

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

The Secreted Aspartic Proteinases of *Candida albicans*

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor
of Philosophy in the Institute of Molecular BioSciences at
Massey University, New Zealand

Michelle E. Sullivan

October 1999

To my parents

ACKNOWLEDGEMENTS

I would like to thank Professor Pat Sullivan, not only for providing an excellent lab environment, but for his wisdom and his unfailing enthusiasm throughout this project. Thanks also to my second supervisor Associate Professor John Tweedie, for his timely support and advice, and for always having an open door.

I would also like to thank all the past and present members of the Mainland lab who not only contributed advice and support, but who gave the lab spirit and made it a great place to work; Rochelle Ramsay who taught me the effectiveness of the 'hands on hips' approach, and who will always be a friend (but I'm not scared of you, even if the whippersnappers are!); Rebecca Sanders for fashion advice, friendship and fun; Linda Giblin for handy shortcuts and raucous Irish laughter; Donna Cooper for her endless kindness; and to Paul McJarrow for always taking the time to explain about tricky carbohydrates!

Thanks also to all the members of the institute who provided advice and inspiration throughout - particularly Catherine Day for her encouragement and direction when everything looked pear-shaped, Carolyn Young for cloning hints and tips, Rachel Page for hot-room help, and Carol Flyger for seeds and gardening advice. For technical support I would like to thank Gill Norris, Diana Carne, Liz Nickless, the staff at SAPU, and of course the ever-present David McSweeney for all the things that no-one else knew about.

Special thanks must go to Liz Hutton who always believed in me, even when I didn't believe in myself, and who will always be a source of inspiration. I hope you rest in peace my darling friend. Thanks also to my parents for their constant love and friendship. Lastly, I would like to thank Shaun Lott, for love, patience, kindness and for making me happier than I ever dreamed I could be.

ABSTRACT

The common human fungal pathogen, *Candida albicans*, possesses at least nine genes encoding secreted aspartic proteinases (*SAPs*). Saps are widely regarded as virulence factors, despite historical controversy surrounding their actual roles in the onset and development of candidosis. While Sap1, Sap2 and Sap3 had been previously studied at the biochemical level, Sap4 and Sap5 had not been detected, purified or characterised. To facilitate analysis of the proteins, *SAP4* and *SAP5* were amplified by PCR and cloned. The nucleotide changes in *SAP4* were silent, and *SAP5* contained one conserved amino acid substitution, compared with the published sequences. The methylotrophic yeast *Pichia pastoris* was used as a host for heterologous expression of *SAP4* and *SAP5*, and the respective proteins were purified to homogeneity. Purification of Sap4 involved Mono Q anion exchange chromatography at pH 7.0, while Sap5 was purified by cation exchange chromatography on Resource S at pH 7.0. *C. albicans SAP1* was also over-expressed in *P. pastoris* as a control. Biochemical analysis of the recombinant proteins revealed that Sap4 and Sap5 were optimally active at pH 4.5, 1-2 pH units higher than the optima of Saps1, 2 and 3. At optimum pH, the specific activities of Sap4 and Sap5 were 239 and 33 μg tyrosine equivalents/min, respectively. These isoenzymes also retained significant activity at pH 7.0, which suggested roles in the disease process at host sites of neutral pH. Sap4 and Sap5 showed decreased specific activity towards denatured globin, and increased specific activity towards a fluorocasein substrate, when compared with Saps1-3. Substrate specificity analyses (performed using a peptide substrate, glucagon, and MALDI-TOF of the purified peptide fragments), showed that Sap4 and Sap5 hydrolysed the glucagon at the same sites, but the analysis did not reveal a consensus cleavage sequence. The deduced masses of Sap1, Sap4 and Sap5 were 36,179, 36,995 and 37,256, respectively. However, ES-MS indicated the masses of the recombinant Sap1 and Sap4 were larger than expected, by 2232 and 2041 respectively. Glycopeptide fragment ion analysis suggested the additional mass was due to attached sugar residues. Carbohydrate chromatography confirmed the presence of mannose and N-acetyl glucosamine. The presence of N-acetyl glucosamine species, and the lack of consensus *N*-linked glycosylation sites in the Sap1, Sap4 and Sap5 proteins

suggests a novel pattern of *O*-linked glycosylation in *P. pastoris*. The purified enzymes were subjected to crystallisation trials and some promising crystals were produced. Previous studies showed that *SAP4*, *5* and *6* are expressed during serum-induced germ-tube formation, but this was not confirmed in this study.

TABLE OF CONTENTS

Abstract.....	i
Table of contents.....	iii
List of figures.....	ix
List of tables.....	xii
Amino acid abbreviations and colour coding.....	xiii
Abbreviations.....	xiv
Chapter 1.0 Introduction	1
1.1 <i>Candida albicans</i>	1
1.2 Pathogenicity	3
1.2.1 Adhesion.....	4
1.2.2 Cell wall.....	4
1.2.3 Polymorphism.....	5
1.2.4 Phenotypic variation	6
1.2.5 Secreted hydrolases	6
1.2.6 Hydrophobicity.....	7
1.2.7 Signal transduction networks.....	8
1.3 Other <i>Candida</i> species.....	9
1.4 <i>C. albicans</i> and host immunity	9
1.5 Antifungal therapies	10
1.6 Phenotypic switching.....	12
1.7 Molecular mechanisms of phenotypic switching.....	14
1.8 Molecular mechanisms of polymorphism.....	17
1.9 Aspartic proteinases.....	21
1.9.1 Evolution	21
1.9.2 Substrate specificity.....	22
1.9.3 Structural features.....	23
1.9.4 Catalytic mechanism.....	23
1.10 Retroviral aspartic proteinases.....	24
1.11 Fungal aspartic proteinases.....	25
1.12 Secreted aspartic proteinases	26
1.13 Secretion of Sap.....	27
1.14 Regulation of Sap	28
1.15 Sap and pathogenicity.....	32
1.16 Sap4, Sap5 and Sap6	35
1.17 Sap structure	38
1.18 Project aims	39

2.0 Materials and Methods	40
2.1 Enzymes	40
2.2 Chemicals	40
2.3 Miscellaneous products	41
2.4 <i>Escherichia coli</i> phenotypes	41
2.5 Yeast strains.....	42
2.6 Plasmids utilised during this study	42
2.7 Plasmids developed during this study.....	42
2.8 Growth media	43
2.8.1 Yeast media	43
2.8.2 Bacterial media.....	43
2.8.3 General buffers	44
2.9 General methods.....	44
2.9.1 Restriction endonuclease DNA digestion.....	44
2.9.2 Agarose gel electrophoresis of DNA	44
2.9.3 DNA ligations.....	44
2.9.4 Creating blunt-ended DNA.....	45
2.9.5 Preparation and transformation of <i>E. coli</i> competent cells	45
2.9.5.1 Preparation	45
2.9.5.2 Transformation	45
2.9.6 Selection of transformants	46
2.9.7 Plasmid DNA preparation – small scale	46
2.9.8 Plasmid DNA preparation – large scale.....	46
2.9.9 DNA purification from a TAE agarose gel.....	46
2.9.10 Genomic DNA purification from <i>C. albicans</i>	47
2.9.11 PCR.....	47
2.9.12 Southern blotting and hybridisation analyses	47
2.9.13 Radiolabelling DNA	48
2.9.14 Total RNA extraction from <i>C. albicans</i>	48
2.9.15 Agarose gel electrophoresis of RNA	48
2.9.16 Northern blotting and hybridisation analyses	49
2.9.17 Oligonucleotide synthesis	49
2.9.18 Dideoxy chain termination sequencing.....	49
2.10 Yeast and protein methods.....	49
2.10.1 Preparation and transformation of yeast spheroplasts.....	49
2.10.2 Screening yeast transformants	50
2.10.3 <i>P. pastoris</i> growth and induction.....	50
2.10.4 TCA precipitation of protein samples.....	51
2.10.5 SDS-PAGE analysis	51
2.10.6 Native PAGE analysis	51

2.10.7	Western blotting of proteins	51
2.10.8	N-terminal amino acid sequence analysis of proteins.....	52
2.10.9	Sap enzyme assays.....	53
2.10.9.1	Azocollagen assay	53
2.10.9.2	BSA-based assay	53
2.10.9.3	Globin-based assay	53
2.10.9.4	Fluorocasein-based assay	54
2.10.10	Substrate specificity of Sap4 and Sap5.....	54
2.10.11	Column chromatography	55
2.10.12	Carbohydrate cleavage.....	55
2.10.12.1	β -elimination reaction.....	55
2.10.12.2	Acid hydrolysis.....	56
2.10.13	DIG Glycan/protein analysis	56
2.10.13.1	Labelling.....	56
2.10.13.2	Detection	56
2.11	Antibody production	57
2.12	Crystallisation trials	57
2.13	Microscopy	58
2.14	Computing programs	58
3.0	Cloning and Vector Construction.....	59
3.1	Introduction	59
3.2	PCR amplification of <i>SAP4</i> , <i>SAP5</i> and <i>SAP6</i>	60
3.3	Optimising PCR amplification of <i>SAPs</i>	61
3.4	Restriction endonuclease analysis of PCR products	61
3.5	Analysis of <i>SAP4</i> and <i>SAP5</i> clones	66
3.6	DNA sequence analysis of <i>SAP4</i> and <i>SAP5</i> clones	70
3.7	Construction of heterologous expression vectors for Sap4 and Sap5.....	75
3.8	Construction of a Sap1 heterologous expression vector	78
4.0	Transformation and Expression	80
4.1	Introduction	80
4.2	Transformation of KM71 with heterologous expression constructs.....	81
4.3	Selection of KM71 transformants.....	81
4.4	Confirmation of plasmid integration.....	83
4.5	Preliminary expression of recombinant Sap proteins	83
4.6	Identification of the recombinant proteins.....	90
4.6.1	Identification by western blotting	90
4.6.2	Identification by enzyme assay.....	90
4.7	Large scale expression of Sap1, Sap4 and Sap5	92

4.7.1	Optimising strain storage conditions	92
4.7.2	Optimising starter culture conditions for Sap expression	95
4.7.3	Effect of methanol concentration on Sap expression.....	95
4.7.4	Effect of added buffer on Sap expression	95
4.7.5	Effect of induction culture size on Sap expression	97
4.7.6	Effect of culture aeration on Sap expression	97
4.8	Fermentation scale induction of Sap.....	99
4.9	Development of a new system for Sap expression	100
4.10	<i>E. coli</i> expression of Sap4	102
4.10.1	PCR primer design.....	102
4.10.2	Transformation of host strains	107
4.10.3	Expression of recombinant Sap4	107
5.0 Purification of Sap		108
5.1	Introduction	108
5.2	Purification of recombinant Sap1	109
5.3	Purification of recombinant Sap4	110
5.3.1	Anion exchange chromatography at pH 6.0	110
5.3.2	Cation exchange chromatography at pH 6.0.....	113
5.3.3	Anion exchange chromatography at pH 7.0	113
5.3.4	Anion exchange chromatography at pH 7.0, low ionic strength.....	114
5.3.5	Optimising the anion exchange gradient.....	114
5.3.6	Batchwise purification trials with DEAE sepharose.....	116
5.3.7	DEAE sepharose column chromatography.....	117
5.4	Large-scale purification of mature Sap4.....	118
5.5	Purification of recombinant Sap5	124
5.5.1	Anion exchange chromatography	124
5.5.2	Cation exchange chromatography	124
5.5.3	Gel filtration of Sap5	125
5.6	Sap 5 purification protocol	125
6.0 Biochemical Characterisation of Recombinant Sap Isoenzymes.....		126
6.1	Introduction	126
6.2	Sap enzyme activity assays.....	127
6.2.1	Sap4 substrate preference	127
6.2.2	Enzyme kinetics.....	127
6.2.2.1	Globin assay	127
6.2.2.2	Fluorocasein assay.....	129
6.3	pH optima of the Sap isoenzymes.....	131
6.4	Specific activity of the Sap isoenzymes	133

6.5	Substrate specificity of Sap	135
6.6	Thermal stability of the Sap isoenzymes	143
6.7	SDS-PAGE analysis of the Saps1-5	144
6.8	Native PAGE analysis of Saps1-5	144
6.9	Autodegradation of Sap isoenzymes.....	146
6.10	Pepstatin inhibition of Sap.....	146
6.11	Glycosylation of Sap	149
	6.11.1 Electrospray mass spectroscopy	149
	6.11.2 FACE analysis of Sap.....	153
	6.11.3 Carbohydrate chromatography	155
6.12	Interaction of recombinant Saps with SQAPI.....	162
6.13	Antibody production.....	164
6.14	Production of anti-Sap4	164
6.15	Purification of the IgG fraction.....	165
6.16	Titre of anti-Sap4.....	165
	6.16.1 Comparison of anti-Sap2 and anti-Sap4	165
	6.16.2 Anti-Sap4 titre	167
6.17	Specificity of anti-Sap4	167
7.0 Expression of Sap.....		171
7.1	Introduction	171
7.2	Expression of <i>SAPs 4, 5 and 6</i>	172
7.3	Northern analysis of <i>SAP4-6</i> expression: strains ATCC 10231 and A72.....	176
7.4	Northern analysis of <i>SAP4-6</i> expression: strain ATCC 10261	179
7.5	<i>In vitro</i> expression analyses.....	181
8.0 Crystallisation trials.....		183
8.1	Introduction	183
8.2	Crystallisation of Sap1.....	184
8.3	Crystallisation of Sap4.....	188
8.4	Crystallisation of Sap5.....	189
8.5	Homology modelling of Saps 1, 4 and 5	193
9.0 Discussion.....		198
9.1	Future work	208
10.0 Appendix.....		210
10.1	<i>Sequence alignment of SAPI-SAP9 of C. albicans</i>	211
10.2	Dionex analysis of Sap isoenzymes.....	214

10.2.1	Dionex calculations	214
10.2.2	Dionex analysis of Sap5 carbohydrate moieties	215
10.3	Purification and analysis of Sap5	216
10.3.1	Purification of Sap5	216
10.3.2	Anti Sap5 antiserum	218
10.4	ESMS data	219
10.5	MALDI-TOF analysis of glucagon peptide fragments	220
11.0	References	224

LIST OF FIGURES

1.1	Alternative morphologies of <i>C. albicans</i>	2
1.2	Organisation of <i>C. albicans</i> cell wall.....	5
1.3	Variant colony phenotypes in the switching repertoire of WO-1	14
1.4	A model for the regulation of white- and opaque-phase specific genes.....	16
1.5	A model for the regulation of pseudohyphal growth in <i>C. albicans</i>	20
1.6	Schematic representation of the substrate binding cleft of an aspartic proteinase	22
1.7	Expression of the <i>SAP</i> gene family in <i>C. albicans</i>	31
3.1	Protocol for generating <i>SAP</i> clones	62
3.2	Optimised PCR amplification of <i>SAP4</i> , <i>SAP5</i> and <i>SAP6</i>	64
3.3	Schematic of <i>SAP4</i> , <i>5</i> & <i>6</i> showing relative positions of PCR primers and restriction sites..	65
3.4	Restriction endonuclease analysis of <i>SAPs</i> 4, 5 and 6 PCR products	67
3.5	Restriction digest analysis of <i>SAP4</i> and <i>SAP5</i> sequences cloned into pBluescript.....	69
3.6	Nucleotide sequence of <i>SAP4</i> from <i>C. albicans</i> strain ATCC 10261	71
3.7	Nucleotide sequence of <i>SAP5</i> from <i>C. albicans</i> strain ATCC 10261	72
3.8	Alignment of the sequence containing the amino acid substitutions in Sap5 with other members of the <i>C. albicans</i> Sap family	74
3.9	Schematic of expression vector construction.....	76
3.10	Screening pPIC3K clones for the correct orientation of <i>SAP4</i> and <i>SAP5</i> genes.....	77
3.11	Schematic of the <i>SAP1</i> cloning strategy	79
4.1	G418 screen of KM71 transformants.....	84
4.2	Schematic representation of pMS1 and pMS5 plasmids integrated into the KM71 genome..	85
4.3	Southern blot of <i>P. pastoris</i> (KM71) transformants MS4A and MS5A, containing pMS4 and pMS5.....	86
4.4	SDS-PAGE of preliminary Sap1, Sap4 and Sap5 expression trials	88
4.5	N-terminal sequence of recombinant Sap proteins	89
4.6	Western analysis of Sap expression	91
4.7	Comparative induction of Sap4 and Bgl2.....	93
4.8	Effect of storage medium on Sap expression	94
4.9	Effect of starter culture medium on Sap expression	94
4.10	Effect of methanol concentration on Sap expression.....	96
4.11	Effect of buffer on Sap expression.....	96
4.12	Effect of induction culture volume on Sap expression.....	98
4.13	Effect of culture aeration on Sap expression.....	98
4.14	Sparger system for induction of recombinant Sap	101
4.15	PCR primers for <i>E. coli</i> expression.....	103
4.16	Strategy for cloning <i>SAP4</i> into pRSETa, an <i>E. coli</i> expression vector	104

4.17	SDS-PAGE and western blot analysis of <i>E. coli</i> expression of Sap4	107
5.1	Purification of Sap1 using ion-exchange chromatography	111
5.2	SDS-PAGE of Sap1 fractions from ion exchange chromatography	112
5.3	SDS-PAGE of Sap4 fractions from ion exchange chromatography (pH 7.0).....	112
5.4	SDS-PAGE of Sap4 fractions from anion exchange chromatography at low ionic strength (pH 7.0).....	115
5.5	Separation of the misprocessed species of Sap4 using Resource Q ion exchange chromatography	115
5.6	SDS-PAGE of Sap4 fractions from optimised anion exchange gradient.....	116
5.7	Batchwise DEAE Sepharose purification trial results.....	117
5.8	DEAE Sepharose chromatography at pH 7.0.....	119
5.9	Graph showing protein and Sap activity fractions from the DEAE Sepharose column.....	121
5.10	SDS-PAGE analysis of fractions from the DEAE Sepharose column	121
5.11	h.p.l.c. trace of Mono Q purification of Sap4	122
5.12	Sap4 at each stage of purification	123
5.13	Purified recombinant Sap4.....	123
6.1	Relative Sap4 activity toward various protein substrates.....	128
6.2	Time course of Sap activity in the globin assay.....	128
6.3	Plot of reaction velocity vs. substrate concentration.....	130
6.4	Hanes plot for Sap4 activity using the globin assay.....	130
6.5	The fluorocasein Sap assay	132
6.6	pH activity profiles of the Sap isoenzymes.....	132
6.7	Specific activities of Sap isoenzymes in the globin (A) and fluorocasein (B) assays	134
6.8	Reversed phase h.p.l.c of glucagon hydrolysates with Sap4.....	136
6.9	Reversed phase h.p.l.c of glucagon hydrolysis by Sap5	137
6.10	Model of cleavage of glucagon by Saps 1-5	140
6.11	Thermal stability profiles of Sap isoenzymes	143
6.12	SDS-PAGE analysis of Saps 1-5	145
6.13	Native PAGE analysis of Saps 1-5	145
6.14	Sap autodegradation.....	147
6.15	Pepstatin inhibition of Sap	148
6.16	ES mass spectra of Sap1, Sap4 and Sap5.....	149
6.17	Glycopeptide fragment ion analysis.....	150
6.18	DIG glycan/protein analysis of Saps 1, 2, 3 and 4	154
6.19	Sap Dionex analysis.....	158
6.20	Standard curve for quantitation of carbohydrate species	159
6.21	O-linked oligosaccharides of <i>S. cerevisiae</i> and <i>P. pastoris</i>	161
6.22	Biacore analysis of Sap interactions with SQAPI.....	163
6.23	Antibody titre after second booster	166
6.24	Western analysis of anti-Sap4 vs. <i>P. pastoris</i> supernatant.....	166

6.25	Comparison of the sensitivity of anti-Sap2 and anti-Sap4	168
6.26	Anti-Sap4 titre.....	168
6.27	Comparison of the specificity of anti-Sap2 and anti-Sap4.....	169
7.1	Time course of serum-induced hyphal formation.....	173, 174
7.2	Clumped <i>C. albicans</i> strain ATCC 10231 hyphae.....	175
7.3	RNA extracted from strain ATCC 10231	177
7.4	Northern hybridisations for <i>SAP4-6</i> and <i>ACT</i>	178
7.5	Northern hybridisations for <i>SAP4-6</i> and <i>LEU2</i>	180
8.1	Crystals of native Sap5 from the sparse matrix screen and crystal screen II	190
8.2	Crystals of Sap4 complexed with A70450 inhibitor	191
8.3	Crystals of Sap4 complexed with SQAPI inhibitor	192
8.4	Homology models using the Sap2 structure as a template.....	195
8.5	Active site models of Saps 1, 4 and 5	196
8.6	Sap2 active site structure (in complex with A70450) overlaid with the Sap4 model.....	197
9.1	Model of proposed roles of Sap isoenzymes in the onset and development of candidosis ...	206
10.1	Alignment of the amino acid sequences of Sap1-Sap9 of <i>C. albicans</i>	211
10.2	Dionex chromatograms of Sap5.....	215
10.3	Purification of Sap5 using ion-exchange chromatography	216
10.4	SDS-PAGE of Sap5 following ion-exchange chromatography	217
10.5	Analysis of the specificity of anti-Sap5	218

LIST OF TABLES

1.1	Antifungal agents currently used for treatment of <i>Candida</i> infections	12
1.2	Table of identity between <i>SAP</i> sequences from <i>C. albicans</i>	27
3.1	Summary of the PCR conditions for <i>SAP4</i> , <i>SAP5</i> and <i>SAP6</i> PCR amplification reactions	63
3.2	Comparison of the expected and actual band sizes from restriction enzyme analysis of <i>SAP4</i> , <i>SAP5</i> and <i>SAP6</i> PCR products.....	68
5.1	Purification of recombinant Sap1	111
5.2	Purification of mature Sap4	122
6.1	Glucagon fragments produced by Sap4 and Sap5.....	139
6.2	Consensus sequences of Saps 1-5 cleavage sites in glucagon.....	141
6.3	Amino acid side chains which form subsites S ₁ and S ₁ '	142
6.4	Predicted and experimental masses of Saps.....	152
6.5	Carbohydrate destruction during acid hydrolysis.....	156
8.1	Sparse matrix crystal screen.....	185
8.2	Crystal screen II.....	186
8.3	Additional crystallisation conditions for Sap1 and Sap4	187

AMINO ACID ABBREVIATIONS & COLOUR

Amino acid	Abbreviation	Symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asparagine or Aspartic acid	Asx	B
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glutamine or Glutamic acid	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Grey	General non-polar	A,M,I,L,V
Green	General polar	S,T,N,Q
Blue	Positively charged	H,K,R
Red	Negatively charged	D,E
Purple	Aromatic	F,W,Y
Brown	“Structure breakers”	G,P
Yellow	Cysteine	C

ABBREVIATIONS

Amp	Ampicillin
BCIG	5'-bromo-4-chloro-3-indolyl- γ -D-galactopyranoside
BCIP	5'-bromo-4-chloro-3-indolyl phosphate
bp	base pair
BSA	Bovine serum albumin
CAPS	3-cyclohexylamino-1-propanesulphonic acid
cDNA	Complementary DNA
C-terminal	Carboxy-terminal
Da	Dalton
DEAE	Diethylaminoethyl
DIG	Digoxigenin
DMF	Dimethylformamide
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
dGTP	2'-deoxyguanosine-5'-triphosphate
dTTP	2'-deoxythymidine-5'-triphosphate
dNTP	deoxynucleoside triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid
ES-MS	Electrospray mass spectrometry
G418	Geneticin – disulphide salt
GalNAc	N-acetyl galactosamine
GlcNAc	N-acetyl glucosamine
HEPES	N-2-hydroxyethylpiperazine-N'-ethanesulphonic acid
h.p.l.c.	High pressure liquid chromatography
IPTG	Isopropyl- γ -D-thiogalactopyranoside
kb	kilobase
kDa	kiloDalton
MOPS	3-[N-morpholino]propanesulphonic acid

NBT	Nitro blue tetrazolium
mRNA	messenger RNA
N-terminal	Amino-terminal
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pI	Isoelectric point
PIPES	Piperazine-n,n'-bis(2-ethanesulphonic acid)
PVDF	Polyvinylidenedifluoride
RNA	Ribonucleic acid
<i>SAP</i>	SAP gene
Sap	Sap protein
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TEMED	N,N,N',N'-tetramethylethylenediamine
TFA	Trifluoroacetic acid
Tris	Tris(hydroxymethyl)aminomethane
Tween 20	Polyoxyethylene-sorbitan monolaurate
v/v	volume/volume
w/v	weight/volume
w/w	weight/weight

C. albicans GENETIC CODE

		<i>S E C O N D P O S I T I O N</i>								
		T	C	A		G				
<i>F</i>	T	TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys	T
		TTC	Phe	TCC	Ser	TAC	Tyr	TGC	Cys	C
		TTA	Leu	TCA	Ser	TAA	*	TGA	*	A
		TTG	Leu	TCG	Ser	TAG	*	TGG	Trp	G
<i>S</i>	C	CTT	Leu	CCT	Pro	CAT	His	CGT	Arg	T
		CTC	Leu	CCC	Pro	CAC	His	CGC	Arg	C
		CTA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A
		CTG	Ser	CCG	Pro	CAG	Gln	CGG	Arg	G
<i>O</i>	A	ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser	T
		ATC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C
		ATA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A
		ATG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G
<i>O</i>	G	GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly	T
		GTC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C
		GTA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A
		GTG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G

Codons are as they appear in the coding strand of DNA, reading 5'→3'. Termination codons are represented by *. Note the unusual decoding of CTG as a Serine in *C. albicans*, rather than the standard amino acid Leucine.

1.0 INTRODUCTION

1.1 *Candida albicans*

Since the earliest known report of candidosis by Hippocrates in the 4th century B.C., assigning species to the taxonomic group *Candida* has been plagued with difficulty, largely due to the significant diversity of these yeasts. Today, this genus contains over 150 species, and the common feature linking this disparate group is the absence of a sexual cycle (Odds, 1988). While the genus is sizeable, medically significant *Candida* species represent a minority, and of these *Candida albicans* is the most common human fungal pathogen, and consequently has been the most extensively studied.

Frequently described as being dimorphic, *C. albicans* is essentially polymorphic, existing *in vivo* as budding yeasts (blastospores), true hyphae and pseudohyphae (which are chains of elongated budding cells) (Fig. 1.1). A fourth, rarer morphology is chlamydo spores which are large thick-walled cells often found at the end of hyphae or pseudohyphae and to date, have only been isolated *in vitro* (Odds, 1988).

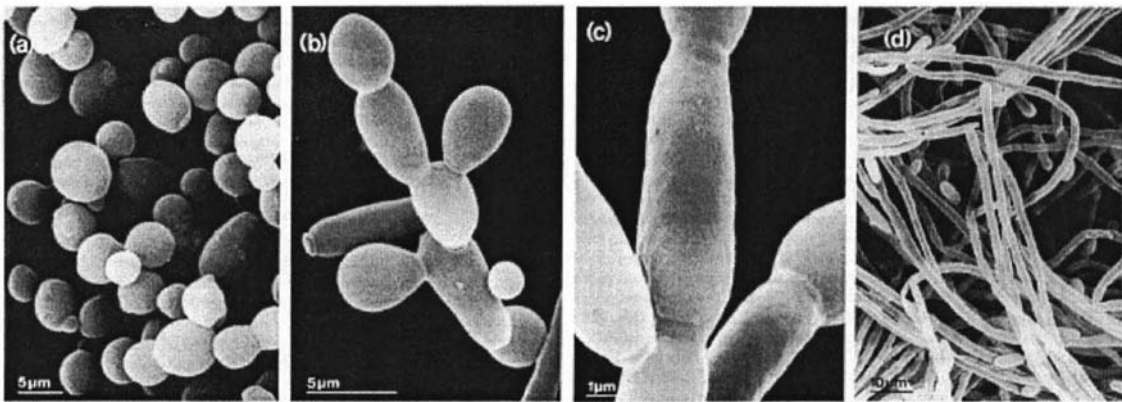


Figure 1.1 *Alternative morphologies of C. albicans*
 (A) Ovoid budding yeast cells (blastospores). (B) Pseudohyphae; chains of blastospores. Distinguished from true hyphae by the obvious constrictions at septal junctions. (C) Pseudohyphae; cells are even more elongated and septal constrictions are less pronounced. (D) True hyphae with few lateral blastospores. No apparent septal constrictions. Figure from Odds, (1988).

C. albicans is also capable of high-frequency, reversible phenotypic switching, the characteristics of which are quite different from those of morphological variation. Switching generally occurs spontaneously in contrast to the yeast-hyphal transition which is an *en masse* conversion in response to environmental cues (Soll, 1992). Advances in elucidating the mechanisms of pathogenicity in *C. albicans* have traditionally been hampered due to its diploid nature and the lack of any sexual cycle. *C. albicans* is an imperfect yeast, lacking a meiotic cycle (Magee, 1990), and many early investigators consequently suspected *C. albicans* was haploid. However, debate over the ploidy was resolved when it was found that a strongly biased spectrum of auxotrophic mutants was produced following UV-irradiation-induced mitotic recombination, suggesting *C. albicans* was heterozygous at several loci (Whelan *et al.*, 1980; Whelan & Magee, 1981). In another series of experiments, including comparisons of the DNA content between *C. albicans* and *S. cerevisiae*, Olaiya and Sogin (1979) also concluded *C. albicans* was diploid.

Recently, *C. albicans* homologues of genes involved in the *S. cerevisiae* mating-response pathway have been cloned (Leberer *et al.*, 1996; Kohler & Fink, 1996). These genes can functionally complement *S. cerevisiae* mating-defects, suggesting that *C.*

albicans may have a signal transduction system analogous to that controlling mating type in *S. cerevisiae*. However, further analysis of these genes has prompted the proposal that in *C. albicans*, the mating pathway components have been adapted for virulence and hyphal formation (Leberer *et al.*, 1997).

1.2 Pathogenicity

C. albicans commonly exists as a harmless commensal in humans and other warm-blooded animals, where it is found in the oral cavity, vaginal or digestive tracts as part of the normal mucosal microflora. The shift from commensal to pathogen is due to an imbalance of the host immune system which occurs as a result of cancer, immunosuppressant therapy following transplant surgery, acquired immune deficiency syndrome (AIDS) or drug addiction (Odds, 1988). *C. albicans* infections can range from mild superficial lesions of the skin and mucosal membranes (including oral and vaginal thrush) to life-threatening systemic or disseminated disease affecting virtually any organ, but especially the liver, kidney, lungs and brain. In addition, difficulties in diagnosing and treating immunocompromised patients have contributed to the increasing incidence of *C. albicans*-related illness (Cutler, 1991).

Epidemiological studies indicate that development of candidosis is usually preceded by either replacement of the commensal with a more pathogenic strain, or by the triggered response of an endogenous strain to express virulence characteristics (Agabian *et al.*, 1994). While the signals responsible for triggering such an adaptive response are yet to be described, several factors have been identified which may have a role in tissue invasion by *C. albicans*.

It is widely accepted that virulence in *C. albicans* involves the interplay of many complex developmentally and environmentally regulated factors (Odds, 1988; White *et al.*, 1995). In addition to the capacity for growth at 37 °C, the various attributes of *C. albicans* which are considered to be virulence determinants are described below.

1.2.1 Adhesion

Adhesion of *C. albicans* to host cell surfaces is regarded as a crucial prerequisite for colonisation and subsequent invasion of host tissues (Fukazawa & Kagaya, 1997). *C. albicans* adheres to a range of surfaces including epithelial cells, fibrin platelets, endothelial cells, skin and plastic. It is the outermost layer of the *C. albicans* cell wall which contains mannoproteins that possess distinct adhesion activity (Casanova *et al.*, 1992; Klotz *et al.*, 1994). Studies on the comparative adhesion of *Candida* spp. to host cells suggest a hierarchy amongst species, with the more commonly presenting pathogens such as *C. albicans* and *C. tropicalis* exhibiting greater *in vitro* adherence than relatively non-pathogenic species such as *C. guilliermondii* (King *et al.*, 1980). In addition, pseudohyphae and true hyphae are typically more adhesive than blastospores from the same isolate, as filamentous forms of the fungus appear to express more adhesive moieties (Calderone *et al.*, 1988).

1.2.2 Cell Wall

The cell wall of *C. albicans* is composed of multiple layers, which differ in chemical composition. The inner layers are responsible for maintenance of the structural integrity of the wall and contain predominantly β -1,3-glucan and chitin which can be covalently cross-linked to further increase structural rigidity. In contrast, the outer layers are composed mainly of mannoprotein (mannan) and β -1,6-glucan (Fig. 1.2). As previously mentioned, some mannoproteins in the outer cell wall facilitate adhesion (Klotz *et al.*, 1994), while others are involved in provocation of the host immune response (Calderone *et al.*, 1994). Variation in expression of these antigens occurs in response to available nutrients, growth phase and to morphogenic form, both *in vitro* (Chaffin *et al.*, 1988) and *in vivo* (Brawner & Cutler, 1987). The ability of *C. albicans* to rapidly alter the presenting cell-wall surface components might provide a mechanism for fungal evasion of host immune systems.

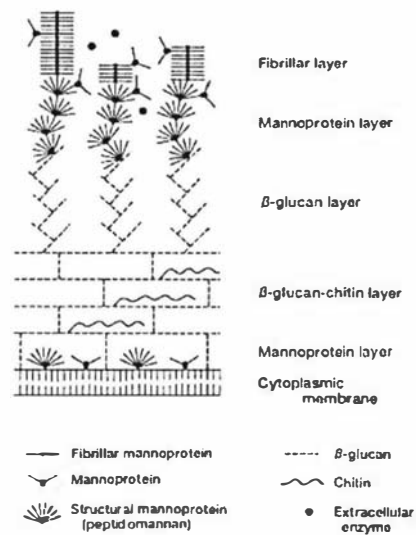


Figure 1.2 **Organisation of *C. albicans* cell wall**
Schematic diagram of the cell wall structures of *C. albicans* showing location of adhesion macromolecules (Fukazawa & Kagaya, 1997).

1.2.3 Polymorphism

In the majority of candidosis lesions, *C. albicans* exists as a medley of yeast cells, pseudohyphae and true hyphal cells. Hyphal forms of *C. albicans* adhere better to epithelial cells than blastospores (Calderone *et al.*, 1988) and evidence suggests that it is hyphal cells which facilitate penetration of the epithelial layer (Howlett & Squier, 1980; Rotrosen *et al.*, 1985). For instance, Sobel *et al.* (1984) noted a reduction in pathogenicity of a variant strain of *C. albicans* which was incapable of hyphal production. Despite this, overall evidence seems to indicate that both yeast and hyphal forms of *C. albicans* have equivalent capacities to initiate and sustain infection (Odds, 1988) thereby implying that *Candida* hyphae are no more virulent than blastospores in animal infections. So it is perhaps the ability for interconversion between yeast and hyphal forms that enhances the pathogenicity of *C. albicans*, given that different morphologies may be better adapted *in vivo* to particular microniches.

1.2.4 Phenotypic variation

C. albicans is capable of high-frequency reversible phenotypic switching, a heritable trait which produces an assortment of variant colony morphologies (Fig. 1.3) (Slutsky *et al.*, 1985). This means that cells emanating from a single progenitor are capable of expressing a variety of colony morphologies, and that this capacity for phenotypic variation is passed on to progeny in a heritable manner. Also, expression of these variant phenotypes occurs at frequencies well above those of point mutation, hence the term 'high-frequency' (Soll, 1992). It is important to note that the mechanism of phenotypic variability is independent to, and can be superimposed upon, that of *C. albicans* polymorphism. Switching has been shown to occur at sites of *C. albicans* infection (Soll *et al.*, 1987) and can alter many aspects of cellular physiology in addition to effecting morphological changes. The strain WO-1 (Slutsky *et al.*, 1987) undergoes a white-opaque transition which is probably the most extensively studied of all switching systems. Evidence of differential gene expression (Morrow *et al.*, 1994), coupled with reports of antigenic variability between the surfaces of white and opaque cells (Anderson & Soll, 1987), led researchers to examine the effects of switching on other putative virulence traits. Kennedy *et al.* (1988) observed striking differences between adhesion of white and opaque cells to epithelium, and Soll *et al.* (1989) reported different profiles of drug susceptibility for different switch phenotypes. Environmental constraints for yeast-hyphal transition have also been found to vary between phenotypes (Anderson *et al.*, 1989) and lastly, expression of members of the secreted aspartic proteinase (Sap) family are differentially regulated during the white-opaque transition of WO-1 (White *et al.*, 1995; White & Agabian, 1995). A more detailed discussion of differentially expressed genes, and the molecular basis of phenotypic switching appears later in this Introduction.

1.2.5 Secreted hydrolases

Many pathogenic fungi, including *C. albicans*, secrete a range of hydrolytic enzymes such as proteinase, lipase and phospholipase. These extracellular enzymes contribute to cell proliferation by providing nutrients for the fungus, and may also cause damage to

host cells allowing subsequent penetration and tissue invasion by the fungus (Tsuboi *et al.*, 1994). Secretion of proteinase, lipase and α -glucosidase by various *Candida* strains correlates with the relative pathogenicity ranking of these strains, and each of these enzymes appears to have a key role for fungal growth in specified environments (Tsuboi *et al.*, 1994). Phospholipase and N-acetyl glucosaminidase are also potentially involved in the breakdown and invasion of tissues, and recently, it was reported that disruption of a phospholipase B gene in *C. albicans* resulted in significant attenuation of virulence, by reducing the capacity of the yeast to penetrate host cells (Leidich *et al.*, 1998). However the most overwhelming evidence for a pathogenic role for secreted hydrolases comes from studies of the secreted aspartic proteinases of *C. albicans* (Hoegl *et al.*, 1996). Evidence linking these enzymes with virulence in *C. albicans* will be discussed in greater detail later in this Introduction.

1.2.6 Hydrophobicity

The precise contribution of cell surface hydrophobicity (CSH) in the development and dissemination of *C. albicans* infections is unclear. Typically, the outer cell wall of *C. albicans* is hydrophilic, but it can rapidly acquire a hydrophobic character (Hazen *et al.*, 1991). However, this conversion does not appear to involve *de novo* protein synthesis as hydrophilic cell walls possess buried hydrophobic proteins which can become exposed (Hazen *et al.*, 1990). The finding that hydrophobic cells were more virulent in mice than their hydrophilic counterparts, led Antley and Hazen (1988) to suggest that CSH might contribute to virulence by reducing the susceptibility of the yeast to elimination by the host immune system. More recently, Hazen *et al.* (1991) reported that hydrophobic cells are capable of a broader pattern of binding to host tissues, compared to hydrophilic cells which require macrophage-rich tissue for adhesion. This intimates that the enhanced pathogenicity of hydrophobic cells may be attributable to their greater potential for tissue-binding following clearance from the bloodstream. Lastly, the relative CSH of an organism contributes to the physical forces involved in surface adhesion and strong positive correlation between adhesion and relative CSH has been noted (Panagoda *et al.*, 1998). This enables adhesion of the fungus to plastic surfaces, such as catheters and other indwelling devices (Klotz *et al.*,

1985), which provides a means of direct entry to the vascular system and predisposes the host to disease.

1.2.7 Signal transduction networks

The suggestion that a common signal transduction pathway regulates pathogenic growth in a variety of fungi is supported by evidence from a diverse range of organisms (Hamer & Holden, 1997). Infection-related morphogenesis is a feature of many human fungal pathogens, including *C. albicans*, and recent evidence suggests that dimorphism is accompanied by changes in gene expression (White *et al.*, 1995; Gow *et al.*, 1995). The pheromone signal transduction pathways of budding and fission yeasts utilise functionally conserved mitogen-activated protein (MAP) kinase cascades. These cascades are also common to multicellular eukaryotes, as conservation of signalling pathways across disparate species is not uncommon (Herskowitz, 1995). However, Liu *et al.* (1993) discovered that components of this signalling pathway are also required for invasive growth and pseudohyphal formation in *Saccharomyces cerevisiae*. Now, *C. albicans* and the diverse plant pathogens *Magnaporthe grisea* and *Ustilago maydis* have also been found to require homologues of the kinases found in the pheromone signalling pathway for infection-related morphogenesis (for review, see Hamer & Holden, 1997). These studies intimate that pathogenic growth in a variety of fungi involves signal transduction via common MAP kinase pathways, evoking speculation that such a pathway might regulate other virulence factors. Indeed, evidence for MAP kinase involvement in cell wall integrity is provided by analysis of the *C. albicans* *MKCl* (mitogen-activated protein kinase) mutant, a viable strain which shows attenuated virulence in murine models due to cell wall defects (Diez-Orgas *et al.*, 1997).

1.3 Other *Candida* species

Although most episodes of candidosis are caused by *C. albicans*, other species in the genus *Candida* are also pathogenic. Reports documenting the increasing frequency of infections due to the non-*C. albicans* species during the 1980's (Horn *et al.*, 1985; Meunier-Carpentier *et al.*, 1981) foreshadowed a shift in the epidemiology of candidosis in the 1990's. The recognised frequency of isolation of *Candida* species in humans ranks *C. albicans* highest, followed by *C. tropicalis*, *C. parapsilosis*, *C. glabrata*, *C. krusei*, *C. lusitaniae* and *C. pseudotropicalis*. However, *C. albicans* and *C. tropicalis* infections accounted for only 61% of all clinical isolates in 1997 (Pfaller *et al.*, 1998), compared with 65% in 1996 (Nguyen *et al.*, 1996) and 80% in 1994 (Wright, 1994). The emergence of non-*C. albicans* species as increasingly frequent pathogens has major clinical implications given the higher rates of complication, antifungal resistance and mortality associated with these opportunists, in addition to the increasing number of immunocompromised hosts. Consequently it is important to identify any discrete virulence factors which distinguish these pathogenic *Candida* species from their non-pathogenic counterparts, thereby increasing the spectrum of targets for antifungal drug design.

1.4 *C. albicans* and host immunity

Despite the prevalence of *C. albicans* infections, our understanding of the interplay of yeast pathogenicity factors with host defence systems, which determine the progression from mucosal colonisation to symptomatic infection, is limited. Up to 80% of individuals may host *C. albicans* in the mucosal surfaces of the mouth, gut and vagina, and most infections arise endogenously from this source, yet only a small proportion of these individuals develop candidosis (Odds, 1988). Therefore host defence systems at mucosal surfaces are obviously important in preventing invasion by the yeast. Innate immunity, humoral immunity and cell-mediated immunity (CMI) are all involved in host defence against *C. albicans*. Each mechanism appears to be site-specific, and clinical and animal model studies have highlighted the roles of each defence mechanism in preventing infection (for review see Greenfield, 1992; Fidel & Sobel, 1994).

Innate immunity, such as phagocytosis by macrophages, natural killer cells and polymorphonuclear neutrophils, is associated with protection against systemic candidosis and defects in this system (e.g. neutropenic patients) results in a high incidence of invasive and systemic candidosis (Odds, 1988). In contrast, humoral immunity - the bactericidal and opsonizing ability of serum - has a role in protection against mucosal *C. albicans* infections. Indeed, components of the *C. albicans* cell wall are known to activate the complement which in turn stimulates phagocytosis (Kagaya and Fukuzawa, 1981), and anti-*Candida* IgG or IgA antibodies in surface secretions may have a protective role (De Bernardis *et al.*, 1997). However, systemic and/or local CMI comprises the major host defense against mucosal candidosis. Tissue-specific CMI is an important defence mechanism at mucosal sites and clinical studies have shown that individuals with chronic oral candidosis are not generally susceptible to vaginal candidosis and *vice versa* (Fidel, P.L. Jr. *et al.*, 1994). CMI is acquired immunity and involves CD4⁺ T-cell activation, which is diminished in immunocompromised individuals, such as HIV⁺ patients, resulting in the high incidence of recurrent mucosal infections in these individuals. This might be due to a shift from the normal T_{H1}-type protective response of immunocompetent individuals to the hypersensitivity associated with the T_{H2}-type response immunocompromised patients, which results in susceptibility to infection (Clerici & Shearer, 1993; Shearer & Clerici, 1992).

1.5 Antifungal therapies

The challenge in developing therapies against eukaryotic pathogens such as *C. albicans* is identification of enzymes or processes which are both essential and novel to the pathogen, therefore providing targets for antimycotics without the risk of drug-associated toxicity in the host. Current anti-*Candida* treatments are largely limited to 5-fluorocytosine and two other main classes of compounds (Table 1.1), and of these none are completely satisfactory, either due to toxicity associated with prolonged oral use or easily acquired resistance by the fungus. While some progress has been made towards identification of the molecular mechanisms of resistance in *C. albicans*, such as various efflux pumps (for review see White *et al.*, 1998), no revolutionary advances have been

made in the treatment of candidosis. Therefore the search for better diagnostics and new antifungal agents is continuing, and a brief summary of some current research strategies follows:

- Pradimicins constitute a novel class of broad-spectrum antifungal compounds currently undergoing clinical trials. They act by disrupting the integrity of the cell wall, and initial *in vivo* studies indicate an excellent therapeutic index with no major toxicity (Walsh & Giri, 1997).
- Sikorski *et al.* (1997) are focusing on design of selective inhibitors of Myristoyl CoA:protein N-myristoyltransferase (NMT), an enzyme essential for *C. albicans* viability (Weinberg *et al.*, 1995). A 50% reduction in N-myristoylation of ADP ribosylation factor (Arf) in *C. albicans* has been shown to result in cellular growth arrest (Lodge *et al.*, 1997).
- Other researchers are investigating lipopeptide inhibitors of fungal glucan synthase, which prevent synthesis of 1,3- β -D-glucan; an essential cell wall polysaccharide in *C. albicans* (Kurtz & Douglas, 1997).
- One of the most promising potential therapeutics is exploitation of the role protective antibodies play in host defence against disseminated candidosis (Matthews & Burnie, 1996). Both anti-mannan and anti-Sap (secreted aspartic proteinase) antibodies have been shown to enhance host immunity against *Candida* infections in mouse models (De Bernardis *et al.*, 1997).

COMPOUND	CLASS	MODE OF ACTION	LIMITATION
5-Fluorocytosine		Disrupts nucleic acid biosynthesis	Fungus can easily acquire resistance
Fluconazole Ketoconazole	Azole	Indirect impairment of fungal plasma membrane results from interaction with ergosterol biosynthesis	Limited selectivity for fungal demethylase
Amphotericin B Nystatin	Polyene antibiotic	Directly impairs function of fungal plasma membrane, causing permeability to protons and leakage of ions.	Nephrotoxicity Amphotericin B is prohibitively expensive - mean annual cost of US\$12 000 per patient

Table 1.1 *Anti-fungal agents currently used for treatment of Candida infections*

Data from Vanden Bossche, (1991).

1.6 Phenotypic switching

The capacity for spontaneous phenotypic variation in response to environmental signals, such as developing host immunity, is a feature of many successful pathogens (Soll, 1992). Most strains of *C. albicans* are capable of reversible phenotypic switching between a variety of distinct colony morphologies, which occur at frequencies well above those of point mutation (Soll, 1992). However, in contrast to the switching systems of bacteria and other eukaryotes, switching in *C. albicans* can simultaneously affect an extensive range of cellular characteristics (Soll, 1997).

Switching was discovered in 1985 in *C. albicans* strain 3153A (Slutsky *et al.*, 1985) and strain 1001 (Pomes *et al.*, 1985), and was initially identified by reversible changes in colony morphology. Subsequently, many switch phenotypes have been shown to manifest at the single-cell level as changes in cellular phenotypes, in addition to morphological colony variation. Switching can also affect expression of a number of putative virulence factors (Anderson & Soll, 1987; Kennedy *et al.*, 1988; Soll *et al.*, 1989), and researchers have found elevated rates of phenotypic switching in pathogenic strains compared to commensal strains of *C. albicans* (Jones *et al.*, 1994). Taken together, this supports the proposal that phenotypic switching contributes to the

pathogenicity of *C. albicans* by providing a mechanism by which the fungus may evade environmental challenges and changes.

In 1985, Slutsky *et al.* discovered that colonies of *C. albicans* strain 3153A could exhibit various phenotypes, which include 'star', 'ring', 'stippled', 'hat' and 'original smooth'. Similarly, Pomes *et al.* (1985) demonstrated the segregation of strain 1001 into rough and smooth colonies following mild UV-irradiation. Following this, large numbers of commensal and pathogenic isolates were tested for high-frequency reversible switching. Researchers identified several switching systems, each with a unique repertoire of phenotypes, and found that each strain could express only one of these systems. However, it was the extraordinary white-opaque transition of strain WO-1, a blood-isolate from a bone marrow transplant patient in Iowa city (Slutsky *et al.*, 1985), which proved to be the most manageable, stable system and hence became the most extensively characterised.

The smooth white colonies of strain WO-1 frequently switch to flat grey (opaque) colonies and, less frequently, produce 'irregularly wrinkled', 'medusa' and 'fried egg' phenotypes (Fig. 1.3). The white-opaque transition is reversible, although Rikkerink *et al.* (1988) and Bergen *et al.* (1990) found the switch frequency from white to opaque was lower than that for the opaque to white transition. Notably, these frequencies are influenced by a number of environmental factors, including mild UV-irradiation (Morrow *et al.*, 1989), temperature (Slutsky *et al.*, 1987), colony age (Soll *et al.*, 1991) and the presence of neutrophils and oxidants (Kolotila & Diamond, 1990).

The slightly ellipsoidal shape of white-phase budding cells is similar to that of most other strains of *C. albicans*. In contrast, budding opaque cells are elongate and bean-shaped, are twice as large and exhibit peculiar budding patterns. They also possess variant wall and cytoplasmic morphologies, with punctate 'pimple' structures on the surface and giant internal vacuoles (Soll, 1992). Opaque cells retain the capacity for

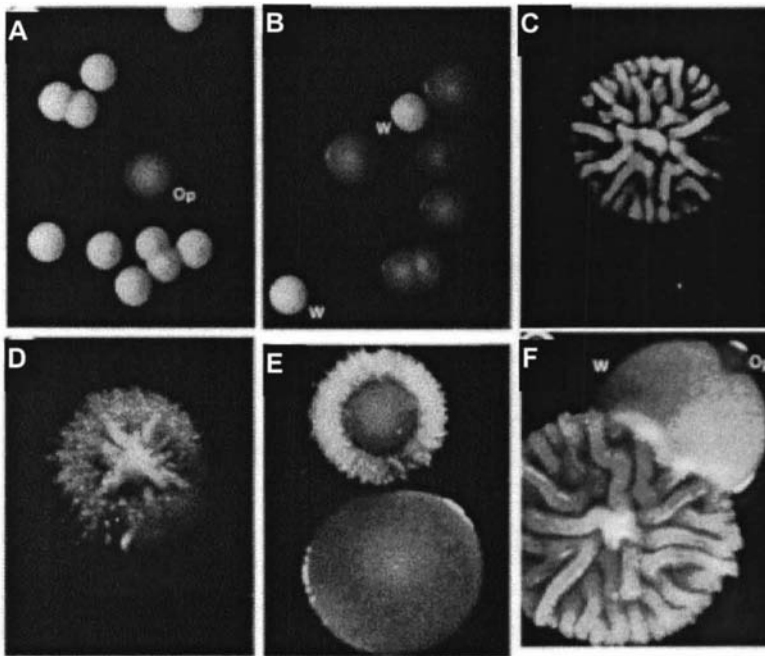


Figure 1.3 *Variant colony phenotypes in the switching repertoire of WO-1*

(A) Switch from white to opaque. (B) Switch from opaque to white. (C) Irregular wrinkle colony. (D) Medusa colony. (E) Fried egg and opaque colonies. (F). Irregular wrinkle colony next to a white colony with an opaque sector. Figure from Soll (1992).

hyphal production, although the environmental constraints on the transition differ markedly from those for white cells. The resultant opaque-cell hyphae lack the punctate pattern of the budding cells and are essentially indistinguishable from white-cell hyphae (Anderson *et al.*, 1989).

1.7 Molecular mechanisms of phenotypic switching

The pleiotropic effects of phenotypic switching can be readily seen by comparison of white- and opaque-phase cells. This reveals differences in antigenicity (Anderson *et al.*, 1990), bud-hypha transition cues (Anderson *et al.*, 1989), adhesiveness (Kennedy *et al.*, 1988), sensitivity to neutrophils and oxidants (Kolotila & Diamond, 1990) and susceptibility to antifungal agents (Soll *et al.*, 1991). Two-dimensional PAGE analysis of white- and opaque-phase polypeptides has also indicated phenotype-specific

expression (Soll *et al.*, 1991). These data indicate that phenotypic switching involves the differential regulation of a variety of unrelated genes. Conceivably then, a reverse genetics approach could be employed to elucidate the switch mechanism by cloning phase-specific genes and tracing the regulation of these genes back to the “master” switch event. *C. albicans* strain WO-1 is an ideal experimental model for this approach, and several phase-specific genes have recently been cloned from this strain.

The first opaque-specific gene cloned was *PEP1*, named for homology to known pepsinogens (Morrow *et al.*, 1992). This gene has subsequently been renamed *SAP1*, and now belongs to an aspartic proteinase gene family which investigators have demonstrated to be differentially regulated by phenotypic switching (White *et al.*, 1993; Hube *et al.*, 1994). A second opaque-phase specific gene, *OP4*, was cloned by Morrow *et al.* in 1994 and shows no homology to any known genes or proteins. Analysis of the upstream regions of the opaque-phase-specific genes *OP4* and *PEP1*, reveals a common C box which may be involved in gene regulation (Soll *et al.*, 1995). More recently, a strong upstream activation sequence has been identified, which contains a palindromic MADS box consensus sequence, most closely related to the *S. cerevisiae* Mcm1 binding site. In *S. cerevisiae*, Mcm1 functions as a global regulator (Kuo & Grayhack, 1994), which interacts with MAT α 1 and MAT α 2 to activate and repress transcription, respectively (Sprague, 1990). Mcm1 also interacts with Ste12 to activate genes in the *S. cerevisiae* mating process and cell cycle (Oehlen *et al.*, 1996). Three different DNA-protein complexes have been shown to form at the *C. albicans* MADS box, none of which appear to opaque-phase specific, which intimates the possibility that these complexes represent interactions between the MADS box and a *C. albicans* Mcm1-type protein. This suggests that regulation of *OP4* might occur via protein-protein interactions or phosphorylation, rather than by direct interaction with a phase-specific transcription factor (Lockhart *et al.*, 1998).

The first white-phase specific gene isolated, *WH11*, shows homology to the *S. cerevisiae* heat shock protein Hsp12. The gene is regulated by both phenotypic switching and the bud-hypha transition; being active in the white budding phase, turned off in the opaque budding phase, and also switched off by the bud-hyphal transitions of both phases (Srikantha & Soll, 1993). Analysis of the *WH11* promoter regions using

reporter gene constructs revealed two white-phase specific transcriptional activation domains; a distal activation domain (DAD) between -475 and -388 bp and a proximal activation domain (PAD) between -307 and -230 bp. Deletion of either domain caused a reduction in reporter gene expression and deletion of both domains completely abolished expression, indicating that *WH11* is positively regulated in the white phase. Subsequent analyses of PAD and DAD found no evidence for negative regulation of *WH11* in the opaque phase and showed that the *cis*-acting regulatory sequences in both elements are unique (Srikantha *et al.*, 1995).

WH11 is known to be regulated by both switching and dimorphism (Srikantha & Soll, 1993), and although these represent distinct developmental programmes, there is evidence to suggest these programmes may overlap at the regulatory level (Soll, 1997). Further functional characterisation of the *WH11* promoter indicates that *WH11* regulation during the bud-hypha transition is mediated by the same DAD and PAD sequences as the white-opaque transition (Soll, 1997).

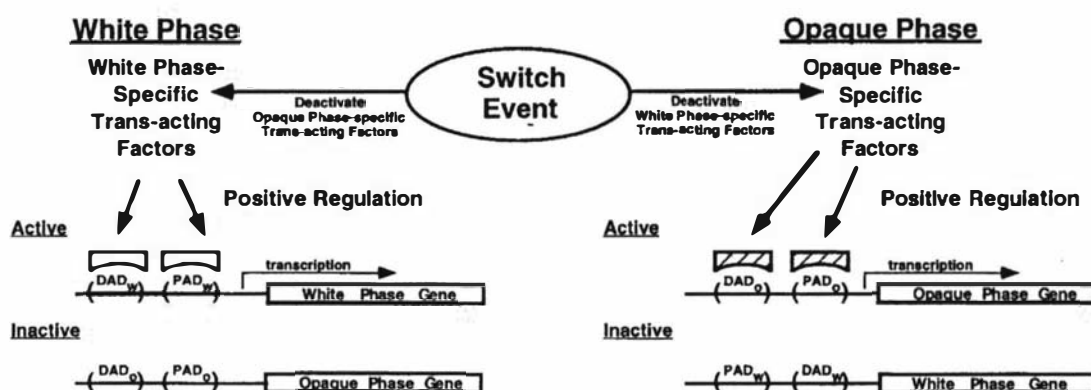


Figure 1.4 A model for the regulation of white- and opaque-phase specific genes

DAD and PAD, distal and proximal activation domains for white (w) and opaque (o) genes, respectively. The inverted unfilled or hatched cups represent white and opaque transcription activation factors, respectively. Diagram from Soll (1997).

Soll (1997) proposed a model for phase-specific gene regulation (Fig. 1.4) which suggests the existence of a 'master switch locus'. The trigger event for such a switch might involve a precise, reversible, heterochromatic change at the master locus (Pilus & Rine, 1989; Aparicho *et al.*, 1991), or a precise, reversible reorganisation of DNA sequences at the master locus (Glasgow *et al.*, 1989). Either mechanism is plausible, given both events have been seen to occur at frequencies similar to those observed in *Candida* switching.

1.8 Molecular mechanisms of polymorphism

A complete definition of the molecular determinants governing the morphological transitions between budding yeasts, pseudohyphae and hyphae in *C. albicans* remains to be elucidated. However, analysis of the environmental cues which induce these transitions, coupled with molecular genetic dissection of the pathways involved, are contributing to our understanding of the links between polymorphism and pathogenesis in *C. albicans*.

Growth of *C. albicans* as ovoid-shaped budding yeast cells is favoured by variables such as low pH, high inoculum, low culture incubation temperature (less than 35 °C) and glucose as carbon source. Conversely, factors which favour filamentous forms of the fungus include neutral pH, high temperature, low inoculum, non-glucose carbon sources, a defined amino acid medium and the addition of serum to the culture medium (Odds, 1988). Similar to phenotypic switching, various reports document differential protein expression during *C. albicans* morphogenesis (Brown & Chaffin, 1981; Niimi *et al.*, 1996). The bud-hypha transition also involves differential expression of many disparate genes, including those for cell wall proteins (Bailey *et al.*, 1996), members of the secreted aspartic proteinase family (Hube *et al.*, 1994) and genes involved in cell elongation (Birse *et al.*, 1993).

Current research is focused on identification and analysis of genes involved in the regulation of *C. albicans* morphology (Magee, 1997). In 1985, Odds proposed the existence of several signalling pathways which could trigger morphogenesis in

C. albicans (Odds, 1985). One pathway regulating *C. albicans* cell shape has now been discovered, and involves *CPH1* (Kohler & Fink, 1996), *HST7*, *CST20* (Leberer *et al.*, 1996) and *CEK1* (Csank *et al.*, 1998); homologs of *S. cerevisiae* *STE12*, *STE7*, *STE20* and *ERK*, respectively. The *S. cerevisiae* genes encode mitogen-activated protein kinase (MAPK) signal cascade components involved in the mating response of haploid cells; interestingly, these signalling components also control pseudohyphal formation in *S. cerevisiae* (Leberer *et al.*, 1996; Kohler & Fink, 1996). The *C. albicans* homologs were cloned by functional complementation of mating-defect mutants of *S. cerevisiae*, and also act in a protein kinase cascade to stimulate the yeast-hypha transition. Deletion of *HST1*, *CPH1* and *CST20* in *C. albicans* results in mutants unable to form hyphae except in response to serum (Leberer *et al.*, 1996; Kohler & Fink, 1996), in contrast to disruption of the *CEK1* gene, which also adversely affects the growth of serum-induced mycelial colonies. These data suggest that in asexual *C. albicans*, this protein kinase cascade might have become adapted to the control of mycelial formation, rather than serving to produce a mating response as it does in *S. cerevisiae*. This adds credence to Odds' hypothesis that more than one pathway can trigger hyphal development in *C. albicans*. Lo and coworkers have recently disrupted both *CPH1* and *EFG1* (refer Fig. 1.5), and the resulting strain was locked in the yeast form, unable to produce hyphae or pseudohyphae in response to many stimuli, and which was avirulent in a mouse model (Lo *et al.*, 1997). This data provides evidence that morphogenesis pathways governing hyphal and pseudohyphal formation, are likely to include some discrete and some overlapping components.

More recently, Leberer *et al.* (1997) again used functional complementation to clone *CaCLA4*, a homologue of *S. cerevisiae* *CLA4* gene. *CaCLA4* encodes a member of the Ste20p family of serine/threonine protein kinases which appear to be involved in triggering morphogenetic processes in response to external signals. It is thought that CaCla4p plays a role in controlling and promoting polarised growth during *C. albicans* budding and hyphal development, respectively. Indeed, deletion of the *CaCLA4* gene showed it to be indispensable for virulence and morphogenesis in *C. albicans*; abolishing hyphal formation both *in vitro* and *in vivo*, and reducing *C. albicans* virulence in a murine model. These results intimate that the bud-hypha transition does in fact contribute to *C. albicans* virulence *in vivo* - perhaps by allowing evasion of the

host immune system or facilitating penetration of host cells during infection (Leberer *et al.*, 1997).

Another system of regulating filamentous growth involves the *C. albicans* *TUP1* gene (a homologue of the *S. cerevisiae* *TUP1* gene) which functions as a general transcriptional repressor, targeting a diverse set of genes in *S. cerevisiae*. Braun and Johnson (1997) discovered that *C. albicans* requires *TUP1* to maintain the yeast form, as deletion of both *TUP1* alleles results in exclusively pseudohyphal growth. Similarly, disruption of the *C. albicans* putative transcription factor *RBF1* (RPG-box binding factor 1), produces pseudohyphal growth on a variety of media. Another report documents the finding that overexpression of the *EFG1* (enhanced filamentous growth) gene results in a very strong pseudohyphal phenotype (Stoldt *et al.*, 1997). A tentative model has been proposed to account for these findings (Magee, 1997), although further investigations will be required to establish its validity (Fig. 1.5).

Interestingly, the above-mentioned deletions, which affect pseudohyphal formation, do not exert the same effects on hyphal formation, suggesting that independent pathways are involved in these two processes. It also seems probable that alterations in chromatin structure represent an important mode of control in *C. albicans* morphogenesis (Fig. 1.5). Taken together, these results suggest that pseudohyphal formation is under negative control and that the yeast form is the result of such negative control over the pseudohyphal state.

In *S. cerevisiae*, the interaction of Tup1 with MAT α 2 and Mcm1 is required for transcriptional repression. A Tup repressor complex (as proposed in Fig. 1.5) might also be involved in repression of *OP4* transcription in the white phase of *C. albicans* WO-1. This could occur through interaction of a multi-protein repressor complex, involving Tup1 and a *C. albicans* Mcm1-like protein, with the MADS box, upstream of the *OP4* gene (Lockhart *et al.*, 1998).

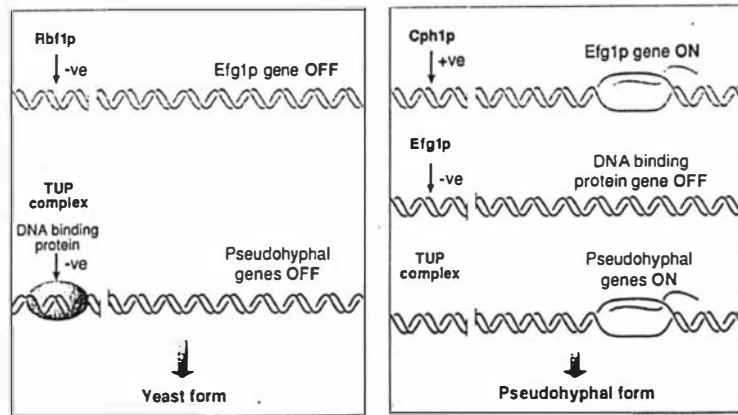


Figure 1.5 *A model for the regulation of pseudohyphal growth in C. albicans.*

Negative regulation of genes required for pseudohyphal growth by a complex consisting of Tup1 and probably other proteins. When this complex is targeted to the pseudohyphal genes by a DNA binding protein, the genes are off (left). When the DNA binding protein is absent, the genes are transcribed and the cell grows in the pseudohyphal form (right). The DNA binding protein is negatively regulated by Efg1; when *EFG1* is overexpressed, synthesis of the DNA binding protein is prevented and the cells are in the pseudohyphal state, as shown on the right. *EFG1* is negatively regulated by Rbf1p and positively regulated by Cph1p when the latter is activated by the MAP kinase cascade. When *RBF1* is deleted, Efg1 is not made and the cells cannot turn on the pseudohyphal genes (from Magee, 1997).

In addition to defining the role of polymorphism in *C. albicans* pathogenicity, these results suggest that regulators of hyphal formation may eventually provide targets for antifungal therapies. For example, Gale and coworkers (Gale *et al.*, 1998) have described linkage of adhesion, filamentous growth and virulence, to a single gene, *INT1*. Overexpression of *INT1* in *S. cerevisiae* is sufficient to confer adhesive capacity, and disruption of *INT1* decreases adherence in *C. albicans*, but does not abolish it completely, indicating other factors are also involved in this process. While deletion of *INT1* also suppressed filamentous growth, Int1p was not required for hyphal formation, and might represent a sensor that triggers the morphogenic decision process in response to a range of environmental cues. Therefore Int1p is involved in both adhesion and morphogenesis, and has a surface location, which makes it an attractive target for potential therapeutics.

1.9 Aspartic proteinases

The four main classes of proteinases are aspartic, cysteine, metallo and serine proteinases, and enzymes are classified according to the nature of their active sites (Kay, 1985). While aspartic proteinases constitute a relatively small group, they are nevertheless ubiquitous in nature, and are involved in a myriad of physiological processes (Davies, 1990). Aspartic proteinases are named for the presence of two active-site aspartic acid residues, and are characterised by their acidic pH optima, inhibition by pepstatin, and having molecular weights in the range of 35-40 kDa (Kay, 1985).

Eukaryotic aspartic proteinases are monomeric, bilobal proteins, which are synthesised as inactive zymogens (Davies, 1990). The N-terminal pro-sequences of zymogens are usually 45-50 amino acids in length and have been shown to participate in protein folding. Cleavage of the pro-sequence can occur by self-processing, and results in enzyme activation (Koelsch *et al.*, 1994). Frequently, aspartic proteinases are secreted, and consequently the pro-sequence is preceded by a signal peptide which is removed during passage into the endoplasmic reticulum, as part of the secretory process (Koelsch *et al.*, 1994).

1.9.1 Evolution

Comparison of the 3-dimensional structures of the N-terminal and C-terminal lobes of eukaryotic proteinases reveals extensive homology. This internal symmetry prompted the hypothesis that these enzymes evolved through gene duplication and fusion (Tang *et al.*, 1978). Retroviral aspartic proteinases are homodimers, which are structurally related to their eukaryotic counterparts, and their catalytic mechanisms are virtually identical. Therefore it has been postulated that the retroviral aspartic proteinases represent the primordial form of the eukaryotic enzymes. As such, the capacity of retroviruses for integration into foreign genomes provides a mechanism for eukaryotic acquisition of the retroviral homodimeric aspartic proteinase gene. The acquisition of this gene might then have conferred a selective advantage, leading to its retention in the host genome. The 'modern' aspartic protease would have ultimately evolved through

gene duplication and fusion to a single-chain, internally symmetrical enzyme (Tang & Lin, 1994).

1.9.2 Substrate specificity

Aspartic proteinases usually possess broad substrate specificities, and generally cleave between two hydrophobic residues (Fruton, 1976). An exception to this is renin, which has an exquisite specificity for the Leu-Val bond of angiotensinogen, and catalyses the conversion to angiotensin I (Sielecki *et al.*, 1989). Some differences in the specificity of cleavage between various members of the family can be attributed to the active site cleft, which is found at the interface between the two lobes of the enzyme. The relatively long active site accommodates at least 7 amino acids, allowing extended binding of β -strand conformation peptide substrates. This facilitates multiple points of interaction between the substrate and enzyme and results in maximal cleavage efficiency (Fig. 1.6) (Dunn *et al.*, 1995).

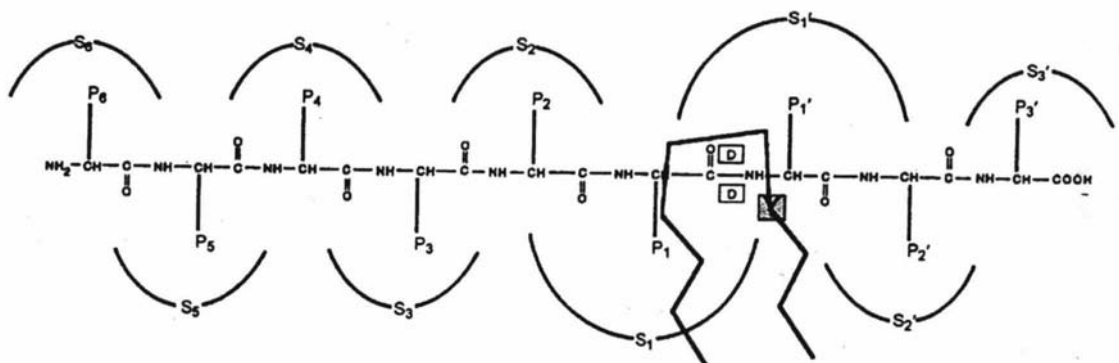


Figure 1.6 Schematic representation of the substrate binding cleft of an aspartic proteinase.

The scissile peptide bond is juxtaposed with the catalytic residues. The two active site aspartic acid residues are shown (D), and the 'flap' which contributes to substrate binding is superimposed over the active site. The position of the conserved tyrosine 75 (pepsin numbering) is depicted by a shaded box. From Dunn *et al.* (1995).

1.9.3 Structural features

The 3-dimensional structures of several aspartic proteinases have been published, including porcine pepsinogen (James & Sielecki, 1986), penicillopepsin (James & Sielecki, 1983), rhizopuspepsin (Suguna *et al.*, 1987), endothiapepsin (Jenkins *et al.*, 1977), human renin (Sielecki *et al.*, 1989), progastricsin (James *et al.*, 1995), yeast proteinase A (Badasso *et al.*, 1993) and cathepsin D (Baldwin *et al.*, 1993). The structural similarity of these enzymes is considerable, and the overall secondary structures are composed primarily of two antiparallel β sandwiches, with very little α -helix in the mature enzyme, which contrasts with the predominantly α -helical prosequences. The two lobes are related by an approximate twofold rotation axis, and are separated by a 40 Å long cleft. In the middle of this extensive cleft are situated two psi loops (Davies, 1990), which extend from each domain and contain the catalytic aspartic acid residues. These catalytic aspartic residues appear in the characteristic sequence motif: hydrophobe-Asp-Thr-Gly-Ser. The active-site region of the molecule is quite rigid, due to connection of the active-site carboxyl groups by a complex network of hydrogen bonds. Low crystallographic *B* factors have confirmed active-site rigidity in this region, and participants in this extensive hydrogen bonding include the conserved active-site threonines (Davies, 1990).

In another highly homologous region (residues 72-85, pepsin numbering), the residues form a conserved hairpin loop, known as the 'flap'. This flap is highly flexible and projects out over the cleft at the active site (Fig. 1.6). It enhances substrate binding by closing down over bound peptides and further excluding solvent from the hydrophobic cleft (Cooper *et al.*, 1987).

1.9.4 Catalytic mechanism

The mechanism by which aspartic proteinases catalyse the hydrolysis of peptide bonds involves general base catalysis, rather than formation of covalent intermediates (James *et al.*, 1977; Pearl, 1987). The catalytic aspartic residues normally share a proton, but this is withdrawn by one of the residues when a substrate binds the active site, which in

turn activates the water molecule (which lies between the substrate and the carboxyl group of Asp32). As previously stated, substrate binding also induces a conformational change in the flap, which effectively brings the conserved Tyr75 (pepsin numbering) close enough to permit involvement in hydrolysis. Attack of the polarised carbonyl carbon atom on the scissile peptide bond is through the activated water molecule, rather than direct nucleophilic attack by the Asp32. This reaction produces a tetrahedral intermediate which is electrostatically stabilised by the nearby phenolic Tyr75 residue. Departure of the products from the binding site is dependent upon the relative affinity of the extended substrate binding pocket for each peptide. The rate of hydrolysis is dependent on substrate amino acid composition and peptide fragment length (Pearl, 1987).

1.10 Retroviral aspartic proteinases

The retroviral aspartic proteinases are encoded as part of a gag-pol fusion protein, and are vital for processing the polyprotein into the separate core proteins. Perturbation of this retroviral proteolytic activity either by site-directed mutagenesis (Kohl *et al.*, 1988) or selective inhibition (Roberts *et al.*, 1990; Sham *et al.*, 1991; Lam *et al.*, 1994) results in the formation of immature, noninfective virions, making these integral proteinases good antiviral drug targets (Kay & Dunn, 1990). The retroviral aspartic proteinases are homodimeric enzymes, which are only active in the self-assembled dimer and are significantly smaller than their mammalian counterparts (Davies, 1990). The presence of the characteristic hydrophobe-Asp-Thr-Gly motif led to their classification as aspartic proteinases. Structural studies of the RSV (Miller *et al.*, 1989) and HIV (Navia *et al.*, 1990) proteinases revealed the active site is formed at the interface of the two monomers, but closely resembles that of the other aspartic proteinases, and showed that the two identical subunits are related by a twofold axis of symmetry (Navia *et al.*, 1990). The active site binding cleft runs across the width of the molecule and due to the symmetric and dimeric nature of these enzymes, there are two flaps which extend into the active site (Davies, 1990). The enzymes require a peptide of at least 7 residues for effective catalysis, and have an unusual propensity to cleave the Phe-Pro and Tyr-Pro bonds found in the gag-pol polypeptide. Since these bonds are not cleavage targets of

mammalian endopeptidases, this unique specificity forms the basis of selective, inhibitor-based drug design. One example of this rational drug design strategy is the nonpeptide cyclic urea inhibitors of HIV proteinase, which displace and mimic the hydrogen bonding features of the catalytic water molecule, thus preventing gag-pol substrate hydrolysis. The success of HIV proteinase inhibitor-based therapies in halting progression of AIDS is reflected in the widespread clinical use of potent inhibitors from various pharmaceutical companies (Bugg *et al.*, 1993). It is hoped on the basis of these results that the strategies employed in targeting the HIV aspartic proteinase can be extrapolated with equivalent success to the enzymes of other human pathogens, including *C. albicans*.

