The majority of information available concerning the life-cycle of *P.brasiliensis* relates to the yeast form or to the morphological transition from mycelia or propagule to yeast. The result of a temperature shift from 26^0C to 37^0C can be observed in Fig1.1. The entire transition process occurs over a period of approximately 15 days. During the transformation there is a cessation of respiratory function and oxidative phosphorylation that re-establishes in the presence of cysteine and the arrival at the yeast form (Medoff *et al.*, 1987). A number of researchers have performed preliminary analyses of the biochemistry underlying the transition event. San-Blas and San-Blas (1986) noted a shift in the chemical linkage of the glucan molecules of the cell wall with the yeast wall

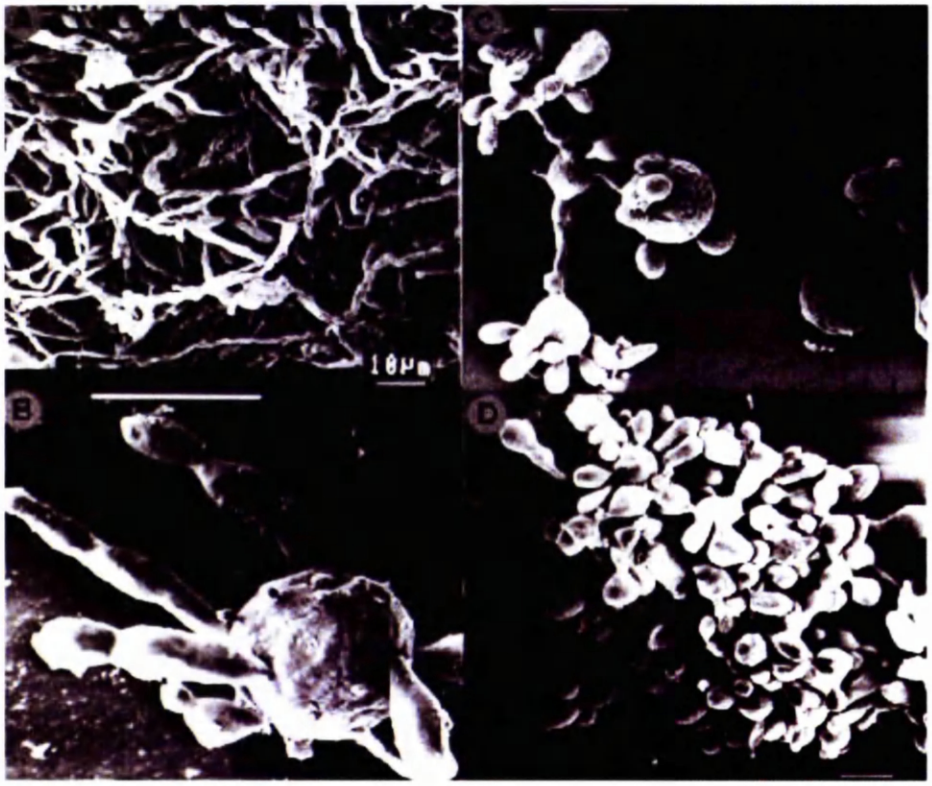


Fig 1.1

(Adapted from (Da Silva *et al.*, 1999)) The temperature induced morphological transition of *P. brasiliensis.*

A: Mycelia at 26^{0}C **B**: Transitional cells at 36^{0}C for 11 days,

C: Transitional cells 36^{0}C for 15 days **D**: Yeast cells at 37^{0}C

exhibiting α-glucan linkages while the myceial cell wall is predominantly β-glucan. Presumably this denotes the differential expression or activation of α-glucan synthases and β-glucan synthases during the phase transition. Indeed a β-glucan synthase from *P.brasiliensis* has recently been cloned and sequenced (Pereira *et al.*, 2000).

The composition of the lipid component of the cell surface also varies during the transition. The study of the glycosphingolipid repertoire of fungi receives attention because the structure and biosynthesis of these components differs from those in mammalian cells, thus providing a potentially useful drug target. The mycelial form of *P.brasiliensis* exhibits a 1:1 ratio of β-glucopyranolsylceramides to (*E*)-Δ 3-unsaturated fatty acids. The yeast form presents only 15% of the glycosphingolipid component as the (*E*)-Δ 3-unsaturated fatty acid (Toledo *et al.*, 1999).

At a more general level, discrepancies in both qualitative and quantitative profiles of protein expression have been noted between the two forms. Da Silva *et al.* (1994) record the differential synthesis of a number of stage specific proteins during the

mycelium to yeast shift. Most of these proteins were denoted only by merit of their difference in size while a 70 kDa protein, found only in the yeast form, was shown to react with the monoclonal rat anti-Hsp70 antibody. (The molecular biology associated by this differential expression has since been elucidated and will be discussed later.) This partly concurred with the investigations of Goldani *et al.* (1994) where a number of proteins were shown to be expressed or repressed in response to temperature stress on both yeast and mycelia. Similarly, the N-termini of two differentially expressed proteins were sequenced by automated protein sequencing, one specific to the mycelial form and one to the yeast. BLAST analysis of the peptide sequences so obtained demonstrated a mycelial enolase while the protein of the yeast was most probably an allergen protein (Cunha *et al.*, 1999). While it was suggested that this demonstrated differential gene expression no physiological significance was inferred.

Other metabolic functions have also been investigated with a differential profile of polyamine synthesis observed on transition. This occurs as a consequence of a modulation in ornithine decarboxylase activity which has been shown to be important in a number of fungal transition systems. In particular this ornithine decarboxlase activity has been mapped to the onset of the budding process rather than throughout the entire transition (San-Blas *et al.*, 1997).

It is worth noting here that specific observations of differential protein expression in a particular strain of *P.brasiliensis* does not infer that this will be the case for all isolates. Protein expression profiles in both mycelia and yeasts of different isolates demonstrates that heterogeneous protein synthesis can be observed. This is particularly true of the mycelial form (Salem-Izacc *et al.*, 1997). The issue is further complicated by the apparently unpredictable nature of *P.brasiliensis* expression profiles. An analysis of antigen batches, both inter- and intra-strain, showed that there is considerable instability in the synthesis of antigens even by the same isolate under tightly controlled incubation conditions (Franco *et al.*, 1996). If this observation of antigenic moieties is mirrored by other protein molecules then an accurate determination of the proteome of the organism at any given time could prove complicated.

It is most likely that the mechanics of the transition process will resemble those observed in other fungi and involve kinase dependant cell signalling mechanisms. In particular, it is likely that *P.brasiliensis* possesses a MAP (mitogen activated protein) kinase cascade and a cAMP kinase cascade similar to those observed in related species. This subject, in the context of pathogenic fungi, has been reviewed recently by Borges-Walmsley and Walmsley (2000). An initial indication that this is indeed the case comes from the isolation and sequencing of a *ras*-like gene from a cDNA clone derived from

the yeast form of *P.brasiliensis* (Borges-Walmsley, unpublished data). The influence of increased intracellular cAMP levels has been shown to stabilise the proliferation of the yeast form. It has been postulated that the induction of expression of the Hsp70 stabilises components of the cAMP signalling pathway in the face of temperature stress thus supporting the biochemistry of the transition process (Borges-Walmsley and Walmsley, 2000).

The yeast cells of *P.brasiliensis* have been shown to be quite metabolically active, despite the apparent fastidious nature of the organism on *in vitro* culture. Among the numerous enzymatic processes assessed in active yeast cells has been the measurement of phosphatase, esterase, dehydrogenase, and triphospatase activities. A comparison of cellular organisation between a number of isolates shows little variation in ultrastructure. A thin but dense cell wall composed of two sublayers encloses a three layered plasma membrane with extensive invaginations. Within the cells, mitochondria are profuse while endoplasmic reticulum tends to be poorly represented. Cells also present a number of vacuoles and multiple nuclei. Additional studies have shown that a number of lipid droplets appear within the cells and that there is a wide distribution in the size range of the oval elongated yeast when grown on plates (Brummer *et al.*, 1990).

Diagnosis of an active population of *P.brasiliensis* in infection has often relied on the microscopic observation of the "pilots wheel" appearance of cells, with a central mother cell possessing a number of associated peripheral daughter cells. A direct immediate diagnosis is difficult using this technique if the microbial load of the infection is low or sampling is difficult. In cases where numbers are low, diagnosis is complicated by the slow growth of the organism with cream coloured yeast colonies requiring 10 to 15 days of incubation.

As previously mentioned, nutritional stress can lead to the formation of conidia or spores. It has been suggested that the response resulting in the formation of conidia is specific to mycelial cells under nutritional or water stress (San-Blas, 1986). The beginning of the formation of conidia is a commitment that invariably results in the collapse of the hyphae and the persistence of the propagule. Investigations of this form demonstrate that these cells are also receptive to thermal dimorphism and reversion to either mycelial or yeast form is determined by the temperature applied. At 37^0C , either on plates or *in vivo* (in mice), the conidia revert to the yeast form. Microscopic observations see the uni-nucleated conidial cell change to the multi-nucleated yeast which is virulent and produces a pulmonary fibrosis in mice that equates to the same sequelae observed in humans (McEwan *et al.*, 1987).

The conidia are rather small, 3.5 to 5.0μm, which may explain why the organism can penetrate to the alveolus on inhalation thus further implicating conidia in the infection process. The intracellular organisation is little different from any eukaryotic cell with all major organelles present and correct within the protective environment of a thick spore wall. Nutrient stores are present, for example large lipid bodies are readily visible in the cytoplasm (Brummer *et al.*, 1990).

The other important propagule is the chlamydospore that forms in response to low oxygen tension and poor nutrient availability. Under certain culture conditions this can also reform the parent phase of growth. Like the yeast form, cells possess multiple nuclei and a high density of mitochondria meaning that further development may be possible (Franco *et al.*, 1989). Alternatively, this increase in mitochondria may be an oxygen scavenging response. Like the mycelial phase, this form of the organism is not extensively characterised.

1.1.2 Paracoccidiodiomycosis; disease and infection

PCM can occur in two forms, the severe and rare "acute", or "subacute", juvenile form or the chronic adult form. Both disease scenarios are associated with extensive sequelae, mediated by systemic infections, including lesions and an abnormality in the cell-mediated immune response. Infection can occur many years prior to the onset of clinical symptoms, disease occurring only on the outgrowth of active yeast cells. Disease onset usually coincides with a period of immuno-compromisation (a dramatic PCM has been found associated with HIV (Miranda Aires *et al.*, 1997)). A failure to administer therapy once PCM is recognised usually precedes mortality (Brummer, *et al.*, 1990).

The juvenile disease is the harshest of the two forms. It accounts for 3-5% of recorded cases and the progression of the disease is relatively rapid (weeks to months) as the disease is disseminated by the reticuloendothelial system. The prognosis for diseased children and young adults is poor (Tapia *et al.*, 1991). In contrast to the adult form the lungs do not appear to be the primary focus for the disease, although the organism can be detected in pulmonary exudate.

The adult form of the disease accounts for more than 90% of all patients. In this case the lungs are certainly the site of initial infection. Chest X-rays reveal lesions in the lower lung that can be extensive even while the initial clinical symptoms are mild (Campos *et al.*, 1991). Indeed some cases of PCM are only identified when the patient

reports with symptoms in a second organ while the pulmonary lesions are silent (Brummer, *et al.,* 1990). PCM is associated with a concurrent episode of tuberculosis in 10% of cases (Angulo-Artega & Pollack, 1972). Most of the patients are adult males. This may be due to the link between the deforestation industry and the contraction of the disease. However, an extensive case exists demonstrating that oestrogens in the female actually inhibit the conversion from propagules to yeast cells. The bias against disease in females and the effects of the female hormone status have been shown at the biochemical level (Feldman *et al.*, 1983; Loose *et al.*, 1983) and *in vivo* in model systems (Aristizabal *et al.*, 1998).

The actual method of disease contraction is still poorly understood. There is certainly no evidence for person to person transmission. It is widely suspected that an individual acquires the organism through inhalation of material from the environment. Where this environmental niche actually lies, at local or geographical scale, is not clear. The disease is definitely associated with tropical, humid, forested areas in South America. Any cases outside South America can generally be traced back to the endemic area, with the suspected contraction occurring on a visit by the patient often many years before (Brummer, *et al.,* 1993). Even within Brazil the contraction of disease in individuals can be correlated with visits to areas considered favourable to the survival of *P.brasiliensis.* For example, all 36 cases of juvenile PCM reported in the state of Rio de Janeiro between 1981 and 1996 were linked with children who had visited a nearby area of abundant watercourses and rainforest (Goncalves *et al.*, 1998).

Studies on the distribution of disease are geographically dispersed. The incidence of exposure and disease in Northwest Argentina has been studied and presented in a series of papers. The lower age limit for the immune response to paracoccidiodin in a skin hypersensitivity test appeared to be two years old but the numbers of children displaying this sensitivity were low (5 out of 344 children tested) (Mangiaterra *et al.*, 1996). The sequential assessment of paracoccidiodiomycosis, or exposure to *P.*brasiliensis, in the individual regions of Northwest Argentina suggests that 10% of individuals are seropositive in this area (Van Gelderen, *et al.*, 1999; Van Gelderen & Duran, 1995). In certain areas of Brazil as much as 80% of the population is thought to have been exposed (Brummer, *et al.,* 1993).

The association of disease with certain environmental features has led to the concept of "reservarea" (Goncalves, *et al.,* 1998). Whether it is the fauna, flora or soil that harbours the organism is unclear and in the dynamic ecology of the forest it is difficult to dissect. *P.brasiliensis* has been linked to animals. It has been suggested that freshwater fish and amphibians represent a significant enzootic reservoir for the fungus. Spread

from this pool could be exacerbated by the excretions of birds which feed on such fish (Conti Diaz & Rilla, 1989). Strong evidence links the armadillo (*Dasypus novemcinctus*) to the survival of the organism in the forest. Organisms can be isolated from the internal organs of such animals in the area of endemicity. This has been confirmed in a number of ways including immunological reactions with anti-gp43 antibody (a *P.brasiliensis* specific antigen) (Vidal *et al.*, 1995) and PCR (random amplified polymorphic DNA analysis) (Sano *et al.*, 1998). It would appear that the growth of *P.brasiliensis* in armadillos is quite common. The fungus was cultured from the organs of three out of nine armadillos captured in the area of Botucatu, Brazil (Bagagli *et al.*, 1998). Additionally, armadillos have been captured that possesses pulmonary lesions consistent with clinical PCM (Vergara & Martinez, 1998).

As an extension of the armadillo concept, it has been proposed that the soil can also act as a habitat for *P.brasiliensis* cells. Indeed the co-isolation of the fungus from soil and armadillos in the south-eastern area of Brazil has been reported (Silva-Vergara *et al.*, 2000). There are numerous examples of cases where *P.brasiliensis* has been isolated from soil samples. The screening of 887 soil samples from Botacatu resulted in the isolation of five strains of the organism, including three from armadillo holes (Montenegro *et al.*, 1996). Similarly a strain was isolated from soil around a coffee plantation in Ibia, South Brazil, and it was proposed that the organism was associated with the coffee industry (Silva-Vergara *et al.*, 1998). Many of these reports are greeted with scepticism however as the confirmatory tests employed are not always conclusive and the reproducibility of these isolations is often impossible (Franco *et al.*, 2000). Nevertheless, the confusing issue of the precise reservarea of *P.brasiliensis* and the route of infection remains unconfirmed. This is further complicated by the appearance of the organism in samples from sources as diverse as Antarctic penguin faeces and dog food (Sandhu *et al.*, 1997)

1.1.3 Molecular studies of *P.brasiliensis*

An increase in the application of current molecular techniques in the study of *P.brasiliensis* has coincided with its elevated profile within the scientific community. While the genomes of important Western pathogens are being sequenced, most of the key genes encoding virulence factors and antigens of *P.brasiliensis* remain undiscovered. A number of researchers in South America, the USA and in the UK are taking steps to remedy this situation. The sequencing of the genome of *P.brasiliensis* may be some time away but the isolation and characterisation of novel and important genetic determinants from this organism is underway. In addition, the application of

molecular biology in the study of the ecology of this organism, as well as the field of diagnosis, has allowed rapid progress and shows great promise for data gathering and practical application.

At the time of writing there are 89 submissions to Genbank which report sequences derived from *P.brasiliensis*. The majority of these sequences (52 of the 89) are multiple submissions reporting the sequence of one particular gene (*gp43*) from a number of individual strains. The group with secondary prominence relates not to protein encoding sequences but to rRNA genes and DNA sequences that have been used as genetic markers. The actual number of distinct, protein encoding genes is 13. Most of these genes are not the subject of publications and there is no biochemical data relating to their putative protein products. Those that have been investigated to some degree are discussed below.

As with so much of the biology of *P.brasiliensis*, recent molecular publications continue to deal with fundamental questions. Even the size of the genome and the chromosomal complement of the organism has been hard to pin down. The first analysis was performed in 1998 where 4 chromosomal bands were recognised on pulse field gel electrophoresis. These bands, ranging from 2.0 to 10.0Mbp, give a genome size of 30Mbp. This is consistent with the genome size and chromosome numbers of *Neurospora crassa* and *Coccidiodes immitis*. Microfluorimetry of DAPI stained yeast nuclei suggested a nuclear DNA content of approximately 60Mbp implying a diploid genome. Hybridisation of chromoblots with selected gene probes gave the interesting result that the laminin binding protein gene, *gp43*, was positioned on different chromosomes in two separate isolates (Nogueira Cano *et al.*, 1998). This intra-species variation was reinforced by a later electrophoretic karyotype study of environmental and clinical isolates. In all five environmental isolates and four clinical isolates tested, five chromosomal bands were observed. Four of these were common but the environmental isolates had a fifth chromosome of 7.2Mbp rather than the 8.8Mbp of the clinical strains (Montoya *et al.*, 1999). It is possible that the apparent discrepancy in the number of chromosomes reported in these two papers can be explained by differences in the electrophoresis technique used. It seems certain though that some sort of deletion and rearrangement occurs among isolates resulting in a change in the map of the genome. Chromosomes appear variable in size and the location of any one gene on any one chromosome seems variable also. This may have important ramifications in the event that the genome of this organism is sequenced.

The molecular cloning of genes from *P.brasiliensis* remains reportable. Included in sequence reports have been the sequences of cell wall synthesis genes including the

chitin synthases (CHS1 to CHS5) (Nino-Vega *et al.*, 1998; Nino-Vega *et al.*, 2000) and a putative β-glucan synthase (Pereira, *et al.*, 2000). Analysis of the mRNA levels of all five of the chitin synthase genes during the morphological transition illustrated that CHS3 was not expressed at any detectable level at any point. The other 4 were peculiar to the mycelial form reflecting the prominent chitin component of the mycelial cell wall (Nino-Vega, *et al.*, 2000). Such studies are of particular importance because it is in areas such as cell wall biosynthesis that we will find new drug targets and points of therapeutic intervention. Indeed testament to the importance of cell wall biosynthesis in fungal drug targeting is the almost exclusive attention that this area receives from current drug regimes (Georgopapadakou, 1998; Vanden Bossche *et al.*, 1998).

An *hsp70* gene, translated only by the yeast form of the organism, is reported by Da Silva *et al* (1999). This study delves deeper than the gene sequence and notes a putative novel mechanism of regulation of expression of the Hsp70 protein. It would appear that the expression of this gene is constitutive and transcripts are observed in both mycelia and yeast forms. However, it is only in the yeast form that the two introns are spliced allowing successful translation of the transcript. This ties in well with the previously mentioned study where an Hsp70 protein was differentially expressed in the yeast form only (Da Silva, *et al.*, 1994).

Two genes have been cloned that encode the 27kDa (McEwan *et al.*, 1996) and 43kDa (Cisalpino *et al.*, 1996) antigenic proteins. The study of gp43 in particular has generated a wealth of research relating to the immunobiology of PCM (discussed in section 1.4). The *gp43* gene was isolated from a lambda DNA library using short DNA hybridisation probes deduced from the protein on N-terminal sequencing. The isolated clone demonstrated a 1329bp ORF interrupted by one 78bp intron which coded for the 416 amino acid protein that is the immuno-dominant antigen of *P.brasiliensis*. The 27kDa antigenic protein was isolated on screening a lambda display library with *P.brasiliensis* antisera. The isolated clone possessed a 1kb cDNA insert which appeared to code for a 259 amino acid protein.

Diagnosis of *P.brasilisensis* by culture from clinical samples requires an extensive incubation of around 2 weeks before an accurate observation can be made by microscopy (Hadad, *et al.*, 1992). There has been considerable interest in the application of diagnostic PCR or nucleic acid hybridisation techniques based on newly discovered gene sequences. Such tests serve not only to identify *P.brasiliensis* in a rapid and more definitive manner but also to aid distinction from other pathogenic and closely related organisms like *Paracoccidiodes loboi* or *Paracoccidiodes cerebiformis.* The specificity and effectiveness of hybridisation to the 28S ribosomal RNA has been

demonstrated (Sandhu, *et al.*, 1997). The sequencing of 800bp of the 5' terminus of the *P.brasiliensis* 28S rRNA gene identified a 14bp motif that was specific for the species. This probe was shown to hybridise to all 7 isolates of the organism investigated, which ranged from clinical isolates to samples from soil and penguin faeces. Hybridisation to an extensive panel of other fungi were negative confirming specificity. This probe was later used in a more clinically relevant method, by *in situ* hybridisation in tissue samples. The 14bp probe and it's complementary sequence resulted in the labelling of 2-3% of the total cells present in the material (in comparison to analysis with Gridley stain) (De Brito *et al.*, 1999). This rather disappointing level suggests that this technique may not be suitable from routine use in the diagnostic lab.

The more sensitive application of PCR has proved more promising. PCR approaches include the amplification of sequences of particular genes and the use of random amplification polymorphisms to determine a characteristic amplicon signature for the organism. The application of these techniques has contributed to both clinical and environmental studies. A potentially global PCR target is the Internal Transcribed Spacer (ITS) genes which contribute to the 5.8S rRNA. The amplification of a 418bp product from this locus has resulted in identical outputs from 29 strains from diverse sources (Imai *et al.*, 2000). A similar result was reported by targeting sequences from both the 5.8S and 28S rRNA sequences (Motoyama *et al.*, 2000).

The sequence of the *gp43* gene allowed the design of primers which detected as little as 10 cells per ml of sputum (Gomes *et al.*, 2000). This was extended in a nested PCR method to correctly identify 21 of 23 culture positive infected BALB/c mice where as little as 0.5fg of *P.brasiliensis* DNA was present in the samples. This technique was specific for this fungus and failed to detect *Histoplasma capsulatum* in other mice (Bialek *et al.*, 2000). An equally promising ratio of positive results was reported in which PCR (with 10pg of template per sample) detected five of five positive blood samples while culture only suggested that two were positive. In this case the sequence of a previously identified PCR fragment was used that has not yet been attributed to any open reading frame (Goldani & Sugar, 1998).

Random amplified polymorphic DNA analysis has been used to group different clinical isolates of *P.brasiliensis*. The grouping suggested by the amplicon patterns mirrored the grouping made on the basis of virulence. In this case two groups were identified with a suggested genetic relatedness of only 17% (Molinari-Madlum *et al.*, 1999). Comparative analysis of amplicons between human and armadillo isolates has suggested a link in transmission to man from the armadillo. Three major clusters were observed in the RAPD analysis. Cluster II was exclusive to the armadillo but cluster I and III

patterns were observed in both human and animal isolates. This suggests that humans may contract the organism in exposure to the armadillo environment (Sano, *et al.*, 1998). Discrimination of isolates on a geographical basis by RAPD illustrated that a single primer was sufficient to divide isolates into five groups. Each of these groups related to the area where the strain was isolated with Venezuelan, Brazilian, Peruvian, Colombian and Argentinean strains all providing a different polymorphism (Calcagno *et al.*, 1998).

It is evident that the further application of molecular techniques in the study of paracoccidiodiomycosis will continue to provide new information. The culmination of these studies will hopefully be the sequencing of the *P.brasiliensis* genome. The degree of relatedness of genes between *P.brasiliensis* and those of other well characterised fungal pathogens means that information derived from the genome sequence would be immediately useful. The extensive biochemical data available for important fungal proteins could be suggestive of a phenotype for many of the genes discovered in this venture. It should be noted that the apparent heterogeneity in the overall structure of the genome among *P.brasiliensis* isolates may complicate this matter. Until the genome is sequenced the classical approaches to gene identification and characterisation will remain important.

1.1.4 The immunobiology and gp43 of paracoccidiodiomycosis

Studies of the mammalian immune response to paracoccidiodiomycosis focuses on two main areas. Firstly, the influence of immuno-effector molecules and immune cell functions have been addressed. However, the bulk of information relates to a single protein, the immuno-dominant antigen of *P.brasiliensis,* gp43. The literature dealing with gp43 represents the most comprehensive area of research relating to this organism.

The gp43 was first recognised in 1986 when the major antigenic exocellular components of the yeast form were identified after immuno-preciptitation with human antisera. The component that appeared to have the major antigenicity was purified and included a 43kDa protein glycoconjugate (Puccia *et al.*, 1986). This protein was digested by proteolysis and N-terminally sequenced. The sequence allowed the design of degenerate oligonucleotide probes. Screening of a lambda library using these probes resulted in the isolation of a 1329bp ORF, interrupted by two introns, encoding a 416 amino acid protein with a deduced molecular mass of 45,947Da (Cisalpino, *et al.*, 1996). A 43kDa component was found in circulating immune complexes (CICs) from the sera of PCM patients, probably representing the gp43 of *P.brasiliensis* (Unterkircher *et al.*, 1996).

The T-cell epitope, which induces a Th-1 response in BALB/c mice, has been mapped to a 15 amino acid motif with the essential core of "HTLAIR". This region, denoted as P10, is sufficient to induce a significant immune response that is comparable to that observed with gp43 (Taborda *et al.*, 1998).

As with so many of the proteins of *P.brasiliensis*, there is variation in the patterns of expression between isolates. The phase-specific expression of gp43 in yeast cells alone has been observed in some strains while a constitutive expression in both major forms is true for others (Mattar-Filho *et al.*, 1997). Polymorphisms of gp43 in individual isolates have been noted on the basis of differing isoelectric points. These polymorphisms arise from substitution sites within the 578-1166bp region of the gene. This is not the region containing the epitope highlighted in P10 and should not effect the antigenicity (Morais *et al.*, 2000). This means that vaccination with gp43 may still be an appropriate endeavour. Related to this notion is the partially successful application of purified gp43 as a paracoccidiodin in a delayed hypersensitivity test to assess the anti-*P.brasiliensis* immune status of animals and patients. In PCM patients the detection rate of the immune response was 48%, compared to the 28% success when using the traditional Fava Netto antigen (AgFN) (Saraiva *et al.*, 1996). Even more encouraging are the results of a DNA vaccination trial in BALB/c mice where the intramuscular and intradermal injection of a plasmid encoding the gp43 gene resulted in a characteristic T-CD4(+) Th1 response and a decrease in both the number of cells in the lungs and the extent of the dissemination of disease (Pinto *et al.*, 2000).

In addition to its immuno-stimulatory properties, a role for gp43 as an important virulence factor is emerging. The progression of most invasive microbial disease relies on an initial process of adhesion to the surfaces of mammalian cells. The notion of micro-organisms binding to extracellular matrix components is well established for many bacterial and fungal pathogens. It was discovered that a monoclonal antibody raised against the laminin-binding protein of *Staphylococcus aureas* binds to gp43 (Vicentini *et al.*, 1994). It would appear that gp43 is present not only as exoantigens but also on the cell surface. Electron microscopy of strains expressing gp43 on the cell surface shows binding of the cells to coated epithelial cell layers (Lopes *et al.*, 1994). Recently a model system, using Vero cells, has demonstrated that permanent attachment is observed after 30 minutes but this attachment is inhibited on addition of anti-gp43 antisera (Hanna *et al.*, 2000). This correlates with *in vivo* observations as the binding of free laminin to gp43 on the *P.brasiliensis* yeast cell surface, prior to injection into hamster testicles, greatly enhances virulence (Vicentini, *et al.*, 1994). The progression from protein production, cellular display and secretion has been monitored by immuno-electron microscopy. The gp43 is directed to the plasma membrane inside large dense

core vesicles. On fusion with the plasma membrane the gp43 extrudes into the cell wall and the immuno-reactive epitope is displayed on the cell surface. The gp43 is eventually released from the cell as membrane free material at scattered locations (Straus *et al.*, 1996). As well as adhesion to tissues, the intracellular parasitic nature of *P.brasiliensis* will require an initial attachment to macrophages. Anti-gp43 antibodies have a competitive effect on *P.brasiliensis* association to macrophages and consequent phagocytosis (Almeida *et al.*, 1998).

The development of a murine model of paracoccidiodiomycosis has assisted in the reinforcement of many of the postulations made on the basis of biochemical and epidemiological data. The first demonstration that PCM arises from the inhalation of conidia was performed in BALB/c mice. The intranasal introduction of 10^6 viable conidia cells into healthy animals was sufficient to induce symptoms of disease. Transformation to yeast commenced after 12 hours in the animal model with budding after 18 hours. Histology demonstrated a sequential recruitment of leukocytes to the alveolar cavity. The preliminary response saw the localisation of PMNs with lymphocytes, macrophages and plasmocytes appearing after 6 days (McEwan, *et al.*, 1987).

Much of the work performed since has attempted to dissect the dynamics of the cellular response and the involvement of certain immuno-modulators in the process. This has included systems using either gp43 or whole organisms. The presentation of purified gp43 to two isogenic mouse strains has demonstrated that the cellular response mounted against the antigen determines the degree of susceptibility of the individual to infection. In the resistant A/Sn mouse strain the antigen is processed and presented by macrophages. This induces a largely Th1 lymphokine response which correlates with the resilience against infection. Conversely, the susceptible B10.A mice use B cells as the main antigen-presenting cell, mounting a Th2 response. Thus, the nature of the APC component of the immune response would seem critical in determining whether the disease progresses (De Almeida *et al.*, 1998). The same comparative experimental system has been used to demonstrate that a decrease in CD8(+) cells allows a rapid and dramatic progression of fungal dissemination to the major organs (Cano *et al.*, 2000).

The clearance of fungal cells by PMNs is important and the fungicidal activities of such murine cells has been measured *in vitro*. The efficacy of PMNs in killing yeast cells would seem to vary according to fungal strains. The extent of fungal killing is enhanced by Interferon-γ (IFN-γ) (Kurita *et al.*, 1999; Kurita *et al.*, 1999) and granulocyte-macrophage colony-stimulating factor (GM-CSF) (Kurita *et al.*, 2000). The importance of IFN-γ has been shown *in vivo*. The depletion of IFN-γ in the mouse lung using

monoclonal anti-IFN-γ antibodies increased the susceptibility of both A/Sn and B10.A mice. The normally defined focal lesions of PCM were transformed into coalescent granulomata that destroyed the fabric of the lung (Cano *et al.*, 1998).

Finally, as suggested by the epidemiology studies of PCM, the female hormone status appears to convey protection from disease progression. A comparative study of the susceptibility of male BALB/c mice and female BALB/c mice at various stages of the oestrus cycle demonstrated that the lowest rates of disease were observed in oestrus female mice that presented high levels of blood oestrogen (Sano *et al.*, 1992). The previous observation that oestrogen's inhibit the mycelium to yeast transition *in vitro* (Restrepo *et al.*, 1984) also holds true *in vivo* allowing early low level infection to clear in females (Aristizabal, *et al.*, 1998).

In conclusion, *Paracoccidiodes brasiliensis* represents an important pathogen in South America. While the volume of published data relating to this organism is increasing the answers to several fundamental questions persist. It will be vital to address these and further issues if a comprehensive program of disease control is to be established. In so doing an often fatal disease could be circumvented and the wealth of data relating to the medical control of fungal pathogenesis could be enriched.

1.2 ATP-Binding Cassette Transporters

The relationship between a cell and the surrounding milieu requires dynamic but selective interactions. Processes at the cell boundary convey the ability to take up essential nutrients and expel toxic and/or waste products. This is universally achieved by membrane transport proteins incorporated within the lipid membrane barrier that defines the cell.

It is now recognised that such transport proteins can be limited to a number of families (Quentin & Fichant, 2000). Members of a given family show a degree of structural and functional relatedness, even though they can be widely distributed among prokaryotes, simple eukaryotes and higher eukaryotes. Despite a wide diversity in transport function, the relatedness of proteins within a family suggests a common evolutionary origin. The largest of these families, which is currently the focus of intense research interest, is the ATP-binding cassette (ABC) superfamily (also called the traffic ATPases). These proteins convey an energy dependent active transport process, by which substances can cross the cell membrane on the hydrolysis of ATP (Higgins,

1992). They have been noted not only on cell surfaces but also on the membranes of organelles (Allikmets *et al.*, 1999; Csere *et al.*, 1998; Leighton & Schatz, 1995). Many experiments have demonstrated that deletion of these transport proteins has a pleiotropic effect on cell function. Numerous ABC transporters display the ability to translocate a number of apparently unrelated substrates, with multidrug resistance associated transporters providing a prime example (Golin *et al.*, 2000; Van Helvoort *et al.*, 1996).

The first ABC transport proteins were found in bacteria in the 1970s, although they were not denoted as ABC transporters at this time. The histidine periplasmic permease system, HisJQMP, of *Salmonella typhimurium* was first isolated by Ames (Ames, 1972), whose laboratory continues to pioneer advances in the study of this complex. This permease, and the maltose transport system, MalEFGK, of *Escherichia coli* (Davidson & Nikiado, 1991) marked the beginning of the gradual discovery of the family. Almost 30 years later, the number of ABC transporters known is extensive and these molecules represent the largest accepted superfamily of homologous proteins. Research in this area remains broad but a few individual transporters have become representative of the ABC system due to the extent of the data relating to their function and/or their medical importance. The original pioneering system, the *S.typhimurium* histidine permease, remains prominent with the first X-ray crystal structure of an ABC transporter component coming from HisP (Hung, *et al.,* 1998). At a similar time the crystal structure of ATPase of RbsA was solved (Armstrong, 1998). The crystal structure of the ATPase subunit of the *E.coli* arsenite transporter, ArsA, has also been solved providing a structural insight into the detailed data concerning the mechanism of metalloid binding and ATP hydrolysis (Zhou *et al.*, 2000). Also, the homodimeric MalK of *Thermococcus litoralis*, has been solved to 1.9Å resolution (Deidericheş, *et al.*, 2000)

Yeast transporters are often used as model systems for eukaryotic transport processes. Notable proteins include the STE6 a-pheromone transporter, involved in mating of *S.cerevisiae* (Kuchler *et al.*, 1989). The PDR5 drug transporter represents the pivotal component in a network of proteins relating to drug transport in yeast, including ABC transporters and transcription factors (Balzi & Goffeau, 1995). The PDR5 drug transporter of *S.cerevisiae* is the fungal equivalent of mammalian P-glycoprotein in the respect that it now represents the archetypal transporter in fungal studies. A number of other ABC transporters have been found to convey a drug resistant phenotype in pathogenic fungi. Two ABC protein drug resistance determinants, CDR1 and CDR2, are now recognised in *C.albicans* (Balan, *et al.,* 1997; Franz *et al.*, 1998; Hernaez *et al.*, 1998; Sanglard *et al.*, 1997). Three ABC transporters have been described in *Aspergillus nidulans* (AtrA, AtrB & AtrC (Angermayr *et al.*, 1999; Del Sorbo *et al.*,

1997)). Thornewell *et al.* and Tobin *et al.* have successfully identified MDR transporters in *Cryptococcus neoformans* (Thornewell *et al.*, 1997), *Aspergillus fumigatus* and *Aspergillus flavus* (Tobin, *et al.*, 1997) using degenerate oligonucleotides for PCR. In mammalian cells the medically important *mdr1* gene product, P-glycoprotein (Pgp), and the cystic fibrosis associated CFTR transporter contribute heavily to the literature (Higgins, 1992).

Recently, the rate at which new ABC transporters have been reported is comparatively high and a further understanding of both structure and function promises scope for applied biology. Genome sequencing projects have allowed the assembly of comprehensive inventories of ABC transporters for certain organisms. For instance the current accepted figure for *Saccharomyces cerevisiae,* the first eukaryote for which the genome was complete, is 31 (Decottignies and Goffeau, 1999). The ABC transporter complements for *S.cerevisiae* and other organisms are further classified into a number of distinct subfamilies. While all of the proteins involved use ATP hydrolysis to pass a solute through the membrane, sequence comparisons show an underlying subdivision of relatedness within the family. With the draft of the human genome sequence now complete the putative inventory of human ABC transporters will surely be assembled. Indeed Klein *et al.* present a pre-emptive inventory in their recent review (Klein *et al.*, 1999). This review perhaps demonstrates that a large number of novel sequences of human ABC transporters remain to be mined from the human genome sequence. The authors note that only 30 human ABC proteins are represented by full sequences in databases. If *S.cerevisiae* has 31 genes then it is probable that the human complement will be much larger. Even within the 30 available we can see a much greater diversity in the array of subfamilies in comparison with yeast. Broccardo *et al.* (Broccardo *et al.*, 1999) note that ABC transporters tend to represent 2-5% of all ORFs in the genomes that have already been sequenced. Given that the human genome appears to possess around 30,000 genes we can expect in excess of 1000 ABC transporters if this trend holds true.

Due to a frequent bias in the literature, any given transporter is often associated with a single phenotype, e.g. Pgp is a drug transporter. This practice is perhaps misleading as many of these molecules display pleiotropic function. The YRS1 transporter of *S.cerevisiae*, which transports reveromycin, also allows tolerance to a broad range of organic anions not used as drugs (Cui *et al.*, 1996) and the mammalian P-glycoprotein has been implicated in the disposition of cholesterol suggesting a role in membrane biosynthesis (Wang *et al.*, 2000). It is thus possible that the phenomenon of multidrug resistance has appeared due to fortuitous substrate tolerance by existing transporters acting in normal physiological processes. Also any given function of these proteins is

often performed by additional transporters within the cell, ABC or otherwise. Thus, the cellular transporter arsenal is a complex system of both pleiotropy and pluralism that often complicates *in vivo* analysis. In lower organisms this can be partially circumvented. For instance Goffeau has created a multiple deletion mutant of *S.cerevisiae* which lacks all known drug transporters and associated transcription factors (unpublished data). Thus, candidate drug resistance mechanisms can be studied more accurately in this system. However, a similar approach in a mammalian cell line is not possible.

Biochemical studies have provided much information about the function of each subunit of ABC transporters and, to a lesser extent, the nature of the interactions between them. As the wealth of functional information is now complemented by crystal structures we may hope to understand these interactions in more detail.

1.2.1 The structural organisation of ABC transport proteins

All known ABC transport proteins are composed in a modular manner consisting of an arrangement of nucleotide-binding domains (NBDs) and membrane spanning domains (MSDs) (Fig1.2). It is generally accepted that there are four domains present in the functional transporter, two nucleotide-binding and two membrane spanning. The membrane spanning portions are highly hydrophobic and most often portrayed as six transmembrane helices. The two hydrophobic subunits form a "two-times-six" arrangement, thought to provide a pathway through which the substrate (or allocrite) can be transported. The other two domains are peripherally located on the cytoplasmic face of the membrane hydrolysing ATP and coupling the resultant energy to the transport process (Higgins, 1992).

Some ABC transporters possess additional domains to the four that form the core of the complex. The ATP-binding domain of the bacterial maltose transporter has a C-terminal extension that possesses an enzymatic activity totally unrelated to the transport process. The mammalian CFTR chloride ion channel possesses an additional domain, the R domain, which plays a regulatory role (based on phosphorylation) but has no apparent function in the transport process *per se*. This motif lies between the first nucleotide-binding domain (NBD1) and the second membrane spanning domain (MSD2). This gives CFTR an overall MSD1-NBD1-R-MSD2-NBD2 topography (Klein, *et al.*, 1999). A similar organisation exists in the emergent ABCA subfamily where the R domain is proposed to form an additional membrane associated region (Broccardo, *et al.*, 1999).

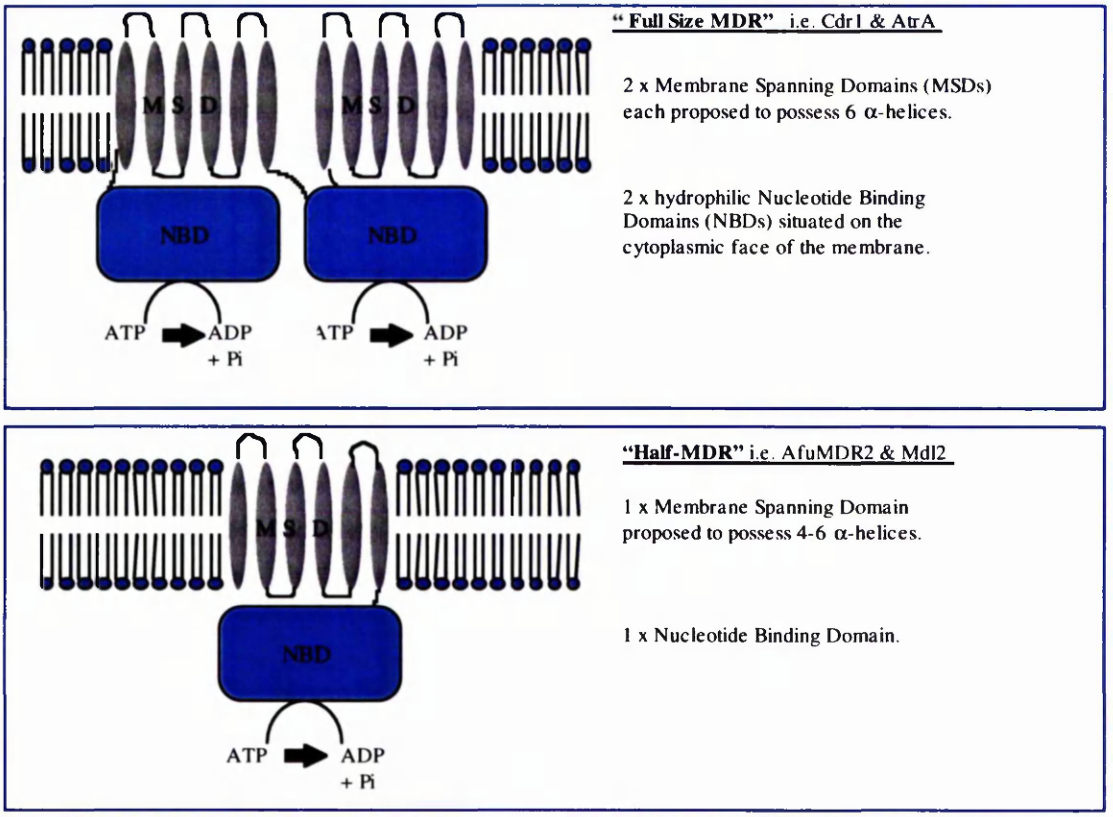


Fig 1.2

The general modular structure of ABC transporters. It is thought all transporters require 2 NBDs and 2 MSDs to function. A number of transport systems require associations between separate proteins the complete complex. This can include heterodimeric or homodimeric arrangements of half transporters. In bacteria the functional transporter can include up to four separate proteins (Higgins, 1992).

The distinctive MRP (multidrug resistance protein) family, related to CFTR, illustrates a set of proteins which appear to be extensions of a Pgp-like organisation. The most important of these is MRP1, identified by Cole and Deele in 1992 (Cole *et al.*, 1992). There now 5 other assured members of the family (MRP2-6). MRP1, 2, 3 and 6 have a similar structure with a large C-terminal extension providing 5 additional transmembrane helices. MRP4 and MRP5 restrict themselves to only 12 helices but have an N-terminal cytoplasmic loop as the extension. As well as having a role in the extrusion of cytotoxic drugs these proteins have a particular affinity for glutathione conjugated substances and have been shown to participate in organic anion balance and to transport immune effector molecules like leukotriene C4 (Borst *et al.*, 1999).

The logistics of expression of ABC proteins is found to be variable. The expression of the domains can occur as separate polypeptides or as fused combinations of subunits.

Almost all possible combinations of subunits have been reported. Generally, all eukaryotic ABC transporters show some fusion of subunits. Either the entire transporter is expressed as a single polypeptide or a half transporter is produced having one NBD and one MSD usually with six helices. It is thought that these half-transporters dimerise in a homologous or heterologous unit (Michaelis & Berkower, 1995).

In bacteria almost every conceivable combination can be found. The oligopeptide permease of *S.typhimurium* is expressed in four separate protein modules (MSD;MSD;NBD;NBD). In the *E.coli* ribose transporter the NBDs are fused giving (MSD;MSD;NBD_2) while the iron hydroxamate transporter has fused MSDs (MSD_2;NBD;NBD) (Young & Holland, 1999). The LmrA transporter of *Lactococcus lactis* is a dimer of $(NBD;MSD)_2$ (Van Veen, *et al.,* 1996). Many ATP dependant uptake systems of Gram negative bacteria have the additional requirement of a periplasmic binding protein (Linton & Higgins, 1998). Solutes can enter the periplasmic space through a non-specific uptake mechanism, e.g. by porin. This solute is then bound by a soluble periplasmic protein which directs the molecule to the correct transporter. The *S.typhimurium* uptake system is typical of this type with the periplasmic binding protein, HisJ, binding to the heterodimeric MSDs, HisQ & HisM. The histidine uptake is then performed on hydrolysis of ATP by the homodimeric HisP (Nikaido *et al.*, 1997).

Regardless of the organisation of expression, four modular units must be present in the functional end product. A number of experiments have been reported whereby the expression of normally fused polypeptides has been altered so as to cause them to be expressed as single units. It is found that these manipulated strains continue to produce functional transporters even though the organisation of expression has been changed. Berkower *et al.* (Berkower & Michaelis, 1991) used the STE6 a-factor transporter to illustrate this property. This protein is normally expressed as a single four subunit transporter. The experiment demonstrated that, if the protein was expressed as two half molecules, the a–pheromone continued to be transported as normal. In a later paper it was demonstrated, by immuno-precipitation techniques, that the two half molecules interacted physically, supporting the idea that such proteins spontaneously associate *in vivo* (Berkower *et al.*, 1996). It is probable that the fusion of subunits into larger arrangements has occurred over time as the single units were constantly co-expressed on the same operon. The above experiment demonstrates that the physical mechanism of protein-protein association persists even now the polypeptides are fused. As previously mentioned, no transporter has been found which functions with only two domains. For the moment, it is accepted that this form is the minimal requirement for

function. This does not rule out the notion of ABC transport processes occurring in larger, more complex arrays of subunits.

The necessity for two MSDs is apparent, with both domains contributing to the structure of the channel or pore. The requirement for two NBDs is perhaps less obvious. What is the significance and functional necessity for the duplication of the energy generating subunit of the transporter? The significance of this arrangement is contentious with investigators reporting different conclusions from different systems.

1.2.2 The Nucleotide-Binding Domain (NBD)

The provision of energy, which allows the transport of solutes against a concentration gradient, is achieved by ATP hydrolysis. The nucleotide-binding domains of ABC transporters, which possess this enzymatic activity, tend to be the most conserved modules of the complex. Typically values of 30% to 50% identity are observed depending on the proteins compared. The typical location for the nucleotide-binding domain is the cytosolic face of the cell membrane in peripheral association with the bilayer (Jones & George, 1999)

Each NBD is approximately 200 residues in length, possessing two highly conserved Walker A and Walker B sequence motifs (Michaelis and Berkower, 1995). These sequences are common to, and characteristic of, ATP-binding proteins. The Walker A motif is generally presented as being of sequence "GxxGxKS/T" where x is any amino acid. This is a general sequence and should not be considered as completely stringent for all proteins. The region is also known as the phosphate (P-) loop as it forms a direct contact with the α, β and γ phosphates of ATP (Schneider & Hunke, 1998), and is a common motif in a number of ATP and GTP binding proteins, not just ABC transporters. The P-loop is also conserved in ATP synthases, the myosin head piece, many kinases including adenylate kinase, RecA protein, the *ras* oncogene product p21 and the heterodimeric G proteins (Saraste *et al.*, 1990). It should be noted that other proteins which have an ATP/GTP-binding capacity do so without a P-loop structure.

The Walker B site is less conserved with the consensus sequence of "$Rx_{(6\text{-}8)}\phi\phi\phi\phi D$" where ϕ is any hydrophobic residue. The conserved aspartic acid residue is thought to interact with the magnesium of the Mg^{2+}ATP complex (Michaelis and Berkower, 1995). In the native protein the Walker A site and the Walker B site are orientated so as to form a nucleotide-binding fold. The A site interacts with the phosphates while the B site stabilises the association by interaction with the Mg^{2+} ion of the complex. While a lysine plays the critical role in the Walker A site an aspartate residue is central to the

function of the Walker B motif. These residues, and others important in ATP binding, will be discussed below in the context of the crystal structure.

Between the two Walker sites lies a helical domain which is followed by a third conserved motif known as the "signature sequence", or linker peptide, of the general sequence "LSGGQK/R/QR". Some non-ABC transporter proteins also exhibit this sequence, such as the UvrA of *E.coli* and the EF-3 elongation factor of yeast (Higgins, 1992). A fourth region within the NBD is the "centre region" which tends to be conserved among subsets of ABC proteins and hosts the CFTR phenylalanine mutation (ΔF508) that is most commonly associated with cystic fibrosis. Also, towards the carboxy terminus, after the Walker B site, lies a switch domain with the consensus x_9H. While there is no clear appreciation of the functional significance of these regions they have proved useful for identifying other members of the family (Schneider and Hunke, 1998).

In early studies of the energy dependent transport functions it was not clear that ATP was the actual energy source. In 1974, Berger and Heppel (Bergher & Heppel, 1974) stated that ATP was the energy source for the transport of glutamine and several other amino acids by shock sensitive bacterial permeases. The key to this conclusion was that a loss of transport function occurred on mutation of the F_0F_1ATPase or disruption of the electron transport chain by arsenate. This apparently clear cut result was conflicted however by later studies of glutamine transport in other labs. Membrane potential (Plate, 1979), direct electron transport and the use of acetylphosphate (Hunt & Hong, 1983) as a source of chemical energy were proposed.

As research progressed, the use of additional uncouplers allowed a better definition and it again became clear that ATP was in fact the energy source. Dean *et al.* (Dean *et al.*, 1990) successfully studied the maltose transporter of *E.coli* in membrane vesicles using a number of chemical agents to disrupt or enhance particular physiological conditions in the model. They found that transport was always observed in the presence of a concentration of ATP or a functional ATP generating capacity. Experiments like this, coupled with the extensive characterisation of a nucleotide-binding site in the transport protein lead to confidence in ATP as the true energy source.

The elucidation of the structure and function of NBDs is best illustrated using current examples of well characterised systems, in particular the two solved structures of bacterial NBDs. These two proteins, as well as the NBDs of MDR1 P-glycoprotein are currently the front runners in this field.

ArsA is the ATP hydrolysing component of the *E.coli* ArsA/B heavy metal transporter. This protein complex is responsible for the active efflux of the toxic trivalent metalloid anions arsenite (As(III)) and antimonite (Sb(III)). The *arsA* gene encodes a 583 residue protein which exhibits two nucleotide binding folds termed A1 and A2 and is thus equivalent to a fused dimer of NBDs. The hydrolysis of ATP by these domains is stimulated directly by the presence of arsenite or antimonite providing a system of positive regulation of nucleotide hydrolysis in the presence of the allocrite (Rosen *et al.*, 1999). It is, in part, the ArsA component of this transport complex that is beginning to unveil the necessity for two nucleotide-binding domains in the functional architecture.

The two ATP binding sites appear to interact (Li *et al.*, 1996). This was first indicated when a mutant was generated that was sensitive to arsenite. This mutation is constituted by the substitution of Gly^{15} to Cys. A screen for silencer mutations which would restore the phenotype to this strain generated a second mutant where Ala^{344} was altered to valine. This suggests a physical interaction between these two residues which are located in the proximity of the A1 and A2 NBDs respectively. Further substitutions in the area of Ala^{344} demonstrated that larger residues further increased the affinity for both antimonite and ATP illustrating the potential for physical interaction between the two nucleotide-binding lobes.

As well as the affinity for ATP, the ArsA binds the metalloid ions resulting in an allosteric activation of ATP hydrolysis. Antimonite is especially potent in this case with a concentration dependant induction of hydrolysis of between 10 and 20 fold. The binding site for antimonite was mapped by mutation of the four cysteine residues of ArsA. ArsA appeared to co-ordinate the Sb(III) ion within a pyramidal array of cysteines 113, 172 and 422. Mutation of any one of these residues substantially decreased the affinity for the metalloid. The feasibility of this conformation was further enforced when the distance between the thiol groups of the respective cysteines was mapped. Using the homobifunctional thiol linker dibromobimane the distance between the cysteine residues was proven to be 6Å (Bhattacharjee & Rosen, 1996), a distance consistent with the crystal structure of arsenite and antimonite bound to small molecule dithiols. The lengths of the As(III)-S Sb(III)-S bonds are 2.23Å and 2.45Å respectively. An inter-cysteine gap of 6Å allows the ion to be perfectly co-ordinated.

The interaction between the allosteric and catalytic sites of ArsA has been proposed to occur in the following way. The two NBDs within ArsA are independent with respect to mobility and are only connected by a 25 residue flexible linker region. In order for the aforementioned interaction to occur some sort of spatial constraint must force the two sites into close proximity. The co-ordination of the metalloid anion within the three

cysteines pulls the two lobes of the ArsA protein together to achieve this (Rosen, *et al.*, 1999). This is perhaps an over simplification however. The contribution of ATP to conformational change cannot be ignored. By monitoring changes in intrinsic tryptophan fluorescence it has been shown that the presence of MgATP is sufficient to induce a change in conformation. This is thought to be associated with the hydrolysis process as it is not observed with MgADP or with non-hydrolysable ATP analogues (Zhou & Rosen, 1997).

Fluorescence based techniques, including the measurement of variations in intrinsic tryptophan fluorescence, have been used to great effect in the dissection of the catalytic process of ArsA. Walmsley *et al,* have published detailed kinetic mechanisms for the hydrolysis of ATP. In so doing a more detailed scheme of the transient dynamics of conformational change has been assembled. For instance the rate-limiting step in the catalysis is a conformational change that occurs after hydrolysis and MgADP release. It is this conformational change that returns the molecule to the original state allowing it then to receive another MgATP. It is further proposed that the presence of the allosteric activator increases the rate of this conformational change thus increasing the rate of catalysis. The experimental evidence for this process is extensive and complex and is presented elsewhere (Walmsley *et al.*, 1999; Walmsley *et al.*, 2000).

Much of the biochemical observations of the mechanism of ArsA have now been explained on the publication of the 3D crystal structure of the entire protein, with bound MgADP and Sb(III), at 2.3Å resolution (Zhou, *et al.*, 2000). The homology in amino acid sequence between the two domains is reflected in the structure. Each NBD is "L" shaped with each arm extensively α-helical in composition. The binding site for the antimonite is distinct to that of MgADP, lying in the upper half of the protein, close to the linker peptide. There appears to be three clustered Sb(III) ions in ArsA. Both the ATP/ADP binding sites were occupied and positioned in close proximity in the lower half (see Fig1.3). The two P-loops are only 8Å apart at the closest point. As suggested by the earlier experiments, each nucleotide-binding site involves residues from the other domain. A1 relies on local residues for interactions with the α-and β-phosphates and Mg^{2+} of the ADP. However, two threonine residues (Thr501 and Thr502) of A2 are involved, Thr501 hydrogen bonding with a hydroxyl group of the ribose structure and Thr502 with the β-phosphate. The A2 binding site is similar to the A1 in that it also relies on some input from A1 residues. In this case the proximity of the interacting A1 residues is less pronounced with Ser210 positioned 6Å away from the β-phosphate. This is thought to give the A2 binding site an open conformation allowing easier access for the nucleotide than A1. This supports earlier biochemical

suggestions that A1 has a high affinity site with a low turnover while the affinity of the A2 P-loop is lower and more processive (Kaur, 1999).

The metalloid binding site does arise from the co-ordination of the anion by a number of cysteine residues, but not in the way previously suspected. Rather than a single metal ion co-ordinated by 3 Cys residues, 3 metal ions appear to be present, each anchored by 2 residues only. Indeed, one of the metal ions has no Cys associated, interacting with H148 and S420. The second Sb(III) involves C113 and C422 and the third C172 and H453. Thus the metal binding site of ArsA requires residues from both A1 and A2 and is not entirely cysteine dependant as first thought.

The allosteric activation of the ArsA ATPase activity on metal binding requires a signal transduction mechanism. The metal and ATP binding sites are too far removed to warrant any direct interaction. The crystal structure suggests an internal signal transduction adapter for each ATP binding site. At the upper, metal ends of the transduction paths are the histidines H148 and H453. Each serves A1 and A2 respectively. A stretch of 7 residues connects the ATP binding site to the metalloid binding site ($D_{142/447}TAPTGH_{148/453}$). The D142 and D447 residues co-ordinate with the Mg^{2+} ion via a water molecule and it is proposed that, when MgATP is bound, this interaction (prompted by metalloid binding) will induce ATP hydrolysis.

The *Salmonella typhimurium* histidine transporter represents not only the first ABC transporter to be identified but also the first transporter for which a component has been solved structurally. The histidine permease was first noted in 1970 (Ames & Lever, 1970) simply as a protein related to histidine acquisition. On the discovery of associated membrane components (Ames & Nikaido, 1978) the His transporter complex proved to be an ABC transporter belonging to the Gram negative bacterial periplasmic permeases. This group of proteins is now very well characterised. The currently accepted organisation for the histidine permease is that of an oligomer composed of four separate subunits making up a functional ABC transporter. A heterodimer of two membrane spanning domains (HisQ and HisM) is coupled to a homodimer of the nucleotide binding domain, HisP, this providing the organisation of $HisQMP_2$ (Kerppola *et al.*, 1991). The biology of the HisP component has been the focus of research in recent years and a wealth of information has emerged. This has been complemented by the elucidation of the 3D crystal structure of the protein in 1998 (Hung, *et al.*, 1998)

The study of the biochemistry of HisP was made increasingly possible by the purification of recombinant soluble HisP. The production of this protein in this way

had proved very problematic previously because of its tendency to be insoluble on over-expression. Indeed even in the aforementioned crystallography publication the pure recombinant HisP displayed a tendency to aggregate in solution even in the presence of detergents and could be partially stabilised using high concentrations of ATP and glycerol (Nikaido, *et al.,* 1997). The recombinant HisP displayed a high intrinsic rate of cation-dependant ATP hydrolysis, implicating the membrane HisQM components in a negative regulatory role. The HisP preparation appeared to be in a dynamic equilibrium between dimer and monomer states with the distribution between the two states being concentration dependant.

The ATPase activity of HisP in the context of an intact permease has been investigated (Liu *et al.*, 1997). The ATPase activity of the complex under various conditions was examined using a Mg^{2+} permeabilised proteoliposome system. It was found that a degree of co-operativity existed between all components and a high elevation in the rate of ATP hydrolysis was observed in the presence of a ligand coupled HisJ protein. HisJ is the periplasmic binding protein that directs the substrate to the permease across the periplasmic space. It had been proposed that the binding of this molecule to the permease complex would induce hydrolysis and this was indeed the case. This induction was specific to liganded HisJ as free histidine or unliganded HisJ had no effect when used alone or in competition assays with liganded HisJ.

Thus a picture begins to emerge regarding the regulation of NBD mediated ATP hydrolysis in relation to the membrane components of the permease and the presentation of the ligand. Free HisP displays a high level of ATPase activity even in the absence of liganded HisJ. This is not the case in the $HisQMP_2$ complex where ATP hydrolysis is inhibited except on association with liganded HisJ. The structural organisation of the entire complex is unknown so the spatial organisation of the components cannot be surmised. It has been proposed that the HisP dimer emerges through the HisQM component to display motifs on the periplasmic face of the plasma membrane (Baichwal *et al.*, 1993). Some have suggested the HisQM dimer is not only a regulator of ATP hydrolysis but is also protective, shielding the hydrophilic HisP from the lipid environment. Supporting evidence for this is the fact that the ATPase activity of HisP is inhibited in the presence of most phospholipids (Liu & Ames, 1998).

The relationship between substrate presentation and ATP hydrolysis has been reinforced by recent work involving the maltose ABC transporter of *E.coli.* The inhibition of maltose transport by vanadate occurs on the trapping of an ADP molecule in one of the nucleotide-binding folds. Vanadate does so by mimicking the transition state of the γ-phosphate of ATP during hydrolysis. The consequence of this trapping is

a high affinity association between the periplasmic maltose-binding protein (MBP) and the transporter and the loss of maltose affinity by the MBP (Chen, *et al.*, 2001). It would appear that the periplasmic binding protein is not merely a delivery system. It also stabilises an ATP hydrolysis transition state stimulating ATP hydrolysis.

As previously mentioned the crystal structure of HisP has been solved. This high resolution structure (1.5Å) shows the protein to be "L" shaped, having two thick arms (ARM I and ARM II) (Fig1.3). ARM I is predominantly β sheet with the nucleotide binding fold positioned at the bottom. A detailed analysis of the nucleotide-binding fold allowed the elucidation of the role of particular residues in hydrogen bond formation and water molecules co-ordination in the stabilisation of the ATP molecule. This P-loop region wraps around the β phosphate of the ATP to form extensive hydrogen bonds between the main chain nitrogens and the oxygens of the phosphates. A lesser degree of association is observed for the α and γ phosphates. The structure had no included cation as the presence of Mg^{2+} promotes catalysis of the ATP which would thus negate the possibility of acquiring the structure. The residues implicated in the ATP interaction in the crystal structure concur with previous point mutants that were unable to bind ATP.

The second arm is mainly α-helices and contains the conserved "LSGQQQRV" signature motif and is proposed to interact with the HisQM membrane component. The role of the signature motif remains unclear but mutation in this region greatly reduces the processivity of the permease. It may be involved in the signalling of liganded HisJ binding to the outer face of the transporter. A number of other ARM II residues, outside of this region, have been mutated and these mutants release HisP from the negative control of HisQM, so the regulation cannot be due solely to the signature motif. Alternatively, this motif may be crucial for HisP folding.

In keeping with previous studies on soluble HisP, the majority of protein present in the crystalline form was monomeric. Dimeric forms were found and a presentation of this organisation can be seen in Fig1.3. Interaction between the two domains occurs at a region of surface hydrophobicity, via the anti-parallel β-sheet on the outer face of ARM I. The negatively charged Asp11 is stabilised in some way but it is not clear whether this is through protonation or binding of a metal ion. Nevertheless, this represents an energetically unfavourable status and it is likely that the dimer is further stabilised on interaction with the HisQM dimer.

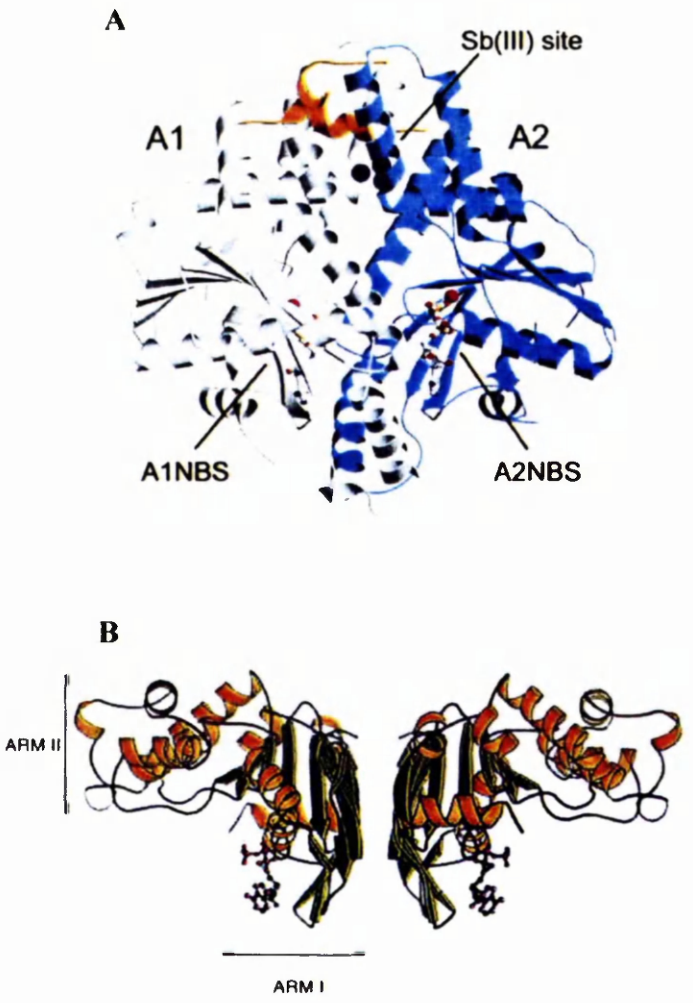


Fig 1.3

The structures of the ATP hydrolysing domains of ArsA (**A**) and HisP (**B**). Taken from (Zhou, *et al.*, 2000) and (Hung, *et al.*, 1998) respectively. HisP is shown in the dimeric form.

The crystal structure strongly supports the idea of dimerisation that has long been suspected for HisP. In keeping with findings for ArsA, it would appear that the two NBDs in the $HisP_2$ complex interact to promote hydrolysis of ATP. Nikiado and Ames recently demonstrated that only one functional NBD was necessary to fuel the translocation process and that the second NBD served to activate this hydrolysis by interacting with the hydrolysing fold (Nikaido & Ames, 1999). This role would appear interchangeable within the dimer as the system with only one functional hydrolysing domain catalysed at half the rate observed for the wild type complex.

The majority of information regarding the NBDs of eukaryotic transporters concerns MDR P-glycoproteins. The MDR1 protein was associated with a drug resistant phenotype when it was seen to be over-expressed in drug resistant human cell lines (Juliano & Ling, 1976). The ATPase activity of MDR1 was predicted later on the basis of homology with the NBDs of bacterial permeases and the realisation that there is a predicted P-loop in the amino acid sequence (Chen *et al.*, 1986). A number of similar genes have been found in humans and other mammals including mice. The MDR1 is expressed as an entire transporter with two NBDs, each NBD possessing one ATP binding fold (Higgins, 1992).