1.11 Fungal aspartic proteinases

Fungal aspartic proteinases are a distinct subgroup of the aspartic proteinase family, being more closely related to each other than to their mammalian equivalents. Nevertheless, a considerable degree of homology does exist between the fungal and mammalian enzymes, supporting the suggestion they all evolved from a common ancestral gene (North, 1982). The enzymes themselves generally have broad substrate specificities, although they preferentially hydrolyse the peptide bonds between two bulky amino acids. These enzymes are frequently unstable at neutral pH or above, but many fungal species which produce extracellular aspartic proteinases also have a tendency to acidify the medium in which they grow. A number of fungal aspartic proteinases have been characterised at the structural level, including penicillopepsin from *Penicillium janthinellum* (James & Sielecki, 1983), Sap2 from *C. albicans* (Cutfield *et al.*, 1995), rhizopuspepsin from *Rhizopus chinensis*, and endothiapepsin from *Endothia parasitica* (Jenkins *et al.*, 1977). Despite significant primary sequence divergence (e.g. 39% identity between rhizopuspepsin and porcine pepsin), structural comparisons have revealed a high degree of secondary and tertiary structural homology, particularly in the region of the active site (Lowther & Dunn, 1995).

1.12 Secreted aspartic proteinases

C. albicans extracellular proteolytic activity was first described by Staib (1965), and enables the yeast to utilise exogenous protein as the sole nitrogen source. This proteolytic activity is now attributed to secreted aspartic proteinases (*SAP* gene; Sap protein), which constitute a multigene family with at least nine members (Monod *et al.*, 1998).

Saps exhibit broad substrate specificity, hydrolysing a range of proteins including albumin and haemoglobin (Ruchel, 1981), and unlike all other aspartic proteinases, *Candida* Saps can also hydrolyse collagen and keratin (Negi *et al.*, 1984). The isozymes are typical members of the pepsin family of aspartic proteinases, being inhibited by pepstatin and having a pH optimum between 2.5-4.5. They are still active up to pH 7.0, but increasing the pH above this causes irreversible 'alkaline denaturation', which involves dimerisation (Wagner *et al.*, 1995). Thermal stability data shows denaturation also occurs rapidly at temperatures above 45 °C (Smolenski *et al.*, 1997). *C. albicans* Saps exhibit anomalous migration on SDS-PAGE, showing apparent molecular masses of 3-6 kDa higher than the deduced masses, which generally range from 37-42 kDa (Smolenski *et al.*, 1997). This phenomenon is not accounted for by glycosylation, as *C. albicans* Saps do not contain covalently linked carbohydrates (Ogrydziak, 1993). Indeed most *Candida* Saps are not significantly glycosylated, with carbohydrates constituting less than 3% of their mass (Ruchel, 1981).

To date, nine *C. albicans* *SAP* genes have been cloned and sequenced (*SAP1-SAP9*) (Hube *et al.*, 1991; Wright *et al.*, 1992; White *et al.*, 1993; Miyasaki *et al.*, 1994; Monod *et al.*, 1994; Monod *et al.*, 1998), and alignments show *SAP1*, *SAP2* and *SAP3* are approximately 75% identical. *SAPs* 4, 5 and 6 represent a distinct subgroup, showing 90% identity to each other, which drops to 65% when compared with *SAPs* 1-3. *SAP8* is most similar to *SAP1* (65%), and *SAP9* has a C-terminal extension which appears to encode a glycosylphosphatidylinositol (GPI)-anchor. *SAP7* is the most

<i>SAP1</i>	<i>SAP2</i>	<i>SAP3</i>	<i>SAP4</i>	<i>SAP5</i>	<i>SAP6</i>	<i>SAP7</i>	<i>SAP8</i>	<i>SAP9</i>	SAP
100%	76%	75%	68%	67%	68%	43%	65%	36%	<i>SAP1</i>
	100%	76%	63%	62%	63%	45%	53%	37%	<i>SAP2</i>
		100%	65%	63%	64%	44%	51%	36%	<i>SAP3</i>
			100%	87%	93%	43%	46%	36%	<i>SAP4</i>
				100%	89%	44%	47%	36%	<i>SAP5</i>
					100%	43%	48%	35%	<i>SAP6</i>
						100%	40%	34%	<i>SAP7</i>
							100%	36%	<i>SAP8</i>
								100%	<i>SAP9</i>

Table 1.2 *Table of identity between SAP sequences from C. albicans*

Identities expressed as percentages based on alignment of the open reading frames of each *SAP* gene using the GCG computing program. See Appendix 10.1 for alignment of the amino acid sequences of Sap1-Sap9.

diverged member of the family; it exhibits only 44% sequence identity to *SAPs* 1-6, and has an open reading frame (ORF) which is 500 nucleotides longer than the other *SAPs* (Table 1.2).

1.13 Secretion of Sap

Studies of protein export in eukaryotic cells have generally been performed with yeast, in particular *S. cerevisiae*, and evidence suggests the Sap proteins in *Candida* spp. follow the classical yeast secretory pathway (Togni *et al.*, 1996). The general features of the process are summarised: Secretion begins with an amino terminal signal sequence, which is both necessary and sufficient to achieve protein secretion. The minimal functional sequence is tripartite, with a basic N-terminal segment, hydrophobic core and a polar C-terminal section (Nothwehr & Gordon, 1990). During translation of the protein on cytoplasmic ribosomes, the signal sequence is recognised by the signal recognition particle (SRP), which prevents further translation until the complex has docked with the SRP receptor of the endoplasmic reticulum (ER). Then, co-

translational insertion of the nascent polypeptide chains into the membrane of the ER occurs, and is accompanied by removal of the signal sequence, which is performed by endopeptidases with high cleavage specificities. Other post-translational covalent modifications occur in the ER, including core glycosylation on asparagine residues, O-glycosylation and disulphide bond formation. The modified polypeptide then folds into its native conformation before it is transported the Golgi body by vesicle intermediates. The efficiency of vesicle-mediated translocation from the ER to the Golgi compartment is often enhanced by the presence of a glycosylated pro-sequence following the protein signal sequence (eg. α -factor pre-pro, see Caplan *et al.*, 1991).

In the Golgi, these pro regions are cleaved by the Kex2 protease (a member of the subtilisin-related proprotein convertases), which recognises the Lys-Arg dipeptide cleavage signal. Recently, a *C. albicans* *KEX2* homologue has been cloned by functional complementation of *S. cerevisiae* mutants (Newport and Agabian, 1997). The role of this Kex2-like endopeptidase in activation of Sap preproteins was investigated by analysis of Sap2 expression by a *kex2/kex2* null mutant. Sap2 expression by the mutant was considerably reduced and the enzyme was abnormally processed, which suggests dependence on Kex2 for maturation. However, the presence of low levels of mature Sap2 suggested that the aspartic proteinase could also be activated by an alternative pathway. The homozygous *kex2/kex2* mutant strain was also unable to produce hyphae, an observation which might be due to either interference with a direct role of Saps in hyphal formation, or ablation of other proteins dependent on Kex2-processing (Newport and Agabian, 1997). The Golgi is also the site for complex outer-chain carbohydrate addition, prior to transport of the proteins through the yeast cell wall.

1.14 Regulation of *SAP* expression

Elucidating the role of Sap isoenzymes in the pathogenesis of *C. albicans* infections requires analysis of the regulation of individual *SAP* genes and characterisation of the protein products. *SAP* gene expression is differentially regulated according to strain, cell type and environmental conditions (Hube *et al.*, 1994). Major environmental

factors which affect Sap expression include pH, temperature, nitrogen source and glucose concentration (Ogrydziak, 1993). Proteinase production is pH-dependent, and Sap expression has been shown to be inversely related to initial culture medium pH (Germaine and Tellefson, 1981), with optimum Sap production in growth media at pH 4.4 (Crandall & Edwards, 1987). Moreover, in culture media above pH 6.0, no Sap activity was detectable, regardless of nitrogen source (Homma *et al.*, 1993). Crandall and Edwards (1987) also demonstrated the temperature-sensitivity of Sap expression, showing increased production at lower temperatures.

Secretion of protease is induced by high molecular weight nitrogen sources such as bovine serum albumin (BSA), haemoglobin, collagen and peptide mixtures (Remold *et al.*, 1968; Ruchel, 1986; Kaminishi *et al.*, 1988; Banerjee *et al.*, 1991), and suppressed in response to low molecular weight nitrogen sources (Remold *et al.*, 1968). The mechanism by which Sap expression is regulated is not well understood, and wide acceptance of Sap induction by BSA as the sole nitrogen source, and repression by ammonium ions (Ross *et al.*, 1990), does not appear to be absolute. Lerner and Goldman (1993) reported that induction of proteinase activity still occurred in the presence of $(\text{NH}_4)_2\text{SO}_4$, and Banerjee *et al.* (1991) have shown a gradual decrease in Sap expression following addition of $(\text{NH}_4)_2\text{SO}_4$.

It appears Sap induction elicits secretion of newly synthesised proteinase rather than stimulating release of stored proteinase from exocytotic secretory vesicles (Ross *et al.*, 1990; Banerjee *et al.*, 1991), although intracellular precursors of Sap have been shown to exist (Homma *et al.*, 1992). Fully intact protein can induce Sap production, although high-level induction of proteinase requires smaller peptides. This indicates that Sap expression is stimulated by the peptides generated by proteolysis of BSA, rather than by BSA itself. Lerner and Goldman (1993) reported Sap induction by peptides of 8 or more residues in length, but not in response to smaller peptides. Preliminary data suggested that induction was favoured by the presence of a His/Lys-Pro motif in the peptide, and inhibited by glycine-containing peptide sequences. Interestingly, the peptide permeases of *C. albicans* do not efficiently transport peptides greater than 7 residues in length (Milewski *et al.*, 1988), indicating that internalisation of the peptides themselves is not the trigger for Sap induction. This supports the notion of a nitrogen

regulatory mechanism where exogenous protein is somehow detected, followed by transduction of a signal resulting in upregulation of Sap secretion (Lerner & Goldman, 1993). The role such a mechanism might play in *C. albicans* pathogenicity is unclear, as disruption of the major nitrogen regulatory gene in the phytopathogen *M. grisea* revealed it was dispensable for the virulence of that organism (Froeliger & Carpenter, 1996).

The literature concerning the properties of *C. albicans* Sap contains many conflicting reports; variations in apparent molecular weight, pI, inhibitor profile, substrate specificity and N-terminal sequence (White & Agabian, 1995). However the recent discovery of a *SAP* multigene family, comprising at least nine members (Monod *et al.*, 1998), accounts for these discrepancies. As a result of these findings, analyses of the regulation of expression of individual Sap isoenzymes have recently been conducted at the transcriptional level, comparing the abundance of isozyme-specific mRNA transcripts between strains and as a function of cell type and environmental conditions.

When BSA is the sole nitrogen source for budding yeast cells *in vitro*, *SAP2* mRNA is the dominant transcript in most strains, and mRNA levels are unaffected by temperature or carbon source (Hube *et al.*, 1994). *SAP2* mRNA is produced in early to mid-log phase, with high levels 2-4 h after induction by BSA, indicating tight regulation of this gene. Also, yeast-form cells of some strains produce low levels of *SAP3* mRNA in response to standard proteinase-inducing medium (Wright, 1994). In strain WO-1, *SAP1* mRNA expression is coupled to the white-opaque switch; log-phase opaque cells contain high *SAP1* mRNA levels which rapidly decrease upon cell conversion to the white-phase (Morrow *et al.*, 1993). *SAP3* transcript can also be detected in both the log and stationary phases of opaque cells, and *SAP2* mRNA is expressed in the log phase of both white and opaque switch phenotypes. Expression of *SAP1* and *SAP3* is strain-specific, in comparison to *SAP2* which is expressed by all strains tested so far (Hube *et al.*, 1994).

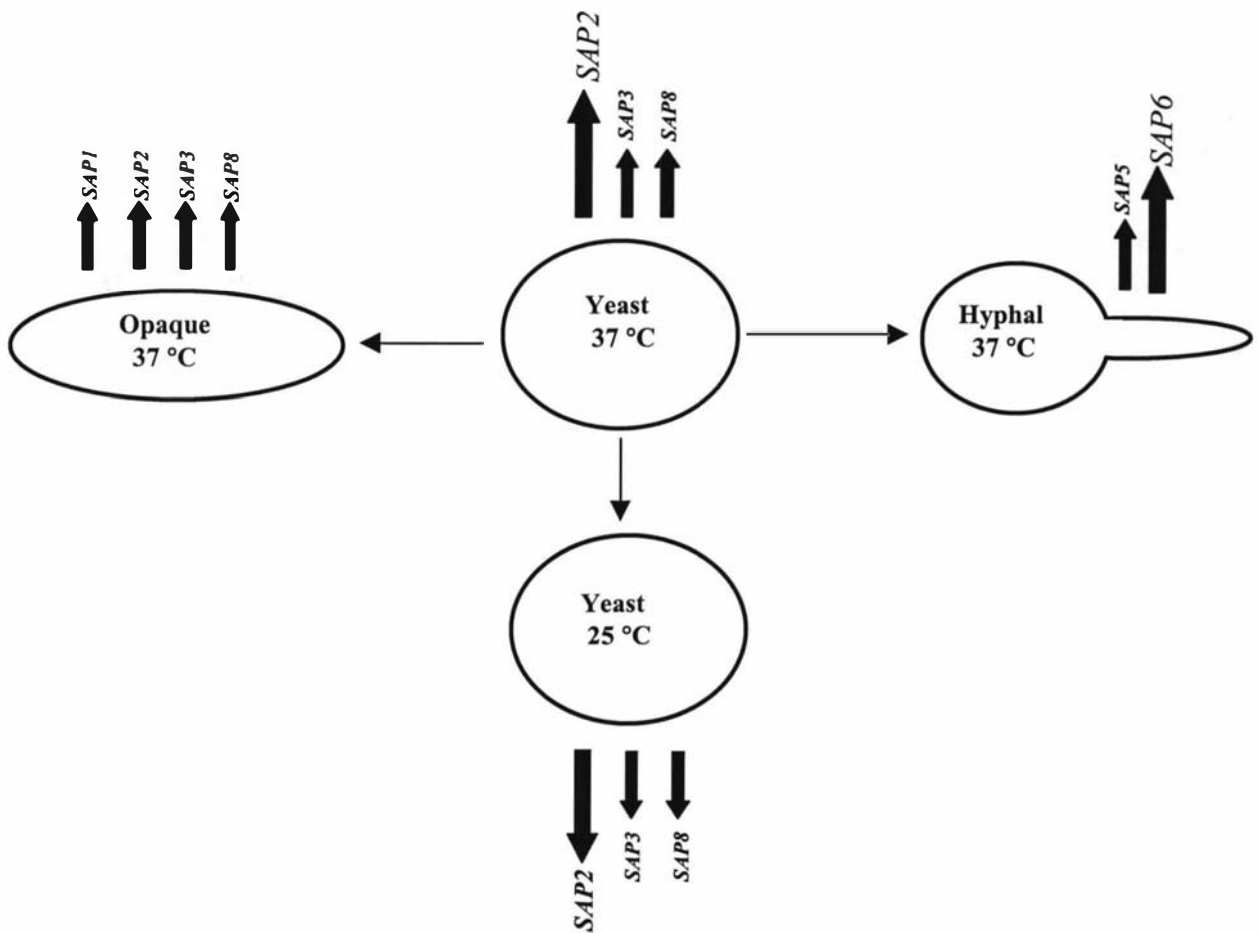


Figure 1.7 Expression of the SAP gene family in *C. albicans*

SAP1, *SAP2* and *SAP3* are expressed in the opaque form of strain WO-1, *SAP2* mRNA is the dominant transcript in the yeast form of *C. albicans*, *SAP3* is also expressed in the yeast form of some strains, while the expression of *SAP4-SAP6* is closely correlated to the hyphal formation of most strains. *SAP7* was found to be silent in each strain and under all conditions tested, while *SAP8* is preferentially expressed in the yeast form at low temperature. Diagram adapted from Hube *et al.* (1998).

Transcription of *SAP4*, *SAP5* and *SAP6* is also strain dependent, and pH 6.0 or above is required to achieve high levels of mRNA. Transient expression of *SAP4-6* mRNA has been detected during serum-induced hyphal production, and more recently, Sap6-specific antibodies have been used to demonstrate high expression of Sap6 antigen on the surface of *C. albicans* cells, following phagocytosis by murine macrophages (Borg-von Zepelin *et al.*, 1998). The expression of *SAPs 4-6* will be addressed in greater detail in Section 1.16.

To date, *SAP7* mRNA remains undetected under any of the conditions, or in any of the strains tested, although expression under other conditions cannot be excluded. *SAP8* mRNA is expressed preferentially in the early exponential growth phase of yeast cells at 25 °C, but is also detectable at lower levels in cultures at 37 °C. *SAP8* transcript is also abundant in WO-1 opaque phase cells at 25 °C. In contrast, a low level of *SAP9* mRNA has been detected in culture conditions tested (Monod *et al.*, 1998).

1.15 Sap and pathogenicity

As previously discussed, *C. albicans* can become pathogenic in response to diminution of the host immune system. It is widely recognised that *C. albicans* has several virulence determinants which contribute to pathogenesis (Cutler, 1991), and although evidence supporting a role for Sap in *C. albicans* virulence is extensive, it is not without controversy.

The *C. albicans* Saps have broad substrate specificities which appear to augment pathogenesis. For example, Colina *et al.* (1996) demonstrated Sap-mediated degradation of gastrointestinal mucin, which suggests a role for Sap activity in *C. albicans* penetration of the mucin barrier. *C. albicans* proteinases also appear to break down the host humoral defence mechanisms through degradation of host IgG and C3 proteins, thereby facilitating debilitation of the infected host (Kaminishi *et al.*, 1995).

Secreted proteinase has been detected in the tissues of infected hosts, and a high titre of anti-Sap antibodies is typical of patients with disseminated candidosis (MacDonald &

Odds, 1980), indicating that Sap is produced *in vivo*. While this is not proof of a role for proteinase in virulence, the finding that anti-Sap antibodies provide increased protection against *C. albicans* challenge in mice (De Bernardis *et al.*, 1997) is less equivocal.

Proteinase expression is correlated with the ability of strains to adhere to host tissue (Cutler, 1991), and Sap has been shown to associate with the cell wall, which suggests a role for Sap in the attachment of *C. albicans* to host surfaces. This proposal is reinforced by pepstatin A-inhibition of *C. albicans* adhesion to mucosal surfaces and epidermal cells (Borg & Ruchel, 1988; Ray & Payne, 1988; Ollert *et al.*, 1994). Potential mechanisms by which Sap might promote adherence includes functioning non-enzymatically as a ligand for attachment to host cells, or by enzymatic modification of either host- or *C. albicans*-cell surfaces to facilitate subsequent fungal attachment.

The neutral pH of saliva inactivates secreted aspartic proteinases, and this led researchers to believe proteolytic activity was insignificant in oral candidosis (Cutler, 1991). However, expression of *SAP4-SAP6* during hyphal induction at neutral pH (Hube *et al.*, 1994), coupled with the discovery of Sap antigens on hyphal cell surfaces (Borg & Ruchel, 1988), implies that Sap4-6 might be involved in mediating adhesion or invading host tissues. While the pH dependence of Sap activity precludes a role in *C. albicans* virulence at sites of neutral pH, association with the cell wall might allow Sap enzymatic activity due to a localised decrease in pH (as a consequence of carbohydrate utilisation and production of acidic by-products) (Samaranayake *et al.*, 1984).

Several early studies document a correlation between levels of Sap expression and virulence of the strain (MacDonald & Odds, 1980; Ruchel *et al.*, 1992; Louie *et al.*, 1994). A comparison of proteolytic activity from vaginal isolates showed the strains from patients with vaginitis produced the highest levels of Sap activity (Cassone *et al.*, 1987), although in direct contrast, other researchers found no such correlation in an equivalent experiment (Schreiber *et al.*, 1985). More recently, Ollert *et al.* (1995) have shown increased proteinase activity in strains isolated from human immunodeficiency virus (HIV)-positive patients, compared with a negative control group. De Bernardis *et*

al. (1996) have also reported elevated levels of proteolytic activity produced by strains of *C. albicans* isolated from patients with AIDS. Following from these results, they concluded that during the course of HIV infections, more virulent strains of *C. albicans* are selected, as determined by direct analysis of virulence factors such as Sap (De Bernardis *et al.*, 1996), and this proposal is supported by molecular epidemiological studies (J. Schmid, personal communication).

Analysis of proteinase-deficient mutants provides the most convincing evidence for Sap activity as a pathogenicity determinant. Kwon-Chung *et al.* (1985) reported reduced virulence of a nitrous acid-generated, proteinase-deficient mutant in mice and restoration of virulence in a proteinase-producing revertant. Similarly, Ross *et al.* (1990) produced a stable Sap-deficient mutant which was 1000-fold less virulent than wild-type strains in animal infections. Despite these convincing data, proteinase-deficient strains obtained by chemical mutagenesis may harbour genetic lesions at other loci which account for the observed decrease in virulence of the strain, and therefore do not provide unequivocal support for the role of Sap in pathogenesis.

The recently developed Ura-blaster protocol facilitates sequential disruption of targeted alleles by using Ura3⁻ auxotrophy as a recyclable, selectable marker (Fonzi & Irwin, 1993). In diploid, asexual yeast such as *C. albicans*, this method permits the construction of genetically defined mutants in an isogenic background. Recently, this technology has been applied to members of the *SAP* multigene family, in an attempt to characterise the role of these enzymes in *C. albicans* virulence (Hube *et al.*, 1997; Sanglard *et al.*, 1997). Hube *et al.* (1997) have reported construction of three homozygous null mutant strains by targeted disruption of either the *SAP1*, *SAP2* or *SAP3* genes. Simultaneously, Sanglard *et al.* (1997) document construction of a strain harbouring a triple deletion of the *SAP4*, *SAP5* and *SAP6* genes in *C. albicans*. Examination of these strains revealed the *sap1* and *sap3* mutants grew slowly in proteinase-inducing medium, and proteinase production was delayed compared with the parent strain. In contrast, *sap2* null mutants grew poorly, tended to clump and produced the least proteolytic activity, supporting the view that Sap2 is the dominant isoenzyme (Hube *et al.*, 1997). Surprisingly, the *SAP4-6* disruptant showed the same growth deficiencies as the *sap2* mutant in Sap2-inducing medium, conditions under which

SAP4-6 expression had not previously been detected. This suggests that Sap4, Sap5 or Sap6 might play an important role in the process of Sap2 induction. This strain also showed significantly attenuated virulence in animal models, indicating the importance of Sap4, Sap5 and Sap6 isoenzymes in the normal progression of animal candidosis (Sanglard *et al.*, 1997). Using the same animal models, Hube *et al.* (1997) showed attenuated virulence for each of the *sap1*, *sap2* and *sap3* mutant strains. Interestingly, virulence of the *sap2* mutant was attenuated to a lesser extent than virulence of the *sap1* and *sap3* null mutants, indicating a lack of correlation between proteolytic activity *in vitro*, and reduction of virulence of the mutants *in vivo*. Reasons for this are unclear, although it may be due to factors such as *in vitro* Sap expression being somehow dependent on Sap2 activity, or *in vivo* upregulation of Sap1 and/or Sap3 expression during the infection process. Nonetheless, these results clearly show that in addition to Saps 4-6, each of Sap1, Sap2 and Sap3 also contribute to *C. albicans* pathogenicity *in vivo*.

In addition to their presumed function facilitating tissue penetration, evidence suggests certain Saps may fulfil an auxiliary role in promoting adherence of *C. albicans* to various host cell types. Analysis of the adherence of each of the *sap1*, *sap2* and *sap3* null mutants to human buccal epithelial cells was reduced, whereas the adherence of the *sap4-6* mutant strain was substantially increased. The enhanced adherence of the Sap4-6 mutant may be due to an absence of proteolytic activity, which normally degrades components which would otherwise promote host-fungus adhesion (Watts *et al.*, 1998).

1.16 Sap4, Sap5 and Sap6

The existence of Sap4 was discovered by Miyasaki *et al.* (1994) while attempting to identify upstream sequences responsible for the differential transcription of the *SAP1* gene in various laboratory strains. The gene was identified while subcloning and sequencing fragments from a WO-1 genomic DNA library, and was shown to exist upstream in tandem to *SAP1* in all laboratory and clinical strains tested. As a prelude to addressing the role of each Sap enzyme in virulence, Monod *et al.* (1994) sought to

isolate all the members of the recently identified multigene family. The approach involved screening a genomic library (constructed from *C. albicans* clinical isolate strain C74) at low-stringency, using a *SAP1* probe. Three new members were identified, *SAP5*, *SAP6* and *SAP7* and were localised to chromosome 6 (*SAP5* & *SAP6*) and chromosome 1 (*SAP7*). *SAP5* and *SAP6* were found to be very closely related to each other and to *SAP4* (refer Table 1.2), while *SAP7* was highly diverged from the other *SAPs*, and possessed a much longer prosequence, although the active site aspartic acid domains were well-conserved. In addition to these findings, the existence of two further *SAP* genes was proposed based on Southern blot hybridisation patterns. These genes have now been cloned and sequenced, and identified as *SAP8* and *SAP9*. *SAP8* encodes the 41 kDa protein previously detected by Morrison *et al.* (1993a, 1993b), while *SAP9* encodes a putative GPI-anchor at the C-terminus which suggests this Sap would remain membrane-bound.

The existence of at least nine Sap isoenzymes in the *C. albicans* genome suggests that these proteins might each have discrete roles in the growth and virulence of the yeast. The differential expression of *SAP4*, *SAP5* and *SAP6* has been studied *in vitro*, using northern analyses (Hube *et al.*, 1994). Strains were induced to form germ-tubes by incubation in 5% calf serum, which was either unbuffered (pH 7.0), or buffered to pH 6.0 or pH 4.5. Under these conditions, germ-tube formation was 90%, 50% and 5%, respectively. *SAP4-6* transcripts were detected 60 min after inoculation into the germ-tube-inducing medium, and transcript levels increased up to 120 min before decreasing markedly 180 min after inoculation. Unlike *SAP1*, *SAP2* and *SAP3*, expression of *SAP4-6* was not repressed by addition of amino acids to the culture medium, suggesting a different mechanism of gene regulation. Also, the expression of *SAP4-6* was not universally linked to germ-tube formation, as pH/temperature-regulated yeast to hyphal transition failed to induce detectable expression of *SAP4-6*. However, there was an apparent correlation between percent germ-tube formation (>90%) and *SAP4-6* expression, whereas strains which formed less than 5% germ-tubes did not express *SAP4-6* transcript, although two strains did form germ tubes efficiently yet expressed *SAP4-6* poorly, suggesting that *SAP4-6* expression was normally, but not universally coupled to serum-induced hyphal formation. Hube *et al.* (1994) did not distinguish between transcripts of *SAP4*, *SAP5* and/or *SAP6*.

Subsequently, White and Agabian (1995) used S1 nuclease analysis to determine which of the *SAP4-6* genes was being expressed during serum-induced hyphal production at neutral pH. Their results indicated that *SAP6* was the dominant transcript, with lower levels of *SAP5* mRNA, and no detectable *SAP4* mRNA. In addition, they report transcription of *SAP5* and *SAP6* in at least one clinical isolate which was induced to form mycelia in a defined amino acid medium (Lee *et al.*, 1975) at neutral pH. This suggests that, at least in some instances, pH and hypha production alone is sufficient for expression of *SAP5* and *SAP6* mRNA. While this contradicts the dependence on serum components for *SAP* induction reported by Hube *et al.* (1994), neither system appears to universally induce expression of *SAP5* and *SAP6* in all strains. Taken together, these findings suggest the existence of a complex system of *SAP* induction, which is not strictly dependent on a single factor, but the interplay of a variety of signals.

More recently, Borg-von Zepelin *et al.* (1998) reported expression of high levels of Sap4, Sap5 and Sap6 following phagocytosis of *C. albicans* by murine peritoneal macrophages. Fresh mouse peritoneal macrophages were infected with either wild-type *C. albicans* (SC5314) or Δ sap mutants which had at least two of the *SAP4*, *SAP5* or *SAP6* genes disrupted. After 2 h incubation, the macrophages had engulfed almost all the blastoconidia, after 4 h the phagocytosed *C. albicans* began to resist the macrophages by forming germ tubes and after 8 h cells were examined for expression of proteinase using immunofluorescence staining and anti-Sap4-6 antibodies. This demonstrated the presence of Sap4-Sap6 antigens on engulfed blastoconidia and pseudomycelia of SC5314, and an absence of Sap4-Sap6 expression in either SC5314 pseudomycelia outside the macrophages or phagocytosed Δ sap_{4,5,6} mutant cells. The Δ sap_{4,5, Δ sap_{5,6} and Δ sap_{4,6} mutants all showed significant immunofluorescence, albeit at reduced levels, indicating that each of Sap4, Sap5 and Sap6 are produced in phagocytosed *C. albicans*. No expression of Sap1, Sap2 or Sap3 was detected. A comparison of the viability of SC5314 and the Δ sap_{4,5,6} mutant cells following phagocytosis revealed just 18.6% of the mutant cells survived compared with 43% of wild-type cells. This two-fold decrease in the survival of the mutant after contact with the macrophages indicates a lower pathogenic potential for this strain and further implicates Sap4, Sap5 and Sap6 in the virulence of *C. albicans*.}

1.17 Sap structure

The increasing frequency of *C. albicans* infections, and the emergence of antifungal resistance among strains, has prompted the search for new antifungal therapies. As *C. albicans* virulence factors, Saps represent a logical antifungal drug target, and analysis of the crystal structure of Sap2 has implications for the design of potent anti-*Candida* drugs.

The Sap2 structure has been determined to 3.0 Å in complex with pepstatin A, and to 2.1 Å in complex with the tight-binding inhibitor A-70450 (Cutfield *et al.*, 1995). Recently, the crystal structure of Sap2X (purified from strain val-1, Sap2X has 96% amino acid identity with Sap2), also complexed with A-70450, has been determined to 2.5 Å (Abad-Zapatero *et al.*, 1996). These crystallographic studies have revealed variations between the Saps and the archetypical aspartic proteinase structures, which dramatically affect the specificity of this family of isoenzymes. The main structural alterations, as reported by Cutfield *et al.* (1995) are:

- An 8-residue insertion near the first disulphide (Cys45-Cys50; pepsin numbering), creates a broad flap which consequently forms a new boundary for the extended binding cleft.
- An enlarged S3 pocket results from the deletion of a small helical segment which normally projects into the binding cleft (residues 110-116, pepsin numbering).
- A C-terminal extension of 11 residues brings the carboxy terminus in close proximity to the N-terminus.
- Deletion of a short (4-7 residue) looped segment, alters the relative position of the two lobes, and this affects the boundary of the extended binding site. The residues thought to be involved in this are conserved in the *SAP1-3* sequences, but not in *SAP4-6*, suggesting structural differences between these two Sap sub-groups.

Of all the variations apparent from the Sap2 and Sap2X crystal structures (when compared with classic aspartic proteinase models), the enlarged S3 subsite is potentially the most important for antifungal drug design. The substantially extended Sap S3 pocket can accommodate significantly larger P3 inhibitor groups, compared to the S3

clefts of mammalian aspartic proteinases, and therefore represents the greatest source of specificity for drug design (Abad-Zapatero *et al.*, 1998).

1.18 Project aims

Sap 4, 5 and 6 have to date only been detected at the transcriptional level, elucidation of the roles these enzymes play in the onset and development of candidosis, requires characterisation of the properties of the proteins, as well as an understanding of the regulation of the encoding genes. The aims of this project were firstly, to clone the genes encoding Saps 4, 5 and 6, followed by heterologous expression in a suitable system. Secondly, to devise protein purification strategies, so that the enzymes could be characterised, and crystallisation trials would be possible. Thirdly it was envisaged that production of polyclonal antibodies (which recognise Sap 4-6) would facilitate further studies of the regulation of these genes, and localisation of the native proteins in *C. albicans*.

2.0 MATERIALS & METHODS

2.1 Enzymes

T4 DNA ligase, T4 DNA polymerase T4 polynucleotide kinase were purchased from New England Biolabs. Expand™ PCR polymerase, proteinase K, calf intestine alkaline phosphatase, and DNA polymerase (Klenow fragment) were supplied by Boehringer Mannheim. Zymolyase 20T and 100T were provided by Seikagaku Kogyo Company. Ribonuclease A (from bovine pancreas, type I-AS) was obtained from Sigma Chemical Company. Restriction enzymes and their appropriate buffers were supplied by Amersham International, Boehringer Mannheim, Gibco BRL and New England Biolabs.

2.2 Chemicals

Ampicillin (sodium salt), 5-bromo-4-chloro-3-indolyl- β -galactosidase (BCIG), dGTP, dCTP, dATP, dGTP were provided by Boehringer Mannheim. Isopropyl-1-thio- β -D-galactoside (IPTG) was sourced from Interned Scientific. Agarose was obtained from

Seakem. Difco Laboratories supplied bacto-tryptone. Xylene cyanol FF, ethidium bromide, Nitro blue tetrazolium (NBT), Coomassie Brilliant Blue R-250, dithiothreitol (DTT), azocoll, ficoll (type 400), pepstatin, phenylmethylsulfonyl fluoride (PMSF) and 3-(N-morpholino)propane sulfonic acid (MOPS) were all purchased from Sigma. Morpholinoethane sulfonic acid (MES) was from Serva. Geneticin (G418), 1kb DNA sizing ladder, bacteriological agar, yeast extract, casein hydrolysate and Freund's adjuvants were supplied by Gibco BRL. Acrylamide Acrylogel 2.6 and PEG 6000 were bought from BDH Chemicals. $\alpha^{32}\text{P}$ -dCTP (3000Ci/mmol) was obtained from Amersham International. Acetylated BSA was procured from Promega. All other chemicals were of analytical grade. Glucagon was purchased from Eli Lilly, Inc.

2.3 Miscellaneous products

Hybond N+, Hybond C+ and Redi-prime DNA labelling kits were from Amersham International and Ready-to-Go DNA labelling beads (used in the latter stages of this work) were from Pharmacia Biotech. Probe-Quant G-50 Micro columns were obtained from Pharmacia Biotech. Qiagen supplied the Qiaex II gel extraction kit for purification of size fractionated DNA fragments. Fujifilm RX X-ray film was purchased from Fuji Photo Film Co., Ltd.

2.4 *Escherichia coli* phenotypes

DH5 α : *supE44* Δ *lacU169* (ϕ 80 *lacZ* Δ *M15*) *hsdR17* *recA1* *endA1* *gyrA96* *thi-1* *relA1*

BL21(DE3)^c: *hsdS* *gal* (λ *cl**ts857* *ind1* *Sam7* *nin5* *lacUV5*-T7 gene 1)

AD202: *F-* *ompT::kan* *araD139* Δ *argF-lacU169* *rpsL150* *relA1* *flbB5301* *deoC1* *ptsF25* *rbsR*

Further information concerning strains DH5 α and BL21 can be found in Brown (1991), and AD202 is described by Akiyama & Ito (1990).

2.5 Yeast strains

<i>Yeast Strain</i>	<i>Description</i>
<i>Candida albicans</i> ATCC 10261	American Type Culture Collection
<i>Candida albicans</i> ATCC 10231	American Type Culture Collection
<i>Candida albicans</i> A72	Clinical isolate provided by Dr. A. Cassone
<i>Pichia pastoris</i> KM71	Invitrogen

2.6 Plasmids utilised during this study

<i>Plasmid</i>	<i>Features</i>	<i>Reference</i>
pBluescript KS+ phagemid	<i>E. coli</i> cloning vector	Stratagene
pPIC3K	<i>P. pastoris</i> integrative plasmid for cytoplasmic protein expression	Invitrogen
pRSET (A)	<i>E. coli</i> expression vector for strain BL21, 6x His tag at N-terminus	Invitrogen

2.7 Plasmids developed during this study

<i>Plasmid</i>	<i>Feature</i>
pMS5b	PCR-generated <i>SAP 5</i> ORF in pBSKS+
pMS4b	PCR-generated <i>SAP 4</i> ORF in pBSKS+
pMS5	1.4 kb <i>SAP5</i> ORF ligated into pPIC3K at <i>Bam</i> HI site upstream of ATG
pMS4	1.4 kb <i>SAP4</i> ORF ligated into pPIC3K at <i>Bam</i> HI site upstream of ATG
pMS1	1.2 kb <i>SAP1</i> ORF ligated into pPIC3K
pMS4Ec	<i>SAP4</i> ORF ligated into pRSET (A) at <i>Bam</i> HI- <i>Pst</i> I

2.8 Growth media

2.8.1 Yeast media

YD: 0.2% yeast extract, 2% dextrose.

YPD: 1% yeast extract, 2% peptone, 2% glucose.

YPG: 1% yeast extract, 2% peptone, 1% glycerol.

Modified YPG: 0.2% yeast extract, 2% peptone, 1% glycerol

For plates, agar was added to 2% prior to autoclaving

10x Wickerhams minimal medium (pH 4.0) stock solution (1 L):

10 g KH_2PO_4 , 5 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g NaCl , 2 g $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$, 50 g $(\text{NH}_4)_2\text{SO}_4$, glucose, vitamins and trace elements (final concentration 2%, x1 and x1 respectively) were added to 1x Wickerhams medium.

200x Vitamins stock (1 L):

20 mg biotin, 400 mg Ca pantothenate, 2 mg folic acid, 2 mg inositol, 200 mg nicotinic acid, 400 mg pyridoxine hydrochloride, 200 mg riboflavin, 400 mg thiamine.

1x Minimal methanol:

1x Wickerhams minimal medium, 1% caseamino acids, 0.5% methanol added every 24 h.

2.8.2 Bacterial media

LB (1 L): 10 g tryptone, 5 g yeast extract, 5 g NaCl .

TB (1 L): 12 g tryptone, 24 g yeast extract, 4 mL glycerol, 100 mL 0.17 M KH_2PO_4 /0.72 M K_2HPO_4 .

2.8.3 General buffers

20x SSC: 3 M NaCl, 0.3 M Na citrate adjusted to pH 7.0 with NaOH.

10x TAE: 40 mM Tris-acetate pH 6.7, 1 mM EDTA.

Church and Gilbert buffer (1 L):

0.5 M sodium phosphate buffer ($\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ pH 7.2), 200 μL

0.5 M EDTA pH 8.0, 7 g SDS per 100 mL buffer.

TE buffer: 10 mM Tris-HCl pH 8.0, 1 mM EDTA.

2.9 General methods

2.9.1 Restriction endonuclease DNA digestion

Restriction endonuclease digests were performed using the supplied buffers, according to the manufacturers' instructions.

2.9.2 Agarose gel electrophoresis of DNA

Ficoll loading buffer (0.25% xylene cyanol FF, 0.25% bromophenol blue and 15% Ficoll type 400) was mixed with DNA prior to size fractionation in electrophoresis-grade agarose with 1xTAE running buffer as outlined by Sambrook *et al.* (1989). Gels were run at 100V until the bromophenol blue dye front was three-quarters down the gel.

2.9.3 DNA ligations

DNA ligations were performed as described by Ausubel *et al.* (1992).

2.9.4 Creating blunt-ended DNA

5' Overhangs were repaired using DNA polymerase (Klenow fragment) while 3' termini were removed with T4 DNA polymerase as specified by Sambrook *et al.* (1989).

2.9.5 Preparation and transformation of *E. coli* competent cells

2.9.5.1 Preparation

Preparation of competent *E. coli* cells according to the method of Inoue *et al.* (1990) involved inoculation of 10-12 large colonies into 300 mL of culture medium (2% tryptone, 0.5% yeast extract, 0.05% NaCl, 0.25 mM KCl, 10 mM MgCl₂) in a 1 L flask which had previously been sterilised containing MQ water. The culture was grown at 18 °C, 200 rpm until the OD₆₀₀ reached 0.6 then the cells were placed on ice for 10 min before centrifugation in sterile 50 mL tubes at 4 000 rpm for 10 min. Each pellet was resuspended in 4 mL ice-cold TB (10 mM PIPES, 55 mM MnCl₂, 15 mM CaCl₂, 250 mM KCl pH 6.7, filter sterilised) using sterile, blunt 5 mL autopipette tips. Following centrifugation at 4,000 rpm for 10 min 0.3 mL cool DMSO was added dropwise to each 4 mL of cells and stored on ice for 10 min. Aliquots were stored in microcentrifuge tubes, snap frozen in liquid nitrogen and stored at -70 °C.

2.9.5.2 Transformation

Aliquots of competent cells were thawed on ice, and either 5 µL plasmid DNA or 10 µL of ligation mix was mixed with the cells and incubated on ice for 15-30 min. Cells were heat-shocked in a 42 °C waterbath for 45 s then chilled on ice for 2 min. Pre-warmed LB medium (0.9 mL) was added to each aliquot of cells, then incubated at 37 °C for 2 h with shaking. Cells were centrifuged at 14,000 g for 30 s, the supernatant was removed and the cells were resuspended into 50 µL of medium and plated onto selection medium.

2.9.6 Selection of transformants

Blue-white selection was performed by plating transformed cells onto agar plates containing ampicillin at a final concentration of 75 $\mu\text{g}/\text{mL}$, BCIG to 30 $\mu\text{g}/\text{mL}$ and IPTG to 20 $\mu\text{g}/\text{mL}$ final concentration. Plates were incubated overnight at 37 °C. Selection with ampicillin and kanamycin was performed by incorporating the antibiotics into the agar plates at a final concentration of 75 $\mu\text{g}/\text{mL}$. Selection with G418 was carried out by spreading 10 mg/mL G418 (in MQ water) onto agar plates to the desired final concentration and allowing the plates to dry.

2.9.7 Plasmid DNA preparation - small scale

Small-scale *E. coli* plasmid DNA preparations were performed using the alkaline lysis mini-preparation method of Ausubel *et al.* (1992).

2.9.8 Plasmid DNA preparation - large scale

Alkaline lysis (Ausubel *et al.*, 1992) was used for large scale preparation of *E. coli* plasmid DNA with omission of the lysozyme treatment and substitution of absolute ethanol for isopropanol during precipitation of the plasmid DNA.

2.9.9 DNA purification from a TAE agarose gel

Following electrophoresis in a TAE agarose gel, DNA purification was performed using a Qiaex II gel extraction kit according to the manufacturer's instructions.

2.9.10 Genomic DNA purification from *C. albicans*

With minor modifications, the method of McEachern and Hicks (1993) was employed for small-scale purification of *C. albicans* genomic DNA. Small cultures (5 mL) were grown to an OD₆₀₀ of 2 - 4. The cells were harvested by centrifugation at 5,000 g for 5 min and the pellet was resuspended into 0.5 mL of 1 M sorbitol, 0.1 M EDTA, pH 7.5 and 14 mM β-mercaptoethanol. Zymolyase 100T (100 μL, 1 mg/mL) was added to the suspension and incubated at 37 °C with occasional mixing for 1 h. The spheroplasts were collected by centrifugation at 3 000 g for 5 min then resuspended into 0.5 mL of 50 mM EDTA, pH 8.5, containing 0.2% SDS. Addition of Proteinase K (25 μL, 2 mg/mL) preceded incubation at 37 °C for 2 h, then heat-treatment of the preparation at 65 °C for 15 min. Following addition of potassium acetate (50 μL, 3 M) the preparation was incubated on ice for 1 h and precipitated with isopropanol. The genomic DNA was resuspended in 100 μL TE pH 8.0, extracted with phenol, precipitated with ethanol and resuspended in 30 μL of TE pH 8.0. This method was also used for purification of genomic DNA from *P. pastoris*.

2.9.11 PCR

Each PCR reaction (total volume 20 μL in a 200 μL thin-walled PCR tube) contained 0.4 μL 1.25 mM dNTP's, 20 pmol of each primer and MgCl₂ to a final concentration of 1.5 mM unless otherwise stated. To this mixture, template DNA and freshly-sterilised MQ water was added, with either 0.3 μL Taq polymerase or 0.15 μL Expand™.

2.9.12 Southern blotting and hybridisation analyses

Southern blotting and hybridisation analyses were carried out using nylon filters (Ausubel *et al.*, 1992). Prehybridisation and hybridisation was performed with Church and Gilbert buffer. Stringency washes for filters probed with oligonucleotides were performed as described by Miyada and Wallace (1987) otherwise as outlined by Ausubel *et al.* (1992).

2.9.13 Radiolabelling DNA

DNA was labelled using ^{32}P dCTP isotope (Amersham) using Ready-To-Go DNA labelling beads, according to the manufacturer's instructions. Radiolabelled DNA was then purified away from unincorporated isotope with Probe-Quant G50 micro spin columns as outlined by the manufacturer (Pharmacia Biotech). Oligonucleotide probes were labelled according to the procedure of Sambrook *et al.* (1989).

2.9.14 Total RNA extraction from *C. albicans*

The method outlined by Ausubel *et al.* (1992) with modifications based on the method of Hube *et al.* (1994) was used for total RNA extraction. Approximately 2×10^8 cells were harvested by centrifugation (5 000 g, 5 min) then resuspended in 1 mL of RNA extraction buffer (0.1 M Tris-HCl, 0.1 M LiCl, 0.01 M DTT) prior to centrifugation at 13 000 rpm for 30 s. The pellet was resuspended into 340 μL of RNA extraction buffer and mixed with 220 μL of glass beads (pre-washed in nitric acid, rinsed in Milli Q water and dried in a baking oven), 20 μL of 10% SDS and 230 μL of phenol/chloroform/isoamyl alcohol 25:24:1 equilibrated with TE pH 8.0. This suspension was immediately mixed vigorously on a vortex mixer for 2 min followed by centrifugation for 1 min at 13 000 rpm. The top aqueous layer was extracted with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) and again mixed on a vortex mixer, centrifuged at 13 000 rpm for 1 min and the top aqueous phase was removed to a clean microcentrifuge tube. The RNA was precipitated with 3 volumes of absolute ethanol then resuspended in 50 μL of TE pH 8.0.

2.9.15 Agarose gel electrophoresis of RNA

RNA was electrophoresed in formaldehyde-containing agarose gels as outlined by Sambrook *et al.*, (1989).

2.9.16 Northern blotting and hybridisation analyses

The formaldehyde agarose gels were blotted to nylon membranes and hybridised as described by Sambrook *et al.* (1989). The filters probed with oligonucleotides were washed as outlined by Miyada and Wallace (1987), otherwise the filters were washed according to Sambrook *et al.* (1989).

2.9.17 Oligonucleotide synthesis

Oligonucleotide synthesis was performed using cyanoethylphosphoramidite chemistry and standard protocols on an Applied Biosystems synthesiser (model 380B).

2.9.18 Dideoxy chain termination sequencing

DNA sequencing was carried out on an Applied Biosystems 373A automated DNA sequencer. DNA was prepared for sequencing as recommended by ABI.

2.10 Yeast and protein methods

2.10.1 Preparation and transformation of yeast spheroplasts

The method of Burgers and Percival (1987) was utilised for preparation and transformation of *Pichia pastoris* spheroplasts. A 50 mL YPD culture was grown at 30 °C (OD₆₀₀ 0.7-1.3), harvested by centrifugation (5 000 rpm for 5 min), washed in 20 mL of 1 M sorbitol, centrifuged, resuspended in 20 mL SCE (1 M Sorbitol, 10 mM CaCl₂, 10 mM Tris-HCl pH 7.5) to which 43 µL β-mercaptoethanol and 33 µL Zymolyase 20T (9 mg/mL) had been added, and incubated at 30 °C for 20 min with occasional mixing. Spheroplast formation was monitored spectrophotometrically at A₈₀₀. Yields of 70-80% spheroplasts were achieved. The spheroplasts were centrifuged at 2,000 rpm for 5 min,

washed in 10 mL filter sterilised STC (1 M Sorbitol, 6.5 mM CaCl₂, 0.25% yeast extract, 0.5% bactopectone) then resuspended in 2 mL STC. Aliquots (100 µl) of the spheroplasts were mixed with linearised plasmid DNA (1-5 µg), incubated at room temperature for 10 min, mixed with 1 mL PEG (10 mM Tris-HCl pH 7.5, 10 mM CaCl₂, 20% w/v PEG 6000/8000) and then incubated for a further 10 min at room temperature before centrifugation at 1 500 g for 10 s. The pellet was resuspended in 150 µL of SOS (1M Sorbitol, 6.5 mM CaCl₂, 0.25% yeast extract, 0.5% bactopectone, filter sterilised) and incubated at 30 °C for 30-60 min. The transformants were plated on minimal agar containing 1 M sorbitol and colonies were counted after 4-5 days.

2.10.2 Screening yeast transformants

Transformants were picked and streaked onto YPD agar plates for overnight growth. Colonies were then restreaked onto selection medium. *P. pastoris* transformants were plated onto minimal medium, and YPD plates spread with 0.25, 0.75, 1.0, 1.5 and 2.0 mg/mL G418 (geneticin). G418 was made up to 10 mg/mL in freshly sterilised MQ water and stored at 4 °C.

2.10.3 *P. pastoris* growth and induction

YPG cultures (50 mL in a 250 mL flask) were inoculated with a single transformant colony, and grown overnight at 30 °C, shaking at 250 rpm. Growth of the cultures was monitored at A₆₀₀, until attenuation reached approximately 20/mL. Cultures were then centrifuged at 3,000 g for 20 min, and the supernatant was discarded. Pellets were resuspended into 50 mL of minimal methanol medium and incubated at 30 °C with shaking at 250 rpm. Further methanol was added every 24 h. For time-course analyses, samples were taken at time zero, then every 24 h. All samples were centrifuged at 3,000 g for 30 min then the supernatants were stored at -20 °C prior to analysis. Larger-scale cultures were performed as described here, but with modifications noted in the text. Induction was monitored by SDS-PAGE and enzyme assays.

2.10.4 TCA precipitation of protein samples

Protein samples (up to 800 μL) were mixed with 100 μL of trichloroacetic acid (72% w/v) and incubated on ice for 30 min, centrifuged at 13 200 rpm for 10 min, washed with 70% ethanol, air-dried and resuspended in 10 μL loading buffer. Dilute protein samples were mixed with a carrier protein (10 μg BSA) prior to precipitation.

2.10.5 SDS-PAGE analysis

SDS-PAGE analyses were performed according to the method of Laemmli and Favre (1973) using 10% separating and 5% stacking gels. Gels were either reductive silver stained (Nesterenko *et al.*, 1994) or Coomassie Brilliant Blue stained (50% v/v methanol, 7.5% v/v acetic acid, 0.025% w/v Coomassie Brilliant Blue R-250) and destained in acetic acid:methanol (7.5%:20% v/v). Molecular weight standards were: β -galactosidase (116 kDa), phosphorylase b (97 kDa), BSA (66 kDa), ovalbumin (45 kDa) and carbonic anhydrase (29 kDa).

2.10.6 Native PAGE analysis

The conditions were identical to those for SDS-PAGE except the SDS was omitted from the separating and stacking gels and both the SDS and β -mercaptoethanol were omitted from the lysis buffer.

2.10.7 Western blotting of proteins

Polyacrylamide gels were electroblotted to Hybond C+ by overlaying the gels with membrane then sandwiching between filter paper (all presoaked in transfer buffer: 10 mM NaHCO_3 , 3 mM Na_2CO_3 , 20% methanol). The electroransfer was performed in transfer buffer at 100V for 1h in a Biorad mini-trans blot electrophoretic cell then checked by staining in Ponceau S (0.3% w/v in 0.1% v/v acetic acid) for 5 min and

destaining in MQ H₂O. The membranes were then incubated for 1-2 h at room temperature in 25 mL block solution with gentle shaking (1% w/v non-fat milk powder in 1x TBS pH 8.0: 25 mM Tris-HCl pH 8.0, 14.5 mM NaCl), then incubated for a further 1.5 h in fresh block solution containing the primary Sap-specific antiserum at a final dilution of 1:5,000 unless otherwise stated (Ross *et al.*, 1990). The blots were washed three times in block solution (5 min each) then goat anti rabbit antiserum-phosphatase conjugate was added (final dilution 1:10,000) to fresh block solution and incubated for 1.5 h. Blots were washed four times in block solution, once with 1x TBS/10 mM MgCl₂ pH 8.8, then incubated in the dark with developing solution (20 mL 1x TBS/10 mM MgCl₂ pH 8.8, 132 µL 50 mg/mL NBT in 70% DMF, 68 µL 50 mg/mL BCIP in DMF). When the bands were visible (10-60 min), the blots were rinsed well with MQ H₂O and left to dry in the dark.

2.10.8 N-terminal amino acid sequence analysis of proteins

All proteins to be sequenced were electrophoresed on SDS-PAGE gels prior to electrotransfer to Problott (PVDF) sequencing membrane. Membranes were prewet in 100% methanol (5 min), washed in MQ (15 min) then washed 4x in CAPS transfer buffer (10 mM CAPS pH 10.0, 20% methanol) and PAGE gels were equilibrated with CAPS transfer buffer for 30 min before transfer to the membrane. Transfer was performed at 100V for 1h using CAPS buffer in a Biorad mini-trans blot electrophoretic cell. Membranes were stained with Coomassie Brilliant blue for 10 s, destained in 50% methanol and dried overnight. The stained protein bands were directly applied to the sequencer. Sequence analyses was performed on an Applied Biosystems (ABI) 470A/170A on line gas-phase sequencer using standard ABI protocols.

2.10.9 Sap enzyme assays

Unless otherwise stated the enzyme activity is given in units for the globin assay (2.10.9.3).

2.10.9.1 Azocollagen assay

Each assay contained 5 mg azocollagen (Sigma), 800 μL 50 mM citrate buffer pH 3.2, and enzyme sample made up to 200 μL . Assays were incubated overnight with shaking, then centrifuged at 14,000 g for 10 min. 500 μL was removed to a fresh microcentrifuge tube and 1 mL ice-cold 50 mM NaOH was added. The absorbance was measured at 520 nm, and results were categorised as no activity, + (light pink supernatant), ++ (medium pink supernatant) and +++ (dark pink supernatant).

2.10.9.2 BSA assay

The assay was based on the method of MacDonald and Odds (1980) and comprised 0.6 mL of 1% BSA in 50 mM sodium citrate, pH 3.2 and up to 0.2 mL of enzyme preparation. Following a 30 min incubation of the reaction at 37 $^{\circ}\text{C}$, 0.2 mL of 2 M perchloric acid was added and the tubes were placed on ice for 10 min then centrifuged at full speed for 5 min. The absorbance of the supernatant was read at 280 nm using a zero time control as the reference (perchloric acid added to reaction prior to incubation). One unit of enzyme catalysed a ΔA_{280} of 1.0 per min.

2.10.9.3 Globin assay

A modification of the procedure given by Jones (1991) was used. The reaction mixture contained 150 μL of 2.0% solubilised globin, pH 3.2, 130 μL 200 mM glycine, pH 3.2 and 20 μL enzyme sample. After 30 min incubation at 37 $^{\circ}\text{C}$, samples were placed on ice and 150 μL perchloric acid added. After centrifugation at 14,000 g for 5 min, a 75 μL sample of supernatant was removed and added to a microcentrifuge tube

containing 75 μL 0.5 M NaOH. Tyrosine-containing peptides in the neutralised sample were determined with Folin's reagent according to Lowry *et al.*, 1951. To each 150 μL sample, 750 μL of reagent consisting of 2.0% (w/v) NaHCO_3 in 1.0 M NaOH, 1.0% (w/v) CuSO_4 and 2.0% (w/v) sodium tartrate was added and incubated at room temperature for 10 min before the addition of 75 μL of 50% diluted Folin's reagent. After 30 min incubation at room temperature the absorbance was determined at 750 nm. One enzyme unit for this system corresponds to the production of 1.0 μg of tyrosine equivalents min^{-1} .

Solubilised globin: 6 volumes of acetone were added to 2% haemoglobin solution, to precipitate the globin. The acetone-soluble haem was decanted off and discarded. The precipitate was then spun at 9,000 rpm for 20 min and the supernatant discarded before resuspending the precipitate in MQ water. The solubilised globin was then dialysed against 3 changes of MQ water, and the final solution was adjusted to pH 3.2 (with HCl) and diluted to 2% w/v concentration.

2.10.9.4 Fluorocasein assay

Assays were performed using the Enz-Check™ Protease Assay Kit (catalogue E-6638, Molecular Probes) according to the manufacturer's instructions. Serial dilutions of enzyme samples (20 μL) were pre-mixed with 100 μL of buffer (either 50 mM citrate, pH 3.2 or 50 mM citrate-phosphate pH 4.5) in microtitre plates. Substrate (1 μL) was added immediately prior to measurement of the fluorescence in the plate-reader, and measurements were taken every minute for 5 min. Data was analysed using a regression program written and kindly provided by Dr. W. Laing.

2.10.10 Substrate specificity of Sap4 and Sap5

This protocol was performed according to the method of Hermann (1996). For Saps 4 and 5, time-course incubations of Sap with glucagon were carried out at two different enzyme: substrate ratios. The reaction mixes consisted of 48 μL of 0.005% TFA,

155 μL citrate-phosphate pH 4.5, 12 μL of 5 mg/mL glucagon and either 0.0164 units (low enzyme:substrate) or 0.328 units (high enzyme:substrate) of Sap enzyme. Each mixture was incubated at 37 °C and 50 μL samples were taken at times as indicated. For the low enzyme:substrate ratio, samples were taken at 2 min, 15 min and 1 h and for the high enzyme:substrate ratio samples were taken at 2 min, 30 min and 3 h. Each sample was frozen in liquid air prior to peptide fragment separation by reversed-phase h.p.l.c. (Gilson 811c), using a linear gradient of 0.1% TFA as mobile phase to 50% acetonitrile/0.08% TFA over 45 min, at a flow rate of 0.6 mL/min. A zero-time control was provided by carrying out reversed-phase h.p.l.c. on a 50 μL sample of the reaction mix taken before the enzyme was added. Identity of purified fragments was analysed by MALDI-TOF mass spectrometry at the Protein Microchemistry Facility, Department of Biochemistry, University of Otago. Analyses were undertaken using narrow bore h.p.l.c. of PTC derivatives.

2.10.11 Column chromatography

All buffers were prepared with fresh MQ H_2O , filtered through 0.22 μm filter membranes and degassed for 1 h immediately prior to use. Chromatographic profiles were recorded on Maclab (version 3.5) and with a chart recorder. Absorbance of the eluate was measured online at 280 nm.

2.10.12 Carbohydrate cleavage

2.10.12.1 β -elimination reaction

The conditions were based on the method described by Piller and Piller (1993). Glycoprotein samples (125 μg in 250 μL) were mixed rapidly with an equal volume of freshly prepared 2 M NaBH_4 in 0.1 M NaOH . Reactions were incubated at 45 °C for 16 h then cooled to room temperature. Five volumes of 0.25 M acetic acid in methanol was added dropwise to destroy the excess NaBH_4 , then evaporated under a stream of

nitrogen. This was repeated twice with acidified methanol and twice with methanol alone.

2.10.12.2 Acid hydrolysis

Glycoprotein samples (150-250 μg) were made up to 400 μL with MQ water. Either HCl or TFA was mixed to a final concentration of 6 M or 2 M respectively. Then samples were incubated at 100 $^{\circ}\text{C}$ for 4 h. Following this, samples were dried under vacuum and resuspended into 100 μL MQ water. Control samples of 1 mg/mL mannose and N-acetyl glucosamine (total volume 400 μL) were treated in the same manner.

2.10.13 DIG Glycan/Protein analysis

2.10.13.1 Labelling

A DIG Glycan/Protein Labelling Kit (Boehringer Mannheim, catalogue number 1500783) was used to label the sugar and protein fractions of glycoproteins. The kit was used according to the manufacturer's instructions. SDS-PAGE gels were run and blotted to Hybond C as previously described. Filters were washed in 50 mL of PBS, pH 6.5 (0.05 M potassium phosphate buffer pH 6.5, 0.15 M NaCl) then incubated in 10 mL 0.1 M sodium acetate, pH 5.5 containing 0.01 M sodium metaperiodate for 20 min. Filters were then washed three times with 50 mL PBS, 5 min each. The immobilised samples were then labelled with 2 μL of DIG hydrazide in 10 mL of 0.1 M sodium acetate pH 5.5 for 1 h at room temperature.

2.10.13.2 Detection

Filters were washed three times (5 min each) in 0.05 M potassium phosphate, pH 8.5 then incubated in a mixture of 5 μL of FLUOSTM in 10 mL 0.05 M potassium phosphate pH 8.5 with 0.01% Nonidet P40 for 1 h. The filters were then washed three times in 50 mL TBS pH 7.5 (0.1 M Tris-HCl, 0.15 M NaCl), 5 min each, then incubated for 30

min in the blocking reagent provided by the manufacturer. The filters were again washed three times in 50 mL TBS (5 min each) then incubated for 1 h in 10 mL TBS to which 20 μ L of each of anti-digoxigenin-POD and anti-fluorescein-AP had been added. Filters were washed in 50 mL TBS, three times for 5 min then incubated with shaking in the following staining reagent: 10 mL Tris buffer pH 8.0 (0.1 M Tris-HCl pH 8.0, 0.05 M MgCl₂, 0.1 M NaCl) containing 50 μ L INT and 50 μ L X-phosphate solution, 50 μ L TETON and 6 μ L H₂O₂. Colour development occurred within 20 min resulting in reddish-brown for proteins and blue for glycans, and the reaction was stopped by rinsing with MQ water.

2.11 Antibody Production

All animal handling was performed by staff of the Massey University Small Animal Production Unit (SAPU). New Zealand white rabbits were used for antibody production. The primary challenge was 200 μ g protein in a 1 mL of emulsion of 50% Freund's complete adjuvant, and subsequent boosters (100 μ g) were given in an emulsion of 50% Freund's incomplete adjuvant. Boosters were given at four week intervals after the primary challenge and ear bleeds were taken one week after each booster to check the antibody titre. Animals were sacrificed by heart puncture and blood samples were coagulated by incubation for 2 h at room temperature, then at 4 °C overnight. Serum was then obtained by centrifuging the sample at 3,000 g for 20 min. Ammonium sulphate precipitation was used to obtain the antibodies as described by Harlow and Lane (1988).

2.12 Crystallisation trials

Crystal trials were performed using 24-well sterile tissue culture plates (Falcon brand) and siliconised glass coverslips. Coverslips were treated with 2% dimethyldichlorosilane in chloroform overnight in a fume hood. The siliconising solution was then drained off and stored for future use. The coverslips were rinsed thoroughly with ethanol, then MQ water, then oven-dried. Prior to setting up hanging

drops, the coverslips were polished with lint-free tissue, and traces of dust were removed by brushing with filtered air. The hanging drop vapour diffusion method was used for the trials, and 1+1 drops were set up. These drops had a total volume of 2 μL , with 1 μL precipitant from the well and 1 μL of concentrated protein solution. Protein solutions were concentrated to at least 5 mg/mL and buffer-exchanged into MQ water immediately prior to setting up the screens. Inhibitors were added to a 50 x molar excess, and incubated with the sample before buffer-exchanging. The wells each contained 600 μL precipitant and were sealed with either vacuum grease or vaseline. Trays were stored at either 17 °C (Sap1 and Sap4) or room temperature (Sap5) and observed at intervals. Crystal screen solutions were kindly provided by Dr. S. Cutfield and Dr. G. Norris.

2.13 Microscopy

All microscopy work was performed with Dr. A Rowland and E. Nickless at Massey University, using a Zeiss microscope, (model Axiophot). Photographs were taken using Kodak Techpan technical film using a Normaski (interference contrast) optical system.

2.14 Computing programs

- The GCG package (Genetics Computing Group), Wisconsin was used for all DNA sequence analysis and alignments.
- DNA and protein gel analyses was performed using an Alphaimager 2000 and Alpha Ease TM v2.3 software, from Alpha Innotech Corporation.
- The Swiss Pdb Viewer (SpdbV: <http://www.expasy.ch/spdbv/mainpage.htm>) was used to produce homology models. The resulting Pdb files were manipulated in SPdbV and figures were produced using QuickDraw 3D.
- Sequence alignments were produced in Jalview© (Michele Clamp, 1998).

3.0 CLONING & VECTOR CONSTRUCTION

3.1 Introduction

The primary aim of this project was the heterologous expression and subsequent characterisation of Sap4, Sap5 and Sap6 from *C. albicans*. Successful over-production of proteins in a foreign host involves several stages. First, the genes must be cloned and a suitable expression system chosen. Then the expression construct is generated and its integrity checked before introduction into the host strain. Finally, an experimental check of the recombinant system will reveal whether or not protein expression has been successful. This chapter describes the first stage of heterologous Sap expression, namely cloning the genes and construction of appropriate expression vectors.

Genomic DNA sequences for *SAPs* 4, 5 and 6 had already been submitted to Genbank (accession numbers L25388, Z30191 and Z30192 respectively) at the outset of this work. The availability of these sequences meant a PCR-based cloning approach could be attempted in the first instance, rather than more laborious methods such as screening a *C. albicans* genomic DNA library. Introns occur relatively infrequently in the *C. albicans* genome and analysis of the *SAP* 4, 5 and 6 Genbank entries confirmed the absence of introns in all three sequences. Therefore genomic DNA was chosen as a

suitable template for PCR. A major pitfall of PCR-based cloning strategies is the propensity of *Taq* polymerase to introduce errors into the DNA sequence being amplified. Fortunately this has now largely been circumvented by the availability of 'proofreading' polymerases (eg. *Pwo*, *Pfu*) and commercial mixtures of *Taq* and proofreading enzymes (eg. Expand™ from Boehringer Mannheim, which was used throughout this work). Nonetheless, careful sequencing of the PCR product is still required to ensure the ORF is free of any nucleotide changes, such as premature stop codons, which might preclude subsequent over-expression. The relative success of an over-expression strategy is also largely dependent on choosing a suitable heterologous system. There is a considerable array of systems available, but in this case, the yeast *Pichia pastoris* was selected as the most appropriate host for Sap 4-6 production. This system will be discussed in greater detail later in this chapter.

3.2 PCR amplification of *SAP4*, *SAP5* and *SAP6*

The PCR-based strategy employed is summarised in Fig. 3.1. The *SAP4*, *SAP5* and *SAP6* DNA sequences are very similar (see Table 1.2) which meant that designing primers specific for each gene was not feasible. As a result, a single upstream primer (homologous to the coding strand of the genomic DNA) was initially designed for use in all PCR reactions. Two downstream primers (homologous to non-coding strand) were designed, based on different sequences in the regions 3' of the termination codons of the *SAP4* and *SAP5* ORFs. These primers were designed to give selective amplification of *SAP4* and *SAP5* open reading frames. However, given the high percentage identity between the *SAP4* and *SAP6* sequences in these regions, it was anticipated that the primers designed to amplify *SAP4* would also amplify *SAP6*. Therefore, an additional *SAP6* PCR primer was designed further upstream of the start codon in an attempt to use a region of lower identity between *SAP4* and *SAP6* to enhance the specificity of the PCR reactions. Despite these attempts to increase the specificity of the PCR reactions, it was anticipated that the products would be a mixture of *SAP4*, *SAP5* and *SAP6* sequences. Details of the PCR primer pairs, expected PCR amplification product sizes and optimal annealing temperatures are summarised in Table 3.1.

3.3 Optimising PCR amplification of *SAPs*

PCR reactions were performed as described in Methods 2.9.11, and each PCR reaction was optimised with respect to magnesium concentration, genomic DNA template concentration and primer annealing temperature (data not shown). PCR reactions were judged as optimised when all bands other than the target band were eliminated, background smear was minimal and the desired product appeared as a tight band on an agarose gel. Optimised *SAP4*, *SAP5* and *SAP6* PCR reactions are shown in Fig. 3.2.

3.4 Restriction endonuclease analysis of PCR products

A restriction endonuclease analysis of the PCR products was performed prior to cloning, to check the relative specificity of the PCR reactions. To select appropriate restriction enzymes, the *SAP* gene sequences were mapped using the GCG DNA analysis package and the enzymes *Eco* RI and *Hind* III were chosen as diagnostic tools. This indicated that *Eco* RI digestion should cleave each *SAP* PCR product once yielding two different sized fragments, and this cleavage pattern should be the same for each reaction, indicating the amplimers are likely to be *SAP* genes. *Hind* III should provide a more informative digest pattern for each PCR product with a unique number of bands for each of *SAP4*, *SAP5* and *SAP6* and therefore confirm the identity of the PCR products. A schematic representation of the *SAP* open reading frames and the relative positions of the restriction endonuclease cleavage sites is shown in Fig. 3.3.

Analysis of the restriction digest patterns of the *SAP* PCR products (Fig. 3.4) involved a comparison of the experimentally determined fragment sizes with the expected band sizes (summarised in Table 3.2). In each case the *Eco* RI digest of the *SAP4*, *SAP5* and

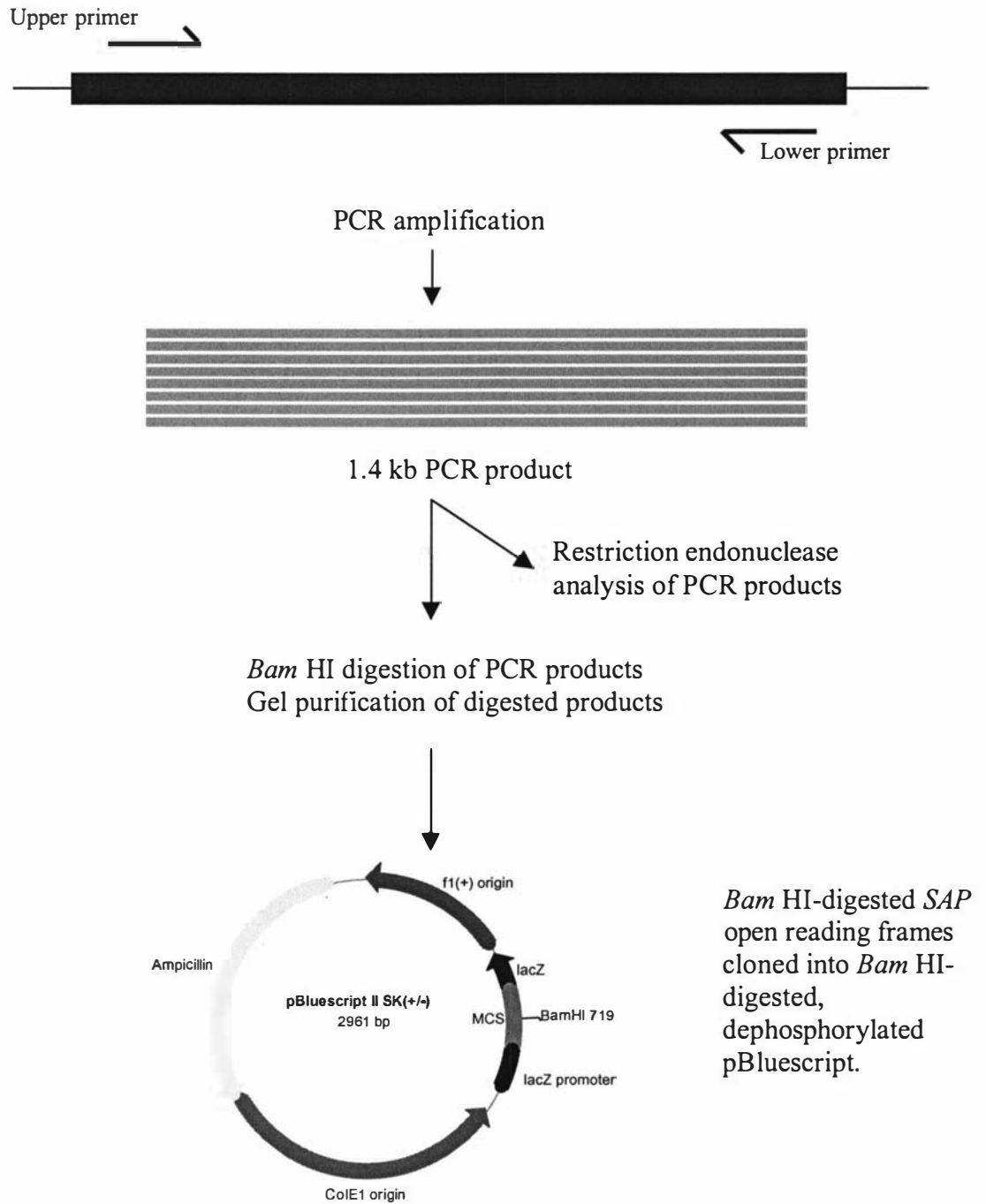


Figure 3.1

Protocol for generating *SAP* clones

Schematic diagram of strategy for PCR-cloning of *SAP* genes. All PCR primers contained *Bam*HI restriction sites to facilitate subcloning and all reactions were carried out using ATCC 10261 genomic DNA as template.

GENE	<i>SAP4</i>	<i>SAP5</i>	<i>SAP6</i>
Upstream Primer	Primer 1: 5'CGG GAT CCA AAG TTC AAC TCA ACC3'	Primer 1: 5'CGG GAT CCA AAG TTC AAC TCA ACC3'	Primer 4: 5'TAC TGC AGA CTA TCA TTC ACT C3'
Upstream Primer Position	Nucleotides 110 – 133	Nucleotides 240 - 263	Nucleotides 167 - 188
Downstream Primer	Primer 3: 5'CGG ATC CCT AAG ATG CAA ACA ACC3'	Primer 2: 5'CGG GAT CCC AAT AGA GAC AAA AA3'	Primer 3: 5'CGG ATC CCT AAG ATG CAA ACA ACC3'
Downstream Primer Position	Nucleotides 1513 - 1491	Nucleotides 1598 - 1574	Nucleotides 1602 - 1580
Annealing Temperature	51.6 °C	52.9 °C	50.8 °C
PCR Product Size	1404 bp	1359 bp	1436 bp

Table 3.1

Summary of PCR conditions for *SAP4*, *SAP5* and *SAP6* amplification reactions

The upper and lower primers contained *Bam* HI (primers 1, 2 & 3) or *Pst* I (primer 4) restriction sites to facilitate subcloning. Nucleotides are numbered as indicated in the database (Miyasaki *et al.*, 1994; Monod *et al.*, 1994). The primer annealing temperatures listed are the optimised temperatures in each case.

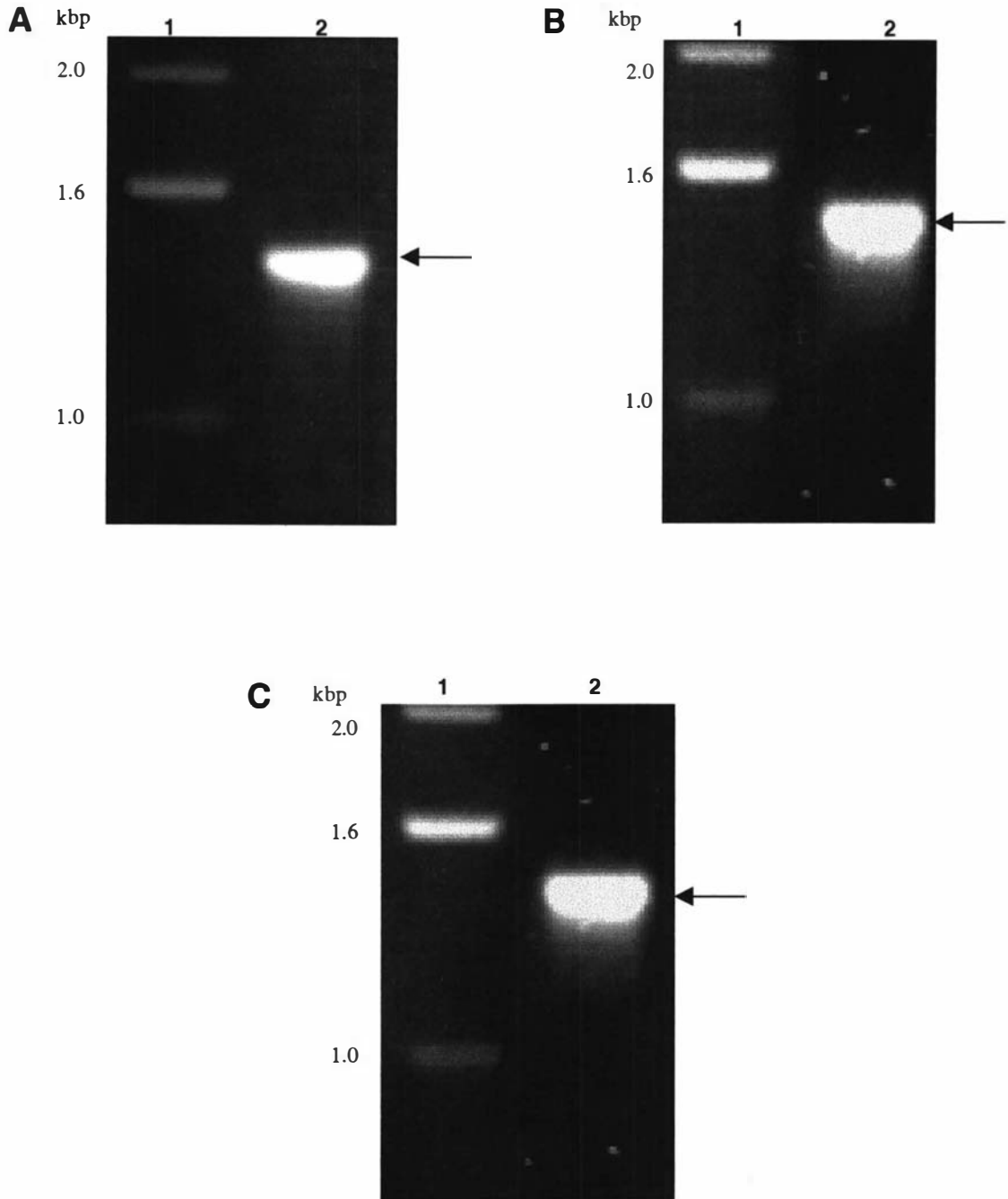


Figure 3.2

Optimised PCR amplification of SAP4, SAP5 and SAP6

Optimised PCR reactions (total volume 20 μ L) were electrophoresed on 0.8% agarose gels and visualised with ethidium bromide. Arrow indicates the optimised PCR product.

- A. Lane 1: 1 kb ladder. Lane 2: *SAP4* PCR reaction (primers 1 & 2).
- B. Lane 1: 1 kb ladder. Lane 2: *SAP5* PCR reaction (primers 1 & 3).
- C. Lane 1: 1 kb ladder. Lane 2: *SAP6* PCR reaction (primers 4 & 2).

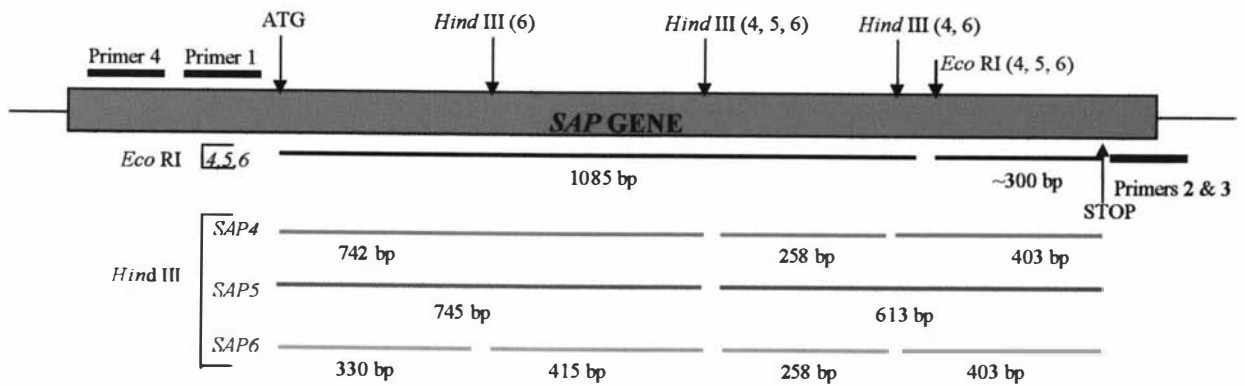


Figure 3.3 Schematic of *SAP 4, 5 & 6* showing relative positions of PCR primers and restriction sites.

The blue bar represents the *SAP 4, 5 and 6* PCR products. The coloured bars represent the upper and lower primers and indicate their relative positions with respect to the start and stop codons. The diagnostic restriction sites are indicated with a black arrow, and are labelled with a 4, 5 or 6 denoting presence of the cleavage site in that *SAP* gene sequence.

- Calculated fragment sizes from *Eco* RI digestion of *SAP 4, 5 and 6* gene sequences.
- Calculated fragment sizes from *Hind* III digestion of the *SAP4* gene
- Calculated fragment sizes from *Hind* III digestion of the *SAP5* gene
- Calculated fragment sizes from *Hind* III digestion of the *SAP 6* gene

SAP6 PCR reactions yielded two bands (Fig. 3.4, lanes 5, 9 & 13), which matched the expected band sizes. This was consistent with the amplification of *SAP4*, *SAP5* and *SAP6* or a mixture of two or more of these targets. The *Hind* III digest revealed anomalous band patterns for each of the PCR products. The *SAP4* primer pair (primers 1 & 2) was expected to amplify a mixture of *SAP4* and *SAP6*. However, the presence of an extra band of approximately 700 bp in the *Hind* III digest was unexpected (Fig. 3.4, lanes 3 & 4). This band could be the 613 bp band expected from a *Hind* III digest of *Sap5* (refer Fig. 3.3), which infers amplification of both *SAP4* and *SAP5* is occurring (refer Fig. 3.3). The *SAP5* amplification reaction contained two additional smaller bands (Fig. 3.4, lanes 6 & 7). This fragment pattern was similar to the *SAP4* *Hind* III digestion (compare observed bands in Table 3.2), despite the apparent differences in migration of the smaller bands. This suggested that either the *SAP5* primer pair (primers 1 & 3) amplified both *SAP4* and *SAP5*, or that one of the *SAP5* alleles possessed an additional *Hind* III restriction site. Some evidence also suggested that non-specific amplification had occurred in the *SAP5* reaction (an undigested second

band visible at approximately 1.4 kb, Fig. 3.4, Lanes 7 & 8), but it was anticipated that this would not interfere with subsequent cloning. The *SAP6* restriction digest results showed a similar pattern to the *SAP4* and *SAP5* analyses which indicated that there was no increase in specificity using the upstream *SAP6* primer (primer 4). First, the expected digest pattern for *SAP6* (Fig. 3.3) would have produced an increase in band intensity around 400 bp, and this is not evident from the gel results (Fig. 3.4, lanes 11 & 12). Second, there should have been no bands greater than 500 bp present if the reaction was specific for *SAP6*. Finally, the absence of a band around 330 bp, suggested that the primers were not amplifying *SAP6*, but instead generated a mixture of *SAP4* and *SAP5*.

This analysis indicated that the *SAP5*-specific primer pair (primers 1 & 3) probably amplified a mixture of *SAP4* and *SAP5*, and that both the *SAP4* and *SAP6* primer pairs (primers 1 & 2, and 4 & 2 respectively) were also producing a mixture of *SAP4* and *SAP5*. Therefore, PCR products from these reactions were used for subsequent cloning, because although the reactions were not absolutely specific for their respective gene sequences, they did appear to contain *SAP4* and *SAP5* DNA sequences. There was no evidence of amplification of *SAP6* using either the original *SAP4* primer pair (homologous to the *SAP6* sequence) or primer 4 (which was further upstream and more specific to the *SAP6* sequence). Other diagnostic experiments also failed to show successful amplification of *SAP6* (data not shown), and attempts to optimise PCR amplification of *SAP6* were discontinued at this stage, so that the work on heterologous expression of Sap4 and Sap5 could be progressed.

3.5 Analysis of *SAP4* and *SAP5* clones

Cloning the *SAP* PCR products involved ligation of *Bam* HI-digested inserts into *Bam* HI-digested, dephosphorylated pBluescript (Fig. 3.1, Methods 2.9.3), transformation of competent *E. coli* cells with the ligation mixture (Methods 2.9.5), and blue-white selection for transformants (Methods 2.9.6).

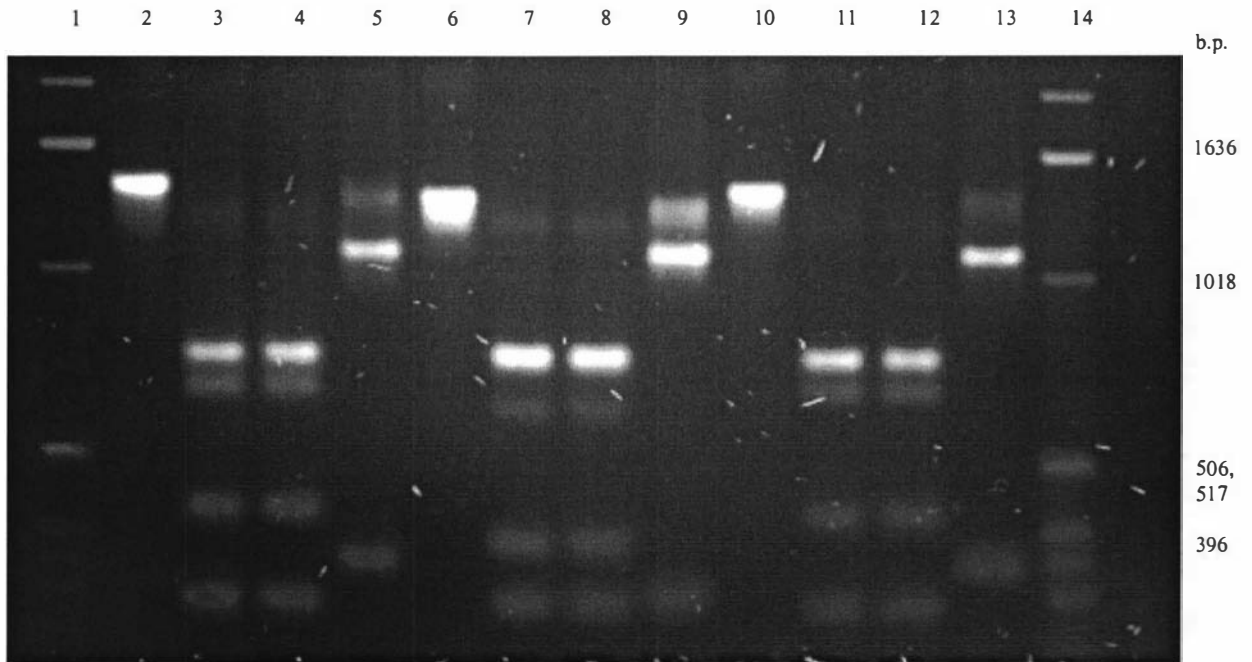


Figure 3.4 *Restriction endonuclease analysis of SAP4, SAP5 and SAP6 PCR products*

PCR reactions (20 μ L, performed as described in Methods 2.9.11) were treated as described below then separated by gel electrophoresis on a 1.5 % agarose gel and visualised with ethidium bromide (Methods 2.9.2).

Lane 1:	1 kb ladder
Lane 2:	<i>SAP4</i> -specific PCR reaction, undigested
Lane 3:	<i>SAP4</i> -specific PCR reaction, digested with 1 μ L <i>Hind</i> III
Lane 4:	<i>SAP4</i> -specific PCR reaction, digested with 2 μ L <i>Hind</i> III to ensure complete digestion of the PCR product
Lane 5:	<i>SAP4</i> -specific PCR reaction, digested with 1 μ L <i>Eco</i> RI
Lane 6:	<i>SAP5</i> -specific PCR reaction, undigested
Lane 7:	<i>SAP5</i> -specific PCR reaction, digested with 1 μ L <i>Hind</i> III
Lane 8:	<i>SAP5</i> -specific PCR reaction, digested with 2 μ L <i>Hind</i> III to ensure complete digestion of the PCR product
Lane 9:	<i>SAP5</i> -specific PCR reaction, digested with 1 μ L <i>Eco</i> RI
Lane 10:	<i>SAP6</i> -specific PCR reaction, undigested
Lane 11:	<i>SAP6</i> -specific PCR reaction, digested with 1 μ L <i>Hind</i> III
Lane 12:	<i>SAP6</i> -specific PCR reaction, digested with 2 μ L <i>Hind</i> III to ensure complete digestion of the PCR product
Lane 13:	<i>SAP6</i> -specific PCR reaction, digested with 1 μ L <i>Eco</i> RI
Lane 14:	1 kb ladder

GENE	DIGEST	EXPECTED BANDS	OBSERVED BANDS
<i>SAP4</i>	<i>Eco</i> RI	1084	- - - - - 1130
		319	- - - - - 340
	<i>Hind</i> III	742	- - - - - 800
		403	- - - - - 420
	258	- - - - - 260	
<i>SAP5</i>	<i>Eco</i> RI	1086	- - - - - 1130
		272	- - - - - 270
	<i>Hind</i> III	745	- - - - - 790
		613	- - - - - 650
		370	
		260	
<i>SAP6</i>	<i>Eco</i> RI	1085	- - - - - 1150
		321	- - - - - 320
	<i>Hind</i> III		810
			720
	415, 403	- - - - - 420	
	330		
	258	- - - - - 260	

Table 3.2 Comparison of the expected and actual band sizes from restriction enzyme analysis of *SAP4*, *SAP5* and *SAP6* PCR products.

A comparison of the calculated sizes of the restriction fragments derived from GCG mapping of the published *SAP* sequences with the experimentally observed bands. The observed fragment sizes were calculated by comparison with the relative migrations of known molecular weight markers. Where an observed band was a similar size to an expected band, the band was considered a match. Additional observed bands, and expected bands which were not present are in bold italicised type.

Initially, considerable difficulty was encountered with the attempts to clone the *SAP4* and *SAP5* sequences into pBluescript. Repeatedly, apparent *E. coli* transformants (white colonies) were screened for the presence of *SAP* inserts by mini-scale DNA preparations (Methods 2.9.7), but these only revealed rearranged pBluescript as indicated by DNA sequencing and restriction digest analysis (data not shown). Eventually the correct transformants were obtained and these clones were identified by restriction analysis with various enzymes: *Bam* HI identified clones which contained a 1400 bp insert,

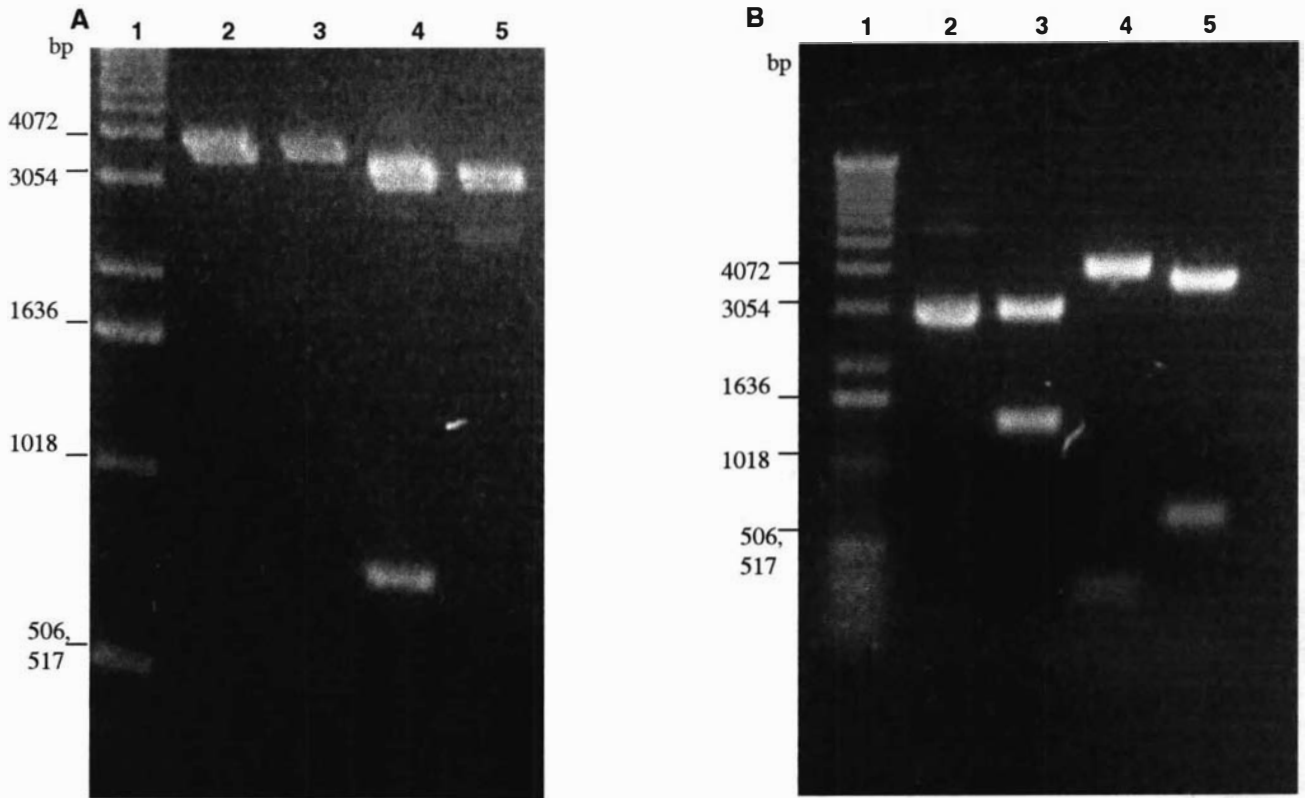


Figure 3.5 Restriction digest analysis of SAP4 and SAP5 sequences cloned into pBluescript

A: Restriction analysis of four *E. coli* transformants.

Lane 1: 1 kb ladder

Lanes 2-5: *Hind* III digestion

B: Restriction digest analysis of a single SAP5 clone, pMS5b.

Lane 1: 1 kb ladder

Lane 2: Uncut DNA

Lane 3: *Bam* HI digest

Lane 4: *Eco* RI digest

Lane 5: *Hind* III digest

* Band visible on original photograph.

Eco RI indicated insert orientation (relative to the T7 primer site of pBluescript) and *Hind* III discriminated between *SAP4* and *SAP5* clones (refer Fig. 3.3). Fig. 3.5A shows *Hind* III digests of four *E. coli* transformants analysed to identify plasmids containing the *SAP4* insert, and to indicate the orientation of the *SAP4* sequence relative to the T7 primer site of the pBluescript vector. These digests were also used to discriminate between clones containing *SAP4* and those containing *SAP5*. Bands at 3750 bp, 450 bp and 250 bp indicated 3'-5' orientation of the insert (Fig. 3.5A, Lanes 2, 3) and 3250 bp, 800 bp and 250 bp bands indicated 5'-3' orientation of *SAP4* (Fig. 3.5A, Lane 4). Vector only with no insert produced a band at 2960 bp (Fig. 3.5A, Lane 5). Fig. 3.5B shows restriction digest analysis of a single *SAP5* clone, pMS5b. In this case, *Bam* HI digestion yields the 1.4 kb *SAP* gene (Fig. 3.5B, Lane 3). *Eco* RI digestion reveals fragment orientation, with bands at 4.2 kb and 300 bp for 3'-5' orientation (Fig. 3.5B, Lane 4), and bands at 3.4 kb and 1.1 kb for 5'-3' orientation. Restriction with *Hind* III identified the insert as *SAP5* (Fig. 3.5B, Lane 5), due to the presence of two restriction sites in the *SAP4* ORF, one site in the *SAP5* ORF and three sites in the *SAP6* ORF (refer Fig. 3.3).

3.6 DNA Sequence Analysis of *SAP4* and *SAP5* clones

Complete double-stranded sequencing of the *SAP4* and *SAP5* pBluescript clones (pMS4b and pMS5b respectively) was required to confirm that these genes had been cloned, to compare the sequences of the cloned genes with the published sequences and to check for errors, deletions or rearrangements introduced by PCR. This involved sequencing in from each end of the construct with T7 and T3 primers, and designing internal primers, which would enable double-stranded sequencing of the internal regions of the genes. Given the high degree of nucleotide sequence identity between *SAP4* and *SAP5*, it was possible to design a single set of primers which were used to sequence both pMS4b and pMS5b. DNA sequencing was carried out using Method 2.9.18, and the resulting DNA sequences were edited and compiled into contiguous sequences using the GCG package (Methods 2.14). These sequences were then aligned with the published sequences to check for errors or strain differences (Fig. 3.6, Fig. 3.7).

```

1   M F L Q N I L S V L A F A L L I D A A P V K R S T G F 27
1   ATGTTCTTACAAAATATCTTGAGTGTCTTGCTTTCGCTTTATTAATTGATGCTGCTCCAGTTAAAAGATCTACAGGTTT 81
-----
28  V T L D F N V K R S L V D P K D P T V E V K R S P L F 54
82  GTTACCTTAGACTTTAATGTCAAAAAGATCCCTTGTGATCCAAAAGATCCAACGTGCGAAGTTAAAAGATCACCTTTATTT 162
-----
55  L D I E P T E I P V D D T G R N D V G K R G P V A V K 81
163  TTAGATATGAGCCCACAGAAATCCCGTCGACGATACTGGTAGAAATGATGTGGGCAAAGAGGACCTGTTGCAGTTAAA 243
-----
82  L D N E I I T Y S A D I T I G S N N Q K L S V I V D T 108
244  TTGACAATGAAATATTACTTATTCTGCTGATATTACGATTGGTTCAAATAACCAAAAACCTTAGCGTTATTGTTGACACT 324
-----
109 G S S D L W V P D S N A V C I P K W P G D R G D F C K 135
325  GGCTCTTCTGACTTGTGGGTTCCAGATTCAAATGCCGTTGTATTCCAAAATGGCCTGGTGACAGAGGAGACTTCTGTAAG 405
-----
136 N N G S Y S P A A S S T S K N L N T P F E I K Y A D G 162
406  AATAACGGTTCCTATTCTCCAGCTGCTTCTAGCACTTCCAAAATTTGAATACTCCTTTTGAAATCAAATATGCCGATGTT 486
-----
163 S V A Q G N L Y Q D T V G I G G V S V R D Q L F A N V 189
487  TCTGTTGCACAAGGTAACCTGTATCAAGATACCCTTGGTATTGGTGGTGTTCCTGTTAGAGATCAATTATTGCTAACGTT 567
-----
190 R S T S A H K G I L G I G F Q S N E A T R T P Y D N L 216
568  AGGCTACTAGTGCTCATAAAGTATTTTAGGTATTGGTTCCAAAGCAACGAAGCCACCAGGACTCCTTACGACAATCTT 648
-----
217 P I T L K K Q G I I S K N A Y S L F L N S P E A S S G 243
649  CCTATTACTTTGAAAAACAAGGCATTATTTCTAAAAATGCTTATTCCTTTTCTTAACTCTCTGAAGCTTCTTCTGGA 729
-----
244 Q I I F G G I D K A K Y S G S L V D L P I T S D R T L 270
730  CAAATTATTTTGGTGGTATTGACAAGGCCAAGTACAGCGGCTCTTTAGTTGATTGGCAATCACTTCTGATAGAACATTA 810
-----
271 S V G L R S V N V M G Q N V N V N A G V L L D S G T T 297
811  AGTGTCCGTTTAAAGATCTGTCAATGTTATGGGACAAAATGTTAATGTCAACGCTGGTGTCTCTTAGATTCTGGTACTACT 891
-----
298 I S Y F T P N I A R S I I Y A L G G Q V H Y D S S G N 324
892  ATCAGTTATTTCACTCCAATATTGCTCGTAGCATTATCTATGCCTTAGGTGGTCAAGTGCATTATGATTCTTCTGGTAAT 972
-----
325 E A Y V A D C K T S G T V D F Q F D R N L K I S V P A 351
973  GAAGCTTATGTTGCTGATTGTA AAACTTCAGGTACCGTTGATTCCAATTCGATAGAAACCTCAAGATTCCGTTCTCTGCT 1053
-----
352 S E F L Y Q L Y Y T N G E P Y P K C E I R V R E S E D 378
1054  TCGGAATTCCTTTACCAATTATATTACACTAATGGTGAACCTTATCCAAAATGTGAAATTCGTGTTTCGTGAAAGTGAAGAT 1134
-----
379 N I L G D N F M R S A Y I V Y D L D D R K I S M A Q V 405
1135  AATATTCTTGGTGACAACCTCATGAGATCAGCTTATATTGTCTACGATTGGATGATAGAAAGATCTCCATGGCTCAAGTT 1215
-----
406 K Y T S Q S N I V G I N * 417
1216 AAATACACTTCCCAGTCTAACATTGTTGGTATTAATTAG 1254
-----

```

Figure 3.6 Nucleotide sequence of *SAP4* from *C. albicans* strain ATCC 10261

Complete nucleotide sequence of the coding strand is shown with the translated protein sequence above. Below is the *SAP4* sequence from WO1. A dash indicates identical nucleotides and silent changes are shown in bold italicised type. The arrow shows the mature N-terminus of Sap4.

```

1   M F L K N I L T V L A F A L L I D A A P V K R S P G F 27
1   ATGTTCTTGAAAAATATCTTGACTGTTCCTGCTTTCGCTTTATTAATTGATGCTGCTCCAGTTAAAAGATCTCCAGGGTTT 81
   -----G-----C-----G-----
28  V T L D F N V K R S L V D P D D P T V E A K R S P L F 54
82  GTCACCTTAGATTTCATGTCAAGAGATCCCTTGTGATCCAAATGGTCCAAGTCTCGAAGCTAAAAGATCGCCTTTATTT 162
   --T--C--T-----T-----G-----T-----
55  L E F A P S E F P V G E T G R D G D V D K R G P V A V 81
163  TTAGAGTTCCTCCTCAGAATTTCCCGTCCGTGAGACTGGTAGAGATGGTGTGTGGACAAAAGAGGACCTGTGCAGTT 243
   -----A-----A-----
82  T L H N E A I T Y T A D I T V G S D N Q K L N V I V D 108
244  ACTTTCACAAATGAAGCTATTACTTATACTGCTGATATTACTGTTGGTTCAGATAACCAAAAACTTAATGTTATTGTTGAC 324
   -----
109  T G S S D L W I P D S N V I C I P K W R G D K G D F C 135
325  ACTGGGCTTCTGACTTGTGGATTCCAGATTCAAACGTTATTTGTATTCCAAAATGGCGTGGTGACAAAGGAGACTTCTGT 405
   -----
136  K S A G S Y S P A S S R T S Q N L N T R F D I K Y G D 162
406  AAGAGTGC CGGTTCTTATTTCCCCAGCATCTCCCGCACTTCCCAAAATTTGAATACCCGTTTTGTATTAATAATGGTGAC 486
   --T-----C-----
163  G S Y A K G K L Y K D T V G I G G V S V R D Q L F A N 189
487  GGTTCCTACGCCAAAGGTAAGTTGTATAAAGATACCGTTGGTATTGGTGGTGTCTCTGTTAGAGATCAATTATTTGCTAAC 567
   -----T-----
190  V W S T S A R K G I L G I G F Q S G E A T E F D Y D N 216
568  GTTGGTCTACTAGTCTCGTAAGGGTATTTTAGGTATTGGTTTTCAAAGCGGCAAGCTACCGAGTTTGATTACGACAAT 648
   -----
217  L P I S L R N Q G I I G K T A Y S L Y L N S P E A S T 243
649  CTTCTATTAGTTTGAGAAATCAAGGTATTATTGGTAAACTGCTTATTCCTCTACCTTAACTCTCTGAAGCTTCTACT 729
   -----C-----C-----
244  G Q I I F G G I D K A K Y S G S L V D L P I T S E K K 270
730  GGGCAAATATTTTTGGTGGTATTGACAAGCCAAGTACAGTGCTCTTTAGTTGATTGCCAATCACTCCGAAAAA 810
   -----A-----
271  L T V G L R S V N V R G R N V D A N T N V L L D S G T 297
811  TTAAGTGTGCTGTTAAGATCTGCAATGTTAGGGGACGAAATGTTGATGCTAACACTAATGTCCTTTTAGATTCTGGTACT 891
   -----
298  T I S Y F T R S I V R N I L Y A I G A Q M K F D S A G 324
892  ACTATCAGTTATTTCACTAGAAGTATTGTTGTAACATTCTCTATGCCATAGGTGCTCAAATGAAGTTTGACTCTGCTGGT 972
   -----
325  N K V Y V A D C K T S G T I D F Q F G N N L K I S V P 351
973  AATAAAGTTTATGTTGCTGATTGTA AAACTTCAGGTACCATTGATTTCCCAATTTGGTAATAACCTCAAGATCTCCGTTCTCT 1053
   -----
352  V S E F L F Q T Y Y T S G K P F P K C E V R I R E S E 378
1054  GTTCCGAATTCCTTTTCCAAACATATTACACTAGTGGTAAACCTTTCCCAAAATGTGAAGTTCGTATTCTGTAAAGTGAA 1134
   -----
379  D N I L G D N F L R S A Y V V Y N L D D K K I S M A P 405
1135  GATAATATTCTGGTGACAACTTTTAAGATCAGCTTATGTTGTCTACAATTTGGACGATAAGAAAATCTCCATGGCTCCA 1215
   -----
406  V K Y T S E S D I V A I N * 418
1216  GTTAAATACACTCCGAGTCTGACATTGTTGCTATTAATTA 1257
   -----

```

Figure 3.7 Nucleotide sequence of *SAP5* from *C. albicans* strain ATCC 10261

Complete nucleotide sequence of the coding strand is shown with the translated protein sequence above. Below is the *SAP5* sequence from Ca74. A dash indicates identical nucleotides, bold italicised type shows silent changes, and changes causing an amino acid substitution are highlighted in white type. The arrow shows the mature N-terminus of Sap5.

Comparison of the *SAP4* (pMS4b) sequence from strain ATCC 10261 with that from strain WO1 (Genbank accession #L25388) revealed four nucleotide base changes in the open reading frame (Fig. 3.6). Two of the nucleotide changes found in pMS4b were A/T→G/C transitions. These are the predominant errors introduced by *Taq* polymerase during PCR (Keohavong & Thilly, 1989). However, all four substitutions lie in the 3rd-base codon position. Together, this suggested that the changes were unlikely to be PCR-introduced, but may represent true sequence variation between strains ATCC 10261 and WO1. In any case, no amino acid changes are introduced by these differences and consequently the *SAP4* clone was considered suitable for subsequent heterologous expression of Sap4 from *P. pastoris*.

An analogous alignment of the *SAP5* (pMS5b) sequence with its counterpart from strain Ca74 (Genbank accession #Z30191) showed significantly more changes at the nucleotide level than the *SAP4* clone. In total eighteen nucleotide substitutions were identified, five of which resulted in changes to the Sap5 amino acid sequence. Nine of these mutations were A/T→G/C transitions, and hence may be *Taq*-mediated PCR errors. However, the rate at which *Taq* introduces errors in PCR (1 per 20,000 bases every round of amplification) is much lower than the frequency of observed changes in pMS5b and this indicated that it was unlikely all of these mutations were PCR-introduced. Also, it is significant that 75% of the mutations in pMS5b occurred in the region encoding the pro-segment. While zymogen pro-regions are important for correct folding and activation of the enzyme (Khan & James, 1998), once cleaved, they have no role in the catalytic activity of the enzyme. This considered, perhaps the clustering of mutations in the pro-region of Sap5 is indicative of fewer constraints on the structure/function relationship of this part of the protein than in the catalytic domains. Interestingly, three of the five amino acid sequence changes affect the pro-sequence, leaving just two amino acid substitutions in the mature active enzyme. An alignment of Sap5 with other members of the *C. albicans* Sap family at the positions of the two amino acid substitutions in the mature protein (Fig. 3.8) reveals these mutations would be unlikely to have a negative impact on the function of the proteinase. First, the position of the mutations is not near the active site in the tertiary structure of the enzyme, and

secondly, the amino acid substitutions in the amplicon from strain ATCC 10261 are even more conservative than in Ca74. At position 231 in the amplicon from strain ATCC 10261, asparagine is replaced by a similar polar, uncharged residue (threonine), whereas in Ca74 the substituted residue has a hydrophobic character (alanine). At position 241 the proline found in Saps 1, 2, 4 and 6 is maintained in Sap5 from ATCC 10261, although the substitution of alanine for proline in Ca74 is a conservative change as both residues are hydrophobic. However, the presence of a bulky, positively charged arginine residue in the same position in Sap3 indicates that even significant substitutions in this region do not impact on enzyme function. Indeed there is significant primary sequence divergence across the family of aspartic proteinases, yet despite this, the overall fold of the protein has been conserved across organisms (Blundell & Johnson, 1993), suggesting that considerable changes to the nucleotide and amino acid sequence can be tolerated without detrimental effects.

	229	234	239
Sap5 (10261)	KTAYS	LYLNS	PEAST
Sap1	KNAYS	LYLNS	PNAAT
Sap2	KNAYS	LYLNS	PDAAT
Sap3	KNAYS	LYLNS	RQATS
Sap4	KNAYS	LFLNS	PEASS
Sap5 (Ca74)	KAYS	LYLNS	AEAST
Sap6	KNAYS	LFLNS	PEASS

Figure 3.8 *Alignment of the sequence containing the amino acid substitutions in Sap5 with other members of the C. albicans Sap family.*

The substitutions in Sap5 (at positions 230 and 239) are highlighted in bold italics. In Sap5 from *C. albicans* Ca74, the asparagine has been replaced by an alanine, whereas in ATCC 10261 there is a threonine residue. At position 241, the conserved proline is maintained in ATCC 10261 but is replaced by an alanine in Ca74. Numbering according to the Sap5 sequence in Fig. 3.7.

Therefore, while the *SAP5* sequence has a considerable number of nucleotide changes, they probably represent sequence divergence rather than introduced errors from the cloning technique. Consequently, the clone pMS5b was considered suitable for subsequent heterologous expression of Sap5 from *P. pastoris*.

3.7 Construction of heterologous expression vectors for Sap4 and Sap5

The host chosen for heterologous expression of Sap4 and Sap5 was the methylotrophic yeast *P. pastoris*. This yeast has several features, which made it the over-expression system of choice. *P. pastoris* possesses an inducible alcohol oxidase (*AOX*) gene, which allows it to utilise methanol as a sole carbon source. The *AOX* promoter, which is used to drive the expression of foreign proteins, is tightly controlled, has a high efficiency of transcription and is inexpensive to induce. *P. pastoris* can also be grown to high cell densities and is suitable for fermentation-scale growth. Also, there are a variety of commercially available vectors for *P. pastoris* and the resulting expression constructs are integrated into the *P. pastoris* genome, producing stable transformants.

To facilitate production of mature, active Sap4 and Sap5, it was important to consider the role of the pro-regions of aspartic proteinase zymogens. *SAP4* and *SAP5* (and other aspartic proteinases) encode a signal sequence, followed by a 75 amino acid pro-sequence, which is important for the correct tertiary structure of the enzyme, possibly by creating a scaffold onto which the protein is folded (Koeslch *et. al.*, 1994). Upon cleavage of the pro-region, aspartic proteinases become activated (for review, see Khan & James, 1998). Due to the relative importance of the pro-region for production of mature Sap enzyme, an expression strategy was designed which utilised the native signal sequence and pro-regions of *SAP4* and *SAP5* in combination with an intracellular *P. pastoris* expression vector (Fig. 3.9). It was anticipated that recognition of the *SAP* signal sequence by the host strain would direct the protein to the secretory pathway, where the pro-enzyme would be correctly folded, and that the pro-sequence would subsequently be cleaved by a *P. pastoris* Kex2 in the Golgi complex.

Production of the Sap4 and Sap5 over-expression constructs described in Fig. 3.9 required subcloning the *SAP4* and *SAP5* sequences into the pPIC3K vector. In pPIC3K, the start

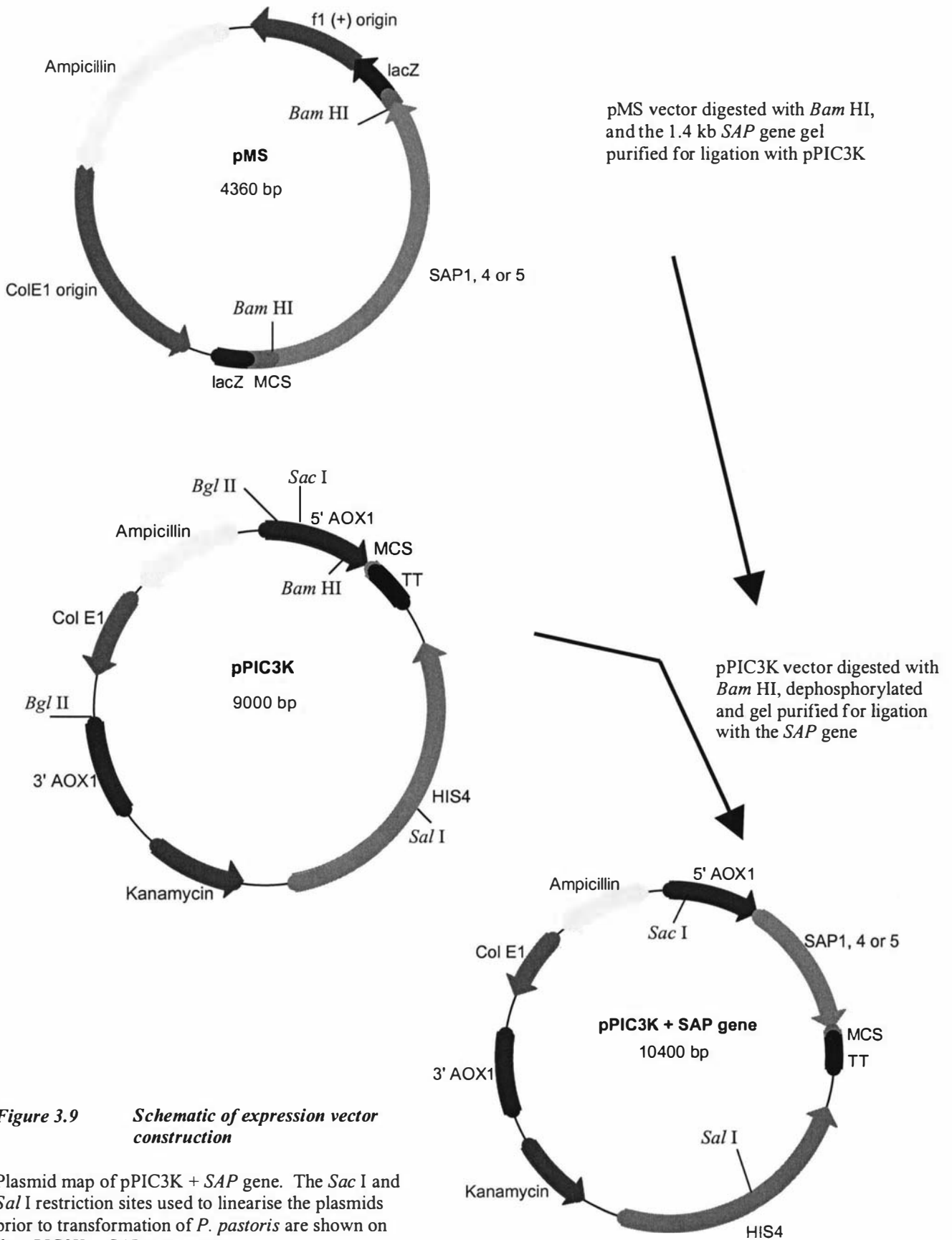


Figure 3.9 Schematic of expression vector construction

Plasmid map of pPIC3K + SAP gene. The *Sac I* and *Sal I* restriction sites used to linearise the plasmids prior to transformation of *P. pastoris* are shown on the pPIC3K + SAP gene map.

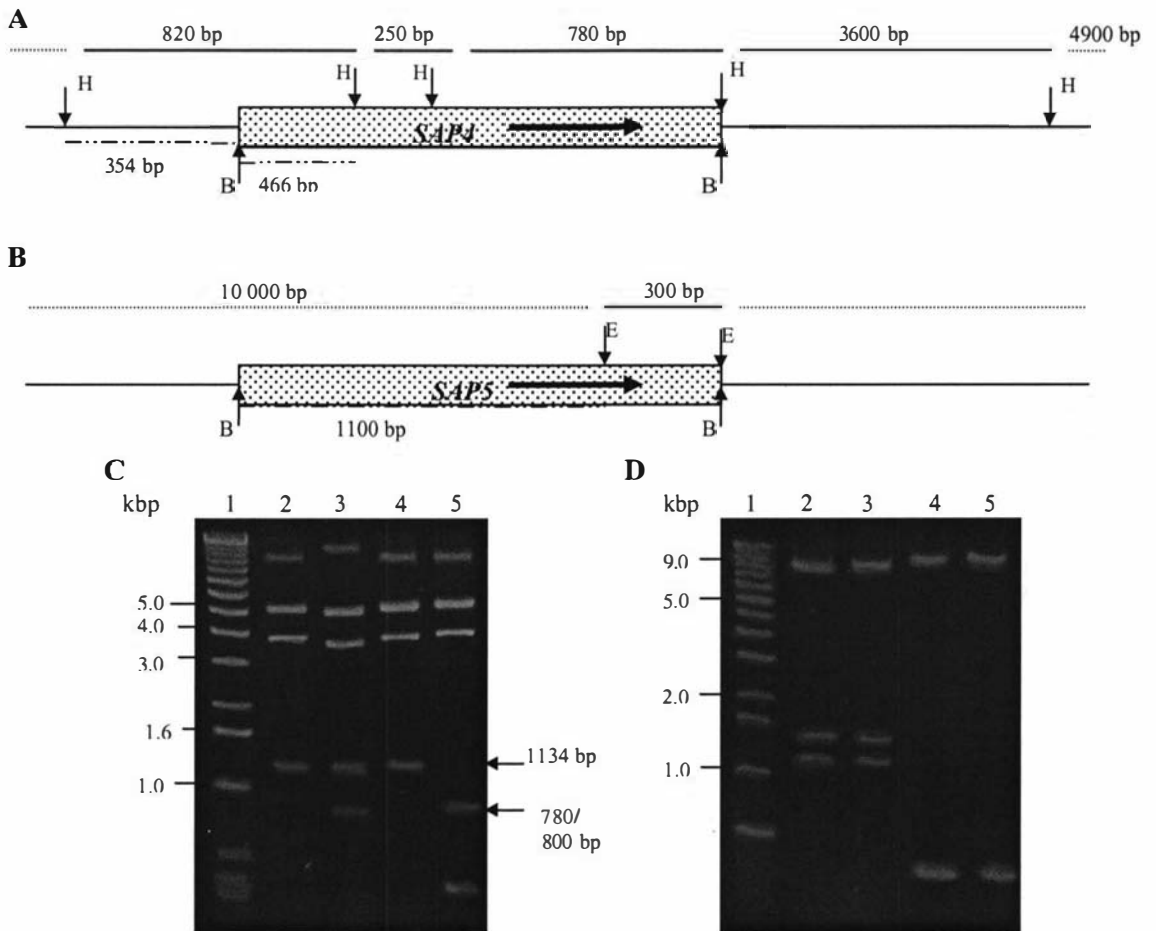


Figure 3.10 Screening pPIC3K clones for the correct orientation of *SAP4* and *SAP5* genes.

A *Hind* III digestion of pPIC3K vector containing the *SAP4* gene in the correct orientation (H represents *Hind* III, B represents *Bam* HI, map not to scale). If the insert is in the reverse orientation, a 466 bp and a 1134 bp band replace the 780 bp and 800 bp bands.

B *Eco* RI digestion of pPIC3K vector containing the *SAP5* gene in the correct orientation (E represents *Eco* RI, B represents *Bam* HI, map not to scale). If the insert is in the reverse orientation, 1100 bp and 9200 bp bands replace the 300 bp and 10 000 bp bands.

C *Hind* III digests of four different pPIC3K/*SAP4* clones, fractionated on a 0.8% agarose gel and visualised by staining with ethidium bromide.

Lane 1: 1 kb ladder.

Lanes 2, 3 & 4: *SAP4* in the reverse orientation in the pPIC3K vector.

Lane 5: *SAP4* in the correct orientation in the pPIC3K vector.

D *Hind* III digests of four different pPIC3K/*SAP5* clones, fractionated on a 0.8% agarose gel and visualised by staining with ethidium bromide.

Lane 1: 1 kb ladder.

Lanes 2 & 3: *SAP5* in the reverse orientation in the pPIC3K vector.

Lanes 4 & 5: *SAP5* in the correct orientation in the pPIC3K vector.

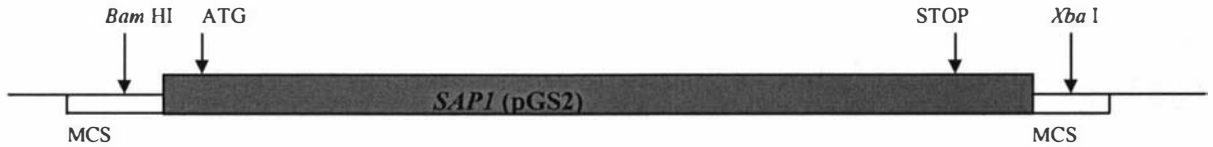
codon is situated downstream of the *Bam* HI cloning site, which meant that the native *C. albicans* start codon would be used for initiation of transcription. This strategy resulted in a small extension (17 nucleotides) between the ATG and the upstream promoter sequences. (see Fig. 3.11, B). A cohesive-end ligation was performed at the *Bam* HI site (see Methods 2.9.3), and the resulting transformants were screened for correct 5'→3' orientation of the insert using restriction endonuclease analysis (Fig. 3.10). Following identification of suitable clones, large-scale DNA preparations were done (Methods 2.9.8) for transformation of the host strain.

3.8 Construction of a Sap1 heterologous expression vector.

Smolenski *et. al.* (1997) expressed *C. albicans* Sap1 in *Saccharomyces cerevisiae*, albeit in modest quantities, where previous attempts at over-production of the Sap proteins by other groups had been unsuccessful (P. Sullivan, personal communication). Therefore in addition to making vectors for recombinant Sap4 and Sap5 production, a vector for heterologous Sap1 expression was also constructed (the plasmid containing the *SAP1* gene, pGS2, was kindly provided by Dr. J Cutfield). Over-expression of Sap1 in *P. pastoris* therefore provided a positive control for the production of the recombinant Sap4 and Sap5.

The strategy used to construct the Sap1 expression vector involved initial digestion of pGS2 with *Xba* I, followed by end-filling with Klenow (2.9.4) to produce a blunt-end fragment. The vector was then digested with *Bam* HI prior to directional ligation with *Bam* HI/*Sna* BI-digested pPIC3K vector (outlined in Fig. 3.11). This produced a construct equivalent to the Sap4 and Sap5 expression vectors, which was purified ready for transformation of the host strain.

A.



B.

VECTOR

GAA GGA **TCC** ACC **ATG**
 CTT CCT ACC **TGG** **TAC**

3 nucleotides

INSERT (SAP5)

GG **ATC** CAA AGT TCA ACT CAA CCA ACA **ATG**
 CC **TAG** GTT TCA AGT TGA GTT GGT TGT **TAC**



CONSTRUCT

GA AGG **ATC** CAA AGT TCA ACT CAA CCA ACA **ATG**
 CT **TCC** **TAG** **GTT** TCA AGT TGA GTT GGT TGT **TAC**

20 nucleotides

Figure 3.11 Schematic of the *SAP1* gene and cloning strategy

A. The plasmid pGS2, containing the *SAP1* gene, was digested with *Xba* I, end-filled (see Methods 2.9.4), then digested with *Bam* HI. Following gel purification of the 1.4 kb fragment, the *SAP1* gene was directionally cloned into *Bam* HI/*Sna* BI double-digested pPIC3K vector (Methods 2.9.3).

B. Cloning *SAP* genes into pPIC3K adds an additional 17 nucleotides upstream of the start codon (shown in bold). The *Bam* HI restriction site is shown in grey. All three Sap expression vectors were constructed in this way.

4.0 TRANSFORMATION & EXPRESSION

4.1 Introduction

Some of the advantages of *P. pastoris* as a host for recombinant protein production were discussed previously (Chapter 3) including the stable integration of the foreign DNA into the genome, rather than existing as an episomal plasmid and the availability of several strains of *P. pastoris* for recombinant protein expression. The *P. pastoris* genome contains two *AOX* genes; *AOX1* accounts for 90-95% of the alcohol oxidase in the cell, while *AOX2* produces only 5-10% of the total. In strain KM71, *AOX1* has been disrupted, resulting in slow growth on methanol (as the sole carbon source), and it has been designated as Mut^S (*Methanol Utilisation Slow*). This eliminates the need to screen transformants for Mut⁺ or Mut^S phenotypes, which is otherwise necessary because Mut⁺ phenotypes continue growing rapidly on methanol while Mut^S strains do not, and this would influence the induction strategy. Strain KM71 was selected for use throughout this work.

This chapter describes the process of vector integration and the subsequent selection of *P. pastoris* transformants. Significant consideration was also given to optimising the induction protocol for maximal Sap expression.

4.2 Transformation of KM71 with heterologous expression constructs

P. pastoris can be transformed as spheroplasts and by electroporation of whole cells. The former increases the probability of multicopy plasmid integration, which is associated with higher levels of protein expression, but is more difficult and time-consuming. In contrast, electroporation is quick and simple. Both methods require a linearised vector, and the position of the restriction site determines the site of integration in the host genome.

KM71 was transformed with pMS4 (*SAP4*), pMS5 (*SAP5*) and pMS1 (*SAP1*) by making spheroplasts, adding the DNA, then regenerating the intact cells by plating on osmotically stabilised medium (Method 2.10.1). pMS4 and pMS5 were digested with *Sal* I to linearise the plasmid, while pMS1 was linearised with *Sac* I because of a *Sal* I restriction site within the open reading frame of the *SAP1* gene. Insertion at the *HIS4* locus occurs when the vector is digested with *Sal* I, and insertion at the *AOX1* locus results with plasmids linearised with *Sac* I (Fig. 3.10).

4.3 Selection of KM71 transformants

Appearance of visible colonies following the regeneration took approximately 2 weeks, at which time the colonies were picked and replated onto selection medium. KM71 contains two selection markers, *HIS4* (*Pichia* wild-type histidinol dehydrogenase gene) which complements the histidine auxotrophy of the host strain, and *kan* (kanamycin resistance gene from Tn903) which confers G418 (Geneticin) resistance in the host strain. Therefore, three different media were used: rich medium (YPD), minimal medium, and YPD containing various concentrations of G418 (geneticin) (Methods 2.10.2). The rich medium enables all the transformants to grow, while the minimal medium is used to screen for conversion to histidine prototrophy (His⁺). Spontaneous

reversion to His⁺ is a relatively frequent occurrence in *P. pastoris*, and up to 60% of the putative transformants can be revertants. Because of this, colonies were also streaked onto YPD + G418, where the G418 selects for the presence of the kanamycin resistance gene (*kan*) in yeast.

The presence of multiple gene copies is associated with increased protein expression due to a gene dosage effect (Vedvick, 1991), and multicopy plasmid insertion is loosely correlated with resistance to increasing concentrations of G418. This means that a simple screen can be performed by plating the transformants onto YPD plates containing increasing concentrations of G418, and the isolates which grow on higher concentrations of G418 should contain multiple inserts. Generally, 0.5 mg/mL G418 resistance corresponds to approximately 1-2 copies of the plasmid, while growth on 4.0 mg/mL G418 represents up to 7-12 gene copies (Invitrogen, 1996). However, G418 resistance does not always correlate with the number of gene copies due to the qualitative nature of the screen.

Fig. 4.1 shows a G418 resistance screen of KM71 transformed with pMS1, performed in conjunction with a growth screen on minimal medium (results not shown). Transformants identified by the growth screens were designated MS5A and MS5B (KM71 containing the pMS5 plasmid, Fig. 4.4), MS4A and MS4B (KM71 containing the pMS4 plasmid, Fig. 4.4), and MS1A, MS1B and MS1C (KM71 containing the pMS1 plasmid, Fig. 4.4). Strains MS5A and MS4A were then analysed further to confirm plasmid integration.

All three strains exhibited G418 resistance at 2.0 mg/mL, which is indicative of multiple copy plasmid insertion. However, screening on G418 is sensitive to the cell density and this can result in false positives (Fig. 4.1, plates D & E). Therefore other methods were required to confirm multiple integration events and establish copy number.

4.4 Confirmation of plasmid integration

A Southern blot was performed to confirm plasmid integration and to determine whether multi-copy insertion had occurred in MS5A and MS4A. Genomic DNA was prepared (Methods 2.9.10) and digested to completion with *Bgl* II, *Bgl* II plus *Sal* I, and *Sac* I (Fig. 4.2). The products of the digests were separated by electrophoresis on a 0.7% agarose gel, then blotted onto nitrocellulose and probed with a ^{32}P labelled fragment of *SAP2*. The fragment sizes expected to hybridise to the probe are listed in Fig. 4.2.

Fig. 4.3 confirms the integration of pMS4 and pMS5 into KM71 creating MS4A and MS5A, respectively. The presence of two bands in the *Sac* I digests of both MS4A and MS5A suggested that multicopy insertion of the plasmid has occurred, as one of the bands in each case is 11.4 kb which would only result if two or more copies of the plasmid had integrated head-to-tail. Both digests contained an additional band at approximately 9 kb, which was consistent with multi-copy integration at the same site. Relative band intensity suggested that MS4A has more copies of the *SAP4* plasmid than MS5A has of the *SAP5* plasmid. However, without further copy number experiments, no firm conclusions can be drawn, other than confirmation of plasmid integration in each strain.

4.5 Preliminary expression of recombinant Sap proteins

Following the identification of recombinant strains, preliminary expression trials were performed to establish whether the Sap proteins were being over-produced in *P. pastoris*. Initial trials were performed in 50 mL of induction medium (Methods 2.10.3), and samples (10 mL) were taken either after 48 h or every 24 h for a time course analysis. Samples were centrifuged to remove cells, then analysed by SDS-PAGE (Methods 2.10.5). The expression trial results are shown in Fig. 4.4, and the putative recombinant Sap proteins are labeled with arrows.

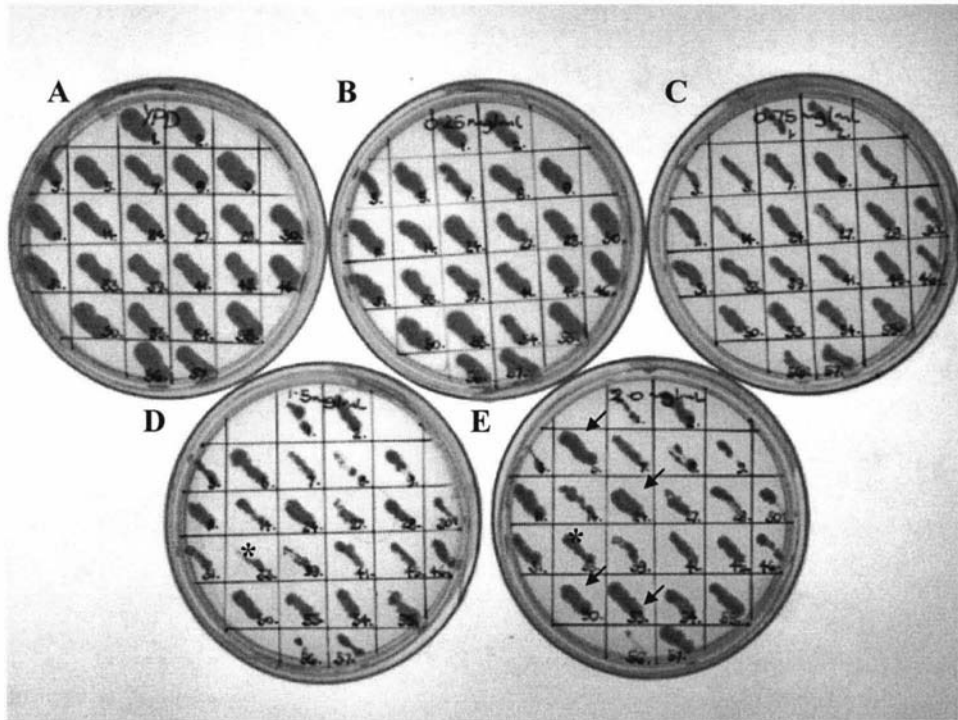


Figure 4.1 *G418 screen of KM71 transformants*

- A: YPD only. B: YPD + 0.25 mg/mL G418
C: YPD + 0.75 mg/mL G418 D: YPD + 1.5 mg/mL G418
E: YPD + 2.0 mg/mL G418

Arrows indicate some of the transformant colonies, asterisk identifies false positive at a G418 concentration of 2 mg/mL.

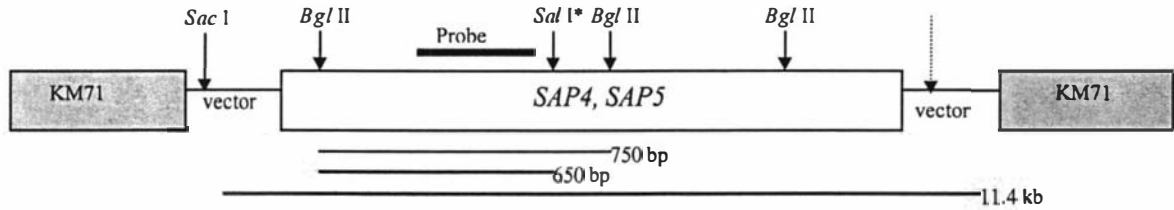


Figure 4.2 Schematic representation of *pMS4* and *pMS5* plasmids integrated into the *KM71* genome

Restriction sites are indicated with an arrow and the region of homology to the probe is shown with a thick bar. *KM71* genomic DNA is represented in dark grey. Dashed arrow indicates insertion point for multiple copies of the plasmid. * Restriction site is only present in *SAP4*. Map is not to scale.

Genomic DNA from *MS4A* and *MS5A* digested and then probed with radiolabelled *SAP2* fragment should give the following fragment sizes:

MS4A	<i>Bgl</i> II	750 bp
	<i>Bgl</i> II + <i>Sal</i> I	750 bp
	<i>Sac</i> I	A single band for single copy integration (unknown size) Two bands for multicopy insertion (one at 11.4 kb, the other band of unknown size).
MS5A	<i>Bgl</i> II	750 bp
	<i>Bgl</i> II + <i>Sal</i> I	650 bp
	<i>Sac</i> I	A single band for single copy integration (unknown size) Two bands for multicopy insertion (one at 11.4 kb, the other an unknown size).

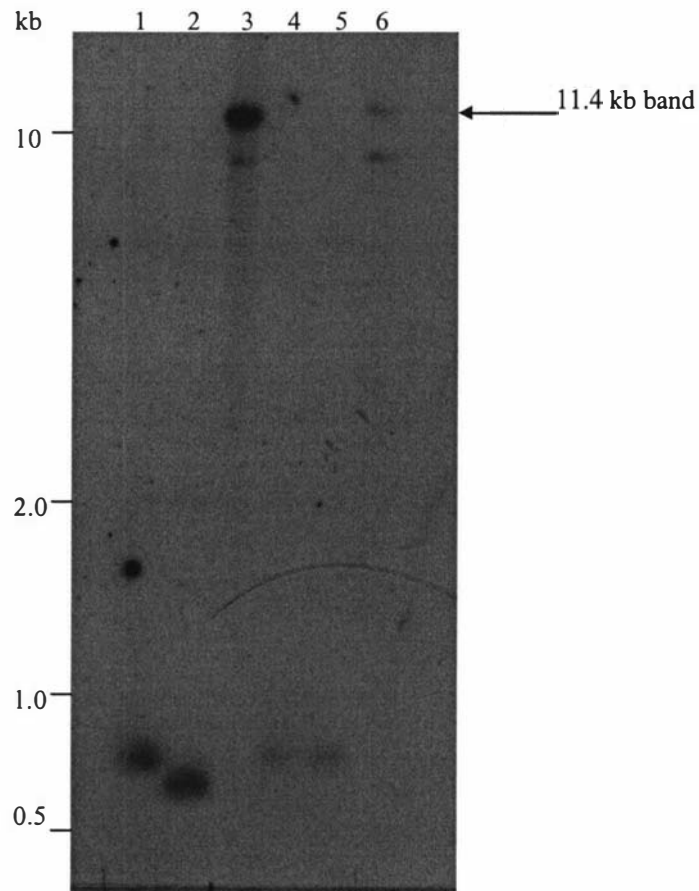


Figure 4.3 Southern blot of *P. pastoris* (KM71) transformants, MS4A and MS5A, containing pMS4 and pMS5

Digested genomic DNA (3 µg per lane) was electrophoresed on a 0.7 % agarose gel at 30 V overnight. DNA was blotted to nitrocellulose and probed at 55 °C with a 500 bp *Pst* I – *Hind* III fragment of *SAP2* (Methods 2.9.12), washed with 2xSSC/0.5% SDS and exposed to film for 2 d.

Lane 1: MS5A digested with *Bgl* II.
 Lane 2: MS5A digested with *Bgl* II/*Sal* I
 Lane 3: MS5A digested with *Sac* I
 Lane 4: MS4A digested with *Bgl* II.
 Lane 2: MS4A digested with *Bgl* II/*Sal* I
 Lane 3: MS4A digested with *Sac* I

Previous studies have shown that Saps 1, 2 and 3 of *C. albicans* migrate anomalously in SDS-PAGE, yielding apparent molecular masses 3-6 kDa larger than the molecular masses deduced from their sequences. For instance, recombinant Sap1 migrates as a 38 kDa protein on SDS-PAGE (Smolenski *et al.*, 1997). It was not known whether Sap4 and Sap5 would also migrate in a similar manner. Interestingly, of the eight putative Sap1 transformants tested, only three expressed recombinant protein of the expected size (Fig. 4.4, A). All eight strains grew on minimal medium and 0.5 mg/mL G418 but only the three transformants which produced recombinant Sap1 (MS1A, MS1B and MS1C) grew on G418 concentrations greater than 1.5 mg/mL.

Both MS4A and MS4B produced three protein bands following the addition of methanol to the culture medium (Fig. 4.4, B), which were absent from the induced control, KM71. The lowest band migrated at a similar size to the native Saps, implying that it might be mature Sap4. This was confirmed by N-terminal sequence analysis (Fig. 4.5B). However, the presence of two higher bands suggested some misprocessing or post-translational modification of the recombinant protein and N-terminal sequence analysis revealed these bands had N-terminal extensions, which accounted for the observed difference in migration (Fig. 4.5B).

A time course analysis of an MS5A induction (Fig. 4.4, C) showed that a protein of the expected molecular mass was produced in large quantities early in the process, with a further increase in the level after 48 h induction. There were also several other protein bands present in the induced samples, and the intensity of these bands increased over time. However, these bands probably arose from cell lysis over time, as they did not immunoreact with the Sap2 antibody preparation (Fig. 4.6C). N-terminal sequence of this product confirmed it was correctly processed Sap5 (Fig. 4.5C).

These results indicated that the *P. pastoris* transformants MS5A, MS4A and MS1A were producing recombinant protein of the expected size in response to methanol induction. Definitive identification of all three recombinant proteins was obtained by N-terminal sequencing (Fig. 4.5) before scale-up and purification of the proteins.

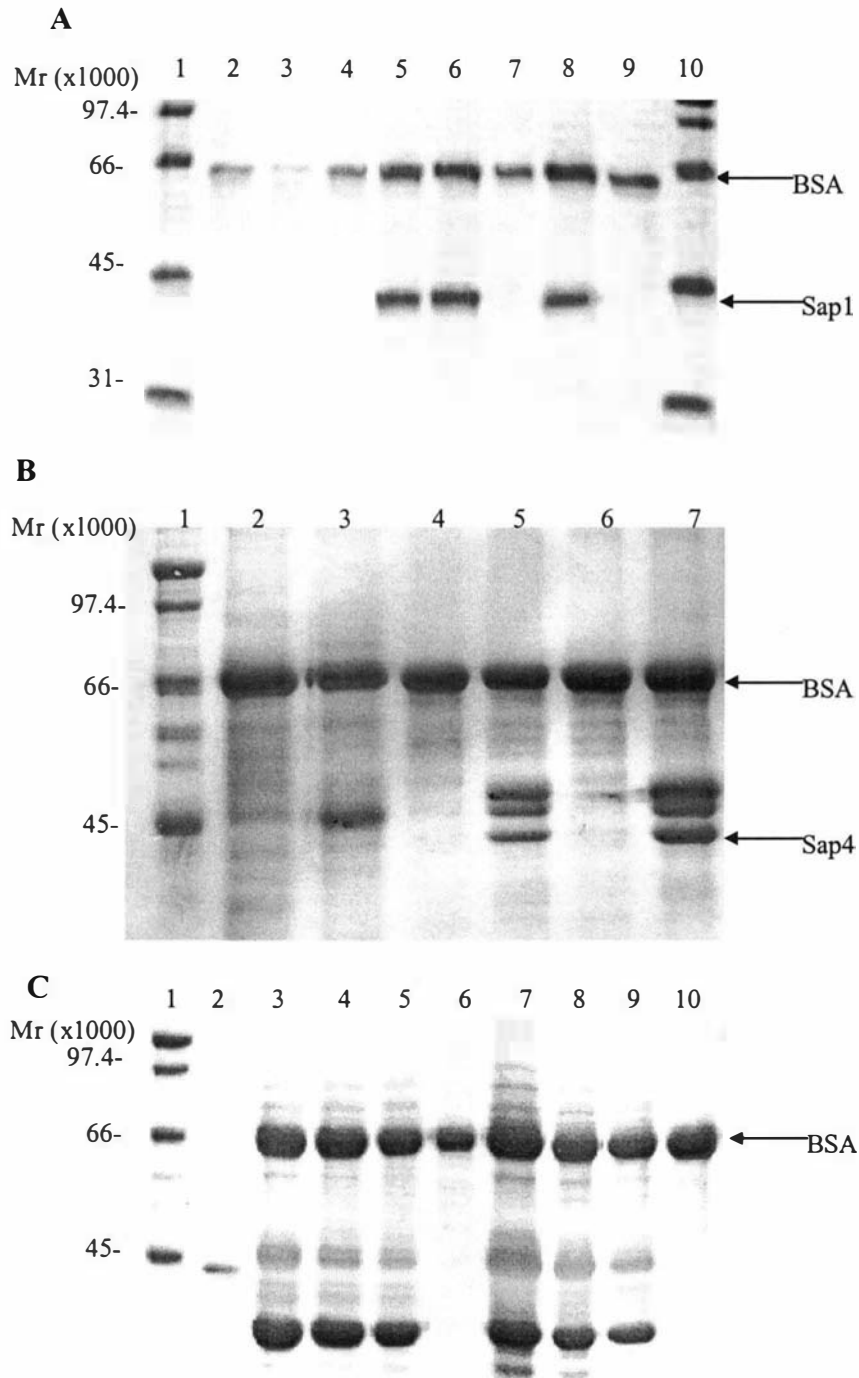


Figure 4.4 SDS-PAGE of preliminary Sap1, Sap4 and Sap5 expression trials

- A. Sap1 expression trial Lanes 1 & 10: Relative molecular mass markers
Lanes 2, 3, 4, 7 & 9: T₄₈ samples from induced putative transformants not expressing recombinant Sap1
Lanes 5, 6 & 8: T₄₈ samples from induced putative transformants expressing recombinant Sap1 (named MS1C, MS1B & MS1A respectively).
- B. Sap4 expression trial. Lane 1: Relative molecular mass markers.
Lanes 2 & 3: 0 h and 54 h samples (respectively) from induced KM71 control
Lanes 4 & 5: 0 h and 54 h time points (respectively) from strain MS4A
Lanes 6 & 7: 0 h and 54 h time points (respectively) from strain MS4B
- C. Sap5 expression trial. Lane 1: Relative molecular mass markers.
Lane 2: native Sap2 control
Lanes 3, 4, 5 & 6: 168 h, 72 h, 48 h & 0 h samples (respectively) from MS5A
Lanes 7, 8, 9 & 10: 168 h, 72 h, 48 h & 0 h samples (respectively) from MS5B

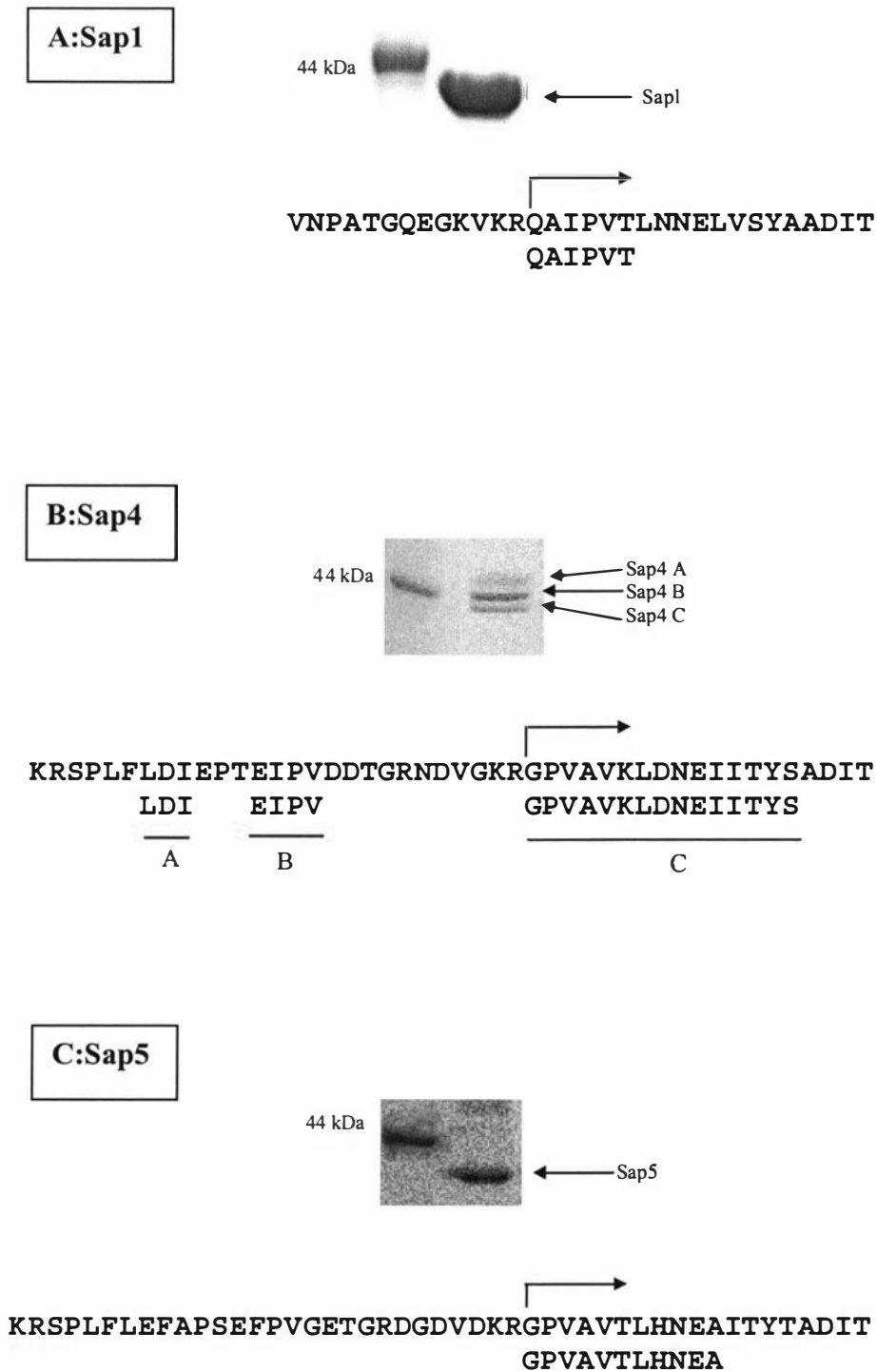


Figure 4.5 *N-terminal sequence of recombinant Sap proteins*

The figure shows an SDS-PAGE of the recombinant proteins and the corresponding N-terminal sequences, obtained by electroblotting the desired band onto PVDF membrane (Methods 2.10.8) and sequencing on an Applied Biosystems 476A protein sequencer. The arrow above the amino acid sequence indicates the start of the mature Sap protein in each case.