The recombinant expression of the N-terminal NBD (395-581) of murine MDR1 allowed the calculation of a millimolar K_m for ATP and demonstrated a functional ATPase activity. This affinity was enhanced to micromolar concentrations when the binding study used the ATP analogue 2',3'-*O*-(2,4,6-trinotrophenyl) (Dayan *et al.*, 1996). The stoichiometry of substrate transport versus ATP hydrolysis was initially measured at a near 1:1 ration of drug to ATP. The uptake of a $^{86}Rb^{+}$-valinomycin complex into vesicles indicated between 0.5 and 0.8 molecules of drug per ATP hydrolysed allowing the accumulation of $^{86}Rb^{+}$ to intra-vesicle concentrations of 8mM (Eytan *et al.*, 1996). As we shall see later, this would appear to be an underestimation in the light of more detailed kinetic analysis.

The use of fluorescence based spectroscopic techniques in ligand binding assays is made possible using fluorescent substrate analogues of ATP, such as MANT-ATP/ADP or TNP-ATP/ADP. These molecules a low level of fluorescence in free solution but when bound, and shielded from the quenching effect of bombarding water molecules, will give a greater, measurable fluorescence output. Experiments have been performed using a GST-MDR1 (NBD2) fusion protein, produced in *E.coli*, which bound both TNP-ATP and TNP-ADP. The resultant blue shift reported the incorporation of the substrate into a hydrophobic environment. Affinities for the ATP and ADP analogues were calculated at 1.8μM and 2.3μM respectively. As with 2',3'-*O*-(2,4,6-trinotrophenyl), the affinities of the analogues are considerably higher than for ATP or ADP (Sharom *et al.*, 1998).

The elucidation of the catalytic mechanism of ATP hydrolysis by P-glycoprotein has been a complex endeavour. Misleading deductions such as the variable estimations of ATP hydrolysis to drug transport stoichiometry, have hampered the rate of progress. The issue has been further complicated by the implications of substrate and modulator interactions which appear to alter the rates of catalysis. Sauna and Ambudkar have recently published two papers that go some way towards clarifying the issue. A more

definitive dissection of the catalytic events on substrate translocation indicates that two ATP molecules are required for the translocation of a single substrate molecule (Sauna & Ambudkar, 2000). The dynamics of the ATP interactions and the sterical consequences of hydrolysis would appear to answer the question of communication between nucleotide-binding sites and substrate-binding sites. In their recent work these authors use both radio-labelled ATP and the Pgp substrate Iodoazidoprazosin (IAAP) to track events at both sites. The catalytic cycle that leads to substrate translocation is very interesting. Initially, recruitment of an ATP molecule to either of the two NBDs is random. The binding of the ATP at one site prohibits nucleotide binding at the second site but greatly increases the affinity for the drug substrate. ATP hydrolysis of this initial bound molecule results in the release of a phosphate and the drug on translocation. This in turn conveys a greater affinity to the second NBD for ATP. The ensuing second ATP hydrolysis event serves not to transport drug but to reset the Pgp into a conformation that restores its high affinity for substrate (Sauna & Ambudkar, 2001). This raises two interesting points. Firstly, ATP hydrolysis does not strictly "energise" drug transport but rather imposes an intra-molecular communication that mediates the degree of affinity at the drug-binding site. Thus we have a dynamic release and capture mechanism. Secondly, it would appear that every other hydrolysis event does not result in substrate transport but in a re-setting of the conformation of the transporter.

The previous work indicates that communication between the two NBDs is essential. We have noted from the crystal structure of ArsA that amino acids from both NBDs play a role in each NBD. The likelihood for a similar strategy in P-glycoprotein is high. The spatial proximity of the two NBDs was recently demonstrated by disulphide cross-linking. Several mutations in the Walker A sites of both NBDs were generated resulting in the substitution of wild-type residues for cysteine. Following oxidative cross linking of the mutant Pgp two mutant forms were found to cross link; G427C(NBD1)/Cys-1074(NBD2) and L439C(NBD1)/Cys1074. The latter mutant retained significant drug stimulated ATPase activity and was studied further. It was discovered that cross linking of this form could be inhibited by MgATP suggesting an alternate "open" conformation in the presence of ATP and a closed conformation in the absence of nucleotide (Loo & Clarke, 2000).

As suggested above, the binding of additional molecules to NBDs can alter the rate of ATP hydrolysis, extent of nucleotide-binding or indeed the affinity or accessibility for transported substrate. While ATP hydrolysis results in substrate transport, the presence of a transported molecule, or a modulator with similar structure, can alter the rate of ATP hydrolysis. This has been demonstrated for a number of molecules. Initially it

was shown that the addition of the anti-cancer drug verapamil dramatically increases the rate of ATP hydrolysis. The phenomenon of drug stimulated ATPase activity for Pgp is now well documented (Loo and Clarke, 2000). Additional modulators, such as flavonoid compounds have been shown to interact with the C-terminal NBD of murine Pgp affecting reducing the binding of nucleotide (MANT-ATP) and steroid molecules (Consiel *et al.*, 1998). Steroid molecules can be transported by Mdr1 depending on their size. Indeed larger steroids, such as antiprogestin RU 486, are used to combat Pgp mediated drug efflux. This lead to the conclusion that flavonoids overlap distinct ATP and steroid binding sites. This has been demonstrated not only for Mdr1 but also the major drug transporter of *S.cerevisiae,* Pdr5 (Conseil *et al.*, 2000). In this case the binding of nucleotide is less affected but drug-binding (rhodamine 123) was severely perturbed. The binding of daunomycin to an ABC transporter of *Leishmania tropica* has been shown to by reduced by modulation by t derivatives of the flavonolignan silybin (Perez-Victoria. *et al.*, 2001)

1.2.3 The Membrane Spanning Domain (MSD)

Information on the membrane spanning modules of ABC transporters is far less than that for NBDs. The main reason for this is the difficulty in producing functional recombinant protein for biochemical study. Investigators must proceed through a lengthy and rigorous program of purification, solubilisation and reconstitution of membrane proteins stabilised in membrane or detergent environments. The yields of these proteins are usually far less than those obtained with the soluble cytoplasmic domains. As discussed below, this has hampered progress in this area but has not ruled it out completely.

Early predictions of protein structure gave rise to the "two-times-six" helix paradigm. That is, the majority of transporters will possess two domains each composed of six membrane spanning α-helices with the C- and N-termini on the cytoplasmic face of the membrane. This results in three extracellular and two intracellular loops (see Fig1.2) (Higgins, 1992). This notion was supported by the binding of epitope specific antibodies to the outer face of P-glycoprotein. These epitopes lie on the glycosylated sites on predicted extracellular loops (Zhang & Ling, 1991). However, even in the early stages of research in this area a few anomalies were noted. MalF of the bacterial maltose transporter has two additional N-terminal helices but these can be deleted without loss of function. On the other hand the HisQ and HisM proteins have only five helices each (Higgins *et al.*, 1982). As we shall see the "two-times-six" paradigm is now only considered a basic and general guide as most well studied proteins show some degree of departure from this organisation.

In contrast to the NBD, conservation of sequence amongst MSDs is not notable. Homologues between species are often highly conserved but comparisons between MSDs of different proteins shows considerable diversity. In some cases, where the membrane component of the transporter is a heterodimer such as in the HisQ/HisM complex, each distinct half relates well to the other suggesting a possible gene duplication (Higgins *et al.*, 1982). Given that ABC transporters are thought to have a common evolutionary origin, it is probable that the functional role of the MSD can be satisfied by a number of amino acid compositions. It may be that gross structural properties play a more important role than any local specifics relating to amino acid side chains. The constraints on mutation would therefore be lessened. The formation of an MSD pathway is generally considered to require the circular organisation of 12 amphipathic helices. This could be provided by several sequences.

A conserved motif has been noted in the MSDs of some ABC transporters including the bacterial periplasmic permeases. This region spans at least 20 residues, is of the sequence "EAA---G---------I-LP", and is located on the cytoplasmic face of the membrane. Mutating this EAA region in MalF and MalG has different effects on the activity of the maltose permease. While there is inequality in the roles of the sequence in both proteins, the simultaneous mutation of both totally perturbs function (Mourez *et al.*, 1997). The structure of HisP suggests that the signature motif of the NBD is involved in communication with the MSD. It is possible that it is the EAA motif that interacts in this case. This is supported by the fact that mutations in the putative helical domain of the MalK ATPase can suppress the mutant MalF/MalG phenotypes. The physical interaction between the MalK helix and MalF and MalG has been demonstrated in proteolysis protection experiments. Trypsin digests of free MalK demonstrates a trypsin digest site in this helix. The addition of MalG and MalF in inverted membrane vesicles protects this particular protease site from digestion (Mourez *et al.*, 1998). It is possible that the communication from one MSD affects only one NBD fold. The limited detrimental effect of mutating one of the two EAA's mirrors the limited effect of mutating one of the ATPase folds of HisP where the ATPase activity is reduced to half the maximal rate (Nikaido and Ames, 1999).

The simplest conception of the MSD is that of a pore or channel for substrate transit. The transport of hydrophilic moieties across the lipid barrier fits well with this notion. Indeed studies of the CFTR channel have contributed in supporting this idea. The channel of this protein was probed by cysteine scanning mutagenesis. The 24 residues of the sixth TM were mutated individually and in turn to cysteine and the accessibility of these residues assessed. The addition of the hydrophilic compound sulphydryl

methanethiosulphonate (MTS) compounds would modify exposed cystiene sulphydryl groups and reduce chloride ion conductance. The mutation of 11 of the 24 residues reduced the function of the transporters in *Xenopus* oocytes indicating that these residues were exposed on the pore lining surface of the channel. By varying the size of the MTS molecule it was calculated that the diameter of the pore was at least 6Å to the depth of Gln353 which is predicted to be the deepest of the positively modified residues (Cheung & Akabas, 1996). A broader view of important residues in the CFTR channel implicates prolines as key components but not necessarily as pore lining residues. The mutation of the 4 prolines in the CFTR MSD resulted in a reduction in the concentration of mature protein produced in HeLa cells. The conductance of any single channel was also affected with each mutant showing a reduced figure in comparison with the wild type protein. Different amino acids introduced at Pro^{99} had varying effects with glycine producing a minimal alteration while lysine and alanine induced very significant reductions in conductance . It would thus appear that these four proline residues play an important role in the stabilisation of the functional architecture of the pore although their precise orientation within that pore is unknown.

In many ABC transporter mediated processes it is hydrophobic substrates that are transported rather than hydrophilic. In these cases the simple notion of amphipathic helices providing a central hydrophilic core is not so appropriate. MDR in cancer cells requires the efflux of hydrophobic drugs such as vinca-alkaloids while antifungal resistance in pathogenic yeasts requires the transport of compounds such as azoles. Intestinal tract bacteria like *E.coli* encounter and efflux toxic bile salts and fatty acids. How are these transported through a hydrophilic channel? It has been suggested that the lipophilic nature of these compounds and their accumulation in the membrane allow some transporters to act as "hydrophobic vacuum cleaners". In this model the transporter would sequester the substrate peripherally from the inner leaflet of the lipid bilayer and into the pore for translocation (Bolhuis *et al.*, 1997). However, this only satisfies part of the problem, substrate acquisition. What about the interior of the pore? The suggestion that ABC transporters including MDR1 are involved in membrane physiology may hold the key. Pgp has been suggested to be a membrane flippase, transferring phosphotidyl choline from the inner to the outer leaflet of the bilayer (Higgins & Gottesman, 1992). It has been postulated that the pathway of drug efflux is similar and that the pore does not really exist but the substrate passed on the side of the MSD within the lipid bilayer. The recruitment from the membrane is supported by the observation that the MSDs of P-glycoprotein can bind drugs in the absence of the ATPase. Deletion construct analysis showed that when the two MSDs were fused and expressed alone in HEK239 cells they interacted with verapimil, vinblastine, caspiacin and cyclosporinA (Loo & Clarke, 1999).

More detailed biochemical investigations required in the over-expression and purification of Pgp. This was first reported by Dong *et al.* (1996), who purified P-glycoprotein from overproducing cell lines, retaining 70% of the starting material at 99% purity. This was achieved by extracting the protein from cells in SDS followed by buffer exchange into a dodecyl-maltoside containing buffer. This detergent solubilised protein was then reconstituted into lipid bilayers and was shown to possess an ATPase activity that was enhanced by verapimil and other drugs. A later adaptation of this procedure has permitted structural analysis by electron microscopy. It is now certain that a central pore exists in Pgp as image analysis has allowed the low resolution structure (2.5nm) to be visualised. The protein is a cylinder, 10nm in diameter, with a large central aqueous pore or pit (about 5nm in diameter) that is closed on the cytoplasmic face of the membrane. The depth of the protein is approximately 8nm, twice that of the lipid bilayer. While the cytoplasmic face is closed there is an opening on the side of the pore which allows access to the lipid bilayer supporting the flippase model (Rosenberg *et al.*, 1997).

In keeping with the ongoing controversy regarding the translocation strategy, the topography of the twelve transmembrane spans is also in question. Early work on the topography of these segments was achieved using fused reporter molecules to putative loops of the MSD. Geller *et al.* (1996) used topogenic reporters to study the topology of the N-terminal half of yeast STE6 in both *E.coli* and *S.cerevisiae*. In this technique the STE6 gene segment was fused to either alkaline phosphatase (in *E.coli*) or invertase (in *S.cerevisiae*). Alkaline phosphatase is active only when extracellular and invertase is heavily glycosylated when expressed extracellularly. Only this glycosylated form allows growth of the yeast on sucrose. This extracellular/ intracellular distinction makes these proteins ideal for topology studies. The result of this study was proposed as proof that the N-terminal MSD of the STE6 gene did contain a typical six helix organisation when expressed in the yeast or in recombinant *E.coli*. The use of a similar approach with Tap1, fused to a β-galactosidase reporters, illustrated the unconventional arrangement of 8 membrane spanning segments with large loops that would be exposed on the lumen face of the endoplasmic reticulum (Gileadi & Higgins, 1997). This was later confirmed using immunological techniques allocating 8 TMs to TAP1 and 7 to TAP2. In addition this study has proposed a three domain organisation for each TAP molecule with an MSD and cytoplasmic peptide-binding domain and a NBD (Vos *et al.*, 1999).

The early experiments may have provided an over simplistic view of the assembly and orientation of MSD components. The most common models involve simple mirror

symmetry of each set of six spanning segments. This gives a head-head/tail-tail orientation. In other words TM1 of the first MSD module is positioned close to TM1 of the second. Some recent papers continue to report biochemical data supporting this organisation. In the TAP1/TAP2 heterodimer complex, TM6 of TAP is stabilised and anchored in the membrane through the influence of the terminal TMs (TM4 +TM5) of TAP2. It is proposed that this interaction will require local interaction between these C-terminal MSD components thus supporting tail-tail interactions (Vos *et al.*, 2000). This is corroborated by cysteine cross-linking experiments with Pgp in which cysteines introduced to TM6 cross-link with TM12 (Loo and Clarke, 2000). This work focused on the fact that this cross-linking inhibits the ATPase activity implicating TM6 and TM12 in intra-molecular communication with the NBDs. It would appear that the cross-linking status of certain residues along the TMs was dependant on the occupation of the NBD by ATP suggesting that the two TMs move on hydrolysis. This cross-linking supports the head-head model.

A recent challenging review by Jones and George (Jones & George, 2000) presents a radically alternative model for MSD organisation in light of current advances that casts confusion on the results of the cross linking data. The essential point to the challenge of the head-head organisation is that the general round pore structure seems impossible in this model. If we accept that the overall structure of each MSD domain is similar then a head-head/tail tail channel would be shaped "((" rather than the required "O". The authors point out that the images of P-glycoprotein from the EM study clearly show an "O" pore. This can only be achieved with rotational head-tail/head-tail symmetry with TM1 being located beside TM12 and TM6 beside TM7. This would appear to contradict the cross-linking data. In response this apparent inconsistency a third symmetry form has been proposed, that of a "cyclone" pore. In this case the two MSD domains are opposed as in the mirror symmetry model but mis-aligned so that the pore arises only from TM4-6 and TM10-12 (Loo & Clarke, 2000). While this brings the TM6 and TM12 closer together it is questionable if it is close enough to reconcile the cross-linking data. Additionally, it seems unlikely that the substantial pore observed in Pgp could be formed by only 6 participating TMs.

The aforementioned radical topology of Jones and George relinquishes all notion of the two-times-six paradigm and forward the possibility that the TMs of ABC transporters are not truly α-helices. In this model the transporter has two pores in a double β-barrel organisation, each formed by the individual 6TMs of each MSD domain This dual channel model has some similarity to the outer membrane porins of Gram negative bacteria. The authors liken this structure to a "two stroke two-cylinder combustion engine".

1.2.4 The fungal ATP-binding cassette transporters

ABC transporters are universal and a number of key physiological process in fungi rely on the activities of this class of protein. *S.cerevisiae*, the first eukaryote for which the genome was entirely sequenced, has 31 genes encoding ABC transporters. (Bauer *et al.*, 1999; Decottignies and Goffeau, 1999). Screening and gene cloning experiments on other fungi has resulted in the identification of homologues throughout the genera. The best characterised of these systems tend to be those associated with multidrug resistance but other major functions include heavy metal resistance, organelle function and biosynthesis, protein synthesis and pheromone secretion. Many of these transporters display a considerable homology to those involved in human disease. The opportunity to use this convenient organism and its transporters as model systems for higher eukaryote processes has been embraced.

The ABC transporters of fungi have been characterised into a series of subfamilies based on sequence homology. The majority of transporters are grouped in the MDR and PDR subfamilies. MDR subfamily members display a considerable homology to mammalian Pgp while the PDR (Pleiotropic Drug Resistance) group are more distinctive as fungal proteins. Interestingly, it is the PDR subfamily rather than the MDR that is responsible for the bulk of drug resistance in *S.cerevisiae* (Balzi and Goffeau, 1995) and *C.albicans* (Cowen *et al.*, 2000). Other transporters belong to the MRP, ALDp, YEF1 and RLI subfamilies as described below (Bauer, *et al.*, 1999).

In *S.cerevisiae* there are four members of the MDR subfamily. The most studied by far is STE6, the transporter responsible for the excretion of a-factor pheromone from haploid *MAT*a cells. On screening mating deficient mutants, a strain was observed that produce pro-a-factor and mature a-factor intracellularly. This strain was termed *ste6* and the mutated gene was cloned and sequenced. It was found to encode a 1290 residue polypeptide (STE6) with considerable homology to P-glycoprotein (Kuchler, *et al.*, 1989). The protein is arranged in an MSD1-NBD1-MSD2-NBD2 manner constant with this homology.

Recently a 1323 amino acid homologue (HST6) was cloned during a *ste6*Δ complementation study using an expression library derived from *C.albicans*. It would appear that *C.albicans* uses a very similar mechanism for pheromone release. As with STE6, HST6 was expressed only in haploid cells of *S.cerevisiae* as the promoter element that facilitates repression of these genes in the diploid yeast is present. However, the gene is constitutively expressed to a high level in all forms of *C.albicans*

cells suggesting that the function of this protein may be different or that it has attained a more pleiotropic role in the pathogenic fungus (Raymond *et al.*, 1998).

The other characterised member of the MDR subfamily is ATM1, until recently *S.cerevisiae's* only known mitochondrial ABC transporter. This is a half transporter that appears to be responsible for the export of Fe-S proteins from the mitochondrial matrix, including cytochromes. An *atm1*Δ strain lacks all of the major cytochromes and accumulates high levels of free iron. Additionally, this strain is unable to use non-fermentable carbon sources. This protein has a human homologue, hABC7, that appears to mediate the X-linked disease sideroblastic anemia with spinocerebellar ataxia (Bauer, *et al.*, 1999).

Two other half transporters are also grouped with the MDR class, MDL1 and MDL2 (SSH1). The genes for these were isolated and cloned in 1994 following the application of degenerate PCR primers designed to target conserved sequences in genes of this type. Deletion mutants of both *mdl1* and *mdl2* showed no reduction in growth so these genes are not essential for viability (Dean *et al.*, 1994). A number of investigators carried out follow up experiments in attempts to ascribe function to these proteins but could not progress (Susan Michaelis; personal communication). Very recently, in closing stages of the writing of this thesis, MDL1 was shown to be responsible for the transport of oligopeptides from the yeast mitochondrion. It would appear that the cleavage of proteins by the mitochondrial m-AAA protease provides substrate for a 200kDa complex including MDL1. This complex can transport the oligopeptides through the mitochondrial inner membrane. The transport function MDL2 was also found to be evident in the membranes of mirochondria but was part of a larger independent 300kDa complex. Deletion analysis demonstrated that the MDL2 function was different than that for MDL1 and it would appear that MDL2 does not facilitate oligopeptide transport (Young *et al.*, 2001).

The PDR subfamily is the most studied in fungi. These proteins are involved in drug resistance and cellular detoxification and are the most numerous in the yeast genome. In addition to ABC transporters, the PDR network now has a number of associated regulatory proteins that control the expression of the ABC proteins. This important class will be discussed below in more detail along with fungal drug resistance.

Of the 7 proteins that appear to belong to the MRP subfamily, only two have been characterised to any degree, YCF1 and YRS1 (Yor1). Initial indications that MRP like proteins could be found in yeast was the detection of the transport of glutathione conjugated substrates into the secretory vesicles of a secretion deficient strain. The

protein responsible was found to be YCF1, a 1515 residue MRP1 homologue. YCF1 also has a weak homology to CFTR as it possesses a rudimentary R domain. It would appear that YCF1 transports not only organic glutathione conjugates into the vacuole but also glutathione associated toxic metal compounds. It has been shown that this protein is responsible for the sequestration of divalent cadmium, arsenite, antimonite and arsenate ions conveying heavy metal resistance to yeast. YRS1, or YOR1, was cloned in a complemetation study in a reveromycin sensitive strain. The *yrs1*Δ knockout strain showed increased sensitivity to this cell cycle arresting drug. Screening a number of anions revealed that YRS1 shared YCF1's ability to transport cadmium. Additionally, YRS1 can translocate a wide number of organic anions containing a carboxyl group (Cui, *et al.,* 1996). A third protein, BAT1, has also been identified that is competent in the transport of bile acids. This can not be the true physiological role in yeast cells but it does introduce BAT1 as a candidate for studying homologues involved in bile acid transport in mammals (Bauer, *et al.,* 1999). The remaining 4 ORFs have not been characterised and are represented only by submissions to the sequence database. Two of these ORFs (YLL015w (Z73120) and YHL035c (U11583)) appear to have the typical $(MSD\text{-}NBD)_2$ organisation. YKR103w and YKR104w (Z28328 and Z28329) appear anomalous as MSD-NBD-MSD and NBD only, respectively.

The two proteins of the ALDp subfamily β-oxidise long chain fatty acids and are homologs of the components of the human peroxisome adrenoleukodystrophy factor. Co-immunoprecipitation demonstrates that these two half-transporters PXA1 (or PAT2) and PXA2 (or PAT1) form a heterodimer. PXA2 is a fairly stable protein but PXA1 tends to degrade in the absence of PXA2. The complex has been hypothesised to form a peroxisomal fatty acyl-CoA transporter (Hettema *et al.*, 1996; Shani *et al.*, 1996; Shani *et al.*, 1995). The YEF3 and RLI subfamilies are the least characterised. These proteins have no TMs and their location within the cell remains unestablished. YEF3 may act as an elongation factor and function at the ribosome and exhibits ribosome stimulated ATPase activity. GCN20 has been implicated in the regulation of amino acid utilisation interacting with a cascade that modulates elongating ribosomes. All other members of these families are appreciated only as ORFs and are of unknown function (Bauer, *et al.,* 1999).

1.2.5 Fungal drug resistance and ABC multidrug transporters

With very few exceptions, antibiotics used in the treatment of fungal disease tend to target the wall of the fungal cell. Treatments gain specificity by directly or indirectly targeting ergosterol, the bulk sterol in the plasma membrane. In mammalian cells it is cholesterol that fills this role. Ergosterol is critical for a number of membrane functions

including fluidity, optimal cell division and the efficient activity of integral membrane proteins including chitin synthases. Antifungals act directly on fungal cells by interacting with ergosterol or by interrupting the ergosterol biosynthesis pathway and promoting defective membrane production and function (White *et al.*, 1998).

Only one therapeutic drug, 5-flucytosine, acts on targets that are not associated with the cells wall. 5-FC is taken up by fungal cells by a cytosine permease and then deaminated to 5-fluorouracil by a fungal specific cytosine deaminase. The fluorouracil is then converted to 5-fluoro-dUMP, a thymidylate synthetase inhibitor, or 5-fluoro-UTP that is incorporated into mRNA and disrupts protein synthesis. Resistance to this drug is widespread and it is rarely administered as a sole therapy. This resistance arises from a defect in the cytosine deaminase synthesis (Vanden Bossche *et al.*, 1997).

Direct targeting of the ergosterol is performed by polyenes, the most common of which is amphotericin B. These molecules are amphipathic allowing efficient ergosterol dependant intercalation with the membrane. The resultant porosity of the membrane results in the leakage of cellular components, especially potassium ions, and the breakdown of the proton gradient (Vanden Bossche *et al.*, 1994). Polyenes are not known to be substrates for ABC transporters. Resistance is not a common occurrence. This may be due to the lower frequency of clinical use. Amphoterecin B cannot be administered orally and usually requires a program of injections making its widespread application logistically difficult. Instances where resistance has been noted tend to be in severely immuno-compromised patients such as cancer sufferers.

As the precise mechanism of action of polyenes is unknown the elucidation of the specific causes of resistance is not possible. It will probably involve some modification of the plasma membrane composition as has been noted in resistant strains of *Candida* spp. and *Cryptococcus neoformans* with a marked reduction in the ergosterol content (White, *et al.,* 1998). Unfortunately, the basis for this variation in biosynthesis is hard to pin down as mutations in biosynthetic enzymes are variable in each isolate examined. It is probable that the interaction of polyenes requires a defined sterol to phospholipid ratio (Broughton *et al.*, 1991). Any mutation that alters this ratio will be beneficial so the range of mutations available is large.

More popular are the drugs involved in the interruption of ergosterol biosynthesis. Ergosterol is produced from squalene, a sterol precursor produced from acetate and acetyl co-enzyme A. The biosynthetic pathway involves the modification of squalene to lanosterol to ergosterol in a pathway involving 10 separate enzymes. Drugs which act on this biosynthetic pathway do so at several intermediate steps and are of four major

groups; allylamines, thiocarbamates, morpholines and azoles. The allylamines and thiocarbamates act on the same target, the first enzyme in the pathway, squalene epoxidase. The napthalene group common to each drug seems to inhibit the action of this *erg1* gene product causing the intracellular accumulation of squalene (Favre & Ryder, 1996). Morpholines act on the *erg24* and *erg2* products, C-14 sterol reductase and C-8 sterol isomerase. The mode of interaction is not appreciated. To date no resistance mechanisms have been reported for these drugs. They are applied infrequently but if this situation changes then resistance mechanisms may develop.

For the moment the favoured antifungal therapy remains the azoles and it is with these drugs that we find the greatest and most problematic drug resistance. Azoles act on C-14 demethylase (lanosterol demethylase), a cytochrome P450 molecule. The active site of this enzyme lies within a pocket located above the prosthetic group. Azole binding to this site is irreversible, thus, blocking sterol modification. Azole resistance occurs by a range of mechanisms but the two main sources are drug efflux or modifications of the C-14 demethylase. ERG11 mutations have been found in a number of clinical isolates of pathogenic fungi. These mutations tend to focus around the active site and would appear to block the access of the azole to the cavity. In many resistant strains, ERG11 mutation is not observed. Instead, the upregulation of expression of certain ABC transporters results in expulsion of the azoles before an effective intracellular concentration is established (White, *et al.,* 1998).

The process of ABC transporter mediated fungal azole resistance is best characterised in the PDR network of *S.cerevisiae* (Balzi and Goffeau, 1995). This system of ABC transporters and transcriptional regulators confers pleiotropic drug resistance, not only to a wide range of antifungals, but other drugs including anti-cancer agents. As well as a model for the MDR phenotype in mammalian cells the network provides a basis for the understanding of clinical drug resistance in pathogenic fungi. As will be discussed, the action of ABC transporters associated with this phenotype have been described in most major pathogenic fungi in both clinical and laboratory situations.

The first ABC drug transporter characterised in yeast was *snq*2. This 160kDa protein is located in the plasma membrane and confers resistance to the mutagen 4-nitroquinolone oxide and is implicated in cell cycle regulation. Deletion of this gene prolongs stationary phase and the biochemical preparation for the onset of exponential phase. (Decottignies *et al.*, 1995). Contemporary to the initial isolation of *snq*2, a second gene encoding an ABC transporter was reported simultaneously by a number laboratories. This gene was named *pdr5*, (Balzi *et al.*, 1994) *sts1* (Bissinger & Kuchler, 1994) and *ydr1* (Hirata *et al.*, 1994) but the accepted0 nomenclature is now *pdr5*. This new

transporter conveyed resistance to cyclohexamide, staurosporine, fluphenazine and sporedesmin when over-expressed. Both of these transporters are full sized proteins predicted to have 12 TMs and 2 NBDs. As drug transporters they are functional homologues of mammalian P-glycoprotein although they are not particularly alike on the basis of sequence homology. Indeed the organisation of these proteins differs in that the MSDs follow the NBDs rather than the other way around. Degeneracy in key locations are also present. The highly conserved lysine residue of the P-loop, conserved in all MDR subfamily proteins, is replaced with a cysteine in the N-terminal NBD of PDR proteins. Similarly the conserved glutamic acid of the N-terminal Walker B is replaced with asparagine. The functional significance of these alterations is not known (Del Sorbo, *et al.*, 1997).

These proteins are localised in the plasma membrane. They are relatively short lived with half-lives of 60-90 min prior to degradation. Degradation does not occur via the proteosome despite ubiquitination prior to internalisation (Egner & Kuchler, 1996). Instead the ubiquitin ligated protein is trafficked to the vacuole by endocytosis for digestion by vacuolar proteases. The dynamic control of the plasma membrane concentration of these two proteins may occur for a number of reasons. Firstly, the broad substrate competence of the transporters could mean that the constitutive presence at high concentration could be a result of the efflux of non-toxic or desirable compounds in the absence of toxins. Alternatively the efflux mechanism of PDR5 and SNQ2 may act at locations other than at the plasma membrane. The inward sequestration of substrate to either the exocytotic or endocytotic vesicle could expel or deliver these molecules to the vacuole respectively. In this case the turnover of the proteins through the vesicle transport system would be essential (Egner *et al.*, 1995).

Cells that lack PDR5 and/or SNQ2 remain viable but exhibit hypersensitivity to a wide range of mycotoxins, including azoles. However, the interest in the ABC transporters of *S.cerevisiae* has tended not to focus on antifungals or azoles particularly, although the basic information about MICs and sensitivities of various strains has been established. The continued use of yeast as a malleable model of higher eukaryotic function predominates. Despite the lack of sequence similarity and the inverted topology, PDR5 will transport substrates transported by P-glycoprotein. Rhodamine dye transport is often used as an assay to demonstrate functional homology between any given protein and Pgp. PDR5 is able to transport rhodamine 123 and rhodamine 6G. This assay affords accurate measurements by fluorometric techniques and allowed the demonstration of competitive interactions on rhodamine binding by taxol, vincristine and vinblastine (anti-cancer drugs) as well as ionophoric peptides and steroids (Kolaczkowski *et al.*, 1996). The proposed non-MDR role for P-glycoprotein, in lipid

transport and membrane maintenance, appears to be true of PDR5 and SNQ2 which transport the steroid oestradiol in intact yeast cells (Mahe *et al.*, 1996). This re-enforces the opinion that drug transport is an unfortunate consequence of substrate tolerance by proteins intended for other physiological processes.

Transported substrates of PDR5 can be structurally unrelated with extensive investigations highlighting hundreds of compounds that are extruded by these proteins (Bauer, *et al.,* 1999). Attempts to define the structure or chemical basis for substrate recognition have not been totally conclusive. The PDR5 transported Tri-*n*-alkyltin chlorides are a range of amphipathic compounds composed of anions and hydrocarbon chains of varying lengths. These chemicals were systematically applied to PDR5 transport assays, by gradually increasing the hydrophobic chain length or varying anion composition. While compounds with ethyl, propyl and butyl chains were transported efficiently, the processing of methyl and pentyl chains was poor. Also, the identity of the anion component had no bearing on the result. It was suggested that substrate competence in PDR5 is governed by molecular size and by the dissociation constant of the tri-*n*-alkytin anion. The three 2, 3 and 4 carbon chain length compounds had significantly higher K_d values than the methyl or pentyl forms. Hydrophobicity did not appear to be a factor although only a limited number of compounds were tested (Golin, *et al.,* 2000).

A recent significant paper implicated TM10 in defining substrate specificity and inhibitor susceptibility (Egner *et al.*, 2000). Mutations introduced in TM10 affected the ability of PDR5 to efflux the azoles itraconazole and ketoconazole as well as rhodamine 6G. Mutants S1360F and T1364F displayed a marked reduction in the transport of all these compounds. Only a slight reduction in ketoconazole transport was retained in the S1360A mutant while T1364A remained greatly impaired in this respect. Itraconazole transport continued to be greatly reduced in both. Thus, TM10 contributes to the binding sites for these compounds and interactions at these binding sites are not identical. FK506, identified as a transport inhibitor of the Tap1/Tap2 complex (Heitman *et al.*, 1993) also proved to be a PDR5 substrate. Mutant S1360A showed a very marked sensitivity to azoles through inhibition by this substrate. These data demonstrate that PDR5 drug export occurs through a common pathway involving TM10 although the precise nature of the interactions at this site varies with substrate. The impact of mutations on certain substrates is variable as is the affinity of interaction, shown by the high inhibition on azole transport by FK506.

PDR5 has two other homologues in yeast, PDR10 and PDR15, which have more than 65% sequence identity with PDR5. These proteins are also drug transporters but are

characterised to a far lesser degree. It would appear that these proteins are expressed in response to stress (Bauer, *et al.*, 1999). PDR12 is a close homologue of SNQ2 (60% identity) but, in contrast to SNQ2, mediates the transport of weak organic acids rather than drugs. This protein was found to confer resistance to water soluble carboxylate anions including sorbic acid, propionic acid and benzoic acid and has been implicated in the ability of yeasts to tolerate food preservatives and cause food spoilage (Piper *et al.*, 1998). Other PDR subfamily transporters have been identified, namely Pdr11, YOR011w and YNR070w, but nothing is known beyond sequence information.

Pathogenic fungi also possess PDR subfamily genes. *C.albicans* has 10 PDR genes encoding ABC transporters although only 2 of these have been linked to a drug resistance phenotype. The most important of these is *cdr1*, the first to be identified. This gene was found by functional complementation study of a *pdr5*Δ *S.cerevisiae* strain. CDR1 conveyed resistance to a wide range of antifungals including cyclohexamide, chloramphenicol and miconazole. The deduced amino acid sequence was consistent with PDR5 providing a full sized PDR type transporter (Prasad *et al.*, 1995). Later studies showed that deletion of both *cdr1* alleles in *C.albicans* resulted in a strain that was highly susceptible not only to azoles but also to allylamines and morpholines (Sanglard *et al.*, 1996). In addition to drug transport, CDR1 also mediates the efflux of human steroid hormones including β-oestradiol (Krishnamurthy *et al.*, 1998). This property could have implications for other fungal pathogens like *P.brasiliensis* as β-oestradiol prohibits outgrowth of the organisms and development of infection. In this case the ability to efflux this hormone could eliminate the resistance to this organism in women (Aristizabal, *et al.*, 1998).

A second gene, *cdr2*, was isolated using the same approach that uncovered *CDR1*. CDR2 has 84% identity to CDR1 and will transport similar substrates including the azoles ketoconazole and itraconazole. The double deletion of *cdr2* did not render *C.albicans* hypersensitive to drugs but the deletion when created in the *cdr1*Δ strain, increased the level of hypersensitivity (Sanglard, *et al.*, 1997). This implies that CDR2 is a secondary drug transporter that merely enhances the action of CDR1. Experimental drug resistance has shown that CDR1 and CDR2 over-expression is frequent in strains developing a hyper-resistant phenotype on exposure to increased concentrations of fluconazole (as well as mutated ERG11 and over-expression of the major-facilitator MDR1) (Cowen, *et al.*, 2000).

While the remaining 8 CDR transporters show sequence homology to these two proteins they do not carry functional homology with respect to drug transport. For instance CDR3 (56% identity to CDR1) fails to confer azole resistance and is found

expressed only in WO-1 opaque phase cells suggesting a role in cell-cycle regulation (Balan *et al.*, 1997).

Not surprisingly, similar transporters have been found in other *Candida* spp. The uptake of [^{3}H]fluconazole efflux in *C.glabrata* proved to be energy dependant with increased accumulation of drug on the interruption of respiratory function (Parkinson *et al.*, 1995). One PDR subfamily gene has now been cloned from this organism, CgCDR1, that confers azole resistance to the *pdr5Δ S.cerevisiae* strain and proves to be directly responsible for azole resistance in clinical isolates (Sanglard *et al.*, 1999). Similarly, *Candida dubliniensis* has CdCDR1 that mediates fluconazole resistance in clinical isolates and the yet uncharacterised CdCDR2 (Morna *et al.*, 1998).

Aspergillus nidulans has at least 3 ABC transporters. Two of these are full-size PDR class proteins designated AtrA and AtrB. AtrA is a constitutively expressed transporter that shows no affinity for any drugs tested while AtrB is upregulated in response the presence of drugs and is shown to confer resistance on over-expression in *S.cerevisiae*. However, the resistance was only 30% of that observed on over-expression of PDR5 (Del Sorbo, *et al.*, 1997). Deletion mutants of *atrB* are hypersensitive to all major classes of antifungals as well as some naturally occurring toxins such as camptothecin (Andrade *et al.*, 2000). The third transporter is AtrC an MDR subfamily transporter. This has not been shown to be responsible for drug efflux but does exhibit a 10-fold increase on expression on the addition of cyclohexamide. The ABC transporters described for *A.fumigatus* and *A.flavus* all belong to the MDR subfamily. Each species has one full size transporter gene, *afumdr1* and *aflmdr1* respectively. On over-expression AfuMDR1 conferred resistance to the antifungal cilofungin, an echinocandin B analog. *A.fumigatus* has a second transporter gene, *afumdr2*, a half transporter with extensive similarity to *mdl1* and *mdl2* of *S.cerevisae*. No function has been ascribed and the intracellular localisation has not been investigated (Tobin, *et al.*, 1997). It is unknown if AfuMDR2 or AflMDR1 will transport drugs as their competence to do so has not been addressed (Paul Skatrud, personal communication).

1.2.6 The regulation of ABC transporter mediated fungal drug resistance

Many papers reporting the sequences of new fungal ABC drug transporters add supporting evidence not only of drug transport but also of upregulation of gene expression in response to drugs. This has often proved a good indicator of the transporters role in drug efflux. When both *atrA* and *atrB* of *A.nidulans* were

examined, drug stimulated expression could only be detected for *atrB*. Only AtrB was able to transport drugs (Del Sorbo, *et al.,* 1997).

The trancriptional response to exposure to drugs is often a dynamic process co-ordinated by specific transcription factors. The PDR network includes five such transcription factors; PDR1, PDR3 PDR4 (Yap1), PDR7 and PDR9. The principal regulator is PDR1, first noted in a multiple nuclear mutation study (Balzi *et al.*, 1987; Meyers *et al.*, 1992). Over 20 independent mutations in the *pdr1* locus have been associated with increases in drug susceptibility and are often coupled with effects in other areas of basic cytoplasmic and mitochondrial function. These include respiratory deficiencies and resilience to stresses such as temperature, pH and sub-optimal osmotic environments. Once cloned and sequenced, PDR1 was recognised as a typical transcription factor containing a Zn_2C_6 binuclear cluster (Balzi and Goffeau, 1995).

The first gene shown to be influenced by PDR1 was *pdr5* and this regulatory relationship was the focus of one of the initial reports recording the sequence of *pdr5* (Balzi, *et al.,* 1994). Gel mobility shifts, using the N-terminal region of PDR1 in association with the promoter region of *pdr5*, demonstrated that a direction interaction occurred strengthening the postulate that PDR1 was a *pdr5* transcription factor. Further mapping of PDR1 responsive genes has associated most of the major ABC transporters of yeast mentioned above including all of the PDR subfamily, *ste6* of the MDR subfamily, transporters not of the ABC superfamily and an additional transcription factor locus *pdr3*.

PDR3 is a similar transcription factor to PDR1 (36% identity with a highly conserved DNA binding domain) (Balzi and Goffeau, 1995). In a separate study this transcription factor was associated with the transcription of *pdr5* explaining why a *pdr1* mutant results only in depressed transcription of *pdr5* and a mild hypersensitivity to cyclohexamide (Meyers, *et al.,* 1992). Not only PDR3 assists PDR1 but also PDR4, PDR7 and PDR9. Only *pdr3* appears to be regulated by PDR1 as well as exhibiting a positive feedback by upregulating it's own expression (Dexter *et al.*, 1994). In contrast to PDR1 these genes appear to influence a limited range of promoters. PDR4, 7, & 9 remain linked only with *pdr5*. PDR3 upregulates *snq2* as a *pdr1*Δ mutant shows a reduction of SNQ2 while in *pdr1*Δ/*pdr3*Δ the expression is abolished (Mahe *et al.*, 1996). PDR3 and PDR1 also contribute to the upregulation of *yor1*. Yor1 has an overlapping yet distinct substrate range in comparison with PDR5 (Kolaczkowski *et al.*, 1998).

The dissection of the *pdr5* promoter established the identity of PDR Response Elements (PDREs). The three PDREs act as binding sites for PDR1 and PDR3 and are of the general sequence "TTCCGCGGAA". While only one of these PDREs is required for PDR3 dependant transcription (Katzmann *et al.*, 1994) all three are required for the correct regulation in the wild type strain involving both transcription factors. PDREs have been recognised on the promoters of other PDR transporters that are responsive to PDR1 including *yor1* (Hallstrom & Moye-Rowley, 1998; Katzmann *et al.*, 1996).

A full understanding of the finer points of this regulatory network is not yet established. A number of other activating influences appear to play subtle roles and can be gene specific. For example, *yor1* possesses a PDRE but also an upstream activating sequence (UAS_{YOR1}) that positively influences expression. The influence of other transcription factors is also possible. The upregulation of expression of *snq2* in response to 4-nitroquinoline oxide requires YRR1, another zinc finger containing transcription factor that binds to a non-PDRE regulatory element (Cui *et al.*, 1998). Functional analysis of the *snq2* promoter denotes four positive sites for transcriptional activation. Both the YRR site and PDRE are in segment IV. The factors influencing transcription in the other three segments remain unknown (Cui *et al.*, 1999). Also the regulation of PDR1 itself appears to be an issue. As would be expected this transcription factor appears in response to stress but the signalling mechanism that stimulates PDR1 production is not understood. The activity of PDR1 is also regulated post-translationally by association with an Hsp70 called PDR13. The interaction of these two proteins further enhances the transcription of PDR1 responsive genes (Hallstrom *et al.*, 1998).

Regulation of ABC transporter gene expression in pathogenic fungi is not well characterised. A transcription factor from *C.albicans* has been isolated by complementation of a *pdr1Δ/pdr3Δ S.cerevisiae* strain. This protein, FCR1 is able to complement the mutations in this strain and provide fluconazole resistance in a PDR5 dependant manner. Conversely, the double deletion mutant of *fcr1* in *C.albicans* results in a hyper-resistant strain suggesting that while FCR1 is a positive regulator of *pdr5* it is a negative regulator of *cdr1* (Talibi & Raymond, 1999). It would seem that a direct transfer of information from *S.cerevisiae* to *C.albicans* might not be so straightforward in this case.

1.2.7 Conclusions

ATP-binding cassette transporters represent an exciting and important group of proteins that have great scope for biomedical and biotechnological application. However, the prominence of these systems in undesirable clinical scenarios, such as cystic fibrosis and multidrug resistance of cancer and infectious diseases, requires expedient and effective intervention. Only on fully understanding the mechanisms and functions of these transporters can we hope to intervene with widespread success. While important advances, including the crystal structures of HisP and ArsA and the extensive biochemical characterisation of P-glycoprotein, moves us towards this goal the questions still outnumber the answers.

Influencing transporter function in the disease situation is possible. For example, the treatment of non-insulin dependent diabetes mellitus has long been achieved using sulphonylurea drugs but the exact mode of action was not known. We now appreciate that the drug binds to an ABC protein to modulate insulin secretion (Aguilar-Bryan *et al.*, 1995). We have already mentioned the therapeutic use of steroids like antiprogestin RU 486 to reduce the anti-cancer drug efflux of P-glycoprotein (Consiel, *et al.,* 1998). Perhaps the increase in the study of transporter modulators that are not actually transport substrates will lead to the discovery of compounds that can competitively inhibit the efflux of beneficial therapeutics.

Intervention may be more difficult in cases where it is a defective molecule rather than an undesirable efflux activity. Current treatments for cystic fibrosis sufferers only improve the quality of life by physically dispersing mucus. The difficulties in positively influencing the CFTR channel with drugs are immense at this stage. Most attention in emerging treatment of this disease goes to the technique of gene therapy. Gene therapists are trying to introduce the functional *CFTR* genes to tracheal epithelial cells in viral vectors.

The consequences of ABC mediated drug resistance in pathogenic fungi may turn out to be severe. Azole antifungals have long been used as the preferred and most effective therapeutic measure to mycoses. The multiple mechanisms of azole resistance developed by fungi now threatens this efficiency. At this stage resistance is not a problem of disastrous proportions. It is most frequently reported in situations of intensive therapy, such as in immuno-suppressed individuals (Sanglard *et al.*, 1995). Mycoses that predominate in such individuals will be particularly amenable to the development and proliferation of these resistance mechanisms. Paracoccidiodiomycosis is one such disease. The added disadvantage in PCM is that there is no sequence

information available for any components of the ERG pathway or for potential efflux ABC transporters. It is prudent that we should attempt to redress this problem with some urgency.

Chapter 2

Materials and Methods

2.1 Reagents and Materials

2.1.1 Origin of materials and reagents

General chemicals were acquired from Sigma (UK) Ltd unless otherwise stated. Microbial growth media was supplied by Oxoid or Difco. Enzymes used in molecular biology were supplied by Roche Molecular Diagnostics unless otherwise stated. Additional materials or reagents for specialised applications are noted in the relevant sections later in this chapter.

2.1.2 Preparation of chemical and media solutions

General chemicals were weighed on a Sartorius digital balance and dissolved in distilled water or milliQ distilled/deionized water where appropriate. Small amounts were weighed on a Mettler Toledo digital fine balance.

2.1.3 Sterilisation of reagents and materials

Media and reagents to be sterilised were either autoclaved or filter sterilised depending on their stability on heating. Heat stable materials were exposed to (121^0C for 15 min at 4 atm of pressure). Heat labile liquid reagents were filter sterilised by syringe filtration through a 0.2μm Gelman syringe filter. Non-autoclavable items (e.g. centrifuge containers) were sterilised by thorough washing in 100% ethanol followed by a drying period in a laminar flow cabinet.

2.1.4 Centrifugation

Routine centrifugation of materials in microfuge tubes was performed in a Mistral Centaur benchtop microfuge. The centrifugation of microfuge tubes at lower temperatures and higher speeds was performed in a Sorval 1K15 benchtop

refrigerated centrifuge. In order to centrifuge larger volumes, a refrigerated benchtop Sorval 4K15 centrifuge was employed, using either Sorval rotors 11150, 12172, or 12256, depending on the type of tube and speed required.

2.2 Microorganisms; growth and storage

A list of all bacterial and fungal strains used in this thesis is presented in table 2.1.

2.2.1 Growth of bacteria

Escherichia coli strains were propagated in or on Luria-Bertani (LB) broth or agar, supplemented with the appropriate additions or antibiotics where appropriate. Such broth was generally prepared from a pre-mixed broth base containing 10g NaCl, 5g of yeast extract and 10g of bactotryptone for each litre of media to be produced. In order to prepare solid media, agar was added to 13.5% w/v prior to autoclaving.

Inoculation in liquid broth was achieved by transferring a single colony from a plate culture to an appropriate volume of medium. The culture was incubated, with shaking at 220rpm and 37^{0}C, in a baffled Erlenmeyer flask or 25ml conical bottomed universal container.

Culture on solid media was achieved by spreading or streaking organisms onto the agar surface using a sterile spreader or microbiological loop. Plates were placed in a 37^{0}C incubator, in an inverted position, for 18h.

2.2.2 Growth of *Saccharomyces cerevisiae*

Saccharomyces cerevisiae strains that did not harbour a recombinant plasmid were propagated in YPD liquid medium or SC minimal medium, with shaking at 30^{0}C, in a baffled Erlenmeyer flask. Culture on solid media was achieved on YPD or SC minimal agar (13.5% w/v) plates incubated at 30^{0}C for 72hr. YPD medium was composed of 5g yeast extract and 10g bactopeptone dissolved in 750ml of distilled water and autoclaved. Prior to use, filter sterilised glucose was added to a final concentration of 2%. The volume was then made up to1L with sterile distilled water. Agar was added at 13% w/v for a solid medium.

Microorganism	Genotype	Origin	Application
Escherichia coli Novablue	*end*A1 *hsd*R17 (r_{K12}^- m_{k12}^+) *sup*E44 *thi* -1 *rec* A⁻ *gyr*A96 *rel*A1 *lac* [F' *pro*AB *lac* I^q ZΔ*M15* Tn*10* (Tet^r)]	Calbiochem Novabiochem	General cloning
Escherichia coli XL1 Blue	*rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F' *pro*AB *lac* I^q ZΔ*M15* Tn*10* (Tet^r)]	Stratagene	General cloning
Escherichia coli XL10 Gold	Tet^RΔ *mcr* A 183 Δ(*mcr*CB-*hsd*SMR-*mrr*) 173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* Hte [F' *pro*AB*lac* I^qZΔ*M15*Tn*10* (Tet^r) Amy Cam^r)	Stratagene	General cloning
Escherichia coli Top10F'	F'{*lac* I^qTet^R} *mcr* A Δ(*mrr-hsd* RMS-*mcr* BC) Ł80*lac*Z ΔM15 Δ*lac*X74 *deo* R *rec* A1 *ara*D139 Δ(*ara* - *leu*)7697 *gal*U gal*K rps*L *end* A1 *nup*G	Invitrogen	pZERO™ cloning
Escherichia coli LE392	*sup*E44 *sup*F58 *hsd*R514 *gal*K2 *gal*T22 *met*B1 *trp*R55 *lac*Y1	Laboratory strain collection	Lambda phage infection
Epicurian coli ® BL21 -CodonPlus™ (DE3) -RIL	*E.coli* B F⁻ *omp*T *had* S($r_B^-m_B^-$)*dcm*⁺ Tet^r *gal end*A Hte [*arg*U *ile*Y *leu*W Cam^r]	Stratagene	pET21b mediated recombinant expression
Saccharomyces cerevisiae AD1-8⁻	Matα *pdr*1-3 *mq*3 *his*7 *ydr*1Δ::*his*G *snq*2Δ::*his*G *pdr*5Δ::*his*G *pdr*10Δ::*his*G *pdr*11Δ::*his*G *ycf*1Δ::*his*G *pdr*3Δ::*his*G *pdr*15Δ::*his*G	Gifted by Prof A Goffeau	Heterologous complemetation
Paracoccidiodes brasiliensis PB01	Wild type; genome not yet defined	Gifted by Prof M.S.Soures Felipe	Source of nucleic acids for PCR & RT-PCR

Table 2.1 Microorganisms and strains thereof, used in this study

SC minimal medium was composed as follows;

0.67% of :- yeast nitrogen base without amino-acids, with ammonium sulphate
0.01% of :- adenine, argenine, cysteine, leucine, lysine, threonine, tryptophan, uracil
0.005% of :- aspartic acid, histidine, isoleucine, methionine, pheylalanine, proline, serine, tyrosine, valine

The solution allowed to cool and filter sterilised glucose or raffinose was added to 20% or 10% respectively. The volume was then made up to 1L with sterile distilled water. Agar was added at 13% w/v for a solid medium.

Propagation of *S.cerevisiae* strains that possessed a recombinant plasmid containing the *ura3A* selectable marker gene was achieved in or on SC-U$^-$ minimal media. This media is prepared in the same way as SC minimal media, but the uracil is omitted.

2.2.3 Long term storage of bacteria

Long term storage of important bacterial strains was achieved by the accumulation of frozen glycerol aliquots of the relevant strains. A 5ml overnight liquid culture of the microorganism, in a relevant medium, was propagated in a 25ml universal container. Replicate stocks were prepared by mixing 850µl of the mature culture with 150µl of sterile glycerol in a 1.5ml microfuge tube. The mixture was then placed at –70^{0}C.

When it was necessary to retrieve the strain, a sterile microbiological loop was used to scrape a small amount of the frozen stock material from the surface of the aliquot that was then used to inoculate an agar plate. The glycerol stock was restored to -70^{0}C before thawing could occur.

2.2.4 Antibiotics

Filter sterilised antibiotics were applied to the following final concentrations unless otherwise stated: ampicillin 100mg/ml, carbenicillin 100mg/ml, kanamycin 50 mg/ml. Chloramphenicol stocks were prepared at 34mg/ml, dissolved in 100% ethanol.

2.2.5 Growth and preparation of *E.coli* LE392 for λ phage infection

A LB agar plate was inoculated with *E.coli* LE392 cells from a glycerol stock. Following incubation, a single colony was used to inoculate 50ml LB broth supplemented with 10mM $MgSO_4$ and 0.2% filter sterilised maltose. Growth was permitted for 6h at 37°C with shaking until the cells were in mid-log phase. The cells were harvested by centrifugation at 5000 rpm and 4°C for 10 minutes. Cells were re-suspended in ice cold 10mM $MgSO_4$ to a final $O.D._{600}$ of 1.0. Cells were then ready for use in infection or could be stored at 4°C for up to 14 days.

2.2.6 Propagation of λ phage on solid media

Plate lysates of individual λ phage clones or λ libraries were prepared on LB agar plates supplemented with 10mM $MgSO_4$ and 0.2% filter sterilised maltose. To a 1.5ml microfuge tube, 200µl of the previously prepared LE392 cells and 100µl of an appropriate dilution of lambda phage (suspended in phage or SM buffer) were added and mixed. The mixture was incubated at 37°C in a waterbath for 20 min to allow phage adsorption to occur. The infection mixture was then added to 5ml of molten (55°C) Top LB agarose (LB+ 6% w/v agarose, 10mM $MgSO_4$ and 0.2% filter sterilised maltose), which was then poured over the surface of the agar plates. The agarose was allowed to harden for 5 min at room temperature prior to inversion and overnight incubation at 37°C. Plaques were usually observed after 16h.

2.2.7 Propagation of λ phage in liquid media

Liquid lysates of phage were also required. These were prepared from either a single phage plaque or from total phage libraries.

A single phage plaque and surrounding LE392 cells was picked using an aseptically cut-off 200µl micro-pipette tip and ejected directly into 200ml of pre-warmed (37°C) LB broth (containing 10mM $MgSO_4$). Overnight shaking at 280rpm and 37°C generally resulted in the lysis of a dense culture of LE392 cells. Lysis was assessed by the propensity of cellular debris in the culture media. In the event that lysis was not complete the culture could be rescued by adding another 200ml of pre-warmed LB broth with supplements and splitting between two 2L Erlenmeyer flasks. A further 3h

of shaking would result in lysis. Once extensive lysis was observed, chloroform was added to 3% v/v and incubation continued for a further 15min to complete the process.

Liquid lysates from whole libraries were performed by infection and pre-adsorption as described in section 2.2.6, followed by inoculation of 200ml of LB (10% $MgSO_4$) in a 2L Erlenmeyer flask. Incubation conditions were as described above for the amplification of single plaques.

Phage particles were isolated from bacterial debris by centrifugation of the lysed culture at 10,000rpm and 4^0C for 10min. Lysate was then stored at 4^0C in the presence of 3% v/v chloroform.

2.3 Analysis and Purification of Nucleic Acids

2.3.1 Agarose gel electrophoresis of DNA

For routine analysis of DNA, electrophoresis in horizontal agarose* gel slabs were used. Agarose was dissolved to an appropriate percentage (0.7 to 2%) in TAE buffer by boiling in a microwave (TAE buffer; 4.84g Tris base, 1.142ml glacial acetic acid, 2ml 0.5M EDTA, made up to 1L in distilled water). Once the agarose solution had cooled to around 65^0C it was poured into a gel-tray. A comb was applied to the gel in order to create wells for sample loading and the gel was allowed to set for 30min at room temperature. Once the gel had solidified, it was overlaid with 10-20ml TAE buffer containing 10µg per ml of ethidium bromide. The ethidium bromide was allowed to diffuse into the agarose gel for 10mins. The gel was then submerged in an electrophoresis chamber containing TAE buffer allowing for approximately 5mm depth over the gel.

The DNA sample for analysis was mixed with 6X DNA loading buffer (15% Ficoll [type 400], 0.25% bromophenol blue, 0.25% xylene cyanol). The sample was then applied to a well of the gel. To each gel would also be added 5µg of 1kb DNA ladder (Roche Molecular Diagnostics) for size reference. Applying 100V of current to the chamber allowed the ionised DNA to migrate to the positive anode. Progression of the material through the gel could be monitored directly by visualisation of the xylene cyanol and bromophenol blue tracking dyes. Once the DNA had migrated a sufficient distance, it was visualised in the gel using an ultraviolet light transilluminator. When

required the illuminated gel was photographed using a Polaroid MP 4^+ instant camera system on Polaroid 667 film.

*On occasion when the purification of high molecular weight genomic DNA was required, LMP (Low Melting Point) agarose (1% in TAE) was used instead of standard agarose.

2.3.2 Small scale preparation of λ DNA

Small scale preparation of DNA from λ phage was achieved from either plate lysates or liquid lysates. Liquid lysates were prepared as previously described.

Plate lysates were prepared as previously described with a sufficient titre of phage so as to make the plaques confluent after 16h incubation. After this incubation the plates were overlaid with 5ml of phage buffer and the phage particles were allowed to diffuse into the buffer overnight at 4°C with gentle shaking. The overlay buffer was then aseptically removed and centrifuged at 10,000 rpm for 10 min. The supernatant was removed stored with 3% chloroform or used immediatley.

For each small scale DNA preparation, 15ml of liquid lysate, or overlay buffer from plate lysates, was added to a 25ml polycarbonate centrifuge tube. Both DNase and RNase were added to a final concentration of 1μg/ml and the lysate was incubated at 37°C for 20min. Following incubation the lysate was mixed with 6ml of ice cold phage precipitant (33% PEG 8000 / 3.3M NaCl) and incubated on ice for 2h. The precipitate was harvested by centrifugation at 12,000rpm for 10mins. The supernatant was discarded and the tubes allowed to dry for 40min at room temperature. The phage pellet was re-suspended in 500μl of phage buffer and transferred to a 1.5ml microfuge tube. Excess PEG was removed by vortexing the phage suspension with 1ml of chloroform for 30s followed by 3min centrifugation at 13,500rpm and 4°C. The upper aqueous layer was removed and processed by extraction with phenol (3 x 500μl treatments), phenol-chloroform-isoamyl alcohol (50:49:1) (1X 500μl treatment) and finally chloroform (2 X 500μl treatments). The DNA was then concentrated from the aqueous phase by ethanol precipitation.

2.3.3 Alcohol precipitation of DNA

DNA was concentrated from aqueous solution by the addition of either 100% ethanol or isopropanol to induce DNA precipitation.

Ethanol precipitation was achieved by the addition of 2.5 volumes of ice cold 100% ethanol and 1/10 volume of 3M sodium acetate (pH4.5) followed by incubation at on ice (or at -20^0C) for 20min to overnight. The precipitate was harvested by centrifugation at 15,300rpm for 35min at 4°C. The pellet was washed with an appropriate volume of ice cold 70% ethanol, re-centrifuged at 15,300rpm and 4^0C for 5min, and allowed to dry at room temperature. The pellet was then dissolved in an appropriate volume of sterile distilled water, 10mM TrisHCl pH8.0 or TE depending on the intended downstream application.

Alternatively precipitation was performed using 0.7 volumes of room temperature isopropanol followed by immediate centrifugation at 15,300 rpm and 4^0C for 35 minutes. The precipitated DNA was then washed with 70% ethanol, re-centrifuged and re-suspended in water or buffer as above.

2.3.4 Small Scale preparation of plasmid DNA

Bacterial strains harbouring plasmid constructs were cultured in 5ml LB broth, containing appropriate selecting antibiotics, in a 25ml universal container. Cells were harvested from 3ml of the turbid culture by centrifugation at 12,000 rpm for 1 min. Plasmid DNA was then isolated using the QIAprep Spin Miniprep Kit (Qiagen) as per manufacturers instructions and as summarised below.

The cell pellet was re-suspended in 250µl of buffer P1 (proprietary composition, containing RNase at 100µg per ml). The cells were then lysed by a modified alkaline lysis procedure (Birnboim & Doly, 1979) upon the addition of 250µl of solution P2 (proprietary composition, containing NaOH and SDS). Precipitation of proteins, genomic DNA and other macromolecules was achieved by the addition of 270µl of solution N3 (proprietary composition, containing high concentration of chaotropic salts). The precipitate was then separated by centrifugation at 14,000rpm for 10min. The supernatant from the mixture was then decanted onto a mini-column housing membranes containing immobilised silica gel.

The supernatant was drawn through the resin by centrifugation at 14,000rpm for 1min, the high salt concentration allowing the plasmid DNA to bind to the resin. The mini-column was then washed with 700µl of solution PE (proprietary composition, containing 80% ethanol). The wash buffer was drawn trough the mini-column by centrifugation for 1min at 14,000rpm. Residual wash buffer was further removed by a second additional centrifugation at 14,000rpm for 1min. The mini-column was then placed in a clean 1.5ml microfuge tube. To the column was added 50µl of hot (80^0C) 10mM TrisHCl pH8.0, directly onto the silica membranes, in order to elute the plasmid DNA under aqueous, low salt conditions. After standing for 1min, the eluate was collected by centrifugation at 14,000rpm for 1min. Plasmid DNA was then stored at -20^0C until required.

2.3.5 Purification if *P.brasiliensis* DNA from Low Melting Point (LMP) agarose gels

The DNA of *P.brasiliensis* (Pb01) was prepared in Brazil and supplied as a solution at various levels of purity. Generally a large quantity of intact *P.brasiliensis* DNA was evident but a background of degraded DNA, RNA and other fungal materials were present, which was undesirable for highly sensitive applications, such as Universal GenomeWalker™ PCR (see section 2.5). In this case a further purification of DNA was performed by electrophoresis of the material through an LMP agarose gel (Flowgen). The high molecular weight band was excised from the LMP agarose gel and dissolved in an equal volume of TE buffer at 60^0C. An equal volume of phenol, equilibrated with 0.1M TrisHCl pH8.0, was mixed with the slice and the aqueous phase separated by centrifugation at 4000rpm for 3min. The aqueous phase was further extracted with an equal volume of phenol /chloroform/ isoamyl alcohol (25:24:1) and then chloroform alone. The DNA from this aqueous phase was then isolated by ethanol precipitation on ice for 20min. The DNA was re-suspended in an appropriate volume of 10mM TrisHCl (pH8.0) and the quality/quantity assessed by agarose gel electrophoresis.

2.3.6 Extraction of DNA from agarose gels

DNA was extracted from agarose gels for the purposes of cloning or probe synthesis. This was generally achieved using the QIAquick Gel Extraction Kit, as per manufacturers instructions and as summarised below.

A volume (1-60µl) of the appropriate DNA solution was subjected to electrophoresis on an agarose gel containing ethidium bromide, as previously described. Visualisation of the resultant DNA band(s) by ultraviolet light allowed the excision of the required bands from the gel using a clean scalpel blade. The excised gel slice was then placed in a clean 1.5ml microfuge tube and the weight of the slice was then assessed. Up to 400mg of DNA/agarose was applied to each individual extraction procedure.

The gel slice was then mixed with a volume (30µl per 10mg of agarose) of solution QG (Gel solubilising buffer, proprietary composition). The agarose slice was allowed to dissolve in a 50^0C water bath for 10min with intermittent mixing. The solution containing the dissolved gel slice was then applied to a mini-column housing membranes impregnated with immobilised silica gel. Binding of DNA was achieved by drawing the solution through the mini-column by centrifugation at 14,000rpm for 1min. The column was washed with 500µl of solution QG to remove any traces of agarose (solution was applied to the column, allowed to stand at room temperature for 5min and then drawn through by centrifugation at 14,000rpm for 1min). The mini-column was then washed with 700µl of solution PE (proprietary composition, containing 80% ethanol). The wash buffer was drawn through the mini-column by centrifugation for 1min at 14,000rpm. Residual wash buffer was further removed by a second additional centrifugation at 14,000rpm for 2min.

The mini-column was then placed in a clean 1.5ml microfuge tube. To the column was added 50µl of hot (80^0C) 10mM TrisHCl (pH8.0), directly onto the membranes, in order to elute the DNA under aqueous, low salt conditions. After standing for 1min, the eluate was collected by centrifugation at 14,000rpm for 1min. Purified DNA was then stored at -20^0C until required.

2.3.7 Purification of DNA from buffered solutions.

On some occasions, for example during sequential digestion of DNA using restriction enzymes with incompatible buffer systems, it was necessary to purify DNA from a buffered solution prior to performing the next application step. This was generally achieved using the Wizard™ DNA Clean-up System (Promega) as per the manufacturers instructions and as summarised below;

The DNA solution was made up to a volume of at least 100µl using sterile TE buffer (pH8.5) in a clean 1.5ml microfuge tube. To the DNA solution was added 1ml of

DNA clean up resin. This was then mixed by inversion of the tube 8 times. The mixture was then applied to a mini-column via a 3ml Luerlok™ syringe. In order to wash the resin, 2ml of 80% isopropanol was transferred through the column, again using the syringe. The column was dried of isopropanol by centrifugation, in a 1.5ml microfuge tube, for 2min at 14,000rpm. DNA was eluted by the application of 50µl of hot (80^{0}C) sterile 10mM TrisHCl (pH8.0) to the column, standing for 1min, and harvesting the eluate by centrifugation at 14,000rpm for 1min.

2.3.8 Alternative Procedure for gel extraction/purification/ concentration and DNA

During the course of the research it was found that certain downstream applications performed better when DNA was extracted, purified or concentrated using the Qiaex II resin (Qiagen). This improved performance was notable in DNA ligations involving sticky ends but not particularly in A-T ligations of PCR products..

DNA was extracted from an agarose gel slice in a 1.5ml microfuge tube. The gel slice was first solubilised in 3 volumes of buffer QX (gel solubilisation buffer, proprietary composition), at 50^{0}C in a water bath, in the presence of 10-30µl of QIAEX II resin. The volume of resin was determined by the quantity of DNA applied. The resin was collected by centrifugation (all centrifugations in this procedure were at 13,000rpm for 1min) and the supernatant discarded. The pellet was re-suspended in 500µl of buffer QX to remove any residual traces of agarose and collected by centrifugation. The wash supernate was discarded and the pellet was washed twice in 500µl of buffer PE (as above) and collected by centrifugation. The buffer PE was completely aspirated by pipette and the pellet was allowed to air dry for 10-30 min. The drying time was variable with the amount of resin used. The DNA was eluted by vortexing the resin in 20µl of 10mM TrisHCl pH8.0 and incubation at 50^{0}C for 10 min followed by centrifugation. The eluate was transferred to a clean microfuge tube and the centrifugation repeated. The supernate was transferred to a clean microfuge tube and stored at -20^{0}C until required.

The procedure for concentrating DNA solutions or removing unwanted contaminants was identical to that of the gel extraction except the 10 min initial incubation time was performed at room temperature rather than 50^{0}C and the wash with buffer QX was omitted.

2.3.9 Purification of mRNA from total nucleic acid

RNA preparations from *P.brasiliensis* were partially prepared in Brazil and supplied as an ethanol submerged pellet of total nucleic acids. In order to purify high quality mRNAs from this pool it was necessary to eliminate genomic DNA and separate mRNA species from other RNAs.

The pellet was dried and re-suspended in 800µl of DEPC treated milliQ water. Once re-dissolved, 50U (50 µl) of RNase-free RQ1 DNase (Promega) was added and the mixture incubated at 37^{0}C for 30min. Inactivation of the RQ1 DNase was achieved by heating to 70^{0}C for 10 min.

The separation of mRNA from total RNA was achieved using Streptavidin Magnesphere® Paramagnetic Particles (SA-PMPs) (Promega). These particles are composed of a magnetite core coated with streptavidin. The streptavidin binds to a biotinylated oligo (dT) probe which is used to specifically interact with the polyA tail peculiar to mRNAs. Magnetic separation of the streptavidin-Magnesphere®-biotinylated oligo (dT)-polyA mRNA complex allows isolation from total RNA.

To a 1.5ml microfuge tube was added total RNA at a concentration in excess of 1.0mg in 500µl of nuclease free water. The tube was placed on heat block at 65^{0}C for 10min. Following this incubation, 3ml of oligo (dT) probe and 13ml of 20X SSC were added. The contents were mixed gently and allowed to cool to room temperature (10min cooling period).

Meanwhile, the SA-PMPs were prepared by re-suspending the particles in a 0.6ml tube as supplied. The re-suspended particles were captured at the side of the tube by exposure to a magnet housed in a magnetic stand (Promega). The supernate was removed and the particles washed three times in 300µl of 0.5X SSC, each time capturing the magnetic particles between washes. The washed particles were then re-suspended in 100µl of 0.5X SSC.

The entire contents of the oligo(dT) annealing reaction were added to the washed SA-PMPs and then incubated for 10min at room temperature. The magnetic particles were re-captured and the supernate removed. The particles were washed four times in 300µl of 0.1X SSC. Care was taken to remove as much wash buffer as possible after the fourth wash without disturbing the particles.

To elute, 100µl of nuclease free water was added to the SA-PMPs and the particles were re-suspended. The particles were re-captured and the supernate transferred to a clean RNase free microfuge tube. The elution process was repeated with 150µl of nuclease free water. The mRNA was then ready for application in RT-PCR.

2.4 Use of the Polymerase Chain Reaction (PCR)

2.4.1 Oligonucleotide primers for PCR

Oligonucleotide primers for PCR were synthesised by Life Technologies and supplied at a desalted purity on a 50nmol scale. Primers were re-suspended in a sufficient volume of 10mM TrisHCl pH8.0 to provide a stock concentration of 60pmol per microlitre.

Primer	Application	Sequence
MDR1	Degenerate PCR (Chapter 3)	GC(CT) CTC GT(GC) GG(GC) CCC TC(GC) GG
MDR2	Degenerate PCR (Chapter 3)	GAT (AG)CG (CT)TG CTT (CT)TG (AG)CC (AG)CC
PDR1	Degenerate PCR (Chapter 4)	GC(TC) GG(TC) AAG AC(TC) AC(TC) CTC CTC
PDR2	Degenerate PCR (Chapter 4)	(AG)GC (CG)GA GGG CTG GTG GAT (AG)GT
T7	General sequencing	ATT ATG CTG AGT GAT ATC CC
T3	General sequencing	AAT TAA CCC TCA GTA AAG GG
SP6	General sequencing	ATT TAG CTG ACA CTA TAG
fXbafwd1	*pft1* sequencing (Chapter 3)	TCT AGA AGA GAT ACA TAG
fXbafwd2	*pft1* sequencing (Chapter 3)	GGA TTG TGG GGG AAC GTA
fXbafwd3	*pft1* sequencing (Chapter 3)	GAA CTG ATG AAG GGC GTT
fXbafwd4	*pft1* sequencing (Chapter 3)	GTC AAT TCA TCA GTG ACT
fXbafwd5	*pft1* sequencing (Chapter 3)	GTG ATG CGG AGC CGG AGA
fXbarev2	*pft1* sequencing (Chapter 3)	CAT ATT TCT CCT TAG AAG
fXbarev3	*pft1* sequencing (Chapter 3)	GAT GAG GAT GTC TGG GTT
fXbarev4	*pft1* sequencing (Chapter 3)	CGT TGC TTC CTT GCG GAA
fXbarev5	*pft1* sequencing (Chapter 3)	CTA TTA TAT CGA TGG ACC

P2SPANF1	*pft2* cloning/sequencing	GCG ACG AAG ACA ATT TCC
P2SPANR1	*pft2* cloning/sequencing	CTA TCC TCT CCC TTA ATC
SPANIF1	*pft2* cloning	CAT GTT AAC TAT GTC TGC
SPANIF2	*pft2* cloning	AGG CTA ATT CCA AGA ACG
SPANIR1	*pft2* cloning	ACC AGA AAC GTG GTC AAG
SPANIR2	*pft2* cloning	TCC CCA GTC AAG AAA ACG
P2F2	*pft2* sequencing (Chapter 4)	AAC GTG AAG GCA AAC GAT TG
P2F3	*pft2* sequencing (Chapter 4)	AAG CGG ATC TTG AAC AAC TT
P2F4	*pft2* sequencing (Chapter 4)	GCT TGG GCT CTC TGG CGC AC
P2F5	*pft2* sequencing (Chapter 4)	GCC GGC GCT GGG AGA ATT GC
P2R2	*pft2* sequencing (Chapter 4)	TGC AAG ATG AAA GAC CAG GT
P2R3	*pft2* sequencing (Chapter 4)	CCA ACG GGA TGT TCA TCC TC
P2R4	*pft2* sequencing (Chapter 4)	GCA ATG CTG ACA CGT TTG CG
P2R5	*pft2* sequencing (Chapter 4)	TGT CTT CCG CCC AAC CCC AA
P2R6	*pft2* sequencing (Chapter 4)	AGA AGA ATC ACC ATT GTT TGA
Adaptor Tag	GenomeWalker™ PCR (See appendix 2)	5′ GTA ATA CGA CTC ACT ATA GGG CAC GCG TGG TCG ACG GCC CGG GCT GGT 3′ 3′-H_2N-cc cga cca 5′
Adapt1	GenomeWalker™ PCR	GTA ATA CGA CTC ACT ATA GGG C
Adapt2	GenomeWalker™ PCR	ACT ATA GGG CAC GCG TGG T
pftwalk1	GenomeWalker™ PCR	GAC GCG CTC GCG CAA CGC ATT TCT TTC GGA GTT
pftwalk2	GenomeWalker™ PCR	GGT ACG CGA GGC ACT GCG CTT TTC AGC AAT GGC
SMART Oligo	5' SMART RACE RT-PCR	AAG CAG TGG TAA CAA CGC AGA GTA A(T_{30})N_{-1}N (N=A,G,C or T & N_{-1}=A, G OR C)
UP1	5' SMART RACE RT-PCR	CTA ATA CGA CTC ACT ATA GGG CAA GCA GTG GTA ACA ACG CAG AGT
UP2	5' SMART RACE RT-PCR	AAG CAG TGG TAA CAA CGC AGA GT
PT72	5' SMART RACE RT-PCR and SSP-PCR	GCA AGC AGA GTA ACG CC

PT7	5' SMART RACE RT-PCR	AGC CGA GCC TCT TCC CGC
P1Pro	*pft1* RT-PCR	CAC ACA TAT CGC GCT ATC
XS62	*pft1* RT-PCR	TTC ACA TTT GCT ATC CAT CT
XT72	*pft1* RT-PCR	TAT TGT TGC CGC ACT AA
XT74	*pft1* RT-PCR	GTT CGG CCT CAC ACT CCC CGT
P1(g)RTfwd	*pft1* RT-PCR / expression	GCC ATG CTT GCC GCC TTC ACT CCA CGA
P1(r)RTfwd	*pft1* RT-PCR / expression	GCC ATG GCG TTA CTC TGC TTG
P1RTrev	*pft1* RT-PCR / expression	TTT TTC TTG GGA CGG CTG TTG
P2RTfwd	*pft2* RT-PCR	CCC GCG CTG CAG CTG CAA AG
P2RTrev	*pft2* RT-PCR	GGA AGT AGA TGA TTA GGT TG
HSPfwd	RT-PCR control	GGT CTA GAA AGC TAC GCC TA
HSPrev	RT-PCR control	CCT CCT CGA CAG TAG GAC CG
P1NBDF	Heterologous expression	CGG GAT CCG AGA ATT TTT GAG CTG GAA AAG AGA GAG
P1NBDR	Heterologous expression	GTG CTC GAG TTT TTC TTG GGA CGG CTG TTG CTT
P1NBDTF	Heterologous expression	CGG GAT CCG ACC AGA CCC GCA GTG ACT GTA TTT
P1NBDTR	Heterologous expression	GTG CTC GAG CGT ATT GTT GCC GCG GAG TAA CGC
P1NBDTWR	Heterologous expression	GTG CTC GAG AAC CAC CGA CTG CAC AGC CCC CCC

Table 2.2 Oligonucleotide primers used for PCR.

2.4.2 Routine PCR

PCR was routinely carried out using HotStarTaq™ thermostable DNA polymerase (Qiagen) and its associated buffers. A number of PCR reactions were performed which required special conditions (optimisation of reagent concentration or the use of Q-solution). Cycling parameters varied with each PCR and primer set but a general example of both cycling parameters and reaction mixture composition is presented. Particular PCR conditions for each reaction will be provided with the appropriate result in later chapters.

Reaction Mixture Composition:	(final amount)
Final volume	100µl
MgCl	1.5mM
dNTPs	200µM
Primers	1.5µM (each)
10X buffer	1X
Polymerase	2.5 units
Template DNA	pg - µg*
Sterile milliQ water	to 100µl

* Template DNA concentration varied depending on the nature of the template. Either pure DNA, ligation reaction mixes, cDNA reverse strand synthesis reactions or untreated lambda phage were used in reactions.

Cycling parameters

1]	1 cycle	95°C	15min (Taq activation)
2]	35 cycles	94°C	30s
		50-65°C	40s
		72°C	30s-5min
3]	1 cycle	94°C	30s
		50-65°C	40s
		72°C	10min

PCR was performed on either an Eppendorf Gradient Thermocycler or a Techne Cyclogene Thermocycler. Both of these thermocyclers possess heated lids. These lids were set to 105°C, removing the need for a mineral oil overlay in each reaction.

The use of HotStarTaq™ reduces the need for a hot-start protocol in PCR. However, reactions were set up on ice and transferred from ice to the thermocycler only when enzyme activation temperature (95°C) had been reached, to enhance specificity.

In order to reduce contamination, reaction reagents were combined using clean micro-pipettes and DNase and RNase free barrier tips (10µl and 100µl capacity; Promega).

Autoclaved, thin-walled PCR reaction tubes (200µl or 500µl capacity; Starsted) were used for all reactions.

2.4.3 Single Specific Primer PCR (SSP-PCR)

SSP-PCR was used to isolated gene fragments from lambda clones that could easily be cloned and then sequenced. The entire genomic DNA of the recombinant phage clone was digested and the fragments ligated into the multiple cloning site (MCS) of a plasmid. A gene specific primer was coupled with a standard sequencing primer that annealed at the boundary of the vector MCS. Amplification then allowed the isolation of the DNA neighbouring the location of the gene specific primer.

Preparation of template for the reaction was achieved via the single digestion of pZERO™ (Invitrogen) vector DNA and lambda clone DNA with identical enzymes. The DNA from both digestions was then purified and a ligation of lambda DNA (750ng) to vector DNA (50ng) was allowed at 16^{0}C for 4–16h in the presence of T4 DNA ligase (section 2.6.5). After the ligation period the ligase was inactivated by heating at 65^{0}C for 18min. The ligation reaction was diluted 1:10 with 10mM TrisHCl pH8.0 and applied as a template for PCR.

2.4.4 PCR based DNA sequencing

DNA sequencing was carried out on DNA cloned into a commercial plasmid vector. Sequencing reactions were out-sourced and performed by BaseClear (Leiden, The Netherlands) on either ABI377 or LiCor automated sequencers. Purified DNA template (and primer if gene specific) were sent to the Company. Raw sequence data was supplied in the form of a trace and text file for manipulation and analysis as described below.

2.4.5 Reverse Trancriptase PCR (RT-PCR)

Standard RT-PCR was performed using the Access™ RT-PCR system (Promega). This kit makes use of an Avian Leukaemia Virus (ALV) reverse transcriptase and *Tfl* DNA polymerase. The use of *Tfl* polymerase allows high fidelity amplification afforded by the enzymes 3'-5' exonuclease proof reading capacity.

A one-tube format reaction was performed using the following reaction components and conditions.

To a 0.5ml thin-walled PCR, on ice tube was added;

Nuclease free water	30µl
AMV/*Tfl* 5X reaction buffer	10µl
dNTP Mix (10mM each)	1µl
Reverse Primer (60pmol per µl)	1µl
Forward primer(60pmol per µl)	1µl
25mM $MgSO_4$	2µl
AMV reverse transcriptase (5U per µl)	1µl
Tfl DNA polymerase (5U per µl)	1µl
mRNA sample	3µl

The following cycling parameters were imposed;

1]	1 cycle	48^0C	45min (RT Step)
2]	1.cycle	94^0C	2min
3]	40 cycles	94^0C	30sec
		60^0C	1min
		68^0C	30sec-3min
4]	1 cycle	68^0C	7min

The success of cDNA amplification was assessed by applying 5µl of the RT-PCR reaction to agarose gel electrophoresis.

2.4.6 Nested 5' SMART™ RACE RT-PCR

In order to assess the transcriptional start site of a gene, nested 5' SMART™ RACE RT-PCR was performed. This procedure is facilitated by the annealing of a 5' adaptor to the 3' end of the first cDNA strand (SMART; Small Molecules Attaching to the 3' end of RNA Trancripts). Reverse transcription for first strand cDNA synthesis

was achieved using SuperscriptII™ reverse transcriptase. This is an MMLV derived reverse transcriptase that exhibits the non-template directed addition of 3 to 5 cytosine nucleotides to the 3' end of the completed first cDNA strand. This polyC motif allows the annealing of the 3' GGGGG sequence of the SMART II™ oligo. Primers directed against the SMART II™ oligo, coupled with the reverse gene specific oligo, allowed 5'RACE. A diagram of the procedure is presented in Appendix I.

The following components and reaction conditions were employed.

First strand synthesis

To a 0.5ml thin-walled PCR tube, on ice, was added;

mRNA sample	3µl
Reverse gene specific primer	1µl
SMART II™ oligo	1µl

Oligo/Primer annealing was permitted at 70°C for 2min.

Following primer annealing the following components were added to the reaction;

5X Reverse transcriptase buffer	2µl
DTT	1µl
dNTP Mix (10mM of each)	1µl
Superscript II™ MMLV reverse transcriptase	1µl

This reaction was placed in a thermocycler block at 42°C and incubated for 1.5h. To terminate the reaction the material was diluted in 250ml of Tricine-EDTA buffer and the block temperature was raised to 72°C for 7min.

Primary PCR

PCR was performed using the Advantage 2 PCR kit (Clontech). The following reagents and conditions were employed.

To a 0.5ml thin-walled PCR tube, on ice, was added;

5' RACE ready cDNA	2.5µl
UPM*	5µl
Gene specific primer (60pmol per µl)	1µl
PCR grade water	34.5µl

10 Advantage 2 PCR Buffer	5µl
dNTP mix (10mM of each)	1µl
Advantage 2 Polymerase Mix	1µl

*UPM; Universal Primer directed against the SMART II™ oligo

The following cycling parameters were imposed;

1]	5 cycles	94°C	5sec
		72°C	3min
2]	5 cycles	94°C	5sec
		70°C	10sec
		72°C	3min
3]	20 cycles	94°C	5sec
		68°C	10sec
		72°C	3min

Secondary PCR

A secondary PCR reaction was performed using a nested gene specific primer, a nested UPM and 1µl of a 1:50 dilution of the primary PCR as a template. The cycling parameters used in the primary reaction were repeated, except the 5 cycles of section 1 were omitted.

The success of both primary and secondary PCRs was assessed by agarose gel electrophoresis.

2.5 Universal GenomeWalker™ SSP-PCR

Specialised SSP-PCR was performed using Universal GenomeWalker™ technology (Clontech.). A diagrammatic overview of the procedure can be found in Appendix II.

2.5.1 Preparation of DNA for library construction

High quality *P.brasiliensis* chromosomal DNA was subjected to five restriction digests with blunt cutting restriction enzymes (*Dra*I, *Eco*RV, *Pvu*II, *Stu*I & *Sca*I).

The following were combined in a 1.5ml microfuge tube;

Genomic DNA (100ng μl^{-1})	25µl
10X restriction enzyme buffer	10µl
Sterile deionised distilled water	57µl
Restriction enzyme (10U μl^{-1})	8µl

The reactions were allowed to proceed for 2h at 37°C. Each reaction was then subjected to brief slow vortexing and restored to 37°C for a further 18h. Following incubation, the efficiency of each endonuclease was assessed by the application of 5µl of each reaction to agarose gel electrophoresis.

The remaining 95µl of DNA was purified and concentrated by ethanol precipitation. The resultant pellet was washed in 70% ethanol and allowed to air dry prior to dissolution in 20µl of 10mM TrisHCl pH8.0. The efficiency of ethanol precipitation was again assessed by agarose gel electrophoresis.

2.5.2 Ligation of GenomeWalker™ adapters to the blunt termini of restriction fragments

To a 0.5ml thin-walled PCR tube was added;

Restriction enzyme digested DNA	4µl
GenomeWalker adaptor	1.9µl
5X T4 ligase buffer	1.6µl
T4 DNA ligase	0.5µl

In total, five ligation reactions were performed, one for each restriction digest. The tubes were incubated at 16°C in a thermocycler block for 16h. The ligation reactions were terminated by elevating the block temperature to 70°C for five minutes followed by the addition of 72µl of 10mM TrisHCl pH8.0. Prior to application in PCR, the contents of each tube were mixed thoroughly.

2.5.3 PCR genome walking using the GenomeWalker™ libraries

Genome walking was facilitated by SSP-PCR using the adapter ligated genomic libraries as template. In order to achieve high specificity a two-step nested PCR

strategy was imposed. Sequentially, two pairs of primers were employed, each pair composed of an adapter directed primer and a gene specific primer. The gene specific primer and adapter directed primer in the primary reactions (GSP1 and AP1 respectively) were designed to allow the application of the second pair of primers (GSP2 and AP2) which would target a sequence internal to the primary PCR product. This allows increased specificity and the monitoring of the nested PCR performance by the visualisation of a decrease in the size of a secondary product in comparison to the primary.

As an additional measure to increase PCR specificity, a touchdown PCR protocol was performed. Essentially, this procedure imposes high stringency on the initial cycles of the PCR using high annealing temperatures that are gradually decreased. Once the target annealing temperature is reached it is held for the remaining cycles of the PCR.

Primary PCR

To a thin-walled 0.5ml PCR tube was added;

GenomeWalker™ library DNA	1µl
GSP1 (100pM μl^{-1})	1µl
AP1 (10µM)	1µl
10X PCR reaction Buffer	5µl
Sterile distilled deionized water	40.5µl
dNTPs (10mM each)	1µl
HotStarTaq™ polymerase	0.5µl

Five tubes were set up in total, one for each library.

The reactions were performed in an Eppendorf gradient thermocycler using the following cycling parameters;

1]	1 cycle	95°C	15min
2]	10 cycles	94°C	30s
		65°C (-0.5°C per cycle)	40s
		72°C	4min
3]	25 cycles	94°C	30s
		60°C	30s

		72^0C	4min
4]	1 cycle	94^0C	30s
		60^0C	30s
		72^0C	10min

For each primary reaction, 5µl of the product was analysed by agarose gel electrophoresis.

Secondary PCR

The primary PCR was diluted 1:50 in 10mM TrisHCl pH8.0 before serving as the template for the secondary reaction. PCR conditions were the same as those used in the primary PCR except primers GSP1 and AP1 were replaced by GSP2 and AP2 respectively. Also, the number of cycles in stage 3 of the PCR was reduced from 25 to 15. For each secondary reaction, 5µl of the product was analysed by agarose gel electrophoresis.

2.6 Enzyme mediated manipulations of DNA

2.6.1 Restriction enzyme digestion

Digestion of DNA with restriction endonucleases was performed in 1.5ml microfuge tube at 37^0C in a water bath. Reaction mixes contained the appropriate enzyme buffer, diluted to 1X, 3-6 units of restriction endonuclease and an appropriate amount (50ng to 10µg) of DNA. Sterile distilled water was added to a final volume of 20µl.

Digestion was allowed to proceed for 20min–4h depending on template concentration and intended application of the DNA. Routinely, the preparation of plasmid DNA for sticky-end cloning was achieved in 20min, the digestion of plasmid DNA for analysis of cloning was performed in 1h while 4h was used for the digestion of larger concentrations of genomic DNA. Reactions were stopped at the endpoint either by immediate electrophoresis, by inactivation of the enzyme by heat (if the enzyme was heat labile) or by a previously described DNA purification technique.

2.6.2 Fluorescein labelling of probes for hybridisation

The production of DNA probes for hybridisation was achieved using the Amersham Gene Images™ kit. An appropriate amount of probe DNA (50-200ng), in a volume of 34µl was denatured by boiling for 5 minutes, followed immediately by incubation on ice.

To a 1.5ml microfuge tube, containing the DNA, was added;

Denatured DNA (50-200ng)	34µl
Flourescein labelled dNTPs	10µl
Random hexamer primers	5µl
Klenow DNA polymerase (Amersham)	5U (1µl)

The reaction was allowed to proceed at 37^{0}C, in a waterbath, for 1hour. The reaction was stopped at the end point by the addition of EDTA to a final concentration of 20mM.

Prior to using a probe for hybridisation, the success of the random prime labelling reaction was assessed. To a strip of Hybond N+ (Amersham) nylon membrane was placed 5µl of the probe labelling reaction and 5µl of fluorescent nucleotides at separated marked positions. The solutions were allowed to air dry for 60s and placed in a hybridisation tube containing 20ml of 2xSSC (20X SSC is 3M NaCl and 0.3M Na citrate pH7.0) at 60^{0}C. The membrane was washed under these conditions, in a hybridisation oven that revolved the tube, for 15min. The success of probe labelling could be visualised immediately by placing the membrane, face down on a UV transilluminator. Typically all unincorporated nucleotides would be washed off and a fluorescent signal would remain, which arose from the fluorescein labelled dNTPs incorporated in the DNA probe.

2.6.3 A-tailing of proof read RT-PCR amplicons to allow A-T cloning

Following gel purification, an appropriate amount (c. 200ng) of a PCR amplicon, generated by a proof-reading polymerase was exposed to non-proof-reading Taq DNA polymerase in the presence of 100µM dATP, 1.5mM $MgCl_2$ and 1X PCR reaction buffer. The mixture was incubated at 72^{0}C for 15 minutes. The DNA was

then re-purified by electrophoresis and gel extraction prior to inclusion in an A-T cloning reaction.

2.6.4 Dephosphorylation of vector ends

The sticky ends of single restriction enzyme digested plasmid DNA were dephosphorylated to reduce the frequency of self-self ligation during a cloning procedure. A 20µl single enzyme restriction digest of 1µg pZERO™ vector DNA was allowed to proceed for 15mins. To this reaction, 5 units of calf intestinal alkaline phosphatase (CIAP) was added and the reaction was further incubated for 5min at 37°C.

2.6.5 Ligation of DNA to plasmid vectors

Vector systems used for cloning were linear, or made linear, to facilitate a T4 ligase or topoisomerase dependant insertion of the DNA of interest into the cloning site of the plasmid. Two types of system were used. For the cloning of PCR products (3' adenylated), A-T cloning was used. For cloning of restriction fragments of DNA, sticky end ligations were performed by cutting the vector DNA to generate ends compatible with those of the DNA to be cloned.

A-T cloning

For general propagation and sequencing of PCR amplicons the pGEM-T Easy™ vector system was used (Promega). Reaction components were as follows;

To a 1.5ml microfuge tube was added;

pGEM-T Easy™ Vector DNA	1µl (50ng)
2x Rapid ligation Buffer	5 µl
Gel Purified PCR amplicon	3 µl (To give a 2:1 insert to vector ratio)
T4 DNA ligase	1 µl (5 Units)

Ligation was allowed to proceed for 1hr at room temperature or overnight at 4°C prior to transformation.

For the construction of recombinant protein expression systems in *Saccharomyces cerevisiae*, the pYES2.1-V5-His Topo™ vector system (Invitrogen) was used. Reaction components were as follows;

To a 1.5ml microfuge tube was added;

Topoisomerase activated pYES2.1-V5-His Topo™ vector	1µl
Gel purified PCR product	4µl

The ligation was allowed to proceed for 5min at room temperature. The ligation was stopped at the endpoint on addition of 1µl of stop buffer (proprietary composition). The reaction was used to tranform *S.cerevisiae*.

Cohesive end cloning

For the cloning of DNA from restriction digests of recombinant lambda DNA, or the generation of template DNA for SSP-PCR, the pZERO™ vector system (Invitrogen) was used. Ligation was achieved using the Life Technologies T4 DNA ligase enzyme and buffer system.

Reaction components were as follows;

Digested pZERO™ vector DNA	50ng
Digested DNA to be cloned	200ng
5x Ligase Buffer	4µl
T4 DNA ligase	1µl (5 Units)
Sterile Distilled Water	to 20 µl

Ligation was allowed to proceed for 1-4hr at 16^{0}C (in a PCR thermocycler) prior to transformation.

For cloning of DNA fragments for recombinant expression in *E.coli* strains, digested pET21b (Novagen) was used. Reaction components were as follows;

Digested pET21b DNA	50ng
Digested DNA to be cloned	30-200ng (to a 2:1 insert to vector ratio)
5x Ligase Buffer	4µl
T4 DNA ligase	1µl (5 Units)
Sterile Distilled Water	to 20 µl

Ligation was allowed to proceed for 1hr at room temperature prior to transformation.

2.7 Transformation of *E.coli* and *S.cerevisiae* with plasmid DNA

2.7.1 Transformation of chemically competent *E.coli*

Chemically competent *E.coli* strains acquired from various commercial sources. Transformation protocols varied, as per manufacturer's instructions. A typical protocol is described below;

To a 1.5ml microfuge tube on ice was added 50µl of thawed competent cells. The cells were then mixed with 1.7µl of 100mM β–Mercaptoethanol and incubated on ice for 10min. Following this 50ng-500ng of plasmid DNA or ligation reaction was added to the cells and gently mixed. This mixture was then incubated on ice for 30mins. Heat shock was applied to the transformation by rapidly inserting the tube in a 42°C water bath for 1min. Tubes were rapidly restored to an ice bath for two minutes. To each tube, 450µl of SOC medium was added and the suspension was shaken at 220rpm and 37^{0}C for 1 hour. Cells were then plated onto LB agar containing the appropriate antibiotic.

2.7.2 Transformation of Bacteria by Electroporation

When a higher number of recombinants was required, electroporation of competent cells was used. To an ice-chilled 2mm electroporation cuvette (Equibio) was added 50µl of electrocompetent cells and 50-500ng of DNA. The electroporation cuvette was inserted in a pre-chilled carriage into a BioRad Electroporator and exposed to 200V/75mA of current. Cells were immediately re-suspended in 950µl of SOC media and shaken at 37^{0}C for 1hr. Cells were then plated onto LB agar containing the appropriate antibiotic.

2.7.3 Transformation of *S.cerevisiae* with plasmid DNA

A single colony of the *S.cerevisiae* strain was used to inoculate 10ml of YPD in a 50ml Erlenmeyer flask. The culture was shaken at 200rpm and 30^{0}C overnight. The

following morning the culture $O.D._{600}$ was assessed and a dilution was made so as to bring the $O.D._{600}$ to 0.4 in 50ml of YPD. Outgrowth was permitted at 30°C and 200rpm for 3h. Cells were pelleted at 2,500rpm and 4°C for 10min. The pellet was re-suspended in 40ml of TE buffer and the centrifugation repeated. The pellet was re-suspended in 2ml of Yeast Transformation Buffer (1X Lithium Acetate/ 40% PEG-3350/ 1X TE) and incubated at room temperature for 30min.

For each transformation the following components were mixed in a 1.5ml microfuge tube;

Plasmid DNA	1µg
Denatured Salmon Sperm DNA	100µg
Yeast Suspension	100µl
Yeast Transformation Buffer	700µl

This mixture was incubated at 30°C for 30min. On the addition of 88µl of DMSO (dimethylsulphoxide) the cells were heat shocked at 42°C for 7 min. The cells were harvested at 8000rpm for 1min, re-suspended in 1ml of TE buffer and 50-100µl were spread on a SC-U^- selective plate. Plates were incubated at 30°C in an inverted position for 72h to allow transformant colonies to grow.

2.7.4 Selective screening of recombinant bacterial clones

Blue/White screening of recombinant pGEM-T Easy™ clones

Differentiation between successful and unsuccessful insert cloning in pGEM-T Easy™ was allowed by the insertional inactivation of the β–Galactosidase gene and resultant perturbation of hydrolysis of X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside), a non lactose substrate analogue. Expression from the β–Galactosidase gene is under the control of the *lac* promoter and as such successful transcription requires the presence of the non-utilisable lactose analogue IPTG (isopropyl-β-D-thiogalactopyranoside). To each plate of LB agar (100µg per ml ampicillin) was spread 100µl of sterile 100mM IPTG (Melford laboratories). The β–galactosidase activity was detected by the inclusion of X-gal on the plates. Prior to inoculation 100µl of 2%w/v X-gal, in dimethyl formamide, was spread onto the agar surface. The surface moisture arising from IPTG and X-gal application was allowed to dry on placing the plates at 37°C for a minimum of 30min prior to use.

Plates were incubated, in an inverted position, at 37°C for 16-18h, and white colonies picked and propagated for further analysis.

Analysis of recombinant pZERO™ constructs by insertional inactivation of the *ccd* B gene.

Discrimination between recombinant and self-ligated plasmid constructs in the pZERO™ vector was also achieved by insertional activation of a plasmid borne gene. In this case however the re-ligation of linearised plasmid to the original form (i.e. without insert) reconstituted the *ccdB* gene. This gene, which was also under the control of the *lac* promoter, is lethal to cells that express the encoded protein. The protein acts on the bacterial DNA gyrase and its presence results in perturbation of ATP-dependant negative supercoiling. As such, all clones that do not possess a plasmid containing recombinant insert are rendered inviable. In order to activate the expression of the *ccd* gene, IPTG was added to the molten LB agar (with 50µg per ml kanamycin to a final concentration of 1mM. Following inoculation, the plates were incubated for 16-18hr at 37°C, and evident colonies were picked and propagated for further analysis.

2.8 Southern transfer and hybridisation of DNA

2.8.1 Southern transfer of DNA from agarose gels

DNA samples for Southern blotting were run on an agarose gel as described above (section 2.3.1). To assist in analysis two lanes of molecular weight markers were run, flanking the samples on both sides and scale was applied to the gel photograph by placing fluorescent rulers either side of the gel during photography.

The gel was then transferred to denaturing solution (1.5M NaCl/0.5M NaOH) on a shaking platform for 45 minutes. The gel was then washed in dH_2O and neutralised by incubation in neutralising buffer (1.5M NaCl/0.5M TrisHCl pH8.0) for 30 min. This neutralising buffer was discarded and the gel was placed in fresh buffer for a further 15min.

The transfer of DNA to the Hybond N+ membrane was achieved by the method of Southern (Southern E.M., 1970). Essentially, a platform was made in a reservoir of 20x SSC (as in section 2.6.2). A wick made from Whatman 3M filter paper covered the platform and was draped from the platform into the SSC. Onto the platform was placed the treated agarose gel in an inverted position (well face down). A section of Hybond N+, cut 1mm smaller than the gel at each edge, was placed on to of the gel avoiding the trapping of air. The nylon membrane was overlaid with 2 sheets of Whatman 3M paper that had been soaked in 2X SSC for 10min. Approximately 10cm of absorbent paper towels were placed on top of the assembly to accumulate the transferred SSC. A 500g weight was applied to the top of the stack of towels. Transfer was allowed overnight.

The efficiency of transfer was assessed by re-staining the gel in 40mls of TAE buffer containing ethidium bromide (0.2µg per ml).

2.8.2 Transfer of DNA to nitro-cellulose membranes from λ plaque lawns

The transfer of lambda plaque DNA to nitro-cellulose was performed after the method of Sambrook, Fritsch & Maniatis (1989). Plaque lawns of recombinant λ phage were grown overnight on 10 cm x 10cm square petri dishes. The plates were placed at 4^0C for 2h to allow hardening. DNA transfer was achieved by placing a section of Hybond N+ nitro-cellulose membrane onto the agar surface for 2min to allow phage particles to adsorb. The membranes were cut at several places to provide orientation marks. These were noted on the plate to aid later analysis. The membranes were then treated for 5min in denaturant (1.5M NaCl/0.5M NaOH), followed by 4 min in neutralising solution (1.5M NaCl/0.5M TrisHCl pH8.0). Following a brief wash in 2xSSC the membranes were wrapped in saran wrap. The membranes were then exposed to 12,000 j per cm^2 of ultraviolet light in a cross-linker to secure the DNA to the membrane. Membranes were used directly for hybridisation or stored at 4^0C until required.

2.8.3 DNA:DNA hybridisation

The hybridisation procedure was performed in a Techne hybridisation oven. Generally, the protocols for the hybridisation of membranes derived from Southern

blots or plaque lifts were identical. The membranes were pre-hybridised for 2h at 58-65^{0}C in pre-hybridisation buffer (5xSSC/1%SDS/5% dextran sulphate/1:10 dilution of blocking reagent (Amersham)). Tubes were rotated at c.80rpm to allow mixing. After this period, denatured probe (5-40ng per ml of hybridisation fluid) was added to the tube. Hybridisation was allowed to proceed for 12-16hrs at 55-68^{0}C.

2.8.4 Washing and detection of bound probe

Unbound probe was removed from the membrane by sequential washings at 60^{0}C in decreasing concentrations of SSC in the presence of SDS. Three 10 min washes were imposed; 2xSSC/0.1%SDS, 1xSSC/0.1%SDS, and finally 0.5xSSC/0.1xSDS. The membrane was then blocked for 1h at room temperature using a 1:10 dilution of liquid block in buffer A (100mM TrisHCl, 300mM NaCl pH9.5).

The membrane was washed briefly in buffer A and probe detection commenced. In order to detect bound probe the membrane was exposed to mouse anti-flourescien HRP conjugate. The antibody was applied to the membrane at a 1:5000 dilution in 20ml of bufferA plus 0.5% w/v bovine serum albumin (fraction V), per hybridisation tube. This was incubated at 25^{0}C for 1.5h. Unbound antibody was removed by 3x10min washes in buffer A with 0.3% v/v Tween 20.

To induce signal from the antibody, 2µl of CDP-Star detection reagent (proprietary composition) was applied per cm^2 of membrane (generally 200µl was used). The blot was soaked in the detection reagent for 2min and excess reagent was removed. The blots were then transferred to a plastic envelope within an autoradiography cassette. A sheet of Hyperfilm^{TM} MP (Amersham) was placed into the cassette under safelight illumination so as to be exposed to the signal generating face of the membrane. Exposure was allowed for 15min - 4h before developing the film in an automatic developer.

2.8.5 Analysis of results derived from agarose gels

The image observed on the blot was marked to allow the distance of positive signals to be measured from the originating wells. The relevant DNA on the original gel was

then identified by reference to the scale on the original photograph, as shown by the fluorescent rulers.

2.8.6 Analysis of results derived from λ plaque lifts

The orientation marks on the blot images were marked. These images were then aligned with the markings on the original phage plates, on a white light box to identify positive plaques. Plaques giving a positive signal were marked, on the underside of the plate, for further investigation.

2.9 Heterologous expression of recombinant proteins

2.9.1 Expression of recombinant proteins from pET21b in *Epicurian coli* ® *BL21* CODON PLUS™ (DE3)-RIL

The heterologous production of recombinant proteins derived from *P.brasiliensis* was achieved in *Epicurian coli* ® *BL21* CODON PLUS™ (DE3)-RIL cells. Bacterial strains harbouring expression constructs were streaked onto a LB agar (100μg per ml carbenicillin, 34 μg per ml chloramphenicol) plate to allow the isolation of a single colony. Following overnight incubation and growth, a single colony was used to inoculate a starter culture. The starter culture was grown in a volume of media 1/10 of that used in the final expression culture. Outgrowth was allowed at 37^0C with shaking (200rpm) until the $O.D_{600}$ reached 0.5. The starter culture was then added as a 1/10 volume to the media intended for expression (LB plus 100μg per ml carbenicillin/100μg per ml chloramphenicol) and growth continued under similar incubation conditions until an approximate $O.D_{600}$ of 0.6. was reached.

At this point expression was induced by the addition of IPTG to a final concentration of 1mM. In order to induce expression, the de-repression of two promoters is achieved by the inactivation of the *lac* repressor (*lac*I). This is an allosteric effect occurring upon interaction of this protein regulator with IPTG. The de-repression of the chromosomal *lacUV* 5 promoter permits the production of T7 RNA polymerase that instigates transcription from the T7*lac* promoter, allowing transcription of

recombinant mRNA. The T7*lac* promoter is also under the control of the *lac* repressor, allowing a further level of control of repression prior to induction.

For time-course experiments samples were aseptically removed at 1h intervals to allow monitoring of protein production by SDS-PAGE.

2.9.2 Preparation of Total Cell Protein (TCP) samples for analysis of recombinant expression

In order to assess the extent of heterologous protein expression, a culture of the recombinant strain was grown in the appropriate medium and induced as described above. Prior to induction and at 1h intervals thereafter a 1ml aliquot was removed from the expression culture. The cells were harvested by centrifugation at 13,000rpm for 1 minute. The cell pellet was prepared for analysis by SDS-PAGE as described in section 2.9.6.

2.9.3 Expression of recombinant proteins from pYES2 in *Saccharomyces cerevisiae*

The heterologous production of recombinant proteins of *P.brasiliensis* origin in *S.cerevisiae* was achieved in the following manner. Yeast strains harbouring expression constructs were streaked onto ScU$^-$ Minimal agar (2% glucose) plates to allow the isolation of a single colony. Following 72h incubation and growth a single colony was used to inoculate a starter culture. The media for this starter culture was a volume of ScU$^-$ Minimal broth (2% glucose), 10% of that of the final intended volume for expression. Outgrowth was allowed at 30^{0}C with shaking (200rpm) until overnight. The OD_{600} of the overnight culture was assessed and an appropriate volume of the starter culture was designated for the inoculation of the induction medium; 100ml ScU$^-$ Minimal broth (2% galactose/1%Raffinose). This volume of starter culture was calculated as the amount of cells that were required to bring the OD_{600} of the expression culture to 0.5. The required volume of the starter culture was pelleted at 3000rpm for 5min and washed once in the galactose/raffinose induction medium. The re-suspended pellet was transferred to the pre-warmed induction medium and incubated at 30^{0}C for 24h.

In order to assess protein production over time, samples were taken at 4h intervals during the course of expression.

2.9.4 Electrophoresis of protein under denaturing conditions by SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Composition and preparation of gels

Analysis of protein derived from recombinant systems was achieved by separation of those proteins by electrophoresis. Electrophoresis was carried out on vertical polyacrylamide gel slabs according to the procedure first described by Laemmli (Laemmli, 1970). Each gel system consisted of a separating gel and a stacking gel.

For separating gels the percentage acrylamide ranged from 10% to 12 % depending on the resolution required at certain molecular weights. Typically a 12% gel was used and was composed and polymerised as follows:

1] The following reagents were mixed in a 10ml beaker;

Distilled water	3.5ml
1.5M TrisHCl pH8.8	5.22ml
10% SDS	100µl
30% w/v acrylamide/bis solution (Biorad)	4.0ml

2] This solution was then degassed for 15 min in a vacuum chamber.

3] To initiate polymerisation, the following were added and gently mixed;

TEMED*	5µl
10% APS*[2]	50µl (0.5mM final concentration)
Final volume	10ml

The gel mix was carefully loaded into a BIO-RAD mini-gel apparatus and overlaid with water saturated butanol. Polymerisation was allowed to proceed for 45min at room temperature.

For stacking gels the percentage acrylamide was 7%. This gel was composed and polymerised as follows:

1] The following reagents were mixed in a 10ml beaker;

Distilled water	6ml
0.5M TrisHCl pH6.8	2.5ml
10% SDS	100µl
30% w/v acrylamide/bis solution	1.3ml

2] This solution was then degassed for 15 min in a vacuum chamber.

3] To initiate polymerisation, the following were added and gently mixed;

TEMED*	10µl
10% APS*[2]	50µl (0.5mM final concentration)
Final volume	10ml

(*TEMED; N,N,N'N'-tetramethylethylenediamine ($C_6H_{16}N_2$). *[2]APS; Ammonium Persulphate ($NH_4S_2O_8$), prepared fresh daily.)

Following removal of the overlaying butanol the stacking gel was carefully applied on top of the separating gel. A gel comb (10 or 16 wells) was inserted into the stacking gel and polymerisation was allowed to proceed for 45min at room temperature. When polymerisation was complete, the comb was removed and the wells were gently flushed with running buffer (see section 2.9.8).

2.9.5 Assembly of electrophoresis apparatus

The gels were inserted and clamped into a central manifold and placed inside an electrophoresis tank. The central chamber was then filled with 1X SDS-PAGE running buffer (per litre; 3g of Tris base, 14.4g of glycine, 1g of SDS, often made as a 5X stock) until a depth of approximately 7mm was present above the height of the gel. The outer chamber was also filled with running buffer at a sufficient level to allow the electrode to be completely submerged.

2.9.6 Preparation of samples for SDS-PAGE

Samples for SDS-PAGE were prepared in various ways depending on the origin of the material. Bacterial pellets to be analysed directly were re-suspended in an appropriate volume 10X SDS-PAGE loading buffer (per 8ml; 4ml of distilled water,

1ml of 0.5M TrisHCl, pH8.6, 0.8ml of glycerol, 1.6ml of 10% SDS, 0.4ml of 2 β-mercaptoethanol and 0.05% w/v bromophenol blue). The samples were then boiled for 5min and placed on ice until required.

Protein preparation from yeast cells was achieved using osmotic lysis coupled by mechanical shearing. To a yeast cell pellet was added a volume of PBS (usually 100µl per 1ml of original culture). An equal volume of 10X loading buffer was added and the sample was vortexed for 30s. The sample was then placed in boiling water for 5 min prior to application on a gel.

Samples not derived directly from microorgansism followed a similar pattern except that no vortexing was necessary. An appropriate volume of sample for analysis was mixed with an equal volume of 10X loading buffer and boiled for 5min.

2.9.7 Electrophoresis protocol

To the each sample well of the gel was added 5-20µl of protein sample using a 100µl capacity Hamilton syringe. Each gel included a lane of LMW molecular weight calibration markers (Pharmacia biotech). Current was applied at 200V and the electrophoresis was allowed to run until the required separation was considered to be achieved (usually 40-50min).

2.9.8 Analysis of SDS-PAGE gels

Analysis of polyacrylamide gels was achieved using Gelcode® Blue Protein Stain (Pierce). Initially the gel was removed from the glass plates and placed into a 10cm x 10cm square petri dish filled with distilled water. The gel was washed in the water for 15min with gentle agitation to remove SDS (because SDS inhibits the staining procedure). The water was discarded and the 15min wash repeated. The water was again discarded and 20ml of Gelcode® Blue Protein Stain was applied. The staining was allowed to proceed for 1hr and then the stain was removed and replaced with water to allow destaining. Following destaining protein bands could be visualised with clarity.

2.10 Analysis of recombinant expression by western blotting

Western immuno-blotting was used to confirm the presence of a His_6 tag on recombinant proteins, or simply to demonstrate the presence and identity of those proteins.

2.10.1 Western transfer of proteins onto nitro-cellulose membrane

Electrophoresis of gels intended for western transfer was achieved as previously described for SDS-PAGE gel electrophoresis. Gels were run in duplicate, one for transfer and one for direct analysis by staining. The transfer gel was removed following electrophoresis and immersed in transfer buffer (per litre; 3.03g of Tris base, 14.5g of glycine and 200ml of 100% methanol, made up to 1L with distilled water) for 40min prior to assembly of the transfer apparatus. For each gel, a section of Hybond C (Amersham) nitro-cellulose membrane was cut to the exact size of the separating gel. Two pieces of Whatman 3M filter paper were also cut to this size. Both the membrane and the filter papers were soaked in the transfer buffer for 20min prior to the assembly of the transfer apparatus.

The gel transfer cassette was assembled as follows. Firstly an abrasive nylon fibre pad, soaked in transfer buffer was placed on the cathodic face of the cassette. Onto this pad was placed one of the soaked 3M filter papers. The stacking gel was trimmed from the separating gel and a corner of the separating gel was trimmed removed to allow orientation. The separating gel was then placed on top of the filter paper. A replicate corner of the Hybond C paper was cut and the paper was laid over the gel, avoiding any trapping of air. Finally, the other 3M filter paper was applied and a second abrasive soaked nylon pad. The cassette was then closed and placed in an electrophoresis tank. The cassette was inserted so that the protein samples could migrate from the gel onto the nylon membrane, in the direction of the anode. An ice block was placed in the tank and the chamber was filled with 4^0C transfer buffer . A magnetic stirrer was also added to provided circulation and cooling of buffer during the procedure.

A current of 100V was applied for 1h to allow transfer. Following disassembly of the cassette, the membrane was reversibly stained in 20ml of 0.1% w/v Ponceau S in

0.5% v/v acetic acid until the transferred proteins could be visualised. The membrane was then destained by 2 x 10min washes with 40ml PBS.

2.10.2 Processing of the western blot

The membrane was removed, placed in a 10 x 10cm dish and blocked for 1hr with 50ml 3% w/v skimmed milk in wash buffer (1.21g of Tris base, 8.7g of NaCl and 800µl of concentrated HCl made up to 1L [pH7.4]). Following blocking the membrane was transferred to a heat sealed bag containing 30mls of 3% skimmed milk and a 1:300 dilution of mouse anti-His_6 monoclonal antibody (Pharmacia). The antibody was allowed to bind on incubation at 37^{0}C for 3 hours with agitation. The blot was then washed 4 x 5min with 40ml of wash buffer and gentle agitation. Secondary antibody binding was performed in 30ml of 3% w/v skimmed milk in wash buffer containing a 1:1000 dilution of rabbit anti-mouse HRP conjugated antibody. Again incubation was for 1h at 37^{0}C with gentle agitation. Prior to detection the blot was washed 4 times with wash buffer as described above. Proteins with bound antibody were detected by immersing the gel in a developing solution (30mg of 4-chloro-1-napthol dissolved in 10mls methanol was added to 50mls TrisNaCl followed by 30µl of 30% w/v hydrogen peroxide). The colour signal was allowed to develop in the dark and the reaction halted by washing in distilled water once the desired intensity was reached. The blot was then air dried and stored in the dark between sheets of Whatman 3M filter paper to discourage fading.

2.11 Purification of His_6-tagged proteins

2.11.1 Small scale purification of His_6-tagged proteins under denaturing conditions

Purification of high concentrations of protein derived from inclusion bodies was performed under denaturing conditions. On a small scale this was done using the Ni-NTA Spin kit (Qiagen).

The spin columns were equilibrated by the application of 600µl of equilibration buffer (8M urea / 0.1MNaH_2PO_4 / 0.01M TrisHCL pH8.0). The buffer was passed through the column by centrifugation at 2000rpm for 2min. The column was then loaded with 600µl of protein solution re-solubilised in the equilibration buffer. This was passed through the column by centrifugation at 2,000rpm for 2min.

Unbound proteins were washed from the column by 3 applications of 600µl wash buffer (as equilibration buffer but pH6.3) and eluted by two applications of 200µl of elution buffer (as equilibration buffer but pH4.5). Again centrifugation at 2,000rpm for two minutes was used to pass buffers through the column. The success of purification steps was analysed by SDS-PAGE.

2.11.2 Larger scale purification of His_6-tagged proteins under denaturing conditions

Larger preparations of inclusion body derived proteins were performed using Ni-NTA agarose (Qiagen). The principles used were the same as those imposed on the Ni-NTA spin columns.

Preparation of the insoluble protein fraction

Following the required period of induction the cultures were removed from the incubator and placed on ice for 5min. The cells were then harvested by centrifugation at 5000rpm for 10min at 4^0C. The supernate was discarded and the pellets placed on ice. The cells were re-suspended in an appropriate volume of re-suspension buffer (50mM NaPi, 300mM NaCl, pH8.0) to allow for a 1:200 concentration factor in reference to the original culture volume. The cell suspension was then treated with lysozyme (1mg ml^{-1} final concentration) and DNAse (5µg ml^{-1} final concentration) and incubated on ice for 1h.

The cells were then disrupted using a OneShot™ Cell disrupter (Constant Systems Ltd). Cells were forced, under high pressure, through a small aperture in the base of a cell disrupter cup. This treatment is efficient at breaking cells. A volume of 5ml could be processed at any one time. The disrupted material was recovered from the cell disrupter cup and the insoluble protein and/or cell debris was harvested by centrifugation at 10,000xg for 30min. The pellet was re-suspended in 30ml of PBS containing 0.1% triton X100 and re-centrifuged. This latter procedure served to remove any miscellaneous cell debris from the inclusion body material.

Nickel affinity chromatography

The insoluble pellet was re-suspended in 20ml of Denaturing Re-suspension Buffer (8M Urea, 0.1M NaH_2PO_4, 0.01M TrisHCl pH8.0) by shaking overnight at room

temperature. To 10ml of this solubilised pellet was added 2ml of 50% Ni-NTA agarose (Qiagen). This mixture was loaded directly onto a glass chromatography column (Biorad Econocolumn). The column was allowed to drain at a rate of 1ml per minute. The column was washed 4 times in 8ml of Denaturing Wash Buffer (8M Urea, 0.1M NaH_2PO_4, 0.01M TrisHCl pH6.3) and eluted 4 times in 0.5ml Denaturing Elution Buffer (8m Urea, 0.1M NaH_2PO_4, 0.01M TrisHCl pH4.5).

The success of each step of purification was assessed by SDS-PAGE.

2.11.3 Purification of His_6-tagged proteins under native conditions

Native folded soluble recombinant protein was purified from *E.coli* cultures using the NTA Ni-agarose.

Preparation of soluble protein

The preparation of the soluble fraction was as described for the preparation of the insoluble fraction. However after centrifugation at 10,000 x g for 30min, the supernate was retained for purification rather then the insoluble pellet. It should also be noted that the composition of the re-suspension buffer was 50mM NaPi*, 300mM NaCl, 15% glycerol, 15mM imidazole, pH8.0, containing EDTA free protease inhibitors (Complete™ EDTA free protease inhibitor tablets (Roche Molecular Diagnostics). Care was taken to perform all re-suspension and disruption steps on ice as much as possible. The supernate from the 10,000 x g spin was decanted into clean 9ml ultracentrifuge tubes (Sorvall) and exposed to ultracentrifugation (Sorvall OTD combi ultracentrifuge) at 43,000rpm for 1.5h to remove the cell membranes. The supernate from this spin was applied to native purification.

* NaPi is Sodium Phosphate buffer, made by combining volumes from 0.5M Na_2HPO_4 and 0.5M NaH_2PO_4 stocks. To make 1L of 50mM NaPi, 93.2mls of 0.5M Na_2HPO_4 and 6.8ml of 0.5M NaH_2PO_4 were added to 900ml of distilled water.

Nickel affinity chromatography

To the isolated soluble fraction was added a volume of 50% Ni-NTA agarose (1ml of Ni-NTA agarose was applied per 8mg of anticipated recombinant protein as this was the advised optimal binding capacity of the nickel resin). The mixture was then incubated, with gentle shaking (90rpm), at 4^0C for 1h.

To a glass chromatography column was added the soluble fraction/Ni-agarose preparation. The column was allowed to drain at a rate of 1ml per min until only the agarose was retained. The column was washed with an appropriate column of wash buffer (50mM NaPi, 300mM NaCl, 15% glycerol, 40mM imidazole, pH8.0, containing EDTA free protease inhibitors). Generally 5 washing steps were applied, using 4ml of wash buffer per ml of 50% resin at each stage. Elution of bound protein was achieved using high concentrations of imidazole in elution buffer (50mM NaPi, 300mM NaCl, 15% glycerol, 300mM imidazole, pH8.0, containing EDTA free protease inhibitors). A total of 6 elution steps were applied each imposing 500µl of elution buffer. The success of each step of purification was assessed by SDS-PAGE

2.11.4 Dialysis of purified protein

Purified protein was dialysed prior to downstream procedures generating fluorescence data. This allowed the removal of intrinsically fluorescent imidazole prior to these experiments. Dialysis tubing (12-14kDa pore size) was cut into 8cm sections and boiled for 10 min in 2% w/v $NaHCO_3$/1mM EDTA. The tubing was washed extensively in distilled water and boiled for a second time in 1mM EDTA. The extensive water washing was repeated and the tubing was submerged in 1mM EDTA and stored at 4^0C until required

The purified protein solution was transferred pre-prepared dialysis tubing (12-14 kDa pore size) and secured using dialysis clips. Dialysis of the protein was performed in dialysis buffer (50mM NaPi, 300mM NaCl, 15% glycerol, pH8.0, containing EDTA free protease inhibitors). A ratio of 1:1000 was imposed on the dialysate to dialysis buffer and the dialysis system was incubated with gentle stirring at 4^0C for 48h. The protein was then recovered from the dialysis tubing, aliquoted into 1.5ml microfuge tubes and snap frozen in a dry ice/ethanol bath.

2.12 Ligand binding analysis

2.12.1 Stopped-flow analysis of nucleotide/protein interactions

The transient process of nucleotide interactions with a protein nucleotide-binding domain was measured in a stopped flow device (Applied Photophysics). This

apparatus allows the measurement of transient changes in the fluorescence output following the rapid mixing of protein and ligand solutions. By plotting values derived from alterations in this emission over a selected time-base, a graphical representation of progressive increases or decreases in fluorescence can be obtained. Depending on the filter used on the output light, the fluorescence of either tryptophan (protein) or ligand can be analysed and conformational changes demonstrated. An increase in the fluorescence of a given tryptophan relates to a conformational change that results in the internalisation of the residue to an environment shielded from quenching water molecules. A decrease in fluorescence signal relates to an externalisation and an increased exposure to the quenching phenomenon.

Interactions between protein and ligand occured in a mixing chamber, equilibrated to 20^0C. Delivery of both protein and ligand solutions to this chamber was achieved by pushing equal volumes of each solution (approx.100µl) from external reservoirs using nitrogen-driven hydraulic plungers. Tryptophan excitation light, at a wavelength of 280nm, was constantly applied to the mixing chamber. This light originates from an argon lamp and the wavelength is selected on passage through two monochromators.

Output or emission light was filtered to select for the appropriate wavelength and collected in a photomultiplier tube (PMT). Alterations in the level of fluorescence resulted in a proportional alteration in the voltage output of the PM tube. It was these PM volts that were plotted to relate the fluorescence behavior of the contents of the mixing chamber.

In our experiments we examined the interactions of a protein nucleotide-binding domain with the fluorescent nucleotide analogue MANT-ATP (2'-(or 3')-O-(N-methylanthraniloyl) adenosine 5' triphosphate) (Molecular Probes Inc.). The changes in tryptophan fluorescence were measured on a 50millisecond logarithmic time base collecting emission light at 320nm. Changes in the fluorescence of the MANT-ATP ligand were measured in the same way using a 420nm filter. The data was processed on an Acorn II microcomputer using Applied Photophysics software.

For each experiment, solutions of both protein and MANT-ATP were prepared using the buffer used for the dialysis of the test protein (50mM NaH_2PO_4, 300mM NaCl, 15% glycerol).

2.12.2 Steady-state analysis of protein/drug interactions

The analysis of interactions between a protein nucleotide-binding domain and the azole drug ketoconazole was measured, on the basis of fluorescence, in a Jascow Series II fluorimeter. Excitation light was applied at a wavelength of 285nm via a series of chromating mirrors. This input light serves to excite fluorophores in a quartz cuvette and the fluorescent emission is measured at right angles to the input light.

For the purposes of our experiments a total volume of 200µl was used in a Hellma suprasil quartz cuvette. This volume contained a constant concentration of protein and variable concentrations of ligand. Again, the buffer system used was the same as that used for protein dialysis (section 2.11.4). Ligands applied were ketoconazole (made as a stock solution of 10mg ml^{-1} in DMSO), and ATP (at a stock concentration of 100mM). For the study of azole binding, the drug was applied at concentrations of zero to 100µM. The effect of ATP on this system was examined using a constant concentration of 10mM ATP. All observations were made at 20^0C.

Measurements were made by exciting the system at 285nm and scanning output from 290nm to 500nm, allowing for fluorescence peaks derived from both protein and ketoconazole to be measured. Inner filter effects imposed by ATP or ketoconazole were quantified by positioning a 2ml Hellma suprasil quartz cuvette in front of the test cuvette containing protein or protein + ligand. This 2ml cuvette was filled with buffer containing 10mM ATP or an appropriate test concentration of the azole drug. The resultant decrease in the fluorescence output from the test cuvette was recorded and experimental observations were corrected accordingly. All experimental observations were made with the 2ml cuvette placed in front of the test cuvette to eliminate variations in the transmission of excitation light.

Chapter 3

Isolation and sequencing of *pft1*, an MDR-like ABC half transporter from *Paracoccidiodes brasiliensis*

The objective of the experiments reported in this chapter was to screen the genome of *P.brasiliensis* for genes encoding ABC transporters belonging to the MDR subfamily. This was to be achieved by the application of degenerate PCR primers, a technique which had proved useful for isolating fragments of such genes by other investigators (Bozdech *et al.*, 1996; Tobin *et al.*, 1997). The full length genes corresponding to any relevant generated fragments would be isolated, cloned and sequenced.

The primers used (denoted as MDR1 and MDR2 in table 2.2) had previously allowed the isolation of gene fragments from *Aspergillus fumigatus* and *Aspergillus flavus* that appeared to encode a segment of nucleotide-binding domains of MDR-like proteins. The corresponding full length genes, *afumdr1, afumdr2* and *aflmdr1* were then cloned and completely sequenced (Tobin, *et al.,* 1997).

This chapter reports that a single gene fragment was obtained from *P.brasiliensis* DNA, using this degenerate PCR. This gene fragment appeared to correspond to a segment of an MDR-like gene. This amplicon was used to screen a λ gene library. A single phage clone was isolated which contained the major part of an open reading frame. The complete sequence was attained using SSP-PCR techniques (Venkatakrishna & Ames, 1989) on *P.brasiliensis* genomic DNA. The complete gene sequence revealed a 2627bp open-reading frame, including two putative introns and encoding for an 827 amino acid half-transporter molecule with extensive homology with *afumdr2*.
2.

3.1 PCR amplification of an MDR-like gene fragment using degenerate PCR

In an attempt to identify MDR-like ABC transporter genes in *P.brasiliensis* the primers successfully used in the cases of *A.fumigatus* and *A.flavus* (Tobin, *et al.,* 1997) were adapted. The codon usage in the degeneracies was altered based on the codon usage of *Neurospora crassa* and *Aspergillus nidulans* (Gurr *et al.*, 1987). These oligonucleotides were designed to target the motifs ALVGPSG (MDR1 forward primer) and

GGQKQRIA (MDR2 reverse primer). Both of these motifs are conserved in the nucleotide binding domains of proteins that share sequence homology with human P-glycoprotein (Gottesman & Pastan, 1993).

Primers MDR1 and MDR2 were applied in PCR, over a range of annealing temperatures, as seen in Fig 3.1a. A single band was consistently observed. The amplicon from the reaction performed at 56^0C was extracted from the agarose gel and ligated into pGEM-T Easy™. The ligation reaction was transformed into competent XL1-Blue *E.coli* cells. Recombinant plasmids were selected on the basis of blue/white screening, with 47 of 68 colonies displaying the required white phenotype. The plasmids harboured by five white clones were prepared by mini-prep and analysed for insert by *Eco*RI restriction digestion. Four were found to contain DNA of the appropriate size (Fig 3.1b). The recombinant inserts of pMDR1.1, pMDR1.2 and pMDR1.3 were analysed by automated sequencing using vector directed standard primers (M13 forward and M13 reverse). All three clones possessed an identical 413bp sequence of *P.brasilensis* DNA that showed considerable homology with a number of ABC transporter genes including *afumdr2, mdl1* and *mdl2* (Fig 3.2). This sequence was projected to belong to a putative open reading frame that we termed *pft1* (*pft1*; *Paracoccidiodes* fungicide transporter).

3.2 Screening of lambda libraries for the *pft1* gene

In order to isolate the complete *pft1* gene, λ phage gene libraries would be screened. Three libraries were in our possession (gifts from Prof. Maria Sueli Soures Felipe, Departmento de Biologia Celular, IB, Universidade de Brasilia, Brasilia, Brazil); a λDASHII genomic DNA library and two λZAP cDNA libraries. The cDNA libraries were constructed using the mRNA of either the yeast form or the mycelial form of the organism (all λ library technologies had been constructed using kits from Stratagene). Prior to the execution of the screening procedure the libraries were titred to assess the viable phage population. Serial dilutions of phage stocks ($1x10^{-4}$ to $1x10^{-7}$) were performed and pfu assessed on the lowest dilution that could be accurately counted. The λDASHII genomic library displayed the highest viable phage count with the stock library calculated to be at $4.5x10^8$ pfu per ml. The yeast cDNA and mycelial cDNA libraries titred at $3.2x10^7$ and $2.6x10^7$ pfu per ml respectively.

In order to establish the feasibility of screening the lambda phage libraries for the complete *pft1* gene, the total DNA of each of the libraries was examined by PCR.

Fig3.1a. **Degenerate PCR screening of *P.brasiliensis* DNA for MDR-like ABC transporter gene fragments**

1] 55^{0}C annealing temperature
2] 57^{0}C annealing temperature
3] 59^{0}C annealing temperature
4] 61^{0}C annealing temperature
5] 63^{0}C annealing temperature
6] 65^{0}C annealing temperature

PCR conditions; 100µl reaction volume with 1.5mM $MgCl_2$. Activation at 95^{0}C for 15min / denaturation at 94^{0}C for 30s / annealing at various temperatures for 40s / extension at 72^{0}C for 1min. A total of 25 cycles concluded with a 10min 72^{0}C extension.

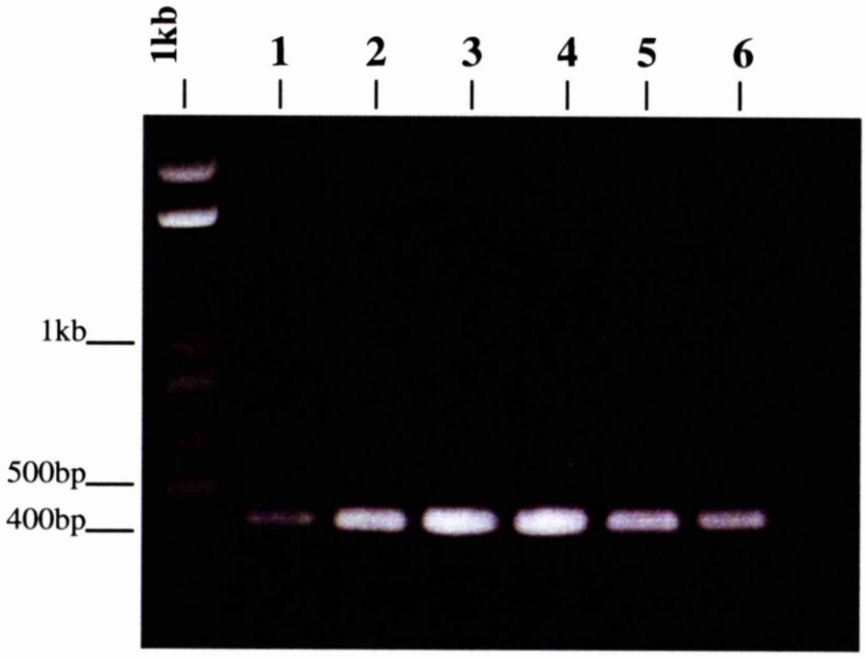


Fig3.1a

Fig3.1b p-GEM-T Easy™ clones containing the MDR1/MDR2 fragment

DNA (500ng) from each of 5 plasmid preparations (pMDR1.1 - pMDR1.5) was digested with 5 Units of *Eco*RI at 37^{0}C for 1h. The digested DNA was examined on a 0.9% agarose gel. The correct clones were confirmed on observation of the c.400bp MDR1/MDR2 fragment.