- A. Sap1 gel and N-terminal sequence
- B. Sap4 gel and N-terminal sequence of all three species
- C. Sap5 gel and N-terminal sequence

4.6 Identification of the recombinant proteins

4.6.1 Identification by western blotting

In parallel with the N-terminal sequence analysis, western analyses were performed using a polyclonal Sap2 antibody preparation to confirm that the recombinant proteins produced by MS4A, MS5A and MS1A were Saps. A strong cross-reactivity between this Sap2 antibody preparation and recombinant Sap1 has been previously reported (Smolenski *et al.*, 1997). This confirmed that the putative Sap1 protein produced by MS1A is indeed recombinant Sap1 (Fig. 4.6A). A western blot of the putative Sap4 and Sap5 proteins also showed an immunoreaction with the Sap2 antibody preparation, albeit considerably weaker than with Sap1 and Sap2. This indicated that the proteins were most likely recombinant Sap4 (Fig. 4.6B) and recombinant Sap5 (Fig. 4.6C).

The three recombinant proteins seen in the Sap4 expression trial (Fig. 4.4, B) all cross-reacted with the Sap2 antibody preparation (Fig. 4.5 C), which suggested that all three bands represented misprocessed or modified forms of Sap4. The Sap1 and Sap5 expression systems did not show these extra species.

4.6.2 Identification by enzyme assay

Further confirmation of the identity of Sap1 and Sap4 was provided by performing enzyme assays with the supernatants from induced cultures. Initially, Sap4 and Sap5 showed no activity towards the BSA substrate, and were thought to be inactive (Methods 2.10.9). However, some Sap4 activity was detected by using denatured globin as the proteolytic substrate. Sap1 proteolytic activity was detected with both the BSA- and globin assays (Methods 2.10.9), but no Sap5 activity was seen with either assay (see Chapter 6.2).

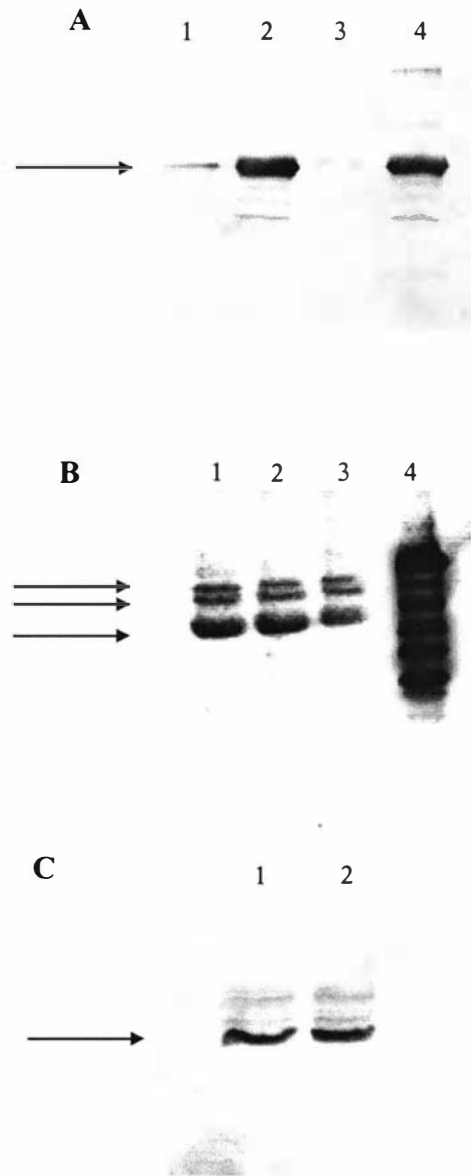


Figure 4.6 *Western analysis of Sap expression*

Aliquots of methanol-induced culture supernatant were TCA precipitated, resuspended and run on SDS-PAGE before electrotransfer to Hybond C. The western blots were performed as described in Methods 2.10.7. Arrows indicate the major immunoreactive bands.

- A. Lanes 1 & 3: 20 L of culture supernatant from MS1A and MS1B respectively (48 h induction)
Lanes 2 & 4: 800 L culture supernatant from MS1A and MS1B respectively (48 h induction)
- B. Lanes 1-3: 1 mL sample from MS4A after 96 h, 48 h and 24 h induction, respectively
Lane 4: 500 ng native Sap2. The laddering effect is caused by Sap2 autodegradation.
- C. Lanes 1 & 2: 1 mL sample from MS5A after 24 h and 48 h induction, respectively

4.7 Large scale expression of Sap1, Sap4 and Sap5

P. pastoris has been reported to produce up to g/L quantities of recombinant proteins (Cregg *et al.*, 1993), and produced 400 mg/L of recombinant exoglucanase of *C. albicans* (R. Ramsay, personal communication). In contrast, the yield of recombinant Saps from the small-scale induction trials was approximately 2-5 mg/L. Although considerably lower than both of these values, 2-5 mg/L would be a reasonable yield from a *S. cerevisiae* expression system, and would be ample for purification and characterisation of the pure enzymes. However, scale-up of the Sap expression system was initially unreliable and several unsuccessful attempts were made at large-scale induction of Sap1, Sap4 and Sap5. The induction protocol was identical each time (Methods 2.10.3), yet significant flask-to-flask variations were noted. A control experiment was performed in which a culture of MS4A was grown and induced in parallel with another *P. pastoris* recombinant strain, which expresses the *C. albicans* glucanosyl transferase, Bgl2 (kindly provided by R. Ramsay). Both cultures were induced under standard conditions, yet the level of expression of Sap4 was strikingly low compared to the level of Bgl2 (Fig. 4.7). This suggested that the expression problem was specific to the recombinant Sap strains, and the standard induction protocol, which had previously been successful with Exg and Bgl2 (R. Ramsay, personal communication), was abandoned in favour of a novel protocol for the Sap isoenzymes.

4.7.1 Optimising strain storage conditions

To investigate whether the storage conditions of the transformants had any effect on the level of induction of Saps, MS4A was streaked onto four freshly prepared plates: YPG, YPG plus G418 (0.5 mg/mL), minimal and minimal plus G418 (0.5 mg/mL). The cells were grown at 30 °C for 2 d, then used to inoculate 50 mL of starter medium in a 250 mL flask. After the culture reached $A_{600} \approx 20$, the cells were resuspended in induction medium and Sap expression was followed for 48 h (Methods 2.10.3). Fig. 4.8 shows that maintenance of the strains on minimal medium gives the best induction results. Therefore, for all subsequent work, the transformed strains were maintained on minimal medium plates for culture inoculation.

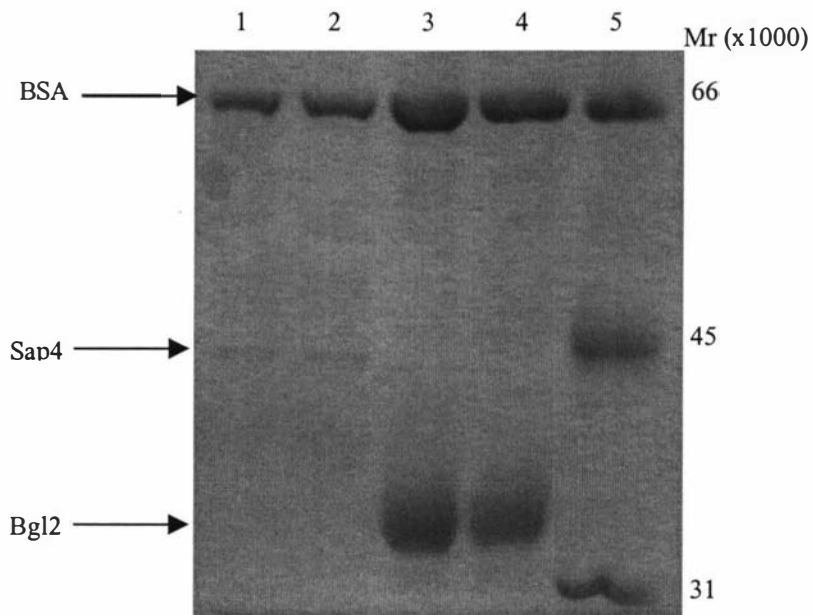


Figure 4.7 *Comparative induction of Sap4 and Bgl2*

Induction of Sap4 and Bgl2 cultures was performed using the standard protocol (Methods 2.10.3) and samples of culture supernatant (1 mL) were taken after 24 h and 48 h. Samples were TCA precipitated (Methods 2.10.4), resuspended, separated by SDS-PAGE and visualised with Coomassie Brilliant Blue™.

Lane 1: 24 h sample from Sap4 induction culture (MS4A)
 Lane 2: 48 h sample from Sap4 induction culture (MS4A)
 Lane 3: 24 h sample from Bgl2 induction culture
 Lane 4: 48 h sample from Bgl2 induction culture
 Lane 5: Molecular mass markers.

Effect of different storage media on Sap expression

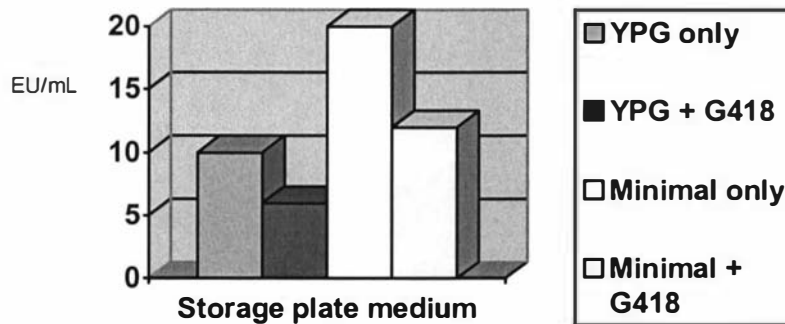


Figure 4.8 *Effect of storage medium on Sap expression*

The graph shows the effect of storage on four different plates on the yield of Sap4. The globin Sap assay was used to determine yield, and enzyme activity is expressed as enzyme units (EU)/mL (μg tyrosine equivalents released/min/mL culture supernatant).

Effect of starter culture medium on Sap expression

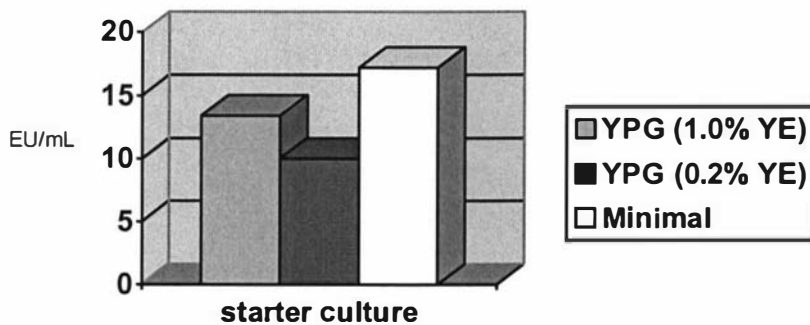


Figure 4.9 *Effect of starter culture medium on Sap expression*

The effects of three induction media were tested. The globin Sap assay was used to determine yield, and enzyme activity was expressed as: μg tyrosine equivalents released/min/mL culture supernatant.

4.7.2 Optimising starter culture conditions for Sap expression

The effect of different starter culture media on subsequent Sap expression was investigated, to determine which was best for enzyme induction. Three different media were tested, YPG (standard recipe, containing 1.0 % yeast extract), modified YPG (containing 0.2 % yeast extract) and minimal medium (see Methods 2.8.1 for recipes). Fig. 4.9 shows the average results of three separate trials using these three media. Minimal medium shows the highest level of Sap induction, followed by standard YPG, then modified YPG. However, growth of the cultures in either standard or modified YPG is twice as fast as in minimal medium, therefore standard YPG was selected for routine use.

4.7.3 Effect of methanol concentration on Sap expression

The standard induction protocol for *P. pastoris* recommends addition of 0.5 % methanol (Invitrogen, 1996) to the culture medium. To investigate whether additional methanol had any effect on Sap induction, four identical cultures were grown in standard YPG then induced with 0.5 %, 1.0 %, 2.0 % or 5.0 % (v/v) methanol, added to the cultures every 24 h. The results of this trial are summarised in Fig. 4.10. The levels of Sap induction varied significantly, but the level of methanol which gave optimum induction was 1.0 %. While there was a noticeable drop in expression in response to 2.0 % methanol, 5.0 % methanol induced Sap to a level equivalent to 0.5 % methanol, indicating that addition of up to 5.0 % methanol was not detrimental to the system. Methanol addition to 1.0 % was selected for routine use in subsequent cultures.

4.7.4 Effect of added buffer on Sap expression

Addition of neutral pH buffer to the induction medium has been reported to increase the yields of some recombinant proteins produced by *P. pastoris* by inhibiting native *P. pastoris* secreted proteases (Cregg *et al.*, 1993). Therefore phosphate buffer (pH 7.0)

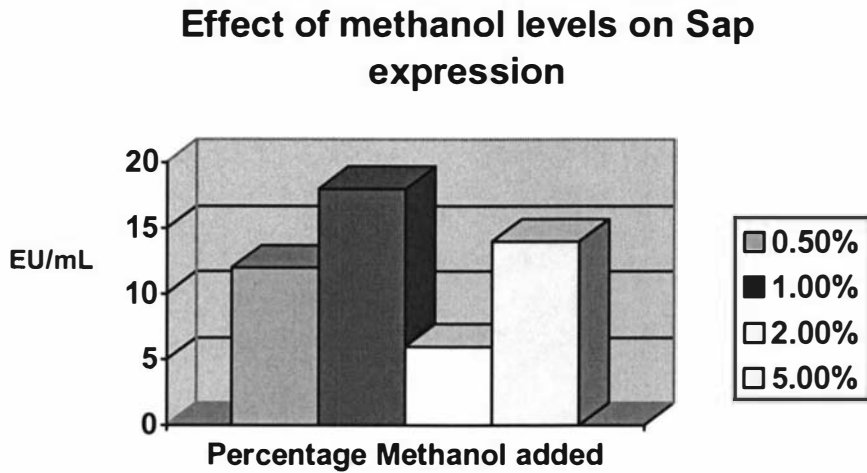


Figure 4.10 *Effect of methanol concentration on Sap expression*

The graph shows the effect of addition of four different amounts of methanol on the yield of Sap4. The globin Sap assay was used to determine yield, and enzyme activity was expressed as EU/mL.

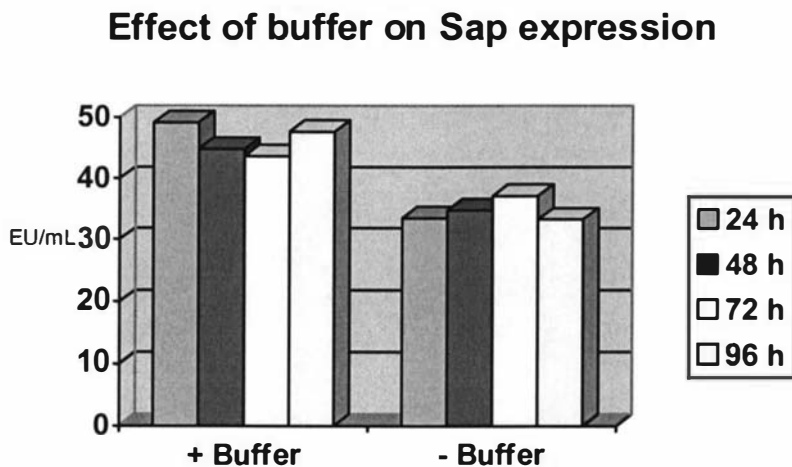


Figure 4.11 *Effect of buffer on Sap expression*

The graph shows the effect of added 100 mM Na phosphate, pH 7.0 on the yield of Sap4 over a 96 h time course. The globin Sap assay was used to determine yield, and enzyme activity was expressed as: μg tyrosine equivalents released/min/mL culture supernatant.

was added to the induction medium to test whether it increased the yield of recombinant Sap.

Fig. 4.11 clearly shows higher overall levels of Sap4 induction with the addition of phosphate, pH 7.0. The reason for this increase might be twofold. Typically, at low pH *C. albicans* Saps undergo a process of autodegradation (refer Fig. 4.6, lane 4; Smolenski *et al.*, 1997). Hence, the increase in Sap4 expression may be due in part to stabilisation of the protein at pH 7.0. The increase in yield could also be due in part to an inhibition of the native proteases. Subsequently, 100 mM Na phosphate (pH 7.0) was always added to the induction medium.

4.7.5 Effect of induction culture size on Sap expression

Whenever Sap expression was scaled up, the yield of protein (in mg/L) appeared to decrease compared to smaller cultures (data not shown). To investigate this, a trial was performed using identical cultures of increasing size. All other growth and induction parameters remained the same, including a 1:5 ratio of culture volume:flask volume. Also all four induction cultures (10 mL, 20 mL, 50 mL and 100 mL) were grown and induced from a single YPG starter culture to eliminate variation between starter cultures. The results of the trial (Fig. 4.12) revealed an inverse relationship between increasing culture size and Sap expression. This needed to be circumvented for subsequent large-scale induction of the Sap.

4.7.6 Effect of culture aeration on Sap expression

One possible explanation for the results in section 4.6.5 is that increasing size of the cultures resulted in decreased culture aeration. This was tested by inducing two 100 mL cultures in two different sized flasks. The control culture was in a standard 500 mL flask, and the second culture was in a 1 L baffled flask. The cultures were induced and samples were taken every 24 h for 72 h. Yields of Sap determined with the globin assay are summarised in Fig. 4.13.

Effect of induction culture volume on Sap expression

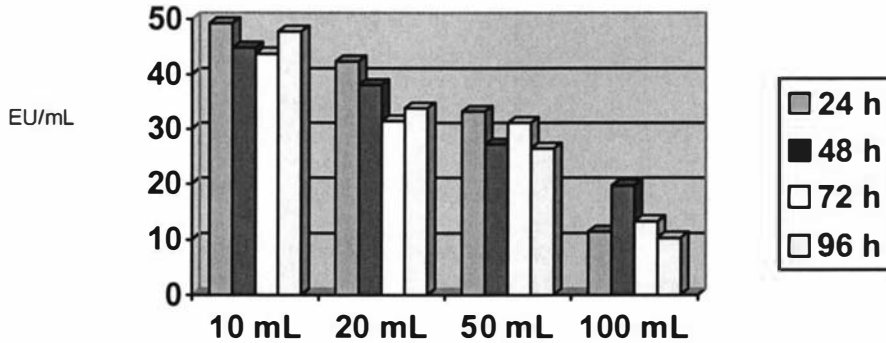


Figure 4.12 *Effect of induction culture volume on Sap expression*

Cultures of 10 mL, 20 mL, 50 mL and 100 mL in 50 mL, 100 mL, 250 mL and 500 mL flasks were induced and assayed for Sap activity at the times indicated. The globin Sap assay was used to determine yield, and enzyme activity was expressed as: μg tyrosine equivalents released/min/mL culture supernatant.

Effect of aeration on Sap expression

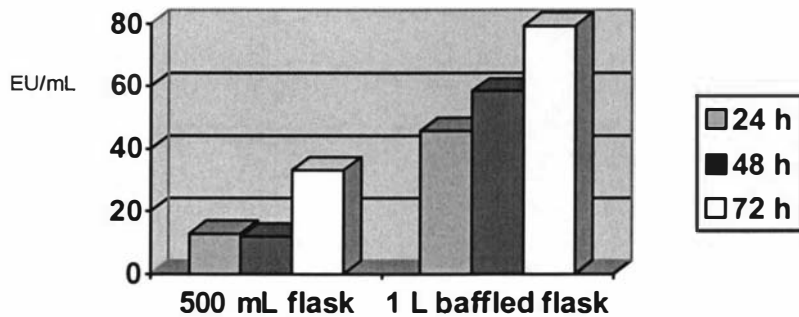


Figure 4.13 *Effect of aeration on Sap expression*

The effect of aeration on the yield of Sap4 was tested over 72 h in three 100 mL cultures in 500 mL standard flasks and in a 1 L baffled flasks. The globin Sap assay was used to determine yield, and enzyme activity was expressed as: μg tyrosine equivalents released/min/mL culture supernatant.

It appears that culture aeration is crucial for successful Sap induction on a large scale. This suggested that an aerated fermenter system could be the best method for protein expression.

4.8 Fermentation scale induction of Sap

P. pastoris is well suited to fermentation-scale growth, and as aeration was shown to be vital for production of significant quantities of Sap, fermentation-scale experiments were attempted. One consideration is the Mut^S phenotype of the host strain, KM71. This means that the culture needs to be grown to high cell density prior to induction with methanol, as subsequent proliferation is extremely slow (typically Mut⁺ hosts are used for fermentation-scale protein expression).

The usual process of protein expression involved growth of the starter cultures to an OD₆₀₀ of approximately 20, sterile centrifugation to harvest the cells, then transfer of those cells into sterile induction medium. As no facilities were available to allow sterile harvest of a 7 L fermentation-scale culture, a different induction strategy was planned. This required the *in situ* induction of the cells in the starter culture medium, which had not previously been tested.

P. pastoris is well suited to growth in deep fermenter cultures with vigorous aeration. A 10 L fermenter vessel was available, but it was not equipped with facilities for efficient retention of volatile constituents such as methanol, or sterile harvesting of the primary culture for the inoculation of induction cultures at high cell density. Several experiments were carried out to adapt the fermenter for Sap induction.

Two attempts at fermentation-scale growth of MS4A were unsuccessful, despite modifications of the growth medium (i.e. using modified YPG with 0.2% YE) and other variables such as induction culture density. These results suggest that it was important to remove the starter culture medium to achieve Sap expression, and therefore a new strategy was required.

4.9 Development of a new system for Sap expression

Following the disappointing fermentation-scale experiments, a small-scale (1 L) fermenter system was tested for Sap production. Starter cultures were grown in regular 2 L baffled flasks, and the cells were harvested under sterile conditions by centrifugation and resuspended in fresh induction medium. Induction was carried out under sterile conditions in a 2 L Schott bottle with vigorous aeration provided by a magnetic stirrer and sparger (Fig. 4.14).

The culture conditions for induction in the sparger system were as follows. Five 400 mL starter cultures were grown in 2 L baffled flasks overnight at 30 °C with shaking (200 rpm). Once the OD₆₀₀ of the culture reached approximately 20 (usually 24 h) the cells were harvested by centrifugation (5 000 rpm, 30 min) then resuspended into 1 L of induction medium (Methods 2.8.1). This suspension was then placed into the sterile Schott bottle, the sterile sparger inserted, and then the system was sealed. The incoming air was filter sterilised, and the outlet tubing was placed in a flask to ensure that no recombinant yeast escaped into the environment. The air was delivered into the culture at a rate of 1 L per min, and the culture was stirred to keep the cells in suspension and to disperse the incoming air.

The sparger system produced approximately 2-5 mg/L recombinant Sap which was comparable to the yield from small shake flask cultures. The sparger system was used for all subsequent Sap expression work. Six consecutive Sap4 expression cultures were induced in the system, and the supernatants were pooled for subsequent purification work (Chapter 5.4). This system was also successfully used for expression of several other secreted proteins from *P. pastoris* (R. J. Sanders, personal communication).

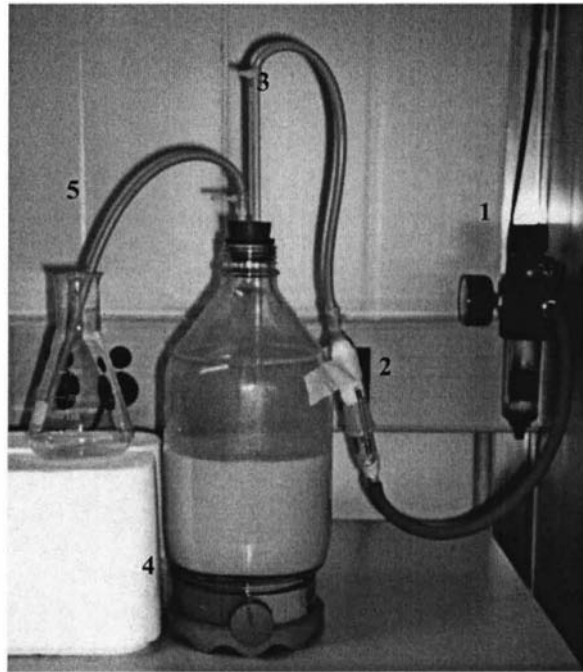


Figure 4.14 *Sparger system for induction of recombinant Sap*

Photograph shows the sparger system running during induction of an MS4A culture.

1. The air inlet valve for adjusting the rate of culture aeration
2. The filter sterilisation unit for the inlet air
3. The sterile sparger
4. The magnetic stirrer, to keep the culture in suspension
5. The outlet hose and receptacle

All components of the system (including tubing) were autoclaved at 120 psi for 20 min prior to addition of the culture.

4.10 *E. coli* expression of Sap4

Preliminary studies (Chapter 4.6.2) showed that *P. pastoris* transformed with *SAP4* or *SAP5* produced protein bands of the expected size but no enzyme activity was detected. In subsequent experiments enzymic activity was detected with both preparations, purified and studied in some detail (Chapter 6). However, at the time it seemed possible that there could be inherent problems with the expression, processing and secretion of active aspartic proteinases by *P. pastoris* and an alternative expression strategy, using *E. coli* as the host, was pursued in parallel with the work on the *P. pastoris* system. *E. coli* was previously used for successful production of recombinant candidropsin, an extracellular aspartic proteinase from *C. tropicalis*. The zymogen (minus its hydrophobic signal sequence) was purified from inclusion bodies, renatured, and then the pro-segment was removed (either by incubation at low pH or limited trypsin digestion), to produce mature, active enzyme (Lin *et al.*, 1993).

This section describes the experimental design and results obtained from attempts to produce active, recombinant Sap4, employing an expression strategy similar to that of Lin *et al.* (1993).

4.10.1 PCR primer design

To facilitate over-production of recombinant Sap4 in *E. coli*, a suitable expression vector was constructed. This required engineering a restriction site between the predicted end of the signal sequence and start of the zymogen pro-sequence. This was done by PCR mutagenesis using pMS4b, the pBluescript-based plasmid which contained the *SAP4* gene as the template.

The amino acid and nucleotide sequences of this region of *SAP4*, and the PCR primers that were designed are shown in Fig. 4.15. The upstream PCR primer resulted in two amino acid mutations in the Sap4 sequence (Fig. 4.15). However, the position of the mutations at the 5' end of the pro-sequence of the protein suggest they would be unlikely to interfere with protein folding. In addition, the mutations would be removed with pro sequence cleavage, when the protein became activated. *Bam* HI and *Pst* I

Upstream Primer

L A F A L L I D A A P V K R S T G F V
 CTTGCTTTCGCTTTATTAATTGATGCTGCTCCAGTTAAAAGATCTACAGGGTTTGT
^{5'}GGGGATCCGCTCCAGTTAAAAGATCT_{3'}

Recombinant Sap4:

G S A P V K R S T G F V

Downstream primer

K Y T S E S N I V A I N *
 AAATACACTTCCCAGTCTAACATTGTTGGTATTAATTAG
_{3'}TGTAACAACCATAATTAATTT**CGACGT**CAAA^{5'}

Figure 4.15 PCR primers for *E. coli* expression

The upstream primer (homologous to the coding strand) contains a *Bam* HI restriction site between the predicted end of the signal sequence and start of the zymogen pro-region. The downstream primer (homologous to the non-coding strand) contains a *Pst* I restriction site. Restriction sites are highlighted in bold type and amino acid mutations at the N-terminus are shown in grey. 5'-3' direction of primers is indicated.

The PCR reaction was performed as described in Method 2.9.11, using 1 ng of pMS4b as the template DNA. The program used to PCR amplify the *SAP4* sequence was designed as follows:

94 °C	5 min	
94 °C	10 s	2 cycles
56 °C	30 s	
72 °C	2 min	
94 °C	10 s	2 cycles
54 °C	30 s	
72 °C	2 min	
94 °C	10 s	2 cycles
52 °C	30 s	
72 °C	2 min	
94 °C	10 s	30 cycles
50 °C	30 s	
72 °C	2 min	

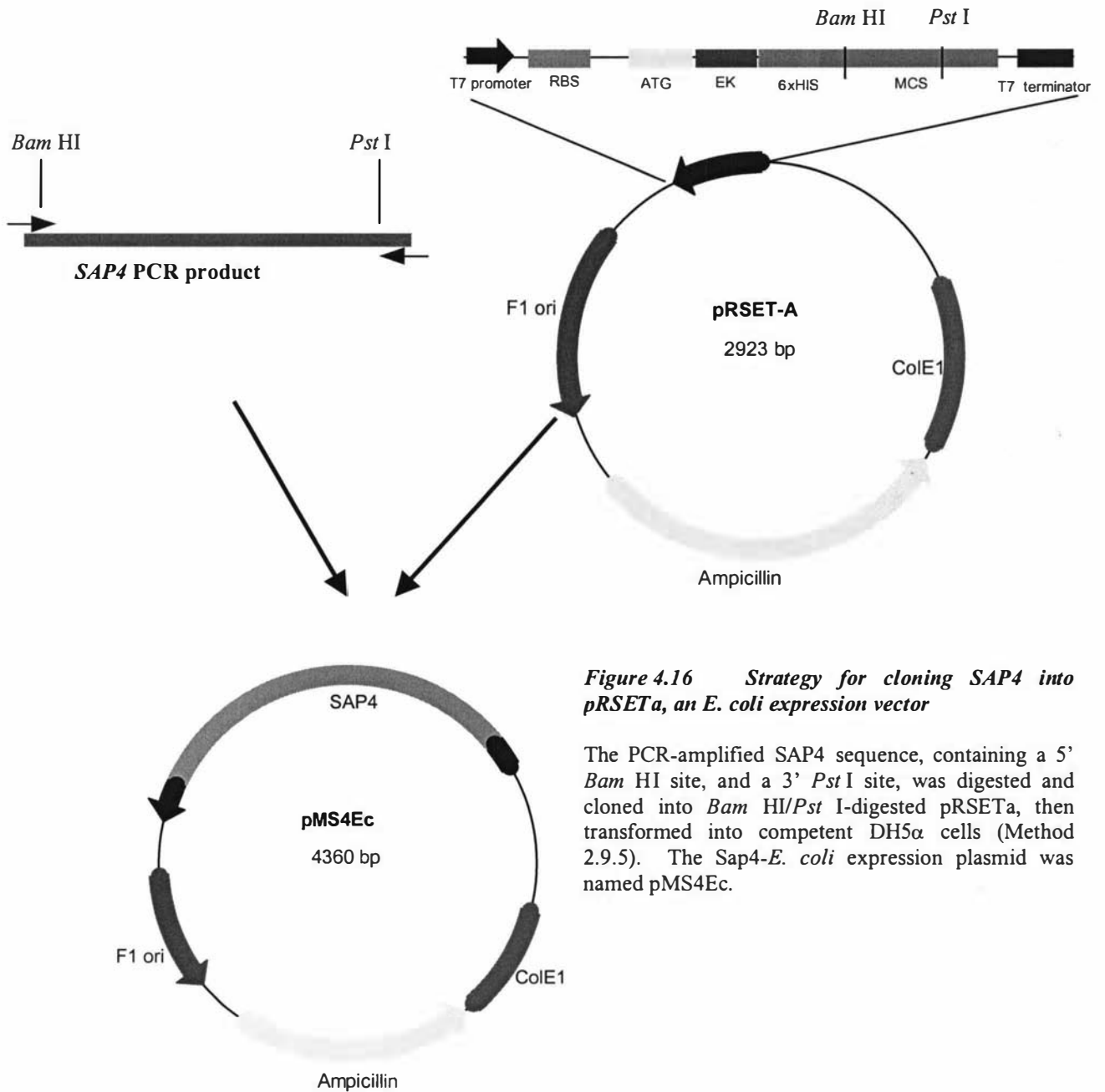


Figure 4.16 Strategy for cloning *SAP4* into *pRSETa*, an *E. coli* expression vector

The PCR-amplified *SAP4* sequence, containing a 5' *Bam* HI site, and a 3' *Pst* I site, was digested and cloned into *Bam* HI/*Pst* I-digested *pRSETa*, then transformed into competent DH5 α cells (Method 2.9.5). The *Sap4-E. coli* expression plasmid was named *pMS4Ec*.

restriction sites were introduced (Fig. 4.15) to enable forced (directional) cloning of the *SAP4* sequence into pRSETa (Invitrogen). This strategy also ensured the *SAP4* ORF was in the correct reading frame with respect to the vector sequences (Fig. 4.16).

4.10.2 Transformation of host strains

Blue-white selection was used to identify putative *E. coli* transformants (Method 2.9.6). These colonies were picked and grown overnight for mini-scale plasmid DNA preparations (Method 2.9.6). Transformants which contained the 1.4 kb *SAP4* insert were identified by digesting the plasmid DNA with *Bam* HI and *Pst* I (data not shown). The *E. coli* strains BL21 and AD202 were chosen for expression of recombinant Sap4. The strains are very similar, with slightly different genetic backgrounds (refer Method 2.4). A large-scale preparation of plasmid DNA (Method 2.9.8) was used to transform these strains.

4.10.3 Expression of recombinant Sap4

Freshly transformed AD202 and BL21 cells were used for the Sap4 expression trials. The trial also included the untransformed strain as a control. Induction was carried out by inoculating 5 mL of TB medium with a single colony (ampicillin was added to maintain the plasmid where necessary). Cultures were incubated at 37 °C, with shaking (200 rpm) until the attenuation of the culture at 600 nm reached approximately 0.5. At this stage a small sample was taken for a plasmid occupancy analysis to check that more than 50% of the cells contained the plasmid pMS4Ec. This was tested by plating dilutions of the culture onto TB and TB + ampicillin, and comparing the number of colonies.

The cultures were then either induced with IPTG or grown as uninduced controls as follows:

Strain only	Uninduced
Strain + plasmid	Uninduced
Strain + plasmid	Induced with IPTG

The cultures were induced for 2 h, with 1 mM IPTG, 1.6 mL samples were centrifuged at 13,000 rpm for 1 min, then resuspended into 300 μ L PBS. Samples were prepared for SDS-PAGE by sonication (3 x 15 s), then either loaded immediately onto the gel or centrifuged at 13,000 rpm for 5 min, and the supernatants and pellets were loaded separately.

In the initial experiments, expression was monitored by SDS-PAGE and western blotting. Unfortunately, the results were equivocal, as the Sap2 antiserum was not specific enough to produce any clear signals corresponding to the recombinant Sap4 (data not shown). Therefore, once the anti-Sap4 became available (See Chapter 6.11), these experiments were repeated to ascertain the level of expression of Sap4 in *E. coli* (Fig. 4.17).

Fig. 4.17 shows an SDS-PAGE gel (A), a western blot using anti-Sap2 (B) and an equivalent western blot probed with anti-Sap4 (C). These western blots revealed how uninformative anti-Sap2 was for analysis of Sap4 expression in *E. coli*, and highlight the benefit of specific antibodies, in this case anti-Sap4.

The western analysis showed that Sap4 was not expressed in strain AD202, but was expressed in strain BL21. This is unusual given the genetic similarity between the strains, and there was no clear explanation for this result. In BL21, Sap4 expression levels were higher in the uninduced control, than in the induced sample. The presence of Sap4 in the uninduced sample suggested that the promoter for Sap4 expression was “leaky”. However, decreased levels of Sap4 in the induced sample indicated that Sap4 expression may somehow be detrimental to the cells. The western blot also revealed that recombinant Sap4 was located in the insoluble fraction.

These results suggested that there could be a number of problems with *E. coli* expression of Sap4. Fortunately, the parallel development of the *P. pastoris* expression system provided a better source of active enzymes, which were secreted and processed by the host.

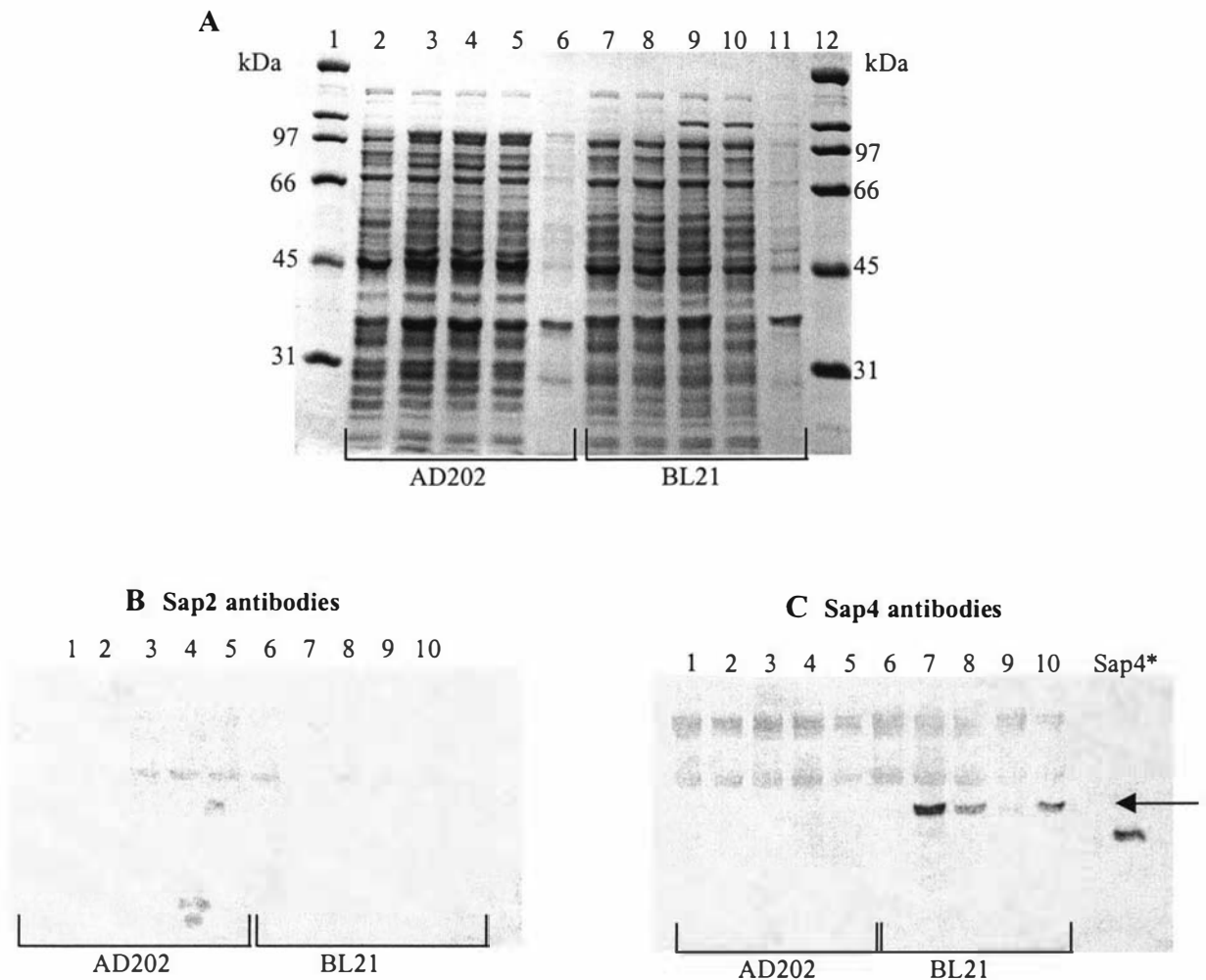


Figure 4.17 *SDS-PAGE and western blot analysis of E. coli expression of Sap4*

A. 10% SDS-PAGE of induction trial using AD202 and BL21 cells.
 Lanes 1 & 12: Molecular weight markers
 Lanes 2 & 7: Whole cells, untransformed
 Lanes 3 & 8: Whole cells, transformed with pMS4Ec, uninduced
 Lanes 4 & 9: Whole cells, transformed with pMS4Ec, induced
 Lanes 5 & 10: Supernatant, induced cells
 Lanes 6 & 11: Insoluble pellet, induced cells

B, C. Western blot of above gel using anti-Sap2 (B) and anti-Sap4 (C) preparations.
 Lanes 1 & 6: Whole cells, untransformed
 Lanes 2 & 7: Whole cells, transformed with pMS4Ec, uninduced
 Lanes 3 & 8: Whole cells, transformed with pMS4Ec, induced
 Lanes 4 & 9: Supernatant, induced cells
 Lanes 5 & 10: Insoluble pellet, induced cells
 *Note: 200 ng Sap4 was run as a positive control for this western.
 Arrow indicates recombinant Sap4 fusion protein.

5.0 PURIFICATION OF Saps

5.1 Introduction

Once the induction strategy had been optimised to produce milligram quantities of recombinant Sap proteins, the next step was to develop purification protocols for the enzymes. Frequently, protein purification can be problematic, and at best, the development of successful purification protocols can be described as semi-systematic. Often protocols are devised from trial-and-error experiments, which can be based around known protocols used for similar proteins. While this is not always successful, it often provides a good starting point, and fortunately, the *C. albicans* Sap1, Sap2 and Sap3 isoenzymes had been previously purified using a variety of techniques (Smolenski *et al.*, 1997; Morrison *et al.*, 1993; Ruchel, 1981). In this case, the simple one-step protocol developed by Smolenski *et al.* (1997) was used as the starting point for purification of recombinant Sap1, Sap4 and Sap5 from *P. pastoris*. While homologous proteins are not expected to behave in an identical manner during purification, they often have some similar properties, depending on the degree of identity between the homologues. Therefore both initial success and failure in purification trials yields a lot of information about the properties of the protein, which can then be used to further

refine the procedure to suit the particular protein. This chapter describes the purification protocols developed for Sap1, Sap4 and Sap5.

5.2 Purification of recombinant Sap1

Smolenski *et al.* (1997) details the purification of recombinant *C. albicans* Sap1 from *S. cerevisiae* using a single-step ion exchange protocol. Therefore to purify recombinant *C. albicans* Sap1 from the supernatant of the *P. pastoris* transformant, (the same protein expressed in a different recombinant host species), the same procedure was followed.

Three *P. pastoris* transformants which expressed Sap1 were identified in section 4.4 (Fig. 4.4). The culture supernatant of these three transformants (MS1A, MS1B and MS1C) was harvested by centrifugation (5,000 rpm, 10 min) then pooled for purification trials. The pooled supernatant (140 mL) was then concentrated in a Filtron™ disposable filtration unit (10 K cut-off) to 15 mL. This concentrate was buffer exchanged with 10 volumes of 20 mM Bis-Tris, pH 6.0 concentrated to 14.5 mL, and filtered through sterile 0.22 µm membranes to remove particulate matter prior to anion exchange chromatography on a Pharmacia Mono Q column (HR5/5 1 mL bed volume).

As described by Smolenski *et al.* (1997), samples (3 mL) were applied to the Mono Q column pre-equilibrated with 10 column volumes of 20 mM Bis-Tris, pH 6.0. A linear gradient from 0 to 300 mM KCl over 30 min was run to elute the bound protein. Fig. 5.1 shows the trace of the chromatography run. The retention time of the single major peak corresponds to the retention time for recombinant Sap1 from *S. cerevisiae* (Smolenski *et al.*, 1997), under the same conditions. Only the unbound eluate and the major peak were assayed as there were no other significant protein peaks in the elution profile, which is typical of *P. pastoris* expression systems (Cregg *et al.*, 1993). The large peak which eluted after 17 min was confirmed to be Sap1 by enzyme activity (Table 5.1) and N-terminal sequence (see Chapter 4.5). SDS-PAGE analysis of the Mono Q fractions (Fig. 5.2) showed the Sap1 protein was purified to apparent homogeneity in a single step.

These observations confirm the identity of recombinant Sap1 from *P. pastoris* and revealed the protein was behaving in the same manner as recombinant Sap1 from *S. cerevisiae*. This recombinant enzyme therefore made a good control for comparison to the previously uncharacterised proteins, Sap4 and Sap5.

5.3 Purification of recombinant Sap4

5.3.1 Anion exchange at pH 6.0

Small samples (150 mL) of induced MS4A culture supernatant were used in trials to develop a purification protocol for Sap4. The calculated isoelectric point (pI) of Sap4 is 5.04, similar to that calculated for Sap1 (4.22). Therefore, it was anticipated that Sap4 would behave in a similar manner to Sap1 during ion exchange chromatography. Culture supernatant (150 mL) was concentrated 5-fold, equilibrated in 20 mM Bis-Tris, pH 6.0 and 5 mL was applied to the Mono Q column (HR5/5 1 mL, pre-equilibrated in the same buffer). A linear gradient from 0-300 mM KCl was run over 30 min, and the eluate was collected as fractions (data not shown). The Sap4 was located by enzyme assay and SDS-PAGE. All of the Sap4 eluted in the unbound fraction, and there was no significant purification. The calculated pI of the protein suggested that at pH 6.0, Sap4 would be negatively charged and would therefore be expected to bind the Mono Q column but this was not the case. The calculated pI of Sap4 was not expected to be the same as the actual pI, and could be up to 1-2 units either side of the calculated value, depending on the number of buried charged residues in the mature folded protein. Therefore, further ion-exchange trials were attempted using different resins at various pH values before abandoning the method (data not shown).

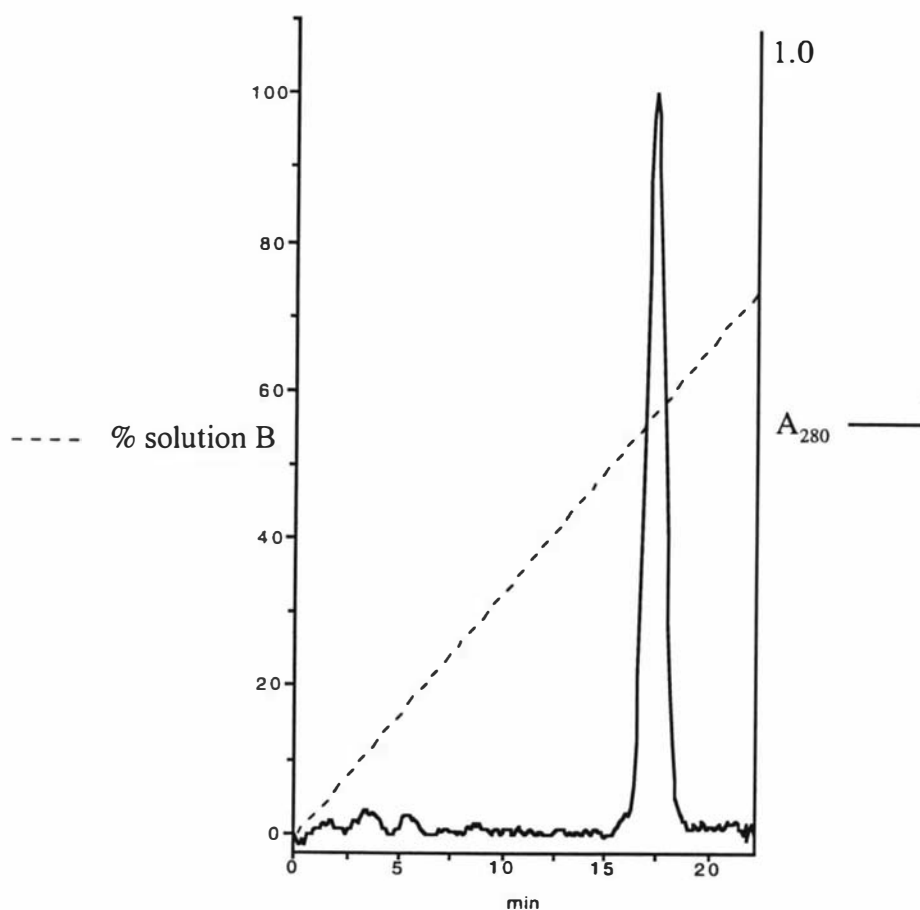


Figure 5.1 Purification of Sap1 using ion-exchange chromatography

The profile of Sap1 ion-exchange chromatography shows a single band eluting at approximately 17 min (130 mM KCl or 45% B). The Sap1 eluted as a sharp single peak. Solution A; 20 mM Bis-Tris pH 6.0, Solution B; 20 mM Bis-Tris pH 6.0, 500 mM KCl. Activity was measured using the globin assay.

SAMPLE	VOLUME	ACTIVITY (Total EU)	PROTEIN (mg/mL)	SPECIFIC ACTIVITY (EU/mg)	PURIFICATION (Fold)
Crude	140 mL	12,320	0.12	733	-
Mono Q	2 mL	8378	0.56	7480	10.2

Table 5.1 Purification of recombinant Sap1

The table summarises the purification of recombinant Sap1 from *P. pastoris*.

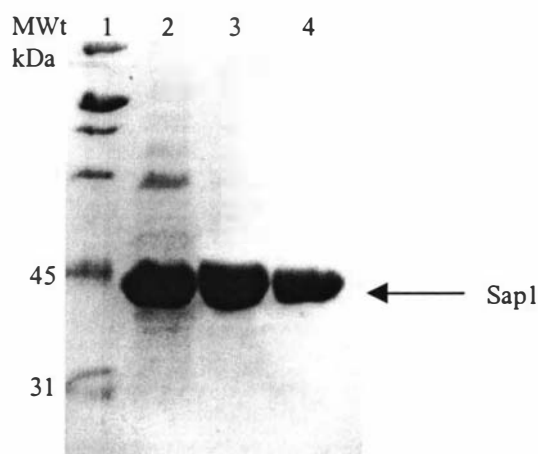


Figure 5.2 *SDS-PAGE of Sap1 fractions from ion exchange chromatography*

Samples were run on 10% SDS-PAGE and visualised by staining with Coomassie Brilliant Blue™.

Lane 1: Molecular weight markers

Lane 2: 40 μ L of concentrated buffer exchanged supernatant

Lanes 3 & 4: 80 μ L fractions of the Sap1 peak from Mono Q chromatography.

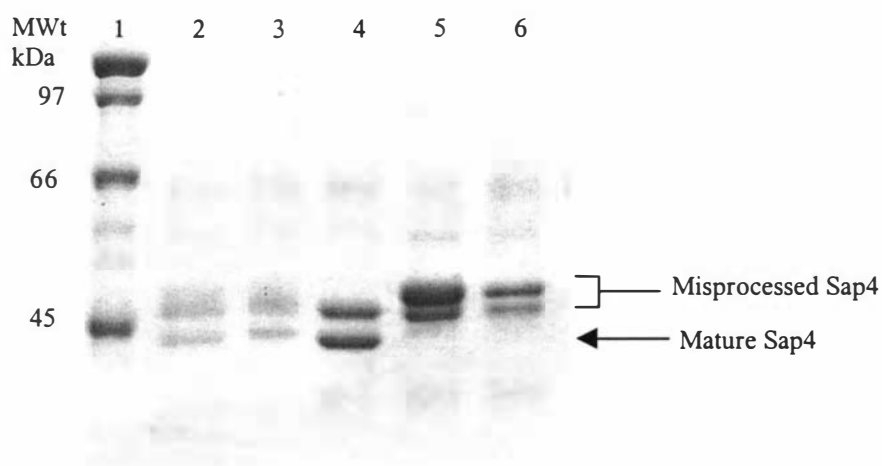


Figure 5.3 *SDS-PAGE of Sap4 fractions from ion exchange chromatography (pH 7.0)*

Samples were run on 10% SDS-PAGE and visualised by staining with Coomassie Brilliant Blue™.

Lane 1: Molecular weight markers

Lane 2: 800 μ L of precipitated supernatant

Lane 3: 10 μ L of concentrated buffer exchanged supernatant

Lane 4: 800 μ L of precipitated unbound eluate.

Lane 5: 800 μ L precipitated sample from the first major peak.

Lane 6: 800 μ L precipitated sample from the second major peak.

5.3.2 Cation exchange chromatography at pH 6.0

Anion exchange involves the interaction of negatively charged proteins with a positively charged medium. As Sap4 did not bind an anion exchange column at pH 6.0, it was possible it might interact with a cation exchange column at the same pH, if the actual pI of the protein was greater than 6.0. This would mean that the protein was positively charged at pH 6.0 rather than negatively charged like Sap1. Therefore cation exchange chromatography was tested using a Resource S column (Pharmacia Biotech, 1 mL bed volume) and the crude enzyme solution equilibrated in 20 mM phosphate, pH 6.0. A small sample (5 mL) was applied to the column and a linear 30 min gradient of 0-1 M NaCl was run to elute the bound proteins. No Sap4 activity was recovered in either the unbound eluate or in the fractions across the gradient (data not shown). Therefore cation exchange was not pursued further.

5.3.3 Anion exchange chromatography at pH 7.0

Due to the lack of success with cation exchange chromatography at pH 6.0, anion exchange was attempted at higher pH values. The pH of the sample was increased from pH 6.0 to pH 7.0, which should be above the pI of the protein, therefore producing an overall negative charge (groups such as histidine are deprotonated). This would then promote the binding of Sap4 to anion exchange columns such as Mono Q and Resource Q. A sample of supernatant (150 mL) was concentrated approximately 5-fold, buffer-exchanged with 10 volumes of 20 mM Tris-HCl, pH 7.0. A Resource Q column (Pharmacia Biotech, 1 mL bed volume) was also pre-equilibrated with 20 mM Tris-HCl, pH 7.0 prior to application of the sample. A linear gradient of 0-500 mM NaCl in 20 mM Tris-HCl, pH 7.0 (over 30 min) was run to elute the bound proteins. Enzyme assays and SDS-PAGE were used to identify fractions containing the Sap4 and to assess the relative purity. The SDS-PAGE indicated that the unbound eluate contained mature Sap4 and misprocessed enzyme while further misprocessed Sap4 was eluted in the gradient. This conclusion is based on the relative mobilities of the bands (Fig 5.3) and preliminary N-terminal sequencing of these bands (see Fig 4.5). The differential binding of the three recombinant Sap4 species to anion exchange medium at pH 7.0 suggested that this method might be optimised to allow separation of the three species.

5.3.4 Anion exchange at pH 7.0, low ionic strength

Candida Saps have been shown to undergo significant pH dependent denaturation above pH 7.0 (Wagner *et al.*, 1995). Therefore, increasing the pH of the supernatant above 7.0 was avoided, and instead the ionic strength of the buffer was lowered in an attempt to increase Sap4 binding to the column. Both the Resource Q column and the protein solution were equilibrated in 5 mM Tris-HCl pH 7.0. Again a linear gradient from 0-500 mM NaCl over 30 min was run to elute the bound proteins and Sap4-containing fractions were identified by enzyme assays and SDS-PAGE. Lowering the ionic strength of the buffer was successful, as it resulted in all three Sap4 species binding the column and eluting in a split peak (data not shown). SDS-PAGE suggested some separation of the misprocessed isoforms of Sap4 (Fig. 5.4), but to purify the mature Sap4 to homogeneity, it was necessary to optimise this separation.

5.3.5 Optimising the anion exchange gradient

Some separation of the Sap4 species was apparent in Fig. 5.4, and this was optimised using identical column conditions (section 5.3.3) and a longer gradient. Instead of a linear gradient from 0-500 mM NaCl, a two-step gradient was run. Initially, a slow linear gradient was run from 0-100 mM NaCl over 30 min, as the Sap4 species eluted from the column at low salt concentration. This was followed by a short (10 min) gradient from 100-500 mM NaCl to remove any remaining bound proteins from the column. The elution profile showed significant separation of the two peaks (Fig. 5.5), and SDS-PAGE confirmed that mature Sap4 was successfully separated from the misprocessed species of Sap4 (Fig. 5.6). This represents a one-step system for small-scale purification of mature Sap4 from induced culture supernatant.

This method, while simple and quick, was not suitable for a large-scale purification of Sap4. Another purification step was needed to enable processing of large quantities of buffer-exchanged culture supernatant prior to purification on Mono Q. As anion exchange was successful, purification trials were carried out using DEAE Sepharose.

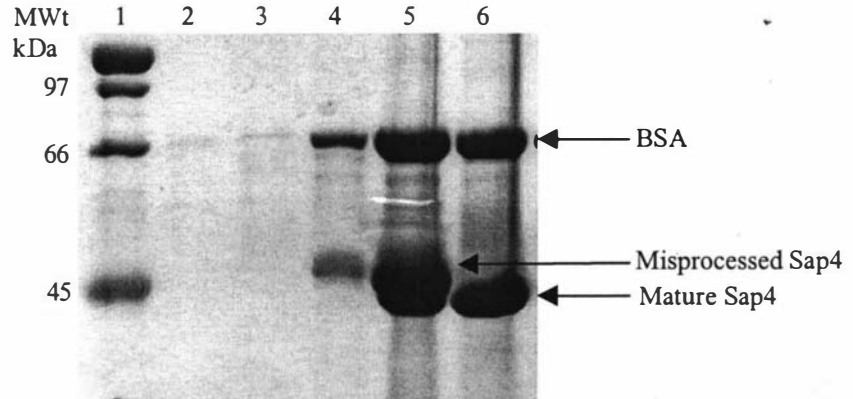


Figure 5.4 *SDS-PAGE of Sap4 fractions from anion exchange at low ionic strength (pH 7.0)*

Samples were run on 10% SDS-PAGE and visualised by staining with Coomassie Brilliant Blue™.

Lane 1: Molecular weight markers

Lanes 2 & 3: 800 µL TCA-precipitated unbound fraction

Lane 4: 800 µL TCA-precipitated sample from the front of the first Sap4 peak

Lane 5: 800 µL TCA-precipitated sample from the back of the first Sap4 peak

Lane 6: 800 µL TCA-precipitated sample from the second Sap4 peak

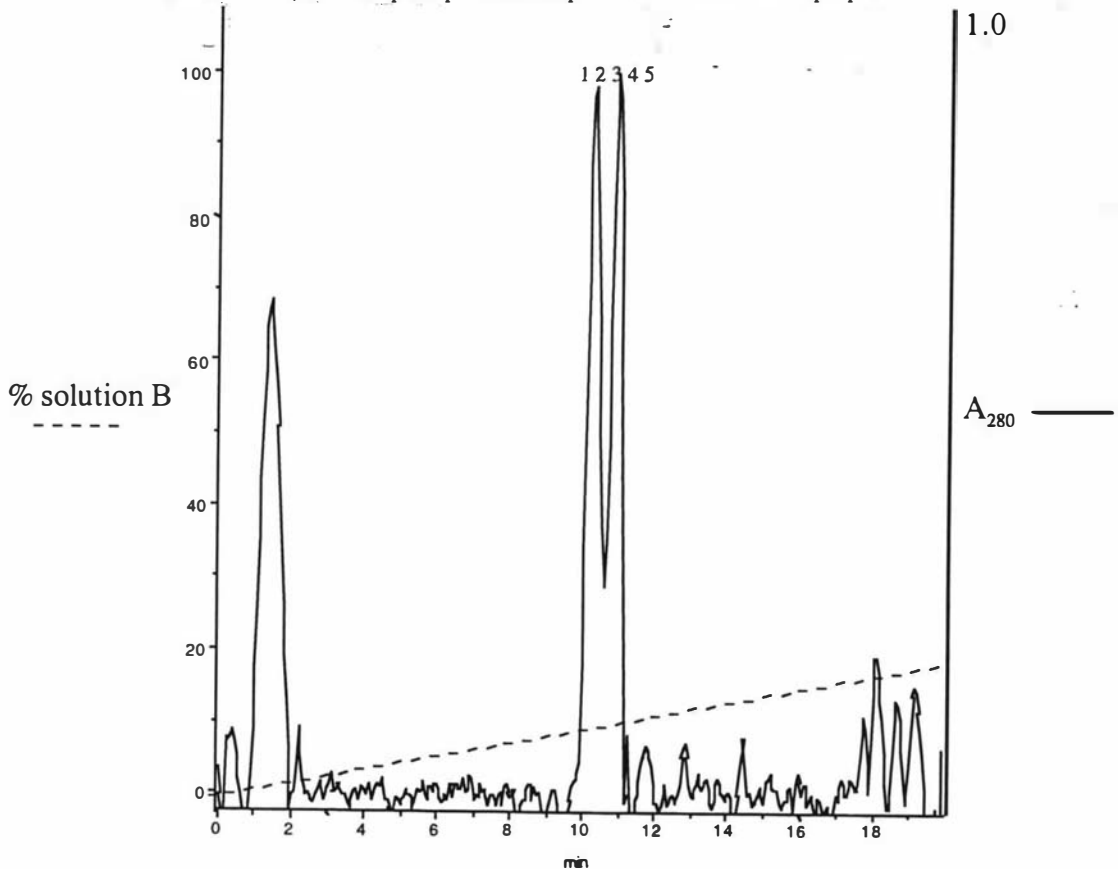


Figure 5.5 *Separation of the misprocessed species of Sap4 using Resource Q chromatography*

The profile shows separation of the Sap4 misprocessed species into two peaks using a shallow salt gradient. The column and sample were equilibrated in 5 mM Tris-HCl, pH 7.0 and a linear 0-100 mM NaCl gradient was run over 30 min. Fractions labelled on the profile were visualised on SDS-PAGE (Fig. 5.6).

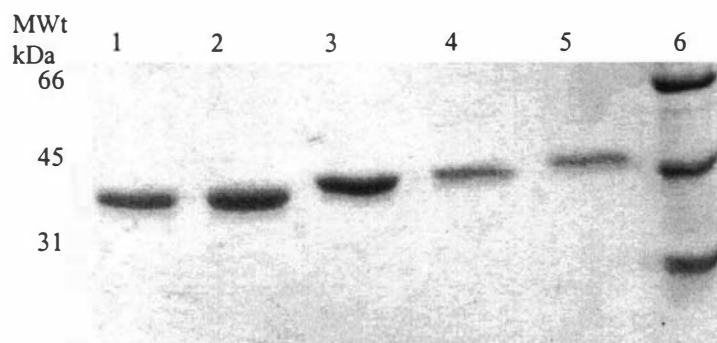


Figure 5.6 *SDS-PAGE of Sap4 fractions from optimised anion exchange gradient*

Samples were run on 10% SDS-PAGE and visualised by staining with Coomassie Brilliant Blue™.

Lane 1: 20 μ L sample from fraction 1 (mature Sap4)

Lane 2: 20 μ L sample from fraction 2 (mature Sap4)

Lane 3: 20 μ L sample from fraction 3 (misprocessed Sap4)

Lane 4: 20 μ L sample from fraction 4 (misprocessed Sap4)

Lane 5: 20 μ L sample from fraction 5 (misprocessed Sap4)

Lane 6: Molecular weight markers

5.3.6 Batchwise purification trials with DEAE Sepharose

To investigate the binding of the recombinant Sap4 species to DEAE Sepharose at pH 7.0, a small amount of loose resin was equilibrated with 5 mM Tris-HCl, pH 7.0. Microcentrifuge tubes containing aliquots of the resin (100 μ L) were mixed with a small volume of equilibrated Sap4 culture medium (500 μ L), then centrifuged at 13,000 rpm for 30 s to pellet the resin. The excess buffer was removed with a pipette, then the resin was washed with 5 mM Tris-HCl pH 7.0. Elution of bound proteins was performed in a stepwise manner, by adding 500 μ L aliquots of buffer containing increasing concentrations of NaCl (i.e. 50 mM, 100 mM, 200 mM, 500 mM & 1 M NaCl), mixing thoroughly, then spinning the tube to separate the resin and the buffer. Following removal of the buffer with a pipette, enzyme assays were performed to assess the binding of Sap4 to the DEAE Sepharose. Enzyme assays (Fig. 5.7) revealed that 26 % of Sap4 did not bind DEAE Sepharose, but that the remainder was eluted in fairly high

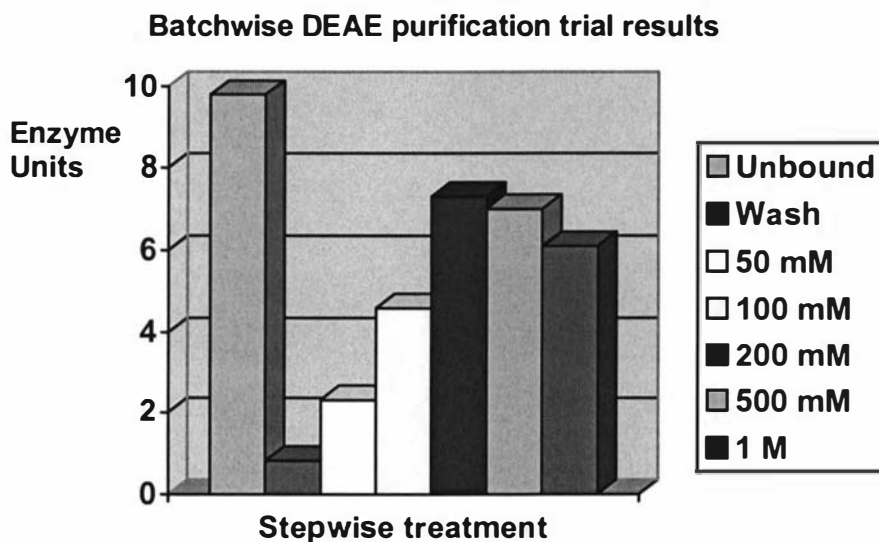


Figure 5.7 *Batchwise DEAE sepharose purification trial results*

The graph shows the recovery of Sap4 (in total enzyme units) in each wash fraction removed from the resin after centrifugation. Significant quantities of Sap4 were still being eluted in the high salt washes indicating a fairly strong interaction between DEAE Sepharose and the Sap4 species.

salt concentrations (100 mM - 1 M). The significant amount of unbound Sap4 activity was presumably due to saturation of the small bed volume of DEAE Sepharose, and this was reduced to residual levels when more resin was used (data not shown). These observations meant that DEAE Sepharose might provide a useful first step for preparative scale purification of Sap4, but further trials were required to establish the best conditions.

5.3.7 DEAE sepharose column chromatography

Having established that Sap4 bound to, and could be eluted from, DEAE Sepharose, it was necessary to test whether this method was appropriate for preparative-scale purification. A DEAE Sepharose column (26 mL bed-volume) was prepared and both the resin and a protein sample (150 mL) were equilibrated with 5 mM Tris-HCl, pH 7.0.

The enzyme solution was applied to the column, and a linear gradient between 0-500 mM NaCl (in 5 mM Tris-HCl, pH 7.0) was run over 125 min. All unbound material and the eluted peaks were collected as fractions and were analysed by SDS-PAGE and enzyme assay. Some separation of the Sap4 species was achieved across the gradient (Fig. 5.8, B), despite no obvious separation of peaks on the elution profile (Fig. 5.8, A). This indicated some differential binding and elution of the various Sap4 species. To investigate whether this differential binding could be enhanced at higher pH values, this trial was repeated under identical conditions at pH 7.5 and the results were analysed by enzyme assay and SDS-PAGE. However, where some separation of the Sap4 species was evident at pH 7.0, there was no separation of these isoforms at pH 7.5 (data not shown). Therefore, all subsequent DEAE Sepharose work was performed at pH 7.0.

5.4 Large-scale purification of mature Sap4

The large-scale purification of Sap4 followed the method developed in section 5.3; DEAE Sepharose chromatography, pH 7.0 (section 5.3.6) followed by fractionation on a Mono Q column at pH 7.0 (section 5.3.4). Using the sparger system (section 4.7), six 1 L cultures of MS4A were induced, harvested by centrifugation (5000 rpm, 30 min) and the supernatants were stored frozen prior to purification. The pooled supernatant (Fig. 5.12, lane 2) was concentrated to 1 L in an Amicon ultrafiltration system (CH2A concentrator) and buffer exchanged against 10 L of 5 mM Tris-HCl, pH 7.0. The volume was then reduced to 550 mL and the concentrate was centrifuged at 9,000 rpm to remove any particulate material (Fig. 5.12, lane 3).

A large DEAE Sepharose column (400 mL bed volume) was equilibrated with 5 mM Tris-HCl, pH 7.0, and the purification step was carried out at 4 °C. The 550 mL sample was loaded onto the column at 4 mL/min, and the column was washed overnight with 5 mM Tris-HCl, pH 7.0 to remove unbound proteins. A linear gradient from 0-800 mM NaCl in 5 mM Tris-HCl, pH 7.0, was run over 500 min, at a flow rate of 4 mL/min. Fractions (8 mL) were collected every 2 min.

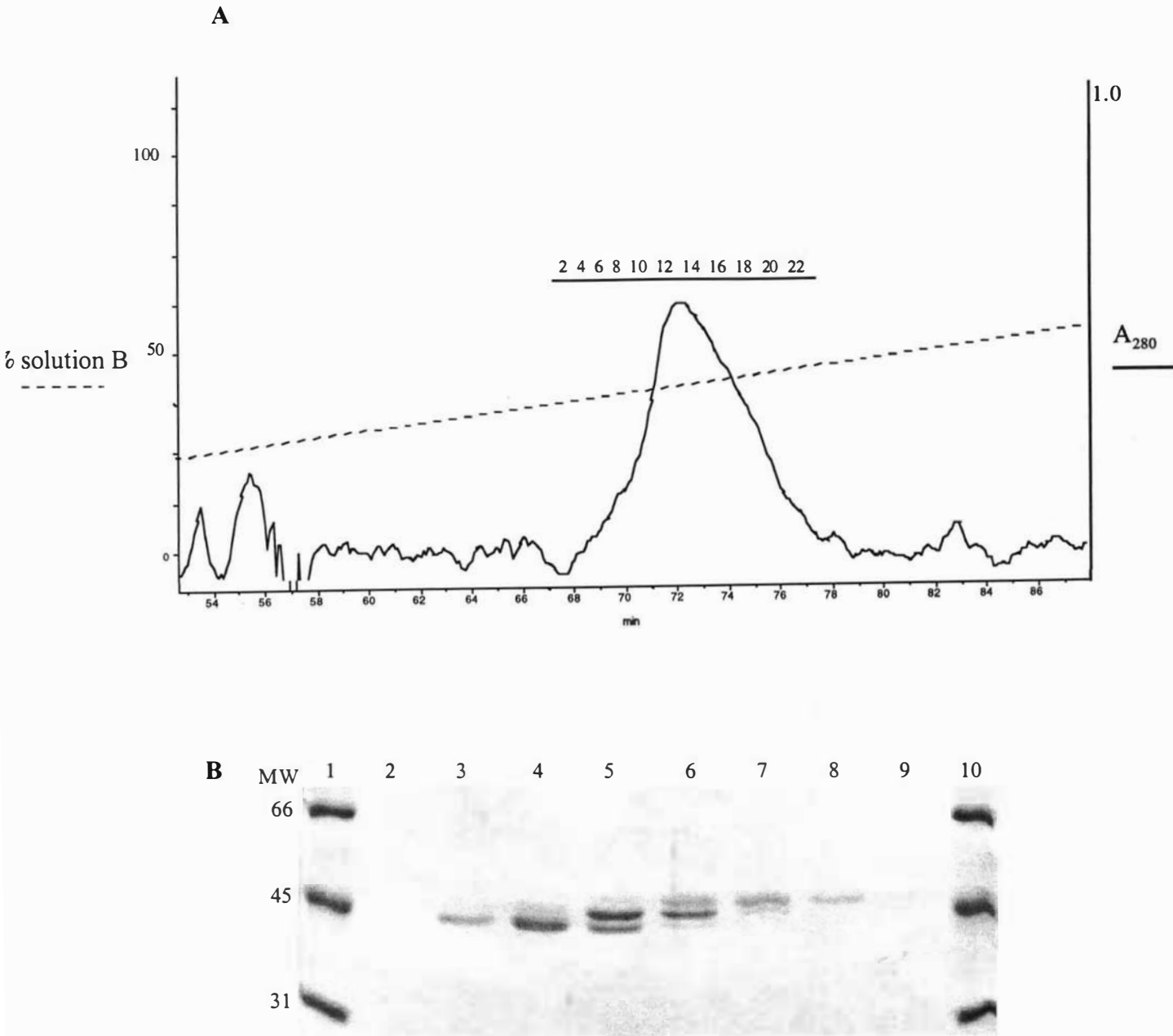


Figure 5.8 *DEAE Sepharose chromatography at pH 7.0*

A. Sap4 elution from DEAE sepharose

Fractions 1-22 were collected across the broad peak which eluted from 68-78 min. Separation of Sap4 species was analysed by SDS-PAGE (Fig. 5.8, B).

B. SDS-PAGE of Sap4-containing fractions from DEAE sepharose

Samples were run on 10% SDS-PAGE and visualised by staining with Coomassie Brilliant Blue™.

Lanes 1 & 10: Molecular weight markers

Lanes 2-9: 20 μ L sample from fractions 4, 6, 8, 10, 12, 14, 16 & 18, respectively

Each fraction was assayed for protein using the BCA method (Methods 2.10.9.2), and of those identified, every second fraction was then assayed for Sap activity (Fig. 5.9). Those fractions containing Sap activity were then analysed by SDS-PAGE (Fig. 5.10). For ease of purification, the fractions were pooled into four groups as follows:

- Pool 1.** Fractions 67-79; contained predominantly mature Sap4
Total volume 104 mL
- Pool 2.** Fractions 80-85; mostly mature Sap4, with some misprocessed forms
Total volume 48 mL
- Pool 3.** Fractions 86-92; contained predominantly misprocessed Sap4
Total volume 56 mL
- Pool 4.** Fractions 93-102; almost exclusively misprocessed Sap4
Total volume 104 mL

Pools 1 & 2 were used for subsequent Mono Q purification of mature Sap4. Pools 3 & 4 were judged to contain too little mature Sap4 for further fractionation (Fig. 5.12, lanes 4-7). Pools 1 & 2 were dialysed 30,000-fold against 5 mM Tris-HCl, pH 7.0 to remove residual salt from the DEAE Sepharose chromatography. Following dialysis, the pools were concentrated 6-fold in a Filtron™ ultrafiltration unit (10 K cut-off), and samples (2 mL) were run on the Mono Q column as described in section 5.3.4 (pool 1: 8 runs; pool 2: 4 runs). Sample h.p.l.c traces from these runs (Fig. 5.11) illustrate the increased abundance of misprocessed Sap4 in pool 2, compared to pool 1. The peaks corresponding to mature Sap4 from each run were also pooled and analysed by SDS-PAGE (Fig. 5.13A, 5.13B). The purified Sap4 was judged to be greater than 99% pure when stained with either Coomassie Brilliant Blue™ or reductive silver stain (Methods 2.10.5).