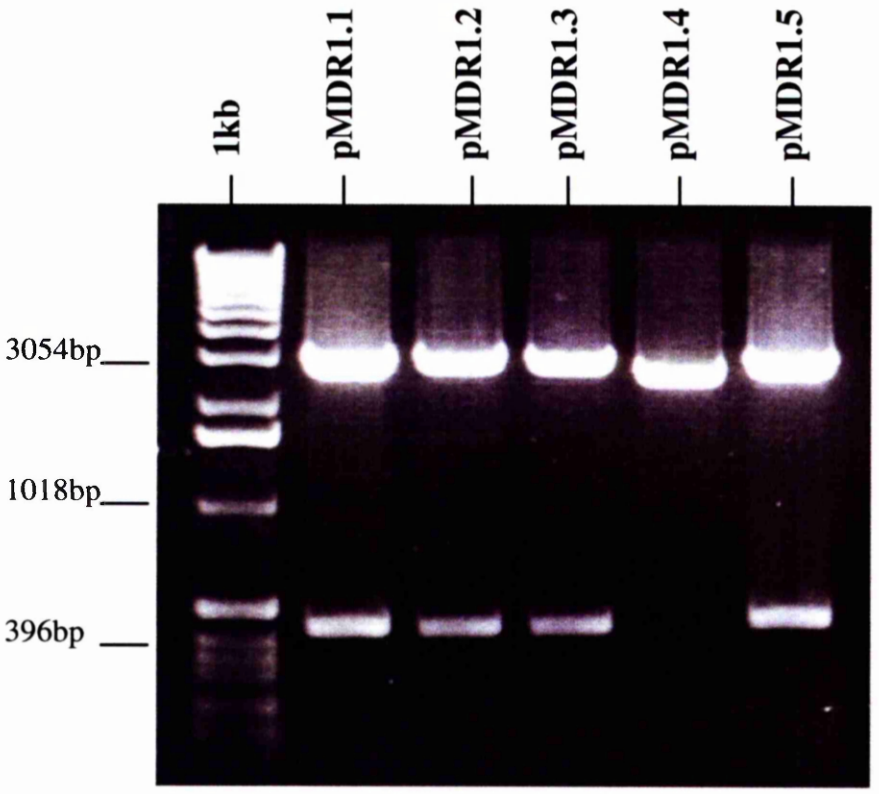

1kb
pMDR1.1
pMDR1.2
pMDR1.3
pMDR1.4
pMDR1.5
3054bp
1018bp
396bp

Fig3.1b

Accession Number	Protein ; Source	Author
1] gi|2673955 (U62935)	AfuMDR2; *Aspergillus fumigatus*	(Tobin *et al.*, 1997)
2] emb|CAB42370.1	Putative permease; *Schizosaccharomyces pombe*	(See note 1)
3] gb|AAB52482.1|	Putative ABC transporter; *C.elegans*	(See note 2)
4] gi|4204793	P-glycoprotein PMDR1; *Solanum tuberosum*	(Wang *et al..*, 1996)
5] emb|CAA94202.1|	P-glycoprotein homologue; *C.elegans*	(See note 2)
6] emb|CAB04797|	YgaD; *Bacillus subtilis*	(Cummings *et al*, 1997)
7] emb|CAA94219.2|	P-glycoprotein homologue; *C.elegans*	(See note 2)
8] emb|CAA94220.2|	P-glycoprotein homologue; *C.elegans*	(See note 2)
9] p|Q00449|MDR4	Mdr4; *Drosophila melonogaster*	(Wu *et al..*, 1991)
10] gb|AAF23176.1|AF216497	P-glycoprotein homologue; *Gossypium hirsutum*	(Cui,X., Unpub.)
11] gb|AAF17668.1|AC009398	P-glycoprotein homologue; *Arabidopsis thaliana*	(Ecker,J.,Unpub.)
12] gi|1166606	P-glycoprotein homologue; *C.elegans*	(Wilson, 1994)
13] pir||A42150	P-glycoprotein Atpgp1; *Arabidopsis thaliana*	(Dudler & Hertig, 1992)
14] pir||S21957	P-glycoprotein pgp1 - *Arabidopsis thaliana*	(Dudler,R., Unpub.)
15] emb|CAA94203.1|	P-glycoprotein homologue; *C.elegans*	(See note 2)
16] gb|AAD28285.1|AF136523	Bile salt export pump BSEP; *Homo sapiens*	(Mol, O., Unpub.)
17] ref|NP_003733.1||	Bile salt export pump BSEP; *Homo sapiens*	(Stautnieks *et al.* 1998)
18] ref|NP_015053.1|MDL2|	MDL2; *Saccharomyces cerevisiae*	(Goffeau *et al.*, 1996)
19] gi|311095 (L16959)	MDL2; *Saccharomyces cerevisiae*	(Dean *et al.*, 1994)
20] gb|AAD22645.1|AC007138	Putative P-glycoprotein; *Arabidopsis thaliana*	(Huang *et al.*, Unpub.)

Fig 3.2

This figure is adapted from a BLAST output generated on searching the NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/) using the MDR1/MDR2 fragment sequence. The 20 genes with highest homology scores are presented, as generated using the BLASTx algorithm. BLASTx analysis allows similarity searching to be performed on the basis of deduced amino acid sequences acquired by translation in all three reading frames, in forward and reverse directions. The description of identity for each hit has been altered from the original output to ease reader identification.

Note 1; Identified by the European *S.pombe* genome sequencing project, Sanger centre.

Note 2; Identified by the *C.elegans* genome sequencing consortium, Sanger centre.

The library DNA was prepared by infecting aliquots of *E.coli* LE392 cells with 10µl of stock phage, as described in section 2.2.7, and the libraries were amplified in liquid media. Following this high multiplicity infection, the DNA was prepared and dissolved to an approximate concentration of 150ng per µl for each library (data not shown). The MDR1 and MDR2 primers were applied in PCR to examine each of the three libraries for the presence of the *pft1* gene fragment. For each test PCR reaction, 300ng of total library DNA was used. As can be seen in Fig 3.3, the required fragment could only be observed in the λ DASHII genomic library. Despite repeating the procedure, evidence of the gene was not observed in the λ ZAP cDNA libraries constructed using the cDNA of either the yeast form or mycelial form of the organism. It was concluded that the *pft1* gene was absent in the cDNA libraries and the complete sequence would be pursued by screening the genomic library only.

The λ DASHII library was plated to allow well separated plaques. The library stock was diluted by a factor of 10^{-6} and 100µl of this dilution was used for infection. This allowed approximately 450 plaques to be screened per 10 x 10 cm plate. Assuming an average insert size of 20kb, each plate would present 9Mb of *P.brasiliensis* DNA. The typical size of the haploid *P.brasiliensis* genome is now known to be 31.3Mb (Montoya *et al.*, 1999) indicating that this genome will be present on a minimum of 4 plates of this titre, assuming that no identical clones appear.

The plaque material from each plate was transferred to nitro-cellulose membranes prior to hybridisation. At least two replicate membranes were prepared for each plate so as to allow the assessment of reproducibility of putative positive signals.

To isolate the required recombinant phage clone, the *pft1* gene fragment, derived from the MDR1/MDR2 PCR, was labelled using fluorescein tagged nucleotides. Successful fluorescein labelling was confirmed by visualisation of bound Klenow polymerase amplicons on nitro-cellulose by UV transillumination. This labelled DNA was applied as a probe in a plaque hybridisation procedure.

After extensive screening a single plaque, λ*pft1*, was found to give a positive signal on two replicate membranes (Fig 3.4a). The plaque was found to be well isolated on the corresponding plate, was picked and a 10^{-3} dilution of the harvested phage re-cultivated. A secondary hybridisation on these lawns of λ*pft1* demonstrated that the initial signal was genuine (Fig3.4b). All signals observed on this image corresponded to plaques on the cultured plate of the isolated phage clone.

Fig3.3 **MDR1/MDR2 PCR screening of phage library DNAs**

1] *P.brasiliensis* genomic DNA (positive control)

2] λDASHII genomic library DNA

3] λZAP cDNA library (yeast form)

4] λZAP cDNA library (mycelial form)

5] No DNA (negative control)

PCR conditions were as stated in Fig3.1a. The annealing temperature was 56^0C.

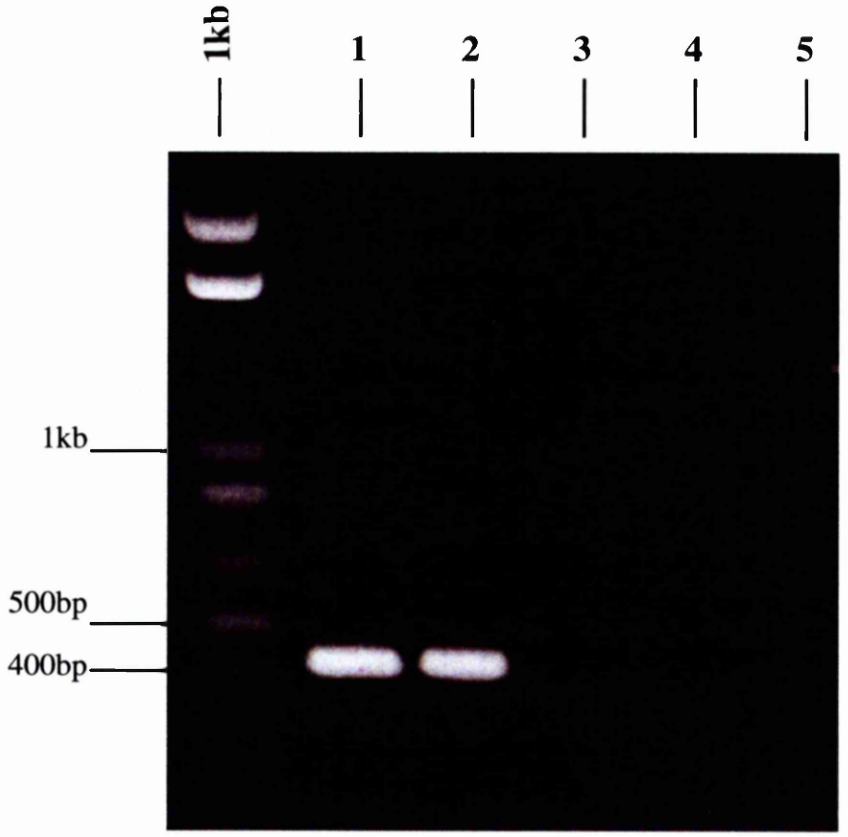


Fig.3.3

Fig3.4a **Primary hybridisation screen of λDASHII genomic library**

The left-hand image represents the signal obtained from the initial lift of plaque material from a phage library culture plate. The right-hand image is the signal obtained from a second plaque lift from the same plate. Thus replicate membranes were available to confirm the position of positive signals.

Hybridisation conditions; 140ng of labelled probe (at 7ng per ml of hybridisation buffer) was applied to replicate filters. The membranes were incubated in the presence of probe for 18h at 58^{0}C. A series of 3 stringency washes were applied using 2xSSC/0.1%SDS, 1xSSC/0.1%SDS and 0.5xSSC/0.1%SDS in succession. Again the solutions were at 58^{0}C. The membrane was blocked for 1h at room temperature. Following this, anti-fluorescien antibody was applied at 1:2000 dilution of stock as supplied. Unbound antibody was washed off, detection reagent applied and exposure to X-ray film was allowed for 2h.

Fig3.4b **Replicate filters of secondary screen on the λ*pft 1* clone**

Hybridisation conditions; As described for Fig3.4a, except hybridisation temperature and stringency washes were performed at 63^{0}C.

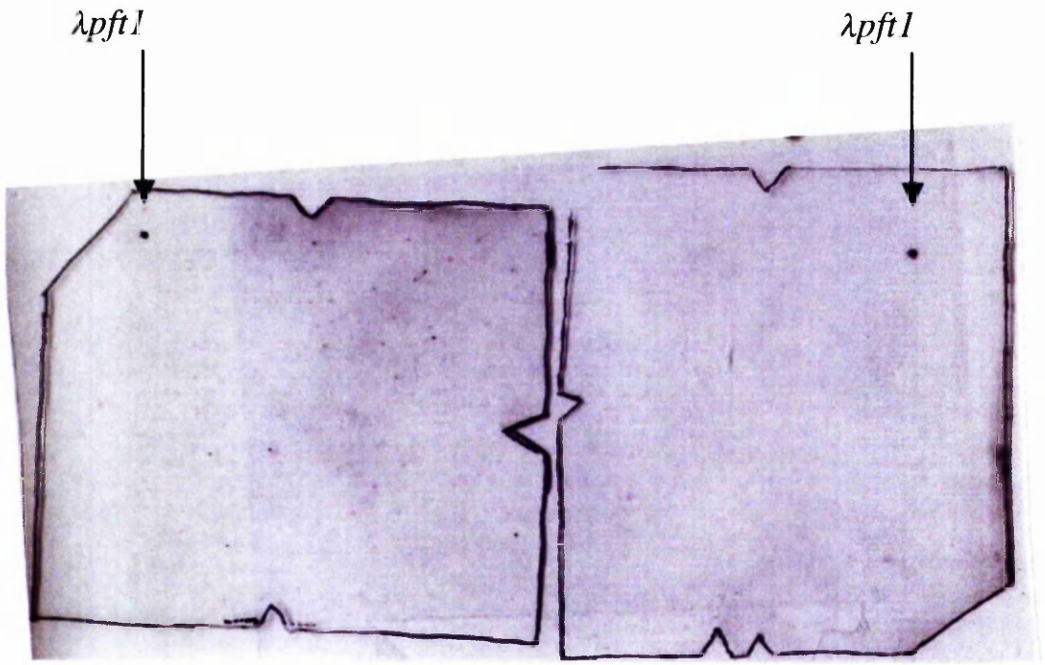


Fig3.4a

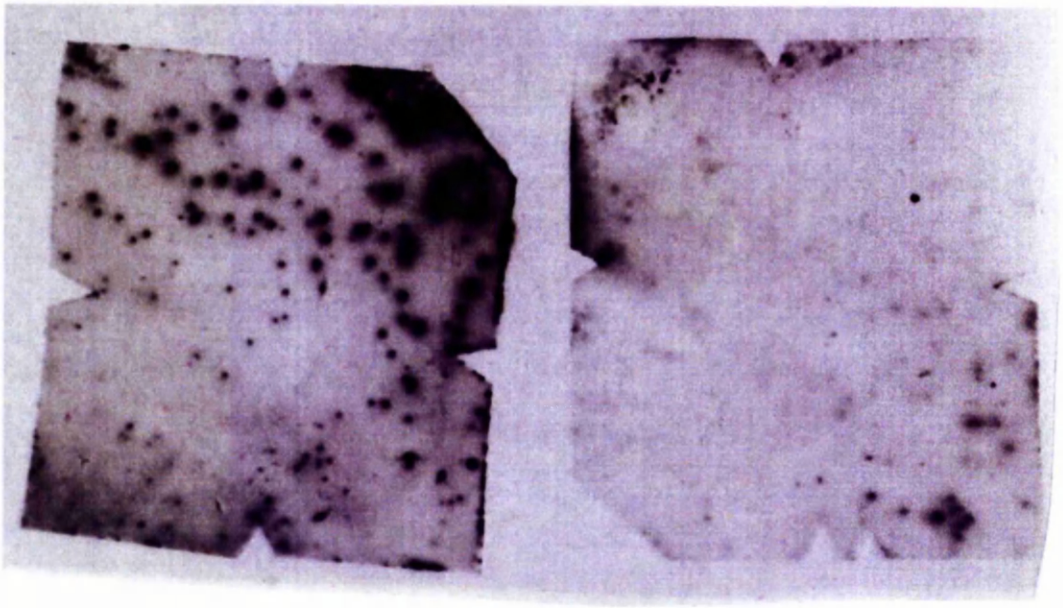


Fig3.4b

3.3 Subcloning *P.brasiliensis* DNA from λ*pft1*

In order to obtain fragments of the *pft1* gene from the lambda clone, a series of subcloning experiments was performed. Initially, a high multiplicity liquid culture of λ*pft1* was prepared and the phage DNA isolated by standard methods to provide 50μg of λ*pft1* DNA. The DNA was then digested, using restriction enzymes that would allow cloning into the multiple cloning site of the pZERO™ vector. Each restriction enzyme was used individually to provide 10 separate reactions, each containing 500ng of phage DNA. The digested DNA was electrophoresed on a 0.9% agarose gel (Fig3.5a). The DNA on this gel was then transferred to a nitro-cellulose membrane and exposed to Southern hybridisation, again using the fluorescein labelled MDR1/MDR2 fragment as a probe. The resultant Southern blot is presented in Fig3.5b. While a signal is present in all lanes, it can be observed that the *Xba*I, *Eco*RI and *Pst*I digests gave positive signals from bands of approximately 2.5, 3.9 and 4.8kb respectively. It was decided that these bands were the most suitable for subcloning.

Initially, 12 individual 30μl *Xba*I digests were performed and electrophoresed. From all digests, the appropriate 2.5kb fragment was gel extracted and the product isopropanol precipitated to concentrate. Of this DNA, 200ng was added to a ligation mix with 50ng of *Xba*I digested pZERO™ and the fragment was consequently cloned (Fig 3.6). Sequencing of this construct is discussed in more detail in section 3.5. The *Xba*I fragment did contain a substantial portion of the *pft1* gene, including a well-defined stop codon representing the 3' terminus of the gene.

The 5' section of the gene was not included in the *Xba*I fragment DNA so the *Eco*RI and eventually the *Pst*I gene fragments were targeted for further investigation. In order to circumvent the requirement for the laborious accumulation of digested fragment for cloning, the portions of each fragment predicted to contain the required 5' sequence were generated using SSP-PCR. *Eco*RI and *Pst*I restriction digests were performed using 2μg of λ*pft1* DNA per reaction. Following 4h incubation at 37^{0}C, the reactions were placed at 60^{0}C for 20min to inactivate the restriction endonucleases. Examination of an aliquot of each digest demonstrated satisfactory digestion. A volume of 7 μl (700ng) of digested DNA was applied to a ligation with 50ng of appropriately cut pZERO™. The total digested fragments of the lambda clone were allowed to ligate at 16^{0}C for 18h. This ligation mix was diluted 1:10, in 10mM TrisHCl pH8.0, and used as a template for PCR. Using the *pft1* specific primer PT72 and either the vector directed

Fig3.5a **Restriction digests of λ*pft1* DNA**

Digests were performed using 500ng of DNA per digest in a volume of 20µl. DNA was incubated at 37^{0}C for 4h in the presence of 5 units of the appropriate restriction endonuclease.

Lane 1]	*Xba*I
Lane 2]	*Hind*III
Lane 3]	*Eco*RI
Lane 4]	*Bam*HI
Lane 5]	*Not*I
Lane 6]	*Pst*I
Lane 7]	*Xho*I
Lane 8]	*Sac*I

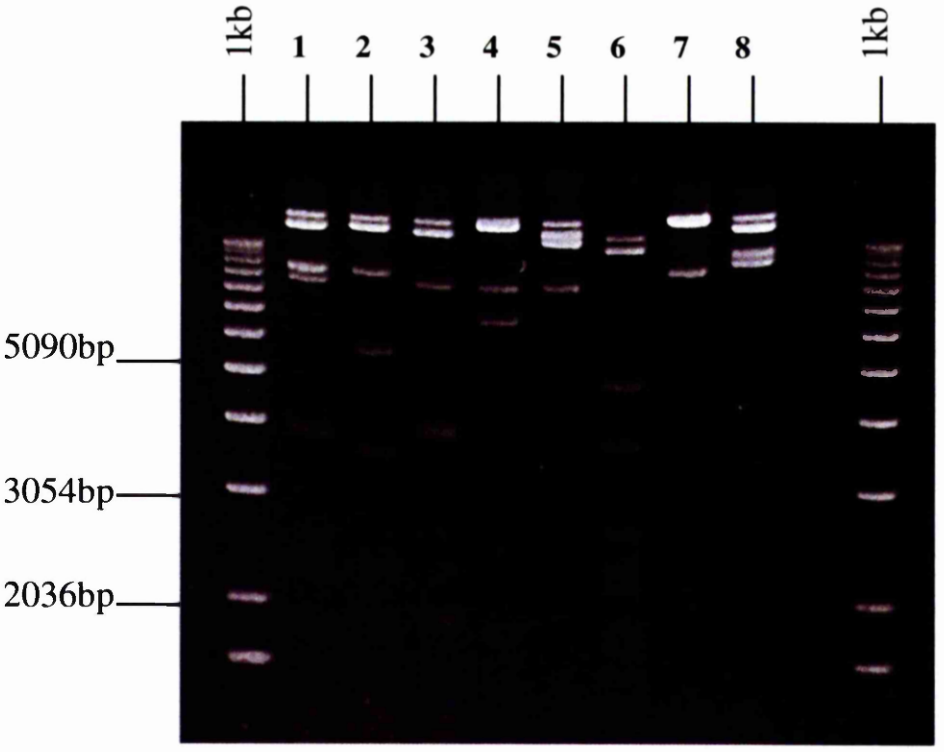


Fig3.5a

Fig3.5b Southern Blot of gel in Fig3.5a

Lanes are as Fig 3.5a

The blot was hybridised at 65^{0}C for 40h with 500ng of the *pft1* gene fragment as a probe (25ng per ml of hybridisation buffer). The blot was washed for 13min with each of 2X SSC/0.1% SDS, 1X SSC/0.1% SDS and 0.5X SSC/0.1% SDS at 60^{0}C. Anti-fluorescein was allowed to bind and the blot was exposed to X-ray film for 40min following the addition of detection reagent.

NB: This figure represents the blot on λ*pft1* DNA. The image was rendered using Adobe Photoshop™ 3.0.5. The original image generated on X-ray film could not be digitally scanned with clarity.

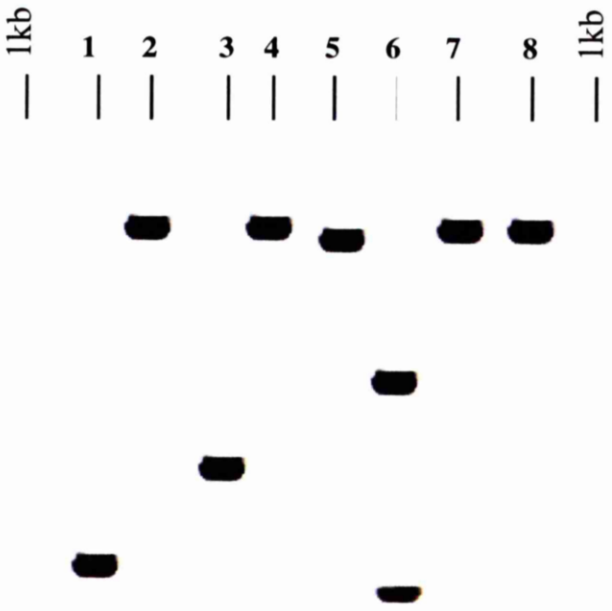


Fig3.5b

Fig3.6 Cloning of the λ*pft1* derived *Xba*I fragment in pZERO™

The 2.5kb *Xba*I fragment was cloned to the *Xba*I site of pZERO™. The ligation was performed using 50ng of linearised vector DNA and 250ng of the purified insert DNA. The ligation was incubated at 16^0C in a thermocycler for 4h prior to transformation into TOP10F' *E.coli* cells.

The insert was found to be cloned in 3 of the 6 putative clones examined. The construct pP1Xba5 was selected for automated sequencing.

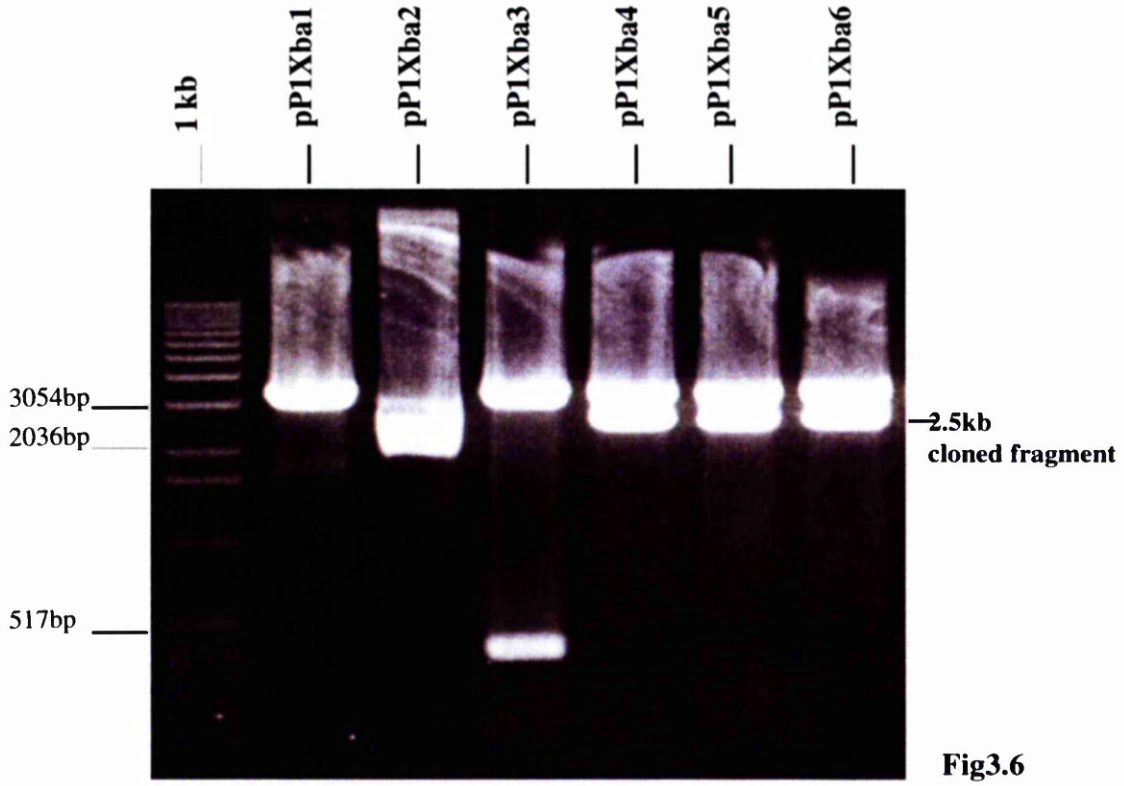

1 kb
pP1Xba1
pP1Xba2
pP1Xba3
pP1Xba4
pP1Xba5
pP1Xba6
3054bp
2036bp
517bp
2.5kb
cloned fragment

Fig3.6

T7 or SP6 primers the 5' region of *pft1* possessed by each fragment was amplified (Fig 3.7a/b). These amplicons were cloned into pGEM-T Easy™ and the relevant region of each insert sequenced (see section 3.5).

It was discovered that the *Eco*RI fragment provided further sequence data for the 5' end of *pft1* but again the complete ORF was not present. The larger *Pst*I fragment also contained additional information but revealed that the *pft1* gene sequence was truncated during library construction. Based upon a comparative alignment with the deduced amino-acid sequence of AfuMDR2, the phage clone λ*pft1* contained the bulk of the open reading frame but the gene appeared to have been severed during the partial *Sau*3A digest used to generate inserts for the library.

3.4 Genome walking to complete the *pft1* gene sequence.

Further screening of the library, using the 350bp sequence at the established 5' region to generate a probe, resulted in the re-isolation of the same λ*pft1* phage clone. The lambda genomic library had undergone several rounds of amplification prior to acquisition, which might result in individual clones being present in multiple numbers. This notion is supported by the high number of plates that had to be screened to find the initial *pft1* phage clone. Previous amplification would have selected for phage that replicated fastest, containing smaller inserts and would thus bias the representation of such clones in the library. In order to generate the region of the *pft1* gene that contained the ATG start codon, two PCR based techniques were employed. 5'RACE RT-PCR was attempted on *P.brasiliensis* mRNA. An amplicon pertaining to *pft1* was generated but did not provide a conventional position for the translational start of the gene. This result, and its potential significance, is discussed in more detail in Chapter 5. Also GenomeWalker™ SSP-PCR was performed using adapter-tag ligated, digested genomic DNA as a template. This would allow genome walking to amplify the sequence upstream of that that already obtained.

A series of 5 individual restriction digests, containing 2μg of *P.brasiliensis* genomic DNA, were performed using the blunt cutting enzymes *Pvu*II, *Eco*RV, *Dra*I, *Stu*I and *Sca*I. The digested DNA was precipitated and re-suspended to allow ligation of adapter tags to the blunt termini of each fragment. PCR was then performed on a 1:10 dilution of each ligation mix using the *pft1* specific primer and an adapter tag specific primer. A nested PCR approach was employed using primers pftwalk1 and Adapt1 for the primary PCR and primers pftwalk2 and Adapt2 for the secondary PCR (see Appendix II).

Fig3.7 SSP-PCR reactions allowing the isolation of the 5' *pft1* sequence contained within the genome of *pft1*.

a) SSP-PCR on ligated λ*pft1* *Eco*RI digested DNA.

Lane1] Gene specific primer SSP1/ Vector specific primer T7.

Lane2] As lane 1 but with no template (negative control).

Lane3] Gene specific primer SSP1/Vector specific primer SP6.

Lane4] As lane 3 but with no template (negative control).

PCR conditions: 100µl reaction volume with 1.5mM $MgCl_2$. Activation at 95°C for 15min / denaturation at 94°C for 30s / annealing at 55°C for 40s/ extension at 72°C for 2min. A total of 35 cycles concluded with a 10min extension at 72°C.

b) SSP-PCR on ligated λ*pft1* *Pst*1 digested DNA.

Lane1] Gene specific primer SSP1/ Vector specific primer T7.

Lane2] Gene specific primer SSP1/Vector specific primer SP6.

PCR conditions: 100µl reaction volume with 1.5mM $MgCl_2$. Activation at 95°C for 15min / denaturation at 94°C for 30s / annealing at 55°C for 40s/ extension at 72°C for 3.5min. A total of 35 cycles concluded with a 10min extension at 72°C.

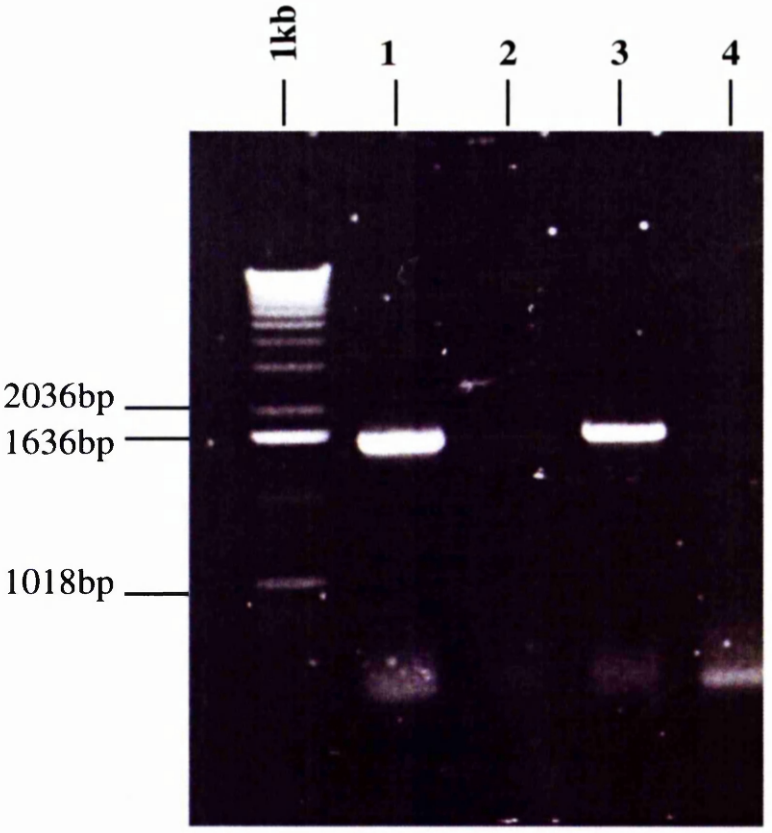


Fig3.7a

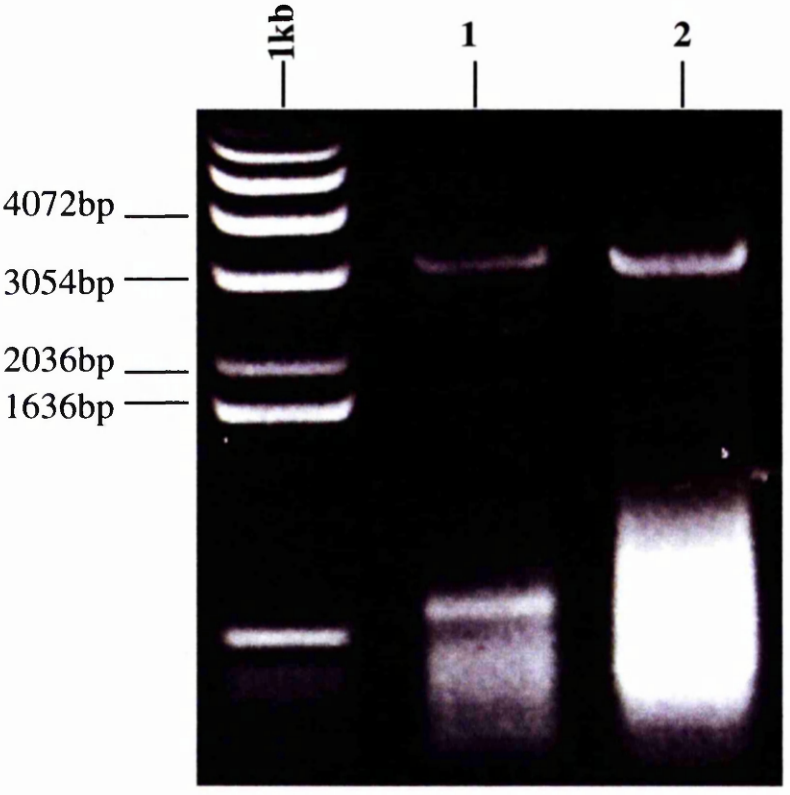


Fig3.7b

The results of both primary and secondary PCRs can be observed in Fig3.8a & Fig3.8b respectively.

A number of amplicons can be observed, of various sizes in the primary PCR. The secondary PCR indicated the prominence of 2 small products in the *Eco*RV and *Pvu*II lanes that displayed a suitable shift in molecular weight to have been generated from a true secondary PCR. Both the *Eco*RV fragment (511bp) and the *Pvu*II fragment (312bp) were cloned into pGEM-T Easy™ and sequenced. Both proved to be from *pft1* and provided information regarding the position of an ATG start codon. Both amplicons were of the same sequence in the region around the ATG start codon and the 5'end of the gene. The larger *Eco*RV fragment contained an additional 199bp that gave information about the sequence of the proposed immediate promoter.

3.5 The sequence of *pft1*

All sequence data received were analysed using the Autoassembler™ (Perkin Elmer) sequence analysis software. This allowed the assembly of the complete contiguous DNA sequence as provided from the data gathered from the *Xba*I, *Eco*RV and *Pst*I fragment clones and the GenomeWalker™ products. A diagrammatical representation is presented in Fig 3.9. Note that each section of the gene is sequenced at least in duplicate ensuring that sequencing is performed in opposing directions. The nomenclature for each sequence is based on its origin. Sequences named "f*Xba*..." are derived from the 2.5kb cloned *Xba*I fragment DNA. The "EVGW" and "PIGW" are from the *Eco*RV and the *Pvu*II genome walker products respectively. Sequences marked as "*Eco*RISSP" and "*Pst*ISSP" are from the *Eco*RI and *Pst*I SSP-PCR products of Fig3.7. The inclusion of sequences generated by 5'RACE will be discussed in Chapter 5.

The complete nucleotide sequence, including the deduced amino acid sequence, can be observed in Fig3.10 while the comparative alignment of amino acids with the closest matching homologue AfuMDR2 is presented in Fig3.11. The comparative alignment was performed using the GenejockeyII sequence analysis software (Copyright; P.L.Taylor, Biosoft).

In Fig3.10, please note the inclusion of 2 putative introns at nucleotide positions 1248-1317 and 1953-2031 (denoted in blue). These sections were proposed to be introns on

Fig3.8a **Primary GenomeWalker™ PCR**

PCR was performed using the following cycling parameters;

1 cycle]	94°C	15:00
35 cycles]	94°C	0:30
	63°C	0:40
	72°C	3:00
1 cycle]	94°C	0:30
	63°C	0:40
	72°C	10:00

Lane 1] *Pvu*II library

Lane 2] *Dra*I library

Lane 3] *Eco*RV library

Lane 4] *Stu*I library

Lane 5] *Sca*I library

Fig3.8b **Secondary GenomeWalker™ PCR**

PCR was performed using the following cycling parameters;

1 cycle]	94°C	15:00
25 cycles]	94°C	0:30
	63°C	0:40
	72°C	3:00
1 cycle]	94°C	0:30
	63°C	0.40
	72°C	3:00

(Lane order as in Fig3.8a.)

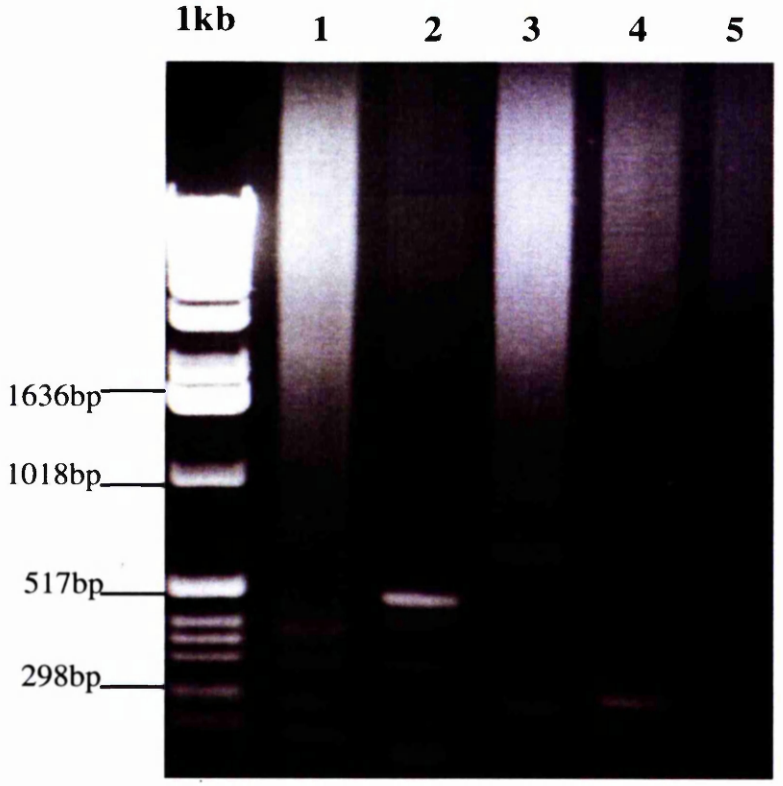


Fig3.8a

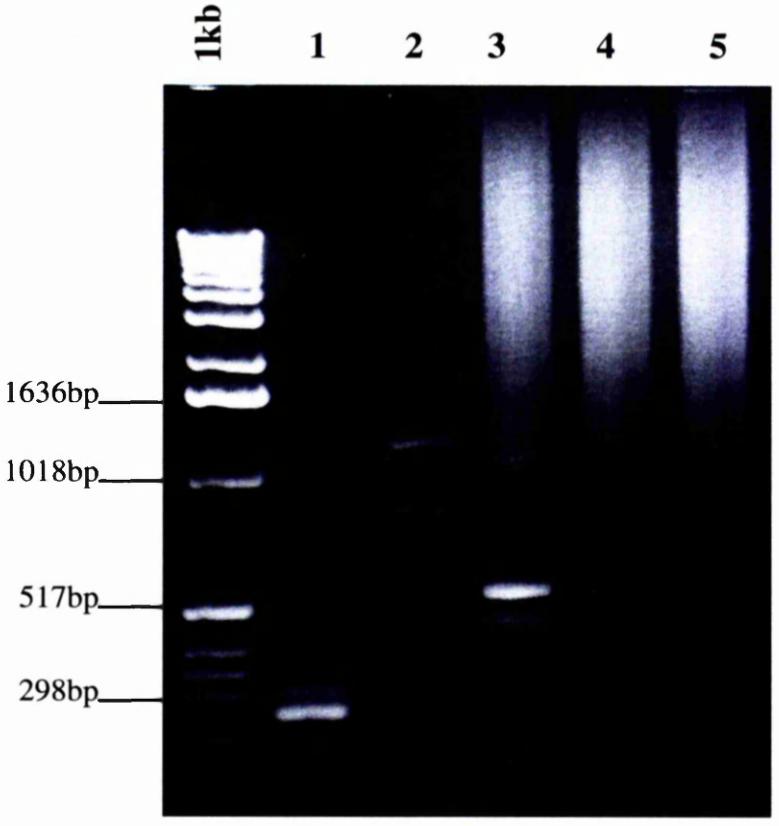


Fig3.8b

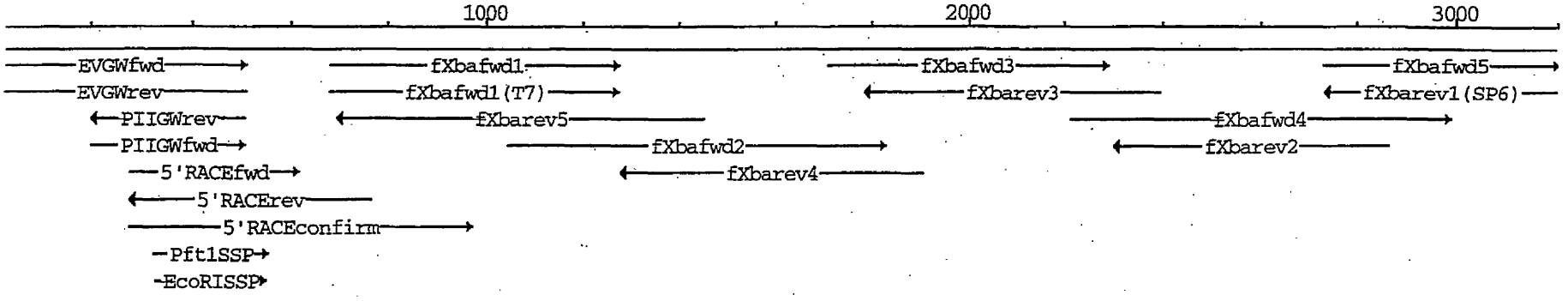


Fig3.9 Contiguous DNA assembly of *pft1* as generated using Autoassembler™ (Perkin Elmer). Sequence annotations are as described in the text. The scale noted above the contiguous assembly relates nucleotide position.

Fig 3.10 The nucleotide sequence and deduced amino acid sequence of *pft1*

```
-276                                atcgcctatagctctcgagatcctcctagaatcctc   -241
-240  aagatcctctagatatatattgtctatccgatctcttagctcttataagcttcggcgccg       -181
-180  caataatcgctataagcgctcctcggatatatacacacatatcgcgctatcttcgctatc       -121
-120  ggggtatctcttcgattctccgattcttatatgccgctaaatccagctgaatcaaatata       -61
-60   agctcttagcccgatcaatactcctgatatcctattattctaatctccttgttccttaat       -1
1     ATGGCACTTGCCGCCTTCACTCCACGAGCATCTTCACGTCGCATTGCTCAGTTTGGATCG       60
(1)     M  A  L  A  A  F  T  P  R  A  S  S  R  R  I  A  Q  F  G  S      (20)
61    CGATCGCTTCCCTTATTTAACTATACCTCCTTCCATGCCCAGAGACCATCGCTTGATGCT       120
(21)    R  S  L  P  L  F  N  Y  T  S  F  H  A  Q  R  P  S  L  D  A      (40)
121   AGCGGTGGCTTAAGACGTCACCAGTCGAACGCCCCCATCGTAACCCAACGACCAGACATC       180
(41)    S  G  G  L  R  R  H  Q  S  N  A  P  I  V  T  Q  R  P  D  I      (60)
181   CTACGGCAGCCAGACATCCGAAAGTCATCTGCAACAAATATACATGTCACTATTTCATCT       240
(61)    L  R  Q  P  D  I  R  K  S  S  A  T  N  I  H  V  T  I  S  S      (80)
241   ATCGTGGGACCTGGTTGTCTTCAGCAGGTCCGATTCCTCTCCTCTACGACAGCACGGCGA       300
(81)    I  V  G  P  G  C  L  Q  Q  V  R  F  L  S  S  T  T  A  R  R      (100)
301   AAAGAAGCCAAACCGGAGCCCGCGAAAAAGCTGCGGAAGCCGCAGAAGCCGCAGACAGAG       360
(101)   K  E  A  K  P  E  P  A  K  K  L  R  K  P  Q  K  P  Q  T  E      (120)
361   CTAGAAAAGAAACAAGAAACCGCTCTAGAAGAGATACATAGAGGATTTGAGCGTACGGAG       420
(121)   L  E  K  K  Q  E  T  A  L  E  E  I  H  R  G  F  E  R  T  E      (140)
421   AAGGCATCACAAGCAGCGAAATTAAACCTTAGCGCAAGATTATCAAAGGATGCGGGCGGG       480
(141)   K  A  S  Q  A  A  K  L  N  L  S  A  R  L  S  K  D  A  G  G      (160)
481   AAGAGGCTCGGCTTTCGCGAGATATGGCGGTTATTGAAAATTGCCCGTCCAGAGGCCAAA       540
(161)   K  R  L  G  F  R  E  I  W  R  L  L  K  I  A  R  P  E  A  K      (180)
541   ATCCTGTCCATGGCGTTACTCTGCTTGCTCATCTCGTCATCTATCACCATGTCGATTCCG       600
(181)   I  L  S  M  A  L  L  C  L  L  I  S  S  S  I  T  M  S  I  P      (200)
601   TTTTCCATCGGCAAGATTCTCGATATCGCAACGCATAGCAGCCCTGAAGGTGGCAATGAA       660
(201)   F  S  I  G  K  I  L  D  I  A  T  H  S  S  P  E  G  G  N  E      (220)
```

```
661   TTGTTCGGCCTCACACTCCCCGTCTTCTACAGTGTCCTGGGCGGTGTTCTTTTGTTGGGT   720
(221)   L  F  G  L  T  L  P  V  F  Y  S  V  L  G  G  V  L  L  L  G   (240)
721   GCTGCTGCAAACTGTGGTCGTATTATAATTCTACGGATTGTGGGGGAACGTATTGTTGCG   780
(241)   A  A  A  N  C  G  R  I  I  I  L  R  I  V  G  E  R  I  V  A   (260)
781   CGACTAAGATCTAAGCTCTTCAGACGTACTTTTATGCAGGATGCGGAGTTTTTCGATGCC   840
(261)   R  L  R  S  K  L  F  R  R  T  F  M  Q  D  A  E  F  F  D  A   (280)
841   AACAGAGTCGGTGACTTGATTTCCCGGCTAAGCTCGGATACGATAATCGTCGGCAAAAGT   900
(281)   N  R  V  G  D  L  I  S  R  L  S  S  D  T  I  I  V  G  K  S   (300)
901   ATCACGCAGAATCTTTCAGATGGACTTAGGGCTGCTGTTAGTGGCGTTGCTGGGTTTGGT   960
(301)   I  T  Q  N  L  S  D  G  L  R  A  A  V  S  G  V  A  G  F  G   (320)
961   CTGATGGCCTTTGTCAGTCTTAAGCTGTCAAGCATTCTCCTTTTGCTGATACCCCCAGTG   1020
(321)   L  M  A  F  V  S  L  K  L  S  S  I  L  L  L  L  I  P  P  V   (340)
1021  TCCCTTGGAGCGTTCTTATACGGCAGATCAATCAGAAATATTAGTCGCAAGATTCAGAAG   1080
(341)   S  L  G  A  F  L  Y  G  R  S  I  R  N  I  S  R  K  I  Q  K   (360)
1081  AACCTGGGTAGCCTGACCAAAATCGCGGAGGAGCGTCTGGGAAATGTGAAGACAAGTCAG   1140
(361)   N  L  G  S  L  T  K  I  A  E  E  R  L  G  N  V  K  T  S  Q   (380)
1141  TCTTTTGCAGGAGAGATTTTGGAGGTCCATCGATATAATAGGCAAACCAGGAGAATTTTT   1200
(381)   S  F  A  G  E  I  L  E  V  H  R  Y  N  R  Q  T  R  R  I  F   (400)
1201  GAGCTGGGAAAGAGAGAGTCTCTCGTTAGCGCAGCGTTCTTTAGCACCgtgagtacttc   1259
(401)   E  L  G  K  R  E  S  L  V  S  A  A  F  F  S  T               (416)
1260  taatcactaattttcatcaaatctggtatatgctgacagctatttgtatctttgcagACC   1319
(417)                                                          T      (417)
1320  GGTTTGATGGGAAATATGACGATTTTGAGTTTACTCTATGTTGGTGGTGGAATGGTCCAG   1379
(418)   G  L  M  G  N  M  T  I  L  S  L  L  Y  V  G  G  G  M  V  Q   (437)
1380  TCCGGGGCCATTTCCATTGGAGACTTGACTTCATTCCTCATGTATGCAGCATACGCTGGA   1439
(438)   S  G  A  I  S  I  G  D  L  T  S  F  L  M  Y  A  A  Y  A  G   (457)
1440  TCAAGCATGTTTGGCTTGTCCAGCTTCTATTCAGAACTGATGAAGGGCGTTGGTGCTGCC   1599
(458)   S  S  M  F  G  L  S  S  F  Y  S  E  L  M  K  G  V  G  A  A   (477)
1500  AGTCGACTATTCGAGCTCCAAGACCGCCATCCCACTATTCCCCTTACTGTAGGAGATAAA   1559
(478)   S  R  L  F  E  L  Q  D  R  H  P  T  I  P  L  T  V  G  D  K   (497)
```

```
1560  GTGGTTTCGGCAAGGGGAGCGATCCGGTTCAAGAATTTGGATTTCAGTTATCCCACCAGA  1618
(498)   V  V  S  A  R  G  A  I  R  F  K  N  L  D  F  S  Y  P  T  R  (517)
1620  CCCGCAGTGACTGTATTTAAAGACCTCGACTTTGAGATTCCGCAAGGAAGCAACGTTGCA  1679
(518)   P  A  V  T  V  F  K  D  L  D  F  E  I  P  Q  G  S  N  V  A  (537)
1680  ATCGTCGGCCCCTCGGGGGGCGGTAAATCCACGATTGCCTCCTTGTTACTGCGCTTCTAC  1739
(538)   I  V  G  P  S  G  G  G  K  S  T  I  A  S  L  L  L  R  F  Y  (557)
1740  AAACATACCAGAGGCCAGATCCTGATAGATGGGAAGGACATCTCCTCAATGAATGCAAAG  1799
(558)   K  H  T  R  G  Q  I  L  I  D  G  K  D  I  S  S  M  N  A  K  (577)
1800  TCTCTACGAAGGAAAATCGGCGTCGTCGCGCAGGAGCCCGTACTCTTCTCTGGAACTATC  1859
(578)   S  L  R  R  K  I  G  V  V  A  Q  E  P  V  L  F  S  G  T  I  (597)
1860  GCTGAGAACATATCCTACGGCAAGCCCCATGCGACAAGGACGGAAATTATCGCAGCGGCG  1919
(598)   A  E  N  I  S  Y  G  K  P  H  A  T  R  T  E  I  I  A  A  A  (617)
1920  CGCAAGGCCAACTGTCAATTCATCAGTGACTTCgtaagtgacatctaccttcctctcca  1978
(618)   R  K  A  N  C  Q  F  I  S  D  F                            (628)
1979  agcattcccacctacccggtttactaatacaagactcgtttctcgattcagCCCGATGGC  2038
(629)                                                  P  D  G  (631)
2039  CTCGACACACACGTAGGCGCCCGCGGCGCTCAACTCTCCGGCGGCCAAAAGCAGCGAATC  2098
(632)   L  D  T  H  V  G  A  R  G  A  Q  L  S  G  G  Q  K  Q  R  I  (651)
2099  GCCATCGCCCGTGCCCTGATTAAAAACCCAGACATCCTCATCCTCGACGAAGCCACGTCC  2158
(652)  A  I  A  R  A  L  I  K  N  P  D  I  L  I  L  D  E  A  T  S  (671)
2159  GCCCTCGATGCCGAATCCGAGACGCTGGTCAACAGCGCTCTTGCCGCGTTACTCCGCGGC  2218
(672)   A  L  D  A  E  S  E  T  L  V  N  S  A  L  A  A  L  L  R  G  (691)
2219  AACAATACGACAATCAGCATCGCCCACCGTCTCTCCACCATCAAGCGCTCCGACACCATC  2278
(692)   N  N  T  T  I  S  I  A  H  R  L  S  T  I  K  R  S  D  T  I  (711)
2279  ATCGTACTCAGCGGTGACGGCCACGTTGCGGAGCAAGGTTCGTATCAGGAACTGAGCGCG  2338
(712)   I  V  L  S  G  D  G  H  V  A  E  Q  G  S  Y  Q  E  L  S  A  (731)
2339  CGACCGGACGGCGCGTTTACGAAGCTTATGGAGTGGCAGATGAGCTCTGAAGGGGGGGCT  2398
(732)   R  P  D  G  A  F  T  K  L  M  E  W  Q  M  S  S  E  G  G  A  (751)
2399  GTGCAGTCGGTGGTTAGGGGACCGCCGTCTGAGAAGGAGGAGTTGCAGCAGATGCTGCAG  2458
(752)   V  Q  S  V  V  R  G  P  P  S  E  K  E  E  L  Q  Q  M  L  Q  (771)
2459  GAAGGGGAGGAAGATTATGGGGAGTATGATGATGATAGTGATGCGGAGCCGGAGAAGCTT  2518
(772)   E  G  E  E  D  Y  G  E  Y  D  D  D  S  D  A  E  P  E  K  L  (791)
```

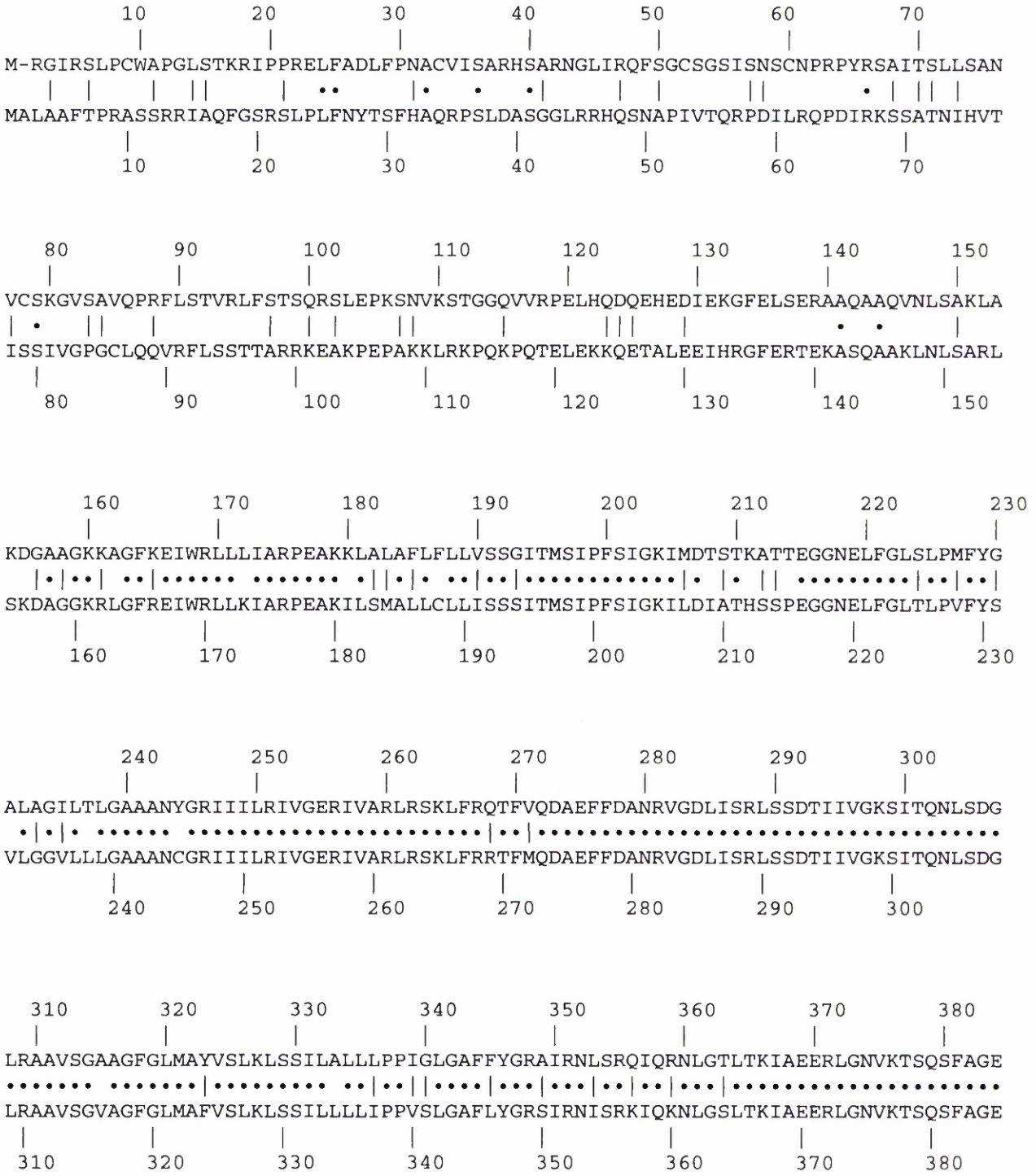
```
2519  GTTGAGAGGGACGGCGTTGCTGAGGGAGCTTCTAAGGAGAAATATGCTGTGGCTGCTGGG  2578
(792)  V  E  R  D  G  V  A  E  G  A  S  K  E  K  Y  A  V  A  A  G  (811)
2579  ATAGAGGCCAGTATTGCTACTTCTAAGCAACAGCCGTCCCAAGAAAAATAGatggatagc  2638
(812)   I  E  A  S  I  A  T  S  K  Q  Q  P  S  Q  E  K  *          (827)
2639  aaatgtgaatactctttgcttgctgtgtcacttccctcctctttgttccccagctcctat  2698
2699  catcccagtcccaaaggtcacgatcctgtggtctcgtccttcaaggcctgaagaccatag  2758
2759  attcgtttctcctttatcctctgcccatcttaccccgtcctccatcgcataggactttgc  2818
2819  taaaccataccttgaactttctctttttttctttccctccatcatctatggatgcatgat  2878
2879  atttggtattttaccgtgcttcctcatgcatttatatattatatctagatgcatgctcga  2938
2939  gcggccgccagtgtgatggatatctgcagaatttccagcacactggcggccggtactagt  2998
2999  gga                                                           3001
```

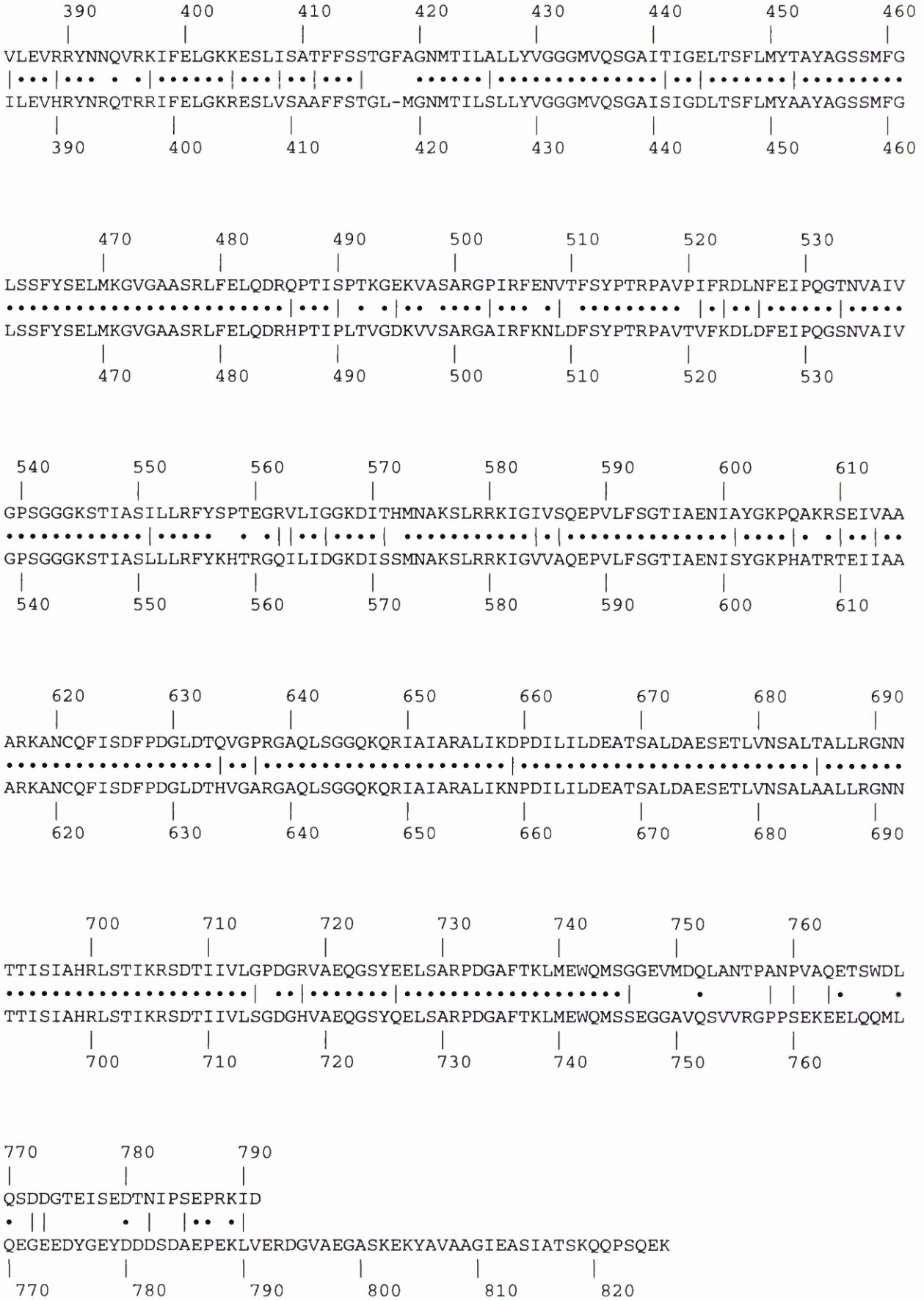
Nomenclature; Nucleotide positions are recorded in black numerals, amino acids in bracketed blue.

Nucleotide data; Coding nucleotides are denoted by uppercase letter. Non-coding bases are denoted by lowercase letters. Putative introns are presented in blue. Potential promoter features (CAAT and TATA elements) are underlined. A putative poly-adenylation signal is present in green.

Amino acid data; The Walker A box is denoted in green. The Walker B box and centre region is denoted in red. The transcriptional stop site is marked as *.

Fig 3.11 Comparative alignment of the deduced amino acid sequences of *afuMDR2* (top) and *pft1* (bottom)




390 400 410 420 430 440 450 460
VLEVRRYNNQVRKIFELGKKESLISATFFSSTGFAGNMTILALLYVGGGMVQSGAITIGELTSFLMYTAYAGSSMFG
ILEVHRYNRQTRRIFELGKRESLVSAAFFSTGL-MGNMTILSLLYVGGGMVQSGAISIGDLTSFLMYAAYAGSSMFG
390 400 410 420 430 440 450 460
470 480 490 500 510 520 530
LSSFYSELMKGVGAASRLFELQDRQPTISPTKGEKVASARGPIRFENVTFSYPTRPAVPIFRDLNFEIPQGTNVAIV
LSSFYSELMKGVGAASRLFELQDRHPTIPLTVGDKVVSARGAIRFKNLDFSYPTRPAVTVFKDLDFEIPQGSNVAIV
470 480 490 500 510 520 530
540 550 560 570 580 590 600 610
GPSGGGKSTIASILLRFYSPTEGRVLIGGKDITHMNAKSLRRKIGIVSQEPVLFSGTIAENIAYGKPQAKRSEIVAA
GPSGGGKSTIASLLLRFYKHTRGQILIDGKDISSMNAKSLRRKIGVVAQEPVLFSGTIAENISYGKPHATRTEIIAA
540 550 560 570 580 590 600 610
620 630 640 650 660 670 680 690
ARKANCQFISDFPDGLDTQVGPRGAQLSGGQKQRIAIARALIKDPDILILDEATSALDAESETLVNSALTALLRGNN
ARKANCQFISDFPDGLDTHVGARGAQLSGGQKQRIAIARALIKNPDILILDEATSALDAESETLVNSALAALLRGNN
620 630 640 650 660 670 680 690
700 710 720 730 740 750 760
TTISIAHRLSTIKRSDTIIVLGPDGRVAEQGSYEELSARPDGAFTKLMEWQMSGGEVMDQLANTPANPVAQETSWDL
TTISIAHRLSTIKRSDTIIVLSGDGHVAEQGSYQELSARPDGAFTKLMEWQMSSEGGAVQSVVRGPPSEKEELQQML
700 710 720 730 740 750 760
770 780 790
QSDDGTEISEDTNIPSEPRKID
QEGEEDYGEYDDDSDAEPEKLVERDGVAEGASKEKYAVAAGIEASIATSKQQPSQEK
770 780 790 800 810 820

the basis of disruption of the expected open reading frame (in comparison with the sequence of homologues), frameshift in the ORF and the presence of semi-conserved splice sites in the sequence which were consistent with those observed in other genes sequenced from *P.brasiliensis* (Da Silva *et al.*, 1997; McEwan *et al.*, 1996; Nino-Vega *et al.*, 1998).