The purification (Table 5.2) shows a 2.6 fold purification and a recovery of 12 %. The low values reflect the exclusion of pools 3 & 4, and all misprocessed forms of Sap4 from pools 1 and 2. Judging the SDS-PAGE by eye (Figs. 5.10 & 5.12), mature Sap4 probably accounted for only 30-40% of the total recombinant Sap4. The specific activity of the purified mature Sap4 was two times higher than that of the combined pools 1-4. This suggests that either the misprocessed enzyme is less active than mature Sap4, or that pools 3-4 also contained other contaminating proteins.

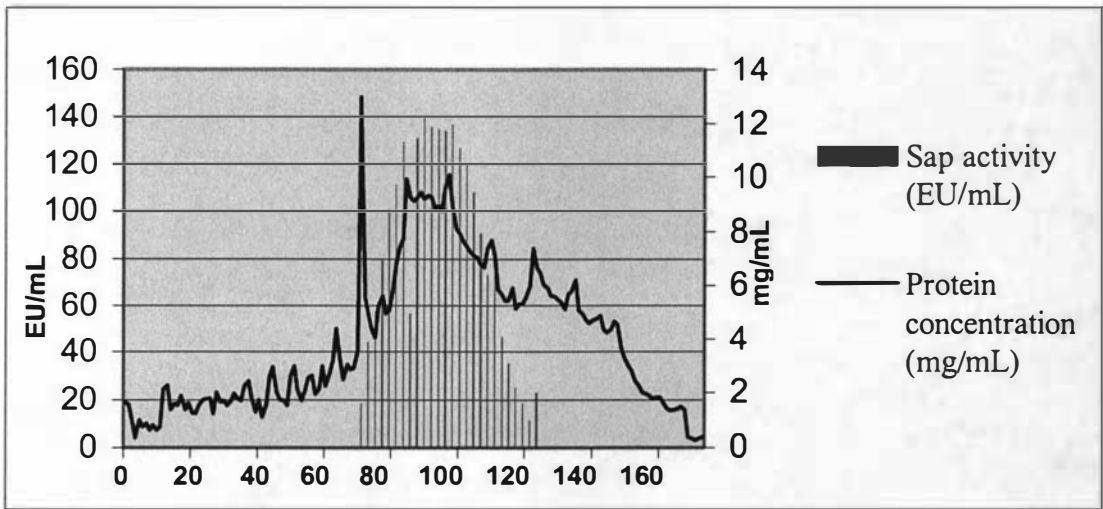


Figure 5.9 *Fractionation of a protein and Sap activity on a DEAE Sepharose column*

Protein concentration was measured using the BCA assay and is expressed in mg/mL. Sap activity is expressed in enzyme units (EU) per mL. Details of the column chromatography are described in Section 5.4.

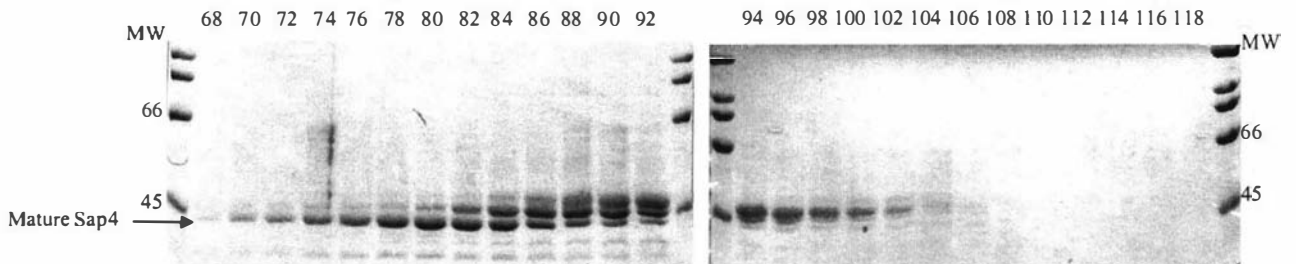


Figure 5.10 *SDS-PAGE analysis of fractions from DEAE Sepharose column*

10% SDS-PAGE was run with 20 μ L samples from each fraction, and visualised by staining with Coomassie Brilliant Blue™. Molecular weight markers were run on the outside lanes of each gel, and the fractions are labelled at the top of each lane.

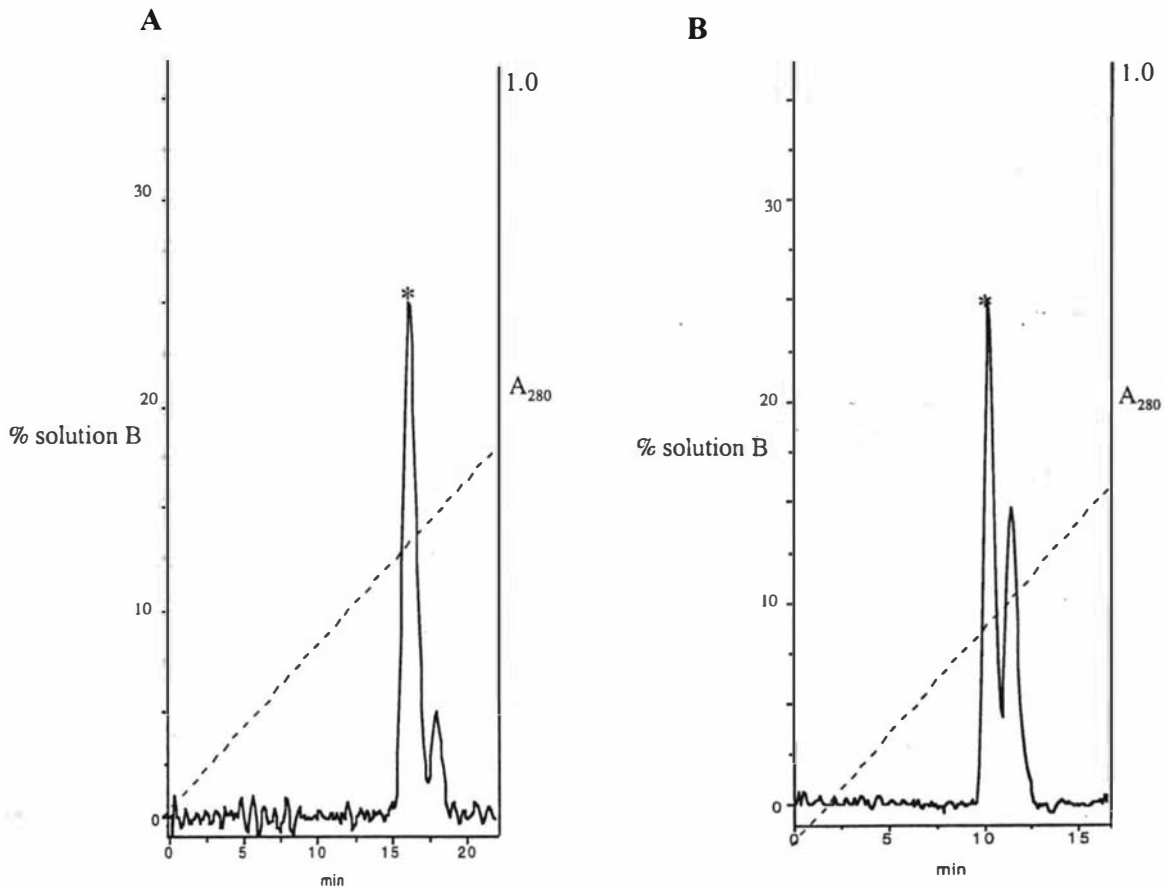


Figure 5.11 *h.p.l.c. traces of Mono Q purification of Sap4*

Mono Q profile from pool 1 (A) and pool 2 (B). 2 mL was injected onto the column in separate 1 mL injections, and eluted by a linear 30 min gradient (0-100 mM NaCl, refer section 5.3.4). Note the increase in peak height of the misprocessed Sap4 from pool 2. An asterisk labels the mature Sap4 peak.

Sample	Volume (mL)	Sap activity (Total EU)	Protein concentration (mg/mL)	Specific Activity (EU/mg)	Purification (fold)
6 L S/N	6,000	N/D	3.79	-	-
Concentrated S/N	550	16,500	0.33	90.9	1
Pool 1	104	3,500	0.14	240.4	2.6
Pool 2	48	3,354	0.34	205.5	2.3
Pool 3	56	4,856	0.62	139.9	(1.5)
Pool 4	104	4173	0.48	83.6	(0.9)
Pool 1-4 total	312	15,883	0.40	127.3	-
Purified Sap4	14	2,008	0.60	239.1	2.6

Table 5.2 *Purification of mature Sap4*

The table summarises the purification of Sap4. S/N is supernatant, N/D is not determined. Each pool is as described in section 5.4. The values in brackets represent the purification (fold) achieved in pools 3 & 4, which contained predominantly misprocessed Sap4.

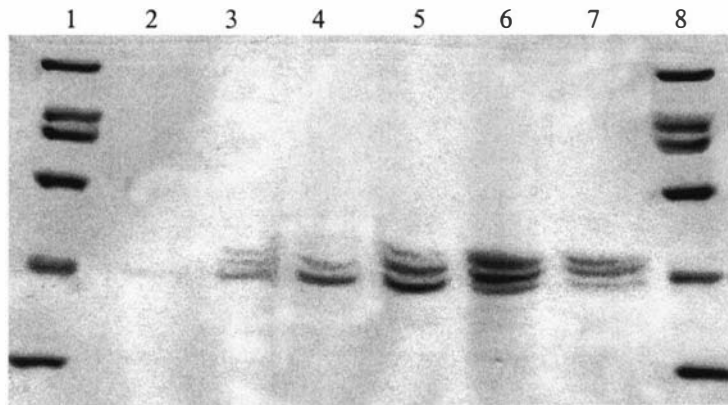


Figure 5.12 *Sap4* at each stage of purification

Samples from each stage of the purification protocol were run on 10% SDS-PAGE to determine relative purity.

- Lanes 1 & 8: Molecular weight markers
- Lane 2: 20 μ L crude 6 L culture supernatant
- Lane 3: 20 μ L concentrated buffer exchanged supernatant
- Lanes 4-7: 20 μ L each of Pool 1-Pool 4, respectively

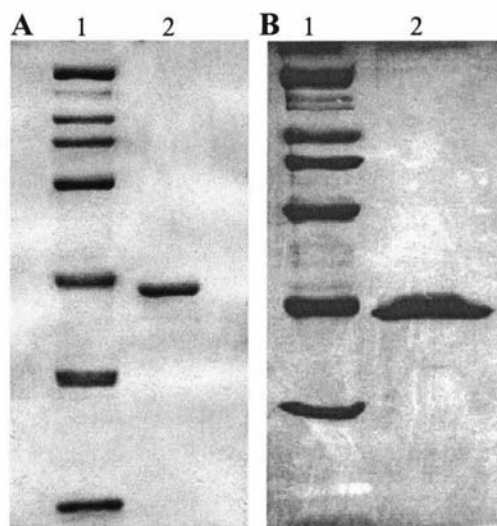


Figure 5.13 *Purified recombinant Sap4*

10% SDS-PAGE gels were run, and visualised with either Coomassie Brilliant Blue™ (A) or reductive silver stain (B). Lanes contained either molecular weight markers (Lane 1) or 3 μ g Sap4 (Lane 2).

5.5 Purification of recombinant Sap5

In the initial stages of the project, some Sap5 purification steps were attempted, as described below. However, the time required to scale-up and purify Sap4 meant there was limited time available to purify Sap5. Consequently, development of the Sap5 purification protocol was carried out by Ms. R. J. Sanders, based on the protocol developed for Sap4.

5.5.1 Anion exchange chromatography

The Mono Q anion exchange chromatography protocol used to purify Sap1 (section 5.2), was used to trial a fractionation of the supernatant of an induced MS5A culture (data not shown). Sap5-containing fractions were identified by SDS-PAGE and the gels revealed that Sap5 was present in both the unbound eluate and fractions which were eluted at 100 mM - 150 mM NaCl. The distribution between unbound and bound Sap5 was puzzling. A significant percentage (40 %) of Sap5 was bound, and this distribution was not affected by varying the amount of protein loaded to the column. This suggested there was either a limiting factor which mediated the interaction, or different populations of Sap5 which all migrated with the same apparent molecular mass on SDS-PAGE.

5.5.2 Cation exchange chromatography

Cation exchange chromatography was used in an attempt to distinguish between the two possibilities. The conditions were as described in section 5.3.2. No Sap5 was recovered in either the unbound eluate, or from the fractions across the gradient (data not shown). Therefore cation exchange was not pursued any further as a potential purification technique.

5.5.3 Gel filtration of Sap5

Unlike MS4A, induced MS5A cultures appeared to produce only one form of Sap5. Therefore gel filtration was a possible purification method. A sample of the culture supernatant was concentrated 50-fold, buffer-exchanged with 10 volumes of ammonium acetate buffer, pH 6.0 then 100 μ L of the sample was injected onto a 26 mL bed volume Superose 12 column (26 mL bed volume). The column was developed at 0.3 mL/min, in 100 mM ammonium acetate, pH 6.0 containing 100 mM NaCl (to prevent interaction of the protein sample with the column). The retention time of the Sap5 on the column was equivalent to a protein much smaller than lysozyme (14 kDa), which suggested significant interactions were occurring (data not shown). Unfortunately this late elution did not enhance the purity of the Sap5 sample and the elution profile was not reproducible. Molecular weight standards were run to check the integrity of the column and all eluted as expected. As a result of these observations, gel filtration was not continued.

5.6 Sap5 purification protocol

The final purification protocol developed by Ms. Rebecca Sanders involved buffer-exchanging the culture medium into 5 mM sodium acetate, pH 5.0, and DEAE anion chromatography (5 mM sodium acetate, pH 5.0) where Sap5 elutes in the unbound fraction. This was followed by a Resource S chromatography step (Pharmacia, 1 mL bed volume). The protein was equilibrated in 5 mM sodium acetate pH 5.0, applied to the column, then eluted with a linear gradient from 0-600 mM sodium chloride over 20 min. The Sap5 eluted in a single peak approximately half way through the gradient. Purification of Sap5 is described in greater detail in Appendix 10.3.

6.0 BIOCHEMICAL CHARACTERISATION OF RECOMBINANT Sap ISOENZYMES

6.1 Introduction

The evidence implicating Saps in the onset and development of candidoses is considerable (for review see Hoegl *et al.*, 1996). With the discovery that the Sap gene family has at least nine members, it has become important to determine the biochemical properties of each isoenzyme, and to elucidate the roles they play in the disease process. Then, if the Saps have been proven to be a suitable drug target, it would be possible to apply a rational approach to designing inhibitors which could be used as drugs.

At the outset of this project, the proteins encoded by *SAP4*, *SAP5* and *SAP6* had not been detected in *C. albicans* cultures. For this reason a recombinant over-expression strategy was chosen, to produce the quantities of the enzymes required to determine their biochemical properties. However, during the course of this work, Borg-von Zepelin *et al.* (1998) reported the expression of Saps 1-6 in *P. pastoris*, as a prelude to the expression studies in murine macrophages. While they did not extensively characterise the properties of the recombinant Sap proteins, some biochemical studies were performed and the results are discussed as appropriate throughout this chapter.

6.2 Sap enzyme activity assays

6.2.1 Sap4 substrate preference

In the initial stages of this project, recombinant Sap4 and Sap5 were thought to be inactive. However, it was subsequently discovered that they do not hydrolyse the usual assay substrate, BSA as rapidly as Sap2. Therefore, a comparison of Sap4 activity towards four different substrates was performed to identify a substrate appropriate for routine assays. The results (Fig. 6.1) showed that globin was the best substrate. This may, in part, be due to the acid-denaturation step used to prepare the globin solution (Method 2.10.9.2), as this could expose many susceptible bonds compared with a folded, globular protein. No activity was detected with lactalbumin, and activity towards BSA and casein was 4-5 fold lower than that observed with globin.

In addition to these substrates, a qualitative assay was performed using azocollagen, an insoluble collagen coupled to a red azo-dye which is solubilised upon hydrolysis of the protein (see Methods 2.10.9.1). Unlike Saps 1-3, no activity was detected with Sap4, even after 24 h incubation.

6.2.2 Enzyme kinetics

6.2.2.1 Globin assay

Having established that globin was the best protein substrate for Sap4, the assay was validated with this enzyme, prior to kinetic studies. Fig. 6.2 shows that for Sap1 and Sap4, enzyme activity was linear over the first 60 min of the assay and no Sap5 activity was detected with the crude enzyme preparation. These data showed that activity was linear for the first 60 min of the assay. A linear plot of velocity versus amount of enzyme was obtained using a 30 min incubation (data not shown).

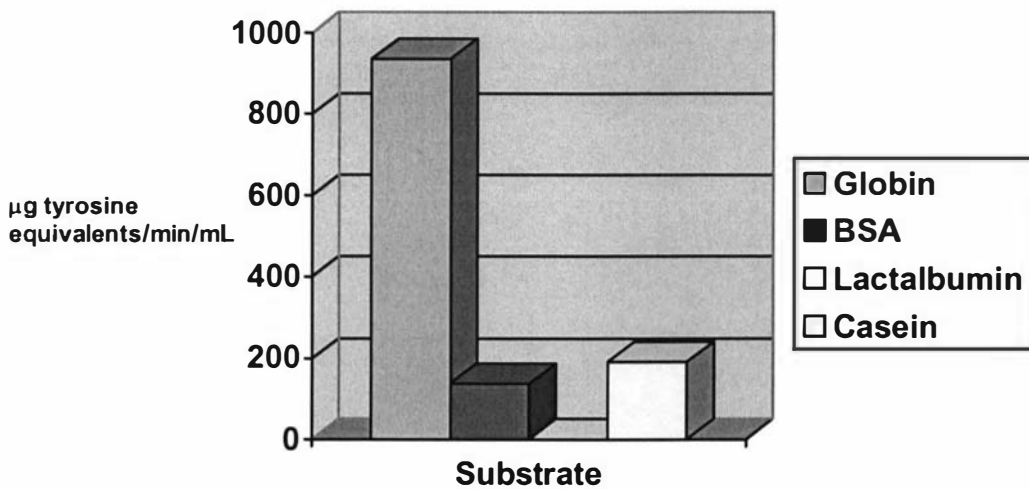


Figure 6.1 *Relative Sap4 activity toward various protein substrates*

Enzyme activity was determined using purified Sap4. Assays were performed as described in Methods 2.10.9.2 using a 1.25 mM solution of each protein. Activity is expressed as: µg tyrosine equivalents released/min/mL enzyme sample.

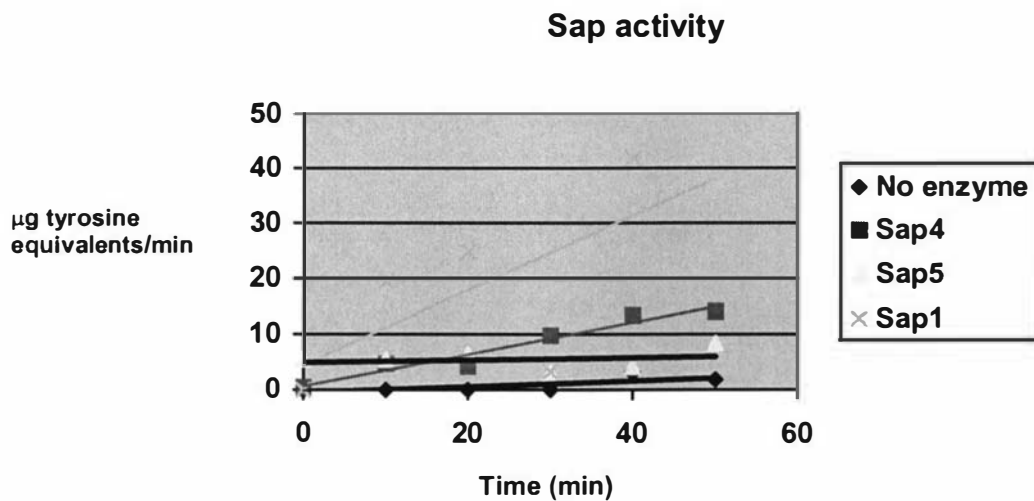


Figure 6.2 *Time course of Sap activity in the globin assay*

Assays were performed with a fixed amount of enzyme, and activity was measured at the times indicated. Activity was calculated as: µg tyrosine equivalents released/mL.

Fig 6.3 shows the effect of the substrate concentrations (globin) on the reaction velocity. A *Hanes* plot (s/v vs. s) was drawn with these data (Fig. 6.4), as the commonly-used *Lineweaver-Burk* plot gives an unbalanced model of experimental error because most weighting is placed on the least accurately measured values, i.e. those at low substrate concentrations (Cornish-Bowden, 1995). The linear *Hanes* plot yielded values of: -0.52 mM for $-K_m$ and 0.75 for $1/v_{max}$. Thus the K_m for globin was approximately $500 \mu\text{M}$. However, there are certainly many cleavage sites in globin upon which the enzyme acts, and the K_m for each of these sites would vary. Further, many of the bonds hydrolysed may not immediately result in the formation of a soluble peptide fragment. Therefore, the K_m constants derived from these data are somewhat arbitrary, as they represent the average K_m of many different hydrolysis reactions occurring simultaneously in the polypeptide substrate. The *Hanes* plot also allows the V_{max} value to be extrapolated, but in this case, the values used for v do not represent the true velocity, for the reasons outlined above, and therefore the V_{max} value has not been calculated. Although the kinetic constants are not informative, this analysis showed that Sap4 exhibited classical saturation kinetics with globin.

6.2.2.2 Fluorocasein assay

A commercially-available fluorocasein substrate (Molecular Probes EnzChek™ Protease Assay Kit E-6638) was also used to measure Sap activity. The fluorocasein substrate contains covalently bound BOPIDY-FL™, which remains quenched in the intact substrate, but fluoresces following enzyme hydrolysis. The sensitivity of this fluorometric microtitre-plate assay was 10-fold greater than the globin-based Sap assay. This meant that the fluorocasein assay could be used, despite casein not being the best substrate (refer Fig. 6.1). Preliminary experiments were carried out to validate the fluorocasein assay with respect to time and the Sap4 concentration. The assay involved collecting fluorescence data at four or more time points, and the increase in fluorescence was shown to be directly proportional to incubation time over the first 6 min of the assay (data not shown). Therefore, these data were used to calculate the reaction rates using linear regression analysis. Fig. 6.5 shows the linear relationship between reaction

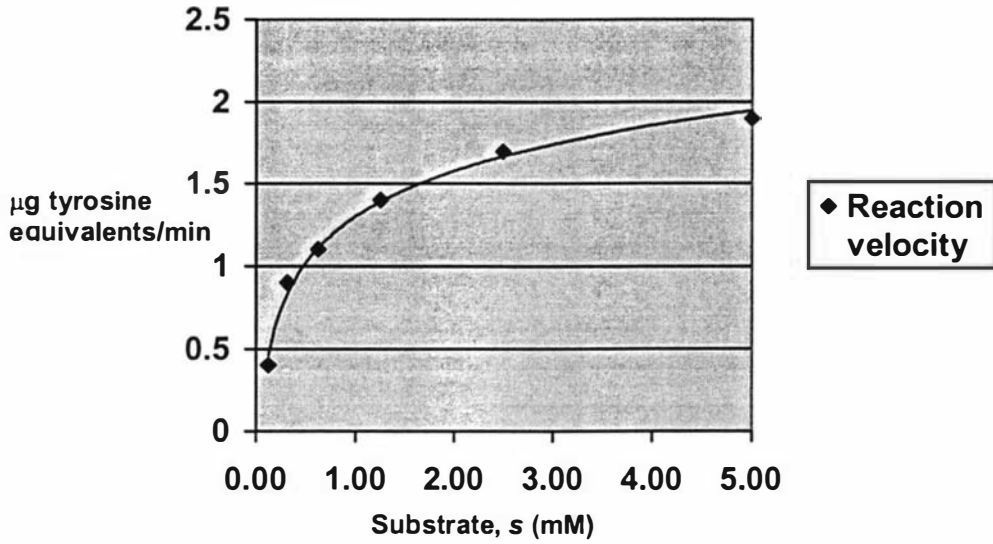


Figure 6.3 Plot of reaction velocity vs. substrate concentration

The globin assay was performed using Sap4 and various substrate concentrations (0.2%, 0.5%, 1.0%, 2.0%, 4.0% & 8.0%) and the activity was measured at five different time points (5 min, 10 min, 15 min, 20 min, 25 min and 30 min) to determine the reaction velocity. Each assay contained 2.0 µg of pure Sap4, specific activity 239 µg tyrosine equivalents/min/mg.

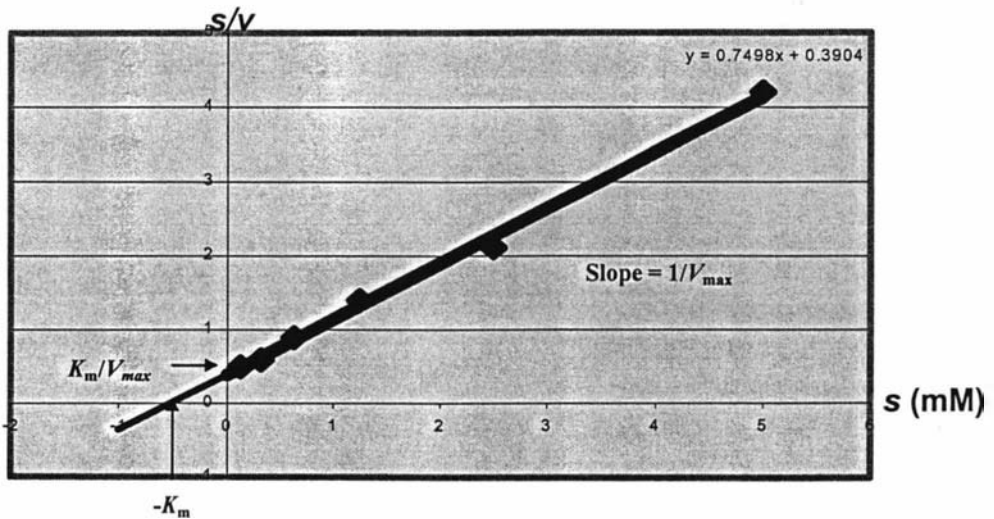


Figure 6.4 Hanes plot for Sap4 activity using the globin assay

A Hanes plot (s/V vs. s) of the data presented in Fig. 6.3. The value of the y intercept is K_m/v , the x intercept is $-K_m$, and the slope of the line is $1/V_{max}$.

rate, expressed as change in fluorescence per min, and the amount of enzyme used in the assay (0.08-2.5 μ g Sap4).

6.3 pH optima of the Sap isoenzymes

Activity assays were performed over a range of pH values, to compare the pH optima and pH activity curves of the recombinant Sap1, Sap4 and Sap5 with native Sap2 and Sap3. Activity was measured at pH values from 2.0–7.0, using 200 mM glycine-HCl (pH 2.0-pH 3.0), 200 mM sodium citrate (pH 3.0-pH 6.0), 200 mM sodium acetate (pH 3.5-pH 5.5), 200 mM sodium phosphate (pH 6.0-pH 7.0), and 200 mM sodium citrate-sodium phosphate (pH 2.5-pH 7.0) buffers. The results (Fig. 6.6) revealed that Sap4 and Sap5 have much higher pH optima (both pH 4.5) than Saps 1, 2 and 3 (pH range 3.2-3.5). In addition, Sap4 and Sap5 still retained greater than 50% of the maximum activity above pH 6.0, in contrast with Sap1 which has less than 20% activity above pH 6.0 and Saps 2 & 3 which showed less than 10% optimum activity above pH 5.0. These results are consistent with the recently reported pH optima of 5.0 for Saps 4 and 5 (Borg-von Zepelin *et al.*, 1998).

Transcription of Sap4 and Sap5 has been shown to be dependent upon neutral pH conditions (Hube *et al.*, 1994; White and Agabian, 1995), and the discovery of Sap activity at pH 7.0 provides the first evidence that Saps might play a role in the infection process at sites of neutral pH. In contrast, the near-neutral pH of most body fluids and the low activity of Saps 1, 2 and 3 above pH 5.0 would preclude involvement of these isoenzymes at most physiological sites, with the exception of some specific sites such as skin, and areas of inflammation. Low culture pH has previously been correlated with increased Sap expression (White & Agabian, 1995; Hube *et al.*, 1994), and the propensity of *C. albicans* to decrease the localised pH by secretion of acidic metabolic by-products has been suggested as an explanation of Saps in pathogenicity (Ruchel *et al.*, 1992). This suggests that Saps might play an auxiliary role in pathogenesis at sites of neutral pH, contributing to the disease process once the infection is established and the local pH decreased. However, possession of Sap isoenzymes, which are functional at pH 7.0, would eliminate the need for a localised decrease in pH prior to expression of Sap activity (Samaranayake *et al.*, 1984).

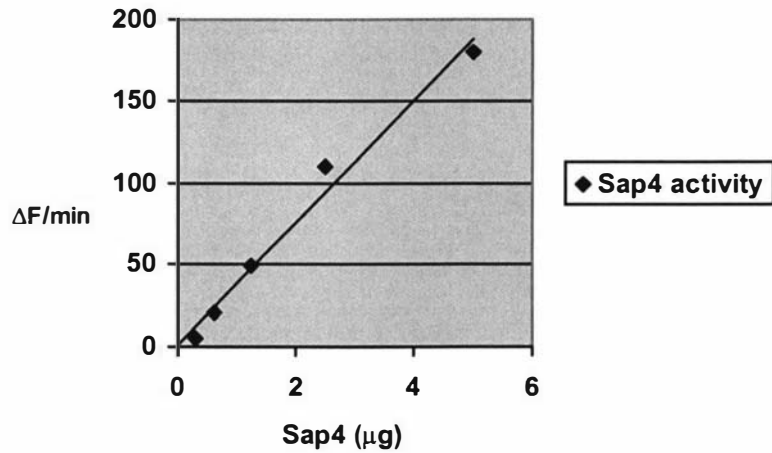


Figure 6.5 *The fluorocasein Sap assay*

Sap4 was assayed with a fluorocasein as substrate (see Methods 2.10.9.4), and the reaction rates were expressed as Δ fluorescence/min.

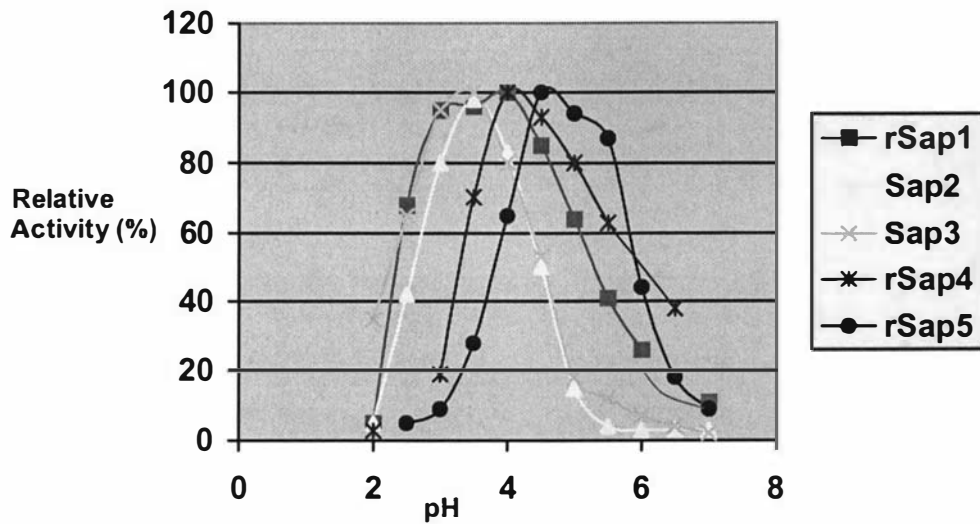


Figure 6.6 *pH activity profiles of the Sap isoenzymes*

Assays were performed at pH values between 2.0 and 7.0 using the globin activity assay. Points shown are the averages of three data sets. The amount of each isoenzyme was constant across the different pH values, but varied between assays. Sap1: 2.5 µg, Sap2: 1.8 µg, Sap3: 1.6 µg, Sap4: 2 µg, Sap5: 10 µg.

pH 2.0 - pH 3.0: glycine-HCl

pH 3.0-pH 6.0: sodium citrate

pH 3.5-pH 5.5: sodium acetate

pH 6.0-pH 7.0: sodium phosphate

pH 2.5-pH 7.0: sodium citrate-sodium phosphate

Interestingly, the pH optimum profile of Sap1 (Fig. 6.6) was similar to that reported recently (Borg-von Zepelin *et al.*, 1998), but differed from the profile determined by Smolenski *et al.* (1997). The only difference between the recombinant protein used for these profiles, was that in the latter case, the recombinant Sap1 was expressed from *S. cerevisiae*, rather than *P. pastoris*. This is discussed further detail in section 6.9.

6.4 Specific activity of the Sap isoenzymes

The specific activities for the Sap isoenzymes were determined using the globin and the fluorocasein assays (Methods 2.10.9), and the results are summarised in Fig. 6.7. Recombinant Sap1 was 3.4 times more active than native Sap2 at pH 4.5 in the globin assay and 3.8 times more active in the fluorocasein assay, although the activity levels were similar at pH 3.6 in the fluorocasein assay. Sap4 showed low activity at pH 3.6 in both assays, and at pH 4.5 using the globin assay. However, considerably higher activity was detected with fluorocasein at pH 4.5. Similarly, Sap5 showed little activity in either assay at pH 3.6, but high activity was detected in the fluorocasein assay at pH 4.5. It was also noted that the relative activities of Sap4 and Sap5 were quite different with the two substrates.

Fig. 6.7A and 6.7B show that Saps 1 and 2 have higher activity than Saps 4 and 5 with globin, but the reverse is true with fluorocasein. Interestingly, underivited casein was a relatively poor substrate for Sap4 and Sap5 compared with globin (Fig. 6.1). This suggests that fluorocasein may have a few substrate sites which are rapidly hydrolysed by Sap4 and Sap5, but not by Saps 1 and 2. In contrast, the absence of such a preferred site might account for the low activity of Sap4 and Sap5 in the globin assay.

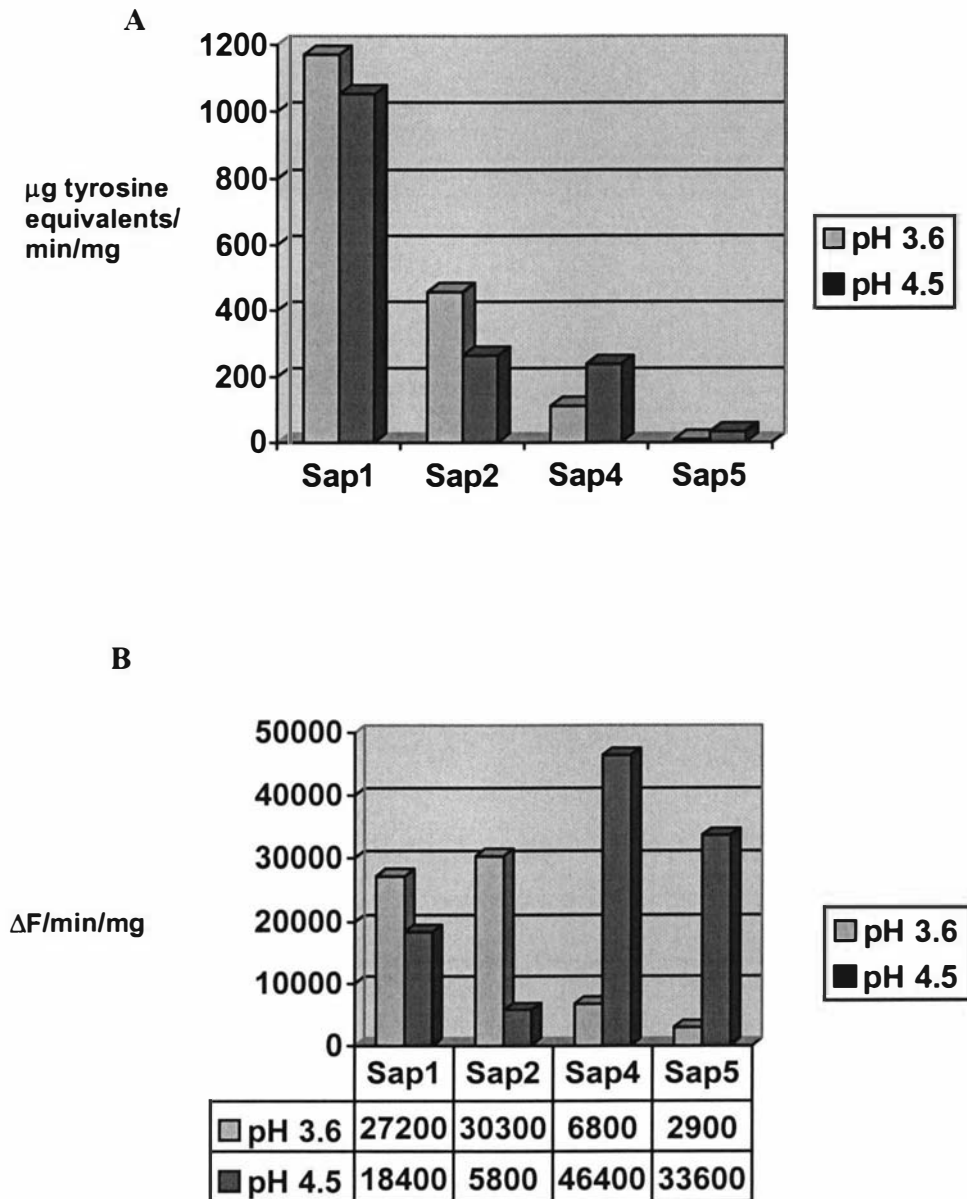


Figure 6.7 Specific activities of Sap isoenzymes in the globin (A) and fluorocasein (B) assays

- A. Specific activities of recombinant Saps 1, 4 and 5 compared with the specific activity of native Sap2 at pH 3.6 and pH 4.5. The assay conditions were as described in Methods 2.10.9.3.
- B. Specific activities of recombinant Saps 1, 4 and 5 compared with the specific activity of native Saps 2 & 3 at pH 3.6 and pH 4.5. The assay conditions were as described in Methods 2.10.9.4.

During the course of this work, Borg-von Zepelin *et al.* (1998) reported expression of Saps 1-6 in *P. pastoris* (strain GS115) using the native *C. albicans* Sap pro-sequences, and the *P. pastoris* α -factor signal sequence (from the pHILS1 expression vector). Saps 1, 2 & 3 were purified using DEAE, and Saps 4, 5 & 6 were purified using hydroxyapatite. It was reported that the specific activities of the six recombinant Saps was similar to each other and to Sap2 isolated from *C. albicans* grown with BSA as sole nitrogen source (data not shown). The proteinase substrate used by Borg-von Zepelin *et al.* (1998) was resorufin-labelled casein. A similar summary could be made for the strikingly different results for the specific activities and pH optima of Saps 1, 2, 3 & 4 with fluorocasein and globin assays.

The data presented in Fig. 6.7 indicated that the relative activities of the recombinant Sap isoenzymes was not the same as that of native Sap2. In addition, the variation between specific activities of Sap4 and Sap5 in the two different assays, and the lack of activity towards azocollagen (section 6.3.1) suggested narrower substrate specificities in the binding sites, in contrast to the broad range specificities of Saps 1, 2 and 3.

6.5 Substrate specificity of Sap

To investigate the substrate specificities of Sap4 and Sap5 further, glucagon (a 29 amino acid peptide) was chosen as a substrate for limited hydrolysis. This involved incubation of either Sap4 or Sap5 with glucagon at 37 °C, with samples being taken at various times. The peptide fragments were fractionated by reversed phase HPLC, and identified using MALDI-TOF mass spectrometry.

Following a protocol developed by Hermann (1996, described in Method 2.10.10), reactions were performed using two different concentrations of Sap, to follow early and late hydrolysis events. Low concentrations of Sap (0.0164 EU) were incubated with glucagon and samples were taken at time zero, 2 min, 30 min and 180 min. High concentrations of Sap (0.328 EU) were also incubated with glucagon and samples were taken at time zero, 2 min, 15 min and 60 min. All samples taken were immediately frozen in liquid air and stored at -80 °C prior to fractionation on reversed phase HPLC. The reversed phase HPLC was performed using a Resource RPC column (Pharmacia

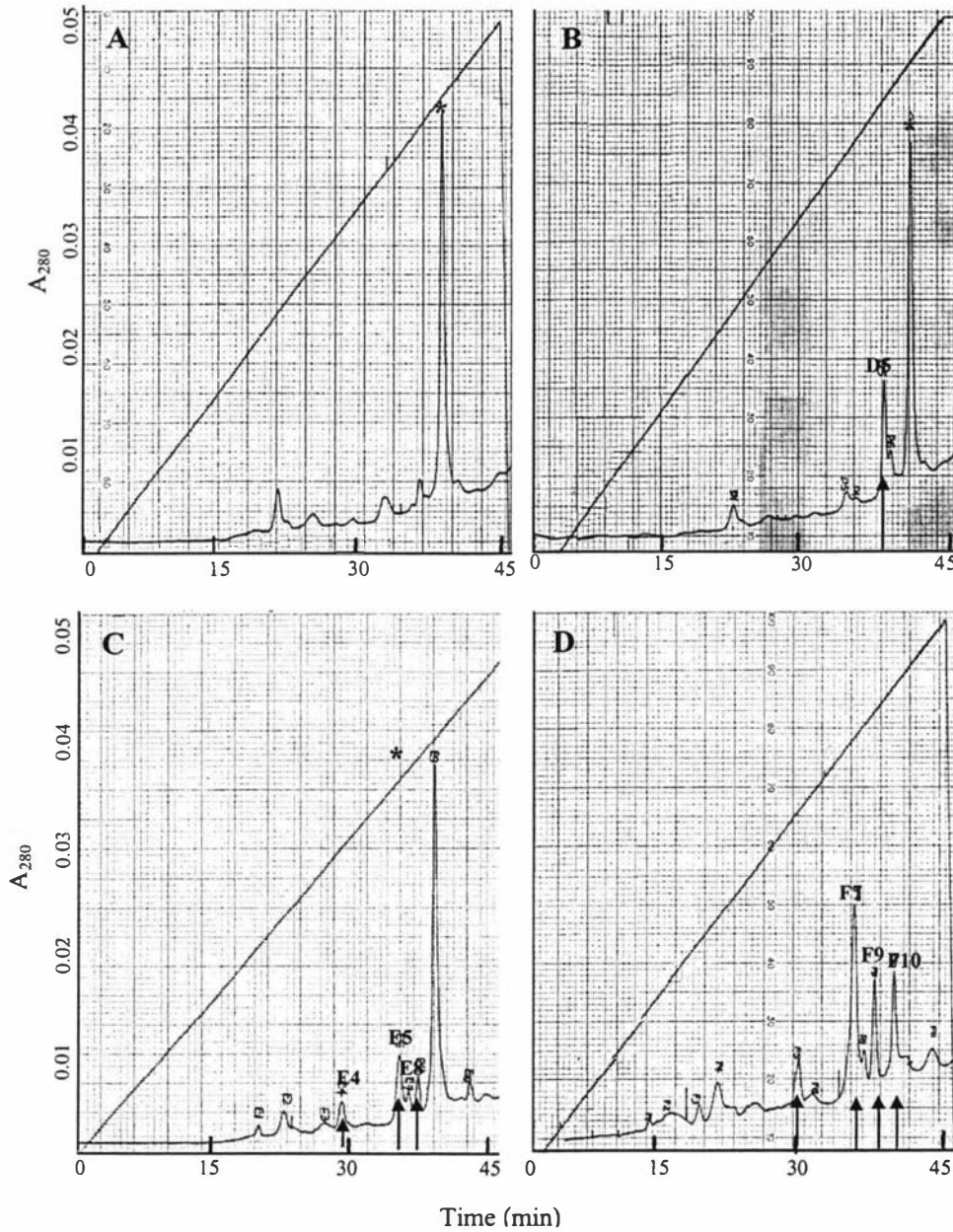


Figure 6.8 *Reversed phase HPLC of glucagon hydrolysates with Sap4*

The series of chromatograms shows the hydrolysis of glucagon over time. These traces were used to deduce the order of appearance of the peptide fragments generated by Sap4, which were subsequently identified by ES-MS. Intact glucagon is labelled with an asterisk and arrows indicate the appearance of hydrolysis products. The peaks were detected by monitoring the eluate at 280 nm, and the column conditions are described in Methods 2.10.10.

- A. Glucagon zero time control
- B. 2 min hydrolysis of glucagon with 0.328 units Sap4
- C. 30 min hydrolysis of glucagon with 0.328 units Sap4
- D. 180 min hydrolysis of glucagon with 0.328 units Sap4

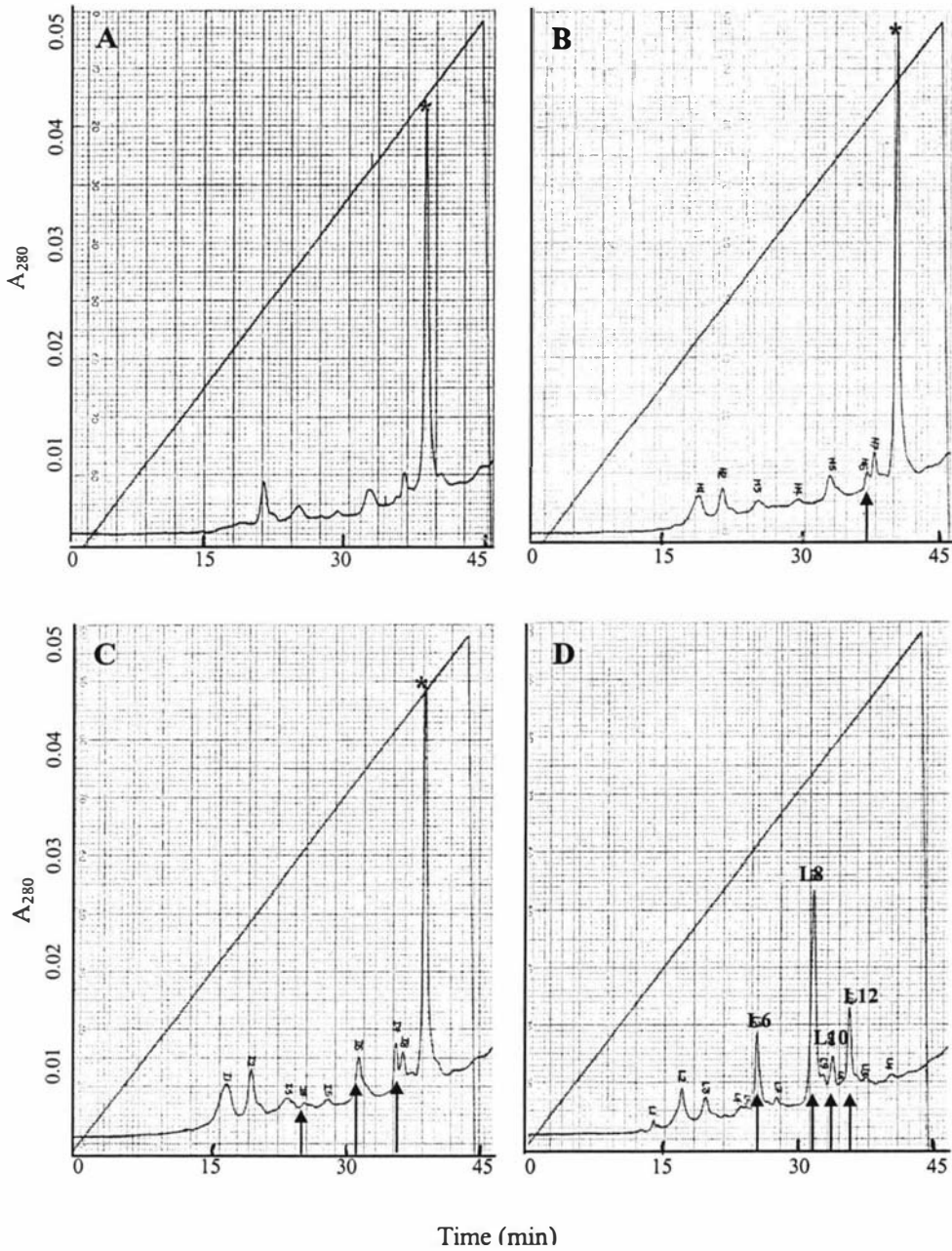


Figure 6.9 *Reversed phase chromatograms of glucagon hydrolysis*

The series of chromatograms shows the hydrolysis of glucagon over time. These traces were used to deduce the order of appearance of the peptide fragments generated by Sap5, which were subsequently identified by ES-MS. Intact glucagon is labelled with an asterisk and arrows indicate the appearance of hydrolysis products. The peaks were detected by monitoring the eluate at 280 nm, and the column conditions are described in Methods 2.10.10.

- A. Glucagon zero time control
- B. 15 min hydrolysis of glucagon with 0.0164 units Sap5
- C. 60 min hydrolysis of glucagon with 0.0164 units Sap5
- D. 180 min hydrolysis of glucagon with 0.328 units Sap5

Biotech, 1 mL bed volume) and monitored at 280 nm. Attempts to monitor the column at 214 nm (which detects peptide bonds) were not successful, due to interference from the acetonitrile solvent. The presence of five fairly evenly spaced aromatic residues throughout the glucagon meant that while some peptide fragments might remain undetected at 280 nm, most should have been detected, and those which were not would be very small peptide fragments.

The reversed-phase h.p.l.c traces of the samples taken at various times (Fig. 6.8 and Fig. 6.9) summarise the progress of the Sap4 and Sap5 hydrolysis reactions. Peaks on the chromatograms were collected and given unique numbers as shown (Fig 6.8, Fig. 6.9). The corresponding MALDI-TOF mass spectra for these peaks are included in Appendix 10.5.

The identity of the peaks was assigned using glucmass, a program written specifically for this purpose by V. Hermann and kindly provided for this work. The mass spectrometry data summarised in Table 6.1, were used to identify the glucagon peptides produced by Sap4 and Sap5. The data in Table 6.1 suggests that Sap4 and Sap5 have the same or very similar substrate specificities, given they both generated the same pattern of peptides from the hydrolysis of glucagon. A model of the pattern of hydrolysis at the major cut sites is shown in Fig. 6.16. An analogous model showing the specificity of Saps 1-3 (kindly provided by Ms. V. Hermann) is also presented. Comparison of these models suggests that the specificity of Sap4 and Sap5 is quite distinct from that of Saps 1, 2 and 3. Remarkably, there appeared to be no consensus between any of the Sap4 and Sap5 cleavage sites based on either amino acid residue size, or charge (Table 6.2). In contrast, Saps 1-3 show a distinct preference for particular types of amino acids at certain positions (Table 6.2).

While there was no apparent consensus specificity for Sap4 and Sap5 hydrolysis, these enzymes cut two sites which are also cleaved by Saps1-3. The cut site between the sequence YSK-YLD was common to all enzymes, but was preferentially cleaved by Saps 1-3 (2nd major cut site), rather than Saps 4 and 5 (4th major cut site). The Sap4 and Sap5 minor cleavage products (listed in italicised type in Table 6.1) indicated some hydrolysis at the sequence DSR-RAQ. This sequence was also cleaved by Sap3

			Intact Glucagon: HSQGTFTSDYSKYLD SRRA QDFVQWLMNT
Enzyme	Peak	Mass (Da)	Identity
Sap4	D5	3136.9	HSQGTFTSDYSKYLD SRRA QDFVQWL---
	E4	2465.2	-----TSDYSKYLD SRRA QDFVQWL---
	E5	1322.7	-----RRAQDFVQWL---
		<i>1409.8</i>	-----SRRAQDFVQWL---
		<i>1168.3</i>	-----RAQDFVQWL---
F9	1800.0	-----YLDSRRAQDFVQWL---	
Sap5	L12/F10	3136.9	HSQGTFTSDYSKYLD SRRA QDFVQWL---
	L6	2462.8	-----TSDYSKYLD SRRA QDFVQWL---
	L8	1319.1	-----RRAQDFVQWL---
		<i>1164.8</i>	-----RAQDFVQWL---
L10	1800.3	-----YLDSRRAQDFVQWL---	

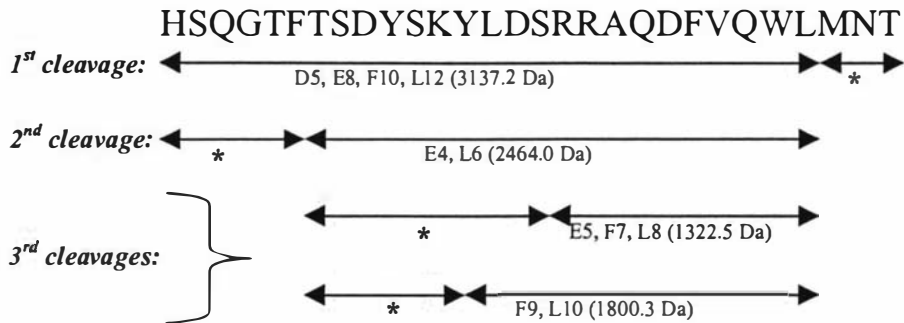
Table 6.1 *Glucagon fragments produced by Sap4 and Sap5*

The table summarises the experimentally determined masses, and the peaks from which the data were obtained. The major pattern of substrate hydrolysis appeared to be the same for both enzymes. Italicised data represent minor peptide fragments.

(Hermann, 1996), although it did not conform to the deduced consensus amino acid sequence for Saps1-3 cut sites.

The lack of obvious consensus for Sap4 and Sap5 cut sites prompted a review of the amino acid residues that make up the S₁ and S₁' active site subsites. As shown in Table 6.3, Sap1 and Sap2 have identical residues in both the S₁ and S₁' subsites, and Sap3 has

A. Sap4 and Sap5 model:



B. Sap1, Sap2 and Sap3 model:

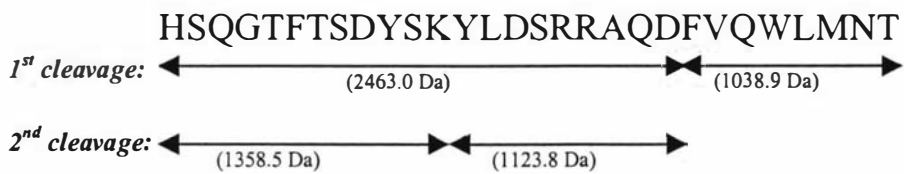


Figure 6.10 Model of cleavage of glucagon by Saps 1-5

- A. The model represents the experimental data derived from hydrolysis of glucagon. Each of the fragments are defined by the arrows, in order of temporal appearance, from top to bottom. The peptides are labelled with the peaks in which they were eluted, and the experimentally determined fragment masses of each peak are shown in parentheses. An asterisk indicates fragments for which no experimental data was obtained.
- B. An analogous model (Hermann, 1996) showing the first two major cut sites of Saps 1-3. The fragment sizes were determined experimentally using MALDI-TOF mass spectrometry and are shown in parentheses.

Sap1, Sap2 and Sap3 cut sites			Saps1-3 Consensus
	1 st cut site	2 nd cut site	
P4	R	D	Hydrophilic
P3	A	Y	Hydrophobic
P2	Q	S	Any amino acid
P1	D	K	Hydrophilic
P1'	F	Y	Aromatic
P2'	V	L	Hydrophobic
P3'	Q	D	Hydrophilic

Sap4 and Sap5 cut sites					Saps 4 & 5 Consensus
	1 st cut site	2 nd cut site	3 rd cut site	3 rd cut site	
P4	V	Q	Y	D	-
P3	Q	G	L	Y	-
P2	W	T	D	S	-
P1	L	F	S	K	-
P1'	M	T	R	Y	-
P2'	N	S	R	L	-
P3'	T	D	A	D	-

Table 6.2 *Consensus sequences of Sap1-5 cleavage sites in glucagon*

The table shows the Sap1-3 consensus amino acid sequences of the major cut sites in glucagon (top). The amino acid sequences in glucagon which are cleaved by Sap4 and Sap5 are also shown. No consensus could be derived from these amino acid sequences.

Enzyme	Subsite S ₁ definition	Subsite S ₁ ' definition
Sap1	Ile30, Asp32, Gly34, Tyr84, Asp86, Ile119, Ile123, Asp218, Gly220, Thr221	Gly34, Tyr84, Gly85, Glu193, Leu216, Asp218, Ile305
Sap2	Ile30, Asp32, Gly34, Tyr84, Asp86, Ile119, Ile123, Asp218, Gly220, Thr221	Gly34, Tyr84, Gly85, Glu193, Leu216, Asp218, Ile305
Sap3	Val 30, Asp32, Gly34, Tyr84, Asp86, Val 119, Ile123, Asp218, Gly220, Thr221	Gly34, Tyr84, Gly85, Glu193, Leu216, Asp218, Ile305
Sap4	Ile30, Asp32, Gly34, Tyr84, Asp86, Ala 119, Ile123, Asp218, Gly220, Thr221	Gly34, Tyr84, Ala 85, Thr 193, Leu216, Asp218, Ile305
Sap5	Ile30, Asp32, Gly34, Tyr84, Asp86, Ala 119, Ile123, Asp218, Gly220, Thr221	Gly34, Tyr84, Gly85, Lys 193, Leu216, Asp218, Ile305
Sap6	Ile30, Asp32, Gly34, Tyr84, Asp86, Ala 119, Ile123, Asp218, Gly220, Thr221	Gly34, Tyr84, Ala 85, Thr 193, Leu216, Asp218, Ile305

Table 6.3 *Amino acid side chains which form subsites S₁ and S₁'*

Residues in the subsites of the Sap enzymes were identified from the alignment of the deduced amino acid sequences with that of Sap2, and numbered according to the Sap2 sequence.

only two changes in the S₁ subsite from this consensus. Sap4 and Sap6 each have only one amino acid substitution in the S₁ subsite and two substitutions in the S₁' subsite while Sap5 had one substitution in each site. These few amino acid changes did not suggest a basis for the observed differences in the substrate specificities of Sap4 and Sap5. However, the basis for the differences may become clear once a structural model can be calculated using x-ray data, or by homology modelling.

It is possible that some peptide fragments which do not absorb at 280 nm have been overlooked during this analysis. While the peptide fragments which were identified at 280 nm provide evidence for the cleavage pattern described above, the possibility exists that there are other cleavage products which were not identified. However, the largest possible fragment which would not be identified is LDSRRAQD, which is an internal fragment requiring glucagon to be cleaved twice, and therefore is not the first cleavage product.

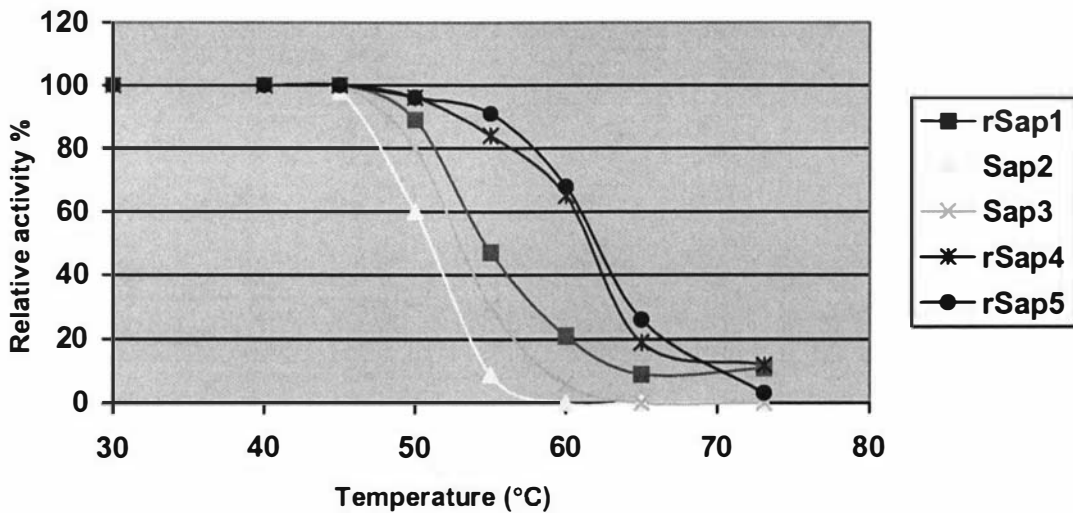


Figure 6.11 Thermal stability profiles of the Sap isoenzymes

The graph shows the relative activities of the Sap isoenzymes at increasing temperatures. Each isoenzyme (Sap1; 2.5 μ g, Sap2; 1.8 μ g, Sap3; 1.6 μ g, Sap4; 2 μ g; Sap5 10 μ g) was incubated with 200 mM buffer (Saps 1-3, glycine, pH 3.2; Saps 4-5, acetate, pH 4.5), heated for 15 minutes at the temperatures indicated, cooled on ice and immediately assayed for residual enzyme activity using the globin assay.

6.6 Thermal stability of the Sap isoenzymes

To compare the thermal stability of recombinant Sap1, 4 and 5 with native Sap2 and 3, each enzyme was incubated at the specified temperature for 15 min immediately prior to the activity assays (using the globin assay). The relative activities were then plotted (Fig. 6.11), and clearly showed that Sap4 and Sap5 were markedly more stable when compared with Sap2 and Sap3. The thermal stability of Sap1 was intermediate between these two groups, and was higher than reported for recombinant Sap1 from *S. cerevisiae* (Smolenski *et al.*, 1997), where the thermal stability profile was virtually identical to that of Sap2. Like the pH activity results, this suggested some difference between the recombinant Sap1 from *P. pastoris* and *S. cerevisiae*. Possible reasons for these differences include glycosylation of the recombinant proteins, and this will be discussed later in this chapter (Chapter 6.9).

6.7 SDS-PAGE analysis of Saps 1-5

The relative mobility of Saps 1-5 in SDS-PAGE investigated, and showed a running order Sap5>>Sap1>Sap2>Sap3>Sap4 (Fig. 6.12). The mobilities of these enzymes corresponded to apparent molecular weights of 36000, 380000, 41000, 42000 and 44000 respectively. The apparent molecular weights of Saps 1-3 were higher than deduced from their sequences. Sap5 migrated as expected for a protein with a deduced mass of 37000, while Sap4 had an apparent mass 8000 greater than expected from the deduced mass.

6.8 Native PAGE analysis of Saps 1-5

A comparison of the relative mobilities of Saps 1-5 on native PAGE (pH 8.8) is shown in Fig. 6.13. Sap5 did not migrate into the gel, while the relative mobilities of Saps 1-4 were Sap3>Sap2>Sap1>>Sap4. Migration in native PAGE is based upon the overall charge of the folded protein, and this order corresponds to the order of increasing negative charge in the Saps. These data are consistent with the previously reported mobilities of Sap1, Sap2 and Sap3 (Smolenski *et al.*, 1996). In addition, the mobilities and charge on Sap4 and Sap5 are consistent with the chromatography data presented in Chapter 5. Anion exchange at pH 6.0 was suitable for Sap1, pH 7.0 was needed for Sap4 and the purification of Sap5 required cation exchange chromatography at pH 7.0.

Recombinant Saps 1 and 4 (Fig. 6.13, Lanes 1 & 4) ran as diffuse bands on native PAGE in comparison to Saps 2 and 3. As Sap5 did not migrate into the gel, no comparison of relative band patterns could be made. However, these observations suggested that Sap1, Sap4, and possibly Sap5, might be glycosylated during secretion from *P. pastoris*. This might explain the marked increase in apparent size of Sap4 on SDS-PAGE compared to the deduced molecular weights (Fig. 6.12), although further analyses were required. Glycosylation of the Sap isoenzymes is addressed in Chapter 6.9.

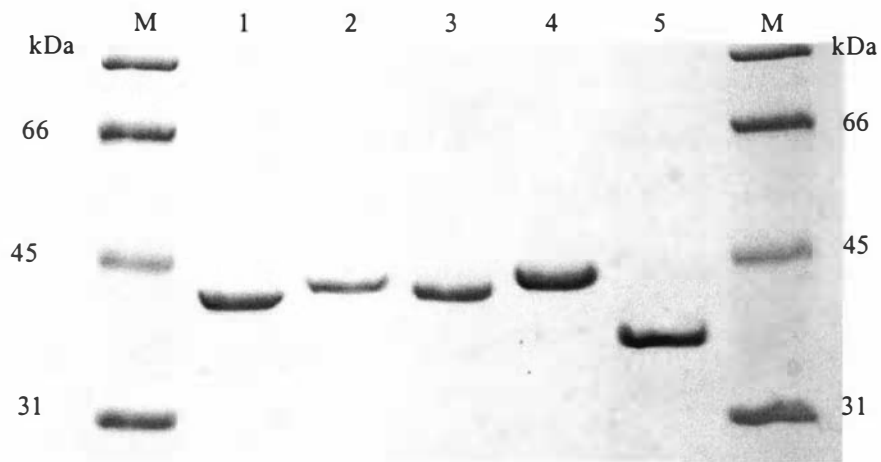


Figure 6.12 *SDS-PAGE analysis of Saps 1-5*

Three μg of each enzyme was run on a 10% SDS gel and visualised by staining with Coomassie Brilliant Blue™.

Lanes M: Molecular weight markers

Lanes 1-5: 3 μg of Sap1 – Sap5 respectively.

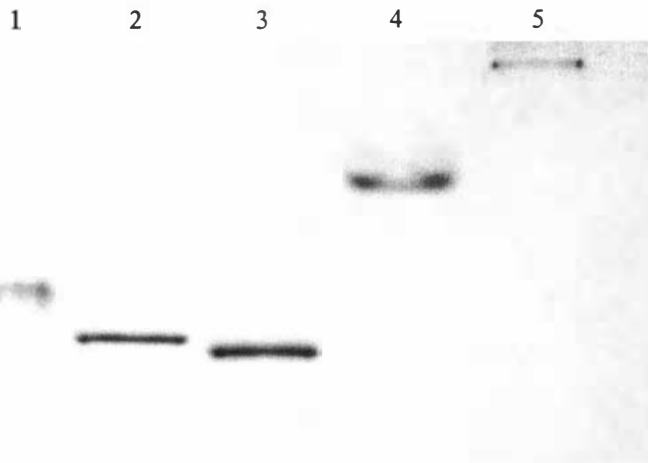


Figure 6.13 *Native-PAGE analysis of Saps 1-5*

Three μg each Sap1-5 was run on a 10% native gel (Method 2.10.6) and visualised by staining with Coomassie Brilliant Blue™.

Lanes 1-5: 3 μg of Sap1 – Sap5 respectively.

6.9 Autodegradation of Sap isoenzymes

The observations in section 6.6 suggested that Sap4 and Sap5 were more stable than other members of the Sap family of isoenzymes. Saps 1, 2 and 3 undergo autodegradation at low pH values, and an experiment was set up to compare the autodegradation of recombinant Saps 4 and 5 with recombinant Sap1 and native Sap2. This involved an 8 h incubation at 37 °C at the pH optimum for each enzyme. Samples were taken at various times and analysed for enzymic activity in the globin assay, and degradation by SDS-PAGE. The results are summarised in Fig. 6.14. SDS-PAGE analysis showed no autodegradation products of Sap4 or Sap5 (Fig. 6.14, A & B) during the 8 h incubation, while both Sap1 and Sap2 produced some smaller protein bands, which increased in abundance over time (Fig. 6.14, C & D). After 8 h, Sap2 activity decreased to 57% of the original value and Sap1 to 66% of the zero time level (Fig. 6.14, E). While there were some variations in the activity levels of all four enzymes during the 8 h, which were probably due to experimental error, there was no significant decrease in Sap4 and Sap5 activities.

These observations showed that Sap4 and Sap5 were not susceptible to autodegradation compared with Sap1 and Sap2. The results also reinforced the potential differences in the substrate specificities of these enzymes (also noted in section 6.5).

6.10 Pepstatin inhibition of Sap

Aspartic proteinases are typically inhibited by pepstatin, and *C. albicans* Saps 1, 2 and 3 are no exception. Pepstatin is a hexapeptide from *Streptomyces* (685.9 Da), which inhibits aspartic proteinases by mimicking the tetrahedral transition state which occurs during hydrolysis of normal substrates (Umezawa *et al.*, 1970). An analysis of pepstatin inhibition was performed with Saps 1, 4 and 5 using both the globin and fluorocasein assays. The results from the globin assay are shown in Fig. 6.15. The results from the fluorocasein assay produced equivalent inhibition curves (data not shown).

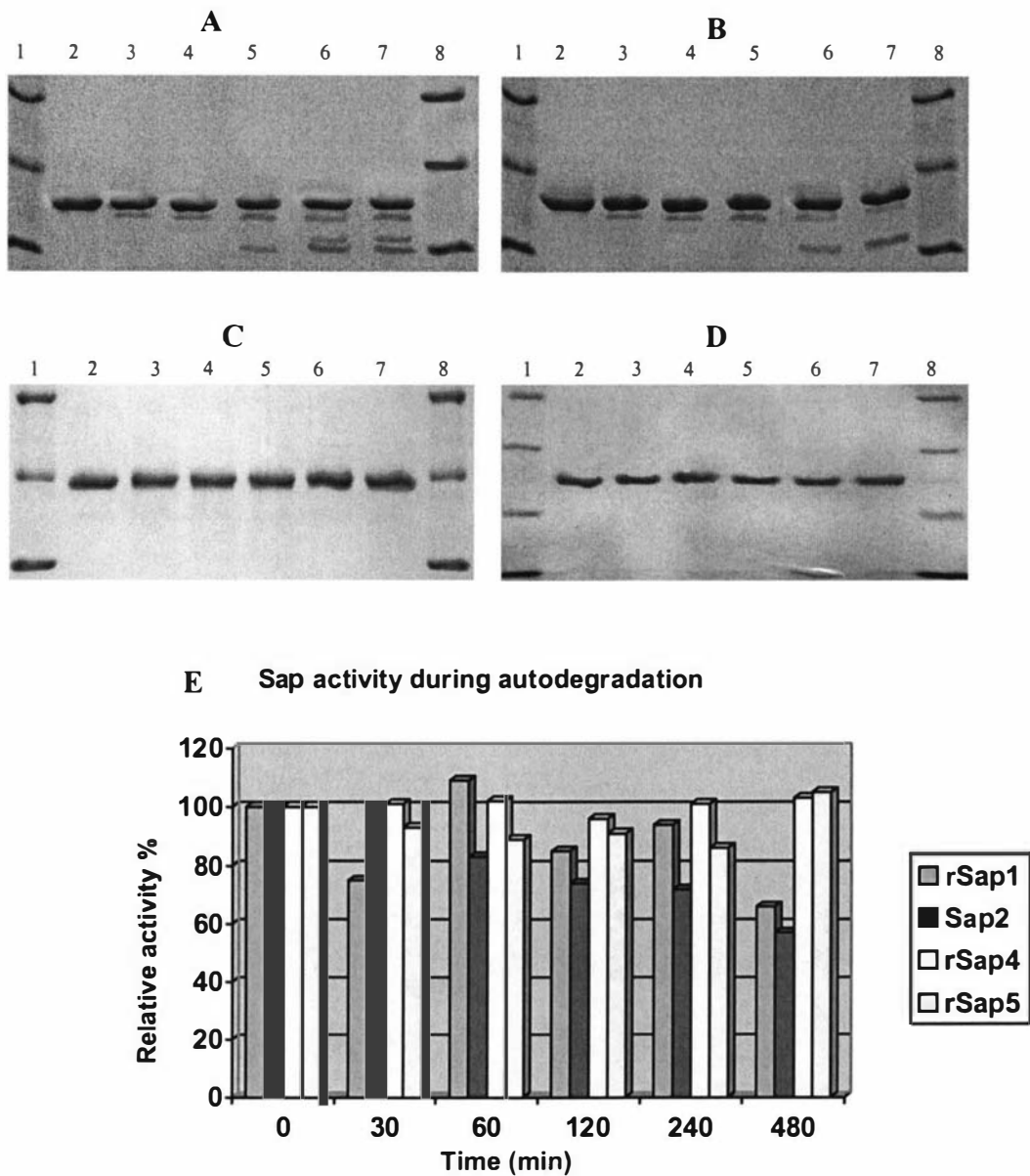


Figure 6.14 Sap autodegradation

A.-D. Autodegradation of native Sap2 (A), and recombinant Sap1 (B), Sap4 (C) and Sap5 (D). Sap1 and Sap2 were incubated in 100 mM sodium citrate pH 3.6, and Saps 4 and 5 were incubated in 100 mM sodium acetate buffer pH 4.5. Samples were taken at the times indicated. Assays contained 0.8 μ g Sap1, 0.8 μ g Sap2, 2.0 μ g Sap4 or 4.0 μ g Sap5.

Lanes 1 & 8: Molecular weight markers
 Lane 2: Time-zero sample
 Lane 3: 30 min
 Lane 4: 60 min
 Lane 5: 120 min
 Lane 6: 240 min
 Lane 7: 480 min

E. The relative enzyme activities during the incubation.

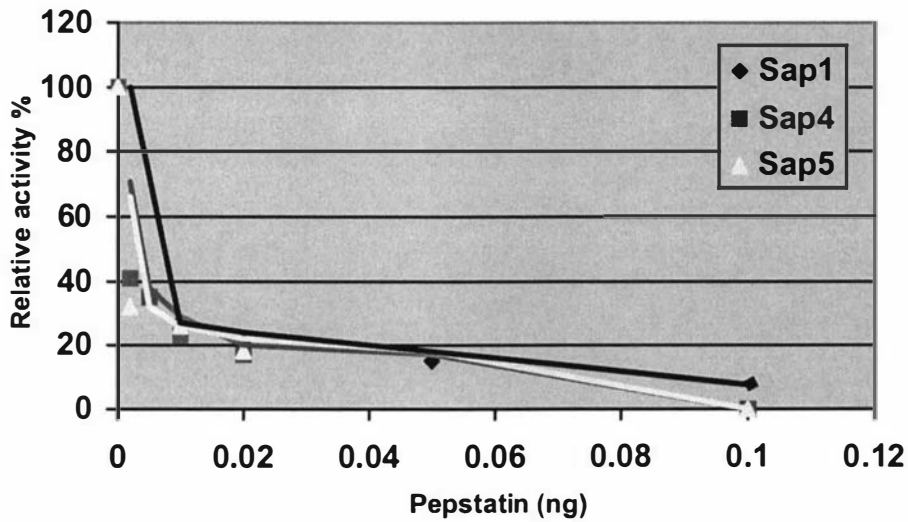


Figure 6.15 *Pepstatin inhibition of Sap*

Sap1, Sap4 and Sap5 (0.5 mg/mL) were incubated with increasing amounts of pepstatin then assayed for activity.

Pepstatin is known to bind and inhibit Saps 1, 2 and 3 in a 1:1 ratio (Cutfield *et al.*, 1993), and the similarity between the inhibition curves for Saps 4 and 5, and those determined for Saps 1, 2 and 3, suggested that pepstatin was probably inhibiting Sap4 and Sap5 in the same manner. That is, one molecule of pepstatin bound in each active site pocket of Saps 4 and 5.

6.11 Glycosylation of Sap

6.11.1 Electrospray mass spectrometry

Electrospray Mass Spectroscopy (ES-MS) was performed to determine the actual mass of the Sap proteins. To obtain ES-MS data, the protein samples were buffer exchanged against 60 volumes of 2 mM sodium bicarbonate, pH 9.0 and concentrated to approximately 20 pmol/ μ L in a total volume of 100 μ L. These samples were submitted for ES-MS analysis on a Sciex API 300 triple quadrupole mass spectrometer, and the experimentally determined masses were then compared with the deduced masses of the proteins.

Mass determination was performed for recombinant Sap1, Sap4 and Sap5. Native Sap2 and an exo- β -(1,3)-glucanase from *C. albicans* were used as controls. These data are summarised in Table 6.4. Both controls returned the expected masses (36329 for Sap2 and 45748 for Exoglucanase) within a margin of error of ± 7 Da (Wright *et al.*, 1991; Mackenzie *et al.*, 1997, and Appendix 10.4), confirming that the data were reliable. In contrast, the molecular weights for Sap1 and Sap4 were approximately 2,000 Da larger than predicted based on amino acid sequences. *C. albicans* deviates from the universal genetic code, and translates the codon CTG as serine, rather than the standard leucine (Santos *et al.*, 1993). Therefore, Sap1 was predicted to be 36,179 Da by amino acid sequence after substituting the mass of one leucine residue with that of a serine, to account for the presence of one CTG-serine codon in the sequence (White *et al.*, 1995a). However, the experimentally deduced mass of 38,411 Da was 2,232 Da larger than expected (Fig. 6.16 A). Similarly, Sap4 was found to be 39,036 Da, which was 2041 Da larger than the predicted value of 36,995 Da (Fig. 6.16, B). Several experiments were performed to determine whether the increase in size of each of these enzymes was due to glycosylation.

An ES-MS analysis for glycopeptide fragment ions was carried out, and this suggested that Sap1 and Sap4 were indeed glycosylated. For both proteins, the major peaks were at 294 and 365 Da (Fig. 6.17; A, B), which is consistent with sugars the size of sialic acid and a hexosyl-N-hexosamine disaccharide, respectively. Sialic acid has not been detected in yeast glycoproteins, but this does not exclude the possibility of a sugar

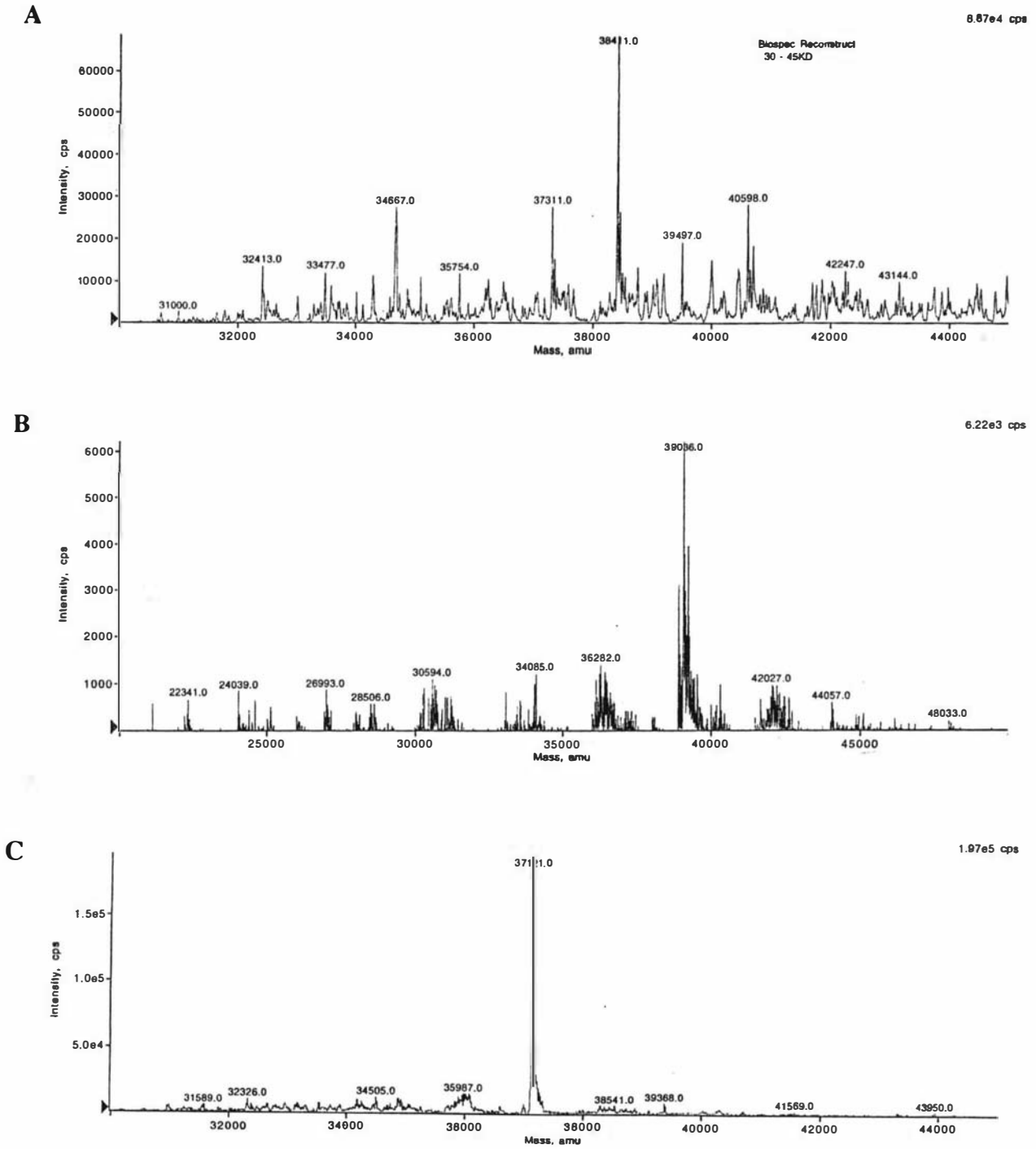


Figure 6.16 *ES mass spectra of Sap1, Sap4 and Sap5*

- A. ES-MS data for Sap1.
- B. ES-MS data for Sap4
- C. ES-MS data for Sap5

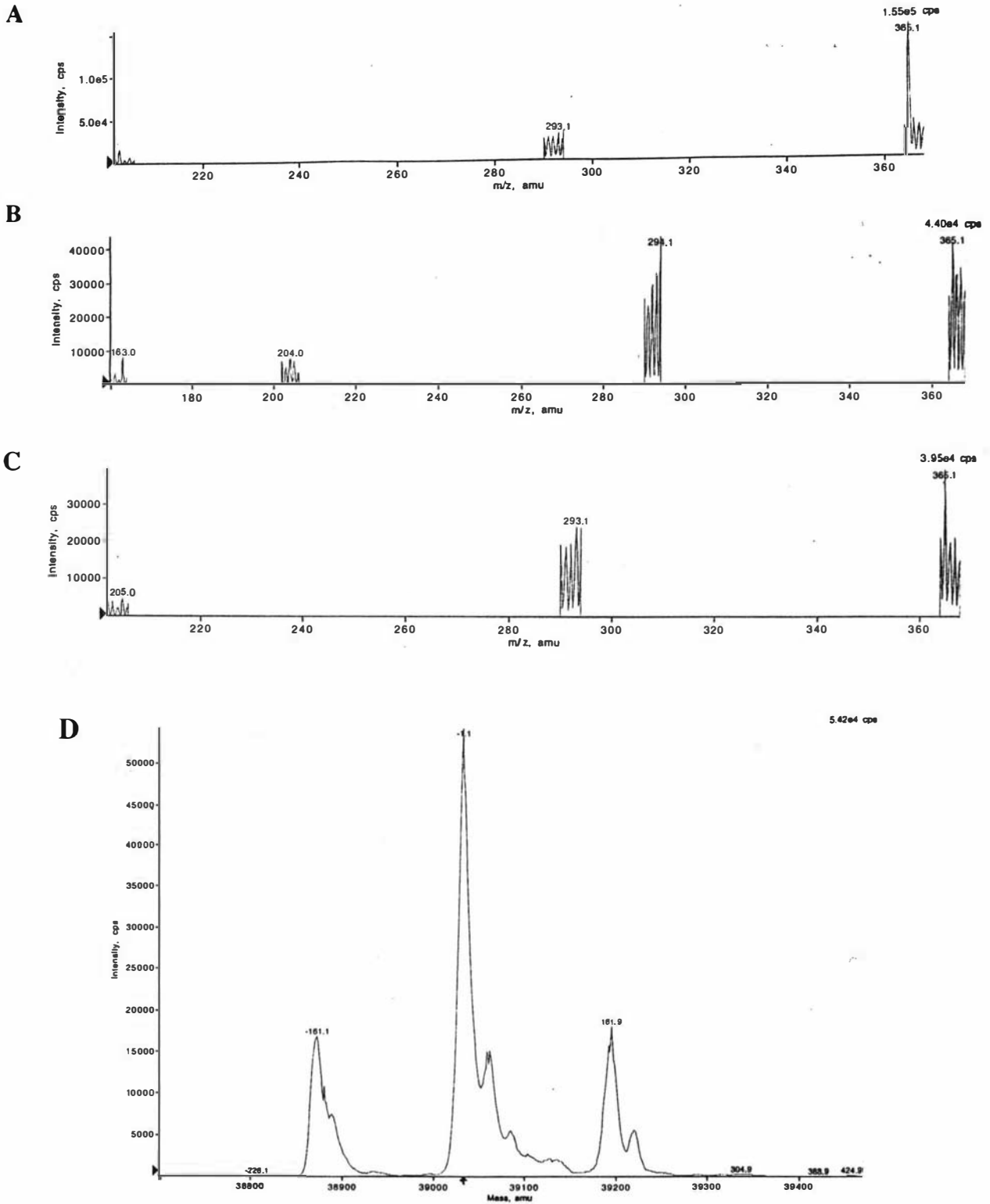


Figure 6.17 Glycopeptide fragment ion analysis

- A. Carbohydrate species from Sap1
- B. Carbohydrate species from Sap4
- C. Carbohydrate species from Sap5
- D. ES mass spectra showing evidence for three different glycoforms of Sap4

Protein	Deduced Mass (Da)	Experimental Mass (Da)	Difference (Da)
Sap1	36179	38411	+2232
native Sap2	36336	36329	+7
Sap4	36995	39036	+2041
Sap5	37256	37121	-135
native Exoglucanase	45745	45748	+3

Table 6.4 Predicted and experimental masses of Saps

The deduced masses of Saps 1, 2, 4 and 5 are shown next to the experimental data. Exo- β -(1,3) glucanase (Exg) from *C. albicans* is included as a control.

derivative of the same apparent mass. More convincing was the Sap4 mass spectrum which showed three different species (Fig. 6.17; D), where the main peak was same size as in the original ES-MS data. The two additional peaks have mass values which differ from the main peak by 162 Da, the size of +/- one hexosyl group. This experiment demonstrated the existence of several different glycoforms of Sap4. The mass increases measured by ES-MS (Table 6.4) suggested that Sap1 and Sap4 each possessed 11-13 sugar residues.

The predicted mass of Sap5 was 37,256 Da, based on the translated nucleotide sequence of pMS5. However, the experimental mass was 37,121 Da (Fig. 6.16, C), is not consistent with glycosylation, but rather a change in the amino acid sequence. The N-terminus was shown to be correct (Chapter 6.2), but an sensitive amino acid analysis is required to confirm the amino acid composition of Sap5 is correct. This puzzle has not been resolved, and poses a dilemma as to why Sap1 and Sap4 are glycosylated, but Sap5 is not.

An ES-MS glycosylation analysis was also performed for Sap5, and results indicated that this protein was also glycosylated (Fig. 6.17, C). However, there was no apparent mass increase corresponding to glycosylation of Sap5. It is unlikely that Sap5 possesses labile carbohydrates which became detached during the ionisation process for mass determination, therefore the most probable origin of these sugar residues was carbohydrate species which co-purified with Sap5.

Further experiments were performed to identify the carbohydrate groups on Sap1 and Sap4, and to confirm whether Sap5 was glycosylated. Carbohydrates can be attached to proteins via the amino acid side chains of asparagine (*N*-linked) and serine or threonine (*O*-linked). The potential attachment site for *N*-linked carbohydrates is the consensus sequence Asn-X-Ser/Thr, so an analysis of the amino acid sequences of all three proteins was performed, to check for the presence of this motif. The absence of this sequence in all three protein sequences indicated that the carbohydrates were probably *O*-linked, although attachment at a non-conservative *N*-glycosylation site is also possible (Jung *et al.*, 1998).

6.11.2 DIG glycan/protein analysis of Sap

A DIG-glycan kit was used to detect glycosylation in the Sap preparations (Method 2.10.13). The principle of the DIG-glycan double-detection system involves labelling the side chains of the sugar groups and the amino groups of the proteins with different haptens, which are then detected with a corresponding antibody. Saps 1-4 were run on 10% SDS-PAGE, electroblotted to Hybond C, followed by the DIG detection protocol (Method 2.10.13.2). The results (Fig. 6.18) indicated that there was no glycosylation whatsoever on either of Sap1 or Sap4, while Sap2 appeared to have some attached sugar residues. Sap2 is not glycosylated, as judged by ESMS, and given the ES-MS data for Sap1 and Sap4, the results were puzzling. However the chemistry of the kit limits the range of carbohydrate groups detected and it is possible that detection of the sugars attached to Sap1 and Sap4 has been precluded in this case.

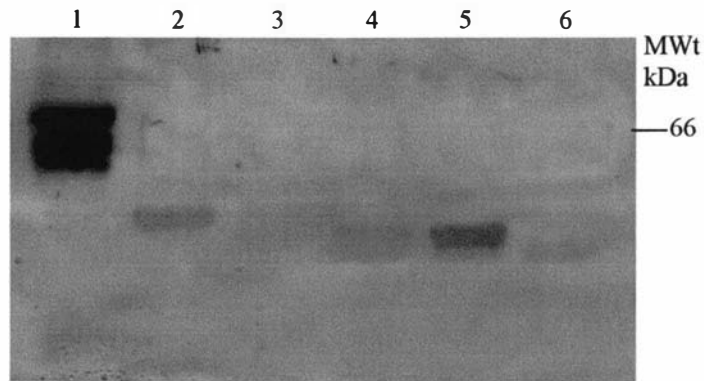


Figure 6.18 *DIG glycan/protein analysis of Saps 1, 2, 3 and 4*

The DIG glycan/protein kit detects glycoproteins (blue) and non-glycosylated proteins (brown). Each lane contains 3 μ g of protein.

Lane 1: Glycosylated protein positive control - Fetuin

Lane 2: Non-glycosylated protein positive control - Creatinase

Lane 3: Recombinant Sap4

Lane 4: Native Sap3

Lane 5: Native Sap2

Lane 6: Recombinant Sap1

6.11.3 Carbohydrate chromatography

Samples of Saps 1, 4 and 5 were subjected to acid hydrolysis to release and covalently linked sugars for identification by carbohydrate chromatography. Two methods of hydrolysis were used for each protein; trifluoroacetic acid and hydrochloric acid (Methods 2.10.12.2), which yield the mannose and the N-acetyl glucosamine residues respectively. The hydrolysates were applied to a CarboPac PA10 column (4 mm, 10-32), and eluted under isocratic conditions with 18 mM NaOH, over 20 min. Under these conditions the sugars have characteristic retention times, and these were used for identification. An ED40 electrochemical detector in conjunction with the Pulsed Amperometry Detection system (PAD) was used to detect the eluting carbohydrates, which were slightly charged in the NaOH solvent (Dionex Corp. Sunnyvale, California).

Five controls were run:

- Untreated mannose
- TFA-treated mannose
- HCl-treated mannose
- Untreated N-acetyl glucosamine
- HCl-treated N-acetyl glucosamine

During the hydrolysis reactions, some destruction of the sugar groups occurred, and comparison between the treated and untreated control samples enabled this level of destruction to be calculated. In addition, a mixture of several standards was run at five different concentrations. These were galactosamine, glucosamine, galactose, glucose and mannose, at 0.01 mg/mL, 0.02 mg/mL, 0.03 mg/mL, 0.04 mg/mL and 0.05 mg/mL. The relative elution time of each was used to identify the unknown carbohydrate species in the Sap preparations, and the areas under the peaks of the standards was used to construct a standard curve for quantitation of the carbohydrate species. N-acetyl galactosamine and N-acetyl glucosamine were not run as standards, because the acid hydrolysis conditions result in deacetylation. Glucosamine and galactosamine were run as controls, in place of their N-acetylamino counterparts.

Sugar residue	Area under peak	% Sugar Remaining
Untreated Mannose	348555373	100 %
Mannose + HCl	115567953	34 %
Mannose + TFA	80548796	23 %
Untreated N-acetyl glucosamine	40606675	100 %
N-acetyl glucosamine + HCl	28383014	70%

Table 6.5 *Carbohydrate destruction during acid hydrolysis*

The area under the peaks of the controls were used to calculate the destruction of the sugar under acid hydrolysis conditions with 400 µg samples of mannose or N-acetyl glucosamine (Methods 2.10.12.2). The destruction is the ratio between the original amount of sugar and the amount left after acid hydrolysis, expressed as a percentage

Table. 6.5 shows that the destruction of mannose in HCl was approximately 66%, while the destruction in TFA was 77%. HCl resulted in approximately 39% destruction of the N-acetyl glucosamine. These levels of destruction were used to extrapolate the original amount of carbohydrate (either mannose or N-acetyl glucosamine) present on the Sap preparations.

The results from the Sap1 analysis were inconclusive, due to the presence of a large peak, which obscured other results. This peak was possibly due to excess salt in the sample or some other contaminant. However there were peaks which suggested that Sap1 was glycosylated. The small peaks at 8.20 and 9.67 min (Fig. 6.19, B) correspond to N-acetyl galactosamine and N-acetyl glucosamine in the original Sap1 sample. No peak at 14.3 min - indicative of mannose - was seen in the Sap1 trace (Fig. 6.19, C).

While the results from the Sap1 analysis did not permit quantitation of the carbohydrate moieties, the Sap4 data was much clearer. From the Sap4/HCl hydrolysis, a peak at 9.57 min suggests the presence of N-acetyl glucosamine in the original sample (Fig. 6.19, D). The data from the HCl hydrolysis was used to estimate the mass of N-acetyl glucosamine by calculating the area under the peak at 9.57 min, then using the destruction figures (Table 6.5) to extrapolate the original amount of carbohydrate present in the sample. The amount of N-acetyl glucosamine present in the original Sap4 protein sample (250 μg or 6.4 nmoles) was then calculated using the data from the standard curves in Fig. 6.20 (see Appendix 10.2.1). The total mass of GlcNAc was 8.4 μg , or 39 nmoles GlcNAc, which corresponded to 6 moles of GlcNAc per mole of Sap4.

Fig. 6.19 (E) showed the outcome of hydrolysis of Sap4 with TFA. The peak at 14.22 min corresponds to mannose. This set of data was used to calculate the mass of mannose present in the original sample by first accounting for mannose destruction of 77%. Then the total amount of mannose present in the original sample was estimated by using the standard curve data (Fig. 6.20). The total mass of mannose was 28 μg , or 156 nmoles mannose, which corresponded to 24 moles of mannose for each mole of Sap4.

The Sap5 results showed very little carbohydrate (Appendix 10.2.2). No peaks corresponding to either mannose or N-acetyl glucosamine were seen. This is consistent with the molecular weight estimation of Sap5 by ESMS (Chapter 6.9.1). However, there were some small peaks of other unidentified sugar residues. The presence of these unidentified peaks in the Sap5 chromatogram, and similar peaks in the Sap1, Sap4 traces is consistent with the possibility that some co-purifying, non-covalently attached sugars are present in each enzyme preparation.

The Dionex data suggested that each molecule of Sap4 possessed 6 molecules of N-acetyl glucosamine, and 24 of mannose. This implied that the actual mass of Sap4 was 42647 Da (the mass from the deduced sequence of 36995 Da + 1327.2 Da GlcNAc + 4324.8 Da mannose). However the experimentally determined size of Sap4 was 39036 Da, 3611 Da smaller than the expected size of the intact glycoprotein based on Dionex analysis. This observation again supported the presence of co-purifying sugar

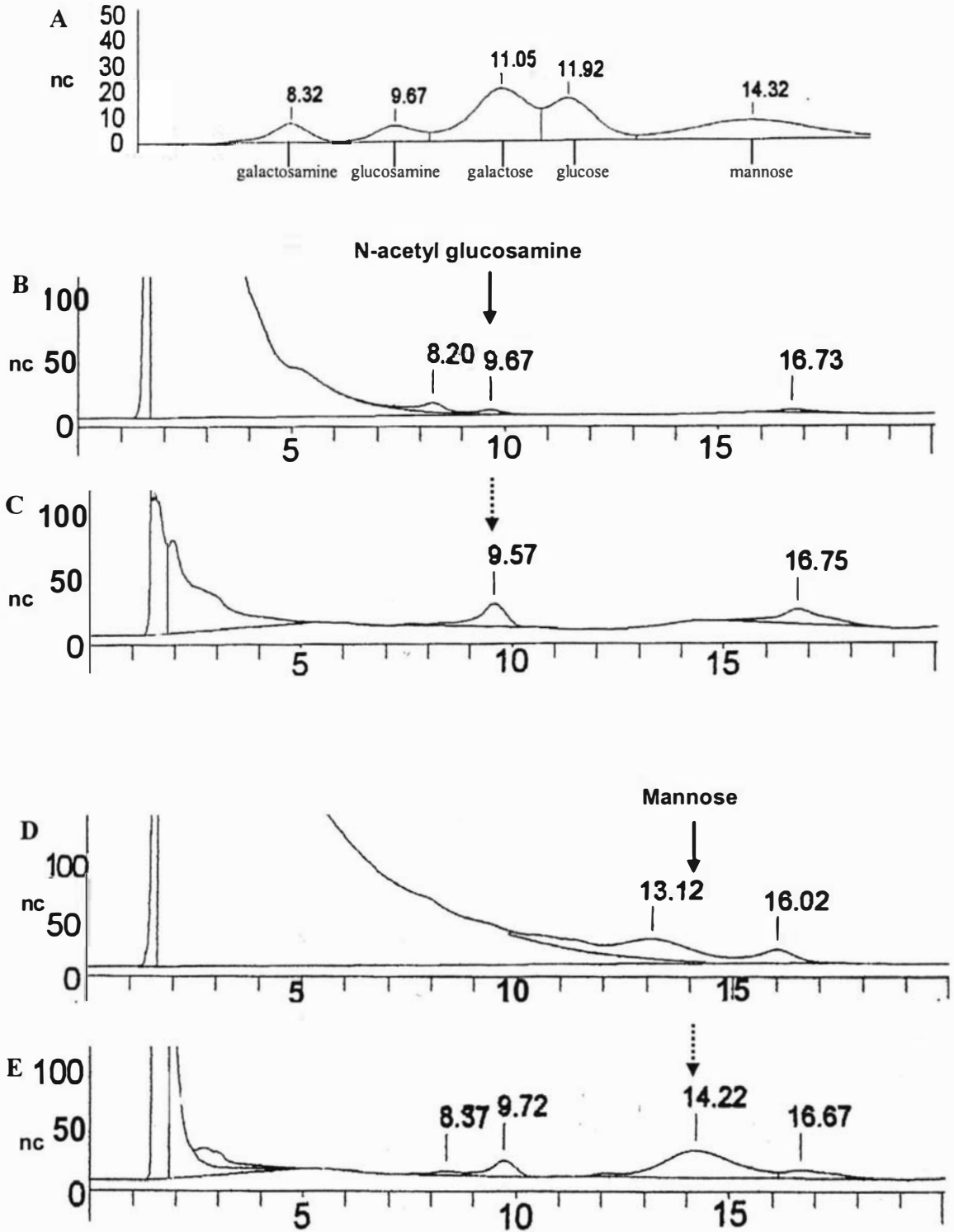


Figure 6.19 Sap Dionex analysis

The chromatograms of the carbohydrate moieties released by either HCl or TFA hydrolysis of the recombinant Sap proteins. 250 µg of each protein was analysed.

- A. Standards (0.02 mg/mL)
- B. Sap1 + HCl
- C. Sap4 + HCl
- D. Sap1 + TFA
- E. Sap4 + TFA

Standard curves for quantitation of carbohydrate species

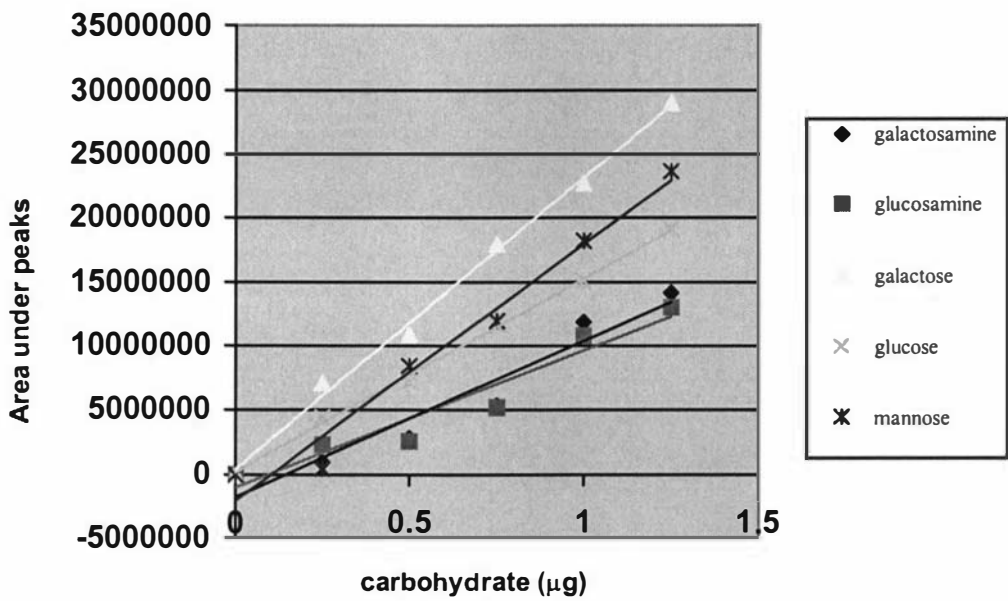


Figure 6.20 *Standard curve for quantitation of carbohydrate species*

The curves were drawn using the area under each peak, for each standard at the five concentrations stated (0.01, 0.02, 0.03, 0.04 and 0.05 mg/mL). This data was used to calculate the amount of glucosamine and mannose present in the original Sap samples.

residues in the enzyme preparations. Therefore, it is likely that the experimental mass data determined by ESMS represents the actual mass of the glycosylated Sap isoenzymes, and the mass implied by the Dionex analysis represents the mass of the O-linked oligosaccharides of Sap4 plus the mass of any other sugar residues present in the preparation.

Dionex analysis of the carbohydrate species in the Sap1, Sap4 and Sap5 preparations has enabled identification of some of the carbohydrate species, and confirmed that Sap5 is unlikely to be glycosylated, as suggested by mass analysis of the enzyme (Fig. 6.16). However, given the likelihood that each Sap isoenzyme preparation contains co-purifying, non-covalently linked carbohydrates, these data are considered qualitative rather than quantitative. Therefore it remains likely that Sap1 and Sap4 each possess between 11-13 sugar residues (section 6.11.1), a mixture of mannose and N-acetyl glucosamine.

Other properties of recombinant Sap1 and Sap4 are also indicative of glycosylation; for example both migrated as fuzzy bands on native PAGE (Chapter 6.8), and recombinant Sap1, Sap4 and Sap5 showed increased thermal stability compared with native Sap1, Sap2 and Sap3 (Chapter 6.6). Glycosylation has been shown to influence protein stability (Wang *et al.*, 1996), and would account for the observed properties of the Saps. Since none of the recombinant Saps possessed the consensus sequence for N-glycosylation, the attached sugars were most likely O-linked through serine and threonine side chains, although N-glycosylation at non-conservative sites remained a possibility (Jung *et al.*, 1998). Little is known about any amino acid consensus sequences or structural features which specify an O-glycosylation site, but yeast O-glycosylation on serine or threonine residues typically consists of two or three mannose residues that are linked (α 1-2) and occasionally (α 1-3) at the non-reducing end (Ballou, 1990). In addition the removal of pro-sequences by Kex2 (or a Kex2 homologue) is frequently a rate-limiting step in protein over-expression. Delays in processing the pro-region result in long resident times in the Golgi complex, which can result in protein hyperglycosylation (Shuster, 1991).

The ES-MS analysis indicated the presence of either GlcNAc or GalNAc on Saps1, 4 and 5, and Dionex analysis of the carbohydrates confirmed the presence of GlcNAc.

This was unexpected for *O*-linked oligosaccharides in fungi. Typically, fungi initiate *O*-glycosylation by the addition of an initial mannose residue in the ER, which is then elongated with either mannose and/or galactose in the Golgi complex (for review see Strahl-Bolsinger *et al.*, 1999). While little has been reported concerning the patterns of *O*-glycosylation in *P. pastoris*, possibly due to the small percentage of proteins which possess *O*-linked glycans (Grinna & Tschopp, 1989), preliminary experiments have shown the *O*-glycans of *P. pastoris* to be similar to those of *S. cerevisiae*. Nevertheless, the presence of more complicated, novel branched structures has also been suggested by Gemmill & Trimble (1999). The typical structure of a yeast *O*-glycan is shown in Fig. 6.21, alongside a model of the proposed structure of the Sap *O*-glycans produced by *P. pastoris*. The mannose(α 1-2)-mannose-Ser/Thr core shown for *S. cerevisiae* (Fig. 6.21, A) is common to all fungi studied to date (Gemmill & Trimble, 1999). The proposed *P. pastoris* *O*-glycan structures are based on the common core, but might include a mixture of *O*-glycans composed of either Mannose or Mannose/GlcNAc residues.

It has been suggested that *O*-mannosylation might play a direct role in protein secretion and increase resistance to intracellular proteases (Strahl-Bolsinger *et al.*, 1999). If this is the case then the results presented here, which have indicated the existence of novel *O*-glycosylation patterns in *P. pastoris*, elude to the possibility of novel mechanisms of protein sorting and secretion in *P. pastoris* and possibly other yeasts.

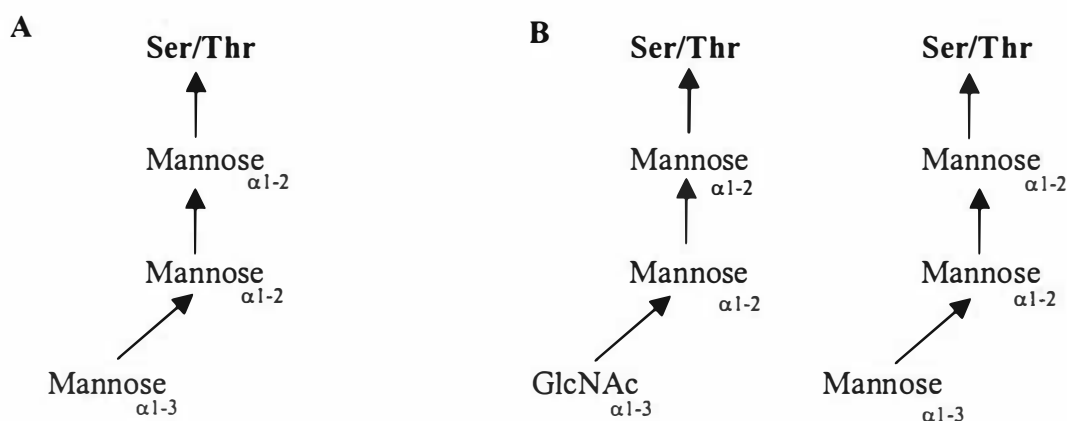


Figure 6.21 *O*-linked oligosaccharides of *S. cerevisiae* and *P. pastoris*

- A. Structure of *S. cerevisiae* *O*-glycans.
 B. Structure of proposed *P. pastoris* *O*-glycans.

6.12 Interaction of recombinant Saps with SQAPI

SQAPI (*Squash Aspartic Proteinase Inhibitor*) is a 10.5 kDa protein, which inhibits aspartic proteinases as part of a plant defence mechanism against pathogens (Christeller *et al.*, 1998). In collaboration with Dr. W. Laing and Dr. P. Farley, an analysis of the binding affinity of the recombinant Sap isoenzymes for SQAPI was performed using a Biacore X.

Biomolecular interaction analysis (BIA) uses optical biosensors to monitor molecular binding processes on the surface of a biosensor chip. One of the interacting partners is attached to the chip, in this case SQAPI, and a solution containing the other partner is passed over the surface. The interaction of the partners is recorded as a function of time. The response, measured in so-called resonance units (RU), is a direct indication of the amount of interacting partner that binds from the solution. Once the chip containing the immobilised SQAPI was inserted into the Biacore 1000, analytes were tested for interactions with the ligand. As a control for non-specific interactions between the analytes and the chip surface, a chip with no bound ligand was tested in parallel with these experiments.

At the time of analysis, Sap5 had not been purified, so the analytes tested were recombinant Sap1, native Sap2 and recombinant Sap4. The analyses performed must be considered qualitative, as insufficient data was collected to allow any reliable kinetic analysis of the interactions to be calculated. However, the results showed yet more evidence of variation between these members of the Sap family of isoenzymes (Fig. 6.22). The Sap1 binding curve had a slow on-rate, and a fast off-rate, which indicated that the affinity of Sap1 for SQAPI was low. The Sap4 binding curve showed a much faster on-rate, and a much slower off-rate, which suggested that Sap4 was binding SQAPI with a relatively high affinity. In contrast, there appeared to be no specific binding of Sap2 to the immobilised SQAPI. Incubating Sap4 with pepstatin prior to the analysis totally abolished the binding curves, which indicated that the interaction was between SQAPI and the Sap active site. The square-shaped Sap2 binding curve suggested that there was no specific interaction between Sap2 and

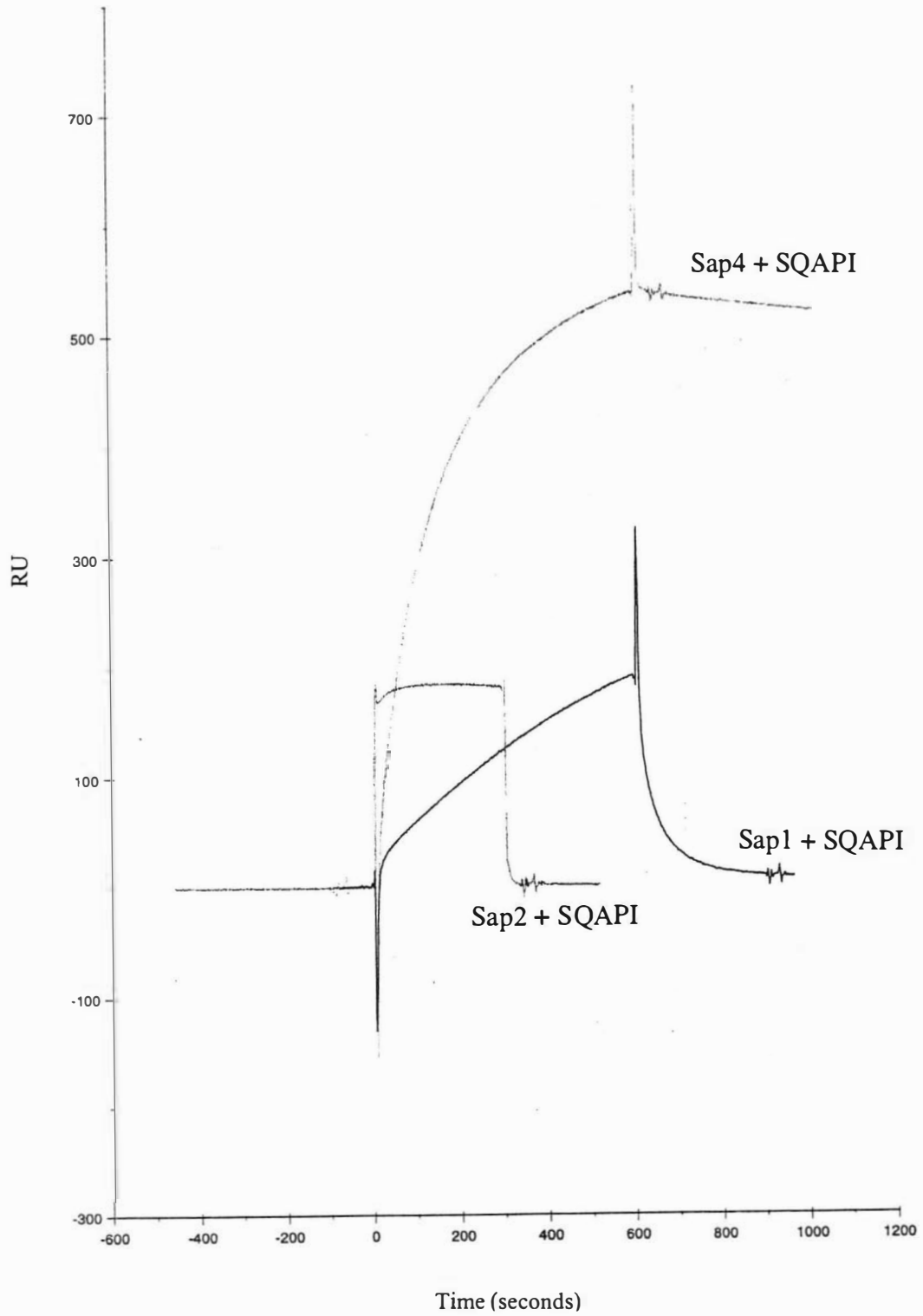


Figure 6.22 *Biacore analysis of Sap interactions with SQAPI*

The figure shows the binding curves of recombinant Saps 1 & 4, and native Sap2, with SQAPI.

SQAPI. However, if there was no interaction at all, the binding curve would have been a straight line, but the square curve suggested that Sap2 was interacting with SQAPI in a non-specific manner.

6.13 Antibody production

During this study, an anti-Sap2 antibody preparation was used to identify recombinant Sap1, 4 and 5. It was previously shown that this anti-Sap2 cross-reacted strongly with Sap1 and Sap3 (Smolenski *et al.*, 1997), but the immunoreaction between anti-Sap2 and recombinant Sap4 and Sap5 was poor at best (Chapter 4.5). Therefore an anti-Sap4 antibody was required for future studies. Due to the high identity between the Sap4, Sap5 and Sap6 at the amino acid level (78-91%), it was anticipated that any polyclonal antiserum raised to one of these enzymes would cross-react with the other two. This high degree of identity has so far prohibited discrimination between transcripts of *SAPs 4-6* in northern analyses (Hube *et al.*, 1994) using full-length DNA probes. While a polyclonal antibody would not be expected to be specific for individual Saps, it would allow localisation of the enzymes within the cell, and other methods could be employed to distinguish between isoenzymes (such as probing northern blots with gene-specific oligonucleotides).

Sap4 was used as antigen for production of anti-Sap4 antiserum. It was planned that this antibody would allow western blotting and immunofluorescence analyses following identification of *SAP4-6* transcript from northern blots. However, *in vitro* expression of *SAP4-6* in *C. albicans* proved to be unsuccessful (section 7.1), and while these experiments were not performed, the anti-Sap4 antibodies, produced as part of this work, will provide a useful tool for future expression analyses.

6.14 Production of anti-Sap4

Sap4 antiserum was produced in New Zealand white rabbits at the Massey University Small Animal Production Unit (SAPU). A primary challenge of 200 µg Sap4,

emulsified in Freund's complete adjuvant (to stimulate the immune system), was given simultaneously to two rabbits (K42 and I45). Two subsequent boosters (100 µg Sap4, emulsified in Freund's incomplete adjuvant) were given to each rabbit, four weeks apart (Method 2.11). The antibody titre was checked by western blotting after each booster according to Method 2.10.7, except that the membranes were cut into strips after electroblotting, and each strip was incubated with a different dilution of antiserum. This experiment revealed that only one rabbit, K42, had mounted an immune response to the Sap4 challenge (Fig. 6.23). The Sap2/anti-Sap2 reaction (lanes 1-4) was a positive control and the serum from rabbit K42 showed good immunoreaction to Sap4 (lanes 5-8). However, rabbit I45 showed no response to Sap4, but some immunoreaction to higher molecular weight proteins (lanes 9-12). Secreted high molecular weight proteins of fungi are often mannoproteins and it is possible that rabbit I45 had mounted an immune response to a fungal infection prior to immunisation with Sap4. Based on this result, rabbit K42 was sacrificed by heart puncture (performed by E. Ormsby, SAPU) and 100 mL blood was harvested for purification of the antibody fraction.

6.15 Purification of the IgG fraction

Following harvest of the serum from rabbit K42, the IgG fraction was purified by ammonium sulphate precipitation (Method 2.11). A western analysis of an induced *P. pastoris* culture supernatant was performed to check that the purified antibody preparation was specific for Sap4. As shown in Fig. 6.24 there was no detectable cross-reactivity between anti-Sap4 antibodies and other extracellular proteins produced by *P. pastoris*.

6.16 Titre of anti-Sap4

6.16.1 Comparison of anti-Sap2 and anti-Sap4

Having confirmed that there was no cross-reaction with background proteins from *P. pastoris*, the anti-Sap4 was analysed for sensitivity. Initially, the anti-Sap4 was

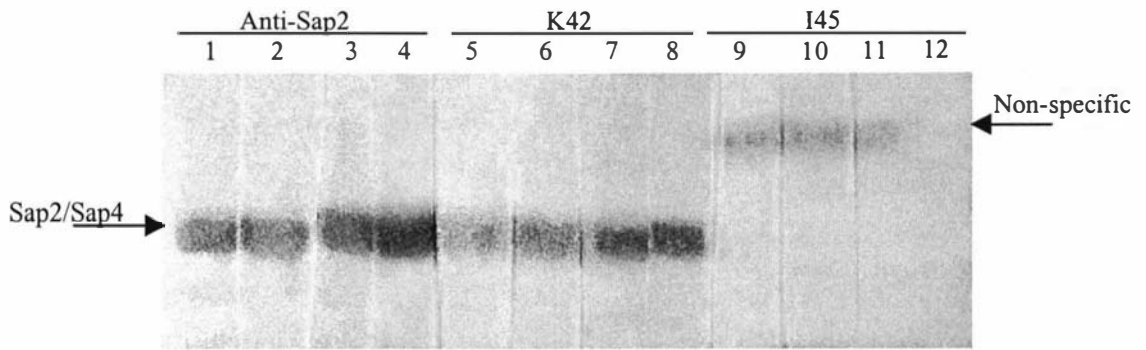


Figure 6.23 *Antibody titre after second booster*

Western blot showing 3 μg of either Sap2 (lanes 1-4) or Sap4 (lanes 5-12), probed separately with anti-Sap2, or K42/I45 serum as follows:

- Lanes 1, 5 & 9: 1/2000 dilution of antibody
- Lanes 2, 6 & 10: 1/1000 dilution of antibody
- Lanes 3, 7 & 11: 1/500 dilution of antibody
- Lanes 4, 8 & 12: 1/100 dilution of antibody

1 2



Figure 6.24 *Western analysis of anti-Sap4 vs. P. pastoris supernatant*

Western analysis to check the anti-Sap4 did not immunoreact with other proteins in the culture supernatant. A 1:2,000 dilution of serum was used.

- Lane 1: Induced KM71 control (20 μL culture supernatant)
- Lane 2: Induced MS4A (20 μL culture supernatant)

compared with the anti-Sap2 at different dilutions. The western blots were performed as described in Method 2.10.7, except that the membrane was cut into strips following the electroblotting and each strip was incubated with a different dilution of primary antibody. Decreasing amounts of each antibody were reacted with the same amount of antigen (200 ng of either Sap2 or Sap4) and blots were developed for the same length of time, using the same reagents to eliminate variation. Fig. 6.25 (A, lane 1) showed the anti-Sap2 at a 1/500,000 dilution gave an immunoreaction with 200 ng of Sap2. In contrast the anti-Sap4 showed no reaction at 1/500,000 dilution, but a band can be seen at 1/50,000 dilution (B, lanes 1 & 2). These observations show that both antibodies have a high titre, but that the anti-Sap2 is either a more concentrated preparation, or is more sensitive than the anti-Sap4 preparation.

6.16.2 Anti-Sap4 titre

To quantitate the sensitivity of the anti-Sap4, a series of westerns were performed against a series of dilutions of Sap4 (1-1000 ng) electroblotted onto a single piece of nylon membrane as described in Method 2.10.7. The results in Fig. 6.26 showed that the anti-Sap4 detected as little as 1 ng Sap4.

6.17 Specificity of anti-Sap4

A comparison was performed of the cross-reactivity of anti-Sap4 and anti-Sap2 against Saps 1-5. Two identical 10% SDS-PAGE gels, each with 200 ng of Sap1, Sap2, Sap3, Sap4 and Sap5, were electroblotted to Hybond C and individually probed with a 1/5,000 dilution of either anti-Sap2 or anti-Sap4. All subsequent steps were performed with the two blots together, to eliminate as much temporal and physical variation as possible. Fig. 6.27 shows the results of the analysis. As previously described, anti-Sap2 cross-reacted strongly with Sap1 and Sap3 (Smolenski *et al.*, 1997), there was a weak cross-reactivity between anti-Sap2 and Sap4, and no visible reaction between anti-Sap2 and Sap5 (Fig. 6.27, A). In contrast, anti-Sap4 showed much less cross-reactivity than the anti-Sap2 and was virtually specific for the Sap4 antigen. There was no reaction

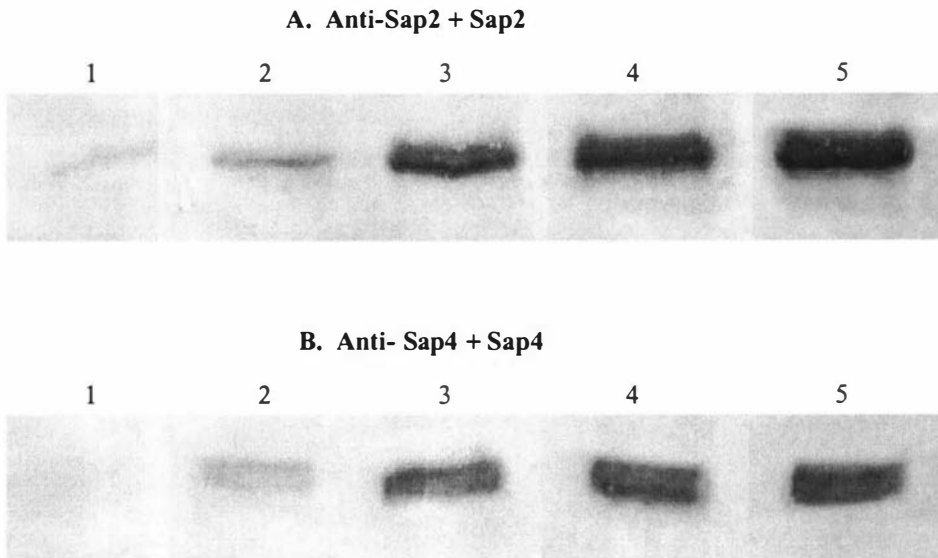


Figure 6.25 *Comparison of the sensitivity of anti-Sap2 and anti-Sap4*

A & B

Lane 1: 1/500,000 dilution of antisera

Lane 2: 1/50,000 dilution of antisera

Lane 3: 1/5,000 dilution of antisera (usual working concentration)

Lane 4: 1/500 dilution of antisera

Lane 5: 1/50 dilution of antisera

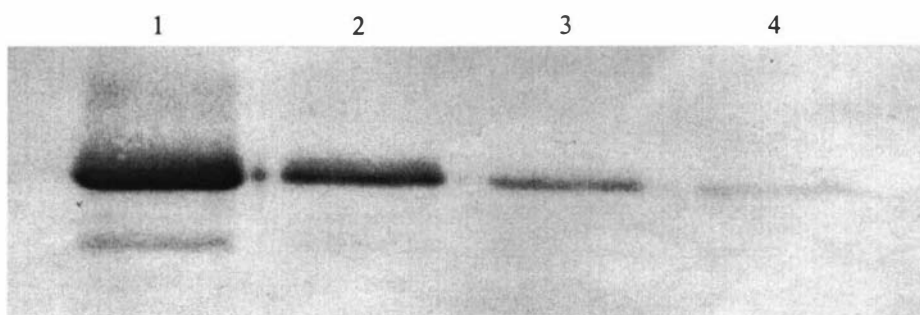


Figure 6.26 *Anti-Sap4 titre*

Lane 1: 1 µg Sap4

Lane 2: 100 ng Sap4

Lane 3: 10 ng Sap4

Lane 4: 1 ng Sap4

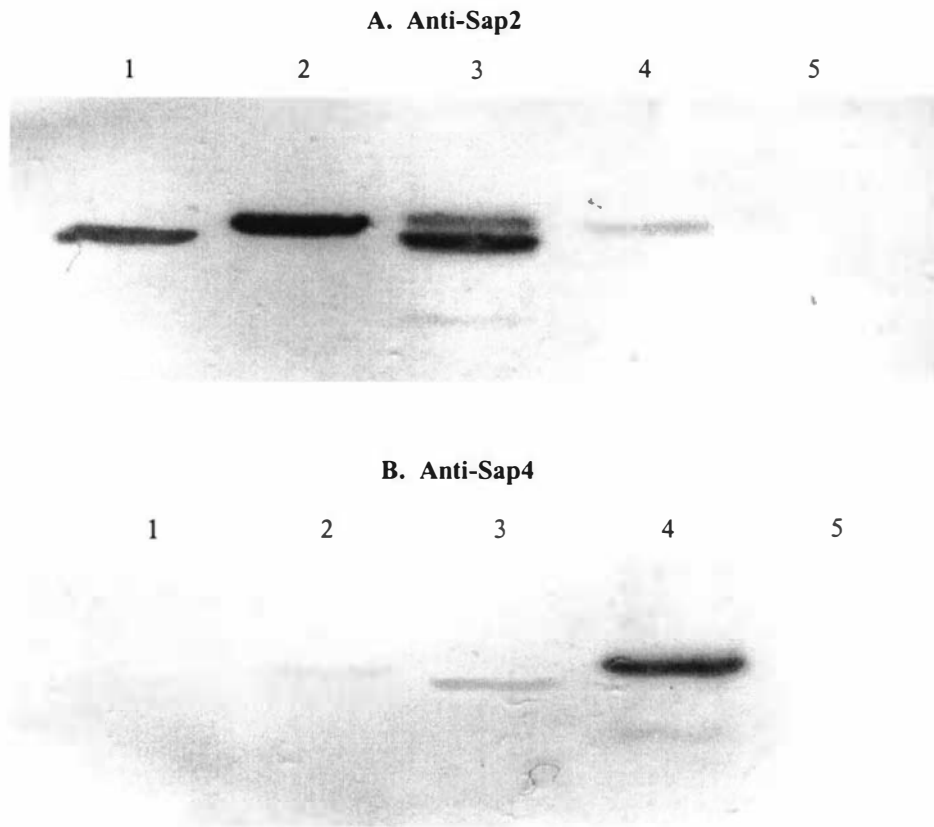


Figure 6.27 *Comparison of the specificity of anti-Sap2 and anti-Sap4*

A & B

Lane 1: 200 ng Sap1

Lane 2: 200 ng Sap2

Lane 3: 200 ng Sap3

Lane 4: 200 ng Sap4

Lane 5: 200 ng Sap5

between the anti-Sap4 and either Sap1 or Sap5, and only a weak reaction with Saps 2 and 3 (Fig. 6.27, B).

These observations indicated that the anti-Sap2 was a useful tool for investigating Saps 1-3, as cross-reactivity with Saps 4 and 5 is negligible. Equally, the anti-Sap4 had essentially no cross-reactivity with the Sap 1-3 group. Further, the anti-Sap4 showed no immunoreaction with Sap5, suggesting it is an extremely specific antibody preparation. Given the degree of identity between the Sap4 and Sap5 proteins (78% at the amino acid level), it is surprising that there is no cross-reaction of anti-Sap4 with Sap5. This suggests the presence of an epitope which is not conserved between Sap4 and Sap5, despite the high sequence identity between these proteins. This specificity makes the anti-Sap4 preparation an excellent tool for Sap4 expression analyses, as *SAP4*, *SAP5* and *SAP6* transcripts have been previously indistinguishable from each other (Hube *et al.*, 1994). More recently, immunofluorescence has been used to analyse expression of Sap4, Sap5 and Sap6 isoenzymes in murine macrophages. However, the antibodies produced for this work cross-reacted across the Sap4-6 group. Therefore, identification of the expressed Sap proteins has required parallel analyses of Sap mutant strains in which two of the three genes were deleted. This method of deducing which isoenzyme is expressed is laborious, and could be circumvented by using more specific antibodies, such as the anti-Sap4 produced in the present project.

7.0 EXPRESSION OF Sap

7.1 Introduction

Previous *in vitro* studies of *C. albicans* demonstrated expression of *SAPs 4-6* during serum-induced hyphal formation (Hube *et al.*, 1994). Analysis of expression involved northern blots with 1100 bp of *SAP5* ORF probe. Since *SAP4*, *5* and *6* are 89-93% identical at the nucleotide level it was not possible to conclude which combination of the three genes was expressed during the yeast-hyphal transition. There are various methods for inducing budding yeast to form hyphae, including addition of N-acetyl glucosamine (Cassone *et al.*, 1985), a pH-temperature transition, growth in a defined amino acid medium (Lee *et al.*, 1975) and addition of serum to the growth medium (Remold *et al.*, 1965). However, expression of *SAPs 4-6* was only associated with the serum-induced yeast-hyphal transition, and was only detected in strains which produce a high proportion of mycelia in response to serum. No expression of *SAPs 4-6* was found in strains which formed little or no mycelia (Hube *et al.*, 1994). White and Agabian (1995) have also detected *SAP* expression in hyphal cells, in response to pH/temperature regulation of the medium. S1 nuclease analysis suggested that *SAP6* was the dominant transcript under these conditions, with some *SAP5* expression and no detectable *SAP4* mRNA. More recently, *SAP4*, *SAP5* and *SAP6* have all been found to

be expressed in both yeast- and hyphal cells after phagocytosis by murine macrophages (Borg-von Zepelin *et al.*, 1998). These results indicated that the expression of *SAPs 4-6* may not be tightly linked to serum induced germ-tube formation.

This chapter describes *in vitro* studies of *C. albicans SAP4-6* expression with three laboratory strains.

7.2 Expression of *SAPs 4, 5 and 6*

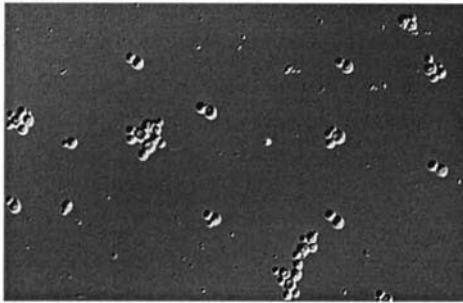
The expression of *SAPs 4, 5 and 6* was tested using two different protocols for serum-induced hyphal formation. The first method involved overnight growth of the yeast in YPD at 30 °C, and resuspension of the cells in 20% (v/v) bovine serum in MQ. Alternatively, the cells were resuspended in YPD containing 20% (v/v) bovine serum. The cell suspensions were then incubated at 37 °C with shaking (200 rpm). Samples were taken at time zero, and after 30 min, 60 min, 90 min, 120 min and 240 min incubation. The two strains used were *C. albicans* ATCC 10231 and *C. albicans* A72 (a clinical isolate kindly provided by Dr. A. Cassone). Cell morphology was examined using Nomarski interference contrast microscopy (Method 2.13, Fig. 7.6). The clinical isolate A72 was selected for use in these experiments because it readily forms hyphae under appropriate conditions and strain ATCC 10231 is well characterised and also forms germ-tubes in serum.

As shown in Fig. 7.1 strain A72 formed hyphae earlier than strain ATCC 10231 in the 20% serum-only medium (compare 60 min and 90 min results), and the strain A72 hyphae are much longer at later time points (compare 120 min and 240 min results). Strain ATCC 10231 formed predominantly pseudomycelia and yeast chains in the first 120 min, whereas strain A72 predominantly formed hyphae over the course of the experiment. In contrast both strains showed an equivalent response to the YPD + 20% serum medium. The restricted ATCC 10231 hyphal transition might be due to limiting nutrients in the serum-only medium, which was overcome by addition of YPD.

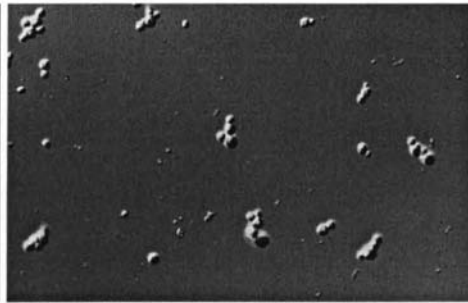
A

Strain ATCC 10231, Serum only

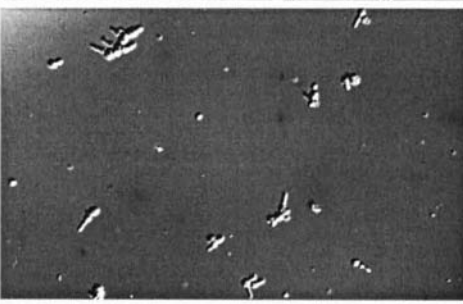
0 min



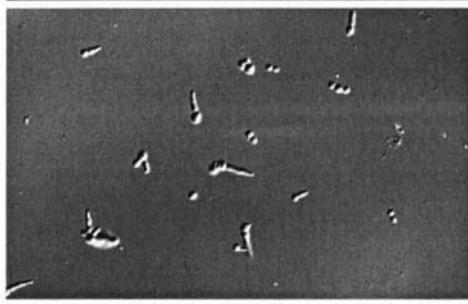
30 min



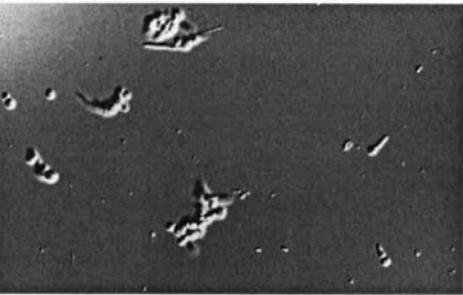
60 min



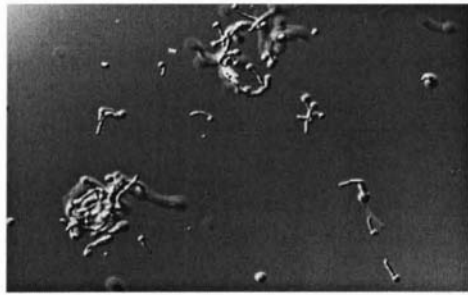
90 min



120 min

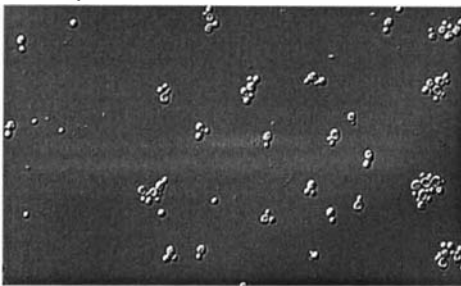


240 min

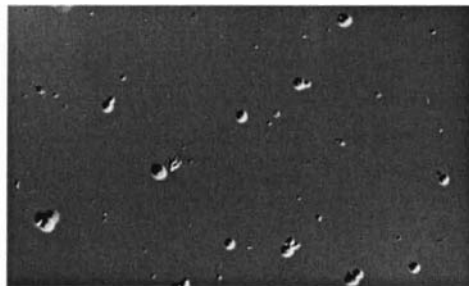


Strain A72, serum only

0 min



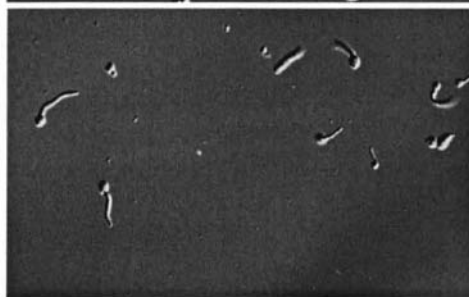
30 min



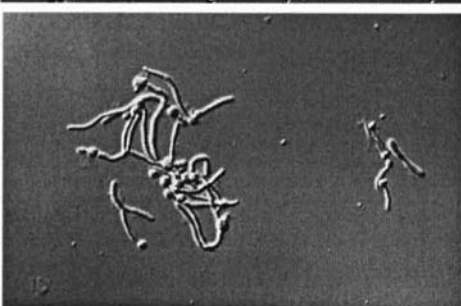
60 min



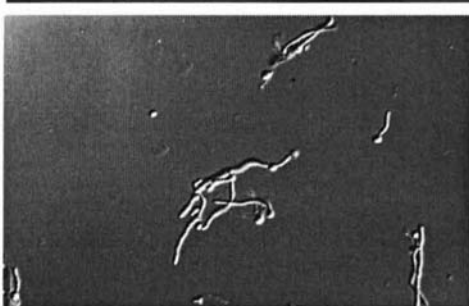
90 min



120 min

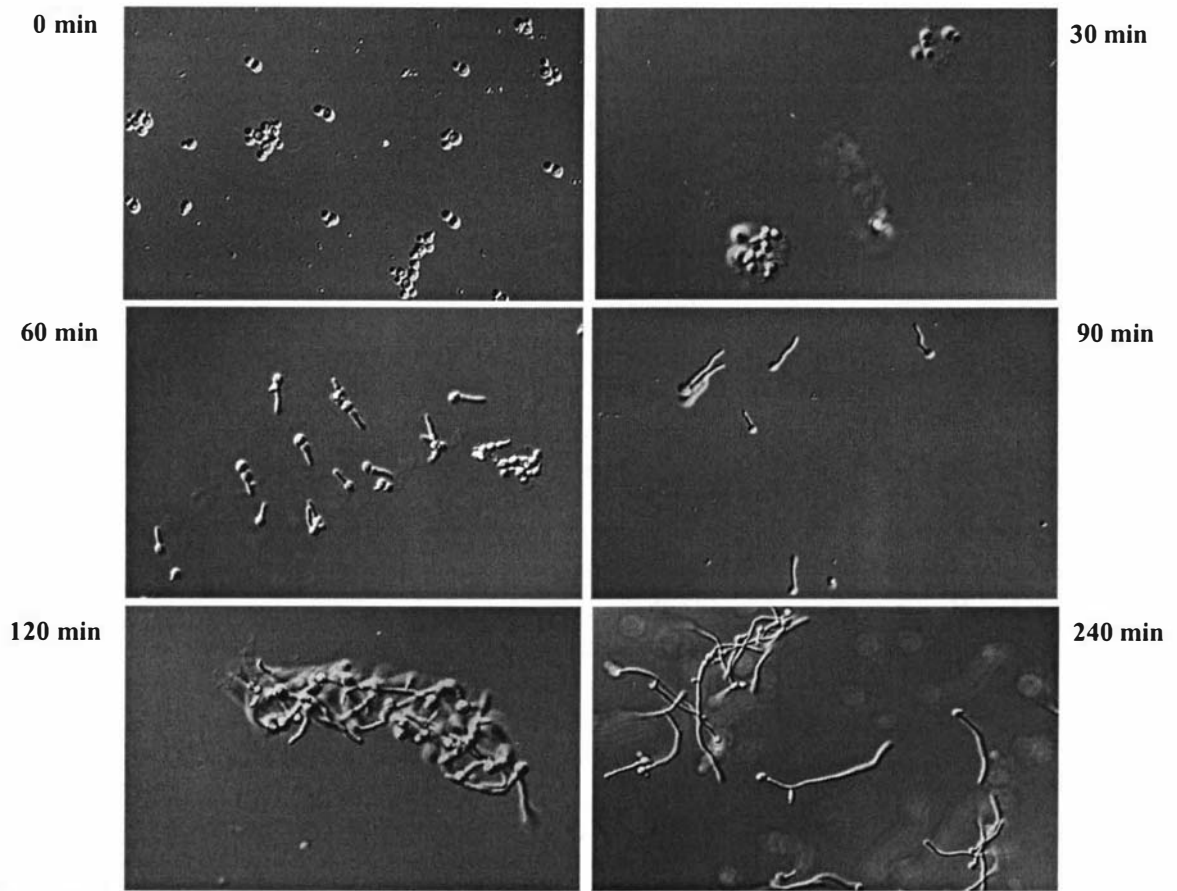


240 min



B

Strain ATCC 10231, YPD + Serum



Strain A72, YPD + Serum

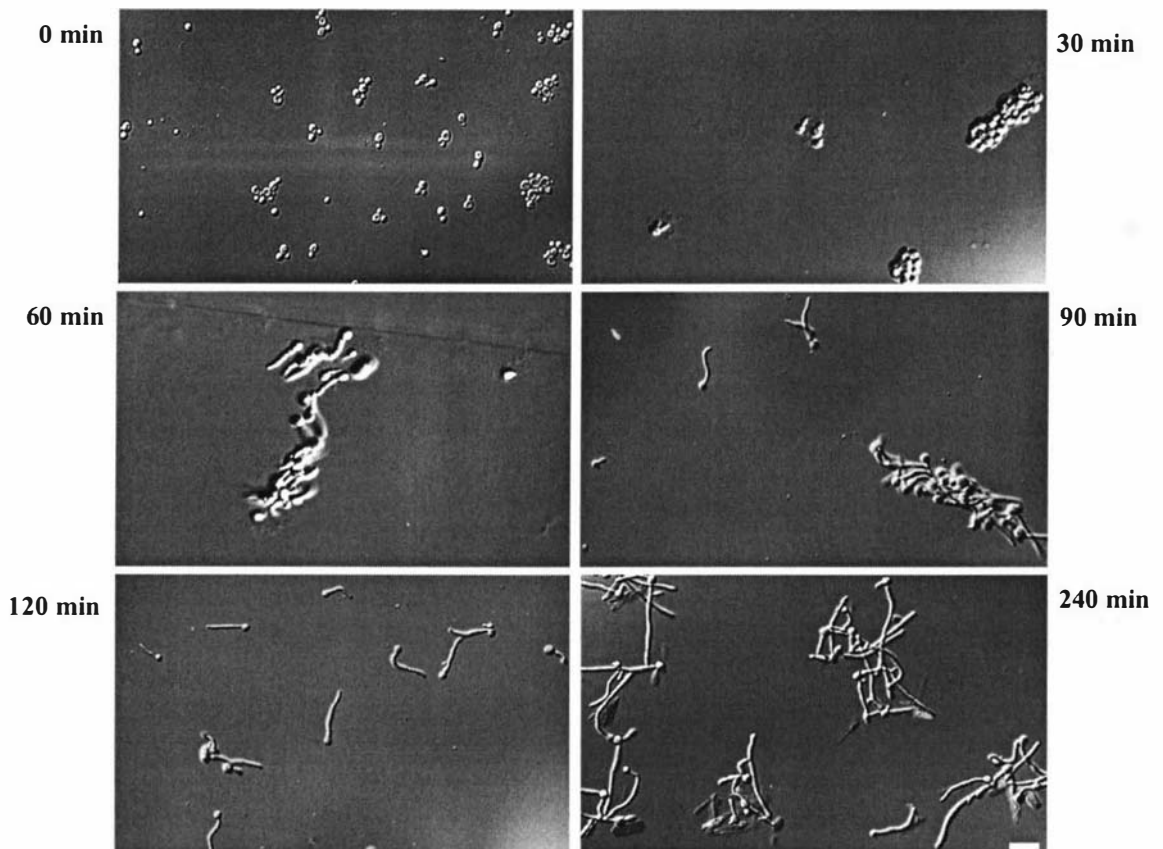


Figure 7.1 **Time course of serum-induced hyphal formation**

Two strains, ATCC 10231 and A72 were induced to form hyphae in 20% serum only (Fig. A) and YPD + 20% serum (Fig. B). The bar at the bottom right hand corner is a 10 μm scale.

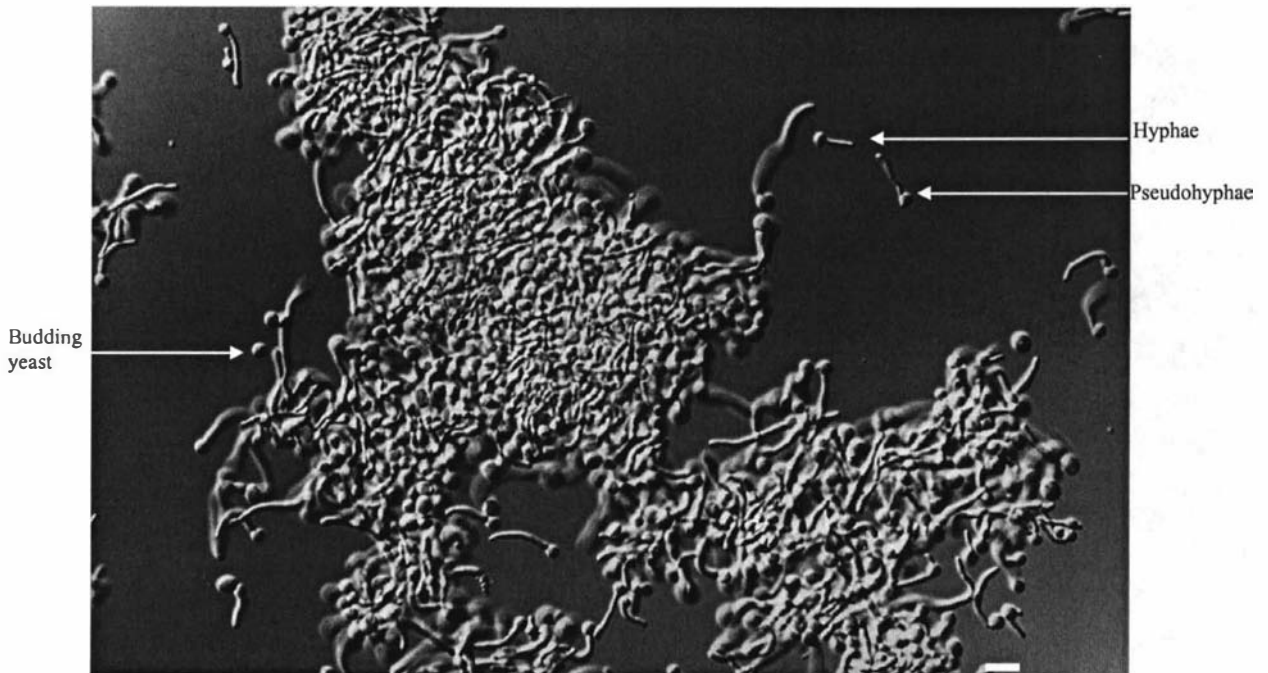


Figure 7.2 **Clumped *C. albicans* strain ATCC 10231 hyphae**

The photograph illustrates the tendency of hyphal cells to flocculate into large clumps. Cells (ATCC 10231) were incubated in YPD + 20% serum for 240 min. Arrows indicate a budding yeast cell, a true hyphal cell and a pseudohypha. Note the constrictions (septae) along the filament of the pseudohypha. The bar at the bottom right hand corner is a 10 μm scale.

One feature of hyphal formation in *C. albicans* is cell flocculation, or cell clumping (Fig. 7.2). It has been proposed that flocculation aids adherence to infection sites, a step which facilitates the disease process (Cutler, 1991). The three common morphological forms of *C. albicans* are also identified in Fig. 7.2, namely budding yeast, pseudohyphae and true hyphae. *SAP 4-6* expression has been associated with the formation of true hyphae, rather than pseudohyphae (Hube *et al.*, 1994).

7.3 Northern analysis of *SAP4-6* expression: strains ATCC 10231 and A72

Northern analyses were performed for the expression of the *SAP4-6* genes, during germ-tube formation. RNA was extracted from strain ATCC 10231 and strain A72 under both induction conditions and at the time points given in Fig. 7.3. Following electrophoresis on formaldehyde-containing agarose gels (Method 2.9.15), the quality of the RNA was assessed under UV light (Fig. 7.3). The RNA was then blotted onto nylon, and probed with radiolabelled full-length *SAP4* cDNA. The various probes used were purified after labelling, to confirm incorporation of radioactive isotope (Methods 2.9.13) and blots were then washed at a moderate stringency (Methods 2.9.12). No bands hybridised to the *SAP4* probe from either strain (Fig. 7.4). The membranes were then stripped and re-probed with the β -actin gene (*ACT*) from *C. albicans* at high stringency. This was successful (Fig. 7.4), and showed that the negative result with the *SAP4* probe was reliable. The membranes were again stripped and re-probed with a 560 bp fragment of *SAP2* to check whether any of the *SAP1 – SAP3* group were expressed. No bands hybridised to this probe (data not shown) indicating that no *SAP* genes were detectably transcribed in the samples tested.