3.6 Discussion

Paracoccidiodes brasiliensis possesses an open reading frame which would appear to encode for an MDR-like ABC half-transporter. A gene, *pft1*, has been sequenced, consisting of 2627 base pairs, interrupted by two putative introns at positions 1248-1317 and 1953-2301. The deduced amino-acid sequence indicates that the gene encodes a 827 residue protein with a predicted molecular weight of 89.5kDa and isoelectric point of 9.95. The *pft1* gene represents the first ABC transporter, indeed the first transporter protein of any kind, to be reported for *P.brasiliensis*.

The application of the MDR1 and MDR2 degenerate primers allowed the amplification of only one gene fragment. The gradient imposed in reaction annealing temperature was imposed in an attempt to amplify fragments from genes with a partial identity to the degenerate primers. Even at relatively low annealing temperatures only a single product is observed. Indeed later sequencing revealed that this product was derived from a partial primer mismatch. The ALVGPSG motif used to design the MDR1 forward primer is degenerated to AIVGPSG in Pft1. This is in keeping with the same observation in *afumdr2*. However in the *A.fumigatus* amplification the additional amplicon revealing *afumdr1* was also generated. The results of the PCR on *P.brasiliensis* genomic DNA suggests that an *afumdr1* homologue is absent in this organism.

The unfortunate truncation of open reading frames during partial digestion is a common shortcoming of this method of library construction. The use of adapter tag SSP-PCR to generate new flanking sequence directly from genomic DNA appears to be an effective and powerful tool for overcoming such problems. The PCR reactions required a degree of optimisation before satisfactory results were observed. Continued optimisation, specifically for each mini-library may have allowed the acquisition of a larger amplicon from libraries that, under our conditions, did not yield a product. It could be considered that successful application of this technology supersedes the need for hybridisation based library screening. It is possible to isolate entire open reading frames using only SSP-PCR. The effort involved in lambda library screening was perhaps

disproportionate in the case of the *pft*1 investigation, the situation being exacerbated by the fact that the gifted library had undergone an undefined number of rounds of amplification prior to acquisition. An unamplified library where selective outgrowth had not occurred would have reduced the probability of repetitive isolation of the same clone as experienced in this investigation.

The gene appears to encode a protein exhibiting a number of features recognised as being typical of ABC transporters. Pft1 is composed of two domains, namely a membrane spanning domain at the N-terminus and a typical P-glycoprotein-like nucleotide-binding domain at the C-terminus. A Kyte-Doolittle plot, indicating hydropathy, was performed by Tobin for AfuMDR2 and was interpreted as proving the membrane spanning domain to have at least four hydrophobic, potentially membrane spanning segments. The same plot, on Pft1, was virtually identical, which is unsurprising considering the degree of amino-acid conservation (data not shown). The plots are not definitive however, and the absence of circular dichroism data at this point prohibits any conclusive enumeration of α-helices that may span the membrane. The implications of the number of spanning helices, and indeed the fact that the gene is a half-transporter molecule, will be further discussed, in the context of other data, in Chapter 8. Despite a widely acknowledged lack of conservation between membrane spanning domains of different ABC transporters (Higgins, 1992) the sequence homology with the other fungal half transporters of this type remains high in the MSD, although is less than observed for the C-terminal nucleotide binding domain. This high degree of identity would imply a selective pressure for conservation implicating the association of these genes with an advantageous function.

Featured within the amino acid sequence of the nucleotide binding domain are motifs that are recognised as being diagnostic of ABC transporters, namely the Walker A (AIVGPSG) and B (DEATSALD) motifs and the signature region in the nucleotide binding domain. While the Walker A and B sites of many typical fungal ABC transporters, such as AtrA (Del Sorbo *et al.*, 1997), PDR5 (Balzi *et al.*, 1994; Bissinger & Kuchler, 1994) and CDR1 (Prasad *et al.*, 1995) are heavily degenerated, those possessed by *pft1* are less so. These particular sequences display far more similarity to mammalian proteins than such fungal drug transporters. Indeed, the structural topology of Pft1 is more closely related to mammalian molecules as the MSD precedes the NBD. In the aforementioned fungal transporters it is the hydrophilic domain that pre-empts the hydrophobic.

A considerable interest has recently emerged regarding appropriate model systems for studying homologues of P-glycoprotein. It has been proposed that similar molecules

could be used to generate useful data that would relate directly to anti-cancer drug resistance. Indeed the use of LmrA of *Lactococcus lactis* to confer an MDR-like phenotype on a mammalian cell line demonstrates the efficiency of ABC half transporters from lower organisms in this endeavour (Van Veen *et al.*, 1998). The similarity of Pft1 to P-glycoprotein, particularly in the NBD, could add additional importance for the investigation of this protein in such a context.

The short 5' sequence flanking the proposed ATG start codon exhibits a number of possible promoter elements. In particular there are a number of sequences that may pertain to "CAAT" or "TATA" boxes. It is especially difficult to ascribe any definitive importance to any individual element observed in the promoters of such fungal genes. It is typically observed that the organism degenerates motifs involved in regulation with respect to both sequence and position. Without extensive further investigation, involving nucleotide protection assays and/or reporter fusion assays on selected promoter mutants, at this time it is impossible to make any significant analysis of the particulars of the transcriptional initiation of *pft1*. However, at this stage it was considered interesting that the *pft1* gene was not observed in either of the cDNA lambda libraries on PCR analysis. Initial explanations could include a technical failure to clone the required cDNA during library construction, the fact that *pft1* is not an expressed reading frame in *P.brasiliensis*, or that induction of *pft1* expression is tightly regulated and that expression was absent under the conditions used to generate cDNA for each library. Clarification of this issue was considered very important and is dealt with in Chapter 5.

Pft1 represents a fourth member of a small but growing subfamily of fungal ABC transporters. AfuMDR2, the two half transporters of *S.cerevisiae* (MDL1 and MDL2) (Dean, *et al.*, 1994), and now Pft1 suggest that an important role exists for these proteins. Until the closing stages of the writing of this thesis no function had been assigned to any of these proteins. Very recently, Young *et al.* (2000) demonstrated that MDL1 and MDL2 can be found on the *S.cerevisiae* mitochondrial membrane. MDL1, but not MDL2, is responsible for the transport of oligopeptides that are products of the action of the m-AAA protease. The sequence similarity between Pft1 and MDL1 could suggest a similar role for this transporter. However the sequence of Pft1 is more akin to that of MDL2 which does not act as an oligopeptide transporter. Later in this thesis we demonstrate an alternative substrate affinity for Pft1 that suggests a possible involvement in other processes.

The *afumdr2* gene of *A.fumigatus* remains uninvestigated and can provide no further information on any potential function of these genes. Further investigation should be

performed. It is possible that these molecules will possess affinity for, and even be involved in the transport of oligopeptides. The extensive sequence homology with Pgp may result in an affinity for hydrophobic substrates that may include drugs, pheromones or lipids. The striking similarity at the amino-acid sequence level with the larger, full sized, counterparts supports this notion.

Chapter 4

Identification and sequencing of *pft2*, a full-sized PDR-like ABC transporter gene from *Paracoccidiodes brasiliensis*

While *pft1* represents the first ABC transporter described for *P.brasiliensis* it is not especially homologous to the majority of transporters known to be involved in fungal drug resistance. Pft1 is a member of the MDR subclass of ABC transporters. While the MDR-like AfuMDR1 has been shown to transport antifungals (Tobin, *et al.* 1997), pleiotropic drug resistance in fungi is best characterised in the PDR subclass of proteins that includes PDR5 (Balzi, *et al.* 1994). These transporters, exhibiting an $(NBD\text{-}MSD)_2$ topology, are heavily degenerated with respect to the conserved motifs of P-glycoprotein-like molecules although conservation between members of the subclass can be considerable (Bauer, *et al.* 1999). These transporters are of significant importance. The PDR-like CDR1 from *C.albicans* has been shown to be involved in clinical azole resistance (Sanglard, *et al.* 2001).

Following the successful application of degenerate primers in PCR and the identification of *pft1* it was decided to develop a second set of primers specifically targeting PDR-subclass genes. This could screen the *P.brasiliensis* genome for the presence of genes encoding these transporters. On considering the existence of 30 ABC transporters in the genome of *S.cerevisiae* (Goffeau, *et al.* 1996) it is almost certain that *pft1* would not represent the comprehensive inventory of *P.brasiliensis* ABC transporters.

This chapter reports the successful application of the degenerate primers PDR1 and PDR2 in amplifying a fragment of a second gene encoding a PDR-like ABC transporter in *P.brasiliensis*. This amplicon was used to screen a lambda genomic library allowing the isolation of a single phage clone. Subcloning of fragments derived from the DNA of this clone allowed a partial sequence to be compiled. The DNA sequenced to this point spans 3792bp and a sequence of 1053 amino acids can be deduced from this data. The partial sequence is thought to belong to an open reading frame encoding an ABC transporter belonging to the PDR subfamily.

4.1 Multiple alignments and the design of degenerate oligonucleotides

Following the success of the MDR subclass oligonucleotides in the identification of *pft1* a similar approach was adopted to identify other transporters. In this approach a multiple alignment of notable proteins of the PDR-subfamily was performed. Amino acid sequences of the C-terminal nucleotide-binding domains of the selected proteins were eventually targeted for primer identification. The results of the multiple alignment, highlighting the motifs selected for primer design, can be seen in Fig4.1. Also illustrated are the nucleotide sequences with degeneracies used as deduced from the sequences of the targeted amino acid motifs. Again the allocation of degeneracies was based on the codon usage of *N. crassa* and *A.nidulans* (Gurr, *et al.*, 1997).

The multiple alignment identified a number of potential targets for primer design. The two targets chosen, AGKTTLL and TIHQPSA, allowed for a minimal amount of degeneracy based on codon usage and represented two regions that were totally conserved in all of the proteins of this class. This is in contrast to the MDR1 primer used for *pft1*, where a single amino acid could have been, and was, degenerated from L to I.

4.2 PCR amplification of a PDR-like gene fragment using degenerate PCR

Primers PDR1 and PDR2 were used in a PCR reaction with *P.brasiliensis* genomic DNA as the template. The results of the PCR can be observed in Fig4.2. A band of c.500bp was observed (the expected size was 495bp, assuming no introns were present) and also a smaller band of c.280bp. Despite the use of a range of annealing temperatures, only these bands were observed. The performance of PCRs over a range in temperatures was variable but higher annealing temperatures favoured consistent high levels of amplification. Both amplicons from the 62^0C reaction were cloned into pGEM-T Easy™ and sequenced using T7 and SP6 sequencing primers. Three individual clones were analysed for each fragment in order to ascertain whether multiple amplicons of similar size were present.

```
AtrA AGKTTLLDVLAHRTTMGVIT-GDMFVNGKGLDASFQRKTGYVQQQDLHLETATVRESLRFSALLRQPASVSIREKHDYVESV
Cdr1 AGKTTLLNCLSERVTTGIITDGERLVNGHALDSSFQRSIGYVQQQDVHLPTSTVREALQFSAYLRQSNKISKKEKDDYVDYV
Cdr2 AGKTTLLNCLSERVTTGIITDGERLVNGHALDSSFQRSIGYVQQQDVHLETTTVREALQFSAYLRQSNKISKKEKDDYVDYV
Cdr3 AGKTTLLNALSERLTTGVITSGTRMVNGGELDSSFQRSIGYVQQQDLHLETSTVREALKFSARLRQPNSVSIAEKDSYVEKI
Pdr5 AGKTTLLDCLAERVTMGVIT-GDILVNGIPRDKSFPRSIGYCQQQDLHLKTATVRESLRFSAYLRQPAEVSIEEKNRYVEEV

AtrA IEMLGMGDFCRACCGTPGEGLNVEQRKLLTIGVELPPSPKLLLFLDEPTSGLDSQSSWAICTFLRKLADSGQAVLCTIHQPSA
Cdr1 IDLLEMTDYADALVGVAGEGLNVEQRKRLTIGVELVAKPKLLLFLDEPTSGLDSQTAWSICKLMRKLADHGQAILCTIHQPSA
Cdr2 IDLLEMTDYADALVGVAGEGLNVEQRKRLTIGVELVAKPKLLLFLDEPTSGLDSQTAWSICKLMRKLADHGQAILCTIHQPSA
Cdr3 IDLLEMRTYVDAIVGVPGEGLNVEQRKRLTIAVELVARPKLLVFLDEPTSGLDSQTAWSICKLIRKLANHGQAILCTIHQPSA
Pdr5 IKILEMEKYADAVVGVAGEGLNVEQRKRLTIGVELTAKPKLLVFLDEPTSGLDSQTAWSICQLMKKLANHGQAILCTIHQPSA
```

Forward Primer (PDR1);	A	G	K	T	T	L	L
	GG(TC)	GG(TC)	AAG	AC(TC)	AC(TC)	CTC	CTC
Reverse Primer (PDR2);	T	I	H	Q	P	S	A
	AC(TC)	CG(GA)	GGG	CTG	GTG	GAT	CG(CT)
{PDR2 reverse compliment	(AG)CG	(CG)GA	GGG	CTG	GTG	GAT	(AG)GT }

Fig 4.1

Fig4.2 PDR1/PDR2 degenerate PCR analysis of *P.brasiliensis* genomic DNA

Lanes 1 to 12 are PCR reactions with annealing temperatures increasing in 1^0C increments (54^0C to 66^0C). Lane 13 is a template free negative control performed with the annealing step at 60^0C.

PCR conditions; 100µl reaction volume containing 1.5mM MgCl. Activation at 95^0C for 15min. Denaturation at 94^0C for 30s / annealing at various temperatures for 40s / extension at 72^0C.A total of 30 cycles concluded with a final 10min extension at 72^0C.

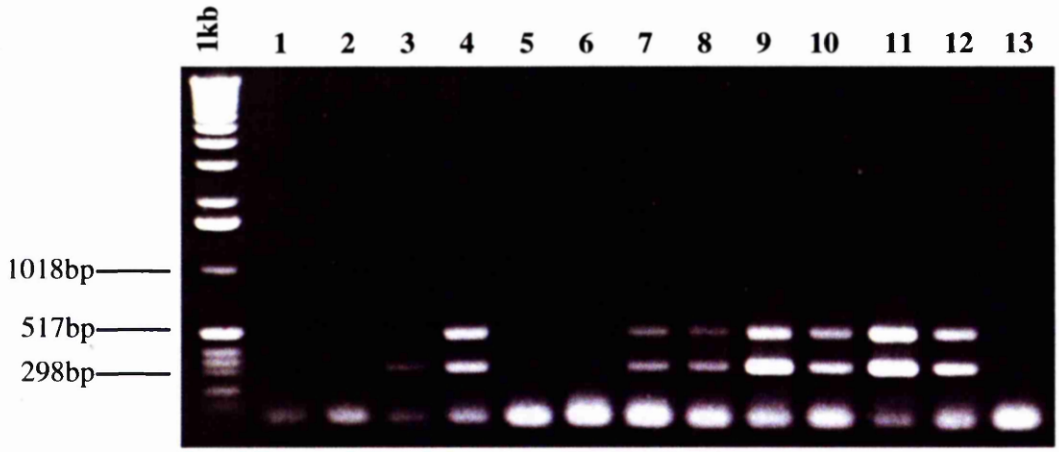


Fig 4.2

The sequence of the smaller product showed no obvious reading frame and no homology with any known sequence when analysed on BLAST at either the nucleotide or amino acid level. The primers were not incorporated into the amplicon in frame. This product was not considered to represent an ABC transporter gene and was not investigated any further.

The replicate sequence results for the clones containing the larger c.500bp fragment showed that all clones contained identical 492bp DNA. When analysed by BLAST the DNA showed considerable homology to other PDR-subclass ABC transporters. The first 20 hits on the BLASTx algorithm are displayed in Fig4.3. Highest homology is observed with AtrA of *Aspergillus nidulans* but considerable identity and similarity is evident for transporters involved in clinical and experimental drug resistance including PDR5 and CDR1. The sequence was predicted to belong to a putative open reading frame that we termed *pft2*.

4.3 Screening of lambda libraries for the *pft2* gene

The PCR using PDR1 and PDR2 was repeated using total library DNAs. Again the λDASHII genomic DNA library (constructed using a partial Sau3A digest of *P.brasiliensis* genomic DNA) and the two λZAP cDNA libraries (constructed using total cDNA from either the yeast form or the mycelial form of the organism) were considered. As in the case of *pft1*, the gene fragment could only be detected by PCR in the λDASHII genomic library and was absent in both cDNA libraries (data not shown).

It was decided to pursue the *pft2* gene in the λDASHII library only. The library was diluted by a factor of 10^{-6} and 100μl of this dilution was used to infect LE392 *E.coli* cells. This allowed for 450 plaques per 10cm x 10cm plate (see section 3.2).

Hybridisation on membranes derived from λ plaque lifts was performed. Following detection, a predominant hybridisation signal was observed in an identical position on two replicate filters as shown in Fig4.4a. Specific hybridisation conditions are as noted in the figure legend. Secondary screening of this lambda isolate also proved

Accession Number	Protein : Source	Author
1] gi|7493926|	AtrA; *Aspergillus nidulans*	(Del Sorbo, *et al.* 1997)
2] gi|1168650|	Bfr1; *Schizzosaccharomyces pombe*	(Nagao, *et al.* 1995)
3] gi|13446213|	Brefeldin A transporter; *Schizzosaccharomyces pombe*	(See note 1)
4] gi|3420017|	Cgr1; *Candida glabrata*	(Miyazaki, H. Unpub. data)
5] gi|3288709|	Pmr1; *Penicillium digitatum*	(Nakaune, *et al.* 1998)
6] gi|5327107|	MgAtr2; *Mycosphaerella graminicola*	(Zwiers, 2000)
7] gi|6175524|	CgCDR1; *Candida glabrata*	(Sanglard, *et al.* 1999)
8] gi|1321667|	Ydr1; *Saccharomyces cerevisae*	(Hirata, *et al.* 1994)
9] gi|6324727|	Pdr5; *Saccharomyces cerevisae*	(Goffeau, *et al.* 1996)
10] gi|7160817|	AtrE; *Aspergillus nidulans*	(Andrade, A. Unpub. Data)
11] gi|913016|	Snq2 homologue; *Schizzosaccharomyces pombe*	(Miyazaki, *et al.* 1998)
12] gi|7493952|	ABC1; *Pyricularia grisea*	(Urban, *et al.* 1999)
13] gi|6320614|	Pdr15; *Saccharomyces cerevisae*	(Goffeau, *et al.* 1996)
14] gi|12248921|	ABC transporter; *Rhizoctonia solani*	(Young Jin, L. Unpub. Data)
15] gi|6324904|	Pdr10; *Saccharomyces cerevisae*	(Goffeau, *et al.* 1996)
16] gi|8101775|	BcatrO; *Botryotinia fuckeliana*	(Del Sorbo, G. Unpub data)
17] gi|5921714|	Cdr4; *Candida albicans*	(Franz, *et al.* 1998)
18] gi|12248923|	ABC transporter; *Alternaria alternata*	(Young Jin, L. Unpub. Data)
19] gi|5921713|	Cdr3; *Candida albican*	(Balan, *et al.* 1997)
20] gi|2492604|	Cdr2; *Candida albicans*	(Sanglard, *et al.* 1997)

Fig 4.3

This figure is adapted from a BLAST output generated when searching the NCBI database (http//www.ncbi.nlm.nih.gov/BLAST/) using the 492bp PDR1/PDR2 amplicon sequence. The 20 genes with highest homology scores are presented as generated using the BLASTx algorithm. The description of identity for each hit has been altered from the original output to ease reader identification.

Note 1; Identified by the European *S.pombe* genome sequencing project, Sanger centre.

Fig4.4a Primary hybridisation screen of λDASH genomic library

The left-hand image represents the signal obtained from the initial lift of plaque material from a phage library culture plate. The right-hand image is the signal obtained from a second plaque lift from the same plate. Thus replicate membranes were available to confirm the position of positive signals.

Hybridisation conditions; 140ng of labelled probe (at 7ng per ml of hybridisation buffer) was applied to replicate filters. The membranes were incubated in the presence of probe for 18h at 62^{0}C. A series of 3 stringency washes were applied using 2xSSC/0.1%SDS, 1xSSC/0.1%SDS and 0.5xSSC/0.1%SDS in succession. Again the solutions were at 62^{0}C. The membrane was blocked for 1h at room temperature. Following this, anti-fluorescien antibody was applied at 1:2000 dilution of stock as supplied. Unbound antibody was washed off, detection reagent applied and exposure to X-ray film was allowed for 20min.

Fig4.4b Replicate filters of secondary screen on the *λpft2* clone

Hybridisation conditions; As described for Fig4.4a, except hybridisation temperature and stringency washes were performed at 65^{0}C.

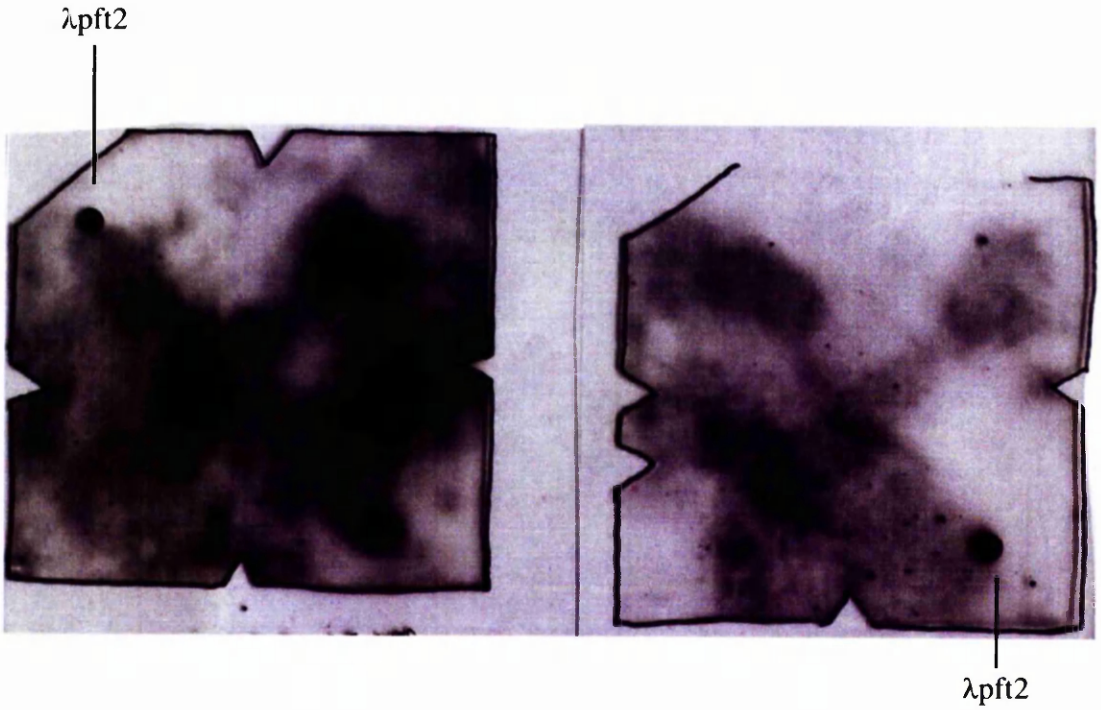


Fig 4.4a

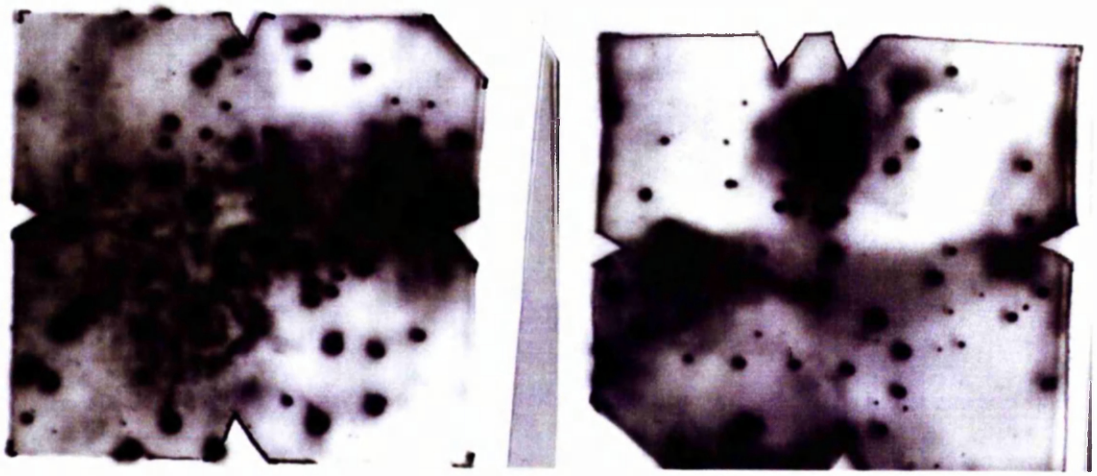


Fig 4.4b

positive as shown by the strong hybridisation signals observed on replicate membranes prepared from plaque lawns of this clone. The results are shown in Fig4.4b. Again, specific hybridisation conditions are noted in the legend.

4.4 Subcloning *P.brasiliensis* DNA from λ*pft2*

In order to obtain fragments of the *pft2* gene from the lambda clone for further sequence analysis a series of subcloning experiments were performed. Following a high multiplicity liquid culture infection of LE392 with λ*pft2*, the DNA from the phage was prepared as previously described. To ten individual restriction enzyme digests was added 500ng of phage DNA. Each restriction digest was performed using a different restriction enzyme, chosen on the basis on their suitability for cloning into the MCS of pZERO™. The restriction digests were electrophoresed on a 0.9% agarose gel (Fig4.5a). The DNA was then transferred to nitro-cellulose membrane and exposed to Southern hybridisation using the PDR1/PDR2 fragment as a probe. Labelling of the probe with fluorescein was assessed on UV illumination and 140ng of labelled probe was applied to the hybridisation (at 7ng per ml of hybridisation solution). The resultant Southern blot is presented in Fig4.5b.

The Southern blot revealed fragments of various sizes that could have been accessible to cloning. (Two bands exist on the *Hin*dIII and *Xba*I lanes as these sites are internal to the sequence of the PDR1/PDR2 fragment probe.) Initially, it was decided to use the c7.8kb *Eco*RI fragment as a principal target for cloning. Following isolation by gel extraction, this fragment could be digested further by *Hin*dIII, providing smaller DNAs with incompatible cohesive ends, allowing more efficient cloning.

*Hin*dIII digestion of the c.7.8kb *Eco*RI fragment resulted in the generation of two separate *Eco*RI/*Hin*dIII fragments (2.2kb/4.2kb) from the ends of the original fragment and a central 1.4kb *Hin*dIII fragment (data not shown). This digestion was cleaned and used in a ligation with *Eco*RI/*Hin*dIII digested pZERO™ DNA. Of the two potential clones (2.2kb insert and 4.6kb insert) only the 2.2kb could be identified among the recombinants, in plasmids pEH4, pEH5 andpEH6 (Fig4.6). This insert of

Fig4.5a Restriction digests of λ*pft2* DNA

Digests were performed using 500ng of DNA per digest in a total volume of 20µl. DNA was incubated at 37^0C for 4h in the presence of 5 Units of the appropriate restriction endonuclease.

Lane 1]	*Hind*III
Lane 2]	*Kpn*I
Lane 3]	*Sac*I
Lane 4]	*Bam*HI
Lane 5]	*Eco*RI
Lane 6]	*Pst*I
Lane 7]	*Eco*RV
Lane 8]	*Not*I
Lane 9]	*Xba*I

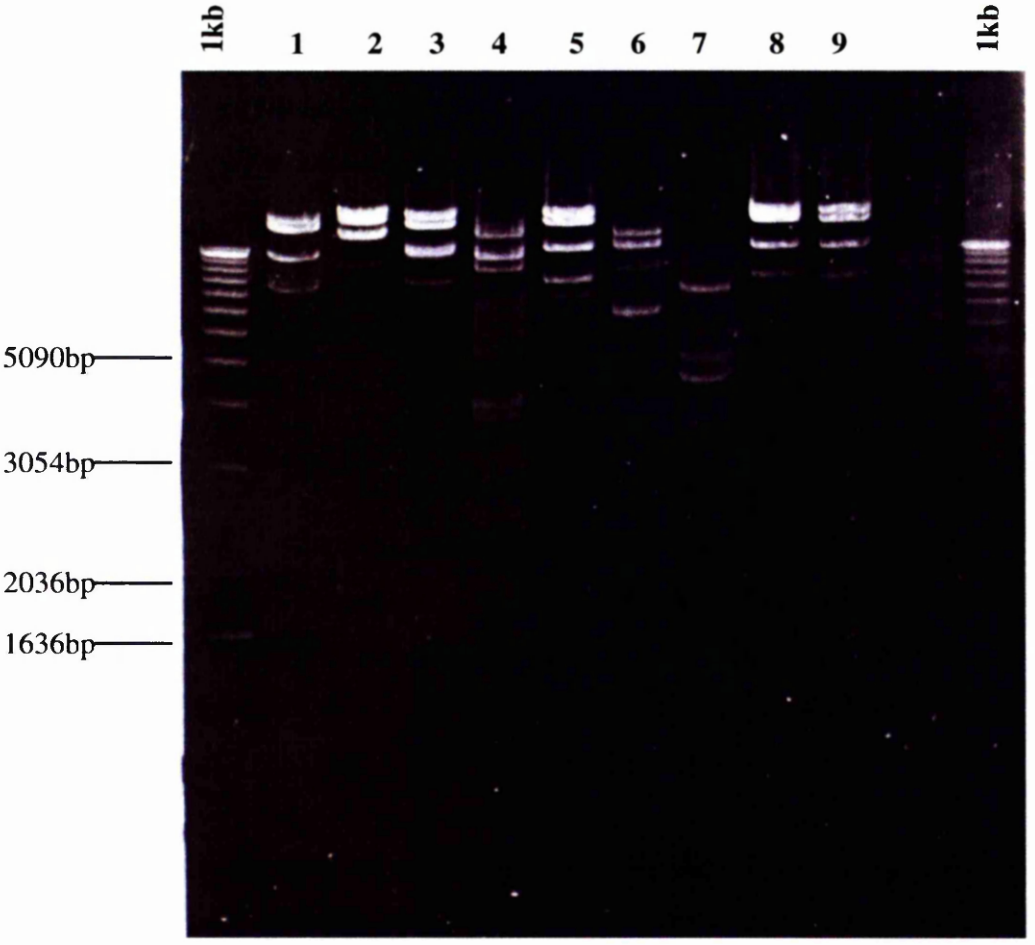


Fig 4.5a

Fig4.5b Southern Blot of gel in Fig4.5a

Lanes are as Fig 4.5a

The blot was hybridised at 65^{0}C for 40h with 500ng of the *pft2* gene fragment as a probe (25ng per ml of hybridisation buffer). The blot was washed for 13min with each of 2X SSC/0.1% SDS, 1X SSC/0.1% SDS and 0.5X SSC/0.1% SDS at 60^{0}C. Anti-fluorescein was allowed to bind and the blot was exposed to X-ray film for 10min following the addition of detection reagent.

Fig 4.5b

pEH6 was then sequenced using T7 and SP6 primers followed by gene specific primers in a walking strategy.

The intervening 1.4kb between the 3' terminus of the pEH6 insert and the original PDR1/PDR2 fragment was amplified by PCR (using a proof-reading polymerase) circumventing the need to accumulate this insert from the lambda DNA. Two opposing primers (SPANR1 and SPANF1) were designed to anneal to sequences from the *pft2* fragment and the pEH6 insert respectively. These primers were applied to the phage clone DNA and the product of the PCR is also illustrated in Fig4.6 on cloning into pGEM-T Easy™. As can be observed, the expected product of c1.6kb was cloned. Once cloned the DNA of pSPAN5 was sequenced.

It was evident that the DNA sequence obtained at this point contained the immediate promoter and the translational start codon for *pft2*. The 3' sequence remained to be sequenced. This would be present in the 4.2kb *Eco*RI/*Hin*dIII fragment that failed to clone in the initial experiment. Despite repeated attempts this larger fragment would not clone to pZERO™ in the presence of the smaller 2.2kb DNA also generated in the restriction digest of the λDNA. In attempts to subclone this 4.4kb fragment, extensive digestion of large amounts of the λ*pft* 2 DNA was performed initially with *Eco*RI and the *Hin*dIII. A total of 20 restriction digests were performed using *Eco*RI. Each restriction digest received 5μg of λ*pft2* DNA. The volume was made to a total of 20μl and the reactions incubated for 3h at 37^{0}C. All digests were electrophoresed on a 0.9% gel and the 7.8kb *Eco*RI fragments was purified in 6 Qiaex II gel extraction procedures. The 20μl eluates from the gel extractions were applied to six *Hind*III restriction enzyme digests and the products were again electrophoresed on a 0.9% gel. The 4.4kb *Eco*RI/*Hin*dIII fragments were gel extracted allowing the accumulation of 250ng of the fragment in a volume of 40μl. Ligation reactions were attempted on a number of occasions using a wide range of ligation conditions and vector:insert ratios. Positive control reactions were performed simultaneously with each test ligation and all were transformed into Novablue competent cells. Despite the consistent good performance of the control reactions in these experiments, no clone containing the required *Eco*RI/*Hin*dIII fragment could be identified.

Fig4.6 **Restriction enzyme digestion of plasmids containing the 2.2kb *Eco*RI/*Hind*III fragment and the SPANR1/SPANF1 PCR amplicon**

Lanes 1 to 6 are *Hind*III digests of potential pZERO™ clones containing the *Eco*RI/*Hind*III fragment. These were denoted "pEH". Only clones pEH3, pEH5 and pEH6 contained the required insert.

Lanes 7 to 11 are *Eco*RI digests of potential pGEM-T Easy™- clones containing the SPANF1/SPANR1 amplicon. This was successfully cloned in pSPAN4 and pSPAN5.

Lane 1]	pEH1
Lane 2]	pEH2
Lane 3]	pEH3
Lane 4]	pEH4
Lane 5]	pEH5
Lane 6]	pEH6
Lane 7]	pSPAN1
Lane 8]	pSPAN2
Lane 9]	pSPAN4
Lane 10]	pSPAN5
Lane 11]	pSPAN6

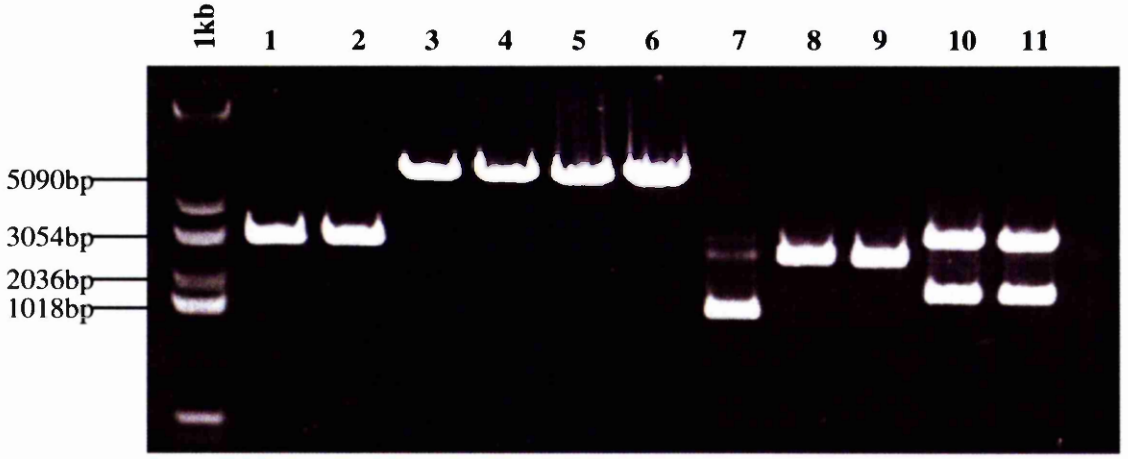


Fig 4.6

The failure to clone this remaining *pft2* fragment could have occurred for a number of technical reasons. While 4.4kb is a larger fragment, the successful cloning of DNA of this size is commonplace. However, it does require a large concentration of insert in order to drive the ligation to a satisfactory result. The extensive repurification of fragments following sequential digests meant that a very high concentration of insert could not be obtained easily. The laborious gel extraction and digestion procedure may also have damaged the termini of the insert DNA making it unreceptive to sticky end cloning. However, it should be noted that the vector DNA was prepared simultaneously and identically to the insert DNA and continued to perform well in control reactions. An alternative explanation is that pZERO™ containing the insert was deleterious to the host strain on transformation and the outgrowth of transformants was inhibited. The application of ligation reactions to SSP-PCR should have demonstrated successful ligation independently of transformation result. All attempts at SSP-PCR on these ligations were unproductive. The recalcitrance of this final fragment of DNA to both PCR and cloning remains unexplained. Attempts to complete the sequence of this gene are ongoing.

4.5 The sequence of *pft2*

All sequences were compiled using AutoAssembler™. The diagrammatical representation of the contig is presented in Fig4.7. It should be noted that any area of the DNA has been sequenced at least in duplicate, and in opposing directions. Sequences derived from pEH6 run between "sp6" and "T7". The two aforementioned sequences were gathered on sequencing pEH6 from the SP6 and T7 sites flanking the MCS. The application of gene specific primers are denoted as "F1" - "F5" and "R1" - "R5", "F" for the forward direction and "R" for the reverse. The sequences derived from the internal *pft2* PCR include "Span" in the annotation. Primers applied to internal sequences of this amplicon are also marked as "IF" or "IR" depending on the

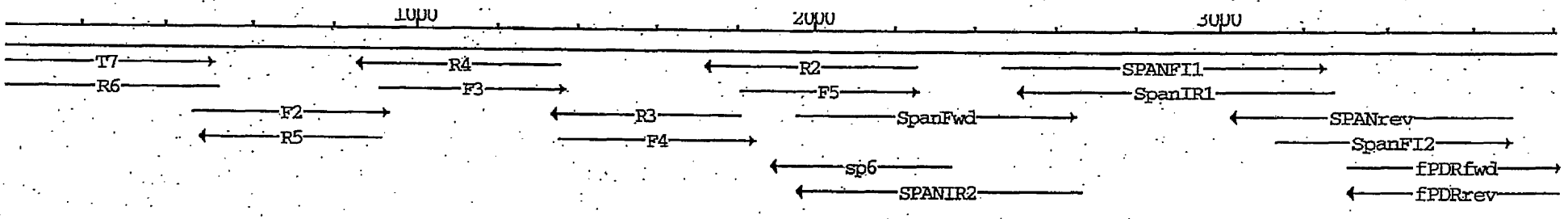


Fig4.7 The contiguous assembly of the partial sequence of *pft2* using sequences derived from ABI sequencing reactions. The assembly was performed using AutoAssembler 2.0™ (Perkin Elmer).

orientation. The sequences of the original PDR1/PDR2 fragment are denoted with "fPDR".

The nucleotide sequence and deduced amino acid sequence of the partial *pft2* contig are presented in Fig4.8. Additional annotations are as described in the legend. Fig4.9 displays the alignment of the partial Pft2 amino acid sequence with that of AtrA.

4.6 Discussion

Following the successful application of degenerate primers in the amplification of the *pft1* gene fragment it was decided to extend this approach to identify other genes for ABC transporters. Multiple alignments of amino acid sequences of proteins that belong to the PDR-subfamily of ABC transporters allowed the identification of two totally conserved motifs in the C-terminal nucleotide-binding domain. The design of degenerate primers for PCR allowed the isolation of a 492bp fragment with extensive similarity to a number of notable transporters, including PDR5 (Balzi, *et al.* 1994), demonstrating the PDR-subfamily in *P.brasiliensis*. A number of clones containing the initial c.500bp amplicon were sequenced and all proved to have an identical insert suggesting that the PCR amplified from only one coding sequence. This sequence was proposed to belong to a larger open reading frame that we termed *pft2*.

A single lambda phage clone was isolated from a λDASHII *P.brasiliensis* genomic DNA library. Subcloning and sequencing of a 2.2kb fragment from this clone revealed a partial sequence of an ABC transporter including a well defined translational start codon and 397bp of the immediate promoter with a number of putative "CAAT" and "TATA" elements. This 2.2kb fragment did not contain the sequence of the hybridisation probe (the PDR1/PDR2 amplicon). The intervening sequence between these two regions was amplified by a proof-reading polymerase and cloned and sequenced.

Fig4.8 **The partial nucleotide sequence and deduced amino acid sequence of *pft2***

```
-397                                gaattcggatccggcgacggaggcgggtccttgggt        -361
-360  tctagctgacaatcgtgaaaccagctgccctcaacacccatcccccattcattcaattggc        -301
-300  aggaccagattcgccttgctcaaacccagagttccaggctggaaaactgcccaatctctg         -241
-240  taataatccccaagagagcaccttggaacacattttccccggggcaatttatactccagt         -181
-180  agtgggcaaataacaccacatctctttcgtctcacagttccaggctgtaattctagttta         -121
-120  ttgaactacctcgttcttttttcttatatacatatatatatatgtatatatatatatctgt         -61
 -60  ataattgctttttgaatccttacaaggctgtgcccgatatcaaactctctccaagaacaa          -1
   1  ATGGCGGGCGATTCGCCGCAATCTTCATTTCACGAAGATCCAGAATACCTTCCGAACGTG           60
 (1)    M  A  G  D  S  P  Q  S  S  F  H  E  D  P  E  Y  L  P  N  V         (20)
  61  AAGGCAAACGATTGCAATGACGATGAAGATGAATCAAACAATGGTGATTCTTCTTCTAAC          120
(21)    K  A  N  D  C  N  D  D  E  D  E  S  N  N  G  D  S  S  S  N         (40)
 121  GAGGATACCAGGTTTAATCTCATTCGAACATTAACCAGAGACCGAAATACAGAGCTCTCG          180
(41)    E  D  T  R  F  N  L  I  R  T  L  T  R  D  R  N  T  E  L  S         (60)
 181  AGAATAGCCTCCGTCTTCTCCCATGCCGACGAGCATTCCACTGCTACTGACTCCCTAGCA          240
(61)    R  I  A  S  V  F  S  H  A  D  E  H  S  T  A  T  D  S  L  A         (80)
 241  CGGATAGATACCCTTGCTGGCCTCCAACTCGGAGATGCCGTGCTGAACCCATCCAGCCCG          300
(81)    R  I  D  T  L  A  G  L  Q  L  G  D  A  V  L  N  P  S  S  P        (100)
 301  CAGTTCGACTTTTACAAATGGGCCCGCATGCTTATGAAGCTGATGGAGGAGGATGGGCTC          360
(101)   Q  F  D  F  Y  K  W  A  R  M  L  M  K  L  M  E  E  D  G  L        (120)
 361  AAGAGACGACGGACCGGCATCACCTTCAGGAACCTGAGTGTCTACGGTTCCGGGCCCGCG          420
(121)   K  R  R  R  T  G  I  T  F  R  N  L  S  V  Y  G  S  G  P  A        (140)
 421  CTGCAGCTGCAAAGCACCGTCTCCACACCGATCATGGCACTCTTCCGGTTCCAAGAGACG          480
(141)   L  Q  L  Q  S  T  V  S  T  P  I  M  A  L  F  R  F  Q  E  T        (160)
 481  TTTGGGGTTGGGCGGAAGACACGCAAGCGGATCTTGAACAACTTCAATGGAGCGTTGAGG          540
(161)   F  G  V  G  R  K  T  R  K  R  I  L  N  N  F  N  G  A  L  R        (180)
 541  GAGGGGGAGATGCTTGTCGTGCTGGGGAGGCCCGGGAGTGGATGCAGCACCTTTCTCAAG          600
(181)   E  G  E  M  L  V  V  L  G  R  P  G  S  G  C  S  T  F  L  K        (200)
 601  ACCATATGTGGAGAAACACATGGGCTTGAGCTGGGCAAGGAGGCCTCGGTTCAATACAAC          660
(201)   T  I  C  G  E  T  H  G  L  E  L  G  K  E  A  S  V  Q  Y  N        (220)
 661  GGTATTCCCCAGACAACGTTCAAGAAGGAATTTCGCGGCGAAGCTGTGTACAGTGCGGAG          720
(221)   G  I  P  Q  T  T  F  K  K  E  F  R  G  E  A  V  Y  S  A  E        (240)
```

```
 721  GATGAGAAACATTTCCCCCATCTTACAGTTGGACAGACGTTGGAGTTTGCTGCTGCTTGT  780
(241)   D  E  K  H  F  P  H  L  T  V  G  Q  T  L  E  F  A  A  A  C  (260)
 781  CGGACTCCTTCTGCAAGAGTGATGGGCATGGAGAGGAAGGAGTTCTCTCACCATATTGCA  840
(261)   R  T  P  S  A  R  V  M  G  M  E  R  K  E  F  S  H  H  I  A  (280)
 841  AGGGTGGTCATGGCCATTTTCGGCTTAAGCCATACCGTCAACGCCAAAGTTGGTGATGAC  900
(281)   R  V  V  M  A  I  F  G  L  S  H  T  V  N  A  K  V  G  D  D  (300)
 901  TACGTCCGAGGCGTCAGTGGCGGCGAACGCAAACGTGTCAGCATTGCGGAGCTTGGGCTC  960
(301)   Y  V  R  G  V  S  G  G  E  R  K  R  V  S  I  A  E  L  G  L  (320)
 961  TCTGGCGCACCGGTAGTCTGCTGGGACAATTCAACCAGGGGGCTTGACTCAGCTACAGCT 1020
(321)   S  G  A  P  V  V  C  W  D  N  S  T  R  G  L  D  S  A  T  A  (340)
1021  CTTGAGTTTACCAAGGCGCTGCGAATCGCCTCTGATGTCATGGGAGCTACACAGGCCGTT 1080
(341)   L  E  F  T  K  A  L  R  I  A  S  D  V  M  G  A  T  Q  A  V  (360)
1081  GCCATCTACCAGGCCAGCCAGGCAATCTACGACTTGTTTGACAAAGCTGTTGTTCTCTAC 1140
(361)   A  I  Y  Q  A  S  Q  A  I  Y  D  L  F  D  K  A  V  V  L  Y  (380)
1141  GAAGGGCGCCAGATCTACTACGGTCCCGCTAACTCAGCCAAGAAGTATTTCGAAGACATG 1200
(381)   E  G  R  Q  I  Y  Y  G  P  A  N  S  A  K  K  Y  F  E  D  M  (400)
1201  GGCTGGTATTGCCCGCCCCGCCAGACAAATGGCGATTTCCTCACCTCTATCACCAACCCG 1260
(401)   G  W  Y  C  P  P  R  Q  T  N  G  D  F  L  T  S  I  T  N  P  (420)
1261  ATGGAGCGGAGAGTCAGGGACGGGTTTGAATCCAAAGTTCCCCGAACCGCCCATGAGTTT 1320
(421)   M  E  R  R  V  R  D  G  F  E  S  K  V  P  R  T  A  H  E  F  (440)
1321  GAGACGTATTGGCGCAACTCGCAGCAATTCAAGGACATGCAGGCTGAAATTGAGCAGTGC 1380
(441)   E  T  Y  W  R  N  S  Q  Q  F  K  D  M  Q  A  E  I  E  Q  C  (460)
1381  GAGGATGAACATCCCGTTGGCGGGCCGGCGCTGGGAGAATTGCGCGAGGCCCACAACCAG 1440
(461)   E  D  E  H  P  V  G  G  P  A  L  G  E  L  R  E  A  H  N  Q  (480)
1441  GCGCAGGCAAAGCACGTTCGCCCGAAATCACCGTATACCATAACCATACTTATGCAGGTG 1500
(481)   A  Q  A  K  H  V  R  P  K  S  P  Y  T  I  T  I  L  M  Q  V  (500)
1501  AAGCTGTGTACTACCCGTGCCTACCAGCGCCTGTGGAATGATAAAGCATCGACGATTTCC 1560
(501)   K  L  C  T  T  R  A  Y  Q  R  L  W  N  D  K  A  S  T  I  S  (520)
1561  AGGGTGATGGCGCAACTGATCATGTCTTTGATTATCGGCTCGCTCTACTTTAACACGCCA 1620
(521)   R  V  M  A  Q  L  I  M  S  L  I  I  G  S  L  Y  F  N  T  P  (540)
1621  CAAGTTACCAGCAGTTTTTTCTCGAAGGGTTCGGTTTTGTTCTTTGCTATTCTGCTCAAT 1680
(541)   Q  V  T  S  S  F  F  S  K  G  S  V  L  F  F  A  I  L  L  N  (560)
1681  GCATTGCTTTCTATTTCTGAAATTAATACGtaagtttacttcgctgctccaataaaataa 1740
(561)   A  L  L  S  I  S  E  I  N  T                                 (570)
1741  ataaaataaaataaaaaaaaaaaaaaaaaaaaaaaaaaagtgatggttactgacagaatct 1800
```

```
1801   ttataggctttgtaaggtctttcccaacctggtctttcatcttgcagtcactaacataac      1860

1861   aaagacTCCCAACGCCCTATCGTCTCCAAACACGTATCATACGCCTTATACTATTCCTGC      1920
(571)          S  Q  R  P  I  V  S  K  H  V  S  Y  A  L  Y  Y  S  C      (588)
1921   GTCGAAGCTTTTGCAGGGATCGTTTCCGATATTCCCATAAAGTTGATAACTAGCACTGTC      1980
(589)    V  E  A  F  A  G  I  V  S  D  I  P  I  K  L  I  T  S  T  V      (608)
1981   TTCAACCTAATCATCTACTTCCTCGGAGACTTGCGACGGCAGGCGGATCATTTCTTCATT      2040
(609)    F  N  L  I  I  Y  F  L  G  D  L  R  R  Q  A  D  H  F  F  I      (628)
2041   TTCTTCCTTTTCACTTTCATCACCATGTTAACTATGTCTGCAATATTCCGGACGCTTGCC      2100
(629)    F  F  L  F  T  F  I  T  M  L  T  M  S  A  I  F  R  T  L  A      (648)
2101   GCTGCGACGAAGACAATTTCCCAGGCTCCCGCATTCGCGGGAGTCATGGTTTTGGCTATT      2160
(649)    A  A  T  K  T  I  S  Q  A  P  A  F  A  G  V  M  V  L  A  I      (668)
2161   GTCATATATACGGGTTTCACAATTCAGCGGTCATATATGtaagtgtctctccgcttaatg      2220
(669)    V  I  Y  T  G  F  T  I  Q  R  S  Y  M                        (681)
2221   gttctggtatcttgctcgttttcttgactggggatgttatgctaaaatcccctgtgacag      2280

2281   GCGCCCATGGATGAATGGATAAGCTGGATTAATCCTGTGGCATATGCATTCGAAGCTATC      2340
(682)    A  P  M  D  E  W  I  S  W  I  N  P  V  A  Y  A  F  E  A  I      (701)
2341   CTGGTAAACGAAGTACATAACCAACGCTACGCATGTGCACTGATTGTCCCGCCATATGGA      2400
(702)    L  V  N  E  V  H  N  Q  R  Y  A  C  A  L  I  V  P  P  Y  G      (721)
2401   GAAGGAATGAACTTCCAGTGTCCAATTGCCGGGGCGGTGCCCGGCGAGAGATCGGTGTCC      2460
(722)    E  G  M  N  F  Q  C  P  I  A  G  A  V  P  G  E  R  S  V  S      (741)
2461   GGCGATGCTTGGGTAGAATCGCAATATGGCTACAAGTACTCCCATATCTGGCGGAATCTC      2520
(741)    G  D  A  W  V  E  S  Q  Y  G  Y  K  Y  S  H  I  W  R  N  L      (761)
2521   GGCTTCATCTTCGCATTCCAGGTCTTTTTCTACGTGGTGTATCTTACTGCGACGCAGCTG      2580
(782)    G  F  I  F  A  F  Q  V  F  F  Y  V  V  Y  L  T  A  T  Q  L      (781)
2581   AACACTGCCTCAGCCTCGACCGCAGAATTTCTTGTTTTCCGCCGCGGTAATGTTCCGATA      2640
(782)    N  T  A  S  A  S  T  A  E  F  L  V  F  R  R  G  N  V  P  I      (801)
2641   TATATGCTGAAACAAGACGATGAAGAGAATGGCAAAGCAGCTCCTCCTGCTGCTGCTGCT      2700
(802)    Y  M  L  K  Q  D  D  E  E  N  G  K  A  A  P  P  A  A  A  A      (821)
2701   GCTGCTGCTGCTGCTGCAGCTGCTGAGGCTAATTCCAAGAACGAAGAAGATAAAACCAAC      2760
(822)    A  A  A  A  A  A  A  A  E  A  N  S  K  N  E  E  D  K  T  N      (841)
2761   GTCCTTCCCCCTCAGACTGACGTTTTCACATGGAGGAATGTAACTTATGACATCACGATT      2820
(842)    V  L  P  P  Q  T  D  V  F  T  W  R  N  V  T  Y  D  I  T  I      (861)
```

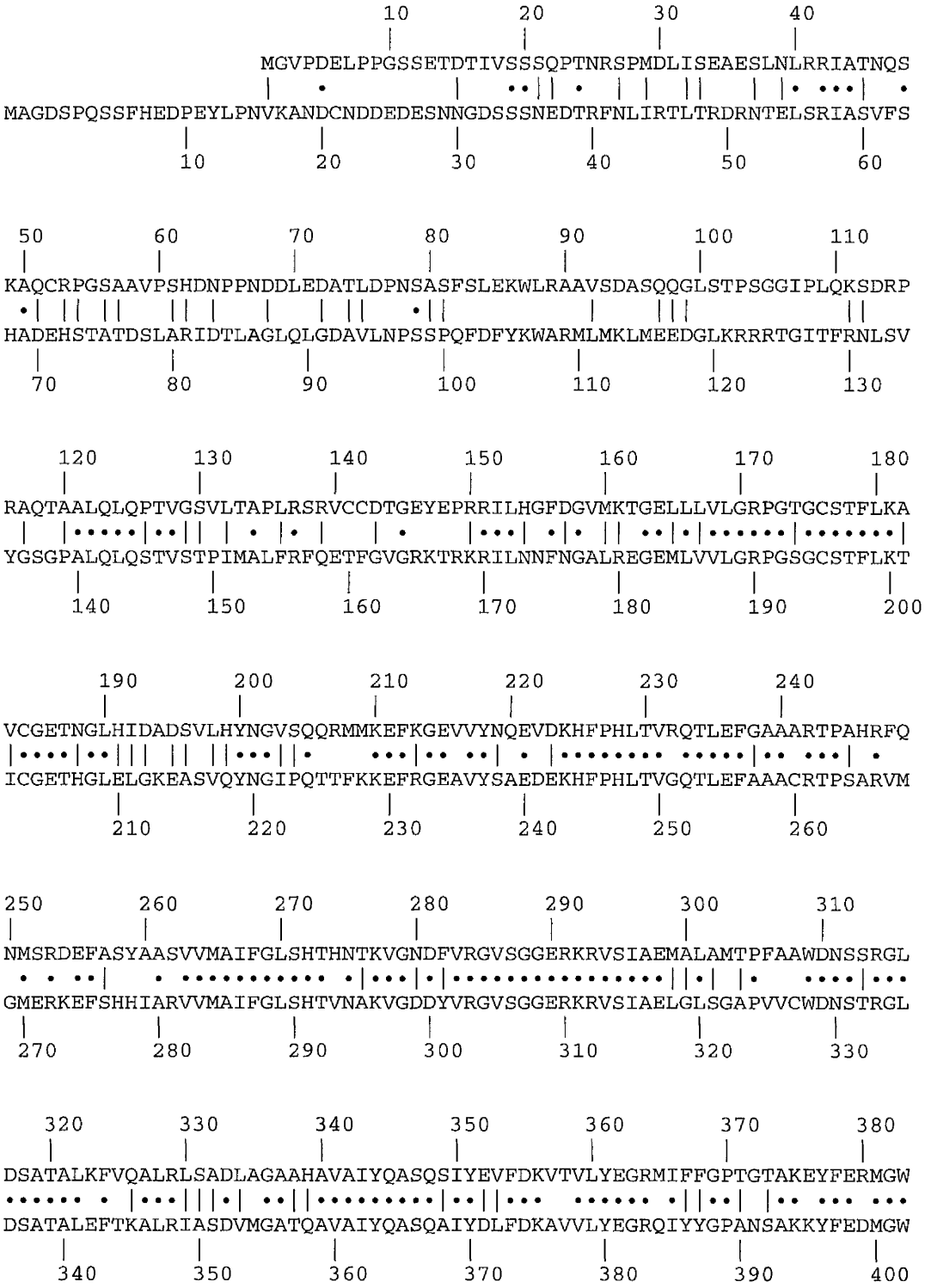
```
2821   AAGGGAGAGGATAGAAGGATCCTTGACCACGTTTCTGGTTGGGTTCGACCCGGTACACTT      2880
(862)    K  G  E  D  R  R  I  L  D  H  V  S  G  W  V  R  P  G  T  L      (881)
2881   ACGGCACTGATGGGCGTCTCTGGAGCCGGCAAAACTACCCTGCTCGACGCGCTCGCGCAA      2940
(882)    T  A  L  M  G  V  S  G  A  G  K  T  T  L  L  D  A  L  A  Q      (901)
2941   CGCATTTCTTTCGGAGTTATCACGGGGGATATGTTTGTCAATGGAAAGCCCCTTGACCTA      3000
(902)    R  I  S  F  G  V  I  T  G  D  M  F  V  N  G  K  P  L  D  L      (921)
3001   TCCTTCCAGCGCAAAACTGGCTACTGTCAACAGCAAGATCTACATCTCGAGACAAGCACG      3060
(922)    S  F  Q  R  K  T  G  Y  C  Q  Q  Q  D  L  H  L  E  T  S  T      (941)
3061   GTACGCGAGGCACTGCGCTTTTCAGCAATGTTAAGACAGCCAAAGAGCGTTTCCAAGCAG      3120
(942)    V  R  E  A  L  R  F  S  A  M  L  R  Q  P  K  S  V  S  K  Q      (961)
3121   GAAAAATACGAGTTTGTTGAAGACGTGATTAAGATGTTGAACATGGAGGATTTTGCGGAA      3180
(962)    E  K  Y  E  F  V  E  D  V  I  K  M  L  N  M  E  D  F  A  E      (981)
3181   GCGGTTGTTGGTAACCCTGGTGAAGGACTCAATGTTGAGCAGCGGAAACTTCTCACCATC      3240
(982)    A  V  V  G  N  P  G  E  G  L  N  V  E  Q  R  K  L  L  T  I      (1001)
3241   GGGGTTGAACTGGCTGCAAAACCGCAACTACTGCTCTTCCTTGACGAACCAACTTCCGGC      3300
(1002)   G  V  E  L  A  A  K  P  Q  L  L  L  F  L  D  E  P  T  S  G      (1021)
3301   CTTGACTCACAAAGTGCATGGGCGATTGTTACTTTTCTCCGGAAGCTTGCTGGCCATGGT      3360
(1022)   L  D  S  Q  S  A  W  A  I  V  T  F  L  R  K  L  A  G  H  G      (1041)
3361   CAAGCTGTGCTCTCCACCATCCACCAGCCCTCCGCA                              3395
(1042)   Q  A  V  L  S  T  I  H  Q  P  S  A                              (1053)
```

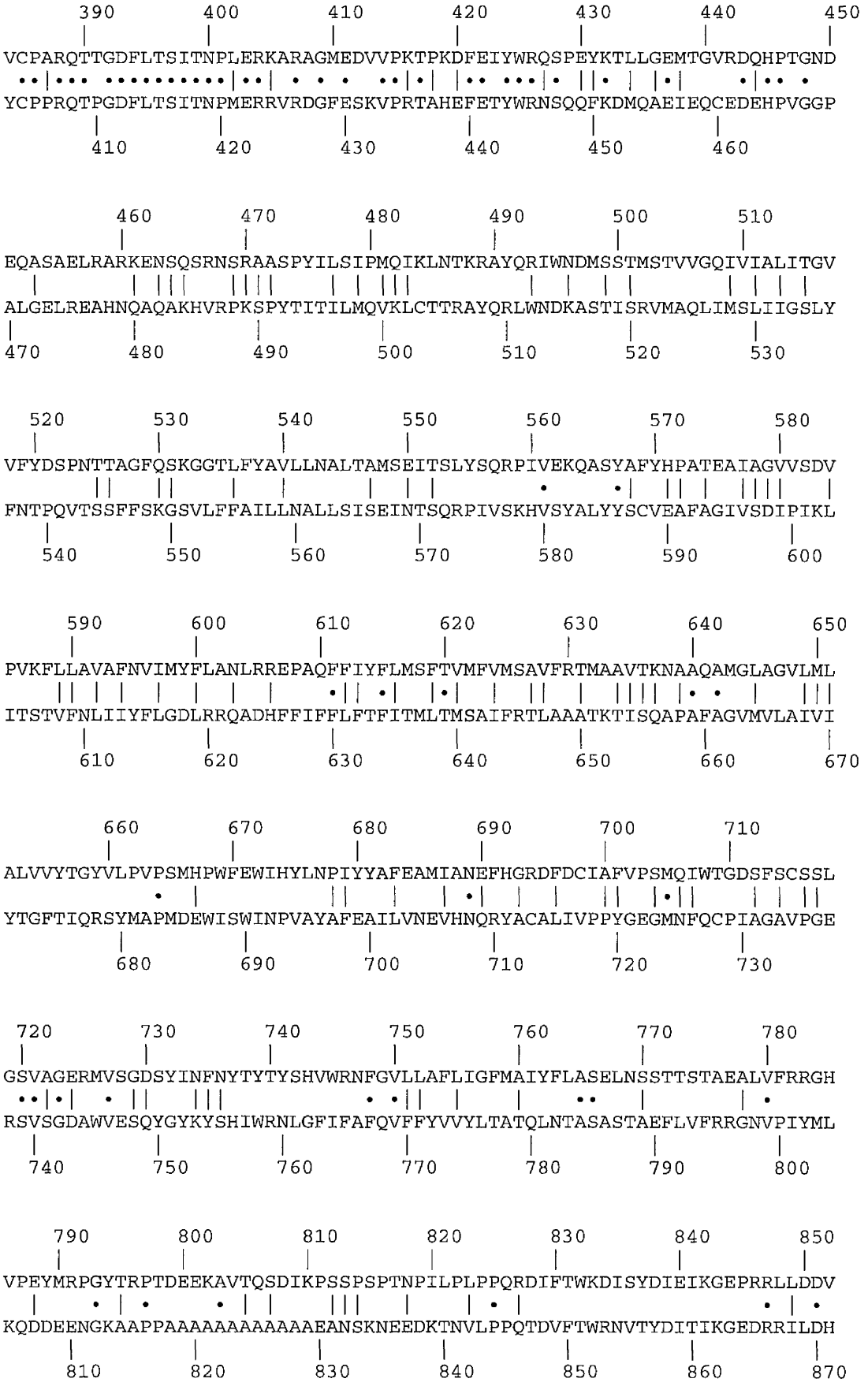
Nomenclature; Nucleotide positions are recorded in black numerals, amino acids in bracketed blue.

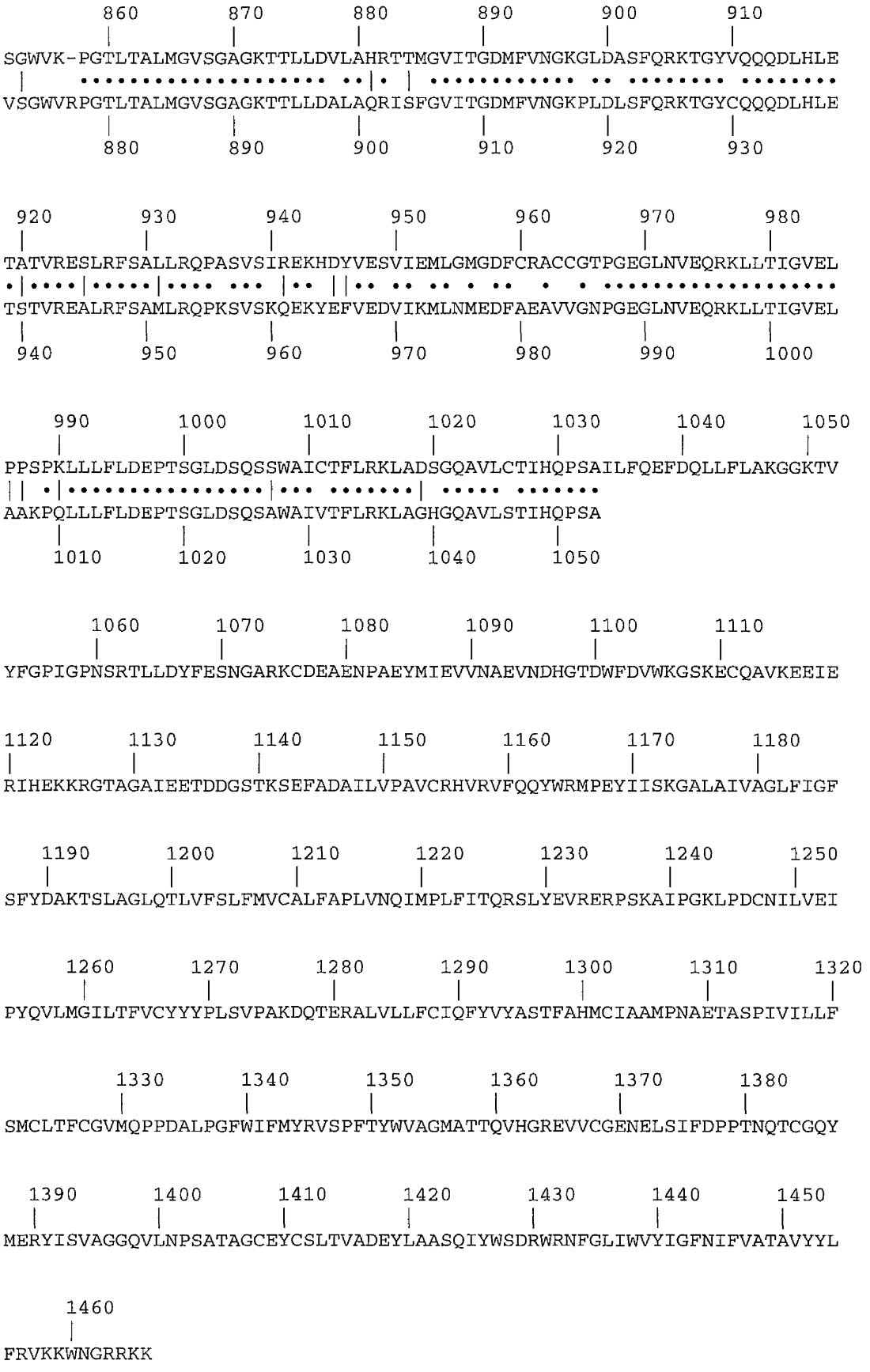
Nucleotide data; Coding nucleotides are denoted by uppercase letter. Non-coding bases are denoted by lowercase letters. Putative introns are presented in blue. Potential promoter features (CAAT and TATA elements) are underlined

Amino acid data; Putative Walker A boxes are denoted in green. The Walker B boxes and centre regions are denoted in red.