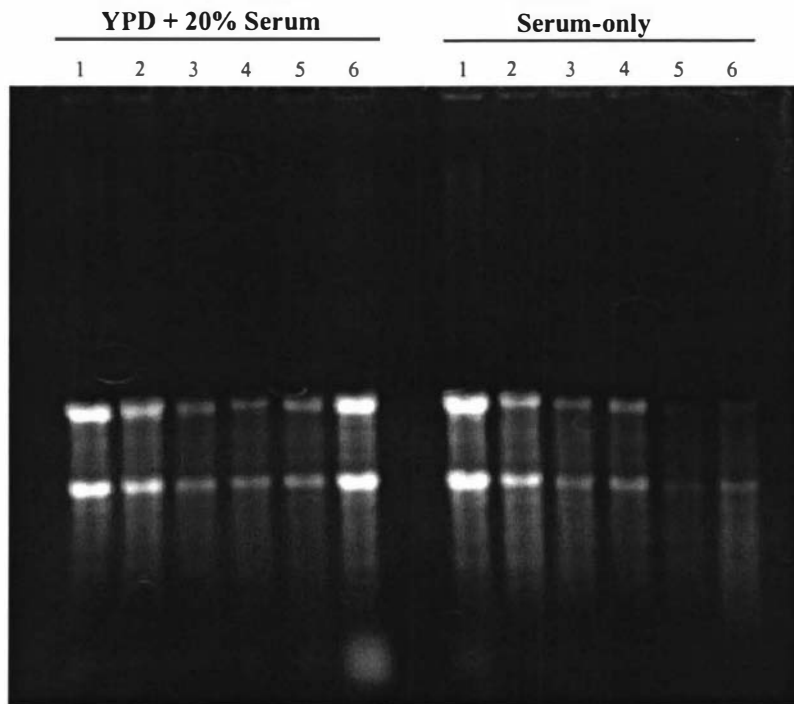


Figure 7.3 *RNA extracted from strain ATCC 10231*

RNA was separated by electrophoresis on a 1.0% formaldehyde-containing agarose gel, and visualised with ethidium bromide. RNA samples were run as follows:

- Lane 1: 0 min
- Lane 2: 30 min
- Lane 3: 60 min
- Lane 4: 90 min
- Lane 5: 120 min
- Lane 6: 240 min

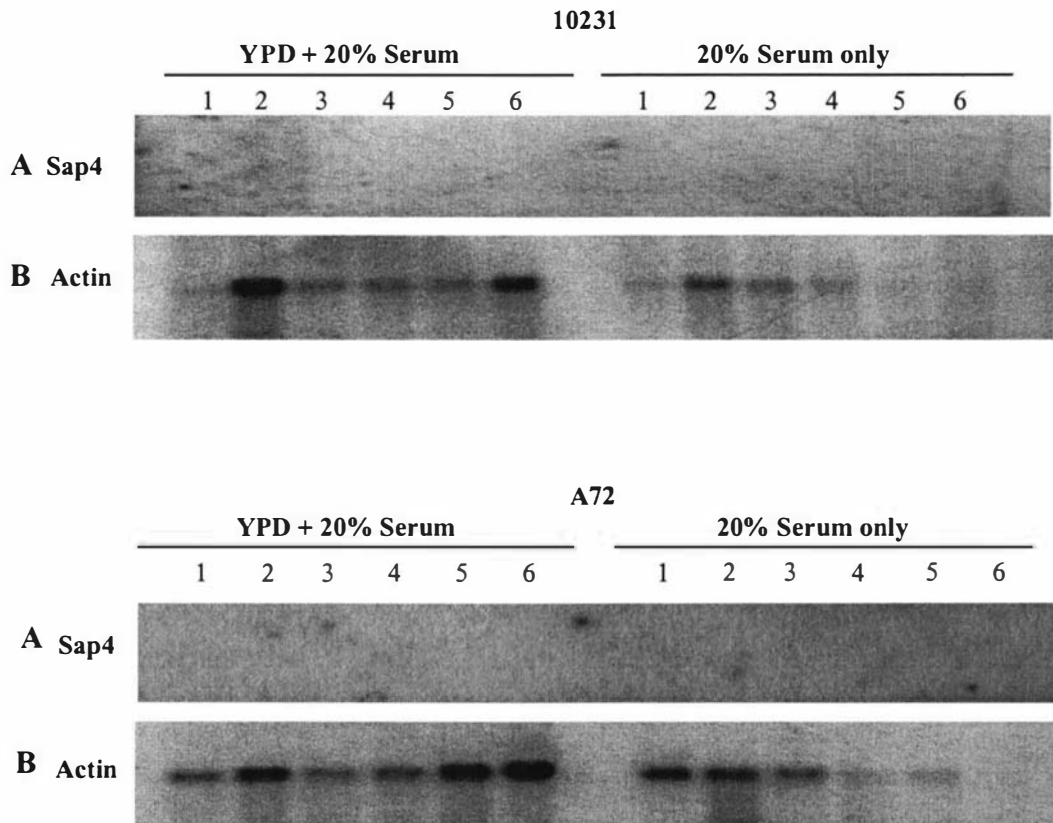


Figure 7.4 Northern hybridisations for *SAP4-6* and *ACT*

Autoradiographs of northern hybridisations from strain ATCC 10231 (A, B) and strain A72 (C, D). Growth conditions and probes are as indicated. Lanes 1-6: Time zero, 30 min, 60 min, 90 min, 120 min, 240 min RNA samples, respectively.

A. An autoradiograph of the ATCC 10231 northern blot (membrane shown in Fig. 7.3) probed with radiolabelled *SAP4* at 60 °C. The blot was washed with 1xSSC/0.1% SDS, at 60 °C and exposed to film for 5 days at -80 °C.

B. An autoradiograph of the northern blot (membrane shown in Fig. 7.3) probed with radiolabelled *ACT* at 65 °C. The blot was washed with 0.1xSSC/0.1% SDS, at 65 °C and exposed to film for 35 h at -80 °C.

C. An autoradiograph of the A72 northern blot, probed with radiolabelled *SAP4* at 60 °C. The blot was washed with 1xSSC/0.1% SDS, at 60 °C and exposed to film for 5 days at -80 °C.

D. An autoradiograph of the A72 northern blot, probed with radiolabelled *ACT* at 65 °C. The blot was washed with 0.1xSSC/0.1% SDS, at 65 °C and exposed to film for 35 hours at -80 °C.

7.4 Northern analysis of *SAP4-6* expression: strain ATCC 10261

A second system for *SAP4-6* expression, using the hyphal induction protocol previously described by Ram *et al.* (1984), was also tested. This involved growing overnight starter cultures of strain ATCC 10261 in YPD, at 30 °C. The cells were harvested, washed under sterile conditions (3 times), resuspended in sterile MQ, and starved for 24 h at 30 °C with shaking (200 rpm). Hyphal formation was then induced in either 5% serum, or N-acetyl glucosamine (2.5 mM GlcNAc, 0.01 M imidazole, 0.1 mM MnCl₂). Yeast cells were resuspended in each medium to a final OD₆₀₀=0.1, and incubated with shaking (200 rpm) at 37 °C. mRNA extractions were performed after 60 min, 90 min and 120 min and percentage germ-tube formation was estimated to be >90% (serum only) and >50% (GlcNAc).

Expression of *SAP4-6* was investigated using a radiolabelled oligonucleotide probe specific to the *SAP4-6* sequences:

5'AAA TTC CCG TCG ATG ATA CTG GTA GAA AT₃'.

The percentage identity of the probe to *SAPs4, 5 & 6* was 97%, 90% and 100%, respectively, whereas identity to *SAPs 1-3* was 70%, 48% and 55%, respectively. The mRNA extracted after 60 min, 90 min and 120 min (serum only) and after 90 min (GlcNAc) was run on a formaldehyde-containing agarose gel (Fig. 7.5, A). Due to the lower percentage germ-tube formation of the N-acetyl glucosamine induction, only the 90 min sample was analysed for *SAP4-6* expression at this stage. The autoradiographs of the northern blot were probed with the *SAP4-6* oligonucleotide (Fig. 7.5, B), but no bands hybridised to the probe. The blots were then stripped, and re-probed with the *C. albicans LEU2* housekeeping gene (Fig. 7.5, C; Cannon *et al.*, 1990). However, due to the 24 h starvation period, the yeast cells were probably in G₀ phase, and *LEU2* mRNA is probably not transcribed under these conditions. Therefore an additional control was performed, using RNA extracted from strain ATCC 10261 cells grown as budding yeast in YPD for 6 h. Fig. 7.5D shows the RNA extracted from these cells expressed *LEU2* mRNA.

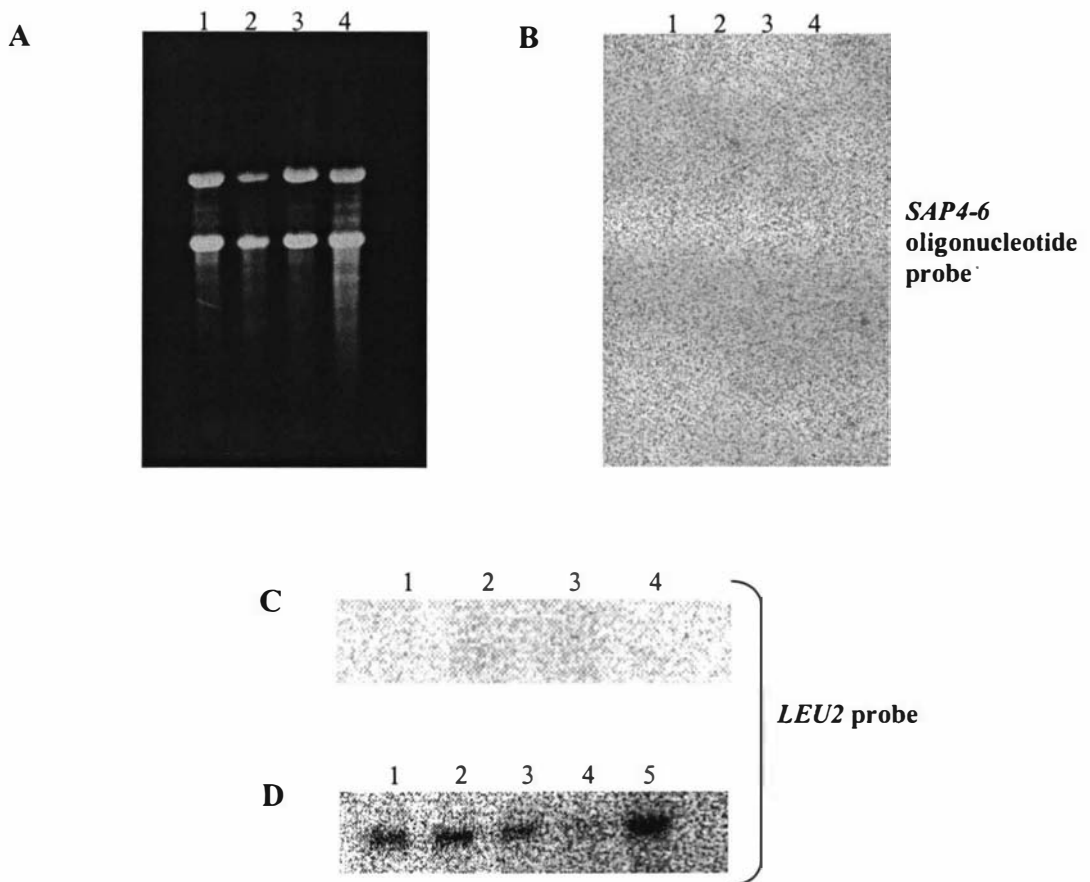


Figure 7.5 *Northern hybridisations for SAP4-6 and LEU2*

Autoradiographs of northern hybridisations from strain ATCC 10261. Growth conditions and probes are as indicated.

A. ATCC 10261 mRNA was separated by electrophoresis on a 1.0% formaldehyde-containing agarose gel, and visualised with ethidium bromide. Lanes 1-4: 60 min, 90 min and 120 min (serum only) and 90 min (GlcNAc), respectively.

B. An autoradiograph of the northern blot (from A) probed with the radiolabelled *SAP4-6* oligonucleotide at 40 °C. The blot was washed with 1xSSC/0.5% SDS, at 40 °C and exposed to film for 5 days at -80 °C.

C. An autoradiograph of the northern blot (from A), probed with radiolabelled *LEU2* at 60 °C. The blot was washed with 1xSSC/0.1% SDS, at 60 °C and exposed to film for 5 days at -80 °C.

D. An autoradiograph of the control northern blot, probed with radiolabelled *LEU2* at 65 °C. The blot was washed with 0.2xSSC/0.5% SDS, at 65 °C and exposed to film for 22 hours at -80 °C. Lanes 1-5: 1 h, 2 h, 3 h, 4 h and 5 h RNA samples from ATCC 10261 grown in rich medium.

7.5 *In vitro* expression analyses

The results shown in Fig. 7.4 and Fig. 7.5 suggested that *SAP4-6* expression was not universally linked to serum-induced hyphal formation, as suggested by Hube *et al.* (1994). Several factors may have contributed to this finding, including the source of the serum, the medium used for growing the starter cultures, or other more subtle effects. Three laboratory strains were tested; ATCC 10231, ATCC 10261 and A72, and four different hyphal induction media were used. No *SAP4-6* expression was seen under any of these conditions, although control northern blots were successful (Fig. 7.4D, Fig. 7.5 D).

Hube *et al.* (1994) used several clinical isolates in addition to laboratory strains 3153 and ATCC 48867, and both the clinical isolates and the laboratory strains produced *SAP4-6* in response to serum-induced hyphal formation. However, no expression of the *SAP4-6* group was seen using other methods of germ-tube formation, such as pH/temperature regulation or N-acetyl glucosamine induction. In contrast to these findings, White and Agabian (1995) showed expression of *SAP5* and *SAP6*, but not *SAP4*, in hyphal cells induced by pH/temperature regulation and serum. *SAP5-6* expression was detected in the *C. albicans* laboratory strain SS, and the *C. albicans* clinical isolate 92-4053. The results from the present study are not consistent with the previous studies. Taken together, this suggests that significant variation exists between *in vitro* experiments, which are affecting the outcome of the analyses.

Until recently, expression of *SAP4-6* had only been documented in hyphal cells *in vitro*, but expression of *SAPs* 4, 5 and 6 has now been detected in yeast cells, in an *in vivo* system (Borg-von Zepelin *et al.*, 1998). Using fresh mouse peritoneal macrophages and immunofluorescence, high level production of Sap4, Sap5 and Sap6 was detected after phagocytosis of *C. albicans* by the macrophages. This suggests that a more physiologically relevant system is more suitable for reliable expression studies. Further, Monod *et al.* (1998) have recently analysed the expression of *SAP8* and *SAP9* in *C. albicans*. *In vitro* studies suggested that *SAP8* was expressed preferentially at 25 °C, although transient expression was detected in the early exponential growth phase at 37 °C, but this disappeared within 3 h. These data indicated that it would be unlikely that this gene would play an important role during infection at the body temperature of the

host. However, *SAP8* expression has subsequently been detected in models of oral candidosis (Schaller *et al.*, 1999; Monod *et al.*, 1998). This shows that extrapolating *in vitro* results to the *in vivo* situation is inherently unreliable when the *in vitro* models do not adequately mimic the *in vivo* situation.

8.0 CRYSTALLISATION TRIALS

8.1 Introduction

The synthesis of Saps during the onset and development of candidoses (Borg-von Zepelin *et al.*, 1998) highlights their potential as drug targets. However, because the Sap family of proteins comprises at least nine members, it will be necessary to analyse the specificity of each enzyme before attempting any rational drug design. To this end, recombinant Sap4 and Sap5 were successfully purified in milligram quantities in this project for further analysis.

While biochemical characterisation has given some insight into the substrate specificity of Sap4 and Sap5, crystallographic studies would permit an understanding of the structural basis of this specificity. The first stage in pursuing a protein structure is the production of diffraction-quality crystals of the protein. The first approach for protein crystallisation depends on whether the protein is novel, or is similar to another protein which has already been crystallised. However, while the conditions under which a similar protein was crystallised might provide a starting point for suitable conditions, it is not common for two related proteins crystallise under identical conditions.

Therefore it is usual to subject the protein to one or more standard “crystal screens”, which test many different parameters (such as pH, buffer, precipitant, salt) in a factorial way. This chapter describes the crystallisation of recombinant Sap1, Sap4 and Sap5. The crystallisation trials with Sap1 and Sap4 were performed at the University of Otago in collaboration with Dr. S. M. Cutfield, while the work on Sap5 was performed at Massey University.

8.2 Crystallisation of Sap1

As Sap2 had already been crystallised and the structure solved (Cutfield *et al.*, 1995), the conditions under which it crystallised were used as a starting point for crystal trials in addition to the standard crystallisation screens. Typically, the *C. albicans* Saps do not form good crystals in the absence of a potent inhibitor, such as the compound A70450 (a synthetic hexapeptide analogue; Capobianco *et al.*, 1992; Jae *et al.*, 1992). This is due (in part) to the propensity of the Sap enzymes to undergo autolysis over time (Cutfield *et al.*, 1993). Therefore, while crystallisation trials were performed with recombinant Sap1 alone, parallel trials included the A70450 inhibitor.

Sap1 was purified by Mono Q anion exchange chromatography (refer Chapter 5.3.4 for details) immediately prior to concentration by ultrafiltration. Once the protein concentration reached approximately 5 mg/mL, it was buffer-exchanged into MQ water to reduce any interference from the buffer in the crystal trials. The hanging-drop vapour-diffusion method was used for crystallisation trials (Method 2.12), with the sparse matrix crystal screen (Jancarik & Kim, 1991; Table 8.1) in addition to some previously successful conditions for crystallisation of Sap2 (Table 8.3). Drops (2 μ L) contained protein at 2.5 mg/mL were incubated at 17 °C. Inspection of the drops after 4 days revealed that both recombinant Sap1 alone, and Sap1 + A70450 produced tiny crystalline shapes (microcrystals) under most conditions, suggesting that crystallisation was possible at the chosen protein concentration.

Well	Precipitant	Sap1		Sap4		Sap5	
		-	+	-	+	-	+
1	30% MPD, 0.1 M sodium acetate pH 4.6, 0.02 M calcium chloride	m	m	-	-	m	m
2	0.4 M potassium-sodium tartrate	m	-	-	-	-	-
3	0.4 M ammonium phosphate	m	-	-	-	m	m
4	2.0 M ammonium sulfate, 0.1 M Tris-HCl pH 8.5	m	-	-	-	m	p
5	30% MPD, 0.1 M sodium hepes pH 7.5, 0.2 M sodium citrate	pt	pt	-	-	pt	-
6	30% PEG 4000, 0.1 M Tris-HCl pH 8.5, 0.2 M magnesium chloride	pt	pt	-	-	m	-
7	1.4 M sodium acetate, 0.1 M sodium cacodylate pH 6.5	m	m	-	-	pt	p
8	30% 2-propanol, 0.1 M sodium cacodylate pH 6.5, 0.2 M sodium citrate	pt	pt	pt	pt	pt	p
9	30% PEG 4000, 0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate	pt	m	m	m	pt	-
10	30% PEG 4000, 0.1 M sodium acetate pH 4.6, 0.2 M ammonium acetate	pt	m	-	m	pt	p
11	1.0 M ammonium phosphate, 0.1 M sodium citrate pH 5.6	m	m	-	-	-	m
12	30% 2-propanol, 0.1 M sodium hepes pH 7.5, 0.2 M magnesium chloride	pt	pt	pt	m	m	p
13	30% PEG 4000.1 M Tris-HCl pH 8.5, 0.2 M sodium citrate	m	m	-	-	-	-
14	28% PEG 400, 0.1 M sodium hepes pH 7.5, 0.2 M calcium chloride	m	m	-	m	-	m
15	30% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M ammonium sulfate	pt	m	-	m	m	-
16	1.5 M lithium sulfate, 0.1 M sodium hepes pH 7.5	m	m	-	-	m	p
17	30% PEG 4000, 0.1 M Tris-HCl pH 8.5, 0.2 M lithium sulfate	m	m	m	m	pt	p
18	20% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M magnesium acetate	m	m	m	m	m	C
19	30% 2-propanol, 0.1 M Tris-HCl pH 8.5, 0.2 M ammonium acetate	m	pt	pt	m	-	p
20	25% PEG 4000, 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate	m	m	-	m	pt	-
21	30% MPD, 0.1 M sodium cacodylate pH 6.5, 0.2 M magnesium acetate	pt	-	m	m	pt	m
22	30% PEG 4000, 0.1 M Tris-HCl pH 8.5, 0.2 M sodium acetate	pt	pt	pt	-	pt	-
23	30% PEG 400, 0.1 M sodium hepes pH 7.5, 0.2 M magnesium chloride	-	m	-	-	m	m
24	20% 2-propanol, 0.1 M sodium acetate pH 4.6, 0.2 M calcium chloride	m	m	m	m	-	m
25	1.0 M sodium acetate, 0.1 M imidazole pH 6.5	m	m	-	-	pt	p
26	30% MPD, 0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate	pt	m	m	-	pt	-
27	20% 2-propanol, 0.1 M sodium hepes pH 7.5, 0.2 M sodium citrate	pt	pt	m	m	-	-
28	30% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M sodium acetate	pt	-	pt	m	m	-
29	0.8 M potassium-sodium tartrate, 0.1 M sodium hepes pH 7.5	-	m	m	m	-	m
30	30% PEG 8000, 0.2 M ammonium sulfate	m	m	pt	pt	m	-
31	30% PEG 4000, 0.2 M ammonium sulfate	pt	D	pt	pt	m	p
32	2.0 M ammonium sulfate	m	m	m	m	m	p
33	4.0 M sodium formate	pt	pt	pt	m	-	m
34	2.0 M sodium formate, 0.1 M sodium acetate pH 4.6	m	m	pt	m	-	m
35	1.6 M sodium-potassium phosphate, 0.1 M sodium hepes pH 7.5	-	-	m	m	pt	p
36	8% PEG 8000, 0.1 M Tris-HCl pH 8.5	m	m	-	-	-	-
37	8% PEG 4000, 0.1 M sodium acetate pH 4.6	m	m	-	m	-	p
38	1.4 M sodium citrate, 0.1 M sodium hepes pH 7.5	m	-	pt	-	pt	p
39	2% PEG 400, 2.0 M ammonium sulfate, 0.1 M sodium hepes pH 7.5	m	-	m	m	-	-
40	20% 2-propanol, 20% PEG 4000, 0.1 M sodium citrate pH 5.6	pt	m	-	m	m	C
41	10% 2-propanol, 20% PEG 4000, 0.1 M sodium hepes pH 7.5	pt	pt	m	m	m	p
42	20% PEG 8000, 0.05 M potassium phosphate	m	m	-	-	pt	p
43	30% PEG 1500	m	m	-	-	-	p
44	0.2 M magnesium formate	m	-	-	-	-	m
45	18% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate	-	-	pt	C	m	p
46	18% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M calcium acetate	m	m	-	m	m	m
47	2.0 M ammonium sulfate, 0.1 M sodium acetate pH 4.6	m	-	m	-	pt	-
48	2.0 M ammonium phosphate, 0.1 M Tris-HCl pH 8.5	m	-	-	-	pt	p
49	2% PEG 8000, 1.0 M lithium sulfate	-	-	-	-	-	-
50	15% PEG 8000, 0.5 M lithium sulfate	-	-	-	-	-	-

Table 8.1 Sparse matrix crystal screen

The table lists the precipitants, and the left-hand columns indicate whether Sap1, Sap4 and Sap5 formed precipitates (pt), microcrystals (m) or showed no change (-). Each enzyme was tested in the presence (+) or absence (-) of inhibitor A70450 (Sap1, 4) or pepstatin A (Sap5) and the columns are labelled as such. Crystals are indicated by either C (large enough to mount) or D (diffraction quality).

Well #	Precipitant:	Sap5 alone
1	10% PEG 6000, 2.0 M sodium chloride	-
2	0.5 M sodium chloride, 0.01 M CTAB, 0.01 M magnesium chloride	-
3	25% ethylene glycol	-
4	35% dioxane	pt
5	5% isopropanol, 2.0 M ammonium sulfate	-
6	1.0 M imidazole pH 7.0	-
7	10 % PEG 1000, 10 % PEG 8000	-
8	10% ethanol, 1.5 M sodium chloride	-
9	2.0 M sodium chloride, 0.1 M sodium acetate pH 4.6	-
10	30% MPD, 0.1 M sodium acetate pH 4.6, 0.2 M sodium chloride	pt
11	1.0 M 1,6-hexanediol, 0.1 M sodium acetate pH 4.6, 0.01 M cobalt chloride	-
12	30% PEG 400, 0.1 M sodium acetate pH 4.6, 0.1 M cadmium chloride	pt
13	30% PEGMME 2000, 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate	pt
14	2.0 M ammonium sulfate, 0.1 M sodium citrate pH 5.6, 0.2 M K/Na tartrate	pt
15	1.0 M lithium sulfate, 0.1 M sodium citrate pH 5.6, 0.5 M ammonium sulfate	-
16	2% polyethyleamine, 0.1 M sodium citrate pH 5.6, 0.5 M sodium chloride	pt
17	35% tert-butanol, 0.1 M sodium citrate pH 5.6	pt
18	10% jeffamine M-600, 0.1 M sodium citrate pH 5.6, 0.01 M ferric chloride	pt
19	2.5 M 1,6-hexanediol, 0.1 M sodium citrate pH 5.6	-
20	1.6 M magnesium sulfate, 0.1 M MES pH 6.5	-
21	2.0 M sodium chloride, 0.1 M MES pH 6.5, 0.2 M Na/K phosphate	-
22	12% PEG 20,000, 0.1 M MES pH 6.5	-
23	10% dioxane, 0.1 M MES pH 6.5, 1.6 M ammonium sulfate	m
24	30% jeffamine M-600, 0.1 M MES pH 6.5, 0.05 M cesium chloride	m
25	1.8 M ammonium sulfate, 0.1 M MES pH 6.5, 0.01 M cobalt chloride	-
26	30% PEGMME 5000, 0.1 M MES pH 6.5, 0.2 M ammonium sulfate	m
27	25% PEGMME 550, 0.1 M MES pH 6.5, 0.01 M zinc sulfate	-
28	1.6 M sodium citrate pH 6.5	-
29	30% MPD, 0.1 M hepes pH 7.5, 0.5 M ammonium sulfate	m
30	10% PEG 6000, 0.1 M hepes pH 7.5, 5% MPD	pt
31	20% jeffamine M-600, 0.1 M hepes pH 7.5	-
32	1.6 M ammonium sulfate, 0.1 M hepes pH 7.5, 0.1 M sodium chloride	-
33	2.0 M ammonium formate, 0.1 M hepes pH 7.5	pt
34	1.0 M sodium acetate, 0.1 M hepes pH 7.5, 0.05 M cadmium sulfate	-
35	70% MPD, 0.1 M hepes pH 7.5	m
36	4.3 M sodium chloride, 0.1 M hepes pH 7.5	m
37	10% PEG 8000, 0.1 M hepes pH 7.5, 8% ethylene glycol	-
38	20% PEG 10,000, 0.1 M hepes pH 7.5	pt
39	3.4 M 1,6-hexanediol, 0.1 M Tris-HCl pH 8.5, 0.2 M magnesium chloride	-
40	25% tert-butanol, 0.1 M Tris-HCl pH 8.5	pt
41	1.0 M lithium sulfate, 0.1 M Tris-HCl pH 8.5, 0.01 M nickel chloride	m
42	12% glycerol, 0.1 M Tris-HCl pH 8.5, 1.5 M ammonium sulfate	-
43	50% MPD, 0.1 M Tris-HCl pH 8.5, 0.2 M ammonium phosphate	m
44	20% ethanol, 0.1 M Tris-HCl pH 8.5	pt
45	20% PEGMME 2000, 0.1 M Tris-HCl pH 8.5, 0.01 M nickel chloride	pt
46	30% PEGMME 550, 0.1 M bicine pH 9.0, 0.1 M sodium chloride	m
47	2.0 M magnesium chloride, 0.1 M bicine pH 9.0	pt
48	10% PEG 20,000, 0.1 M bicine pH 9.0, 0.2% dioxane	pt

Table 8.2 *Crystal screen II*

The table describes the conditions of crystal screen II (Cudney *et al.*, 1994). The left-hand columns list the conditions under which Sap5 (alone) was either precipitated (ppt.), formed small crystals (C). A dash (-) indicates no change. No diffraction-quality crystals were obtained using this screen.

Well #	Precipitant:	Sap1		Sap4	
		-	+	-	+
1	17% PEG 3350, 0.035 M imidazole-malate pH 6.5, 1.7 mM zinc acetate	pt		m	m
2	19% PEG 3350, 0.035 M imidazole-malate pH 6.5, 1.7 mM zinc acetate	pt	m	m	m
3	21% PEG 3350, 0.035 M imidazole-malate pH 6.5, 1.7 mM zinc acetate	m	m	m	m
4	23% PEG 3350, 0.035 M imidazole-malate pH 6.5, 1.7 mM zinc acetate	pt	pt	pt	pt
5	12.5% MPD, 0.05 M sodium acetate pH 4.6, 0.02 M calcium chloride	m	m	m	m
6	16% MPD, 0.05 M sodium acetate pH 4.6, 0.02 M calcium chloride	m	m	m	m
7	12.5% PEG 600, 0.05 M sodium acetate pH 4.6, 0.02 M calcium chloride	-	m	m	m
8	16% PEG 600, 0.05 M sodium acetate pH 4.6, 0.02 M calcium chloride	-	m	m	m
9	2.3 M ammonium sulfate, 0.05 M sodium acetate pH 4.6, 0.6 M sodium chloride	-	pt	pt	pt
10	2.4 M ammonium sulfate, 0.05 M sodium acetate pH 4.6, 0.6 M sodium chloride	pt	pt	pt	pt
11	2.5 M ammonium sulfate, 0.05 M sodium acetate pH 4.6, 0.6 M sodium chloride	m	m	pt	pt
12	2.6 M ammonium sulfate, 0.05 M sodium acetate pH 4.6, 0.6 M sodium chloride	pt	pt	pt	pt

Table 8.3 *Additional crystallisation conditions for Sap1 and Sap4*

The conditions described above were tested for crystallisation of Sap1 and Sap4 at the same time as the sparse matrix screen. Each enzyme was tested in the presence (+) or absence (-) of inhibitor A70450 and the columns are labelled as such. These trials were based on previously successful crystallisation conditions used for Sap-inhibitor complexes (Dr. S. M. Cutfield, personal communication). The left-hand columns show the conditions under which Sap1 and Sap4 (+/- A70450) formed precipitates (pt), microcrystals (m) or showed no change (-).

While most conditions resulted in microcrystals, Sap1 complexed with A70450 produced small but well-formed rhombic crystals in the presence of 30% PEG 4000 and 0.2 M ammonium sulfate (sparse matrix screen - well no. 31, Table 8.1). The largest of these crystals (shown in Fig. 8.1) was mounted and photographed using a precession camera (performed by Dr. S. M. Cutfield, University of Otago). The results from this technique showed that the crystal diffracted, and confirmed it was protein, rather than salt.

8.3 Crystallisation of Sap4

Crystallisation trials with Sap4 were identical to the Sap1 trials, with Sap4 concentrated to approximately 5 mg/mL and buffer-exchanged into MQ water. The hanging drop vapour diffusion method was used, again with 2 μ L drops. The same crystallisation conditions were tested (Table 8.1 and Table 8.3), and inspection of the drops after 4 days (at 17 °C) revealed that both recombinant Sap4 alone, and Sap4 + A70450 produced microcrystals under many conditions.

While most conditions tested resulted in nothing more than microcrystals, the precipitant 18% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate produced small but promising crystals of Sap4 + A70450 (well no. 45, Table 8.1). However, these crystals (Fig. 8.2) were too small for diffraction analysis. In addition, larger misshapen crystals were produced in similar conditions (18% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M calcium acetate, well no. 46, sparse matrix screen).

Sap4 complexed with SQAPI (the squash aspartic proteinase inhibitor; Christeller *et al.*, 1998) has also been crystallised (Dr. S. M. Cutfield, personal communication), and the crystals are shown in Fig. 8.2. Unfortunately, while these crystals are large, they did not diffract and more work will be required to produce similar, diffraction-quality crystals of Sap4/inhibitor complexes.

Crystals of uncomplexed Sap4 were also obtained, in a precipitant solution of 10% 2-propanol, 20% PEG 4000 and 0.1 M sodium hepes pH 7.5 (data not shown). These crystals were too small to mount and hence it is not known whether they will diffract. Previously, crystallisation of native Sap enzymes has been problematic, largely due to the propensity of Saps to undergo autodegradation (Cutfield *et al.*, 1993). Therefore these preliminary native Sap4 crystals were surprising. However, the lack of autodegradation exhibited by Sap4 (refer Chapter 6.7), might account for this relative success. Further work was done to optimise the growth of these crystals, but without success.

8.4 Crystallisation of Sap5

Pepstatin A was used in the Sap5 crystallisation trials in place of A70450, which was unavailable at the time. Sap5 protein was concentrated to 8 mg/mL and buffer exchanged into MQ. The hanging-drop vapour-diffusion method was used and the sparse matrix screen was performed in the presence and absence of Pepstatin A. In addition, a second crystallisation screen (crystal screen II) was carried out, to investigate some more unusual precipitants. Sap5 did not produce crystals of the quality of Sap1 and Sap4 using these screens, but nevertheless some progress was made towards identifying suitable crystallisation conditions for this enzyme.

After 60 days at room temperature, inspection of the trays revealed that Sap5 either precipitated or formed small crystalline shapes in many of the precipitants tested (Table 8.1, Table 8.2). However, of the larger crystals which formed most were bundles of tiny rods and needles (Fig. 8.3A). Some individual crystals were grown but these were small and misshapen (Fig. 8.3B). One large crystal grew in 50% MPD, 0.1 M Tris-HCl pH 8.5 and 0.2 M ammonium phosphate (Fig. 8.3C) although again this crystal was misshapen. None of these crystals were mounted for x-ray analysis to determine whether they were protein or salt.

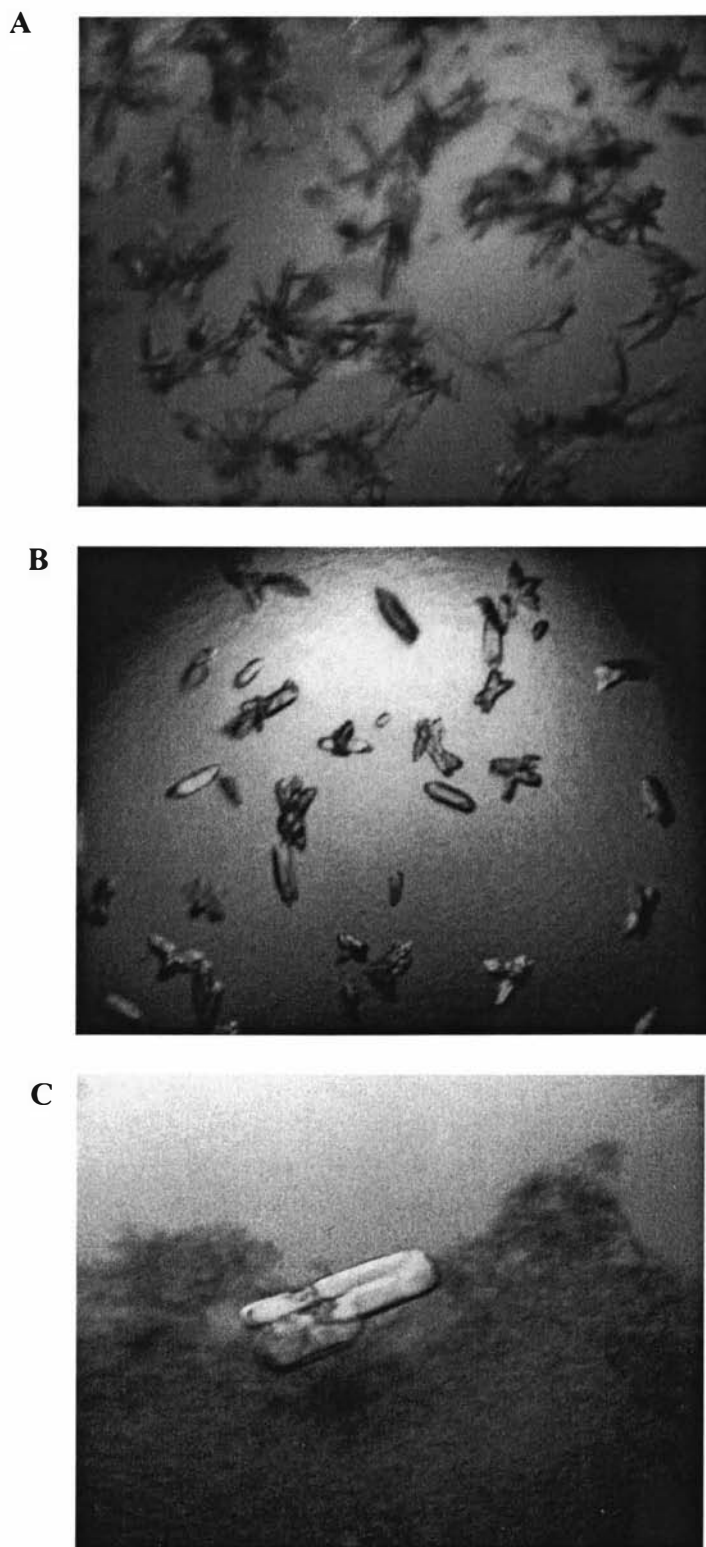
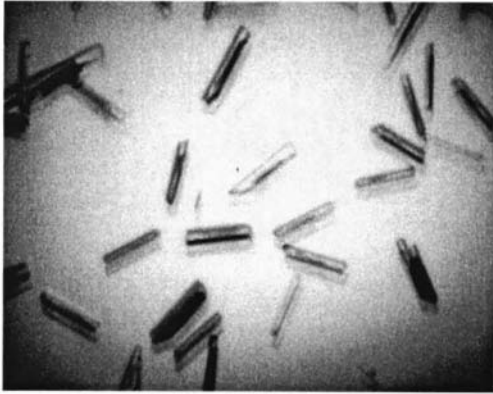


Figure 8.1 *Crystals of native Sap5 from the sparse matrix screen and crystal screen II*

- A. Typical irregular needle & plate crystals. Photographed from well no. 40, but also found in wells 26, 29, 35 and 36 from screen II and 6, 15, 18, 26, 30, 31, 40, 41 45 and 46 (sparse matrix).
- B. Better quality crystals than in photo A, but still irregular shaped and very tiny (well no. 31, sparse matrix).
- C. The largest crystal grown (well no. 43, crystal screen II), but still irregularly shaped. Magnification = 40x

A



B



Figure 8.2 *Crystals of Sap4 complexed with A70450 inhibitor*

- A. Two photos of Sap4 complexed with A70450 inhibitor, grown in 18% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate (well no. 45, sparse matrix crystal screen).
- B. Crystals of Sap4 complexed with A70450 inhibitor, grown in 18% PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M calcium acetate (well no. 46, sparse matrix crystal screen).
Magnification = 40x

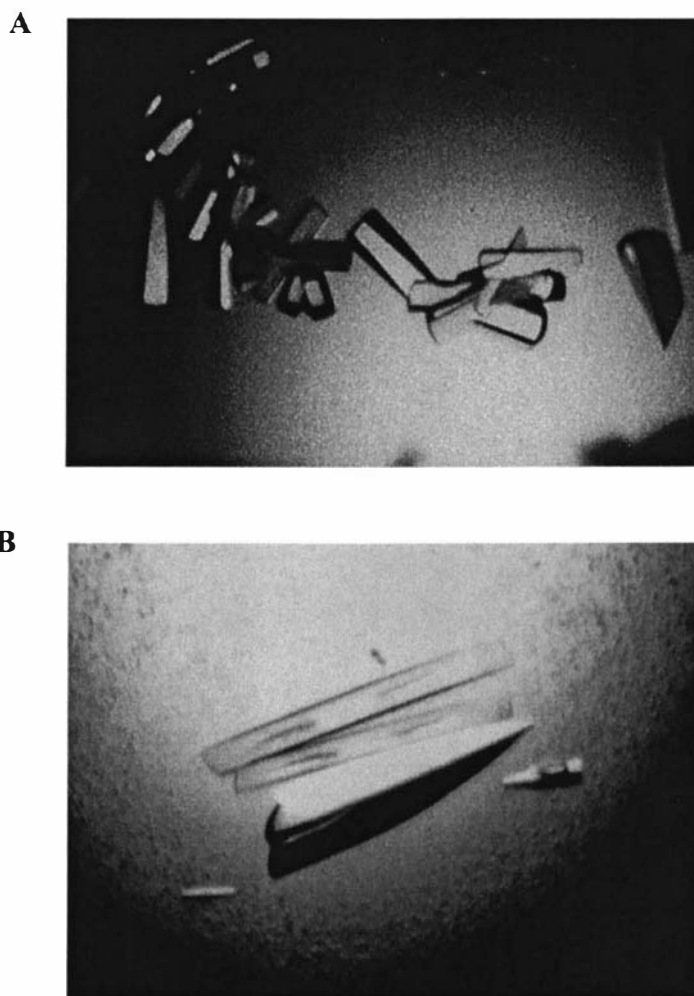


Figure 8.3 *Crystals of Sap4 complexed with SQAPI inhibitor*

- A. Crystals of Sap4 complexed with SQAPI inhibitor
- B. Crystals of Sap4 complexed with SQAPI inhibitor
Magnification = 40x

8.5 Homology modelling of Saps 1, 4 and 5

Homology modelling is now frequently performed to assess either the predicted fold of a protein, or to deduce the structural basis for specificity-function relationships. However, modelling the predicted structure of a protein does not substitute for determination of the structure by x-ray analysis or NMR spectroscopy. Further, modelling is limited by the availability of structures for homologous proteins. Nevertheless, as diffraction-quality crystals were not produced during the course of this work, homology modelling was performed in an attempt to understand the basis for the observed differences in specificity of the Sap isoenzymes.

Until recently, the availability of tools for comparative protein modelling has been limited. However, the advent of freely available Netscape-based homology modelling programs, such as the SWISS MODEL and Swiss-PdbViewer (SPdbV, Guex & Peitsch, 1997: <http://www.expasy.ch/spdbv/mainpage.htm>) has provided useful tools and data bases. These programs were used to produce homology models of the recombinant Sap isoenzymes. SWISS MODEL is a server for automated comparative protein modelling, which takes the input protein sequences, generates templates for homology modelling from Protein Data Bank (Pdb) entries, and threads the input sequence onto these known structures. SPdbV is the Macintosh/PC-based program used to view the resulting models.

Initially, the first approach method was used to predict the structures of mature Saps 1, 4 and 5, using the templates generated by SWISS MODEL. However, this was unsuccessful, and the program failed to produce any comparative models. This was probably due to the inclusion of Pdb templates from the closest structural relatives to Sap2 (rhizopuspepsin among the fungal enzymes and porcine pepsin from the mammalian proteases) in the modelling process. While the *C. albicans* Saps are members of the aspartic proteinase superfamily, they present several unique features in both the amino and carboxyl lobes that place them into a unique sub-class of the monomeric aspartic proteinases (Abad-Zapatero *et al.*, 1998).

Therefore, subsequent modelling was based on the structure of mature Sap2 (Fig. 8.4A) as a template. The active site inhibitor, A70450, was removed prior to modelling, as

inclusion of the inhibitor in the model generated some impossible contacts between atoms, and caused the iterative modelling process to fail. Removing of the A70450 inhibitor co-ordinates resulted in successful models, and the predicted structures are shown in Fig. 8.4B, Fig. 8.5 and Fig. 8.6.

To compare the active sites of the protein models, the active site residues (Asp37, Asp218 and Tyr 84, which stabilises the tetrahedral intermediate during catalysis) were superimposed onto the Sap2 structure (Fig. 8.4B). Superposition of structures can be performed by fitting the C α backbone, the side chains, or using all the atoms in each structure; in this case, the C α backbone was used. Analysis of the overlaid structures showed no significant variations between the Sap2 active site residues and the Sap1, Sap4 or Sap5 active sites. Some structural differences were seen in the region Thr210-Val215 (Fig. 8.5, Fig. 8.6), which is significantly removed ($>15\text{\AA}$) from the substrate binding cleft. These variations would not be expected to affect the primary substrate specificity of the enzymes due to their surface location, and their distance from the active site. In addition, they do not correspond to any of the regions of significant difference between Sap2 and the other monomeric aspartic proteinases (Cutfield *et al.*, 1995).

While homology modelling is often a useful tool, the basis for enzyme specificity, or subtle structural differences between enzymes cannot necessarily be predicted, as seen in this instance. Therefore, high-resolution x-ray structure determination should be pursued, and this will require diffraction-quality crystals. Given the success of the crystallisation trials described here, future trials should be based on these preliminary data.

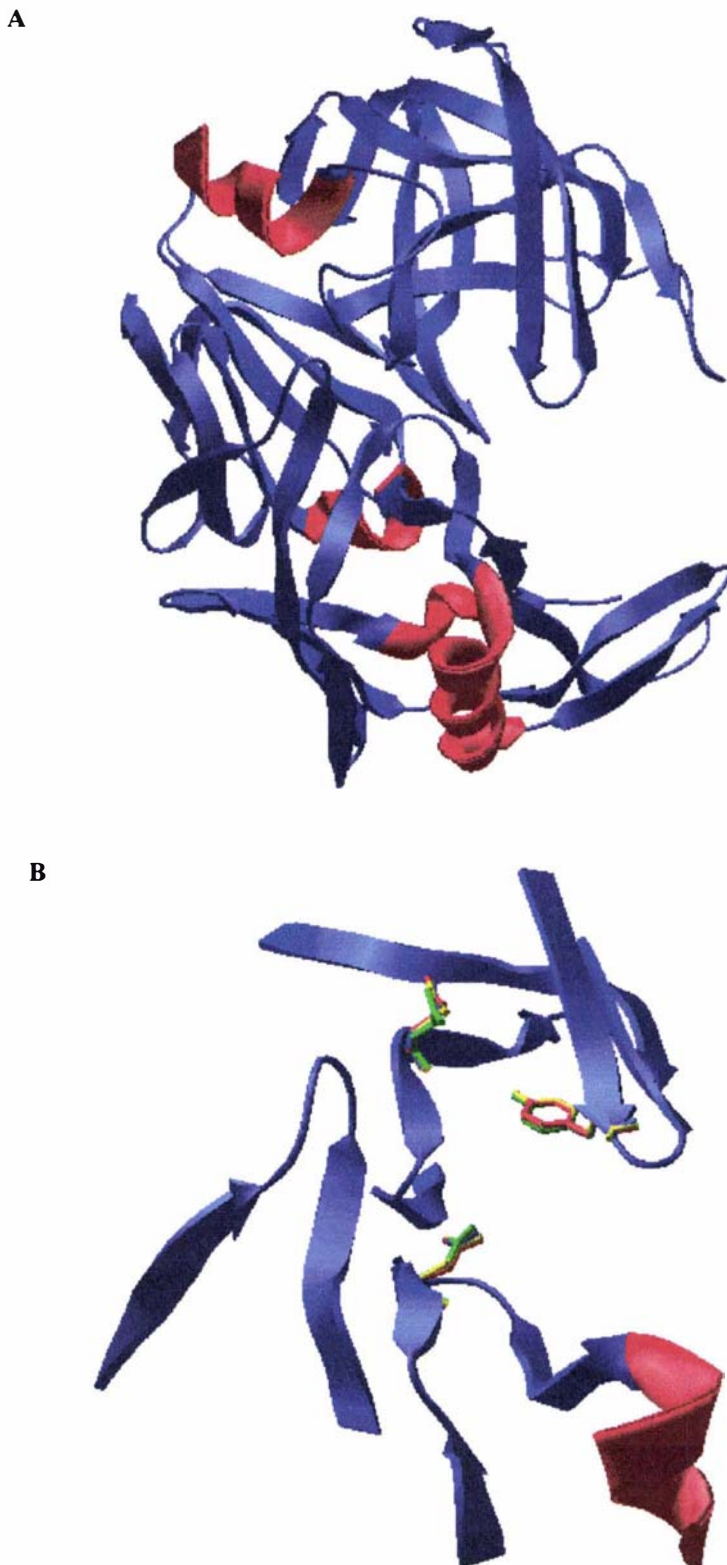


Figure 8.4 Homology models using the Sap2 structure as a template

A. The structure of Sap2 (Cutfield *et al.*, 1995) with β -sheets shown in blue and α -helix in red.

B. The catalytic residues (D37, D218 and Y84, Sap2 numbering) of the active sites of Saps1 (yellow), 4 (red) and 5 (green) shown on the Sap2 ribbon diagram.

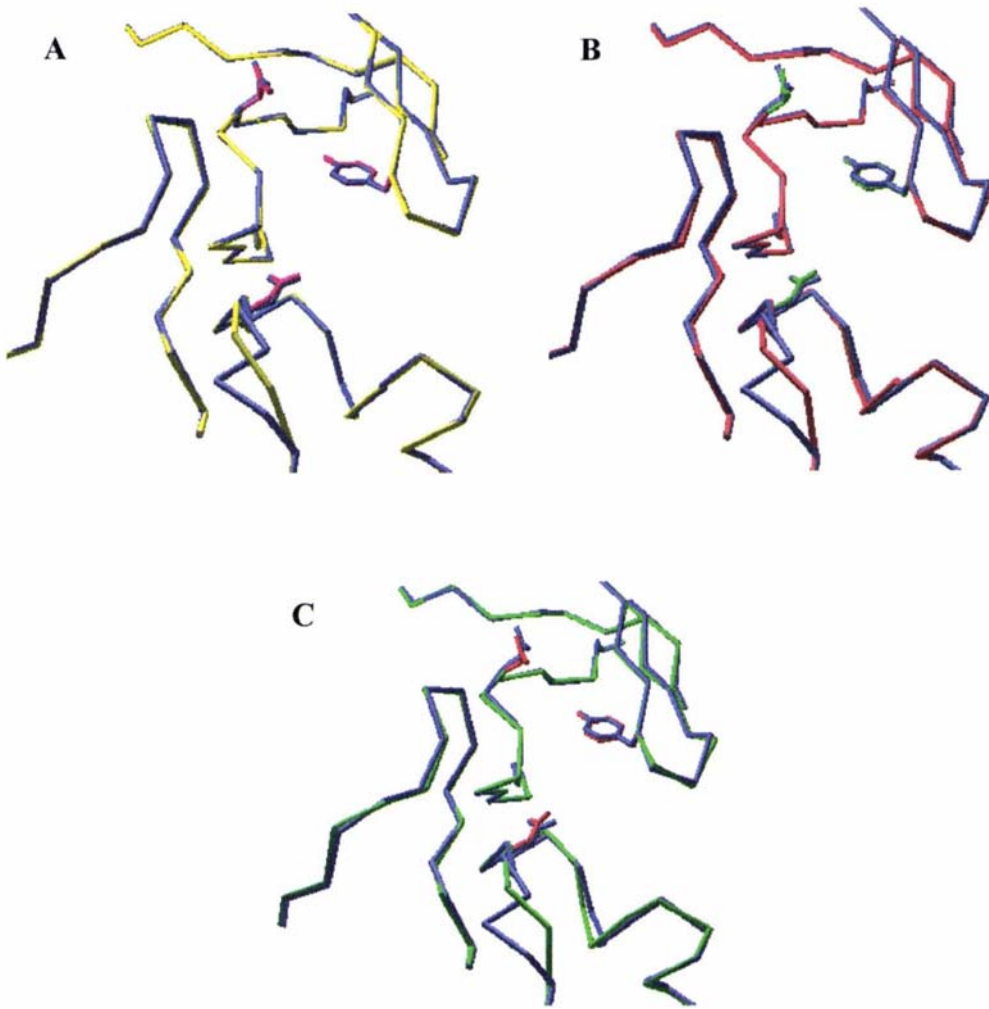


Figure 8.5 *Active site models of Saps1, 4 and 5*

The C- α backbone of Sap2 (blue) is overlaid with the predicted active site structures of Saps1, 4 and 5. The active site residues of Sap2 are shown in blue (D37, D218, Y84).

- A. The predicted C- α backbone of Sap1 is shown in yellow and the active site residues are highlighted in magenta.
- B. The predicted C- α backbone of Sap4 is shown in red, and the active site residues are highlighted in green.
- C. The predicted C- α backbone of Sap5 is shown in green, and the active site residues are highlighted in red.

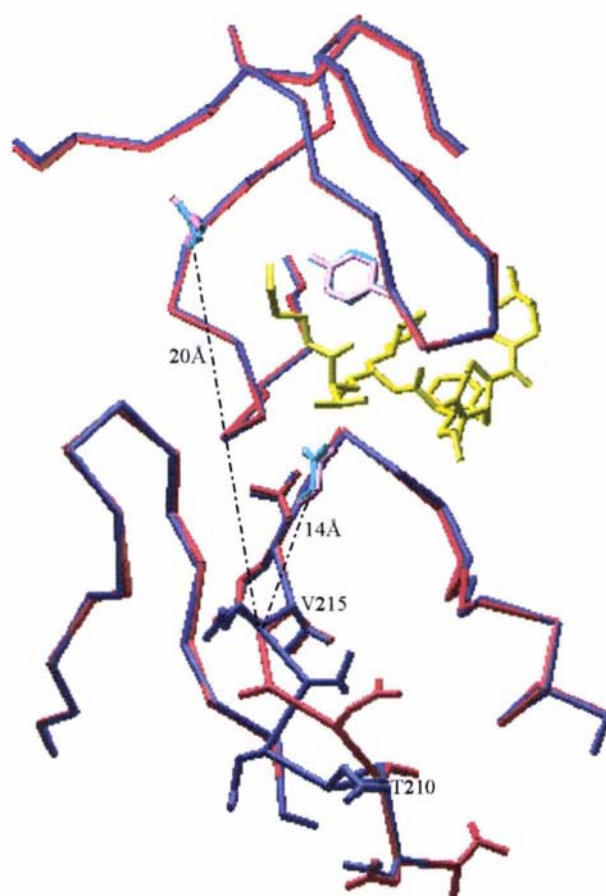


Figure 8.6 *Sap2 active site structure (in complex with A70450) overlaid with the Sap4 model*

The active site structures are shown in blue (Sap2) and red (Sap4), with the active site aspartic residues and tyrosine 84 shown in pink (Sap2) and aqua (Sap4). The tight-binding inhibitor, A70450, is depicted in yellow. The models show the alpha-carbon backbone, except where the backbones diverge, then the sidechains are also shown (same colours as mainchain). The significant predicted changes occur between Valine 215 (V215) and Threonine 210 (T210). The comparison shows the spatial separation between the substrate binding cleft and the divergence between Sap2 and the predicted Sap4 structure.

9.0 DISCUSSION

The contribution of the Saps to the colonisation and invasion of host tissues by *C. albicans* has been debated extensively, with many contradictory findings. For example there is evidence implicating Saps in the disease process (De Bernardis *et al.*, 1996; Borg & Ruchel, 1988; Kwon-Chung *et al.*, 1985), and results which suggest that Saps play little more than a spectator role in the development of candidoses (Togni *et al.*, 1994; Schreiber *et al.*, 1985; Germaine & Tellefson, 1981). One of the most contentious issues has been how isoenzymes, which are active only at low pH values, could be involved in a disease process at sites at or near neutral pH values. Samaranayake *et al.* (1984) have shown that *C. albicans* can decrease the localised pH by secretion of small acid metabolites (predominantly pyruvate and acetate) following carbohydrate utilisation. This would certainly enable Saps to play a role in candidosis at some sites, such as the oral cavity. However, it does not explain how Saps might contribute to infection at sites, such as the bloodstream, where the yeast would be unable to alter the environment. Therefore studying the Sap family is important for two reasons: First, to unravel the roles Sap enzymes play in the *C. albicans* disease process. Second, if these secreted enzymes are indispensable for pathogenicity, they would make excellent drug targets, as they do not require uptake of the drug by the organism itself.

It is for this latter reason that an understanding of the Sap isoenzymes is particularly important.

Elucidating the roles of Saps 4, 5 & 6 in the infection process requires an understanding of the biochemical characteristics of the enzymes, including properties such as pH and temperature optima, and substrate specificities. The expression of recombinant Sap4 and Sap5 from *P. pastoris* in the present project enabled them to be purified in large quantities, for biochemical and enzymatic studies. A significant finding from this work was the higher pH optima of these enzymes compared with other members of the Sap family. Sap4 and Sap5 both showed optimal activity at pH 4.5, whereas Saps1, 2 and 3 were optimally active at pH 3-3.5, and had 85%, 50% and 53% of optimum activity at pH 4.5. Sap4 and Sap5 retained 40% and 20% (respectively) of their optimum activity at pH 7.0, but Saps1, 2 & 3, showed just 5-10% of their optimal activity at neutral pH (section 6.3; Borg-von Zepelin *et al.*, 1998). The significant activity at high pH values suggests that these isoenzymes could indeed be implicated in the disease process at sites of neutral pH, without the need for localised acidification by the fungus. This finding concurs with the observation of *SAP4-6* expression in *C. albicans* hyphae, either in response to pH/temperature transition or addition of serum, at neutral pH.

The recombinant Sap1, Sap4 and Sap5 prepared in this work was more thermostable than native Sap2 and Sap3 and recombinant Sap1. This may be due to glycosylation of the proteins by *P. pastoris*. Glycosylation frequently contributes to protein thermal stability (Wang *et al.*, 1996) and Sap1 expressed in *P. pastoris* (Chapter 6.6, Borg-von Zepelin *et al.*, 1998) shows increased thermal stability when compared with recombinant Sap1 from *S. cerevisiae* (Smolenski *et al.* 1997). Analysis of the sugars associated with the recombinant proteins unexpectedly revealed the presence of *N*-acetylglucosamine as well as mannose in the *O*-linked oligosaccharides. Saps1, 4 and 5 do not contain any consensus Asn-X-Ser/Thr sequences for N-glycosylation and in most yeasts studied to date, N-linked oligosaccharides are of the high mannose type and mannose only is present in *O*-linked oligosaccharides. However, *O*-linked *N*-acetylglucosamine is not unprecedented; it has been detected in mammalian neurofilaments (Dong *et al.*, 1993), as the terminal residue of the mannan chains of *Kluyveromyces lactis* (Abeijon *et al.*, 1996), and in the simple eukaryote *Dictyostelium discoideum* (Jung *et al.*, 1998).

The substrate specificity analyses of Sap4 and Sap5 also performed as part of this project (section 6.5) revealed a surprising result. Neither enzyme showed any apparent sequence consensus for the hydrolysis of glucagon. This contrasts with an analysis of Sap1, Sap2 and Sap3 (Hermann, 1996; refer Chapter 6.5), which revealed a strong preference for an aspartate residue in the P₁ position, and a phenylalanine residue in the P₁' position of the active site. The consensus sequence for Sap1, 2, and 3 was hydrophilic - hydrophobic - any amino acid - hydrophilic - aromatic - hydrophobic - hydrophilic for the P₄-P₃' sites, respectively (refer to Table 6.2). While Sap4 and Sap5 both hydrolysed glucagon at the same sites, the resulting peptide fragments showed no cleavage sequence consensus based on size, hydrophobicity or charge. Sap4 and Sap5 also showed much less activity towards glucagon, compared with Sap1, 2 and 3, as judged by the rate of peptide production, which suggests that glucagon does not contain the preferred amino acid sequences for hydrolysis by Sap4 or Sap5.

The specific activities of Saps1, 2, 4 and 5 are quite different when assayed using either the fluorocasein or globin (refer section 6.4). This suggests that there are preferred cleavage sites in some protein substrates, which are absent in others. The slower hydrolysis of glucagon by Sap4 and Sap5, coupled with the lack of consensus cleavage sequence suggests that these enzymes might have specific physiological targets upon which they act. Substrate specificity analyses of recombinant Sap1, Sap2, Sap3 and Sap6 (expressed as pro-proteins in *E. coli*) suggested that Sap6 has the same specificity as Sap1 and Sap2 in the P₁ position, with amino acid preference Phe>Leu>Tyr, but differed at the P₁' position where the order of preference for Saps2 and 3 was Tyr>Phe>Trp, Tyr>Ala>Phe for Sap1, and Ala>Glu>Tyr for Sap6 (Koelsch *et al.*, 1998). While a direct comparison between the reported specificity of Sap6 with that determined for Sap4 is not strictly valid, it was anticipated they would be similar. One possible explanation for the observed differences between Sap4 and Sap6 might be the source of the recombinant enzymes; Koelsch *et al.* (1998) used renatured enzyme expressed from *E. coli*, in contrast to the yeast expression system described here for the production of Sap4. However, comparison of the specificity of Sap6 expressed in *P. pastoris* by Borg-von Zepelin *et al.* (1998) with Sap4 from the same host would resolve this discrepancy.

One anomaly in the substrate specificity work was the identical pattern of glucagon cleavage products, following incubation with either Sap4 or Sap5. This contrasts with the differences in substrate specificity noted in earlier work, where Sap4 was active towards denatured globin, but Sap5 was not (Fig. 6.7). If Sap4 and Sap5 have the same substrate specificity, it would be expected that they would both cleave the same protein substrates, but low activity of Sap5 towards denatured globin suggests otherwise. While the pH of both the globin assay and the glucagon digestion were the same, other assay components varied (eg. buffer composition; refer Methods 2.10.9), and these variations might have affected the specificity of the enzymes. Indeed, Goldman *et al.* (1995) document the changing substrate specificity of Sap2 towards three different peptide substrates between pH 5.5 and pH 7.3, with up to 98% decrease in specific activity over two pH units. This finding, coupled with the anomalous behaviour of Sap5 in the globin assay, suggests that Sap specificity is modulated by environmental conditions, and therefore may well have substrates and cleavage sites which are preferentially cleaved in the host environment which are not detected in the laboratory as assay conditions do not adequately mimic host physiology.

The most common primary substrate specificity determinant for aspartic proteinases is hydrophobic residues at positions P_1 and P_1' (Fruton, 1976). However, a study of the substrate specificities of aspartic proteinases from three *Candida* species; *C. albicans*, *C. tropicalis* and *C. parapsilosis*, showed the specificities of these enzymes differed from those of major human proteinases, including gastric pepsins, renal renin and cathepsin D (Fusek *et al.*, 1994). Substantial differences in specificity among the *Candida* proteinases were also identified. Most aspartic proteinases prefer a hydrophobic residue in the P_1 position (Fruton, 1970), and this is possibly due to the participation of phenylalanine (residue 159, Sap5 numbering, refer Fig. 3.7) forming part of the S_1 pocket. In Sap1, the phenylalanine is replaced by a proline, which might be expected to result in broad substrate specificity at this position. In the *C. tropicalis* and *C. parapsilosis* proteinases, and Saps 2 & 3, the phenylalanine is replaced by an aspartic acid. This might alter the preference for a hydrophobic amino acid at this position but the results presented in Table 6.2 for Saps1-3 show a preference for aspartic acid and lysine in position P_1 . Sap1 would be expected to exhibit broad substrate specificity in the S_1 binding pocket, and Sap2 and Sap3 would be expected to hydrolyse a basic residue, such as lysine, in this position. However, a charge conflict between

aspartic acid at both P_1 and S_1 in Sap2 and Sap3 would be expected to interfere with substrate binding and hydrolysis at this site. Sap4 and Sap5 have histidine and arginine instead of phenylalanine at this position, which suggests they would discriminate against basic residues in this position, which concurs with the results summarised in Table 6.2.

Specificity at position P_2 is influenced by the characteristics of residue 74, which is found at the margins of pockets S_2 and S_2' . In both the *C. tropicalis* enzyme and Sap3, residue 74 is a glutamic acid, compared with arginine in *C. parapsilosis* enzyme, glycine in Sap1 and Sap2, and lysine in Sap4 and Sap5. This correlates with the observed specificities of these enzymes; the *C. tropicalis* proteinase and Sap3 discriminate against acidic residues, the *C. parapsilosis* proteinase, Sap1 and Sap2 tolerate a variety of side chains, while Sap4 and Sap5 discriminate against basic residues in position P_2 (Table 6.2 and Fusek *et al.*, 1994). The amino acid sequence from residues 190-200 in the protein (Sap5 numbering, Fig. 3.7) contributes to specificity at both S_1 and S_3 sites. In particular, residue 196 (Sap5 numbering, Fig. 3.7) is thought to play a role, so that specificity at the P_3 position is the same as that mentioned above for position P_1 and the experimental data presented in Table 6.2 is in agreement with this. Interactions at position P_4 are essential for tight binding, and several aspartic proteinases show remarkable rigidity in specificity at P_4 . For example, cathepsin D is highly selective for proline at position P_4 (Scarborough *et al.*, 1991). It has been proposed that shape complementarity between the P_4 side chain and the S_4 binding pocket plays an important role. Indeed, while there is no apparent consensus based on amino acid residue type, at position P_4 , Saps 1-5 all prefer amino acids with large side chains (Table 6.2).

An analysis of the residues which compose the S_1 and S_1' subsites of the Saps (Table 6.3; Chapter 6.5) showed few differences between the isoenzymes and did not provide an explanation for the observed differences in specificity, particularly the unusual specificity of Sap2 and Sap3 at position P_1 (Hermann, 1996). This intimates that nearby residues in the extended binding site must affect substrate specificity as suggested by Ridky *et al.* (1996). Indeed, there are some clear amino acid differences between the groups Sap1-3 and Sap4-6, such as the substitution of Ala303 in Saps 1 & 2, and Tyr303 in Sap3 for Asp303 in Saps4-6 (Sap2 numbering). This substitution lies within

the vicinity of one of the unique structural features of the *C. albicans* Saps, a deletion of a short (4-7 residue) looped segment, which alters the relative position of the two lobes, and affects the boundary of the extended substrate binding site (Chapter 1.15). However, analysis of homology models of Sap1, Sap4 and Sap5 showed that this variation did not alter the local structure, and therefore did not provide a structural basis for enzyme specificity. Rather the comparative protein models of Sap1, Sap4 and Sap5 showed that the only structural differences occurred in a small segment of surface peptide (Fig. 8.6), which did not affect the substrate binding cleft, and hence did not explain the altered specificity of these enzymes. The comparison of proteinase sequences, together with the specificity data presented in section 6.5, suggests the participation of as yet unidentified residues of these aspartic proteinases in forming specificity binding pockets (Fusek *et al.*, 1994). Therefore, crystallographic studies will be required to determine the structural basis of Sap isoenzyme specificity. Preliminary crystallographic trials have been successful, producing small, diffraction-quality crystals of Sap1 in complex with A70450, and slightly larger crystals of Sap4 in complex with A70450.

One inherent difficulty in the biochemical analysis of recombinant enzymes is the question of whether native and recombinant enzymes behave in the same manner, or whether non-native modifications might affect one or more properties of the enzymes. For example, there is no evidence of significant glycosylation of the native Sap isoenzymes, but recombinant Sap1 and Sap4 appear to possess 10-13 moles of sugar residues per mole of enzyme. This glycosylation might have no effect on the properties of the enzymes, or it may enhance thermal stability, and even affect substrate specificity. While the lack of native Sap4 and Sap5 means that this question remains unanswered, it must still be considered when comparing these enzymes.

The discovery of a family of at least nine Sap isoenzymes has prompted further *in vivo* studies in an effort to understand the roles each of these enzymes plays in growth and infection. Various experimental approaches have been used to show differential regulation of the Sap isoenzymes during the growth- and morphological transition-phases of *C. albicans* (refer section 1.14).

The differential pattern of Sap expression suggests individual roles for each isoenzyme in the growth of *C. albicans*, and presumably also in augmenting the onset and development of disease. Disruption of *SAP1* ($\Delta sap1$), *SAP2* ($\Delta sap2$), *SAP3* ($\Delta sap3$), and *SAPs4-6* ($\Delta sap4,5,6$), in *C. albicans* strain SC5314, has shown these enzymes play a role in the disease process, as mice and guinea pigs challenged with the null strains showed increased survival times compared with control animals infected with the parent strain (Sanglard *et al.*, 1997; Hube *et al.*, 1997). Forty percent of animals challenged with the disrupted strain ($\Delta sap4, 5, 6$), survived to day 14, compared with four percent of those infected with the parent strain. Also, yeast phase cells of the *sap4, 5, 6* null strain exhibited increased adherence to buccal epithelial cells *in vitro* (Watts *et al.*, 1998), but Sap4, 5 or 6 would not be expected to have an effect on the adherence of budding yeast cells if the proteins were only expressed in the hyphal form. This suggests that the *in vivo* expression of *SAP4-6* might not be limited to the yeast-hyphal transition as previously suggested from *in vitro* studies (refer to Section 1.16). Further support for this comes from evidence that the *sap4,5,6* null strain was unable to grow or express Sap2 in medium containing BSA as the sole nitrogen source. Typically, *SAP2* is the dominant transcript in yeast form cells with protein as the sole nitrogen source (section 1.14), and no *SAP4-6* expression has been previously detected under these conditions (Sanglard *et al.*, 1997). This result suggests that Sap4, Sap5 or Sap6 plays an important role for the process of induction of *SAP2*. The regulation of expression of Sap4, Sap5 and Sap6 is clearly more complicated than previously thought, and is not limited to the transition between yeast and mycelial phases of growth. While these data provide clear evidence implicating Saps in the infection process, they do not contribute to our understanding of either the regulation of individual Saps or the role each isoenzyme plays in contributing to disease.

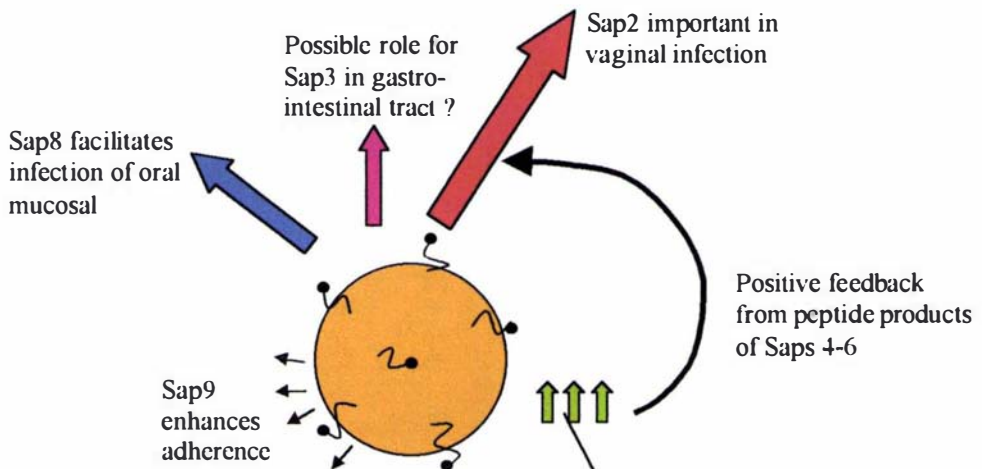
More recent *in vivo* experiments have suggested possible roles for some of the Sap isoenzymes. De Bernardis and others have previously shown a correlation between increased Sap production and pathogenicity (De Bernardis *et al.*, 1996; De Bernardis *et al.*, 1990; Cassone *et al.*, 1987), but have now demonstrated a link between elevated levels of Sap2 and successful vulvovaginal colonisation (De Bernardis *et al.*, 1999). A *sap2* mutant strain was virtually avirulent in a rat vaginitis model, but reinsertion of the *SAP2* gene into this mutant led to recovery of the vaginopathic potential. A role for Sap4, 5 and 6 has also been suggested by Borg-von Zepelin *et al.* (1998) following their

observation of Sap4, 5 and 6 expression in yeast form cells following phagocytosis by murine macrophages. Following fusion of the phagocytes with cellular lysosomes, the intracellular pH of 4.7-4.8 favours not only the host's lysosomal acid hydrolases which degrade the engulfed microbe, but is also optimal for Sap4, 5 and 6 activity (Chapter 6.3, Borg-von Zepelin *et al.*, 1998). Macrophages are a major obstacle to the establishment of systemic *C. albicans* infections, although intracellular persistence of *C. albicans* has been previously demonstrated (Marquis *et al.*, 1991). To this end, a comparison of the viability of the $\Delta sap4,5,6$ null mutant with the parent strain after phagocytosis by macrophages revealed a two-fold decrease in the survival of the mutant strain (survival rates of 18.6% and 43%, respectively), suggesting a role for Saps4, 5 and 6 in disseminated disease.

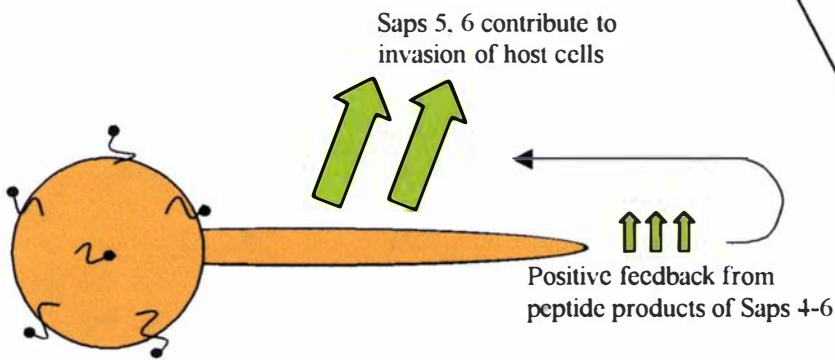
Despite many studies on the regulation, expression and characterisation of various members of the Sap family, a clear picture of the roles of the Sap isoenzymes in the disease process is yet to emerge. However, a tentative model can be proposed, based on the seminal findings reported to date (Fig. 9.1). The model proposes a pivotal role for one or all of Saps 4-6, which involves virtually undetectable levels of expression in most forms of *C. albicans*. As such, the task of Saps4-6 would be hydrolysis of specific target proteins, which would produce distinct peptide inducing agents, which then act to transduce the signal for increased expression of the appropriate Sap isoenzyme; predominantly Sap2 in vaginal mucosa, Sap8 in oral infections and Saps4-6 in macrophages. Sap3 might play a part in more acidic environments, such as the gastrointestinal tract, given that of all the Saps, Sap3 retains the most activity at low pH values. This hypothesis would account for the various observations that deletion of Saps 4-6 results in an altered phenotypes in growth conditions under which these enzymes have not previously been shown to be expressed (Watts *et al.*, 1998; Sanglard *et al.*, 1997). Sap9 has been proposed to assist adherence of *C. albicans* to host surfaces, given the putative GPI anchor encoded at the carboxy terminus of the protein, and the low but persistent level of *SAP9* expression observed under most conditions (Monod *et al.*, 1998).