Fig4.9 Comparative alignments of AtrA (top) and the partial deduced amino acid sequence of Pft2 (bottom)




390 400 410 420 430 440 450
VCPARQTTGDFLTSITNPLERKARAGMEDVVPKTPKDFEIYWRQSPEYKTLLGEMTGVRDQHPTGND
YCPPRQTPGDFLTSITNPMERRVRDGFESKVPRTAHEFETYWRNSQQFKDMQAEIEQCEDEHPVGGP
410 420 430 440 450 460
460 470 480 490 500 510
EQASAELRARKENSQSRNSRAASPYILSIPMQIKLNTKRAYQRIWNDMSSTMSTVVGQIVIALITGV
ALGELREAHNQAQAKHVRPKSPYTITILMQVKLCTTRAYQRLWNDKASTISRVMAQLIMSLIIGSLY
470 480 490 500 510 520 530
520 530 540 550 560 570 580
VFYDSPNTTAGFQSKGGTLFYAVLLNALTAMSEITSLYSQRPIVEKQASYAFYHPATEAIAGVVSDV
FNTPQVTSSFFSKGSVLFFAILLNALLSISEINTSQRPIVSKHVSYALYYSCVEAFAGIVSDIPIKL
540 550 560 570 580 590 600
590 600 610 620 630 640 650
PVKFLLAVAFNVIMYFLANLRREPAQFFIYFLMSFTVMFVMSAVFRTMAAVTKNAAQAMGLAGVLML
ITSTVFNLIIYFLGDLRRQADHFFIFFLFTFITMLTMSAIFRTLAAATKTISQAPAFAGVMVLAIVI
610 620 630 640 650 660 670
660 670 680 690 700 710
ALVVYTGYVLPVPSMHPWFEWIHYLNPIYYAFEAMIANEFHGRDFDCIAFVPSMQIWTGDSFSCSSL
YTGFTIQRSYMAPMDEWISWINPVAYAFEAILVNEVHNQRYACALIVPPYGEGMNFQCPIAGAVPGE
680 690 700 710 720 730
720 730 740 750 760 770 780
GSVAGERMVSGDSYINFNYTYTYSHVWRNFGVLLAFLIGFMAIYFLASELNSSTTSTAEALVFRRGH
RSVSGDAWVESQYGYKYSHIWRNLGFIFAFQVFFYVVYLTATQLNTASASTAEFLVFRRGNVPIYML
740 750 760 770 780 790 800
790 800 810 820 830 840 850
VPEYMRPGYTRPTDEEKAVTQSDIKPSSPSPTNPILPLPPQRDIFTWKDISYDIEIKGEPRRLLDDV
KQDDEENGKAAPPAAAAAAAAAAAAEANSKNEEDKTNVLPPQTDVFTWRNVTYDITIKGEDRRILDH
810 820 830 840 850 860 870


860 870 880 890 900 910
SGWVK-PGTLTALMGVSGAGKTTLLDVLAHRTTMGVITGDMFVNGKGLDASFQRKTGYVQQQDLHLE
VSGWVRPGTLTALMGVSGAGKTTLLDALAQRISFGVITGDMFVNGKPLDLSFQRKTGYCQQQDLHLE
880 890 900 910 920 930
920 930 940 950 960 970 980
TATVRESLRFSALLRQPASVSIREKHDYVESVIEMLGMGDFCRACCGTPGEGLNVEQRKLLTIGVEL
TSTVREALRFSAMLRQPKSVSKQEKYEFVEDVIKMLNMEDFAEAVVGNPGEGLNVEQRKLLTIGVEL
940 950 960 970 980 990 1000
990 1000 1010 1020 1030 1040 1050
PPSPKLLLFLDEPTSGLDSQSSWAICTFLRKLADSGQAVLCTIHQPSAILFQEFDQLLFLAKGGKTV
AAKPQLLLFLDEPTSGLDSQSAWAIVTFLRKLAGHGQAVLSTIHQPSA
1010 1020 1030 1040 1050
1060 1070 1080 1090 1100 1110
YFGPIGPNSRTLLDYFESNGARKCDEAENPAEYMIEVVNAEVNDHGTDWFDVWKGSKECQAVKEEIE
1120 1130 1140 1150 1160 1170 1180
RIHEKKRGTAGAIEETDDGSTKSEFADAILVPAVCRHVRVFQQYWRMPEYIISKGALAIVAGLFIGF
1190 1200 1210 1220 1230 1240 1250
SFYDAKTSLAGLQTLVFSLFMVCALFAPLVNQIMPLFITQRSLYEVRERPSKAIPGKLPDCNILVEI
1260 1270 1280 1290 1300 1310 1320
PYQVLMGILTFVCYYYPLSVPAKDQTERALVLLFCIQFYVYASTFAHMCIAAMPNAETASPIVILLF
1330 1340 1350 1360 1370 1380
SMCLTFCGVMQPPDALPGFWIFMYRVSPFTYWVAGMATTQVHGREVVCGENELSIFDPPTNQTCGQY
1390 1400 1410 1420 1430 1440 1450
MERYISVAGGQVLNPSATAGCEYCSLTVADEYLAASQIYWSDRWRNFGLIWVYIGFNIFVATAVYYL
1460
FRVKKWNGRRKK

The protein encoding sequence of the *pft2* gene characterised in this study spanned a further 3792bp and was interrupted by two putative introns. The first intron is 156bp in length and located at nucleotide positions 1710-1866 while the second intron is 80bp and is found at nucleotide positions 2200-2280. The partial gene sequence appears to encode 1053 amino acids representing the bulk of a PDR-subfamily ABC transporter. This includes both NBDs and one of the MSDs. It would appear that Pft2 conforms to the NBD-MSD-NBD-MSD arrangement that is typical of these proteins and an inversion of the organisation observed in P-glycoprotein (Higgins, 1992).

The protein displays highest homology with AtrA from *Aspergillus nidulans* (Del Sorbo, *et al.*, 199) but also shows considerable homology to CDR1 (Prasad *et al* 1995) and PDR5. As is common in ABC transporters this homology is most pronounced in the nucleotide-binding domain. The conservation of Walker A and Walker B boxes is most notable in the C-terminal domains, the region targeted in the degenerate PCR. The N-terminal NBDs of PDR-subfamily transporters are peculiar in that the normally conserved lysine residue is degenerated to a cysteine. This remains true in Pft2. The Walker B site in these proteins is only slightly conserved. The precise Walker B of Pft2 is not seen in any other ABC transporter but most of these proteins have largely exclusive sequences at this point. The ABC signature sequence in the N-terminal NBD tends to be semi-conserved but the corresponding motif in Pft2 is identical to that of AtrA. Similarly the C-terminal LNVEQ signature consensus can be degenerated but is conserved between AtrA and Pft2.

The membrane spanning domain of the partial gene sequence displays relatively little homology making the *pft2* gene very distinct in this region. It is difficult to note any distinguishing features in this region as the importance of any given sequence is not well characterised. A conspicuous stretch of 12 consecutive alanine residues is observed at amino acid positions 818-829. The functional consequences of this long hydrophilic stretch of amino acids cannot be surmised without further biochemical investigation but its inclusion in a membrane spanning domain is interesting. It may represent an extracellular or intracellular loop between two helices. Alternatively it

may provide a hydrophilic surface within the pore but this is hard to envisage if we accept the most popular model of 12 amphipathic helices.

It is unfortunate that the complete gene sequence of *pft2* was not achieved in this study. Despite continued and varied approaches the final 3' coding sequence proved recalcitrant to cloning. The missing DNA is expected to encode the C-terminal MSD. Assuming that no introns are present the complete sequence should be attained on the isolation of an approximate additional 600bp (if the homology with other transporters continues). The completion of this sequence is an ongoing concern. Despite the incomplete reading frame, the current data allows access to biochemical studies. For example, the over-expression and purification of soluble NBDs is now possible for Pft2 as both NBDs are present in the above sequence.

Ascribing any function to Pft2 on the basis of sequence homology is difficult. AtrA does not seem to be a drug transporter despite a number of investigations applying a wide range of drug substrates (Alan Andrade; Personal Communication). However this does not preclude such a role for Pft2. The sequence homology with AtrA lies in the NBD and the homology with proven drug transporters is almost as high in these domains. It is possible that the diversity of substrate tolerance in these transporters involves the MSDs. If this is the case then the novelty of the Pft2 MSD sequence prohibits any direct comparison with other proteins. When the full sequence is appreciated this gene should be used in complementation studies in *S.cerevisiae* mutants including strains that are hypersensitive to drugs. This is a preferable approach to knockout studies in *P.brasiliensis* as the organism is difficult to manipulate and the multi-nucleated nature of the cells means the construction of a complete knockout would be very difficult.

Pft2 could enhance the current understanding of fungal ABC transporter processes and potentially the phenomenon of multidrug resistance.

Chapter 5

Expression analysis of *pft1* and *pft2* transcripts at mRNA level.

Analysis of any particular genetic determinant can be performed at a number of levels within the biochemical compartments of the cell. The four main levels currently acknowledged are the genome, transcriptome, proteome and metabolome. Molecular biology concentrates extensively on the first two areas. The field of genomics has produced an enormous volume of information that is evident on examination of the literature base. Such data ranges from small gene fragments to analysis of entire genomes. The previous two chapters have demonstrated the successful application of genome level analysis in the case of *P.brasiliensis*.

In order to further the analysis of *pft1* and *pft2* investigations progressed to the level of the transcriptome, namely the RNA population of the cell. For the purposes of this study the questions addressed by transcript analysis would be fourfold. Firstly, the demonstration of a mRNA corresponding to the *pft* genes would demonstrate that these determinants are indeed expressed in this organism thus confirming the validity of the study. Secondly, it was necessary to confirm the position of the proposed introns in the genomic sequences discussed in Chapters 3 and 4 and to ensure that they are successfully spliced. Also, a comparative analysis of the *P.brasiliensis* mRNA populations purified from cultures grown under different conditions could provide some evidence of transcriptional regulation. Finally, in the case of *pft1*, we investigate the 5' boundary of the transcript, the transcriptional start site, and demonstrate a potential ambiguity in the conceptual translational start site for this and other fungal MDR-like ABC half transporters.

It should also be noted that successful isolation of intron free cDNA is central to construction of recombinant plasmids for heterologous protein expression. Many expression systems, especially those relying on prokaryotic technology, are unable to process introns in unspliced transcripts. Splicing is always inadvisable in expression systems as the process will hinder over expression. Therefore, without cDNA material, the analysis of genes cannot graduate from nucleic acid to protein level.

5.1 Introns and splicing in *pft1*.

Initially it was necessary to prepare RNA from total nucleic acids derived from *P.brasiliensis*. The total nucleic acids were prepared elsewhere by another investigator. The total RNA was prepared by the removal of DNA by DNase. It was decided that the mRNA would be isolated from the total RNA fraction in order to enhance the performance of RT-PCR. The preparation of two mRNA pools was successfully performed using a polyA tail hybridisation/ magnetic separation procedure (see section 2.3.10). Of the two pools of mRNA, one was derived from the yeast form of *P.brasiliensis*, treated with 5µg per ml of amphotericin B while the second was derived from the yeast form treated with 1µg per ml fluconazole. Growth in the presence of drugs was requested to allow further specific investigations described and discussed in this chapter. The yield of mRNA was assessed by spectrophotometry at 260nm and the yield was determined to be 965ng and 846ng per ml for the amphotericin B and fluconazole associated preparations respectively.

The expression and splicing of *pft1* was addressed by RT-PCR using oligonucleotides specific to the *pft1* gene. The mRNA used for this reaction was that of the fluconazole treatment pool as it was expected that expression would be highest in this preparation, assuming that *pft1* was associated with drug resistance

The XS62 reverse primer, responsible for first strand synthesis was designed to overlap the proposed translational stop site at nucleotide position 2626-2645 of the *pft1* sequence (see Fig3.10). The forward primers, XT72 and XT74, correspond to the boundary of the 5' *pft1* sequence possessed at this time and anneal at nucleotide positions 771-787 and 663-683 respectively. The isolation of the final 5' sequence of the gene remained problematic at this stage. However, following the complete sequencing of *pft1* it was evident that these RT-PCRs did indeed span both of the putative introns in the genomic sequence.

As can be observed in Fig5.1, a good yield of cDNA amplicon was observed in the RT-PCR reactions that demonstrated that *pft1* was expressed at mRNA level and that splicing did occur. The amplification of the XT72/XS62 genomic fragment is used as an illustration of the decrease in transcript size. The cDNA amplicon generated by the XT72/XS62 RT-PCR reaction was cloned to pGEM-T Easy™ and completely sequenced in both directions (data not shown). The sequence observed was identical to that obtained for the genomic sequence, except the proposed introns were absent. The splicing had occurred at the exact sites proposed previously. In order to enhance the illustration of splicing, the product of the XT72/XS62 RT-PCR was gel extracted and

Fig5.1 **RT-PCR analysis of *pft1* transcripts to assess expression and splicing.**

Lane1]	XT74/XS62 RT-PCR
Lane2]	XT72/XS62 RT-PCR
Lane3]	XT72/XS62 genomic DNA PCR
Lane4]	XT74/XS62 RT-PCR negative control
Lane5]	XT72/XS62 RT-PCR negative control
Lane6]	MDR1/MDR2 PCR (cDNA product)
Lane7]	MDR1/MDR2 PCR (genomic product)
Lane8]	MDR1/MDR2 PCR (negative control).

RT-PCR conditions:	1 cycle]	48^{0}C for 45min
	40 cycles]	94^{0}C for 30s
		60^{0}C for 1min
		68^{0}C for 2.5min
	1 cycle]	68^{0}C for 7min

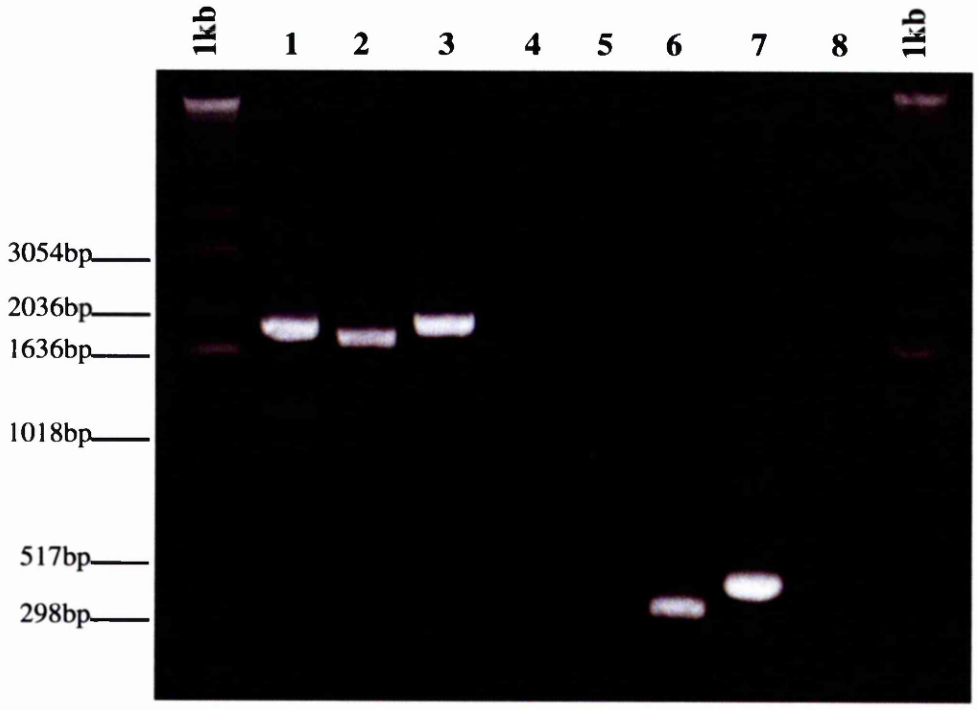


Fig5.1

used in a PCR with the primers MDR1 and MDR2 (see section 3.1). The region of the *pft*1 gene spanned by these primers contains a 77bp intron, thus the expected amplicons from an RT-PCR reaction would be 338bp while the genomic amplicon would run at 415bp. The size difference shown for these amplicons better illustrates that splicing has occurred.

5.2 RT-PCR expression analysis of *pft1* and *pft2* under conditions of drug treatment

An upregulation or complete induction could if observed for *pft1* and *pft2* could provide an initial indication of the association of these genes with hydrophobic drug transport. The preparation of mRNAs from amphoterecin B and fluconazole treated populations could allow an initial assessment of this phenomenon. RT-PCR was again used. The *pft*1 RT-PCR, using primers XT72 and XS62 (see Fig5.1) would be employed to analyse transcriptional regulation.

As a positive control for the integrity of the RNAs and their competence as templates for RT-PCR, an additional reaction was performed for each RNA pool. This control targeted the *hsp70* transcript of *P.brasiliensis*. The trancriptional profile of this gene had been investigated previously and had illustrated the constitutive expression of a completely spliced form of this mRNA in the yeast form. Two *hsp70* oligonucleotides (HSPfwd and HSPrev) were designed to target a 3' intron spanning sequence of the heat shock protein gene (Da Silva, *et al.*, 1999). A successful reaction on this template would result in a 250bp cDNA amplicon (or a 412bp amplicon if using a genomic DNA template). As a PCR positive control, and differential size reference, the XS72/62 PCR was performed on *P.brasiliensis* genomic DNA.

The results of this RT-PCR are presented in Fig5.2. The amplicon was isolated only under conditions of fluconazole treatment and not in the template from amphoterecin B treated cells. Fluconazole is a well known substrate for many fungal ABC drug transporters while amphoterecin B is not an MDR substrate (Bolhuis, *et al.*, 1997). This dramatic and specific induction represents the first indication that a fungal MDR-like ABC half-transporter may be involved in resistance to a hydrophobic drug compound.

A similar RT-PCR was performed to analyse expression of *pft2*. In this case primers P2RTfwd and P2RTrev were used, targeting nucleotide positions 415-435 and 1983-2003 of the *pft*2 gene respectively. This region is spans the first 152bp putative intron

Fig5.2 Azole dependant induction of *pft1* expression.

Lane1]	XT72/XS62 RT-PCR (Amphoterecin B mRNA pool)
Lane2]	Hsp70Fwd/Hsp70Rev (Amphoterecin B mRNA pool)
Lane3]	XT72/XS62 RT-PCR (Fluconazole mRNA pool)
Lane4]	Hsp70Fwd/Hsp70Rev (Fluconazole mRNA pool)
Lane5]	XT72/XS62 PCR (Genomic DNA)

RT-PCR conditions (Lanes 1-4); RT-PCR conditions applied were those optimal for XT72/XS62 amplification and as noted in Fig5.1.

PCR conditions (Lane 5); The PCR on genomic DNA was performed under the same reaction conditionsas the RT-PCRs except the Afl Reverse Transcriptase was omitted. The PCR reaction was added to the thermocycler on termination of the first strand synthesis step of the RT-PCRs.

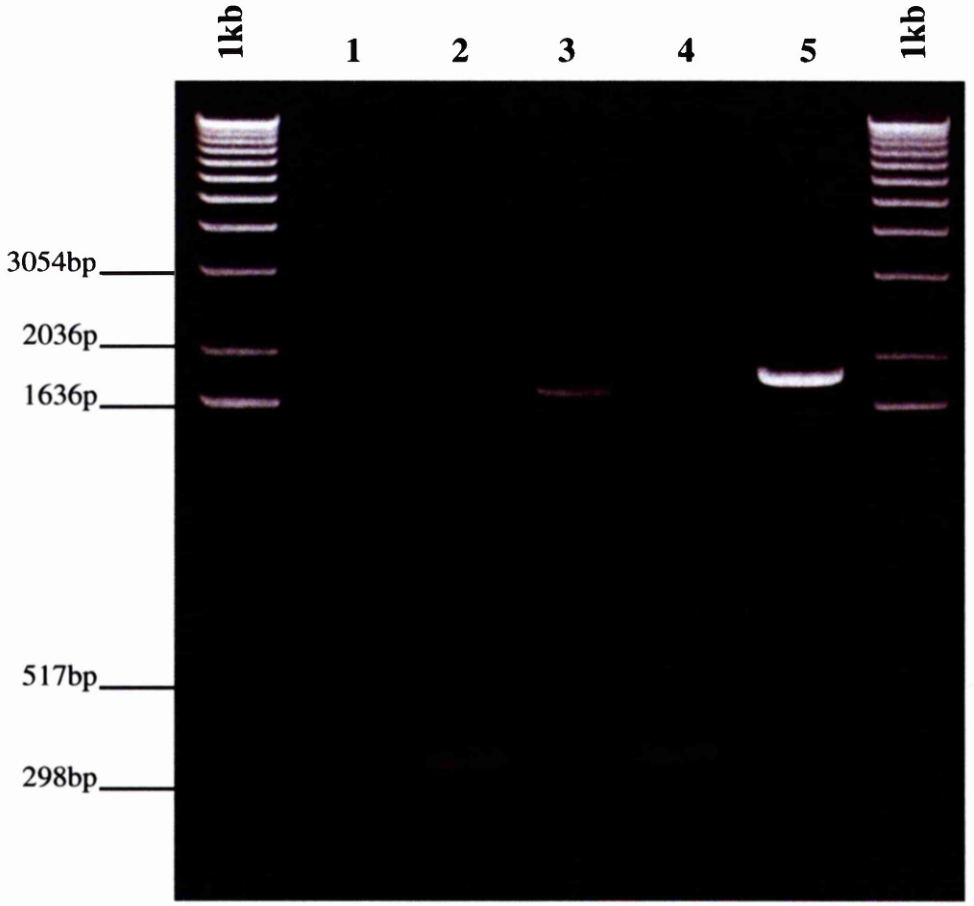


Fig5.2

producing a 1588bp product from genomic DNA and a 1432bp cDNA from spliced mRNA. No specific regulation was observed in this case (Fig5.3) and indeed no particular difference can be observed in the intensity of yield from either RT-PCR, (identical template quantities being used in each case). It must be concluded therefore that the *pft2* gene is not expressed in an azole specific manner. No *hsp70* positive controls were employed in this experiment because the integrity of both RNA pools was demonstrated to be good in the previous *pft1* experiment.

5.3 SMART™ RACE RT-PCR analysis of the 5' terminus of the *pft*1 transcript

As previously mentioned in Chapter 3, the final completion of the *pft1* gene sequence proved problematic due to a truncation of the genomic open reading frame during library construction. In order to solve this problem both GenomeWalker™ SSP-PCR (see section 3.4) and 5' SMART™ RACE RT-PCR would be used. It was decided to include 5' RACE RT-PCR as a strategy to complete the sequence because the apparent robust nature of the *pft1* mRNA observed to this point indicated it was a receptive target for molecular applications. The results for the 5'RACE experiment were superseded by those generated by the SSP-PCR with respect to extending the *pft1* sequence. However, the 5' RACE data did provide a serendipitous discovery relating to an unexpected ambiguity between the expected and observed transcript length.

This particular 5' RACE technique makes use of MMLV reverse trancriptase. This enzyme adds 3 to 5 C nucleotides to the 3' end of the first cDNA strand in a non-template directed manner (similar to the addition of a single A nucleotide to 3' DNA amplicon termini by Taq polymerase). These 3 to 5 C nucleotides are used as a target for the annealing of a SMART™ adapter to the 3' end of the transcript which can serve as a target for a forward facing primer (see Appendix II.). The addition of the C nucleotides by the MMLV RT requires successful termination of reverse strand synthesis and theoretically alleviates the risk of amplification from partial reverse strands.

The 5' RACE RT-PCR was performed using a nested approach using two *pft1* specific primers (PT72 and PT7) and two adapter specific primers (AP1 and AP2). The results of both the primary and secondary nested PCRs show successful and unambiguous products in both cases (fig 5.4a & b). It should be noted that the result presented here was achieved after extensive PCR optimisation with respect to template concentration,

Fig5.3 Induction of *pft2* expression.

Lane1]	RT-PCR (Amphoterecin B mRNA pool)
Lane2]	RT-PCR (Fluconazole mRNA pool)
Lane3]	PCR (Genomic DNA)

RT-PCR conditions:	1 cycle]	48^0C for 45min
	40 cycles]	94^0C for 30s
		65^0C for 1min
		68^0C for 2min
	1 cycle]	68^0C for 7min

PCR conditions (Lane 3); As above but the first 48^0C cycle was omitted. The PCR reaction was added to the thermocycler on termination of the first strand synthesis step of the RT-PCRs.

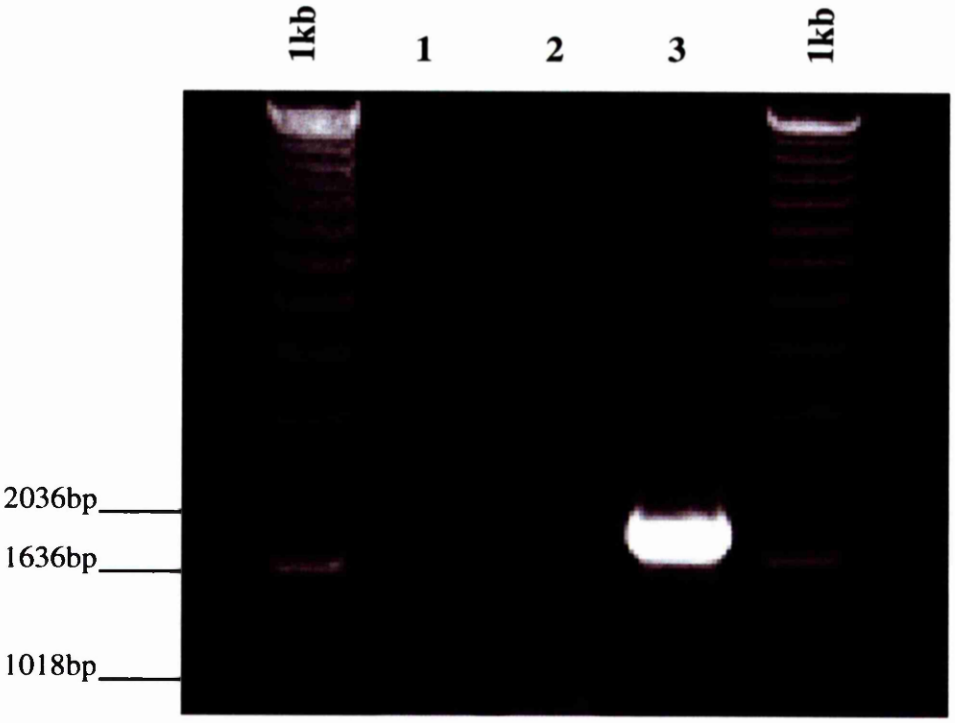


Fig5.3

Fig5.4 5' SMART™ RACE RT-PCR amplification of *pft1* to determine the 5' transcript boundary

a) PT72 primary PCR on *P.brasiliensis* cDNA

b) PT7 secondary PCR on *P.brasiliensis* cDNA

PCR conditions are as noted in Chapter 2, section 2.4.6

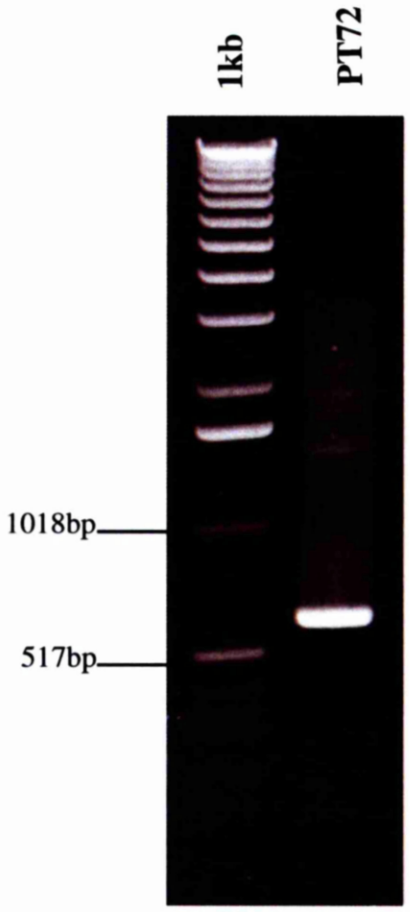


Fig5.4a

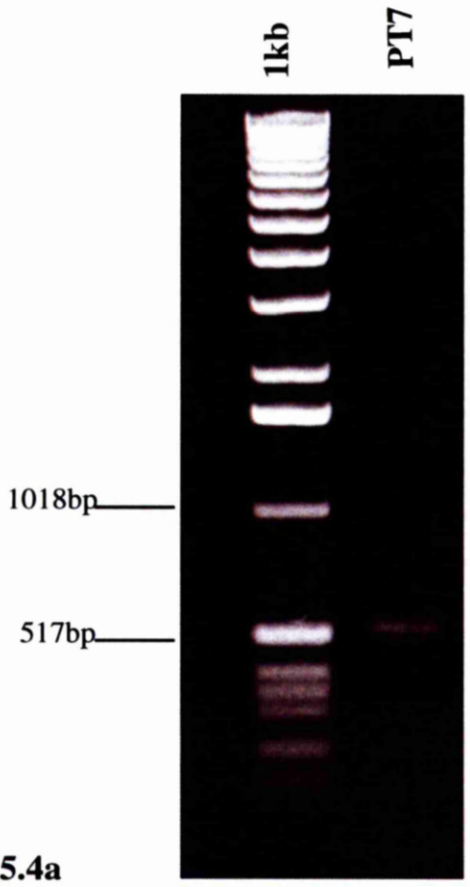


Fig5.4b

annealing temperature and $MgCl_2$ concentration. The 5' RACE RT-PCR reactions performed using alternative downstream reverse primers, which would have generated larger amplicons, did not seem to perform well and no single product could be generated, even in a nested strategy (data not shown).

Despite the difficulty in initial PCR optimisation, this 5'RACE RT-PCR was entirely reproducible under these optimal conditions. The primary (PT72/AP1) fragment was cloned to pGEM-T Easy™ and sequenced in both directions. The DNA was confirmed as originating from the *pft1* transcript. Inspection of the ends of the amplicon readily allowed the identification of the SMART™- adapter and the 5 C nucleotides that allowed annealing. However, translation of the coding sequence of the fragment revealed no new ATG start codon. Critical analysis of the sequence data and the sequencing of a number of clones from replicate PCRs continued to confirm the initial observed sequence as being correct. Application to RT-PCR of the primer Pft1pro (targeting the promoter of the *pft1* genomic sequence) in conjunction with the PT72 reverse failed to generate a product (data not shown).

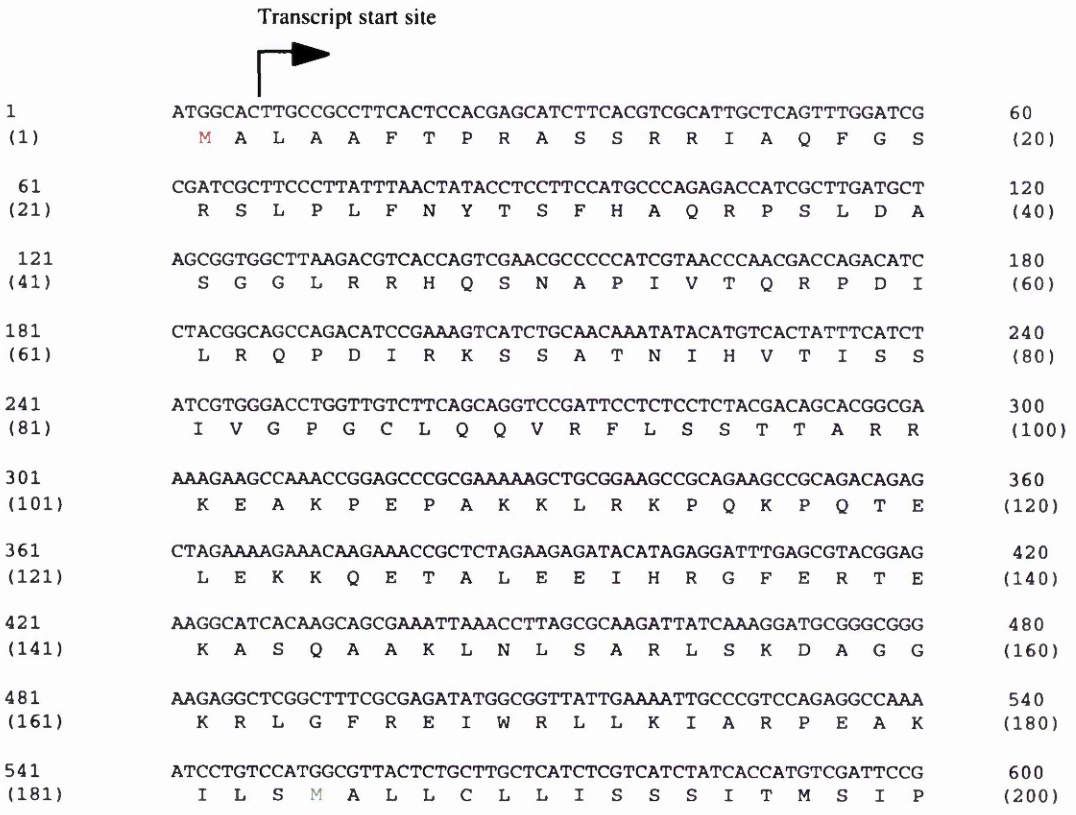
The conclusion had to be that the *pft1* mRNA in the fluconazole pool did not exhibit an ATG start codon in a position that would be consistent with the conceptual start codon of other such fungal half transporters. The actual start of the mRNA transcript shows that the RNA polymerase initiation site was 6bp downstream from the conceptual *pft1* ATG as suggested by the GenomeWalker™ SSP-PCR. Translation from this transcript would implicate the ATG at nucleotide position 550 as the initiating codon (Fig5.5).

5.4 Full length cDNA for *pft1*

The recombinant expression of *pft1* would require complete length cDNA transcripts to be generated for incorporation into an appropriate expression vector. Previous investigators have reported a difficulty in the generation of complete transcripts for ABC transporter genes and many have resorted to a protracted method of partial amplifications and fusion PCRs to piece together the necessary DNA.

In the case of *pft1* there were two potential boundaries to the open reading frame, based on the comparative data from the genomic sequence versus the 5'RACE RT-PCR sequence. It was decided to amplify the cDNA relating to both sites. It is fortunate that the *pft1* transcript in the fluconazole mRNA pool terminates just 6bp downstream from

Fig5.5 A comparison of the ATG start sites of *pft1* based on the conceptual and the experimental positions. The arrow marks the first codon (CTT) of the transcript demonstrated by 5'RACE. The initiating methionine based on the genomic DNA sequence is in red while the ATG suggested by 5'RACE is in green.



the genomic DNA conceptual start codon. A forward primer, including the conceptual ATG of the gene, would still anneal to this site as only a 6bp 5' extension is required. This primer extension is a standard practice and is often used to introduce restriction sites to the termini of amplicons to facilitate downstream cloning.

Successful amplification of both cDNAs was evident (fig5.6). These products were cloned in pGEM-T Easy™ in case they were required for further experiments.

5.5 Discussion

The transcription of *pft1* and *pft2* is assured in *P.brasiliensis.* The presence of these transcripts argues against the possibility that the *pft* genes represent pseudogene sequences and suggests the loci can be functionally active in the yeast cells of this organism. It is also notable that the transcript is fully spliced. Of the small amount of information available on the regulatory process employed by *P.brasiliensis*, the retardation of the splicing process has been implicated. The expression of the Hsp70 protein is specific to the yeast form of the organism. However, a *hsp70* transcript can be detected in the mycelial form (Da Silva *et al.*, 1999). This mycelial transcript remains unspliced thus preventing translation and protein synthesis. It is evident that this does not occur under conditions of yeast growth and drug treatment with respect to the *pft1* and *pft2* transcripts.

The analysis of mRNA populations in response to environmental stimuli is a common and productive technique. It has been employed to good effect in the case of molecular studies on the expression of fungal ABC transporters involved in drug transport often by northern blot. The *atr* genes of *A.nidulans* (Del Sorbo, *et al.*, 1997), the *pdr5* gene of *S.cerevisae* (Bissinger and Kuchler 1994) and the *cdr1* transcript of *C.albicans* (Prasad, *et al.*, 1995) have all been shown to increase in abundance in direct response to various cytotoxic drugs using northern blot analysis. Such induction has been shown to be specific, for instance the *atrA* gene is upregulated by the presence of imazalil but not particularly with cyclohexamide.

For the purposes of the *pft1* study, RT-PCR was substituted for Northern blot. This provided the added advantage of enhanced sensitivity and a requirement for less material, an important aspect considering the logistical difficulties experienced in acquiring *P.brasiliensis* nucleic acids from external sources. The choice of this

Fig5.6 RT-PCR generating cDNA amplicons from the conceptual and experimental ATG positions of *pft1* (see Fig 5.5).

Lane 1] ATG position 1
Lane 2] ATG position 1 (no template negative control)
Lane 3] ATG position 550
Lane 4] ATG position 550 (no template negative control)

RT-PCR conditions:	1 cycle]	48^0C for 45min
	40 cycles]	94^0C for 30s
		56^0C for 1min
		68^0C for 3min
	1 cycle]	68^0C for 7min

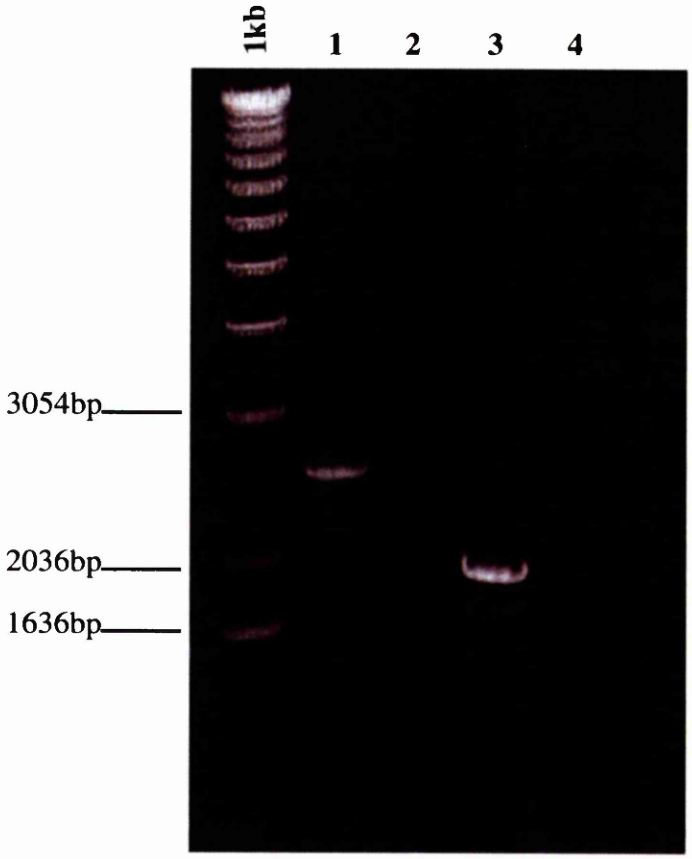


Fig 5.6

technique allowed additional benefits. For indicating complete induction of a gene, like that demonstrated for *pft1*, northern blot is less conclusive. In a situation where low levels of transcript are present, their abundance may lie below the detection limit of the hybridisation procedures employed. Northern analysis can conclude that a transcript lies below the level of detection but not that it is necessarily absent. In a PCR based technique it is theoretically possible to detect of a single copy thus enhancing specificity.

The differential transcription of *pft1*, illustrating fluconazole dependant induction, serves as a preliminary indicator that this gene may be involved, in some way, with azole drug resistance. In keeping with the lack of biochemical characterisation, the dynamics of expression of the MDR-like fungal half transporters has remained unreported to this point. Further analysis of mRNAs from a more extensive range of growth conditions may provide further information regarding *pft1*. Unfortunately, the preparation of these mRNAs was not possible in this study due to the lack of facilities. It is very likely that other azoles, such as ketoconazole (the therapeutic drug of choice for PCM) could elicit a similar profile.

The expression of the transcript of *pft2* did not show an azole specific induction. Indeed, although this RT-PCR was not designed to be particularly quantitative, the yield of product between amphoterecin B and fluconazole reactions doesn't suggest upregulation in the presence of the azole. However, this does not dispel the notion that Pft2 could be involved in drug transport. Many proteins involved in response to potentially damaging environmental stimuli are transcriptionally activated from stress response elements (STREs) (Katzmann, *et al.*, 1996). It may be that the exposure of the organism to amphoterecin B is sufficient to trigger this response. Again, this data can only be considered preliminary and the opportunity to extend this aspect of study to a more extensive array of mRNA populations should be encouraged.

As a contemporary route of investigation to the GenomeWalker™ SSP-PCR we employed 5' RACE RT-PCR, both techniques aiming to complete the 5' sequence of *pft1*. Indeed the 5'RACE RT-PCR reactions have been used by a number of investigators to determine the definitive open reading frame for genes of interest. There are a number of cases where the completion of a 5' sequence during initial gene cloning and characterisation has been achieved solely using 5'RACE RT-PCR and the results of those PCRs were considered definitive. Recent notable examples include the human homologue of the yeast PheRS tRNA synthetase subunit (Zhou, *et al.*, 1999),the gene encoding rat homologue of the connective tissue growth factor (CTGF) (Xu, *et al.*, 2000), the nuclease of the fungus *Syncephalastrum racemosum* (Ho and Liao, 1999) and the *cbhI* and *cbhII* genes for cellulose utilisation by the straw mushroom *Vovariella*

volvacea (Jia, *et al.*, 1999). Scenarios which prompted the use of RACE RT-PCR in these investigations included the completion of ORFs for which an apparent partial sequence was possessed.

The definition of the dimensions of the *pft1* gene were extensively influenced by those of homologues. Those homologues have not been studied at transcript level and all data is therefore putative. The 5' SMART™ RACE RT-PCR data poses a number of significant issues regarding the transcript boundaries of the fungal half-transporter genes. It should be noted that all definitions of the dimensions of these genes has been generated from DNA sequences rather than RNA. In essence the 5' transcriptional start is often identified as the first ATG codon following an extended stretch of apparently non-coding DNA that exhibits stop codons in the appropriate reading frame. It remains possible that these conceptual start codons can be mis-identified. The theory relies of the fact that stop-codons will spontaneously arise in regions of non-coding DNA because the selective pressure required for sequence conservation is absent. However, the lack of selective pressure *per se* does not select for stop codons, it merely doesn't select against them. The 5' RACE data presented for *pft1* would suggest that, in this pool of mRNA at least, any protein arising from this transcript will begin at a later ATG. It is interesting that the region of conservation between these genes becomes apparent around the proposed alternative ATG rather than the conceptual one. It has to be acknowledge that some conservation remains upstream from this position but that may persist as an evolutionary artefact.

Recently, a similar observation was made for the *C.albicans tpk2* gene, which encodes a protein kinase A involved in the regulation of dimorphism. The unconventional electrophoretic mobility of the Tpk2 protein prompted an investigation of the 5' ends of the associated transcript to clarify the precise ATG used. It was found that the correct ATG was positioned 33bp upstream from the conceptual start codon which explained the apparent difference in the expected size of the protein. The proposed Tpk2 protein was 422 amino acids while the correct length is 411 (Sonneborn, *et al.*, 2000). The potential error in *pft1* is even greater with the second candidate ATG positioned at position nucleotide 550 with respect to the conceptual start codon (a difference of 183 amino acids).

A second hypothesis could lie with the use of alternative translational start sites for the same core protein under different growth conditions. Differential transcription, leading to the expression of different isoforms has been reported, particularly in higher organisms. Such differential transcription can arise from alternative transcriptional initiation sites or from differential splicing of pre-mRNAs. The murine *Girk2* transcript

exists as five transcript isoforms each one proposed to encode for a K^+ channel with slightly differing function (Wei, *et al.*, 1998). Likewise the Na,K-ATPase mRNA expressed in the rat myocardium imposes the possibility of two different transcriptional start sites and three different 3' termini arising through 3 different poly-adenylation sites (Shao, *et al.* 1999). If Pft1 is a multifunctional protein then this multifunction could originate from differential transcription and the exclusion of the first 183 amino acids provided by the rest of the 5' sequence.

A final possibility is that this data indicates a negative regulation effect on the expression of Pft1. The origin of the transcript at position +6 in the *pft1* gene removes the conceptual ATG, but it would also remove the ribosome binding site associated with the gene. If a splicing event has occurred which removes this ribosome binding site, translation will effectively be blocked. It remains a possibility that a specific splicing event occurs within the *pft*1 transcript in order to prevent the progression to the translation process. This is a wasteful notion on the part of the organism. Not only does the organism exhibit a direct induction of *pft*1 in response to drug but it also successfully removes the introns towards the 3'end of the transcript. This is a lot of physiological effort to extend to a transcript with no intended function. However, *P.brasiliensis* persists in expressing unspliced forms of the *hsp*70 transcript in the mycelial form, prohibiting protein expression, so the organism is not adverse to this apparently wasteful mechanism of post-transcriptional regulation (Da Silva, *et al.*, 1999).

The generation of a full length cDNA transcript, originating at either ATG, was achieved for *pft1*. This success was in keeping with the excellent performance of this mRNA for other applications and would allow progression to recombinant protein expression.

Chapter 6

Over-expression of Pft1 domains and the purification of the Pft1 nucleotide-binding domain

The successful acquisition of the spliced cDNA of *pft1* allowed the introduction of this gene to a heterologous protein expression system. Any recombinant protein obtained would be purified and examined on biochemical investigation. A number of expression constructs were produced, allowing successful purification of the Pft1 nucleotide-binding domain for biochemical studies.

Construction of expression systems was achieved in a 2 step process. Firstly expression cDNAs were amplified by RT-PCR and cloned to pGEM-T Easy™. The expression fragments were then excised from the pGEM-T Easy™ using *Bam*HI and *Xho*I restriction enzymes. These enzymes acted at restriction sites at the termini of the expression fragment inserts. The sites were incorporated into the RT-PCR amplicons by inclusion of the appropriate sequence in the PCR primers. The restriction fragments were then cloned into the multiple cloning site of appropriately prepared pET21b vector. The scheme is presented in overview in Fig6.1.

It was decided to use the pET system of recombinant expression vectors, in particular pET21b. This vector allows the tightly regulated induction of cloned ORFs, to very high level. In addition, the plasmid affords the fusion of a hexa-histidine (His_6) tag at the C-terminus of the protein. The His_6 tag exhibits high affinity for the nickel anion allowing efficient and rapid purification of protein on a nickel-agarose matrix (ion-affinity chromatography). A number of *E.coli* strains have been generated in which the genome has been optimised to enhance the efficiency of the recombinant expression process. Many of these are available commercially and we used a derivative of BL21 (DE3), *Epicurian coli* ® BL21-CodonPlus™ (DE3)-RIL cells (Stratagene), the benefits of which will be explained below.

The recombinant over-expression and purification of membrane proteins is widely acknowledged to be a very difficult practice. Even homologous over-expression of *E.coli* membrane proteins in *E.coli* can often be difficult. In order to express recombinant *Emr*B, the membrane component of an *E.coli* transporter the inefficient practice of *in vitro* translation had to be employed (Lomovskaya & Lewis, 1992).

Fig 6.1 Strategy for the construction of expression plasmids

The construction of pET21b expression systems was achieved in a two stage cloning procedure. The necessary coding cDNA was obtained by RT-PCR and A-T cloned to pGEM-T Easy™. *Bam*HI and *Xho*I digestion, at sites incorporated by the PCR primers, allowed the isolation of expression fragments for cloning into the corresponding sites in the MCS of pET21b.

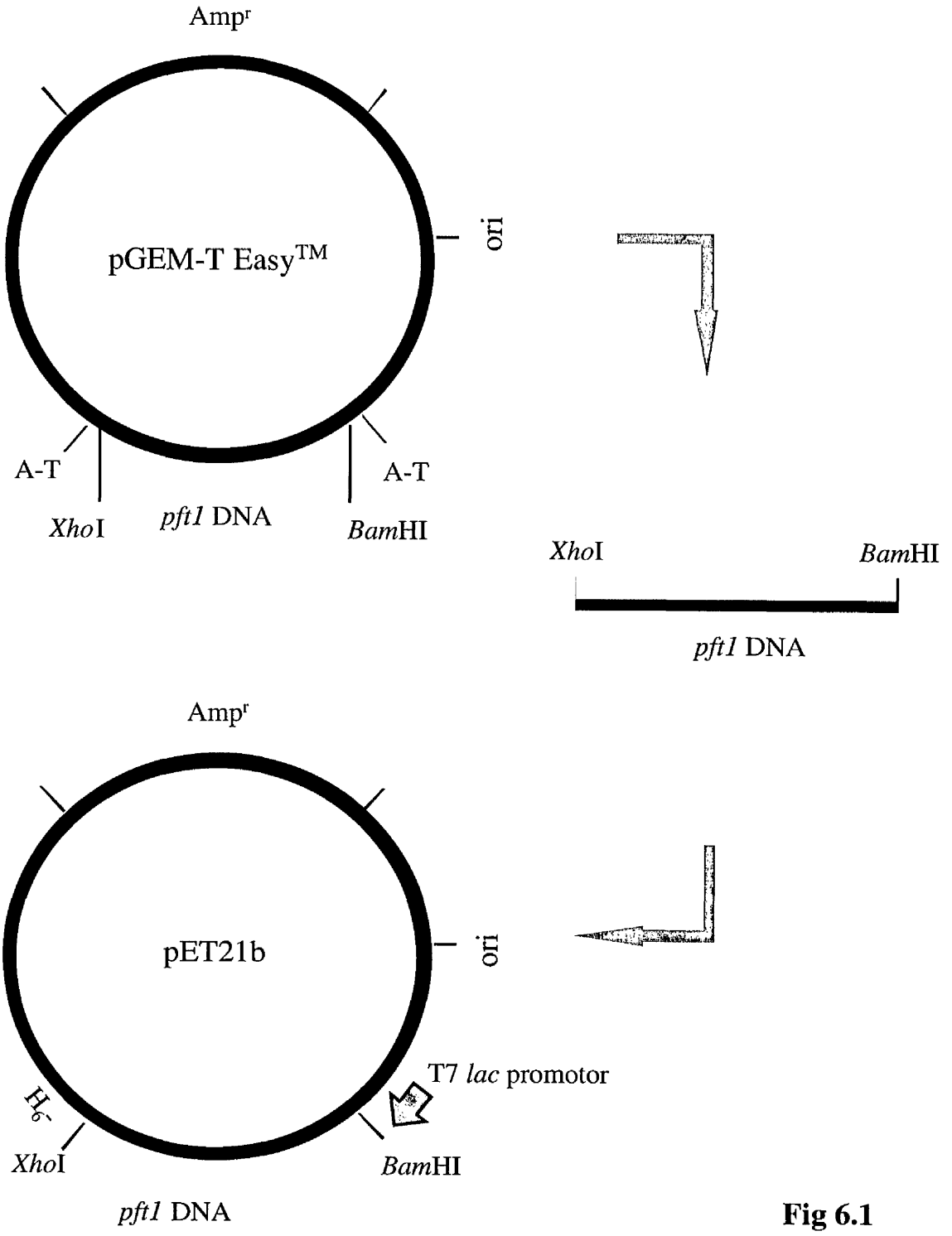

Amp^r
pGEM-T Easy™
ori
A-T
A-T
XhoI
pft1 DNA
BamHI
XhoI
BamHI
pft1 DNA
Amp^r
pET21b
ori
T7 lac promotor
H6
XhoI
BamHI
pft1 DNA

Fig 6.1

Successful over-expression of membrane proteins in prokaryotic systems is often the result of extensive optimisations that can take years to perfect.

It is true that the accumulation of pure membrane proteins is critical to the long-term evaluation of the mechanism of ABC transporters in terms of both biochemistry and structural biology. However, it is important that these difficulties should not retard the progression of research on ABC transporters. As an alternative avenue of introduction to the study of these proteins a number of researchers have reported successful expression of the nucleotide-binding domains of ABC transporters in the cytoplasmic fraction of the *E.coli* cell. These domains tend to exhibit greater solubility and are more accessible to current recombinant systems. Of particular note are the successful investigations of the substrate binding capacity and processivity of the murine *mdr1* N-terminal NBD (Dayan *et al.*, 1996) and C-terminal NBD (Conseil *et al.*, 1998). These two investigations illustrated not only the successful acquisition of a soluble NBD but also that these proteins could be successfully used to study the binding of ATP and other substrates. The functional and structural aspects of NBDs have been previously reviewed (Sharom *et al.*, 1999) and are discussed in detail in Chapter 1.

This chapter reports the over-expression of the Pft1 nucleotide-binding domain. The design of constructs were based on the requirement for soluble, hydrophilic molecules, that could be readily purified and applied to downstream fluorescence based biochemical experiments. As in the aforementioned investigations, the production of a suitable soluble molecule was achieved by partially empirical means.

6.1 Initial construction of a pET21b plasmid that would allow the over-expression of Pft1 NBD

In order to generate the appropriate cDNA for expression it was first necessary to define the boundaries of the Pft1 domains. In the absence of any structural information for this protein the junction was estimated on the basis of a Kyte-Doolittle hydropathy plot (Fig6.2a). As mentioned in Chapter 3 this plot is ambiguous in relation to the position of membrane spanning or associated regions. It was decided to define the nucleotide-binding domain as being the C-terminal portion of the protein. The junction of the MSD and NBD domains was estimated to be at the position where the flux in polarity of hydropathy was markedly reduced, as indicated in the figure. The NBD portion was estimated to span from amino acid at position 398 to the terminal amino acid of the protein at position 827.

Fig6.2a **Kyte-Doolitle Hydropathy plot of the amino-acid sequence of Pft1** (generated using GenejockeyII).

Noted amino acid positions represent the boundaries of the expression constructs prepared from Pft1NBD.

1]	Pft1NBD N-terminus	(amino acid number 398)
2]	Pft1NBDT and Pft1NBDTW N-termini	(amino acid number 516)
3]	Pft1NBDT C-terminus	(amino acid number 634)
4]	Pft1NBDTW C-terminus	(amino acid number 756)
5]	Pft1NBD C-terminus	(amino-acid number 827)

Fig6.2b **Restriction enzyme analysis of pET21bNBD clones.**

The *Bam*HI/*Xho*I Pft1NBD expression fragment was ligated at 16^0C to appropriately prepared pET21b at 16^0C, overnight and then transformed to Novablue competent cells.

Ten potential pET21bPft1NBD recombinants were analysed using *Bam*HI and *Xho*I restriction digestion (37^0C for 2h in Promega multi-core buffer). All recombinant plasmids exhibit Pft1NBD inserts.

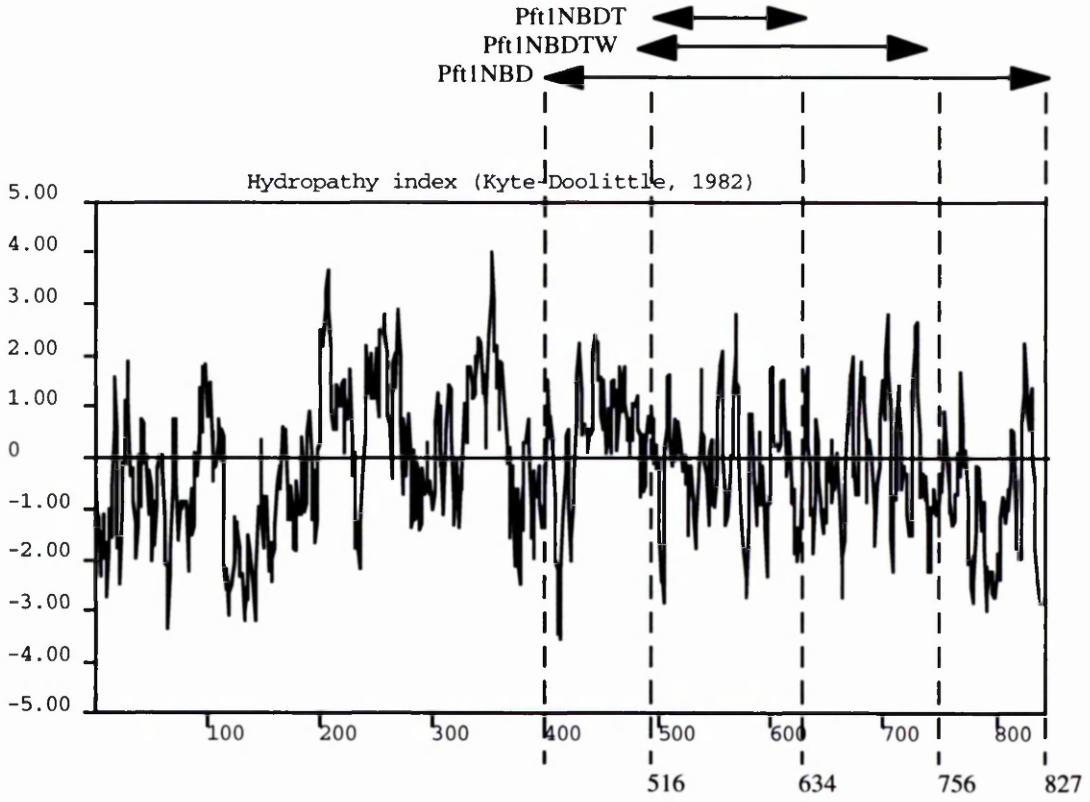


Fig6.2a

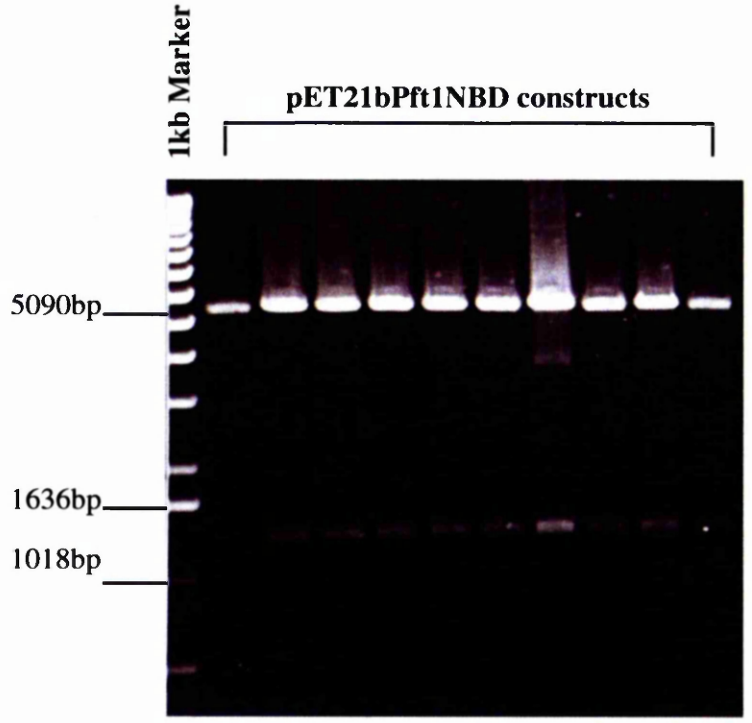


Fig6.2b

Generation of Pft1NBD cDNA for expression was achieved using the primers Pft1NBDfwd and Pft1NBDrev. The forward and reverse primers were designed so as to introduce terminal *Bam*HI and *Xho*I sites respectively. Downstream digestion of these sites with *Bam*HI and *Xho*I (which do not cut the *pft1* gene) would facilitate cloning into the MCS of pET21b. These primers were used in RT-PCR and the required cDNA insert was produced (data not shown). This Pft1NBD was cloned into pGEM-T Easy™ (data not shown). The insert was released from this clone by digestion and ligated to appropriately cut pET21b. The clone pET21Pft1NBD was consequently created (Fig6.2b).

6.2 Over-expression from pET21Pft1NBD

The pET21bPft1NBD expression plasmid was transformed into *Epicurian coli* ® BL21-CodonPlus™ (DE3)-RIL cells. These cells are of the same phenotype as the BL21 (DE3) cells commonly used for recombinant expression but an additional plasmid, based on a pACYC construct, is maintained on a chloramphenicol resistance marker within the cell. This plasmid promotes the expression of three recombinant tRNAs, *argU, ileY* and *leuW*, which although rare in *E.coli*, are optimal for the translation of a number of eukaryotic genes including Pft1.

Expression from the T7 promoter of pET21b is under the control of the *lac* operator and is induced by the non-utilisable lactose analogue IPTG. The Pft1NBD expression clone was grown to an $O.D._{600}$ of 0.5 in 100ml of LB broth and induced by the addition of IPTG to a final concentration of 1mM. Prior to induction and at 1h intervals thereafter, a 1ml sample was removed. The cells were pelleted by centrifugation and prepared for SDS-PAGE analysis.

The over-expression of the nucleotide-binding domain was rather modest. A very weak band, at the expected size of 48.4kDA, could be visualised in induced cultures but could not be observed in photographs of the gels. For this reason no SDS-PAGE gel is presented here. This result was very encouraging and demonstrated that this prokaryotic expression system was capable of expressing *P.brasiliensis* proteins. The small amount obtained was certainly less than would be hoped for and accumulation of large amounts of protein for characterisation would probably require large volumes of culture. Conversely, the lower level of expression may well mean the circumvention of insoluble protein, or inclusion bodies, and this system may afford access to a correctly folded functional molecule.

Prior to purification it was necessary to confirm that the protein was indeed His_6 tagged and localised as a soluble molecule in the cytoplasm. To this end a western blot, using a mouse anti-His_6 antibody, was performed on the total cell protein of the pET21Pft1NBD clone. In addition, the soluble cytoplasmic, membrane and insoluble fractions of the cells were isolated. This blot (Fig6.3) confirmed that the C-terminal His_6 tag was present in the Pft1NBD. However, the protein remained exclusive to the membrane fraction. The estimation of the domain junctions from the hydropathy plot was evidently erroneous and the Pft1NBD persisted as a membrane protein. It was decided to re-assess the strategy for the expression of the Pft1 nucleotide-binding domain in order to circumvent the need to purification of a hydrophobic protein. Such an endeavour is complicated by the need to optimise buffer detergent composition and by the relative poor performance of strains expressing these proteins. It was anticipated that the effort required to accumulate this protein would not be proportionate any additional benefit that may be passed to down-stream experiments.

6.3 Construction of expression systems for Pft1 nucleotide binding domain truncates.

In comparison with the size of the domains purified by other investigators, the Pft1NBD was substantially larger. It was hoped that a larger NBD would allow more extensive investigation of substrate binding, assuming that the bindings sites for molecules other that ATP could be located at some distance from the nucleotide-binding fold. However, as a soluble cytoplasmic preparation would not be readily obtained from this system, it was decided to emulate the strategy used for the murine Mdr1 C-terminal NBD (Conseil, *et al.*, 1998). The homology between this mouse NBD and that of Pft1 readily allowed the comparative alignment of these sequences. The region of Pft1 that corresponded to the previously expressed portion of the mouse protein was noted.

Unfortunately, despite the extensive homology between the two sequences, the single tryptophan residue possessed in this region by the Mdr1 was substituted for serine in Pft1 (Ser^{578}). The analysis of Mdr1 in the aforementioned work relied extensively on the variations in intrinsic fluorescence of this tryptophan as a reporter of conformational change in the protein. The exclusion of this tryptophan residue in our truncate would

Fig 6.3 Western Blot of Pft1NBD in total cell protein (TCP) and subcellular fractions of expressing cells.

Lane 1]	TCP	(0h induction)
Lane 2]	TCP	(1h induction)
Lane 3]	TCP	(2h induction)
Lane 4]	TCP	(3h induction)
Lane 5]	TCP	(4h induction)
Lane 6]	Soluble cytoplasmic fraction	(4h induction)
Lane 7]	Membrane fraction	(4h induction)
Lane 8]	Insoluble fraction	(4h induction)

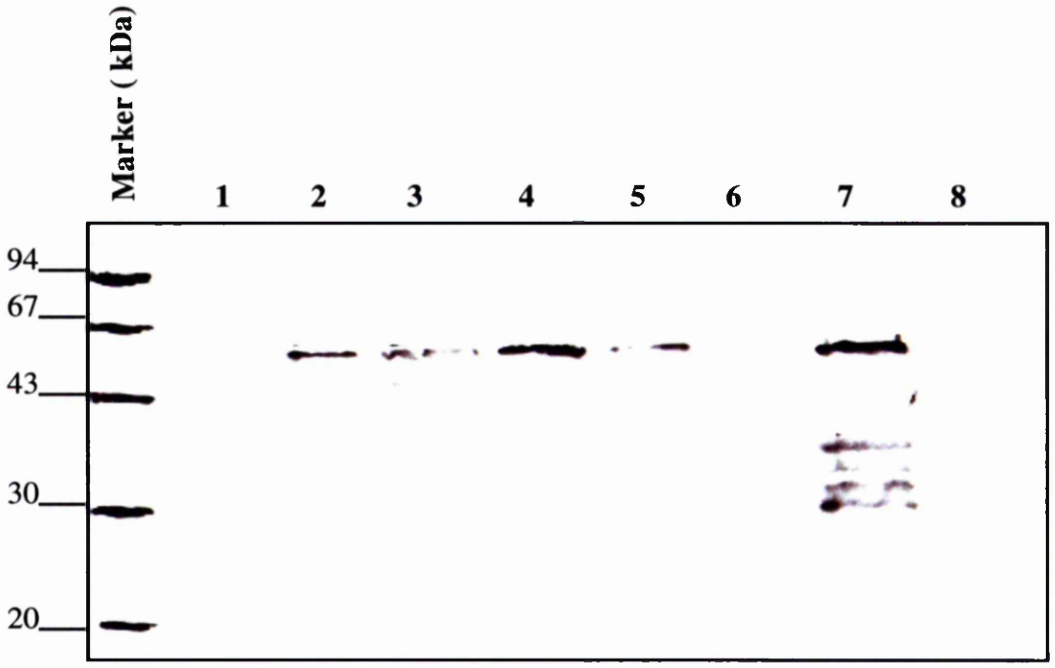


Fig 6.3

s

limit options on the type of fluorescent analysis that could be performed. The differential fluorescence of bound and unbound fluorescent substrate analogues would still be possible but any experiments based on protein tryptophan fluorescence would not.

The inclusion of a tryptophan molecule in the NBD of Pft1 would require an extension of at least 39 amino-acids, the nearest tryptophan being positioned at amino acid 743. It was decided to make two further expression constructs. The construct pET21Pft1NBDT would contain the region of DNA corresponding to the original region used by Consiel *et al.*, (excluding any Pft1 tryptophan), and would consist of amino acids Thr^{516} to Thr^{634}. The second construct, pET21Pft1NBDTW, would also originate from Thr^{516} but would have a further 122 amino acids at the C-terminus (terminating at Val^{756}) thus including the native tryptophan, Trp^{743}. An additional 13 amino acids following the tryptophan were included in the hope that the tryptophan would then be positioned in the correct place in the folded molecule.

The cDNAs were generated using the forward primer Pft1NBDTfwd and either the reverse primer Pft1NBDTrev or Pft1NBDTWrev, the amplicon cloned to pGEM-T Easy™, excised and cloned into the *Bam*HI and *Xho*I sites of pET21b. The resultant pET21b clones are shown in Fig6.4.

Over-expression from these clones, again in *Epicurian coli* ® BL21-CodonPlus™ (DE3)-RIL cells, was performed as for earlier constructs. For the purposes of analysis samples were taken prior to induction and every hour for 3 hours following induction. The resultant SDS-PAGE gel can be seen in Fig6.5. A high concentration of the 21.3kDa Pft1NBDT and 28.0kDa Pft1NBDTW is observed for each respective clone. This result is a sharp contrast to the relatively meagre quantities produced by the earlier construct and initially suggested that purification of large amounts of each protein would be possible. However, a common problem with over-expression to such high levels is the formation of insoluble masses of recombinant protein, or inclusion bodies. Protein in inclusion bodies is incorrectly folded and generally biologically inactive. A number of investigators have reported the restoration of biological activity on the careful refolding of inclusion bodies but any activity observed is often a small percentage of that which is observed from proteins purified directly from a soluble form. On examination of the sub-cellular localisation of the Pft1NBDT and Pft1NBDTW proteins it was found that they persisted almost exclusively in the insoluble form (Fig6.6).

There are a number of strategies that can be employed to alleviate, wholly or partially, the formation of inclusion bodies. The biological mechanisms that induce inclusion

Fig 6.4 Restriction enzyme analysis of pET21bNBD clones

The *Bam*HI/*Xho*I Pft1NBDT and Pft1NBDTW expression fragments were ligated at 16^0C into appropriately prepared pET21b and transformed into Novablue competent cells.

Twelve potential pET21b recombinants (Lanes 1-6 are Pft1NBDT, 7-12 are Pft1NBDTW) were analysed using *Bam*HI and *Xho*I restriction digestion (37^0C for 2h). All recombinant plasmids exhibit Pft1NBDT or Pft1NBDTW inserts.

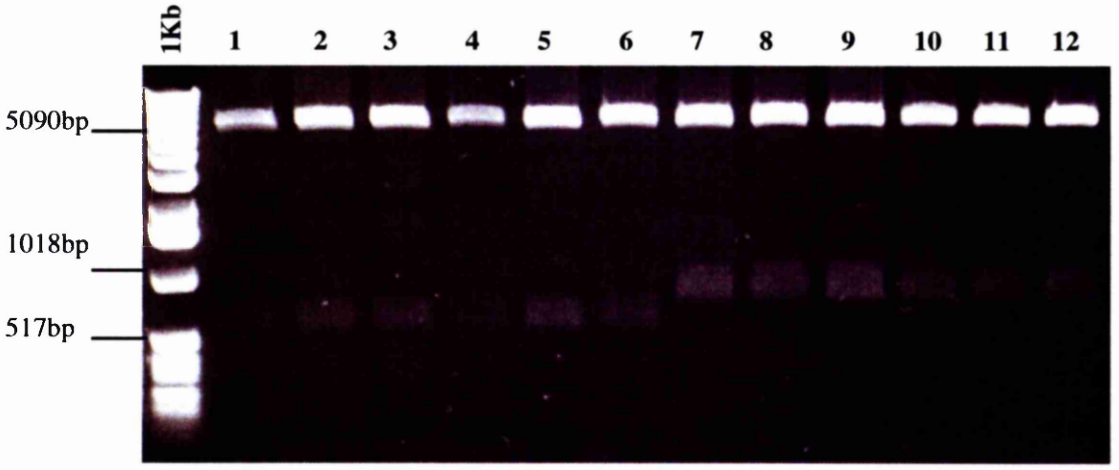


Fig 6.4

Fig6.5 Time-course analysis of expression of Pft1NBDT and Pft1NBDTW

Lane 1]	Pft1NBDT	(0h induction)
Lane 2]	Pft1NBDT	(1h induction)
Lane 3]	Pft1NBDT	(2h induction)
Lane 4]	Pft1NBDT	(3h induction)
Lane 5]	Pft1NBDTW	(0h induction)
Lane 6]	Pft1NBDTW	(1h induction)
Lane 7]	Pft1NBDTW	(2h induction)
Lane 8]	Pft1NBDTW	(3h induction)

i] 21.3 kDa Pft1NBDT

ii] 28.0 kDa Pft1NBDTW

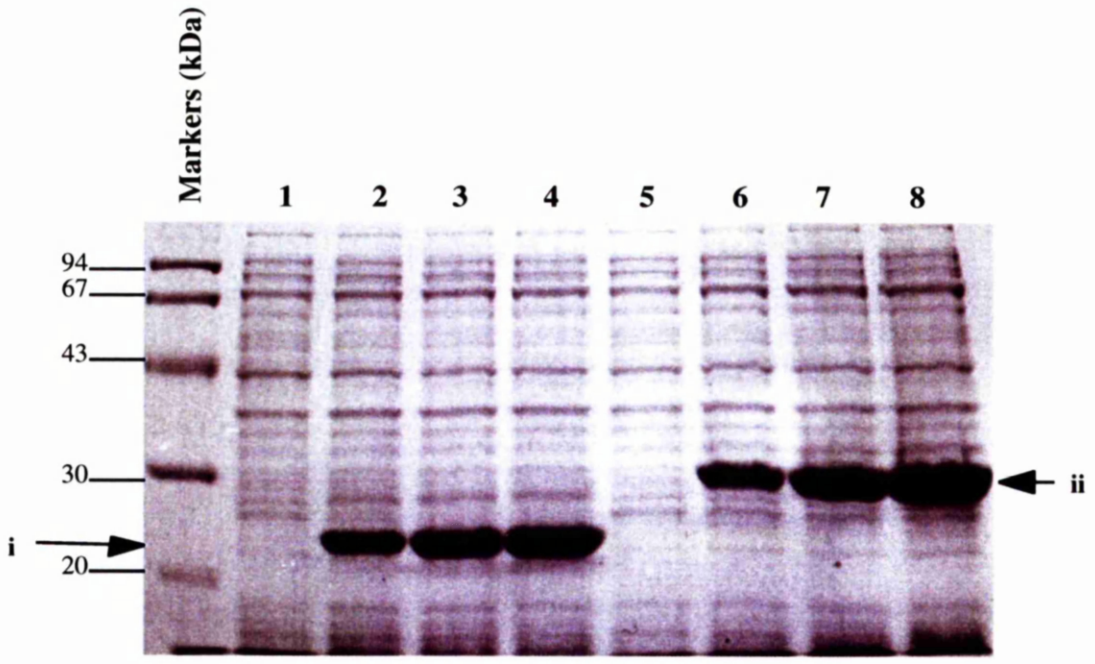


Fig 6.5

Fig 6.6 Subcellular localisation of Pft1NBDT and Pft1NBDTW

Lane 1] Pft1NBDT soluble fraction
Lane 2] Pft1NBDT insoluble fraction
Lane 3] Pft1NBDT membrane fraction
Lane 4] Pft1NBDTW soluble fraction
Lane 5] Pft1NBDTW insoluble fraction
Lane 6] Pft1NBDTW membrane fraction
Lane 7] *Epicurian coli* ® BL21-CodonPlus™ (DE3)-RIL soluble fraction
Lane 8] *Epicurian coli* ® BL21-CodonPlus™ (DE3)-RIL membrane fraction

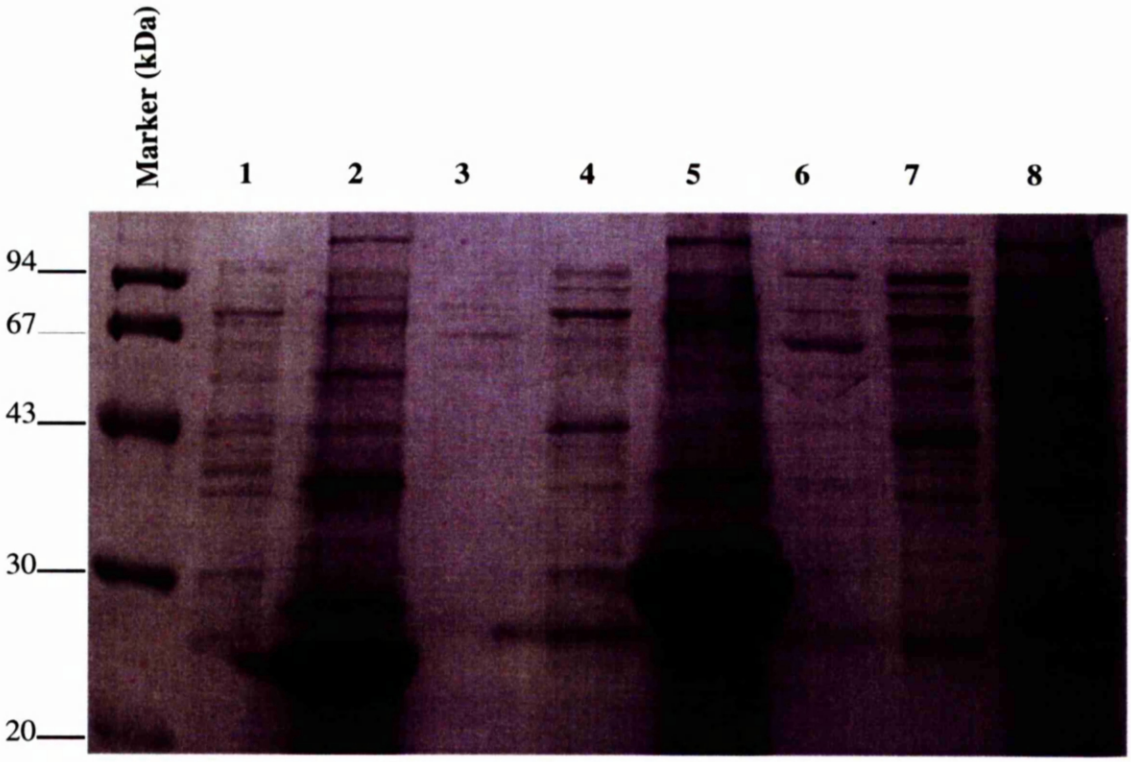


Fig 6.6

body formation are not well characterised but is generally appreciated that their formation occurs as a result of an incapacity in the *E.coli* foldases to process the heterologous or massively over-expressed molecules. Slowing the growth rate of the culture or the rate of expression of the protein can often increase the amount of soluble protein by reducing the stress on the foldase population of the cell. In the case of these Pft1 proteins, conventional measures such as varying the induction point, lowering the culture temperature to 25^{0}C, 20^{0}C or 16^{0}C or reducing the concentration of IPTG used to induce had no effect on the distribution of the protein within the cell (data not shown). However, on occasions where the majority of protein exists in an insoluble form there is often a smaller proportion of the protein folded in the correct manner and soluble in the cytoplasm. This was true for both Pft1NBDT and Pft1NBDTW.

6.4 Denaturing purification of Pft1NBDTW.

It was decided to concentrate purification efforts on Pft1NBDTW. This protein would afford all the features of Pft1NBDT with the additional advantage of the single tryptophan residue that may prove useful as a fluorescent reporter for biochemical studies. Ideally, biochemical studies are performed using protein folded in a native conformation as isolated from the soluble cell fraction. In the case of Pft1NBDTW, which manifested as massive amounts of inclusion bodies material, it appeared prudent to attempt purification of this insoluble protein. A successful purification would demonstrate the inclusion of a His_6 tag on the protein. Also, in the event that a reasonable amount of soluble Pft1NBDTW could not be purified it may be important to resort to this insoluble material for a source of protein, which would be refolded prior to analysis.

In order to purify the protein, the inclusion body fraction of the cells were prepared from cell lysate by centrifugation. Generally, the insoluble pellet from 1L of culture, which had been induced for 3h, was collected and prepared. These inclusion bodies were solubilised in 30ml of 8M urea, 0.1M NaH_2PO_4, 0.01M TrisHCl, pH8.0, and applied to an NTA-agarose affinity column. This procedure of purification under denaturing conditions relies on a decreasing affinity of the His_6 tag for nickel as the buffer pH falls. The column was washed in excess volumes (4 x 80ml) of the aforementioned buffer at pH6.3. It is important to wash extensively when large concentrations of denatured proteins are involved to ensure that the proteins observed in elutions are indeed purified and not merely dilutions of the crude extract applied to the column. The protein content of each wash was assessed by spectrophotometric analysis of A_{280} during the procedure. Only when the A_{280} of the wash flow-through was less than 0.01 was the washing considered to be complete. In this case it was

usually the fourth wash that achieved the satisfactory low level of contaminants. The bound protein was then eluted in 4 x 0.5ml of the above urea buffer at pH 4.5.

The result was the purification of a single band on SDS-PAGE that corresponded to Pft1NBDTW (Fig6.7). It was noted that this procedure would supply large amounts of purified protein, but that protein would not be immediately useful for biochemical studies.

6.5 Native purification of soluble Pft1NBDTW

The soluble fraction of the cell was prepared by the removal of inclusion bodies by centrifugation and of membrane material by ultracentrifugation. This soluble fraction was applied to a column and purification was attempted under native conditions. Initially, the conditions used were those proposed by the Ni-NTA matrix manufacturer. Essentially, the manufacturer advises the use of buffers based on 50mM NaH_2PO_4 and 300mM NaCl at pH8.0. Imidazole is included in each buffer at an appropriate concentration, competing with and eventually displaces the His_6 tag of the protein from the nickel ion. The advised concentrations of imidazole are 10mM for lysis buffer, 20mM for wash buffers and 250mM for elution buffers.

Despite evident Pft1NBDTW protein in the cytoplasmic fraction of the cells, purification was not achieved initially. Precipitation of the protein during the purification process was suspected to be the cause and conditions were varied in an attempt to retain the protein in soluble form. Variation in NaCl concentration (from zero to 500mM) did not seem to have any particular affect. The inclusion of glycerol in the buffers did result in the purification of protein and titration of the glycerol concentration over a number of preparations suggested that 15% was optimal (data not shown). This observation supports the notion that precipitation had been the problem in previous preparations, as glycerol is known to decrease hydrophobic interactions between proteins.

Additional buffer optimisations were required to reduce the number of contaminant proteins associated with Pft1NBDTW in the elutions. In particular an increase in the imidazole concentration of the wash buffer to 40mM allowed the production of Pft1NBDTW of higher purity. Manipulation of the pH of buffers was not attempted as a pH of less than pH8.0 allows the action of metalloproteases. Despite the inclusion of protease inhibitors in all buffers, the action of metalloproteases would not be inhibited (EDTA is generally used to achieved this inhibition but its inclusion would chealate all nickel from the NTA matrix used in the purification procedure).

Fig6.7 Purification of Pft1NBDTW under denaturing conditions

Lane 1] Column flow-through
Lane 2] Wash number 1
Lane 3] Wash number 4
Lane 4] Elution number 1
Lane 5] Elution number 2

Buffer compositions and volumes

Lysis buffer; 8M Urea
0.1M NaH_2PO_4
10mM TrisHCl pH8.0
1L pellet was resuspended in 30ml

Wash Buffer; As lysis buffer but with 10mM TrisHCl at pH6.3
4 x 80ml washes

Elution Buffer; As lysis buffer but with 10mM TrisHCl at pH4.5
4 x 0.5ml elutions

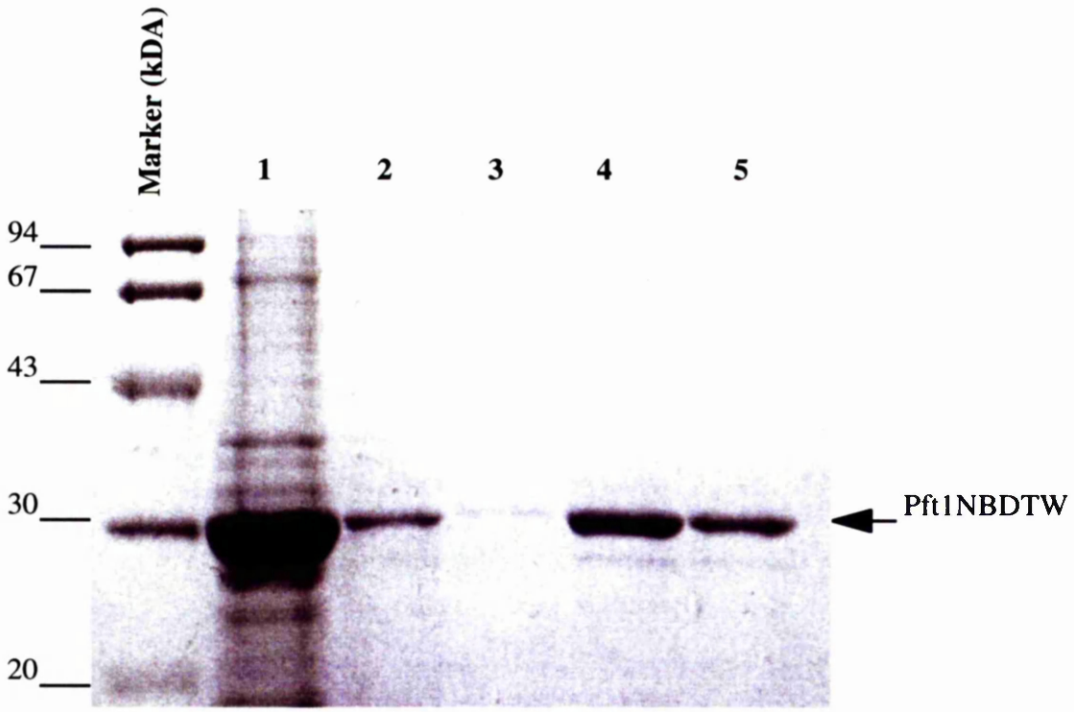


Fig 6.7

The composition and volume of each buffer used is noted in Fig6.8, which shows a typical SDS-page gel of the native purification procedure.

Following purification and SDS-PAGE analysis, the elutions were generally combined and placed in 14kDa pore-size dialysis tubing. The protein was dialysed against 3000 volumes of dialysis buffer (50mM NaPi pH8.0, 300mM NaCl, 15% glycerol plus protease inhibitors). Dialysis was permitted at 4^0C for 36h and the complete removal of imidazole was assured by fluorimeter analysis. Imidazole is fluorescent and is excited at 380nm. It can thus be detected as a fluorescence peak if present in the protein preparation. The protein was then aliquoted into 200µl volumes in sterile cryotubes and snap frozen in an ethanol/dry ice bath. The frozen stocks were stored at -80^0C until required.

6.6 Discussion

Initial expression of the entire C-terminal of Pft1 proved to be successful to a degree with a low level of expression of the 48.4kDa protein observed in the whole cell protein of the expressing cells. Unfortunately, the estimation of domain boundaries using a Kyte-Doolittle hydropathy plot of Pft1 proved to be ineffective in denoting the beginning of the hydrophilic portion of this protein, and the Pft1NBD remained associated with the membrane. This is perhaps of no surprise. No proven consensus is available which defines either the number of position of membrane associated regions in these proteins. Four transmembrane domains are proposed for AfuMDR2 (Tobin *et al.*, 1997) while other attribute six TMs to MDL1 and MDL2 (Bauer *et al.*, 1999). Despite the problematic locale, western blot did demonstrate that this molecule could be expressed in the *E.coli* strain, which was encouraging for the further refinement of the expression systems.

Essentially a recombinant expression system was required which would supply a correctly folded, native molecule which would remain soluble while exhibiting the basic biochemical properties of a nucleotide-binding domain, essentially the ability to bind ATP and, perhaps, other ligands (the eventual ligand binding studies are discussed in Chapter 7). This had been achieved by other investigators, including the study of the modulation of ATP-binding by flavonoids in the murine Mdr1 C-terminal NBD. By adopting the dimensions of this nucleotide binding domain, and a 52 amino-acid C-terminal extension of those dimensions, a further two expression systems were constructed. These were termed pET21bPft1NBDT and pET21bPft1NBDTW respectively (NBD Truncate and NBD Truncate W [tryptophan]). The dynamics of

Fig 6.8 Purification of Pft1NBDTW under native conditions

Lane 1]	Soluble fraction of Pft1NBDTW expressing cells (pre-purification)
Lane 2]	Column flow-through
Lane 3]	Wash number 1
Lane 4]	Wash number 2
Lane 5]	Wash number 3
Lane 6]	Elution number 1
Lane 7]	Elution number 2
Lane 8]	Elution number 3
Lane 9]	Elution number 4

Buffer compositions and volumes

Lysis buffer;	50mM NaH_2PO_4 pH8.0, 300mM NaCl, 15% glycerol, 10mM imidazole. 3L cell pellet was resuspended in 40ml
Wash buffer;	50mM NaH_2PO_4 pH8.0, 300mM NaCl, 15% glycerol, 40mM imidazole. Each wash was applied at a volume of 50ml
Elution buffer;	50mM NaH_2PO_4 pH8.0, 300mM NaCl, 15% glycerol, 250mM imidazole. Each elution was applied at a volume of 0.5ml

All buffers also contained EDTA-free protease inhibitors.

Fig6.8

expression in these systems was in extreme contrast to those observed for the original Pft1NBD construct, with the recombinant protein representing an estimated 80% of the total cell protein. Excellent yield of Pft1NBDT and Pft1NBDTW could be observed after only 1h induction. As is often the case with such high level expression of heterologous proteins, the product was isolated, to a very great extent in inclusion bodies, thus rendering it inaccessible to direct biochemical analysis. Unfortunately, all of the common techniques used to minimise inclusion body formation, such as manipulation of temperature and induction conditions, proved ineffective in this situation.

Previous work by others had suggested that the amount of soluble NBD obtained was inversely proportional to the size of the molecule expressed. However, for these two NBD constructs of Pft1, the expression performance of both constructs was rather similar in terms of both total yield and the yield of soluble protein. It was decided to continue the work only on Pft1NBDTW. This larger protein possessed all of the features of Pft1NBDT and included a native tryptophan towards the C-terminus.

Denaturing purification of the inclusion bodies demonstrated that the Pft1NBDTW was successfully His_6 tagged and very high yields (hundreds of milligrams) could be prepared from 1 litre of culture. This was a re-assuring observation because it meant that large amounts of protein was readily available should refolding experiments be required to provide material for further investigation. This was reinforced by the relative ease of purification which required no optimisation of recommended purification protocols.

As with many cases of inclusion body formation a relatively small yet notable yield of protein could be found in the cytoplasmic fraction that represented a native and soluble form. While this yield was dwarfed by that of the inclusion body material, it was decided to pursue this small amount for the biochemical studies rather than go through what could be laborious and potentially ineffective refolding experiments. Refolding was dismissed for three main reasons. Firstly, as the molecule was a truncated form of a larger protein there was a chance that sub-optimal refolding conditions could result in an aberrant conformation that was not akin to that of the conformation of the domain in the full-sized molecule. Secondly, it would be very difficult to ascertain if the behaviour of the molecule in future experiments would be due to a true biochemical characteristic or an artefact of refolding. This is particularly true in this case where the "activity" is not a measurable enzymic activity but a measure only of ligand binding, where true processive function is not demonstrated. Finally, as the amount of restored biological activity would probably be partial (relative to material from a native source) it

may be difficult to attribute any numerical data to observations without first knowing the proportion of protein yield that had attained the correct conformation. This is impossible without well characterised, active, native protein as a reference.

The native purification of Pft1NBDTW proved difficult at the outset and the suitability of the standard purification protocol was not observed as it was for the denaturing purification. Despite visible evidence of the Pft1NBDTW in the soluble fraction and it's retention of the column (as demonstrated by it's absence from the initial column flow-through) the protein did not appear on elution, even when imidazole concentrations were elevated to 500mM. It was suspected that this was probably due to aggregation or precipitation, perhaps exacerbated by the fact that the molecule in question was a truncate and some exposed faces of the protein may normally be stabilised by interactions with portions of Pft1 which were omitted in this construct? The addition of glycerol to 15% proved effective in stabilising the protein and a purified 28.0 kDa protein was observed by SDS-PAGE. This protein remained stable during dialysis and cryostorage. The protein was then used in a number of biochemical experiments reported in Chapter 7.

Chapter 7

Ligand binding studies for Pft1NBDTW.

In order to attribute a function to Pft1 a number of options for investigation were considered. Initially, and perhaps most logically, the recombinant expression of the entire Pft1 protein in a drug sensitive mutant of *S.cerevisiae* could be attempted. This could provide a direct answer to the question of whether or not Pft1 is involved in multidrug resistance. This option was developed, to a point, and is discussed briefly at the close of this chapter. However, in designing experiments to answer some of the basic questions we had to acknowledge that there were scenarios that could preclude a successful result from such a direct approach.

Firstly, many of the modular components of bacterial ABC transporters are expressed separately and form a dimeric or oligomeric complex. The bacterial maltose transporter is also an oligomeric complex of $MalFGK_2$ (Mourez, *et al.* 1998). Similarly, the histidine permease of *Salmonella typhimurium* is arranged as $HisQMP_2$ (Pei-Qi and Ames 1998). The *Lactococcus lactis,* LmrA is perhaps a good example of an ABC half-transporter which is functional as a homodimer (Van Veen, *et al.* 1998). However, it has to be noted that dimer complexes in eukaryotic ABC transporters are most often heterodimers. The best characterised must be the Tap1/Tap2 heterodimer complex involved in MHCI antigen presentation in mammalian cells (Abele and Tampe 1999). Also a heterodimeric ABC transporter has been described in *S. cerevisiae* which displays considerable structural homology to the human adrenoleukodystrophy transporter complex (Shiani and Valle 1996). The real possibility that Pft1 may represent a component of a hetero-oligomeric complex meant that straight forward heterologous complementation might not provide a phenotype.

It was also decided that the association of Pft1 with drug transport was, in fact, an assumption and the potential that the protein may have no affinity for drug must be addressed. This notion is partly supported by the fact that the *S.cerevisiae* AD1-8$^-$ mutant (a gift from Prof. A. Goffeau) used for the complementation assays retains copies of both *mdl1* and *mdl2* but continues to exhibit a high sensitivity to selected drugs. If these half-transporters are involved in drug resistance then why do the Mdl1 and Mdl2 proteins not result in a persistent Mdr phenotype in this mutant strain. However, it must be noted that the PDR1 and PDR3 transcription factors involved in the Pdr network (Balzi and Goffeau 1995) are inactive in this strain and the transcriptional control of the two ABC transporter genes in relation to these factors has

not been addressed. It is possible that the expression of Mdl1 and Mdl2 is averted due to this deletion.

In addition to the direct question of phenotype (Is Pft1 a drug transporter?), a more subtle biochemical approach was employed. Essentially the issues that we aimed to address related to the competence of the purified native Pft1NBDTW (see chapter 6) to bind substrates. Substrate binding would be confirmed by a demonstration of conformational change (measuring variation of intrinsic tryptophan fluorescence) or by fluorescence resonance energy transfer (FRET) (the transfer of fluorescence energy from the excited tryptophan to a bound fluorescent substrate). In this way we could implicate biochemical function at a protein level even if a phenotype could not be demonstrated at the level of the organism.

7.1 Fluorescence based analysis of ATP binding to Pft1NBDTW.

In order to demonstrate that the Pft1NBDTW retained an affinity for ATP, a stopped-flow device was employed. This device allows the rapid mixing of both substrate and protein in a small chamber and the resultant changes in fluorescence to be measured repeatedly over a very short time base. Thus a transient profile of progressive fluorescence change can be plotted.

For the purposes of these experiments a fluorescent ATP analogue, MANT-ATP, was used. The MANT-group on this molecule can be useful in two ways. This additional group is an efficient receptor of the fluorescence energy generated by tryptophan molecules and is thus a good candidate for FRET based analyses. Additionally, binding of MANT-ATP can result in a tryptophan independent increase in MANT fluorescence which occurs as result of the shielding of the MANT group from the aqueous environment by internalisation within the protein structure. Either phenomena would illustrate ATP binding.

The stopped flow was programmed in the following way. An input beam, at an excitation wavelength of 280nm was applied to the mixing chamber, via two monochromators. Measurements of tryptophan fluorescence were recorded over a 10 second logarithmic time base. These measurements were made on the basis of changes in light detection by a photomultiplier tube (PMT). Essentially, the emission light derived from tryptophan or MANT fluorescence passes from the chamber through

an appropriate filter (320nm for Trp fluorescence, 420nm for MANT-ATP) to the photomultiplier (PM) tube. The gathered emission light causes an alteration in the voltage of the PM tube and this variation in PM volts (PMV) is displayed over the selected timebase. An increase in PMV denotes an increase in fluorescence while a PMV decrease reflects a fluorescence decrease.

Eight replicate traces were generated for each form of fluorescence (tryptophan or MANT). This allowed averaging at each data point and a mean PMV trace to be plotted. For each individual trace a solution of 25μM MANT-ATP in buffer was mixed with an equal volume of protein (at 400μg per ml). The MANT-ATP solution was prepared in the same buffer that was used to dialyse the protein of that preparation. It is the mean trace that is presented for each experiment. It is notable that the reproducibility between each trace was very good and the variation at any data point was minimal.

The analysis of a change in tryptophan fluorescence was the first to be examined. While FRET requires both ligand and tryptophan to be in relative close contact, the observation of tryptophan fluorescence is independant of the location of ligand binding. The degree of tryptophan fluorescence is directly related to the extent to which the tryptophan residue is accessible to bombardment by water molecules. The collision of water with the fluorescent aromatic group dissipates fluorescent energy and thus reduces the observed output of light at 320nm. Thus an increase in tryptophan fluorescence relates to a protein conformational change that results in the internalisation and shielding of the amino-acid. Conversely, a decrease in fluorescence would indicate a conformational change which exposes the tryptophan to the quenching aqueous environment.

The trace presented in Fig7.1 is the mean trace of the change in the tryptophan fluorescence of Pft1NBDTW when the protein was mixed with MANT-ATP. An increasing signal is observed indicating a conformational change and suggestive of internalisation of the tryptophan. The reaction is rapid, occurring over approximately 0.07sec. The trace was fitted to a double expontential equation using the Applied Photophysics software. The initial rapid phase had an amplitude of 0.127V. The PMT was zeroed to a 4V signal so the percentage fluorescence change can be calculated as 0.127/4 x 100. This defines a fluorescence change of 3.175%. The conformational change appears biphasic with a rapid increase in fluorescence from 0 to 0.02 sec followed by a change of slower rate.

Fig7.1 Analysis of Pft1NBDTW tryptophan fluorescence on binding of MANT-ATP

Sample Reservoir A; 25 µM MANT-ATP
50mM NaH_2PO_4 pH8.0
300mM NaCl
15% glycerol

Sample reservoir B; Pft1NBDTW (400µg ml^{-1})
50mM NaH_2PO_4 pH8.0
300mM NaCl
15% glycerol

Input light at 285nm. Output light filtered at 320nm

100µl volumes from each sample reservoir were rapidly mixed in a stopped flow apparatus and fluorescence output measured over a 10s timebase.

This trace represents the averaging of points from 8 individual traces.

X Axis Time (Seconds)

Y Axis PMT Voltage signal

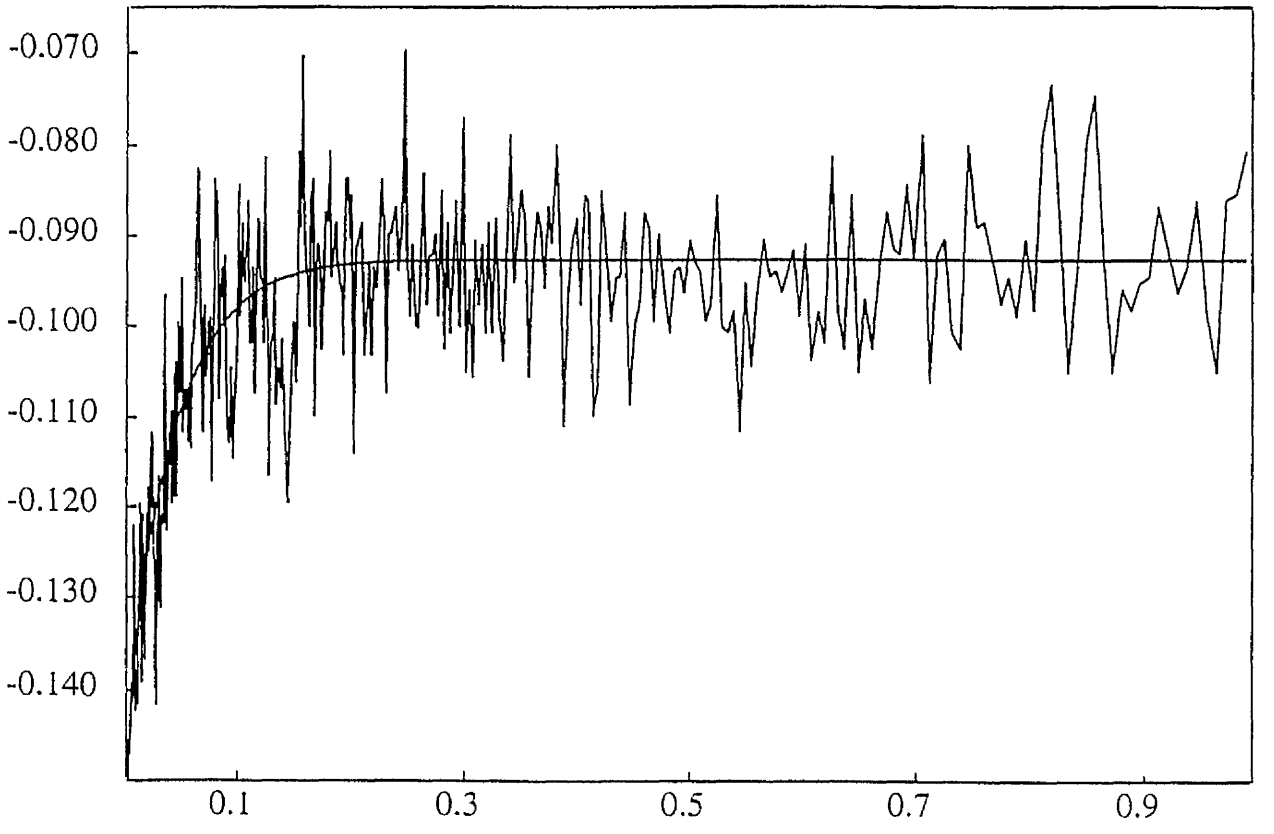


Fig7.1

The amplitude of the second expontential (0.051V) gives a further fluorecence increase of 1.275%. This increase in fluorescence signal was not observed in control experiments where the Pft1NBDTW or the MANT-ATP were mixed with reaction buffer alone. Thus we can conclude that the protein does indeed exhibit affinity for ATP (or at least an ATP analogue). Also this initial experiment demonstrated that the C-terminal tryptophan was an effective reporter for conformational change.

The filter for emission light was replaced and the experiment was repeated, this time analysing emission from the MANT-ATP. Again an increase in fluorescence was observed (Fig7.2). The major fluoresence changes appear to occur over approximately 0.8 seconds. The data has been fitted using a double exponential equation revealing a biphasic configuration. The first rapid increase in fluoresence has an amplitude of 0.107V representing a 2.67% increase in fluoresence. This occurs over 0.02 seconds. The second slower exponential has an amplitude of 0.054V allowing for a further 1.35% increase. This result consolidates the conclusion that conformational change is occurring as a consequence of ligand binding. The increase in output can be attributed to binding of MANT-ATP in such a way that the MANT group was internalised into the protein structure, thus shielding it from the aqueous environment, or as a result of fluorescence energy transfer from the tryptophan. It is not possible on the strength of this data to dissect which of these two mechanisms is responsible or indeed if both are at work. If FRET is involved then the signal of tryptophan fluorescence of Fig7.1 may under-estimate the true output from the tryptophan when the MANT-ATP is bound, because some of this energy would have been donated to the MANT group.

These experiments were repeated in the presence of 10mM $MgCl_2$, ATP binding is usually observed to be greatly enhanced by the presence of the Mg^{2+} ion. In the case of Pft1NBDTW the traces with and without 10mM $MgCl_2$ proved to be identical suggesting no involvement of magnesium in this case (data not shown).

7.2 Steady state binding studies for azole binding to Pft1NBDTW.

The successful binding of the MANT-ATP to Pft1NBDTW implied that the protein was correctly folded and performed the basic function of binding nucleotide. We now wished to examine interactions with a drug ligand. A number of NBDs for ABC transporters display an affinity for the substrates that they transport. These substrates have been shown to modulate the affinity of the NBD for ATP and thus act as modulators on the process of energising the transport of solute (Conseil, *et al.* 1998).

Fig7.2 Analysis of MANT-ATP fluorescence on binding to Pft1NBDTW

Sample Reservoir A; 25 μM MANT-ATP
50mM NaH_2PO_4 pH8.0
300mM NaCl
15% glycerol

Sample reservoir B; Pft1NBDTW (400μg ml^{-1})
50mM NaH_2PO_4 pH8.0
300mM NaCl
15% glycerol

Input light at 285nm. Output light filtered at 420nm

100μl volumes from each sample reservoir were rapidly mixed in a stopped flow apparatus and fluorescence output measured over a 10s timebase.

This trace represents the averaging of points from 8 individual traces.

X Axis Time (Seconds)

Y Axis PMT Voltage signal

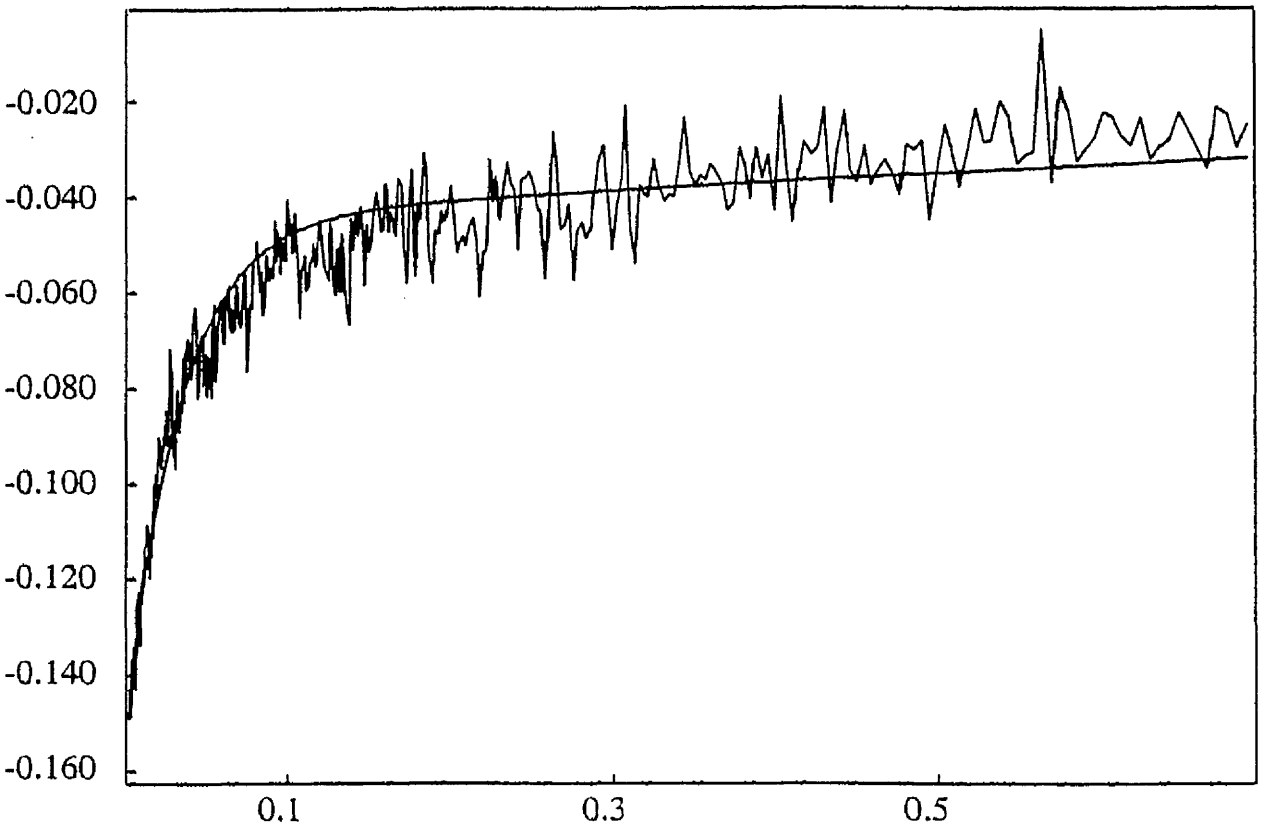


Fig7.2

The upregulation of *pft1* expression in direct response to an azole stimulus was the only evidence pointing towards any phenotype afforded by this protein. Therefore interaction of azoles with the nucleotide binding domain was an appropriate avenue to follow. It was decided to use ketoconazole, the therapeutic drug of choice for paracoccidiodiomycosis. If Pft1 has evolved a drug binding capacity it would have done so in the presence of ketoconazole rather than any other azole.

The intention was to use fluorimetery to measure the steady-state change in tryptophan fluorescence in response to the drug. This technique has proved useful for characterisation of a wide range of kinetic systems. A lamp supplies input light at a required wavelength, in this case 285nm, via a series of chromating mirrors. The input light enters a quartz cuvette containing the test substances and excitation of tryptophan is achieved. The emission light is then measured at right angles to the input light, assuring that the measured light is indeed emission based and not merely transmitted light from the input source (as used in spectrophotometry).

For the experiments described below (in section 7.2 and 7.3), carefully designed conditions were employed in order to ascertain whether the observed changes in emission output had been influenced by inner-filter effect. The inner-filer effect can occur if a non-protein ligand or other substance tends to absorb light at 285nm. This reduces the amount of light available for excitation of the tryptophan and the degree of tryptophan fluorescence would appear artificially low. To this end a second quartz cuvette was placed directly in front of the cuvette housing the protein. All substances used for these experiments were applied to the front cuvette; buffer plus ketoconazole, buffer plus ATP, buffer plus ketoconazole and ATP (at all concentrations used in experiments) and buffer alone. It was noted that a significant and concentration dependant inner filter effect could be observed with ATP. The degree of inner filter effect was noted and all values presented in the following data have been corrected accordingly. Also all observations were carried out with the front cuvette in place (filled with buffer) to eliminate any variation in input light transmission which may have occurred from the use of two parallel cuvettes.

The intention was to monitor the flux in tryptophan fluorescence on addition of drug. However, on execution of the experiment, the serendipitous discovery that ketoconazole was highly receptive to FRET superseded the direct tryptophan based approach. To the test cuvette was added Pft1NBDTW and ketoconazole to a final concentration of 100μM. Ketoconazole exhibits a basal level of fluorescence on excitation at 280nm, emitting at 340nm. This fluorescence is markedly elevated in the presence of protein, indicating binding of the azole to Pft1NBDTW (Fig7.3).

Fig7.3 FRET transfer from Pft1NBDTW to ketoconazole

X Axis Wavelength (nm)

Y Axis Arbitrary fluorescence units

Peak at 280nm and 200 Absorbance units	Pft1NBDTW alone
Peak at 390nm and 200 Absorbance units	Ketoconazole alone
Peak at 280nm and 950 Absorbance units	Pft1NBDTW plus Ketoconazole

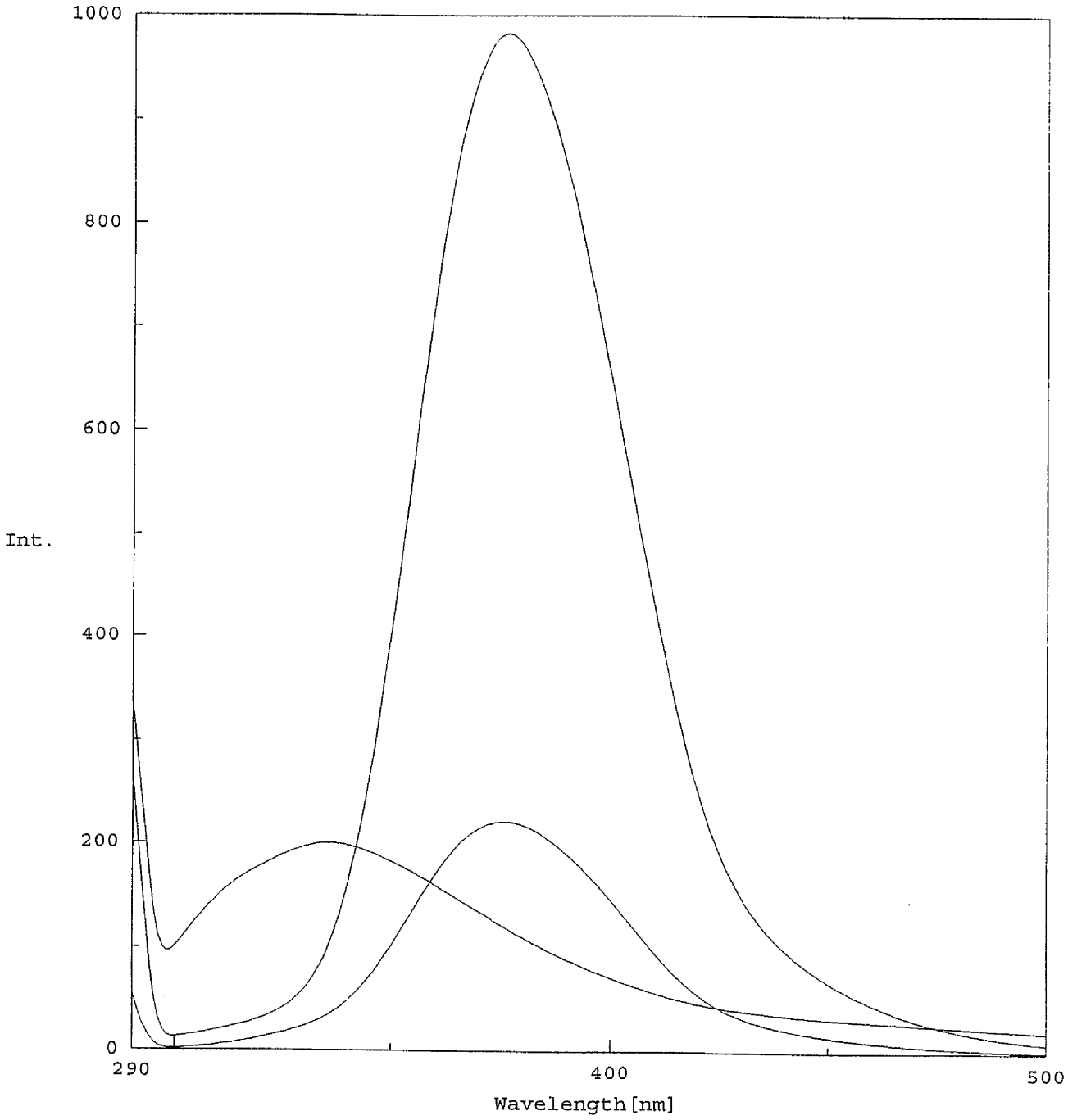


Fig7.3

Tryptophan to ketoconazole FRET results in the removal of all evidence of tryptophan fluorescence, presumably through a high efficiency of energy transfer. This indicates that ketoconazole does bind to the NBD and that the location of this binding site is very local to the tryptophan. The increase in the fluorescence of ketoconazole on the inclusion of protein is much greater than the observed decrease in fluorescence of the tryptophan. It is possible that the extent of ketoconazole fluorescence is enhanced by a FRET independent increase in azole fluorescence that may occur if the bound drug is buried within the protein.

On the basis of this effect, the concentration dependant affinity of binding could be investigated. Ketoconazole was applied to the fluorimeter cuvette at increasing concentrations (0 to 60 μM), thus titrating the drug against protein. Each concentration was measured in triplicate (from three replicate titrations) and the arbitrary fluorescence units attained by the ketoconazole FRET peak were noted. Fig7.4 shows a plot of the azole titration. Each point represents the mean value for replicate observations but no error bars are included because the reproducibility of data was almost exact (fluctuation occurred only in the second decimal place of the fluorescence unit).

A non-linear regression analysis was performed using Sigmaplot™. This process applies the data to the following equation;

$$\Delta F/F_0 = (\Delta F_{MAX}/[Pft1NBDTW])/F_0 \,.\, \{(K_d + [Pft1NBDTW] + [Ketoconazole]) - \sqrt{(K_d + [Pft1NBDTW] + [ketoconazole]^2 - 4\,([PFT1NBDTW][Ketoconazole])}\}/2$$

Where F is the fluorescence.

Using this equation a K_d value of 0.959μM was calculated.

7.3 ATP alters the binding of ketoconazole on Pft1NBDTW.

The experiments on the murine NBD discussed earlier (Conseil, *et al.* 1998) focus on the kinetics of ATP binding and the role of flavonoid ligands in the modulation of the rate of ATP binding. For Pft1 investigations involving ATP and ketoconazole were pursued but from a slightly different perspective. It was decided to focus on the effect of the presence of ATP on the affinity for ketoconazole rather than the presence of ketoconazole on ATP binding. ATP was applied to the fluorimetery experiments at a

Fig 7.4 **Titration of ketoconazole against Pft1NBDTW**

X Axis Ketoconazole concentration (μM)

Y Axis Arbitrary fluorescence units

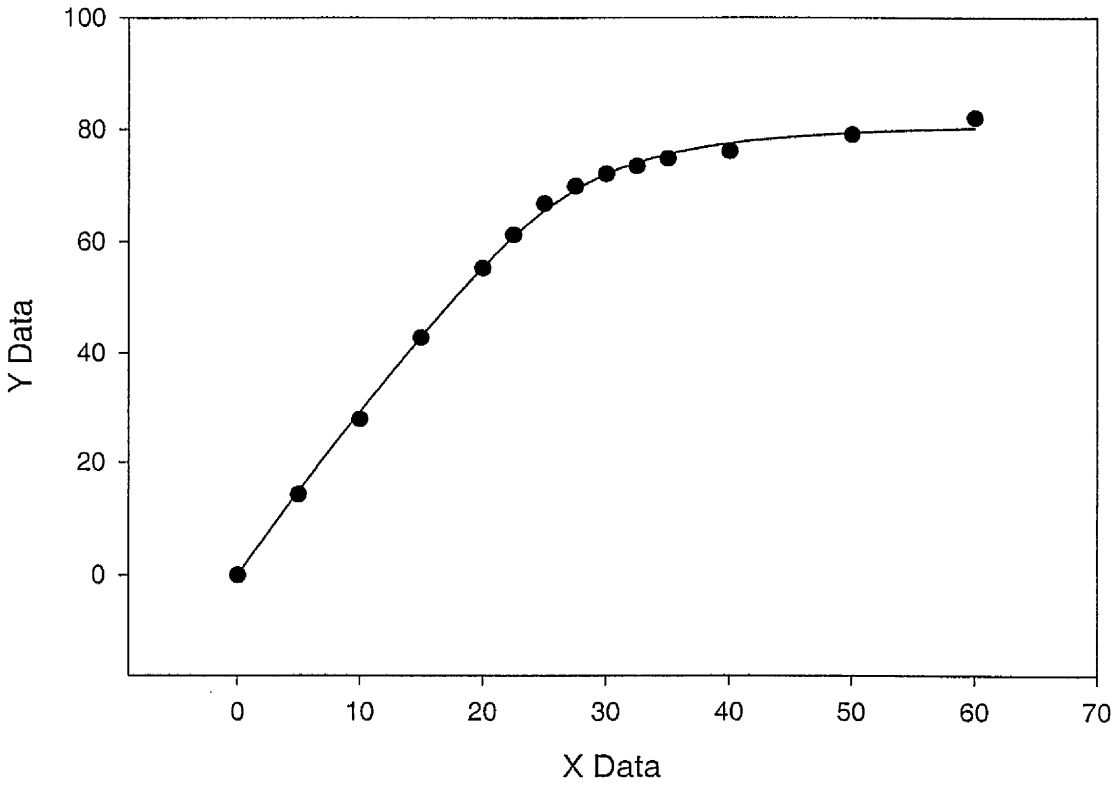


Fig7.4

constant concentration of 10μM, pre-incubated at 20^{0}C with the protein for 10min prior to measurements. The titrations with ketoconazole were repeated, again in triplicate. The results of this titration can be observed in Fig7.5. Increased levels of FRET are observed. Using the equation described above a K_d value of 0.858μM was calculated. This suggests that a slightly higher affinity for the azole was observed on the addition of ATP. Also this increase in fluorescence output is consistent with the internalisation of the drug supporting this hypothesis of tighter binding.

7.4 Heterologous complementation of a drug sensitive phenotype in *S.cerevisiae* by Pft1.

As mentioned in the opening section of this chapter, the direct way to demonstrate a drug transporter phenotype with Pft1 was heterologous complementation. The potential pitfalls of this approach were also discussed. These experiments were introduced but were not followed up to a conclusive level. This section serves to mention the status of this approach and the few observations made so far.

The construction of expression systems for Pft1 in *S.cerevisiae* was achieved in pYES2-TOPO-His. Two cDNAs were amplified, allowing expression from each of the alternate ATG codons implicated in Pft1 transcription (data not shown). These cDNAs were T-A cloned to the pYES2 vector. As A-T cloning is not directional, the orientation of the inserts were confirmed by PCR using the forward Pft1 primer and the reverse T3 sequencing primer of the pYES2 vector. Correct clones were identified for both constructs. These constructs were transformed to *S.cerevisiae* AD1-8$^-$.

7.5 Discussion

The purified Pft1NBDTW proved to be very useful in the initial demonstration of Pft1 ligand affinity. The integrity and correct folding of the Pft1NBDTW was demonstrated by the competence of the recombinant protein to bind ATP. This was demonstrate from both protein (tryptophan fluorescence) and ligand (MANT-ATP fluorescence) perspectives. Stopped-flow analysis showed a rapid increase in fluorescence for both protein and ligand over a short time base, indicating that the binding of ATP results in a rapid conformational change in the NBD. The ability of the NBD to bind ATP was

Fig 7.5 **Titration of ketoconazole against Pft1NBDTW in the presence of 10mM ATP**

X Axis Ketoconazole concentration (μM)

Y Axis Arbitrary fluorescence units

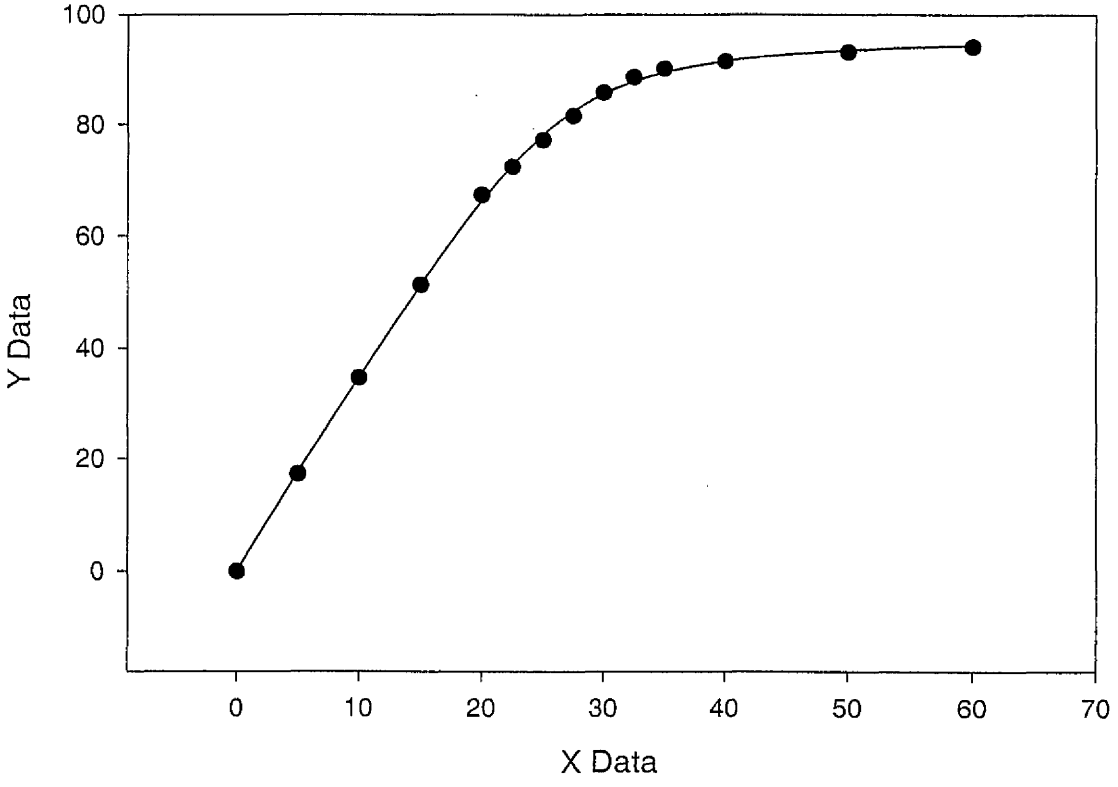


Fig7.5

expected as it had been successfully detected by various means by other investigators using similar constructs (Dayan, *et al.* 1996; Conseil, *et al.* 1998).

Until these investigations, no biochemical analysis had been performed on the Mdr-like half-transporters of fungi that had been proposed to be involved in drug resistance. The experiments above do not prove that Pft1 is a drug transporter either. They do give conclusive proof that Pft1 exhibits an affinity for ketoconazole (K_d=0.959μM). In addition these data suggest that ATP may improve ketoconazole binding providing another example of ATP co-operativity in ligand binding (K_d=0.858μM).

It is important to note that the binding of a substrate to the NBD does not indicate that this substrate will be transported by the protein. Indeed, in other well characterised transport systems where the solute binds to the NBD we find that this precise molecule is not transported but merely served to positively modulate function to facilitate the transport of another. The NBD of bacterial arsenite transporter complex (ArsA) has the capacity to bind both arsenite and antimonite ions. The binding of these ions to the NBD is not a precursor to their transport but a precursor to the transport of another identical ion. Secondly, if some current models are to be believed (eg. the hydrophobic vacuum cleaner model (Bolhuis, *et al.*, 1997)) the extrusion of hydrophobic substances occurs by their direct removal from the lipid bilayer during passage. It is initially difficult to see how the molecule to be transported could bind with such high affinity to the soluble portion of the NBD if this is true.

While there is no experimental evidence to support the following hypothesis, it remains an interesting notion. The high affinity demonstrated for ketoconazole would mean that the drug would bind to the Pft1NBD when the intracellular concentration of the drug is low. The binding of the ketoconazole to the NBD could serve to allosterically activate the membrane portion of Pft1 and enhance the removal of drug molecules passing through the plasma membrane. This environmental sensing mechanism could convey a mechanism of selective activity on the ABC transporter so that, in the absence of the azole, the passage of other hydrophobic molecules through the bilayer is not affected. The considerable homology between the Pft1NBDTW and the corresponding regions of other proteins could mean that this is a widespread phenomenon.

While the heterologous complementation assay for drug transport by Pft1 were not conclusive it is not possible to state that Pft1 one does not act as a drug transporter. There are a considerable number of technical difficulties to be addressed which may have influenced the result and the data briefly described above can only be described as preliminary. For example, it would be prudent to continue this line of research and

make conclusions about MICs for yeast cells in which Pft1 expression is confirmed and quantified. If a transport system is observed then it may be possible to study the subcellular localisation of transport, e.g. whether the transport occurs at the cell membrane directly or whether transport of solute results in sequestration into the vesicle transport system and the intracellular concentration of drug is reduced by exocytosis. While it was hoped that such investigations would form part of our characterisation of Pft1, it was found that to advance this aspect to a conclusive degree was beyond the timescale of this thesis. However, the molecular infrastructure for such experiments is now in place and it is hoped that this avenue will be advanced in the short term.

Chapter 8

Final Discussion

In the present climate of complete genome sequencing, the practice of searching for, cloning and sequencing individual genes might be considered less valid. This is true of course for those involved in the study of organisms where a genome sequencing project is underway and will shortly be complete. However, a number of important microorganisms, including *P.brasiliensis*, have not yet been afforded a genome sequencing program. This oversight does not mean that molecular investigations of such organisms are not valid. Indeed the opposite should be considered true. As drug discovery in the 21st century tends towards rational drug design rather than the high throughput screening of naturally occurring substances, it is important that science provides a contemporary literature base for all organisms of clinical importance. To this end the search for potentially important determinants must continue until all genomes have been sequenced. With the necessity for entry level molecular characterisation in mind, the subject of this thesis was conceived and developed. At the time of writing there are a total of 13 *P.brasiliensis* genes available on the NCBI database. Fewer were available at the outset of the research. This total is rather small when you consider the 3000 ORFs identified in *S.cerevisiae* (Decottignies and Goffeau, 1999).

ATP-binding cassette transporters are among the most ubiquitous of proteins found in nature and the superfamily is now known to be responsible for a number of essential and important physiological functions. In microorganisms these functions include solute uptake (Hung, *et al.*, 1998), virulence factor export (Young and Holland, 1999) and multi-drug resistance (Higgins, 1992). ABC transporters have been identified as candidates for drug targets in both the infectious and physiological disease context and this feasibility is reflected in the intense global interest in the field. With the exception of the two genes identified by this study no ABC transporters are known for *P.brasiliensis*. The universal nature of these proteins would tend to assure that they would be present.

The isolation of the *pft1* and *pft2* gene fragments by PCR confirmed the effectiveness of the degenerate oligonucleotide primers for initial gene identification. This technique has since been used to good effect in other projects within our lab to discover a number of genes, including loci potentially involved in the thermal dimorphism of *P.brasiliensis* (unpublished data). Each set of degenerate primers were designed to be specific for the subclass of ABC transporters concerned; MDR1 and MDR2 targeting P-glycoprotein

(*mdr*-) like genes while PDR1 and PDR2 were intended for PDR-like genes. While others had found that the application of the MDR primers could lead to the amplification of regions from more than one gene, only a single product was observed in our experiments. In retrospect the gene corresponding to this amplicon was the exception to the rule in that Tobin *et al.* isolated a full-sized ABC transporter from both of the *Aspergillus* ssp. they addressed The half transporter, *afuMDR2*, was found in *A.fumigatus* alone (Tobin *et al.*, 1997). In *P.brasiliensis* we could find no evidence of a second gene, full-sized or otherwise, using this technique. The tendency of our degenerate PCR experiments to identify only one gene per primer set carried on in the application of the PDR primers. This primer set should have amplified DNA from coding regions from a number of similar ABC transporters within the genomes of both *S.cerevisiae (pdr5, pdr10, pdr12,* etc.) and *C.albicans* (*cdr1, cdr2, cdr3*). But, in our experiments, only one *pdr*-like amplicon was generated, even over a range of PCR conditions. While we have addressed both of the major classes of ABC transporters it is most likely that *pft1* and *pft2* do not represent a complete inventory of *P.brasiliensis* ABC transporters. It is possible that the degeneracies chosen for these primers sets, as well as the PCR conditions in the initial stages of amplification have selected for the preferential amplification of these fragments in each PCR. However, experiments that altered the degeneracies of the MDR primers and the initial conditions used in this PCR had no effect on the amplicon generated. The *pft1* fragment prevailed even when primer mismatch was required for oligo annealing.

Identification of further ABC transporters in *P.brasiliensis* remains a valid avenue of research. The findings of this study may allow access to this endeavour. The NBDs of both genes show considerable homology to most recognised ABC transporter genes of other fungi. This conservation would allow their employment as heterologous probes for further library screening. The procedures employed for homologous probing of libraries in the search for a lambda clone harbouring the *pft1* and *pft2* genes could be easily adapted to allow heterologous probing. A second option is to approach the issue from the level of phenotype rather than genotype. The construction of an expression library in a yeast vector system could allow screening for typical ABC-transporter mediated phenotypes in an appropriate yeast strain. For example, the drug hypersensitive *S.cerevisiae* strain AD1-8^{-} would be a good host for such expression plasmids if MDR was the phenotype in question. An added advantage of this approach is that a functional protein is assured. No truncation of the gene has occurred that would necessitate the additional demanding PCRs required in this study. Also, no genes would be isolated which represented a single component of an oligomeric complex, although it does not follow that this is a good thing. However, it is always encouraging to know that you have a complete experimental system when designing

future experiments. Unfortunately, the expression library approach requires a reasonable quantity of total cDNA to allow library construction. For logistical reasons this material was not available during the course of this study.

It is important that the identification of further ABC transporters is carried out within a framework of directed research. The work reported in this thesis demonstrates the existence of such genes in *P.brasiliensis*. In order to avoid a "stamp collecting" exercise further investigations should be motivated from the point of view of phenotypic investigation. This makes the second approach, using yeast expression libraries, the most appropriate.

Work at the mRNA level was more extensive than originally intended, primarily because of the need for the 5'RACE RT-PCR in the case of *pft1*. The results for this PCR were rather confusing at first and this prompted a considerable amount of repetition and optimisation. The observed result, indicating a shorter transcript that expected, persisted throughout and it had to be concluded that this was the mRNA species present. This highlights a fundamental issue in the characterisation of genes based on the ORFs observed on DNA sequencing. It is common practice to define the 5' boundary of a gene based on the "first ATG after a stop" philosophy. The prolific post-transcriptional modification and alternative transcriptional start sites observed in higher eukaryotes should serve as a warning that the true boundaries of a gene are not those depicted by the DNA sequence but those exhibited by the transcript on arrival at the ribosome. While this thesis yields to conventional wisdom and sets the first ATG of the genome at nucleotide positions 1-3 of *pft1*, it is just as possible that the ATG at position 550-552 is a more likely start codon for this is the one that has been identified experimentally.

Other mRNA experiments, such as the RT-PCRs examining transcriptional regulation, resulted in the bias for *pft1* displayed in this thesis. The genetic and biochemical characterisation of Pdr5 and Pdr5-like proteins has been carried out to great extent by a number of excellent researchers and the pursuit of data for Pft2 would probably involve a repetition of their experiments and merely reinforce paradigms already firmly in place. Half-transporters remained uninvestigated to any notable degree and we had already uncovered ambiguity between our observations and accepted wisdom. The apparent neglect of these smaller proteins permits easier access to original research thus Pft1 was given priority.

The generation of a Pft1 cDNA allowed progression to recombinant protein expression, and thus the potential for downstream biochemistry was realised. The procedures

employed to obtain protein were standard. As a consequence of the limited number of genes sequenced for *P.brasiliensis*, the recombinant expression of proteins from this organism is not reported frequently in the literature. Cisaplipino *et al.* (1996) successfully expressed the C-terminal portion of the major *P.brasiliensis* diagnostic antigen gp43. This was then transferred to nitro-cellulose rather than purified and reacted with anti-gp43 sera with a view to preliminary epitope mapping. To our knowledge the recombinant expression and purification of Pft1NBDTW represents the first isolation of protein, derived from this organism, for biochemical analysis.

While the expression and purification procedures employed gave satisfactory results there is now scope for further optimisation. The aggressive induction of the pET system is perhaps not the ideal in this case. From the large amount of inclusion bodies observed for both Pft1NBDT and Pft1NBDTW, it is apparent that protein folding remains a problem in the *E.coli* system. Alternative vector systems could be considered. For example a higher level of transcriptional control is afforded by expression vectors like pBAD (Invitrogen) in which the degree of indication can be controlled and optimised by the empirical titration of arabinose to the expression culture (pBAD vectors express from a promoter induced by arabinose rather than IPTG). These alternative strategies may allow a higher yield of soluble protein from Pft1NBDTW expressing cultures.

The allocation of function to Pft1 proved to be rather illusive at first. The problems associated may be the reason for the lack of published information on these proteins. At the time of investigation the fascinating data regarding mitochondrial peptide transport by Mdl1 (Young *et al.*, 2001) was unavailable. The discoverer of the *mdl* half-transporters of *S.cerevisiae* reported that these genes have been supplied to a number of other investigators who were keen to include them in there own research technologies, but these endeavours have never come to fruition in the literature (Paul Dean, personal communication). Most of the approaches were based on heterologous complementation in yeasts of various genotypes, an approach that was not particularly effective in this study. In light of the very recent data, the possibility that Pft1 is involved in mitochondrial processes must now be addressed. It will be interesting to see how the binding of ketoconazole fits in with this activity. Drug binding by Pft1 may represent a totally unrelated process giving the protein a pleiotropic role. The investigation that ascribed function to Mdl1 reported that the protein was found in isolated mitochondria. Other components of the cell, including the cell membrane, were not investigated as mitochondrial function was the initial focus of the study. The truncation of the *pft1* mRNA observed in 5'RACE may be involved in a process of pleiotropy as the exclusion of the 5' end of the transcript might remove any

mitochondrial localisation signal and allow a different Pft1 isoform to progress to the cytoplasmic membrane. Alternatively Pft1 may sequester azole drugs into vesicles during intracellular transport and reduce the free concentration in the cytoplasm. Or perhaps Pft1 does not transport azoles at all but is modulated by them by interactions at the NBD. All these possibilities will require further investigation and promise an interesting insight to the processes of these proteins.

The alternative approach of detailed biochemical characterisation at the level of the protein has proved more appropriate in the case of Pft1. The experiments reported in Chapter 7 demonstrate that the expression of domains homologous to those successfully employed in the study of mammalian proteins are also effective in fungal research. While the binding of ATP to Pft1NBDTW is of no great surprise the results concerning ketoconazole are remarkable. The fact that this ABC transporter has an affinity for this drug does not confirm that it is a substrate for translocation but does suggest that it could be. Alternatively this affinity may afford some allosteric control of activity as discussed in the previous chapter.

A detailed study of Pft1NBDTW mechanism is now possible. The dissection of binding kinetics for other ABC transporter nucleotide-binding domains, such as ArsA, illustrate the depth of information that can be gathered from comprehensive experimentation. The competence of the NBD for interactions with other azole drugs and hydrophobic compounds should be a priority. The potential that the NBD was processive and could hydrolyse ATP was not addressed in this study for technical reasons. The optimised purification system for Pft1NBDTW was based on phosphate buffers which prohibits direct analysis of phosphate ion release as a consequence of ATP hydrolysis. An alternative purification procedure was attempted, based on MOPs buffers, but no protein was obtained in this way. It may be possible to purify the protein in phosphate buffer and dialyse into MOPs or another phosphate free buffer for this analysis.

Despite no demonstration of a transport process *per se* this thesis reports strong evidence for the involvement of Pft1 in the physiological response of *P.brasiliensis* to ketoconazole. This is an important observation as ketoconazole is the therapeutic drug of choice for paracoccidiodiomycosis and these findings may have significant ramifications for the clinical efficacy of this drug. Other hydrophobic ligands may also prove to have a high affinity for this protein. One of the complicating clinical issues associated with the MDR phenotype is the competence of ABC transporters for a wide range of molecules that appear structurally unrelated. Recently the Cdr1 protein of *C.albicans* proven to be a mediator of clinical azole resistance (Sanglard, *et al.*, 1995)

was shown to transport the female hormone oestrogen (Krishnamurthy, *et al.*, 1998). A similar scenario in *P.brasiliensis* could have a deleterious impact. It has been demonstrated that the morphological transition of this organism, which is required for progression to the disease state, is inhibited in females due to such hormones. If the organism is developing a mechanism that will circumvent the intracellular accumulation of this hormone then the protection afforded to females may be alleviated, increasing the susceptible population.

In summary, the results in this thesis reveal the existence of two genetic determinants in *P.brasiliensis* that appear to encode ABC transporters. One of these determinants, termed *pft1*, has been completely sequenced to reveal a 2627bp open reading frame coding for 827 amino acids and interrupted by two introns of 68 and 77bp. The gene exhibits highest sequence homology to *afuMDR2* of *A.fumigatus* (Tobin, *et al.*, 1997). The second gene, *pft2* was partially sequenced. The portion sequenced amount to 3792bp coding 1053 amino acids covering the bulk of the gene (as estimated on comparison with the nearest homologue AtrA (Del Sorbo, *et al.*, 1997)). Both of these genes were found to be expressed loci in *P.brasiliensis*, with a total induction of *pft1* in response to fluconazole. The *pft2* transcript showed no notable up-regulation in the presence or absence of this MDR substrate. Analysis of the 5' region of the *pft1* transcript reveals a shorter than expected transcript, implicating the ATG at position 550 as the true start codon.

Recombinant expression of various C-terminal portions of the Pft1 nucleotide-binding domain was achieved in the pET21b vector. The largest of these constructs expressed Pft1NBD, a 48.5kDa protein. On analysis of the subcellular localisation the protein was found to be exclusively localised to the membrane fraction of the cell. Smaller truncates of the NBD were expressed and the tropism for the membrane was averted. Unfortunately the extensive over-expression of these proteins resulted in the predominate formation of inclusion bodies. Despite a number of attempts to manipulate the dynamics of expression and increase the proportion of the protein in the native soluble form, a comparatively small amount of protein could be observed in the cytoplasmic fraction. Even so, the purification of one of these truncated nucleotide-binding domains, the 28.0kDa Pft1NBDTW, was successfully achieved under native conditions using an empirically optimised protocol. Following dialysis this protein was applied to biochemical characterisation using tryptophan fluorescence based techniques. Transient analysis of MANT-ATP binding, for the perspective of both tryptophan fluorescence and MANT- fluorescence illustrated that the Pft1NBDTW exhibits an affinity for ATP. The highly efficient fluorescence energy transfer from the excited tryptophan of the protein to ketoconazole allowed the measurement of the binding

kinetics of this drug to Pft1NBDTW. A titration of the drug over a range of concentrations allowed the evaluation of K_d value that was measured at 0.959μM which increased to 0.858 μM on the addition of 10mM ATP.

References

Abele, R. and R. Tampe (1999). "Function of the transport complex TAP in cellular immune recognition." Biochim Biophys Acta **1461**: 405-419.

Aguilar-Bryan, L., C. G. Nichols, *et al.* (1995). "Cloning of the beta-cell high-affinity sulfonylurea receptor - a regulator of insulin-secretion." Science **268**: 423-426.

Allikmets, R., W. H. Raskind, *et al.* (1999). "Mutation of a putative mitochondrial iron transporter gene (ABC7) in X-linked sideroblastic anemia and ataxia (XLSA/A)." Hum Mol Genet **8**: 743-9.

Almeida, S. R., C. S. Unterkircher, *et al.* (1998). "Involvement of the major glycoprotein (gp43) of *Paracoccidioides brasiliensis* in attachment to macrophages." Med Mycol **36**: 405-11.

Ames, G. F. and J. Lever (1970). "Components of histidine transport: histidine-binding proteins and HisP protein." Proc Natl Acad Sci U S A **66**: 1096-103.

Ames, G. F. and K. Nikaido (1978). "Identification of a membrane protein as a histidine transport component in *Salmonella typhimurium*." Proc Natl Acad Sci U S A **75**: 5447-51.

Ames, G. F.-L. (1972). Components of histidine transport. Biological membranes. Proceedings of the 1972 ICN-UCLA Symposium in Molecular Biology. C. F. Fox. New York, Academic Press.

Andrade, A. C., G. Del Sorbo, *et al.* (2000). "The ABC transporter AtrB *from Aspergillus nidulans* mediates resistance to all major classes of fungicides and some natural toxic compounds." Microbiol **146**: 1987-97.

Angermayr, K., W. Parson, *et al.* (1999). "Expression of atrC - encoding a novel member of the ATP binding cassette transporter family in *Aspergillus nidulans* - is sensitive to cycloheximide." Biochim Biophys Acta **1453**: 304-10.

Angulo-Artega, A. and L. Pollack (1972). Paracoccidiodiomycosis. The pathological anatomy of the mycosis. Human infections with fungi, actinomycetes and algae. R. D. Baker. Berlin, Springer-Verlag: 507-560.

Arango, M. and A. Restrepo (1976). "Determination of the growth curves of the mycelial and yeast forms of *Paracoccidiodes brasiliensis*." Mycopathol **59**: 163-169.

Aristizabal, B. H., K. V. Clemons, *et al.* (1998). "Morphological transition of *Paracoccidioides brasiliensis* conidia to yeast cells: in vivo inhibition in females." Infect Immun **66**: 5587-91.

Bagagli, E., A. Sano, *et al.* (1998). "Isolation of *Paracoccidioides brasiliensis* from armadillos (*Dasypus noveminctus*) captured in an endemic area of paracoccidioidomycosis." Am J Trop Med Hyg **58**: 505-12.

Baichwal, V., D. Liu, *et al.* (1993). "The ATP-binding component of a prokaryotic traffic ATPase is exposed to the periplasmic (external) surface." Proc Natl Acad Sci U S A **90**: 620-4.

Balan, I., A. M. Alarco, *et al.* (1997). "The *Candida albicans* CDR3 gene codes for an opaque-phase ABC transporter." J Bacteriol **179**: 7210-8.

Balzi, E., W. Chen, *et al.* (1987). "Multidrug resistance gene *PDR1* from *Saccharomyces cerevisiae*." J Biol Chem **312**: 16871-16879.

Balzi, E. and A. Goffeau (1995). "Yeast multidrug resistance: the PDR network." J Bioenerg Biomembr **27**: 71-6.

Balzi, E., M. Wang, *et al.* (1994). "PDR5, a novel yeast multidrug resistance conferring transporter controlled by the transcription regulator PDR1." J Biol Chem **269**: 2206-14.

Bauer, B. E., H. Wolfger, *et al.* (1999). "Inventory and function of yeast ABC proteins: about sex, stress, pleiotropic drug and heavy metal resistance." Biochim Biophys Acta **1461**: 217-236.

Beaudet, L. and P. Gros (1995). "Functional Dissection of P-glycoprotein Nucleotide-binding Domains in Chimeric and Mutant Proteins." J Biol Chem **270**: 17159-17170.

Bergher, E. A. and L. A. Heppel (1974). "Different Mechanisms of Energy Coupling for the Shock-Sensitive and Shock-Resistant Amino-Acid Permeases of *Escherichia coli*." J Biol Chem **249**: 7747-7755.

Berkower, C. and S. Michaelis (1991). "Mutational analysis of the yeast a-factor transporter STE6, a member of the ATP binding cassette (ABC) protein superfamily." Embo J **10**: 3777-85.

Berkower, C., D. Taglicht, *et al.* (1996). "Functional and physical interactions between partial molecules of STE6, a yeast ATP-binding cassette protein." J Biol Chem **271**: 22983-9.

Bhattacharjee, H. and B. P. Rosen (1996). "Spatial proximity of Cys113, Cys172, and Cys422 in the metalloactivation domain of the ArsA ATPase." J Biol Chem **271**: 24465-70.

Bialek, R., A. Ibricevic, *et al.* (2000). "Detection of *Paracoccidioides brasiliensis* in tissue samples by a nested PCR assay." J Clin Microbiol **38**: 2940-2.

Birnboim, H. C. and J. Doly (1979). "A Rapid Alkaline Lysis Procedure for Screening Recombiant Plasmid DNA." Nuc Acid Res **7**: 1513-1523.

Bissinger, P. H. and K. Kuchler (1994). "Molecular cloning and expression of the *Saccharomyces cerevisiae STS1* gene product." J Biol Chem **269**: 4180-4186.

Bolhuis, H., H. W. van Veen, *et al.* (1997). "Mechanisms of multidrug transporters." FEMS Microbiol Rev **21**: 55-84.

Borba, C. D., M. D. S. L. Meirelles, *et al.* (1999). "Ultrastructural differences observed in *Paracoccidiodes brasiliensis* yeast phase cells grown on solid and liquid medium." Mycoses **42**: 281-283.

Borges-Walmsley, M. I. and A. R. Walmsley (2000). "cAMP signalling in pathogenic fungi: control of dimorphic switching and pathogenicity." Trend Microbiol **8**: 133-141.

Borst, P., R. Evers, *et al.* (1999). "The mutidrug resistance protein family." Biochim Biophys Acta **1461**: 347-357.

Bozdech, Z., U. Delling, *et al.* (1996). "Cloning and sequence-analysis of a novel member of the ATP-binding cassette (ABC) protein gene family from *Plasmodium falciparum*." Mol Biochem Parasitol **81**: 41-51.

Braibant, M., P. Lefevre, *et al.* (1996). "A Mycobacterium tuberculosis gene cluster encoding proteins of a phosphate transporter homologous to the *Escherichia coli* Pst system." Gene **176**: 171-6.

Broccardo, C., M.-F. Luciani, *et al.* (1999). "The ABCA subclass of mammailian transporters." Biochim and Biophys Acta **1461**: 395-404.

Broughton, M. C., M. Bard, *et al.* (1991). "Polyene resistance in ergosterol producing strains of *Candida albicans*." Mycosis **34**: 75-83.

Brummer, E., E. Castadena, *et al.* (1993). "Paracoccidiodiomycosis: an Update." Clin Microbiol Rev **6**: 89-117.

Brummer, E., A. Restrepo, *et al.* (1990). "Virulence of *Paracoccidiodes brasiliensis*: the influence of in vitro passage and storage." Mycopathologia **109**: 13-7.

Brummer, E., S. H. Sun, *et al.* (1990). "Ultrastructure of phagocytosed *Paracoccidiodes brasiliensis* in nonactivated or activated macrophages." Infect Immun **58**: 2628-2636.

Calcagno, A. M., G. Nino-Vega, *et al.* (1998). "Geographic discrimination of *Paracoccidioides brasiliensis* strains by randomly amplified polymorphic DNA analysis." J Clin Microbiol **36**: 1733-6.

Campos, E. P., C. R. Padovani, *et al.* (1991). "Paracoccidiodiomycose: estudo radiologico e pulminor de 58 casos" Rev Inst Med Trop Sao Paulo **33**: 267-276.

Cano, L. E., S. S. Kashino, *et al.* (1998). "Protective role of gamma interferon in experimental pulmonary paracoccidioidomycosis." Infect Immun **66**: 800-6.

Cano, L. E., L. M. Singer-Vermes, *et al.* (2000). "Depletion of CD8(+) T cells in vivo impairs host defense of mice resistant and susceptible to pulmonary paracoccidioidomycosis." Infect Immun **68**: 352-9.

Chen, C., J. E. Chin, *et al.* (1986). "Internal Duplication and Homology with Bacterial Transport Proteins in the *mdr1* (P-Glycoprotein) Gene from Multidrug-Resistant Human Cells." Cell **47**: 381-389.

Chen, J., S. Sharma, *et al.* (2001). "Trapping the transition state of an ATP-binding cassette transporter: Evidence of a concerted mechanism of maltose transport." Proc Natl Acad Sci U S A **98**: 1525-1530.

Cheung, M. and M. H. Akabas (1996). "Identification of the Cystic Fibrosis Transmembrane Conductance Regulator Channel-Lining Residues In and Flanking the M6 Membrane -Spanning Segment." Biophysical Journal **70**: 2688-2695.

Cisalpino, P. S., R. Puccia, *et al.* (1996). "Cloning, characterisation and epitope expression of the major diagnostic antigen of *Paracoccidiodes brasiliensis*." J Biol Chem **271**: 4553-4560.

Cole, S. P. C., G. Bhardwaj, *et al.* (1992). "Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line." Science **258**(1650-1654).

Conseil, G., A. Decottignies, *et al.* (2000). "Prenyl-flavonoids as potent inhibitors of the Pdr5p multidrug ABC transporter from Saccharomyces cerevisiae." Biochem **39**: 6910-7.

Consiel, G., H. Baubichon-Cortay, *et al.* (1998). "Flavonoids: A class of modulators with bifunctional interactions at the vicinal ATP- and steroid-binding sites on mouse P-glycoprotein." Proc Natl Acad Sci U S A **95**: 9831-9836.

Consortium, T. C. e. S. (1998). "Genome sequence of the nematode *C.elegans*: a platform for investigating biology." Science **282**: 2012-2018.

Conti Diaz, L. and F. D. Rilla (*1989*). "Hipotesis sobrenico ecologico de *Paracoccidiodes brasiliensis*." Med Rev Uraguay **5**: 97-103.

Cowen, L. E. and D. Sanglard, (2000). "Evolution of Drug Resistance in Experimental Populations of *Candida albicans*." J Bacteriol **182**: 1515-1522.

Csere, P., R. Lill, *et al.* (1998). "Identification of a human mitochondrial ABC transporter, the functional orthologue of yeast Atm1p." FEBS Lett **441**: 266-70.

Cui, Z., D. Hirata, *et al.* (1999). "Functional analysis of the promoter of the yeast *SNQ2* gene encoding a multidrug resistance transporter that confers the resistance to 4-nitroquinoline N-oxide." Biosci Biotechnol Biochem **63**: 162-7.

Cui, Z., D. Hirata, *et al.* (1996). "The multidrug resistance-associated protein (MRP) subfamily (Yrs1/Yor1) of *Saccharomyces cerevisiae* is important for the tolerance to a broad range of organic anions." J Biol Chem **271**: 14712-6.

Cui, Z., T. Shiraki, *et al.* (1998). "Yeast gene YRR1, which is required for resistance to 4-nitroquinoline N- oxide, mediates transcriptional activation of the multidrug resistance transporter gene SNQ2." Mol Microbiol **29**: 1307-15.

Cummings, N. J. and I. F. Connerton (1997). "The *Bacillus subtilis* 168 chromasome from *sspE* to *katA*." Microbiology **143**: 1855-1859.

Cunha, A. F., M. V. Sousa, *et al.* (1999). "Identification, N-terminal region sequencing and similarity analysis of differentially expressed proteins in *Paracoccidiodes brasiliensis*." Med Mycol **37**: 115-121.

Da Silva, S. P., M. I. Borges-Walmsley, *et al.* (1999). "Differential expression of an *hsp*70 gene during the morphlogical transition from the mycelial to the yeast form of the human pathogenic fungus *Paracoccidiodes brasiliensis*." Mol Microbiol **31**: 1039-1050.

Da Silva, S. P., M. S. S. Felipe, *et al.* (1994). "Phase transition and Stage-Specific Protein Synthesis in the Dimorphic Fungus *Paracoccidiodes brasiliensis*." Exp Mycol **19**: 294-299.

Dayan, G., H. Baubichon-Cortay, *et al.* (1996). "Recombinant N-terminal Nucleotide-binding Domain from Mouse P-glycoprotein." J Biol Chem **271**: 11652-11658.

De Almeida, S. R., J. Z. De Moraes, *et al.* (1998). "Pattern of immune response to GP43 from *Paracoccidioides brasiliensis* in susceptible and resistant mice is influenced by antigen-presenting cells." Cell Immunol **190**: 68-76.

De Bedout, C., L. E. Cano, *et al.* (1986). "Water as a substrate for the development of *Paracoccidiodes brasiliensis* mycelial form." Mycopathol **96**: 123-130.

De Brito, T., G. S. Sandhu, *et al.* (1999). "*In situ* hybridisation in paracoccidiodiomycosis." Med Mycol **37**: 207-211.

Dean, D. A., A. L. Davidson, *et al.* (1990). "The Role of ATP as the Energy Source for Maltose Transport in *Escherichia coli*." Res in Microbiol **141**: 348-352.

Dean, M., R. Allikmets, *et al.* (1994). "Mapping and sequencing of two yeast genes belonging to the ATP-binding cassette superfamily." Yeast **10**: 377-83.

Decottignies, A. and A. Goffeau (1999). "Complete inventory of the yeast ABC proteins." Nat Genetics **15**: 137.

Decottignies, A., L. Lambert, *et al.* (1995). "Identification and characterization of SNQ2, a new multidrug ATP binding cassette transporter of the yeast plasma membrane." J Biol Chem **270**: 18150-7.

Del Sorbo, G., A. C. Andrade, *et al.* (1997). "Multidrug resistance in *Aspergillus nidulans* involves novel ATP-binding cassette transporters." Mol Gen Genet **254**: 417-26.

Dexter, D., W. S. Moye-Rowley, *et al.* (1994). "Mutations in the Yeast *PDR3, PDR4, PDR7* and *PDR9* Pleiotropic (Multiple) Drug Resistance Loci Affect the Transcript Level of an ATP Binding Cassette Transporter Encoding Gene, *PDR5*." Genet **136**: 505-515.

Diederiches, K., J. Doez, *et al.* (2000). "Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*." EMBO J **19**: 5951-5961.

Dong, M., F. Penin, *et al.* (1996). "Efficient Purification and Reconstitution of P-glycoprotein for Functional and Structural Studies." J Biol Chem **271**: 28875-28883.

Dudler, R. and C. Hertig (1992). "Structure of an *mdr*-like gene from *Arabidopsis thaliana*. Evolutionary implications." J Biol Chem **267**: 5882-5888.

Egner, R., B. E. Bauer, *et al.* (2000). "The transmembrane domain 10 of the yeast Pdr5p ABC antifungal efflux pump determines both substrate specificity and inhibitor susceptibility." Mol Microbiol **35**: 1255-63.

Egner, R. and K. Kuchler (1996). "The yeast multidrug transporter Pdr5 of the plasma membrane is ubiquitinated prior to endocytosis and degradation in the vacuole." FEBS Lett **378**: 177-81.

Egner, R., Y. Mahe, *et al.* (1995). "Endocytosis and vacuolar degradation of the plasma membrane-localized Pdr5 ATP-binding cassette multidrug transporter in *Saccharomyces cerevisiae*." Mol Cell Biol **15**: 5879-87.

Eytan, G. D., R. Regev, *et al.* (1996). "Functional Reconstitution of P-glycoprotein Reveals an Apparent Near Stoichiometric Drug Transport to ATP Hydrolysis." J Biol Chem **271**: 3172-3178.

Favre, B. and N. S. Ryder (1996). "Characterization of squalene epoxidase activity from the dermatophyte *Trichophyton rubrum* and its inhibition by terbinafine and other mycotic agents." Antimicrob Agents Chemother **40**: 443-447.

Feldman, D., D. Loose, *et al.* (1983). "A receptor-like estrogen binding protein in *Paracoccidiodes brasiliensis* (PB) - possible basis for lower infection rate in women than men." Clin Res **31**: A470.

Franco, M., E. Bagagli, *et al.* (1996). "*Paracoccidioides brasiliensis* antigen batches from the same isolate show immunological and biochemical differences." Mycopathol **135**: 13-9.

Franco, M., E. Bagagli, *et al.* (2000). "A critical analysis of isolation of *Paracoccidioides brasiliensis* from soil." Med Mycol **38**: 185-91.

Franco, M., A. Sano, *et al.* (1989). "Chlamydospore formation by *Paracoccidiodes brasiliensis* mycelial form." Rev Inst Med Trop Sao Paulo **31**: 151-157.

Franz, R., S. Michel, *et al.* (1998). "A fourth gene from the *Candida albicans* CDR family of ABC transporters." Gene **220**: 91-8.

Geller, D., D. Taglicht, *et al.* (1996). "Comparative topology studies in *Saccharomyces cerevisiae* and in *Escherichia coli*. The N-terminal half of the yeast ABC protein Ste6." J Biol Chem **271**: 13746-53.

Georgopapadakou, N. H. (1998). "Antifungals: mechanism of action and resistance, established and novel drugs." Curr Opin Microbiol **1**: 547-57.

Gileadi, U. and C. F. Higgins (1997). "Membrane topology of the ATP-binding cassette transporter associated with antigen presentation (Tap1) expressed in *Escherichia coli*." J Biol Chem **272**: 11103-8.

Goffeau, A., B. G. Barrell, *et al.* (1996). "Life with 6000 genes." Science **274**: 564.

Goldani, L. Z., M. Picard, *et al.* (1994). "Synthesis of heat-shock proteins in mycelia and yeast forms of *Paracoccidiodes brasiliensis*." J Med Microbiol **40**: 124-128.

Goldani, L. Z. and A. M. Sugar (1998). "Short report: use of the polymerase chain reaction to detect Paracoccidioides brasiliensis in murine paracoccidioidomycosis." Am J Trop Med Hyg **58**: 152-3.

Golin, J., A. Barkatt, *et al.* (2000). "Chemical specificity of the PDR5 multidrug resistance gene product of *Saccharomyces cerevisiae* based on studies with tri-n-alkyltin chlorides." Antimicrob Agents Chemother **44**: 134-8.

Gomes, G. M., P. S. Cisalpino, *et al.* (2000). "PCR for diagnosis of paracoccidioidomycosis." J Clin Microbiol **38**: 3478-80.

Goncalves, A. J., A. T. Londero, *et al.* (1998). "Paracoccidioidomycosis in children in the state of Rio de Janeiro (Brazil). Geographic distribution and the study of a "reservarea"." Rev Inst Med Trop Sao Paulo **40**: 11-3.

Gurr, S. J., S. E. Unkles, *et al.* (1987). The structure and function of nuclear genes of filamentous fungi. Gene Structure in Lower Eukaryotic Microbes. J. R. Kinghorn. Glasgow, IRL Press**:** 93-139.

Hadad, D. J., M. d. F. Pieres, *et al.* (1992). "*Paracoccidiodes brasiliensis* (Lutz 1908) isolated by hemoculture in a patient with the aquired immunodeficiency syndrome." Rev Inst Med Trop Sao Paulo **34**: 565-567.

Hallstrom, T. C., D. J. Katzmann, *et al.* (1998). "Regulation of transcription factor Pdr1p function by an Hsp70 protein in *Saccharomyces cerevisiae*." Mol Cell Biol **18**: 1147-55.

Hallstrom, T. C. and W. S. Moye-Rowley (1998). "Divergent transcriptional control of multidrug resistance genes in *Saccharomyces cerevisiae*." J Biol Chem **273**: 2098-104.

Hanna, S. A., J. L. Monteiro da Silva, *et al.* (2000). "Adherence and intracellular parasitism of *Paracoccidioides brasiliensis* in vero cells" Microbes Infect **2**: 877-84.

Heitman, J., A. Koller, *et al.* (1993). "The immunosuppresant FK506 inhibits amino-acid import in *Saccharmyces cerevisiae*." Mol Cell Biol **13**: 5010-5019.

Hernaez, M. L., C. Gil, *et al.* (1998). "Induced expression of the multidrug resistance gene CDR1 in response to fluconazole and other antifungals." Yeast **14**: 517-26.

Hettema, E. H., C. W. van Roermund, *et al.* (1996). "The ABC transporter proteins Pat1 and Pat2 are required for import of long-chain fatty acids into peroxisomes of *Saccharomyces cerevisiae*." Embo J **15**: 3813-22.

Higgins, C. F. (1992). "ABC TRANSPORTERS: From Microorganisms to Man." Ann Rev Cell Biol **8**: 67-113.

Higgins, C. F. (1995). "P-glycoprotein and cell volume-activated chloride channels." J Bioenerg Biomembr **27**: 63-70.

Higgins, C. F. and M. M. Gottesman (1992). "Is the multidrug transporter a flippase?" Trends Biochem Sci **17**: 18-21.

Higgins, C. F., P. D. Haag, *et al.* (1982). "Complete nucleotide sequence and identification of membrane components of the histidine transport operon of *S. typhimurium*." Nature **298**: 723-7.

Hirata, D., K. Yano, *et al.* (1994). "*Saccharomyces cerevisiae* YDR1, which encodes a member of the ATP- binding cassette (ABC) superfamily, is required for multidrug resistance." Curr Genet **26**: 285-94.

Hitchcock, C. A. (1993). "Resistance of *Candida albicans* to azole antifungal agents." Biochem Soc Trans **21**: 1039-47.

Ho, H.-C. and T.-H. Liao (1999). "Protein structure and gene cloning of *Syncephalastrum racemosum* nuclease." Biochemical Journal **339**: 261-267.

Hung, L. W., I. X. Wang, *et al.* (1998). "Crystal structure of the ATP-binding subunit of an ABC transporter." Nature **396**: 703-7.

Hunt, A. G. and J.-S. Hong (1983). "Properties and characterization of binding-protein-dependant active transport of glutamine in isolate membrane vesicles of *Escherichia coli.*" Biochemistry **22**: 844-850.

Imai, T., A. Sano, *et al.* (2000). "A new PCR primer for the identification of Paracoccidioides brasiliensis based on rRNA sequences coding the internal transcribed spacers (ITS) and 5 x 8S regions." Med Mycol **38**: 323-6.

Jia, J., P. S. Dyer, *et al.* (1999). "Cloning of the *cbhi* and *cbhii* genes invloved in cellulose utilisation by the straw mushroom *Volvariella volvacea.*" Mol Gen Genet **261**: 985-993.

Jones, P. M. and A. M. George (1999). "Subunit interactions in ABC transporters: towards a functional architecture." FEMS Microbiol Lett **179**: 187-202.

Jones, P. M. and A. M. George (2000). "Symmetry and structure in P-glycoprotein and ABC transporters: What goes around comes around." European Journal of Biochemistry **267**: 5298-5305.

Juliano, R. L. and V. Ling (1976). "A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants." Biochim Biophys Acta **455**: 152-162.

Karttunen, J. T., J. Trowsdale, *et al.* (1999). "Antigen presentation: TAP dances with ATP." Curr Biol **9**: R820-4.

Katzmann, D. J., P. E. Burnett, *et al.* (1994). "Transcriptional control of the yeast PDR5 gene by the PDR3 gene product." Mol Cell Biol **14**: 4653-61.

Katzmann, D. J., T. C. Hallstrom, *et al.* (1996). "Multiple Pdr1p/Pdr3p binding sites are essential for normal expression of the ATP binding cassette transporter protein-encoding gene PDR5." J Biol Chem **271**: 23049-54.

Kaur, P. (1999). "The anion-stimulated ATPase ArsA shows unisite and multisite catalytic activity." J Biol Chem **274**: 25849-54.

Kelly, A., S. H. Powis, *et al.* (1992). "TAP1/2." Nature **355**: 614-644.

Kerppola, R. E., V. K. Shyamala, *et al.* (1991). "The membrane-bound proteins of periplasmic permeases form a complex. Identification of the histidine permease HisQMP complex." J Biol Chem **266**: 9857-65.

Klein, I., B. Sarkadi, *et al.* (1999). "An inventory of the human ABC proteins." Biochim et Biophys Acta **1461**: 237-262.

Kolaczkowski, M., A. Kolaczowska, *et al.* (1998). "*In vivo* characterization of the drug resistance profile of the major ABC transporters and other components of the yeast pleiotropic drug resistance network." Microb Drug Resist **4**: 143-58.

Kolaczkowski, M., M. van der Rest, *et al.* (1996). "Anticancer drugs, ionophoric peptides, and steroids as substrates of the yeast multidrug transporter Pdr5p." J Biol Chem **271**: 31543-8.

Krishnamurthy, S., V. Gupta, *et al.* (1998). "Characterisation of human steroid hormone transport mediated by Cdr1p, a multidrug transporter of *Candida albicans*, belonging to the ATP-binding cassette superfamily." FEMS Microbiol Lett **158**: 69-74.

Kuchler, K., R. E. Sterne, *et al.* (1989). "*Saccharomyces cerevisiae* STE6 gene product: a novel pathway for protein export in eukaryotic cells." Embo J **8**: 3973-84.

Kurita, N., S. K. Biswas, *et al.* (1999). "Fungistatic and fungicidal activities of murine polymorphonuclear leucocytes against yeast cells of *Paracoccidioides brasiliensis*." Med Mycol **37**: 19-24.

Kurita, N., M. Oarada, *et al.* (1999). "Antifungal activity of human polymorphonuclear leucocytes against yeast cells of *Paracoccidioides brasiliensis*." Med Mycol **37**: 261-7.

Kurita, N., M. Oarada, *et al.* (2000). "Effect of cytokines on antifungal activity of human polymorphonuclear leucocytes against yeast cells of *Paracoccidioides brasiliensis*." Med Mycol **38**: 177-82.

Laemmli, U. K. (1970). "Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4." Nature **227**: 680-685.

Leighton, J. and G. Schatz (1995). "An ABC transporter in the mitochondrial inner membrane is required for normal growth of yeast." Embo J **14**: 188-95.

Li, J., S. Liu, *et al.* (1996). "Interaction of ATP binding sites in the ArsA ATPase, the catalytic subunit of the Ars pump." J Biol Chem **271**: 25247-52.

Linton, K. J. and C. F. Higgins (1998). "The *Escherichia coli* ATP binding cassette (ABC) proteins." Mol Microbiol **28**: 5-13.

Liu, C. E., P. Q. Liu, *et al.* (1997). "Characterization of the adenosine triphosphatase activity of the periplasmic histidine permease, a traffic ATPase (ABC transporter)." J Biol Chem **272**: 21883-91.

Liu, P. Q. and G. F. Ames (1998). "In vitro disassembly and reassembly of an ABC transporter, the histidine permease." Proc Natl Acad Sci U S A **95**: 3495-500.

Lomovskaya, O. and K. Lewis (1992). "*emr*, an *Escherichia coli* locus for mdr." Proc Natl Acad Sci U S A **89**: 8938-8942.

Loo, T. W. and D. M. Clarke (1999). "The transmembrane domains of human multidrug resistance P-glycoprotein are sufficient to mediate drug binding and trafficking to the cell surface." J Biochem**274**: 24759-24765.

Loo, T. W. and D. M. Clarke (2000). "Determining the structure and mechanism of the human multidrug resistance P-glycoprotein using cysteine scanning mutagenesis and thiol-modification techniques." Biochim Biophys Acta **1461**: 315-325.

Loo, T. W. and D. M. Clarke (2000). "Drug-stimulated ATPase activity of Human P-glycoprotein is Blocked by Disulphide Cross-Linking between the Nucleotide-binding Sites." J Biol Chem **275**: 19434-19438.

Loose, D. S., E. P. Stover, *et al.* (1983). "Estradiol binds to a receptor-like cytosol binding protein and initiates a biological response in *Paracoccidiodes brasiliensis*." Proc Natl Acad Sci U S A **80**: 7659-7663.

Lopes, J. D., M. C. Moura-Campos, *et al.* (1994). "Characterisation of glycoprotein gp43, the major laminin-binding protein of *Paracoccidiodes brasiliensis*." Brazilian J MedBiol Res **27**: 2309-2313.

Lutz, A. (1908). "Uma mycose psudococcidioa localisida na bocca e observada no Brazil. Contribuicao ao conhecimento das hyphoblastomycoses americanas." Bras-Med **22**: 121-124.

Mahe, Y., Y. Lemoine, *et al.* (1996). "The ATP Binding Cassette Transporters PDR5 and Snq1 of *Saccharomyces cerevisiae* Can Mediate Transport of Steroids *in Vivo*." J Biol Chem **271**: 25167-25172.

Mahe, Y., A. Parle-McDermott, *et al.* (1996). "The ATP-binding cassette multidrug transporter Snq2 of *Saccharomyces cerevisiae*: a novel target for the transcription factors Pdr1 and Pdr3." Mol Microbiol **20**: 109-17.

Mangiaterra, M., J. Alonso, *et al.* (1996). "Histoplasmin and paracoccidioidin skin reactivity in infantile population of northern Argentina (1)." Rev Inst Med Trop Sao Paulo **38**: 349-53.

Mattar-Filho, R., M. O. Azevedo, *et al.* (1997). "Expression of glycoprotein gp43 in stage-specific forms and during dimorphic differentiation of *Paracocciodiodes brasiliensis*." J Med Vet Mycol **35**: 341-345.

McEwan, J. G., V. Bedoya, *et al.* (1987). "Experimental murine paracoccidiodiomycosis induced by inhalation of conidia." J Med Vet Mycol **25**: 165-175.

McEwan, J. G., B. L. Ortiz, *et al.* (1996). "Molecular Cloning and Characterisation of a 27kDa Antigenic Protein from *Paracoccidiodes brasiliensis*." J Biol Chem **271**: 4553-4565.

McEwan, J. G., B. L. Ortiz, *et al.* (1996). "Molecular Cloning, Nucloetide Sequencing, and characterization of a 27-kDa Antigenic Protein from *Paracoccidiodes brasiliensis*." Fun Genet Biol **20**: 125-131.

Medoff, E., A. Painter, *et al.* (1987). "Mycelial to yeast phase transitions of the dimorphic fungi *Blastomyces dermatitidis* and *Paracoccidiodes brasiliensis*." J Bacteriol **169**: 4055-4060.

Meyers, S. W., W. Schauer, *et al.* (1992). "Interaction of the yeast pleiotropic drug resistance genes *pdr1* and *pdr5*." Curr Genet **21**: 431-436.

Michaelis, S. and C. Berkower (1995). "Sequence comparison of yeast ATP-binding cassette proteins." Cold Spring Harb Symp Quant Biol **60**: 291-307.

Miyazaki, H., Y. Miyazaki, *et al.* (1998). "Fluconazole resistance associated with drug efflux and increased transcription of a drug transporter gene, PDH1, in *Candida glabrata*." Antimicrob Agents Chemother **42**: 1695-701.

Molinari-Madlum, E. E., M. S. Felipe, *et al.* (1999). "Virulence of *Paracoccidioides brasiliensis* isolates can be correlated to groups defined by random amplified polymorphic DNA analysis." Med Mycol **37**: 269-76.

Montenegro, M. R., M. Miyaji, *et al.* (1996). "Isolation of fungi from nature in the region of Botucatu, state of Sao Paulo, Brazil, an endemic area of paracoccidioidomycosis." Mem Inst Oswaldo Cruz **91**: 665-70.

Montoya, A. E., A. L. Alvarez, *et al.* (1999). "Electrophoretic karyotype of environmental isolates of *Paracoccidioides brasiliensis*." Med Mycol **37**: 219-22.

Morias, F. V., T. F. Barros, *et al.* (2000). "Polymorphism in the gene coding for the immunodominant gp43 from the pathogenic fungus *Paracoccidiodes brasiliensis*." J Clin Microbiol **38**: 3960-3966.

Morna, G. P., D. Sanglard, *et al.* (1998). "Identification and Expression of Multidrug Transporters Responsible for Fluconazole Resistance in *Candida dubliniensis*." Antimicrob Agents Chemother **42**:136-141.

Motoyama, A. B., E. J. Venancio, *et al.* (2000). "Molecular identification of Paracoccidioides brasiliensis by PCR amplification of ribosomal DNA." J Clin Microbiol **38**(8): 3106-3111.

Mourez, M., M. Hofnung, *et al.* (1997). "Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permases defines an important site of interaction with the ATPase subunits." EMBO J **16**: 3066-3077.

Mourez, M., M. Jehanno, *et al.* (1998). "In vitro interaction between components of the inner membrane complex of the maltose ABC transporter of *Escherichia coli*: modulation by ATP." Mol Microbiol **30**: 353-363.

Nagao, K., Y. Taguchi, *et al.* (1995). "bfr1+, a novel gene of *Schizosaccharomyces pombe* which confers brefeldin A resistance, is structurally related to the ATP-binding cassette superfamily." J Bacteriol **177**: 1536-43.

Nakaune, R., K. Adachi, *et al.* (1998). "A novel ATP-binding cassette transporter involved in multidrug resistance in the phytopathogenic fungus *Penicillium digitatum.*" Appl Environ Microbiol **64**: 3983-8.

Nikaido, K. and G. F. Ames (1999). "One intact ATP-binding subunit is sufficient to support ATP hydrolysis and translocation in an ABC transporter, the histidine permease." J Biol Chem **274**: 26727-35.

Nikaido, K., P. Q. Liu, *et al.* (1997). "Purification and characterization of HisP, the ATP-binding subunit of a traffic ATPase (ABC transporter), the histidine permease of *Salmonella typhimurium*. Solubility, dimerization, and ATPase activity." J Biol Chem **272**: 27745-52.

Nino-Vega, G. A., E. T. Buurman, *et al.* (1998). "Molecular Cloning and Sequencing of a Chitin Synthase Gene (CHS2) of *Paracoccidiodes brasiliensis.*" Yeast **14**: 181-187.

Nino-Vega, G. A., C. A. Munro, *et al.* (2000). "Differential expression of chitin synthase genes during temperature- induced dimorphic transitions in *Paracoccidioides brasiliensis.*" Med Mycol **38**: 31-9.

Nogueira Cano, M. I., P. S. Cisalpino, *et al.* (1998). "Electophoretic Karyotypes and Genome Sizing of the Pathogenic Fungus *Paracoccidiodes brasiliensis.*" J Clin Microbiol **36**: 742-747.

Parkinson, T., D. J. Falconer, *et al.* (1995). "Fluconazole Resistance Due to Energy-Dependent Drug Efflux in *Candida glabrata.*" Antimicrob Agents Chemother **39**: 1696-1699.

Pei-Qi, L. and G. F.-L. Ames (1998). "*In vitro* disassembly and reassembly of an ABC transporter, the histidine permease" Proc Natl Acad Sci U S A **95**: 3495-3500.

Pereira, M., M. S. Felipe, *et al.* (2000). "Molecular cloning and characterization of a glucan synthase gene from the human pathogenic fungus *Paracoccidioides brasiliensis.*" Yeast **16**: 451-62.

Perez-Victoria, J. M., F. J. Perez-Victoria, *et al.* (2001). "High-Affinity Binding of Silybin Derivatives to the Nucleotide-Binding Domain of a *Leishmania tropica* P-Glycoprotein-Like Transporter and Chemosensitization of a Multidrug-Resistant Parasite to Daunomycin." Antimicrob Agents Chemother **45**: 439-446.

Pinto, A. R., R. Puccia, *et al.* (2000). "DNA-based vaccination against murine paracoccidiodiomycosis using the gp43 gene from *Paracoccidiodes brasiliensis*." Vaccine **18**: 3050-3058.

Piper, P., Y. Mahe, et al. (1998). "The Pdr12 ABC transporter is required for the development of weak organic acid resistance in yeast." Embo J **17**: 4257-65.

Plate, C. A. (1979). "Requirement for Membrane Potential in the Active Transport of Glutamine in *Escherichia coli*." J Bacteriol **163**: 1237-1242.

Prasad, R., P. Dewergifosse, et al. (1995). "Molecular-cloning and characterization of a novel gene of *Candida albicans*, *CDR1*, conferring multiple resistance to drugs and antifungals." Curr Genet **27**: 320-329.

Prasad, R., S. K. Murthy, *et al.* (1995). "Multiple drug resistance in *Candida albicans*." Acta Biochim Pol **4**: 497-504.

Puccia, R., S. Schenkman, *et al.* (1986). "Exocellular components of *Paracoccidiodes brasiliensis*: indentification of a specific antigen." Infect Immun **53**: 199-206.

Quentin, Y. and G. Fichant (2000). "ABCdb: an ABC transporter database." J Mol Microbiol Biotechnol **2**: 501-4.

Raymond, M., D. Dignard, *et al.* (1998). "A Ste6p/P-glycoprotein homologue from the asexual yeast *Candida albicans* transports the a-factor mating pheromone in Saccharomyces cerevisiae." Mol Microbiol **27**: 587-98.

Restrepo, A., M. E. Salazar, *et al.* (1984). "Estrogens inhibit the mycelium to yeast transformation in the fungus *Paracoccidiodes brasiliensis*." Infect Immun **46**: 346-353.

Rosen, B. P., H. Bhattacharjee, *et al.* (1999). "Mechanism of the ArsA ATPase." Biochim Biophys Acta **1461**: 207-15.

Salem-Izacc, S. M., R. S. Jesuino, *et al.* (1997). "Protein synthesis patterns of *Paracoccidiodes brasiliensis* isolates in stage-specific forms and during cellular differentiation." J Med Vet Mycol **35**: 205-11.

Sambrook, J., E. F. Fritsch, *et al.* (1989). Molecular Cloning: A Laboratory Manual. New York, Cold Spring Harbor Laboratory Press.

San-Blas, F. (1986). "Ultrastructure of spore formation in *Paracoccidiodes brasiliensis*." J Med Vet Mycol **24:** 203-210.

San-Blas, G. and F. San-Blas (1986). "Effect of nucleotides on the glucan synthesis in *Paracoccidiodes brasilinsis*." J Med Vet Mycol **24**: 243-245.

San-Blas, G., F. San-Blas, *et al.* (1997). "Polyamines in growth and dimorphism of *Paracoccidiodes brasiliensis*." Arch Microbiol **166**: 411-413.

Sandhu, G. S., R. A. Aleff, *et al.* (1997). "Molecular detection and identification of *Paracoccidioides brasiliensis*." J Clin Microbiol **35**: 1894-6.

Sanglard, D., F. Ischer, *et al.* (2001). "Role of ATP-Binding-Cassette Transporter Genes in High-Frequency Acquisition of Resistance to Azole Antifungals in *Candida glabrata*." Antimicrob Agents Chemother **45**: 1174-1183.

Sanglard, D., F. Ischer, *et al.* (1999). "The ATP binding cassette transporter gene CgCDR1 from *Candida glabrata* is involved in the resistance of clinical isolates to azole antifungal agents." Antimicrob Agents Chemother **43**: 2753-65.

Sanglard, D., F. Ischer, *et al.* (1996). "Susceptibilities of *Candida albicans* multidrug transporter mutants to various antifungal agents and other metabolic inhibitors." Antimicrob Agents Chemother **40**: 2300-5.

Sanglard, D., F. Ischer, *et al.* (1997). "Cloning of *Candida albicans* genes conferring resistance to azole antifungal agents: characterization of CDR2, a new multidrug ABC transporter gene." Microbiol **143**: 405-16.

Sanglard, D., K. Kuchler, *et al.* (1995). "Mechanisms of resistance to azole antifungal agents in Candida albicans isolates from AIDS patients involve specific multidrug transporters." Antimicrob Agents Chemother **39**: 2378-86.

Sano, A., M. Miyaji, *et al.* (1992). "Studies on the relationship between the estrous cycle of BALB/c mice and their resistance to *Paracoccidioides brasiliensis* infection." Mycopathol **119**: 141-5.

Sano, A., R. Tanaka, *et al.* (1998). "Comparison between human and armadillo *Paracoccidioides brasiliensis* isolates by random amplified polymorphic DNA analysis." Mycopathol **143**: 165-9.

Saraiva, E. C., A. Altemani, *et al.* (1996). "*Paracoccidiodes brasiliensis*-gp43 used as paracoccidiodin." J Med Vet Mycol **34**: 155-161.

Saraste, M., P. R. Sibbald, *et al.* (1990). "The P-loop - a common motif in ATP- and GTP-binding proteins." Trend Biochem Sci **15**: 430-434.

Schneider, E. and S. Hunke (1998). "ATP-binding-cassette (ABC) transport systems: Functional and structural aspects of the ATP-hydrolyzing subunits/domains." Biochim Biophys Acta **22**: 1-20.

Shani, N., A. Sapag, *et al.* (1996). "An *S.cerevisiae* peroxisomal transporter, orthologous to the human adrenoleukodystrophy protein, appears to be a heterodimer of two half ABC transporters: Pxa1p and Pxa2p." Ann N Y Acad Sci **804**: 770-2.

Shani, N., P. A. Watkins, *et al.* (1995). "PXA1, a possible *Saccharomyces cerevisiae* ortholog of the human adrenoleukodystrophy gene." Proc Natl Acad Sci U S A **92**: 6012-6.

Shao, Y., T. A. Pressley, *et al.* (1999). "Na,K-ATPase mRNA B1 expression in rate myocardium - effect of thyroid status." European Journal of Biochemistry **260**: 1-8.

Sharom, F. J., R. Liu, *et al.* (1998). "Spectroscopic and biophysical approaches for studying the structure and function of the P-glycoprotein multidrug transporter." Biochem Cell Biol **76**: 695-708.

Sharom, F. J., L. Ronghua, *et al.* (1999). "Insights into the structure and substrate interactions of the P-glycoprotein multidrug resistance transporter from spectroscopic studies." Biochim Biophys Acta **1461**: 327-345.

Shiani, N. and D. Valle (1996). "A *Saccharomyces cerevisiae* homolog of the human adrenoleukodystrophy transporter is a heterodimer of two half ATP-binding cassette transporters." Proc Natl Acad Sci U S A **93**: 11901-11906.

Silva-Vergara, M. L., R. Martinez, *et al.* (2000). "Isolation of *Paracoccidioides brasiliensis* from armadillos (*Dasypus novemcinctus*) in an area where the fungus was recently isolated from soil." Med Mycol **38**: 193-9.

Silva-Vergara, M. L., R. Martinez, *et al.* (1998). "Isolation of a *Paracoccidioides brasiliensis* strain from the soil of a coffee plantation in Ibia, State of Minas Gerais, Brazil." Med Mycol **36**: 37-42.

Smit, L. S., D. J. Wilkinson, *et al.* (1993). "Functional roles of the nucleotide-binding folds in the activation of the cystic fibrosis transmembrane conductance regulator." Proc Natl Acad Sci U S A **90**: 9963-9967.

Sonneborn, A., D. P. Bockmuhl, *et al.* (2000). "Protein kinase A encoded by *TPK2* regulates dimorphism of *Candida albicans*." Mol Microbiol **35**: 386-396.

Stautnieks, S., L. Bull, *et al.* (1998). "A gene encoding a liver specific ABC transporter is mutated in progressive familial intrahepatic cholestasis." Nat Genetics **20**: 233-238.

Straus, A. H., E. Freymuller, *et al.* (1996). "Immunochemical and subcellular localization of the 43kDa glycoprotein antigen of *Paracoccidiodes brasiliensis*with monoclonal antibodies." J Med Vet Mycol **34**: 181-186.

Taborda, C. P., M. A. Juliano, *et al.* (1998). "Mapping of the T-cell epitope in the major 43-kilodalton glycoprotein of *Paracoccidioides brasiliensis* which induces a Th-1 response protective against fungal infection in BALB/c mice." Infect Immun **66**: 786-93.

Talibi, D. and M. Raymond (1999). "Isolation of a Putative *Candida albicans* Transcriptional Regulator involved in Pleiotropic Drug Resistance by Functional Complementation of a *pdr1 pdr3* Mutation in *Saccharomyces cervisiae*." Journal Bacteriol **181**: 231-240.

Tapia, F. J., M. Golhman-Yahr, *et al.* (1991). "Leukocyte immunophenotypes in bronchoalveolar lavage fluid in peripheral blood of paracoccidiodiomycosis sarcodosis and silicosis." Histol Histopathol **6**: 395-402.

Thornewell, S. J., R. B. Peery, *et al.* (1997). "Cloning and characterization of CneMDR1: a *Cryptococcus neoformans* gene encoding a protein related to multidrug resistance proteins." Gene **201**: 21-9.

Tobin, M. B., R. B. Peery, *et al.* (1997). "Genes encoding multiple drug resistance-like proteins in *Aspergillus fumigatus* and *Aspergillus flavus*." Gene **200**: 11-23.

Toledo, M. S., S. B. Levery, *et al.* (1999). "Characterisation of Sphingolipids from Mycopathogens: Factors Correlating with expression of 2-Hydroxy Fatty Acyl (*E*)-Δ^3-Unsaturation in Cerebosides of *Paracoccidiodes brasiliensis* and *Aspergillus fumigatus*." Biochem **38**: 7294-7306.

Unterkircher, C. S., S. C. Yazaki, *et al.* (1996). "Specific components found in circulating immune complexes (CIC) in paracoccidioidomycosis." J Med Vet Mycol **34**: 273-7.

Urban, M., T. Bhargava, *et al.* (1999). "An ATP-driven efflux pump is a novel pathogenicity factor in rice blast disease." Embo J **18**: 512-21.

Van Gelderen, K., A. De Komaid, *et al.* (1999). "Histoplasmosis and Paracoccidioidomycosis in northwestern Argentina III. Epidemiological survey in Vipos, La Toma, and Choromoro - Trancas, Tucuman, Argentina." Eur J Epidemiol **15**: 383-8.

Van Gelderen, K. and E. Duran, L. (1995). "Histoplasmosis in northwestern Argentina. II: Prevelance of *Histoplasmosis capsulati* and paracoccidiodiomycosis in the population south of Chuscha, Gonzalo and Potero in the province of Tucuman." Mycopathologia **129**: 17-23.

Van Helvoort, A., A. J. Smith, *et al.* (1996). "MDR1 P-Glycoprotein is a Lipid Translocase of Broad Specificity, while MDR3 P-glycoprotein specifically translocates phosphatidyl choline." Cell **87**: 507-517.

Van Veen, H. W., R. Callaghan, *et al.* (1998). "A bacterial anti-biotic resistance gene that complements the human multidrug-resistance P-glycoprotein gene." Nature **391**: 291-295.

Van Veen, H. W., K. Venema, *et al.* (1996). "Multidrug resistance mediated by a bacterial homolog of the human multidrug transporter MDR1." Proc Natl Acad Sci U S A **93**: 10668-72.

Vanden Bossche, H., F. Dromer, *et al.* (1998). "Antifungal drug resistance in pathogenic fungi." Med Mycol **36** (Suppl 1): 119-28.

Vanden Bossche, H., P. Marichal, *et al.* (1994). "Molecular mechanisms of drug resistance in fungi." Trends Microbiol **2**: 393-400.

Vanden Bossche, H., G. Willemsens, *et al.* (1997). "Anti-*Candida* drugs - the biochemical basis for their activity." Crit Rev Microbiol **15**: 57-72.

Venkatakrishna, S. and G. F.-L. Ames (1989). "Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR." Gene **84**: 1-8.

Vergara, M. L. and R. Martinez (1998). "Role of the armadillo *Dasypus novemcinctus* in the epidemiology of paracoccidioidomycosis." Mycopathologia **144**: 131-3.

Vicentini, A. P., J. L. Gesztesi, *et al.* (1994). "Binding of *Paracoccidioides brasiliensis* to laminin through surface glycoprotein gp43 leads to enhancement of fungal pathogenesis." Infect Immun **62**: 1465-9.

Vos, J. C., E. A. Reits, *et al.* (2000). "Head-head/tail-tail relative orientation of the pore-forming domains of the heterodimeric ABC transporter TAP." Curr Biol **10**: 1-7.

Vos, J. C., P. Spee, et al. (1999). "Membrane topology and dimerization of the two subunits of the transporter associated with antigen processing reveal a three-domain structure." J Immunol **163**: 6679-85.

Walmsley, A. R., T. Zhou, *et al.* (1999). "The ATPase mechanism of ArsA, the catalytic subunit of the arsenite pump." J Biol Chem **274**: 16153-61.

Walmsley, A. R., T. Zhou, *et al.* (2001). "A kinetic model for the action of a resistance efflux pump." J Biol Chem **276**: 2378-91

Wang, E., C. N. Casciano, *et al.* (2000). "Cholesterol interaction with the daunorubicin binding site of P- glycoprotein." Biochem Biophys Res Commun **276**: 909-16.

Wang, W., D. Takezawa, *et al.* (1996). "A potato cDNA encoding a homolog of mammailing multidrug resistant p-glycoprotein." Plant Mol Biol **31**: 683-7

Wei, J., M. E. Hodes, *et al.* (1998). "Characterization of murine *Girk2* transcript isoforms; structure and differential expression." Genomics **51**: 379-390.

White, T. C., K. A. Marr, *et al.* (1998). "Clinical, Cellular and Molecular Factors That Contribute to Antifungal Drug Resistance." Clin Microbiol Rev **11**: 382-402.

Wilson, R. A. O. (1994). "2.2Mb of contiguous nucleotide sequence from chromosome III of *C.elegans*." Nature **368**: 32-38.

Wu, C. T., M. Budding, *et al.* (1991). "Isolation and chracterisation of *Drosophila* multidrug resistance gene homologs." Mol Cell Biol **11**: 144-9.

Xu, J., S. L. Smock, *et al.* (2000). "Cloning the full length cDNA for rat connective tissue growth factor: implications for skeletal development." J Cell Biochem **77**: 103-115.

Young, J. and I. B. Holland (1999). "ABC transporters: bacterial exporters-revisited five years on." Biochim Biophys Acta **1461**: 177-200.

Young, L., K. Leonhard, *et al.* (2001). "Role of the ABC transporter Mdl1 in peptide export from mitochondria." Science **291**: 2135-8.

Zhang, J. T. and V. Ling (1991). "Study of membrane orientation and glycosylated extracellular loop of mouse P-glycoprotein by in vitro translation." J Biol Chem **266**: 18224-18232.

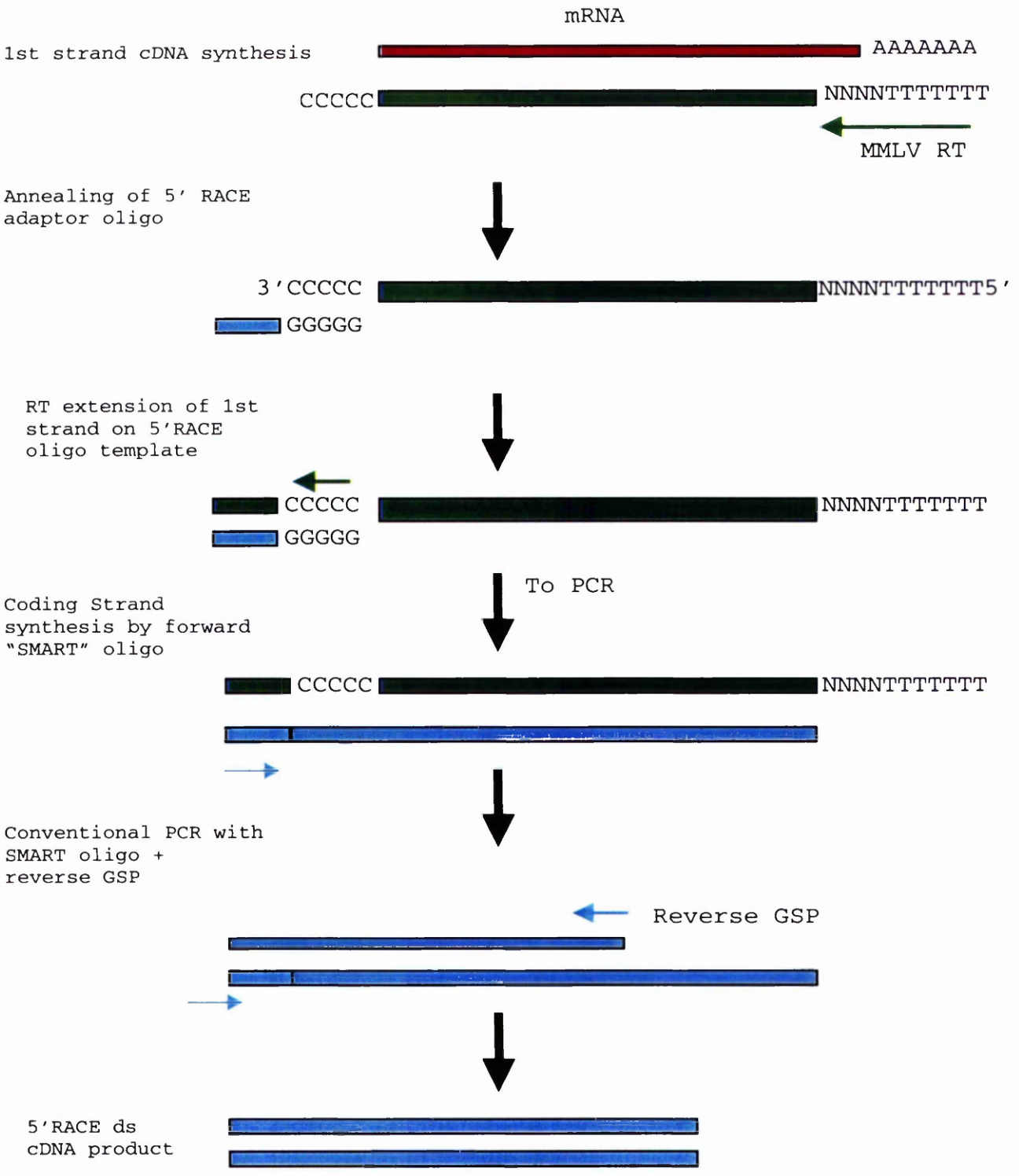
Zhou, T., S. Radaev, *et al.* (2000). "Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump." Embo J **19**: 4838-45.

Zhou, T. and B. P. Rosen (1997). "Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase." J Biol Chem **272**: 19731-7.

Zhou, X., V. M. Richon, *et al.* (1999). "Cloning of the cDNA encoding phenylalanyl tRNA synthetase regulatory a-subunit-like protein whose expression is down regulated during differentiation." Gene **233**: 13-19.

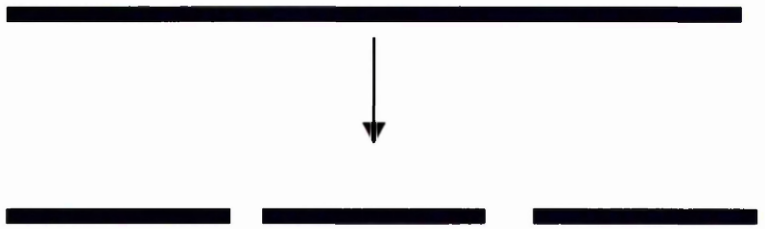
Zwiers, L. H. and M. A. De Waard (2000). "Characterization of the ABC transporter genes *MgAtr1* and *MgAtr2* from the wheat pathogen *Mycosphaerella graminicola*." Fungal Genet. Biol. **30**: 115-125.

Appendix I: 5 SMART RACE RT-PCR



Appendix II: GenomeWalker™ PCR

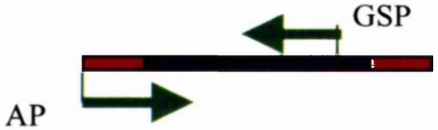
P.brasiliensis Genomic DNA



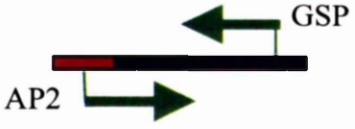
Digest with 1 of 5 blunt cutting restriction endonucleases

Ligate GenomeWalker™ adapter tags

Primary PCR with Gene Specific Primer 1(GSP1) and Adapter Primer 1(AP1)



Secondary Nested PCR with Gene Specific Primer 1(GSP2) and Adapter Primer 1(AP2)



GenomeWalker™ PCR product