This model does not address the roles of the Sap enzymes when expressed in different switch phenotypes of *C. albicans*, such as the opaque phase of WO-1. However, if the proposed roles of the isoenzymes are correct, then these functions could presumably be

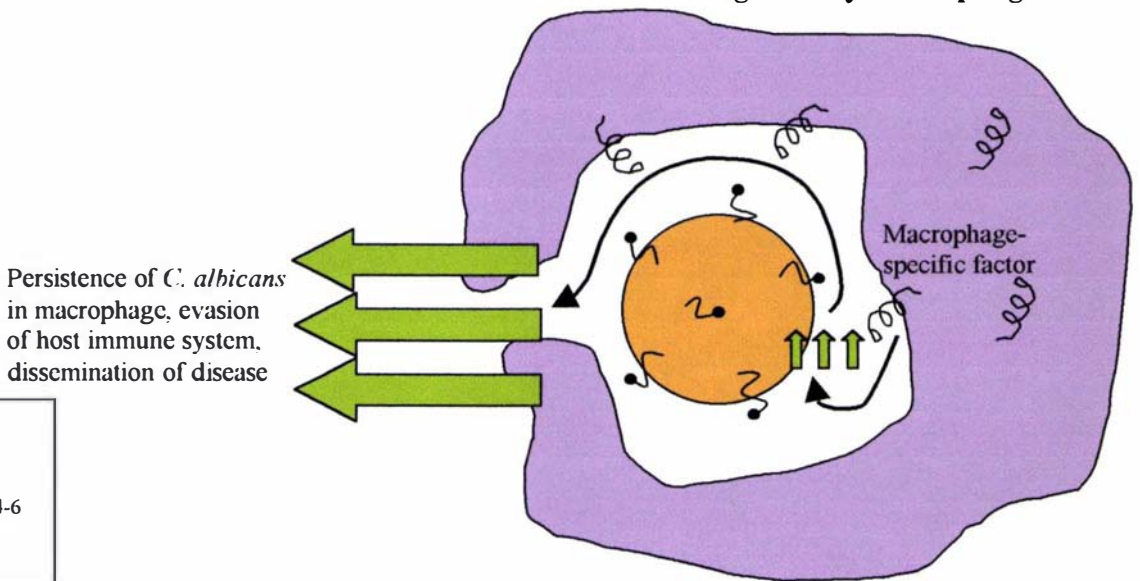
Yeast form *C. albicans*



Hyphal transition at neutral pH



Yeast form engulfed by macrophage



KEY	
	Sap2
	Sap3
	Saps4-6
	Sap8
	Sap9

Figure 9.1 Model of proposed roles of Sap isoenzymes in the onset and development of candidosis

extrapolated to various morphologies and switch phenotypes of *C. albicans*. In addition, the model suggests *in vivo* functions for each of the isoenzymes studied to date, but does not indicate the likely mechanism by which the specific peptide products induce expression of the relevant isoenzyme. In previous studies, Sap4-6 expression has been detected in serum-induced hyphal cells at neutral pH but not at low pH (Hube *et al.*, 1994), which suggests a pH regulation of the induction trigger in addition to the proposed peptide feedback signal. However, the system of induction is probably quite complex, as the present study has shown that Sap4-6 expression is not universally linked to hyphal formation in response to serum, at neutral pH.

In vitro expression studies have provided useful information about the linkage between Sap expression and various traits, and morphologies, such as Sap1 expression in the switching strain WO-1 (Hube *et al.*, 1994; White & Agabian, 1995). However, many *in vitro* studies do not appear to adequately mimic physiological conditions, producing results which differ from the *in vivo* situation. For example, analysis of *SAP8* expression at both 25 °C and 37 °C indicated that significant mRNA levels were only detected at 25 °C, suggesting that it was unlikely to play a significant role in mammalian infection. However, transcripts of *SAP8* have since been found in models of oral candidosis, indicating that there are some important differences between the inducing conditions present in *in vitro* experiments and *in vivo* infection. In addition, expression of *SAP4*, *5* and *6* was not detected in yeast-form cells until a physiologically-relevant model, employing murine macrophages, was used (Borg-von Zepelin *et al.*, 1998). Expression analyses were also performed in this work, using conditions which have been previously shown to induce *SAP4*, *5* and *6*, but these experiments did not induce expression of any of these transcripts (Chapter 7.3 & 7.4). The reasons for this are unclear, but the finding supports the suggestion that *in vitro* models are inherently less reliable, as there are likely many factors governing Sap induction which are as yet unknown, and therefore not controlled in the experiments.

9.1 Future Work

The primary motivation for studying the Sap family of isoenzymes is to understand their contribution to the virulence of *C. albicans* and to examine their potential as therapeutic drug targets. Therefore, structure-function studies should be a priority, because if the Saps do not represent a feasible target for drug design, then analysing the regulation of this family of pathogenicity factors is merely of academic interest. For this reason, the next step in this project should be optimising Sap1, Sap4 and Sap5 crystal growth, followed by solving and refining the Sap structures. These structures should provide a basis for understanding the differences in specificity of the Sap enzymes.

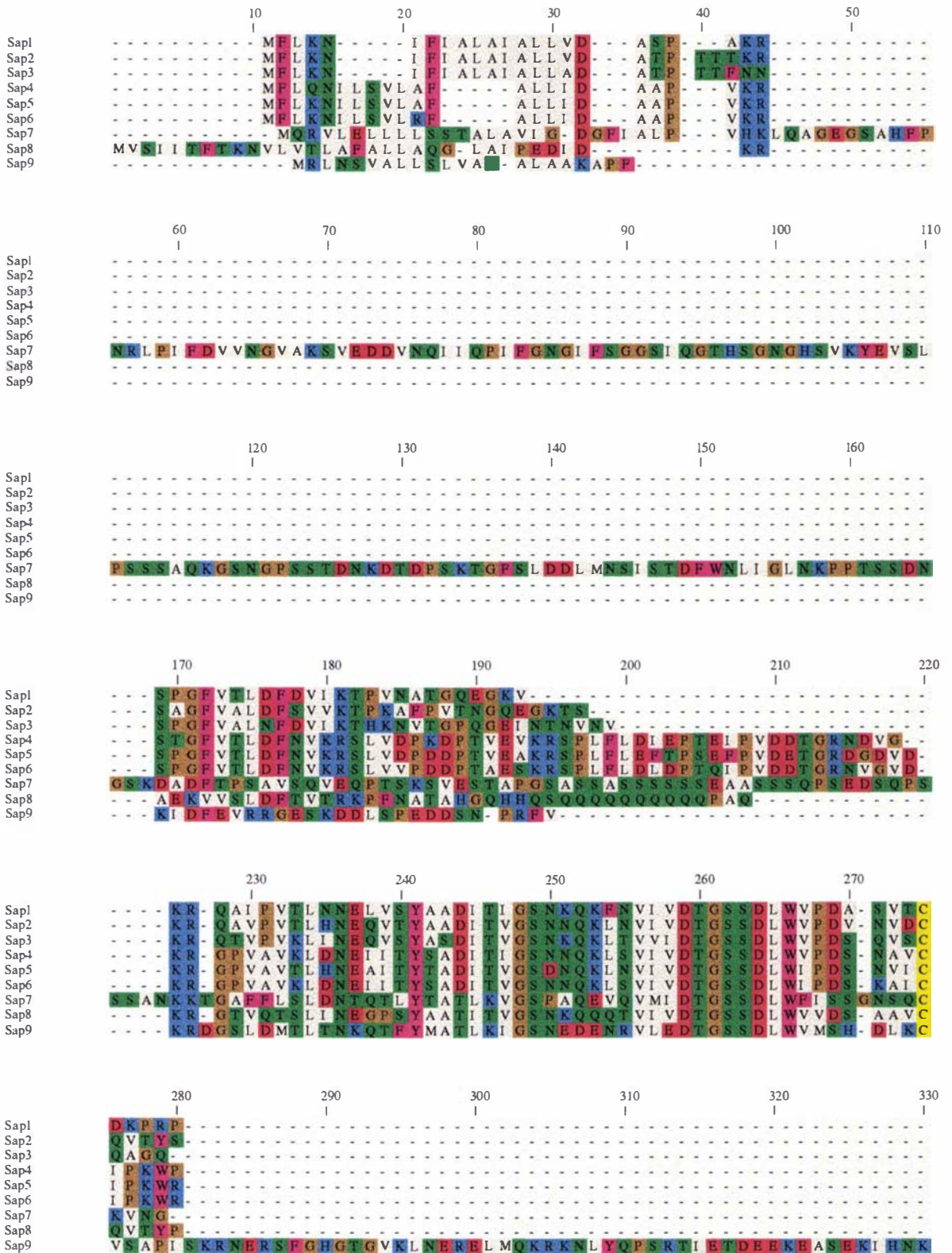
To date, only two aspartic proteinase zymogen structures have been solved (Moore *et al.*, 1995; Hartsuck *et al.*, 1992), therefore expression, purification and structural analysis of either the Sap4 or Sap5 pro-proteins would be of interest for two reasons. Firstly, the structure would provide excellent insight into the mechanism of peptide inhibition of Saps, as the pro sequence has been shown to be a powerful inhibitor of Sap activity *in vivo* (Baker *et al.*, 1992). Also, a high resolution zymogen structure would aid our understanding of substrate binding and perhaps illuminate the structural basis for the observed specificity of these enzymes. Secondly, most aspartic proteinase pro-sequences are approximately 50 amino acids in length, whereas the Sap4, 5 & 6 pro sequences each possess a significant insertion of 20-25 amino acids adjacent to the mature amino terminus. This segment could conceivably be involved in regulation or activation of the proteinases and therefore of interest to study. Zymogen expression would also create the possibility of analysing the binding affinity of the pro-segment for the mature protein.

To further our knowledge of the differential regulation of the Sap isoenzymes, in particular Saps4, 5 & 6, a search for any unique physiological substrates would be useful. That is, using a physiologically-relevant inducer, such as murine macrophages which promote expression of Saps4, 5 & 6 (Borg-von Zepelin *et al.*, 1998), the inducing component could be fractionated to determine whether a specific substrate exists, and if so, the identity of the substrate. Another aspect of Sap expression analysis is determining the relative contribution to virulence of each enzyme at each site of colonisation (such as the oral cavity, vagina, epidermis, etc.). Utilising the previously

described anti-Sap2 antiserum (Ross, 1990), the anti-Sap4 antiserum produced during this work, and the anti-Sap5 antiserum made as a result of this work, western blotting or immunofluorescence analyses could be performed on clinical samples to identify which members of the family were expressed. Comparison of samples from various sites of infection might provide some insight into the relative contribution of specific enzymes at each locus.

10.0 APPENDIX

10.1 Sequence alignment of *SAP1-SAP9* of *C. albicans*



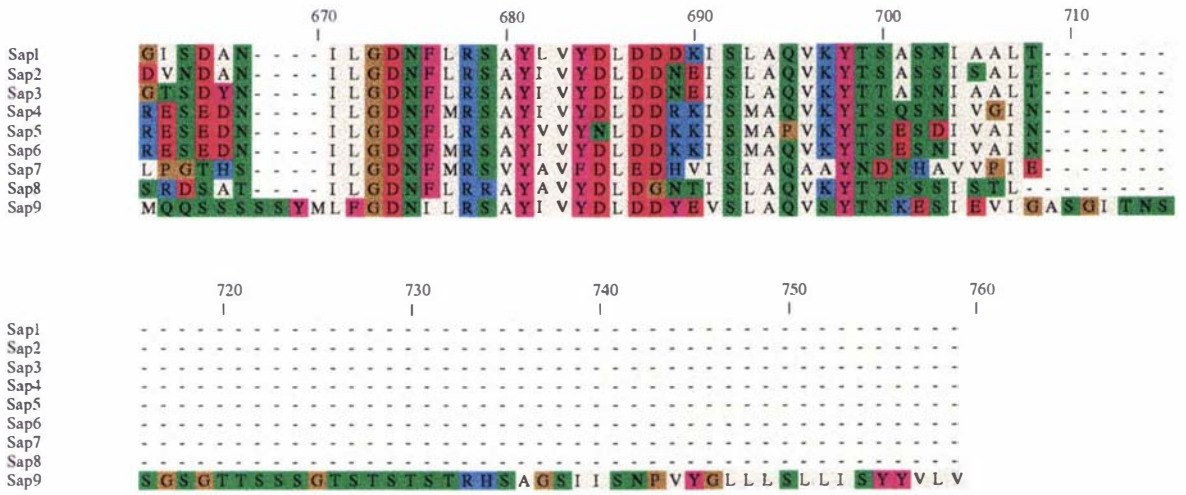


Figure 10.1 Alignment of the amino acid sequences of Sap1-Sap9 of *C. albicans*

The amino acids are coloured according to the colour coding schedule on page xiii.

10.2 Dionex analysis of Sap isoenzymes

10.2.1 Dionex calculations

To calculate the amount of carbohydrate on the Sap enzymes, a standard curve was drawn using known amounts of particular sugars. Areas under the peaks from the chromatograms (Fig 6.19) were plotted, and the line equations used for the calculations.

Line equations:

Galactosamine	$y = (1 \times 10^7)x - (2 \times 10^6)$
Glucosamine	$y = (1 \times 10^7)x - (1 \times 10^6)$
Galactose	$y = (2 \times 10^7)x - 382604$
Glucose	$y = (2 \times 10^7)x + 160182$
Mannose	$y = (2 \times 10^7)x - (2 \times 10^6)$

Sample calculation:

Sap4 + TFA; peak 8 (14.22 min) represents mannose.

Area under peak = 31802201

$$\begin{aligned} \text{Destruction of mannose by TFA} &= 77\%, \text{ therefore original area} = \frac{31802201}{0.23} \\ &= 1.383 \times 10^8 \end{aligned}$$

Substituting into the line equation for mannose:

$$\begin{aligned} 1.383 \times 10^8 &= (2 \times 10^7)x - (2 \times 10^6) \\ \therefore x &= 7.01 \mu\text{g mannose} \end{aligned}$$

Each sample injected onto the column was 25 μL , and the total sample volume was 100 μL . Therefore the total amount of mannose was:

$$7.01 \mu\text{g} / 25 \mu\text{L} \times 100 \mu\text{L} = 28 \mu\text{g mannose}$$

Mr mannose = 180.2

$$\begin{aligned} \text{Calculating number of moles of mannose present:} \quad n &= m/\text{Mr} \\ n &= 2.8 \times 10^{-5} \text{ g} / 180.2 \text{ g/mol} \\ n &= 1.55 \times 10^{-7} \text{ mol} \\ n &= 155 \text{ nmoles mannose} \end{aligned}$$

Number of moles of Sap4 in original sample: $n = 6.4 \text{ nmoles Sap4}$

Number of moles of mannose per mole of Sap4 $n = 155/6.4$

$$\underline{\underline{n = 24.2}}$$

NOTE:

For the GlcNAc calculations, the Mr used was 215.6, which is the mass of GlcN, as the amino group is lost during hydrolysis. Then the number of moles of GlcN = moles GlcNAc. The area under the glucosamine peak was 12220172, which extrapolated to 20033069, and a final value of 39 nmoles GlcNAc per 6 nmoles Sap4.

10.2.2 Dionex analysis of Sap5 carbohydrate moieties

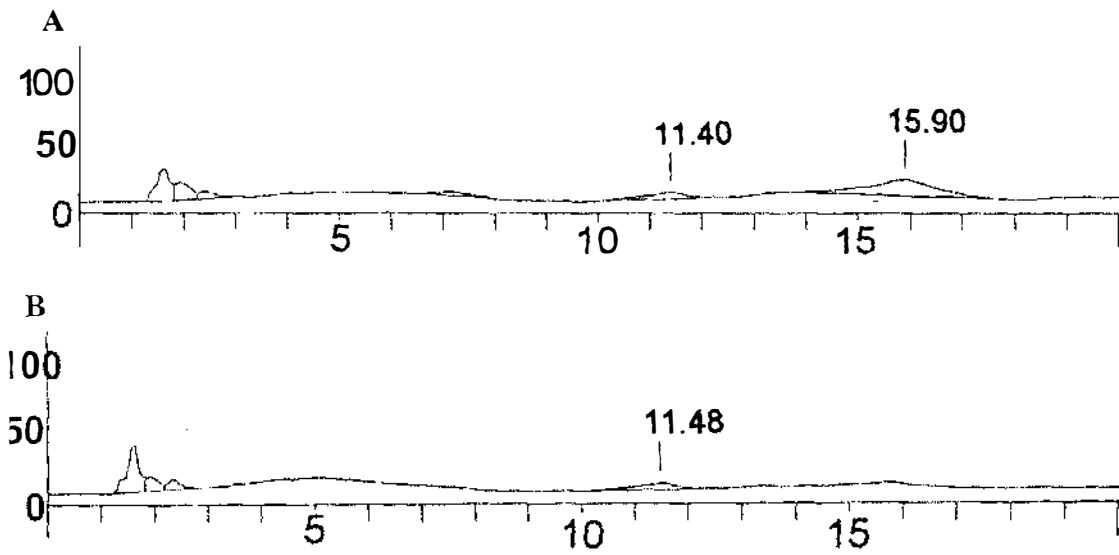


Figure 10.2 *Dionex chromatograms of Sap5*

The traces show the carbohydrate moieties cleaved from Sap5 by either HCl (A) or TFA (B). Each reaction contained 250 μ g of Sap5.

10.3 Purification and analysis of Sap5

10.3.1 Purification of Sap5

Large-scale induction of Sap5 was performed by Ms. R. J. Sanders, using the sparger system described in Chapter 4.9. Four litres of MS5A culture supernatant was reduced to approximately 500 mL in an Amicon ultrafiltration unit (CH2A concentrator), and buffer exchanged against 10 volumes of 5 mM sodium acetate, pH 5.0. The retentate was then centrifuged at 5,000 rpm for 10 min to remove particulate material, then concentrated to 25 mL in a Filtron™ disposable filtration unit (10 K cut-off). This retentate was filtered through a 0.22 μm filter prior to loading onto the Resource S column. A gradient of 0-1 M NaCl in 5 mM sodium acetate, pH 5.0 was run, over 30 min. Sap5 eluted as a sharp peak at approximately 400 mM NaCl (Fig. 10.3).

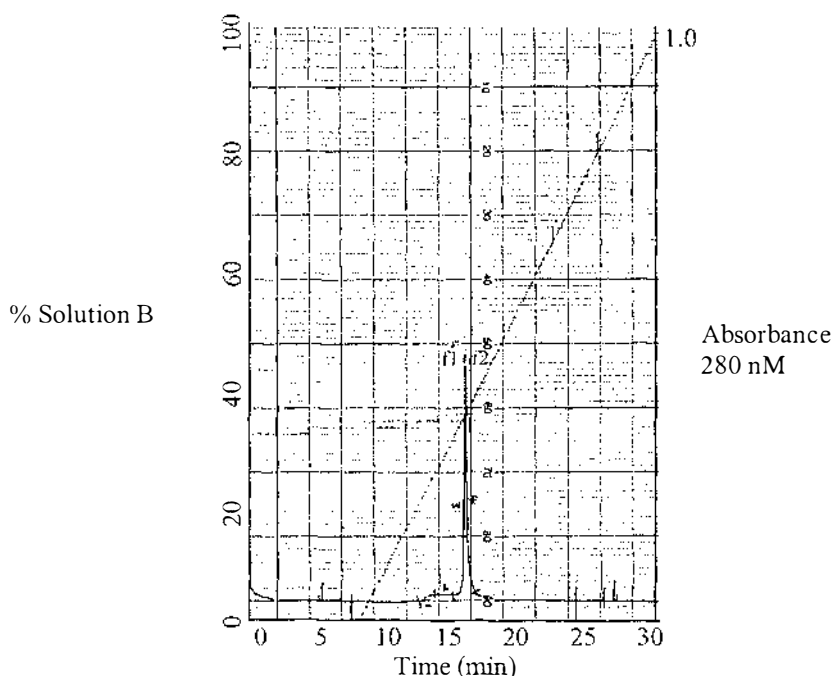


Figure 10.3 *Purification of Sap5 using ion-exchange chromatography*

The profile of Sap5 ion-exchange chromatography shows a single band eluting at approximately 17 min (400 mM NaCl or 40% B). The Sap5 eluted as a sharp single peak. Solution A; 5 mM sodium acetate pH 5.0; Solution B; 1 M NaCl, 5 mM sodium acetate, pH 5.0.

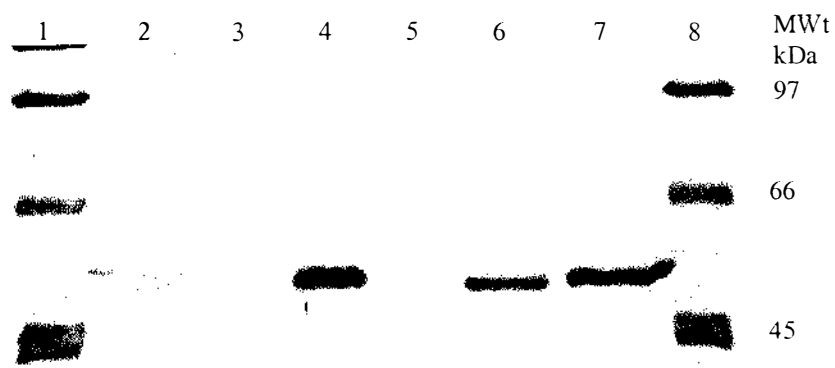


Figure 10.4 *SDS-PAGE of Sap5 following ion-exchange chromatography*

10% SDS-PAGE showing Sap5 at each stage of purification.

Lanes 1 & 8: Molecular weight markers

Lane 2: 50 μ L Sap5 culture supernatant

Lane 3: 50 μ L filtrate

Lane 4: 50 μ L concentrated, buffer exchanged Sap5

Lane 5: 50 μ L unbound fraction from Resource S column

Lane 6: 50 μ L fraction 1 from column

Lane 7: 50 μ L fraction 2 from column

10.3.2 Anti Sap5 antiserum

A polyclonal anti-Sap5 antiserum was also produced during the course of this work. The antiserum was raised in New Zealand white rabbits, at the Massey University Small Animal Production Unit (SAPU), as described in section 6.14. Purification of the IgG fraction, and subsequent titre of the antiserum was performed by Ms. R. J. Sanders. The methods are described in section 6.15.

The cross-reactivity of anti-Sap5 to Saps 1-5 was investigated. A 10% SDS-PAGE gel was run, with 200 ng of Sap1 - Sap5 in each lane. The gel was electroblotted to Hybond C and probed with a 1/5,000 dilution of anti-Sap5. All subsequent steps were performed as described in Methods 2.10.7. Fig. 10.5 shows anti-Sap5 cross-reacted strongly with Sap4 and Sap5, there was a moderate cross-reactivity between anti-Sap5 and Sap2 and Sap3, and no visible reaction between anti-Sap5 and Sap1.

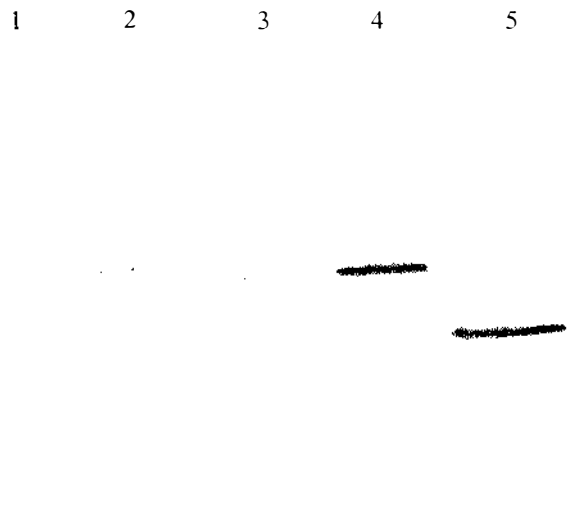
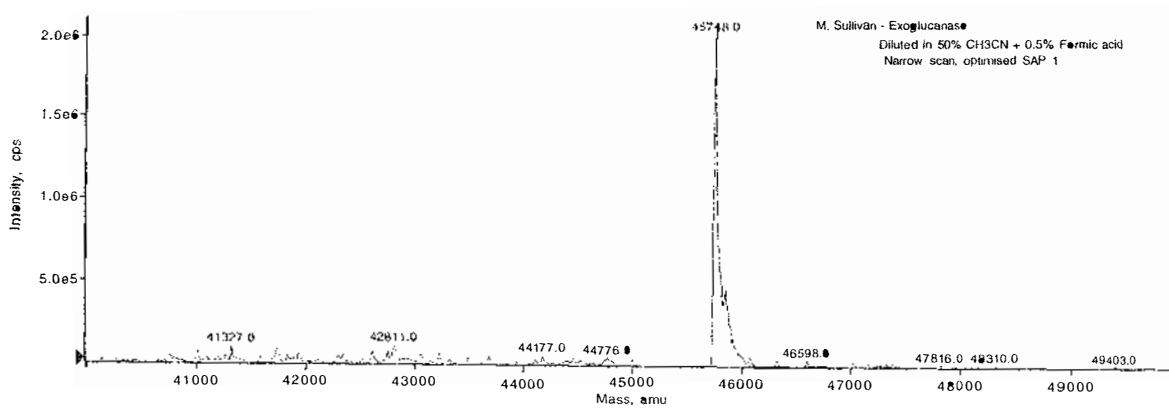
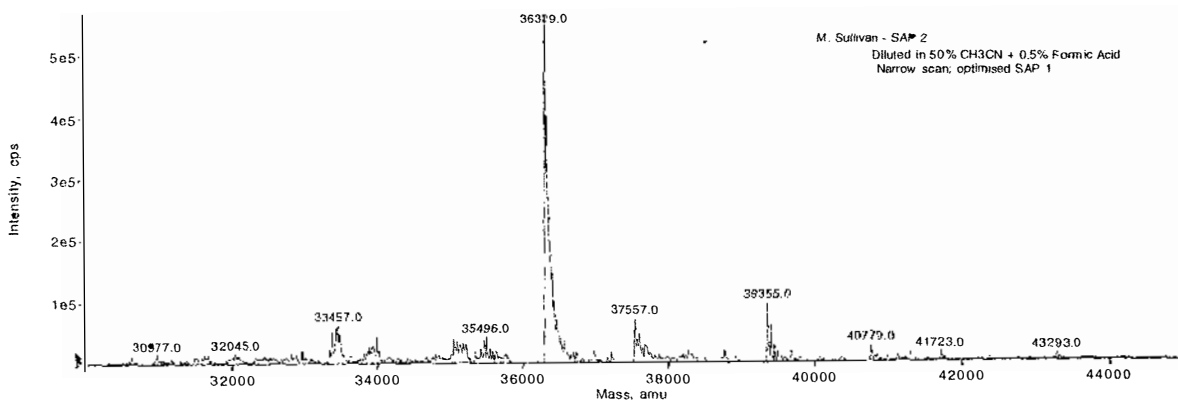


Figure 10.5 *Analysis of the specificity of anti-Sap5*

Lane 1: 200 ng Sap1
Lane 2: 200 ng Sap2
Lane 3: 200 ng Sap3
Lane 4: 200 ng Sap4
Lane 5: 200 ng Sap5

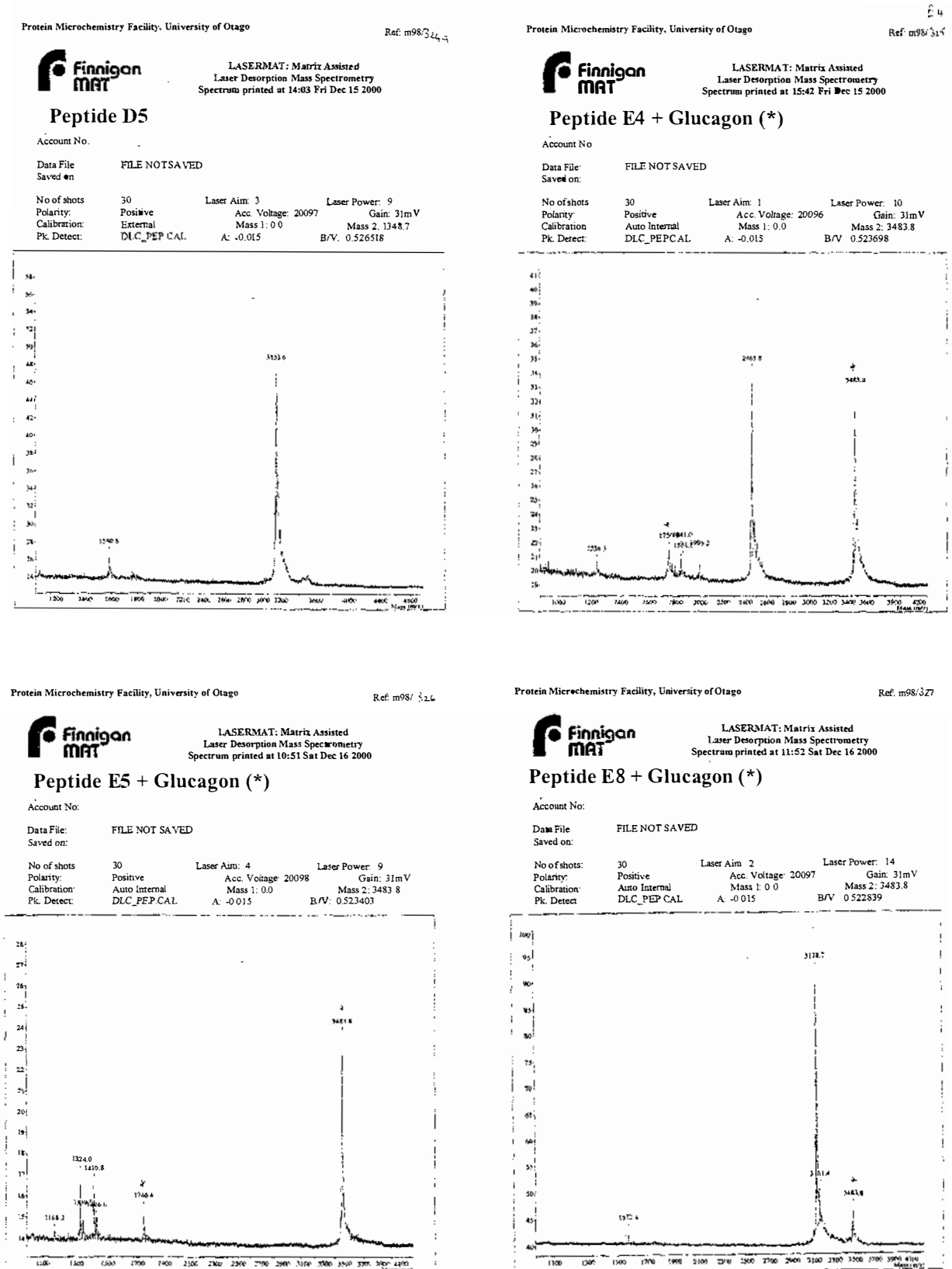
10.4 ESMS data

Native Sap2 and Exg (both from *C. albicans*) were submitted for ESMS analysis. The raw data is shown below.



10.5 MALDI-TOF analysis of glucagon peptide fragments

The glucagon peptide fragments were submitted for MALDI-TOF analysis to determine the mass of the fragments. The data from each of the samples submitted is shown below.



Protein Microchemistry Facility, University of Otago

Ref: m98/539



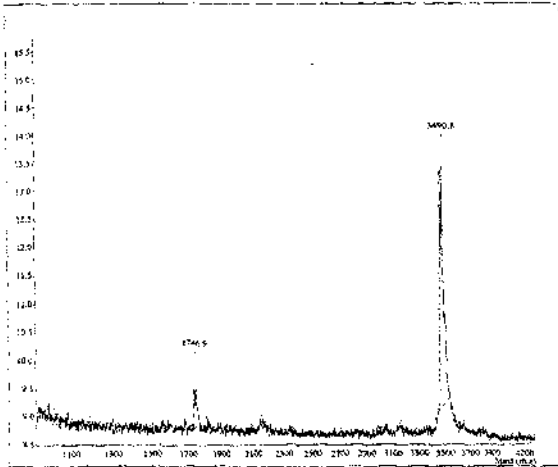
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 12:11 Fri Dec 22 2000

Peptide E9

Account No

Data File: FILE NOT SAVED
Saved on:

No of shots: 20 Laser Aim: 1 Laser Power: 12
Polarity: Positive Acc. Voltage: 20098 Gain: 31mV
Calibration: Manual Internal Mass 1: 0.0 Mass 2: 3794
Pk. Detect: DLC_PEP.CAL A: -0.015 B/V: 0.525194



Protein Microchemistry Facility, University of Otago

Ref: m98/524



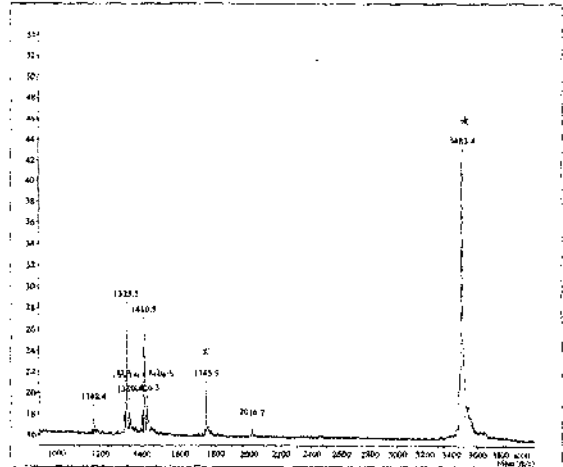
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 13:57 Sat Dec 16 2000

Peptide F7 + Glucagon (*)

Account No

Data File: FILE NOT SAVED
Saved on:

No of shots: 30 Laser Aim: 4 Laser Power: 9
Polarity: Positive Acc. Voltage: 20097 Gain: 31mV
Calibration: Auto Internal Mass 1: 0.0 Mass 2: 3483.8
Pk. Detect: DLC_PEP.CAL A: -0.015 B/V: 0.523811



Protein Microchemistry Facility, University of Otago

Ref: m98/531



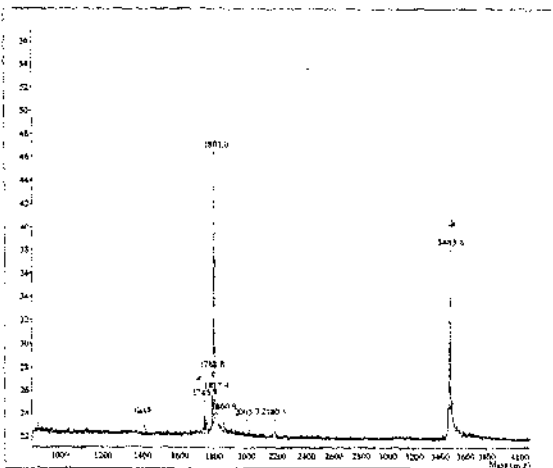
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 15:11 Sat Dec 16 2000

Peptide F9 + Glucagon (*)

Account No

Data File: FILE NOT SAVED
Saved on:

No of shots: 15 Laser Aim: 4 Laser Power: 13
Polarity: Positive Acc. Voltage: 20082 Gain: 31mV
Calibration: Auto Internal Mass 1: 0.0 Mass 2: 3483.8
Pk. Detect: DLC_PEP.CAL A: -0.015 B/V: 0.524389



Protein Microchemistry Facility, University of Otago

Ref: m98/532



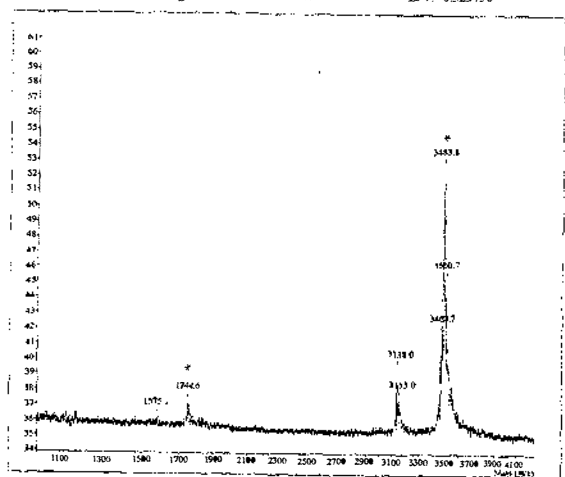
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 15:25 Sat Dec 16 2000

Peptide F10 + Glucagon (*)

Account No

Data File: FILE NOT SAVED
Saved on:

No of shots: 30 Laser Aim: 2 Laser Power: 8
Polarity: Positive Acc. Voltage: 20097 Gain: 31mV
Calibration: Auto Internal Mass 1: 0.0 Mass 2: 3483.8
Pk. Detect: DLC_PEP.CAL A: -0.015 B/V: 0.523750



Protein Microchemistry Facility, University of Otago

Ref: m98/232



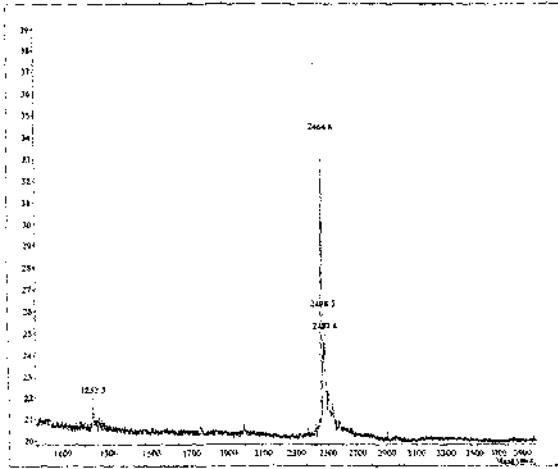
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 12:06 Thu Nov 23 2000

Peptide L6

Account No.

Data File: FILE NOT SAVED
Saved on:

No of shots: 34 Laser Aim: 4 Laser Power: 9
Polarity: Positive Acc. Voltage: 20096 Gain: 31mV
Calibration: Manual Internal Mass 1: 0.0 Mass 2: 379.4
Pk. Detect: MJH_PEP.CAL A: -0.015 B/V: 0.526577



Protein Microchemistry Facility, University of Otago

Ref: m98/233



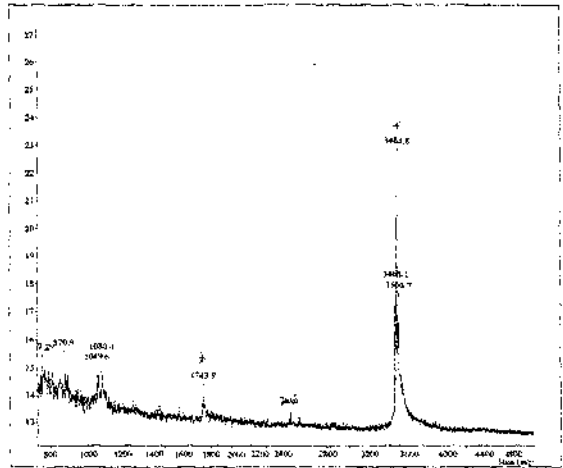
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 16:27 Sat Dec 16 2000

Peptide L7 + Glucagon (*)

Account No.

Data File: FILE NOT SAVED
Saved on:

No of shots: 30 Laser Aim: 4 Laser Power: 13
Polarity: Positive Acc. Voltage: 20097 Gain: 31mV
Calibration: Auto Internal Mass 1: 0.0 Mass 2: 3483.8
Pk. Detect: DLC_PEP.CAL A: -0.015 B/V: 0.523259



Protein Microchemistry Facility, University of Otago

Ref: m98/234



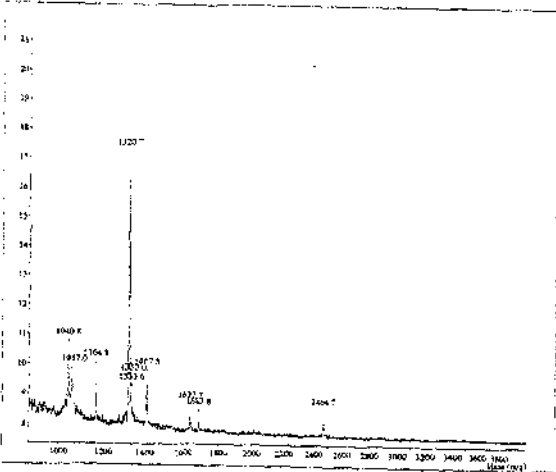
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 12:20 Thu Nov 23 2000

Peptide L8

Account No.

Data File: FILE NOT SAVED
Saved on:

No of shots: 30 Laser Aim: 1 Laser Power: 11
Polarity: Positive Acc. Voltage: 20096 Gain: 31mV
Calibration: Auto Internal Mass 1: 0.0 Mass 2: 379.4
Pk. Detect: MJH_PEP.CAL A: -0.015 B/V: 0.526697



Protein Microchemistry Facility, University of Otago

Ref: m98/235



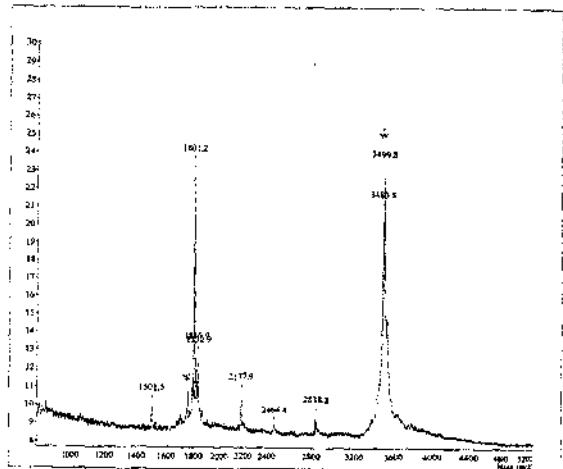
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 10:30 Mon Jan 22 2001

Peptide L10 + Glucagon (*)

Account No.

Data File: FILE NOT SAVED
Saved on:

No of shots: 40 Laser Aim: 2 Laser Power: 11
Polarity: Positive Acc. Voltage: 20092 Gain: 31mV
Calibration: Auto Internal Mass 1: 0.0 Mass 2: 3483.8
Pk. Detect: DLC_PEP.CAL A: -0.015 B/V: 0.523656



Protein Microchemistry Facility, University of Otago

Ref: m98/274



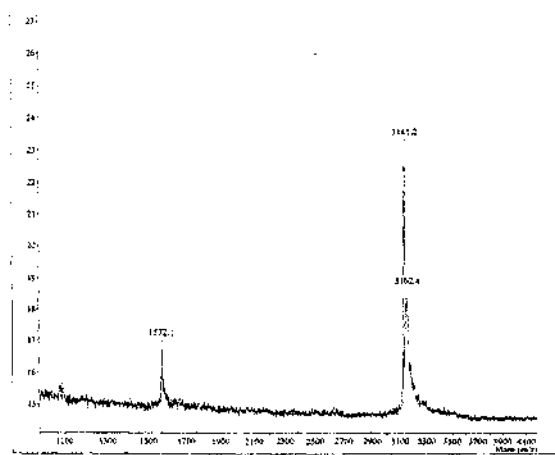
LASERMAT: Matrix Assisted
Laser Desorption Mass Spectrometry
Spectrum printed at 12:35 Thu Nov 23 2000

Peptide L12

Account No

Data File: FILE NOT SAVED
Saved on:

No of shots	30	Laser Aim	1	Laser Power	9
Polarity	Positive	Acc. Voltage	20095	Gain	31mV
Calibration	Auto Internal	Mass 1	0.0	Mass 2	379.4
PK Detect	MJH_P...CAL	A	-0.015	B/V	0.527196



11.0 REFERENCES

- Abad-Zapatero, C., Goldman, R., Muchmore, S.W., Hutchin, C., Stewart, K., Navaza, J., Payne, C.D. and Ray, T.L. (1996). Structure of a secreted aspartic protease from *C. albicans* complexed with a potent inhibitor: Implications for the design of antifungal agents. *Protein Science* **5**, 640-652.
- Abad-Zapatero, C., Goldman, R., Muchmore, S.W., Hutchins, C., Oie, T., Stewart, K., Cutfield, S.M., Cutfield, J.F., Foundling, S.I. and Ray, T.L. (1998). Structure of secreted aspartic proteinases from *Candida*: Implications for the design of antifungal agents. In *Aspartic Proteinases*. Edited by James. Plenum Press, New York.
- Abeijon, C., Robbins, P.W. and Hirschberg, C.B. (1996). Molecular cloning of the Golgi apparatus uridine diphosphate-N-acetylglucosamine transporter from *Kluyveromyces lactis*. *Proceedings of the National Academy of Sciences* **93**, 5963-5968.
- Agabian, N., Odds, F.C., Poulain, D., Soll, D.R. and White, T.C. (1994). Pathogenesis of invasive candidiasis. *Journal of Medical and Veterinary Mycology* **32** (Suppl. 1), 229-237.
- Akiyama, Y. and Ito, K. (1990). SecY protein, a membrane-embedded secretion factor of *E. coli* is cleaved by the ompT protease *in vitro*. *Biochemical and Biophysical Research Communications* **167**, 711-715.
- Anderson, J.M. and Soll, D.R. (1987). Unique phenotype of opaque cells in the white-opaque transition of *Candida albicans*. *Journal of Bacteriology* **169**, 5579-5588.

Anderson, J.M., Cundiff, L., Schnars, B., Gao, M., Mackenzie, I. and Soll, D.R. (1989). Hypha formation in the white-opaque transition of *Candida albicans*. *Infection and Immunity* **57**, 458-467.

Anderson, J.M., Mihalik, K. and Soll, D.R. (1990). Ultrastructure and antigenicity of the unique cell and pimple of the *Candida* opaque phenotype. *Journal of Bacteriology* **172**, 224-235.

Antley, P.P. and Hazen, K.C. (1988). Role of yeast cell growth temperature on *Candida albicans* virulence in mice. *Infection and Immunity* **56**, 2884-2890.

Aparicho, O.M., Billington, B.C. and Gottschling, D.E. (1991). Modifiers of position effects are shared between telomeric and silent mating-type loci in *S. cerevisiae*. *Cell* **66**, 1279-1287.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1992). *Short Protocols in Molecular Biology*. New York: John Wiley and Sons.

Badasso, M., Wood, S.P., Aguilar, C., Cooper, J.B., Blundell, T.L. and Dreyer, T. (1993). Crystallisation and preliminary crystallographic characterisation of aspartic protease A from baker's yeast and its complex with inhibitors. *Journal of Molecular Biology* **232**, 701-703.

Bailey, D.A., Feildmann, P.J.F., Bovey, M., Gow, N.A.R. and Brown, A.J.P. (1996). The *Candida albicans* *HYR1* gene, which is activated in response to hyphal development, belongs to a gene family encoding yeast cell wall proteins. *Journal of Bacteriology* **178**, 5353-5360.

Baker, D., Silen, J.L. and Agard, D.A. (1992). Protease pro region required for folding is a potent inhibitor of the mature enzyme. *Proteins: Structure, function and genetics* **12**, 339-344.

- Baldwin, E.T., Bhat, T.N., Gulnik, S., Hosur, M.V., Sowder, R.C. (2nd), Cachau, R.E., Collins, J., Silva, A.M. and Erickson, J.W. (1993). Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design. *Proceedings of the National Academy of Sciences USA* **90**, 6796-6800.
- Ballou, C.E. (1990). Isolation, characterization, and properties of *Saccharomyces cerevisiae* *mn* mutants with nonconditional protein glycosylation defects. *Methods in Enzymology* **185**, 440-470.
- Banerjee, A., Ganesan, K. and Datta, A. (1991). Induction of secretory acid proteinase in *Candida albicans*. *Journal of General Microbiology* **137**, 2455-2461.
- Bergen, M.S., Voss, E. and Soll, D.R. (1990). Switching at the cellular level in the white-opaque transition of *Candida albicans*. *Journal of General Microbiology* **136**, 1925-1936.
- Birse, C.E., Irwin, M.Y., Fonzi, W.A. and Sypherd, P.S. (1993). Cloning and characterization of *ECE1*, a gene expressed in association with cell elongation of the dimorphic pathogen *Candida albicans*. *Infection and Immunity* **61**, 3648-3655.
- Blundell, T.L. and Johnson, M.S. (1993). Catching a common fold. *Protein Science* **2**, 877-883.
- Borg, M. and Ruchel, R. (1988). Expression of extracellular acid proteinase by proteolytic *Candida* spp. during experimental infection of oral mucosa. *Infection and Immunity* **56**, 626-631.
- Borg-von Zepelin, M., Beggah, S., Boggian, K., Sanglard, D. and Monod, M. (1998). The expression of the secreted aspartyl proteinases Sap4 to Sap6 from *Candida albicans* in murine macrophages. *Molecular Microbiology* **28**, 543-554.

Braun, B.R. and Johnson, A.D. (1997). Control of filament formation in *Candida albicans* by the transcriptional repressor *TUP1*. *Science* **277**, 105-109.

Brawner, D.L. and Cutler, J.E. (1987). Cell surface and intracellular expression of two *Candida albicans* antigens during *in vitro* and *in vivo* growth. *Microbial Pathogens* **2**, 249-257.

Brown, L.A. and Chaffin, W.L. (1981). Differential expression of cytoplasmic proteins during yeast bud and germ tube formation in *Candida albicans*. *Canadian Journal of Microbiology* **27**, 580-585.

Brown, T.A. (1991). *Molecular Biology LabFax*. Oxford: BIOS Scientific Publications.

Bugg, C.E., Carson, W.A. and Montgomery, J.A. (1993). *Drugs by Design*. *Scientific American* **163**, 60-66.

Burgers, P.M.J. and Percival, K.J. (1987). Transformation of yeast spheroplasts without cell fusion. *Analytical Biochemistry* **163**, 391-397.

Calderone, R. A., Linehan, L., Wadsworth, E. and Sandberg, A.L. (1988). Identification of C3d receptors on *Candida albicans*. *Infection and Immunity* **56**, 252-258.

Calderone, R., Diamond, R., Senet, J.-M., Warmington, J., Filler, S. and Edwards, J.E. (1994). Host cell-fungal cell interactions. *Journal of Medical and Veterinary Mycology* **32** (Suppl), 151-168.

Cannon, R.D., Jenkinson, H.F. and Shepherd, M.G. (1990). Isolation and nucleotide sequence of an autonomously replicating sequence (ARS) element functional in *Candida albicans* and *Saccharomyces cerevisiae*. *Molecular and General Genetics* **221**, 210-218.

- Caplan, S., Green, R., Rocco, J. and Kurjan, J. (1991). Glycosylation and structure of the yeast *MATa1* α -Factor precursor is important for efficient transport through the secretory pathway. *Journal of Bacteriology* **173**: 627-633
- Capobianco, J.O., Lerner, C.G. and Goldman, R.C. (1992). Application of a fluorogenic substrate in the assay of proteolytic activity and in the discovery of a potent inhibitor of the *Candida albicans* aspartic proteinase. *Analytical Biochemistry* **204**, 96-102.
- Casanova, M., Lopez-Ribot, J.L. and Monteagudo, C., Llombart-Bosch, A., Sentandreu, R. and Martinez, J.P. (1992). Identification of a 58-kilodalton cell surface fibrinogen-binding mannoprotein from *Candida albicans*. *Infection and Immunity* **60**, 4221-4229.
- Cassone, A., De Bernardis, F., Mondello, F., Ceddia, T. and Agatensi, L. (1987). Evidence for a correlation between proteinase secretion and vulvovaginal candidosis. *Journal of Infectious Disease* **156**, 777-783.
- Cassone, A., Sullivan, P.A. and Shepherd, M.G. (1985). N-acetyl-D-glucosamine induced morphogenesis in *Candida albicans*. *Microbiologica* **8**, 85-99.
- Chaffin, W.L., Skudlarek, J. and Morrow, K.J. (1988). Variable expression of a surface determinant during proliferation of *Candida albicans*. *Infection and Immunity* **56**, 302-309.
- Christeller, J.T., Farley, P.C., Ramsay, R.J., Sullivan, P.A. and Laing, W.A. (1998). Purification, characterization and cloning of an aspartic proteinase inhibitor from squash phloem exudate. *European Journal of Biochemistry* **254**, 160-167

- Clerici, M. and Shearer, G.M. (1993). A TH1→TH2 switch is a critical step in the etiology of HIV infection. *Immunology today* **14**, 107-111
- Colina, A-R., Aumont, F., Deslauriers, N., Belhumeur, P. and de Repentigny, L. (1996). Evidence for degradation of gastrointestinal mucin by *Candida albicans* secretory aspartyl proteinase. *Infection and Immunity* **64**, 4514-4519.
- Cooper, J., Foundling, S., Hemmings, A., Blundell, T., Jones, D.M., Hallet, A. and Szelke, M. (1987). The structure of a synthetic pepsin inhibitor complexed with endothiasepsin. *European Journal of Biochemistry* **169**, 215-221.
- Cornish-Bowden, A. (1995). Fundamentals of enzyme kinetics. Revised edition. Portland Press, London.
- Crandall, M. and Edwards, J.E. (1987). Segregation of proteinase-negative mutants from heterozygous *Candida albicans*. *Journal of General Microbiology* **133**, 2817-2824.
- Cregg, J.M., Vedvick, T.S. and Raschke, W.C. (1993). Recent advances in the expression of foreign genes in *Pichia pastoris*. *Biotechnology* **11**, 905-910.
- Csank, C., Schroppel, K., Leberer, E., Harcus, D., Mohamed, O., Meloche, S., Thomas, D.Y. and Whiteway, M. (1998). Roles of the *Candida albicans* mitogen-activated protein kinase homolog, Cek1p, in hyphal development and systemic candidiasis. *Infection and Immunity* **66**, 2713-2721.
- Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y and McPherson, A. (1994). *Acta Cryst.* **D50**, 414-423.
- Cutfield, S., Marshall, C., Moody, P., Sullivan, P. and Cutfield, J. (1993). Crystallization of inhibited aspartic proteinase from *Candida albicans*. *Journal of Molecular Biology* **234**, 1266-1269.

Cutfield, S.M., Dodson, E.J., Anderson, B.F., Moody, P.C.E., Marshall, C.J., Sullivan, P.A. and Cutfield, J.F. (1995). The crystal structure of a major secreted aspartic proteinase from *Candida albicans* in complexes with two inhibitors. *Structure* **3**, 1261-1271.

Cutler, J.E. (1991). Putative virulence factors of *Candida albicans*. *Annual Reviews of Microbiology* **45**, 187-218.

Davies, D.R. (1990). The structure and function of the aspartic proteinases. *Annual Reviews in Biophysics and Biophysical Chemistry* **19**, 189-215.

De Bernardis, F., Agatensi, L., Ross, I.K., Emerson, G.W., Lorenzini, R., Sullivan, P.A. and Cassone, A. (1990). Evidence for a role for secreted aspartate proteinase of *Candida albicans* in vulvovaginal candidiasis. *Journal of Infectious Diseases* **161**, 1276-1283.

De Bernardis, F., Boccanera, M., Adriani, D., Spreghini, E., Santoni, G. and Cassone, A. (1997). Protective role of antimannan and anti-aspartyl proteinase antibodies in an experimental model of *Candida albicans* vaginitis in rats. *Infection and Immunity* **65**, 3399-3405.

De Bernardis, F., Chiani, P., Ciccozzi, M., Pellegrini, G., Ceddia, T., D'Offizzi, G., Quinti, I., Sullivan, P.A. and Cassone, A. (1996). Elevated aspartic proteinase secretion and experimental pathogenicity of *Candida albicans* isolates from oral cavities of subjects infected with human immunodeficiency virus. *Infection and Immunity* **64**, 466-471.

De Bernardis, R., Arancia, S., Morelli, L., Hube, B., Sanglard, D., Schafer, W. and Cassone, A. (1999). Evidence that members of the secretory aspartyl proteinase gene family, in particular *SAP2*, are virulence factors for *Candida* vaginitis. *Journal of Infectious Diseases* **179**, 201-208.

- Diez-Orejas, R., Molero, G., Navarro-Garcia, F., Pla, J., Nombela, C. and Sanchez-Perez, M. (1997). Reduced virulence of *Candida albicans* *MKC1* mutants: a role for mitogen activated protein kinase in pathogenesis. *Infection and Immunity* **65**, 833-837.
- Dong, D.L., Xu, Z.S., Chavrier, M.R., Cotter, R.J., Cleveland, D.W. and Hart, G.W. (1993). Glycosylation of mammalian neurofilaments. Localisation of multiple O-linked N-acetylglucosamine moieties on neurofilament polypeptides L and M. *Journal of Biological Chemistry* **268**, 16679-16687.
- Dunn, B.M., Scarborough, P.E., Lowther, W.T. and Rao-Naik, C. (1995). Comparison of the active site specificity of the aspartic proteinases based on a systematic series of peptide substrates. In: *Aspartic proteinases: structure, function, biology, and biomedical implication*. Edited by K. Takahashi. Plenum Press, New York.
- Fidel, P.L.Jr. and Sobel, J.D. (1994). The role of cell-mediated immunity in candidiasis. *Trends in Microbiology* **2**, 202-206.
- Fidel, P.L. Jr., Lynch, M.E. and Sobel, J.D. (1994). Effects of preinduced *Candida*-specific cell-mediated immunity on experimental vaginal candidiasis. *Infection and Immunity* **62**, 1032-1038.
- Fonzi, W.A. and Irwin, M.Y. (1993). Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134**, 717-728.
- Froeliger, E.H. and Carpenter, B.E. (1996). *NUT1*, a major nitrogen regulatory gene in *Magnaporthe grisea*, is dispensible for pathogenicity. *Molecular and General Genetics* **251**, 647-656.
- Fruton, J. (1970). TITLE HERE. *Advances in Enzymology* **33**, 401-443.

- Fruton, J. (1976). The mechanism of the catalytic action of pepsin and related acid proteinases. *Advances in Enzymology and Related Areas of Molecular Biology* **44**, 1-33.
- Fukazawa, Y. and Kagaya, K. (1997). Molecular bases of adhesion of *Candida albicans*. *Journal of Medical and Veterinary Mycology* **35**, 87-99.
- Fusek, M., Smith, E.A., Monod, M., Dunn, B. and Foundling, S.I. (1994). Extracellular aspartic proteinases from *Candida albicans*, *Candida tropicalis*, and *Candida parapsilosis* yeasts differ substantially in their specificities. *Biochemistry* **33**, 9791-9799.
- Gale, C.A., Bendel, C.M., McClellan, M., Hauser, M., Becker, J.M., Berman, J. and Hostetter, M.K. (1998). Linkage of adhesion, filamentous growth, and virulence in *Candida albicans* to a single gene, *INT1*. *Science* **279**, 1355-1358.
- Gemmill, T.R. and Trimble, R.B. (1999). Overview of *N*- and *O*- linked oligosaccharide structures found in various yeast species. *Biochimica et Biophysica Acta* **1426**, 227-237.
- Germaine, G.R. and Tellefson, L.M. (1981). Effect of pH and human saliva on protease production by *Candida albicans*. *Infection and Immunity* **31**, 323-326
- Glasgow, A.C., Hughes, K.T. and Simon, M.I. (1989). Bacterial DNA inversion systems. In *Mobile DNA*, pp. 637-660. Edited by D.E. Berg & M.M. Howe. American Society for Microbiology, Washington D.C.
- Goldman, R.C., Frost, D.J., Capobianco, J.O., Kadam, S., Rasmussen, R.R. and Abad-Zapatero, C. (1995). Antifungal drug targets: *Candida* secreted aspartyl protease and fungal wall β -glucan synthesis. *Infectious Agents and Disease* **4**, 228-247.

- Gow, N.A.R., Hube, B., Bailey, D.A., Schofield, D.A., Munro, C., Swoboda, R.K., Bertram, G., Westwater, C., Broadbent, I., Smith, R.J., Gooday, G.W. and Brown, A.J.P. (1995). Genes associated with dimorphism and virulence in *Candida albicans*. *Canadian Journal of Botany* **73**(Suppl. 1), S335-S342.
- Greenfield, R.A. (1992). Host defense system interactions with *Candida*. *Journal of Medical and Veterinary Mycology* **30**, 89-94.
- Grinna, L.S. and Tschopp, J.F. (1989). Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. *Yeast* **5**, 107-115
- Guex, N. and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723.
- Hamer, J.E. and Holden, D.W. (1997). Linking approaches in the study of fungal pathogenesis: a commentary. *Fungal Genetics and Biology* **21**, 11-16.
- Harlow, E. and Lane D. (1988). Storing and purifying antibodies. In *Antibodies: A laboratory manual*, pp. 283-318. Edited by E. Harlow and D. Lane. Cold Spring Harbor Press.
- Hartsuck, J.A., Koelsch, G. and Remington, S.J. (1992). The high-resolution crystal structure of porcine pepsinogen. *Proteins: Structure, function and genetics* **13**, 1-25.
- Hazen, K.C., Brawner, D.L., Riesselman, M.H., Jutila, M.A. and Cutler, J.E. (1991). Differential adherence of hydrophobic and hydrophilic *Candida albicans* yeast cells to mouse tissues. *Infection and Immunity* **59**, 907-912.
- Hazen, K.C., Lay, J.-G., Hazen, W., Fu, R.C. and Murthy, S. (1990). Partial biochemical characterisation of cell surface hydrophobicity and hydrophilicity of *Candida albicans*. *Infection and Immunity* **58**, 3469-3476.

- Hermann, V.M. (1996). Studies on inhibitor binding to secreted aspartic proteinases of *Candida albicans*. BSc. (Hons) thesis (Biochemistry), University of Otago, Dunedin, New Zealand.
- Herskowitz, I. (1995). Map kinase pathways in yeast: for mating and more. *Cell* **80**, 187-197.
- Hoegl, L., Ollert, M. and Korting, H.C. (1996). The role of *Candida albicans* secreted aspartic proteinase in the development of candidoses. *Journal of Molecular Medicine* **74**, 135-142.
- Homma, M., Chibana, H. and Tanaka, K. (1993). Induction of extracellular proteinase in *Candida albicans*. *Journal of General Microbiology* **139**, 1187-1193.
- Homma, M., Kanbe, T., Chibana, H. and Tanaka, K. (1992). Detection of intracellular forms of secretory aspartic proteinase in *Candida albicans*. *Journal of General Microbiology* **138**, 627-633.
- Horn, R., Wong, B., Kiehn, T.E. and Armstrong, D. (1985). Fungemia in a cancer hospital: changing frequency, earlier onset, and results of therapy. *Reviews in Infectious Diseases* **7**, 646-655.
- Howlett, J.A. and Squier, C.A. (1980). *Candida albicans* ultrastructure: colonization and invasion of oral epithelium. *Infection and Immunity* **29**, 252-260.
- Hube, B., Monod, M., Schofield, D.A., Brown, A.J.P. and Gow, N.A.R. (1994). Expression of seven members of the gene family encoding secretory aspartyl proteinases in *Candida albicans*. *Molecular Microbiology* **14**, 87-99.
- Hube, B., Sanglard, D., Odds, F.C., Hess, D., Monod, M., Schafer, W., Brown, A.J.P. and Gow, N.A.R. (1997). Disruption of each of the secreted aspartyl proteinase

- genes *SAP1*, *SAP2*, and *SAP3* of *Candida albicans* attenuates virulence. *Infection and Immunity* **65**, 3529-3538.
- Hube, B., Sanglard, D., Schaller, M., Ibrahim, A., Odds, F.C. and Gow, N.A.R. (1998). What are the functions of six different genes encoding secreted proteinases in *Candida albicans*? *Mycoses* **41**, 47-50
- Hube, B., Turver, C.J., Odds, F.C., Eiffert, H., Boulnois, G.J., Kochel, H. and Ruchel, R. (1991). Sequence of the *Candida albicans* gene encoding the secretory aspartate proteinase. *Journal of Medical and Veterinary Mycology* **29**, 129-132.
- Inoue, H., Nojima, H. and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23-26.
- Invitrogen. (1996). A manual of methods for expression of recombinant proteins in *Pichia pastoris*. Pichia expression kit: Protein expression version A.
- Jae, H.-S., Baker, W., Goldman, R., Capobianco, B., Plattner, B. and Plattner, J. (1992). The discovery of potent inhibitors of *Candida albicans* aspartyl proteinase. *Abstracts of the American Chemical Society* **204**, 150.
- James, M., Moore, S., Sielecki, A., Chernaia, M. and Tarasova, N. (1995). The molecular structure of human progastricsin and its comparison with that of porcine pepsinogen. *Advances in Experimental Medicine and Biology* **362**, 11-18.
- James, M.N.G. and Sielecki, A. (1983). Structure and refinement of penicillopepsin at 1.8 Å resolution. *Journal of Molecular Biology* **163**, 299-301.
- James, M.N.G., Hsu, I-N. and Delbaere, L.T.J. (1977). Mechanism of acid protease catalysis based on the crystal structure of penicillopepsin. *Nature* **267**, 808-813.

- Jancarik, J. and Kim, S.H. (1991). Sparse matrix sampling: A screening method for crystallization of proteins. *Journal of Applied Crystallography* **24**, 409-411.
- Jenkins, J.A., Tickle, I.J. and Blundell, T.L. (1977). Homology among acid proteases: comparison of crystal structure at 3 Å resolution of acid proteases from *Rhizopus chinensis* and *Endothia parasitica*. *Proceedings of the National Academy of Sciences* **74**, 556-559.
- Jones, E.W. (1991). Tackling the protease problem in *Saccharomyces cerevisiae*. In *Methods in Enzymology*, pp. 444-446, Vol 194. Edited by C. Guthrie & G.R. Fink. Academic Press Inc. San Diego.
- Jones, S., White, G. and Hunter, P.R. (1994). Increased phenotypic switching in strains of *Candida albicans* associated with invasive infections. *Journal of Clinical Microbiology* **32**, 2869-2870.
- Jung, E., Gooley, A.A., Packer, N.H., Karuso, P. and Williams, K.L. (1998). Rules for addition of ●-linked N-acetylglucosamine to secreted proteins in *Dictyostelium discoideum* - *in vivo* studies on glycosylation of mucin MUC1 and MUC2 repeats. *European Journal of Biochemistry* **253**, 517-524.
- Kagaya, K. and Fukuzawa, Y. (1981). Murine defense mechanism against *Candida albicans* infection. II Opsonization, phagocytosis and intracellular killing of *Candida albicans*. *Microbiology and Immunology* **25**, 807-818.
- Kaminishi, H., Hagihara, Y., Tanaka, M. and Cho, T. (1988). Degradation of bovine achilles tendon collagen by *Candida albicans* proteinase. *Journal of Medical and Veterinary Mycology* **58**, 2139-2143.
- Kaminishi, H., Miyaguchi, H., Tamaki, T., Suenaga, N., Hisamatsu, M., Mihashi, I., Matsumoto, H., Maeda, H. and Hagihara, Y. (1995). Degradation of humoral host defence by *Candida albicans* proteinase. *Infection and Immunity* **63**, 984-988.

- Kay, J. (1985). Aspartic proteinases and their inhibitors. In: *Aspartic proteinases and their inhibitors*. pp 1-17. Kostka, V. (ed.), Berlin: Walter de Gruyter and Co.
- Kay, J. and Dunn, B.M. (1990). Viral proteinases: weakness in strength. *Biochimica et Biophysica Acta* **1048**, 1-18.
- Kennedy, M.J., Rogers, A.L., Hanselman, L.R., Soll, D.R. and Yancey, R.J. (1988). Variation in adhesion and cell surface hydrophobicity in *Candida albicans* white and opaque phenotypes. *Mycopathologia* **102**, 149-156.
- Keohavong, P. and Thilly, W.G. (1989). Fidelity of DNA polymerases in DNA amplification. *PNAS* **86**, 9253-9257.
- Khan, A.R. and James, M.N.G. (1998). Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Science* **7**, 815-836.
- King, R.D., Lee, J.C. and Morris, A.L. (1980). Adherence of *Candida albicans* and other *Candida* species to mucosal epithelial cells. *Infection and Immunity* **27**, 667-674.
- Klotz, S.A., Drutz, D.J. and Zajic, J.E. (1985). Factors governing adherence of *Candida* species to plastic surfaces. *Infection and Immunity* **50**, 97-101
- Klotz, S.A., Hein, R.C., Smith, R.L. and Rouse, J.B. (1994). The fibronectin adhesin of *Candida albicans*. *Infection and Immunity* **62**, 4679-4681.
- Koelsch, G., Mares, M., Metcalf, P. and Fusek, M. (1994). Multiple functions of pro-parts of aspartic proteinase zymogens. *FEBS Letters* **343**, 6-10.
- Koelsch, G., Tang, J., Monod, M., Foundling, S.I. and Lin, X. (1998). Primary substrate specificities of secreted aspartic proteinases of *Candida albicans*. In *Aspartic Proteinases*. Ed. James. Plenum Press, New York.

- Kohl, N.E., Emini, E.A., Schleif, W.A., Davies, L.J., Heimback, J.C., Dixon, R.A.F., Scolnick, E.M. and Sigal, I.S. (1988). Active human immunodeficiency virus protease is required for viral infectivity. *Proceedings of the National Academy of Sciences* **85**, 4686-4690.
- Kohler, J.R. and Fink, G.R. (1996). *Candida albicans* strains heterozygous and homozygous for mutations in mitogen-activated protein kinase signaling components have defects in hyphal development. *Proceedings of the National Academy of Science* **93**, 13223-13228.
- Kolotila, M.P. and Diamond, R.D. (1990). Effects of neutrophils and in vitro oxidants on survival and phenotypic switching of *Candida albicans* WO-1. *Infection and Immunity* **58**, 1174-1179.
- Kuo, M. and Grayhack, E. (1994). A library of yeast genomic *MCM1* binding sites contains genes involved in cell cycle control, cell wall and membrane structure, and metabolism. *Molecular and Cellular Biology* **17**, 348-359.
- Kurtz, M.B. and Douglas, C.M. (1997). Lipopeptide inhibitors of fungal glucan synthase. *Journal of Medical and Veterinary Mycology* **35**, 79-86.
- Kwon-Chung, K.J., Lehman, D., Good, C. and Magee, P.T. (1985). Genetic evidence for role of extracellular proteinase in virulence of *Candida albicans*. *Infection and Immunity* **49**, 571-575
- Laemmli, U.K. and Favre, M. (1973). Maturation of the head of bacteriophage T4. *Journal of Molecular Biology* **80**, 575-600
- Lam, P.Y.S., Jadhav, P.K., Eyermann, C.J., Hodge, C.N., Ru, Y., Bacheler, L.T., Meek, J.L., Otto, M.J., Rayner, M.M., Wong, N., Chang, C-H., Weber, P.C., Jackson, D.A., Sharpe, T.R., Erickson-Viitanen, S. (1994). Rational design of potent,

bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **263**, 380-384.

Leberer, E., Harcus, D., Broadbent, I.D., Clark, K.L., Dignard, D., Ziegelbauer, K., Schmidt, A., Gow, N.A.R., Brown, A.J.P. and Thomas, D.Y. (1996). Signal transduction through homologs of the Ste20p and Ste7p protein kinases can trigger hyphal formation in the pathogenic fungus *Candida albicans*. *Proceedings of the National Academy of Science* **93**, 13217-13222.

Leberer, E., Ziegelbauer, K., Schmidt, A., Harcus, D., Dignard, D., Ash, J., Johnson, L. and Thomas, D.Y. (1997). Virulence and hyphal formation of *Candida albicans* require the Ste20p-like protein kinase CaCl4p. *Current Biology* **7**, 539-546.

Lee, K.L., Buckley, H.R. and Campbell, C. (1975). An amino acid liquid synthetic medium for development of mycelial and yeast forms of *Candida albicans*. *Sabouraudia*, **13**: 148-153.

Leidich, S.D., Ibrahim, A.S., Fu, Y., Koul, A., Jessup, C., Vitullo, F., Fonzi, W., Mirbod, F., Nakashima, S., Nozawa, Y. and Ghannoum, M.A. (1998). Cloning and disruption of caPLB1, a phospholipase B gene involved in the pathogenicity of *Candida albicans*. *Journal of Biological Chemistry* **273**, 26078-26086

Lerner, C.G. and Goldman, R.C. (1993). Stimuli that induce production of *Candida albicans* extracellular aspartyl proteinase. *Journal of General Microbiology* **139**, 1643-1651.

Lin, X., Tang, J., Koelsch, G., Monod, M. and Foundling, S. (1993). Recombinant candidopsin, an extracellular aspartic proteinase from yeast *Candida tropicalis*. *Journal of Biological Chemistry* **268**, 20143-20147.

Liu, H., Styles, C. and Fink, G.R. (1993). Elements of the yeast pheromone response pathway required for filamentous growth of diploids. *Science* **262**, 1741-1744.

- Lo, H-J., Kohler, J.R., DiDomenico, B., Loebenberg, D., Cacciapuoti, A. and Fink, G.R. (1997). Nonfilamentous *C. albicans* mutants are avirulent. *Cell* **90**, 939-949.
- Lockhart, S.R., Nguyen, M., Srikantha, T. and Soll, D.R. (1998). A MADS box protein consensus binding site is necessary and sufficient for activation of the opaque-phase-specific gene *OP4* of *Candida albicans*. *Journal of Bacteriology* **180**, 6607-6616
- Lodge, J.K., Jackson-Machelski, E., Devadas, B., Zupec, M.E., Getman, D.P., Kishore, N., Freeman, S.K., McWherter, C.A., Sikorski, J.A. and Gordon, J.I. (1997). N-myristoylation of Arf proteins in *Candida albicans*: an *in vivo* assay for evaluating antifungal inhibitors of myristoyl-CoA:protein N-myristoyltransferase. *Microbiology* **143**, 357-366.
- Louie, A., Dixon, D.M., El-Maghrabi, E.A., Burnett, J.W., Baltch, A.L. and Smith, R.P. (1994). Relationship between *Candida albicans* epidermolytic proteinase activity and virulence in mice. *Journal of Medical and Veterinary Mycology* **32**, 59-64.
- Lowther, W.T. and Dunn, B.M. (1995). Site-directed mutagenesis of rhizopuspepsin: an analysis of unique specificity. In: *Aspartic proteinases: structure, function, biology, and biomedical implications*. Takahashi, K. (ed.), Plenum Press, New York.
- MacDonald, F. and Odds, F.C. (1980). Inducible proteinase of *Candida albicans* in diagnostic serology and in the pathogenesis of systemic candidosis. *Journal of Medical Microbiology* **13**, 423-435.
- Mackenzie, L.F., Brooke, G.S., Cutfield, J.F., Sullivan, P.A. and Withers, S.G. (1997). Identification of Glu-330 as the catalytic nucleophile of *Candida albicans* exo-beta-(1,3)-glucanase. *Journal of Biological Chemistry* **272**, 3161-3167.

- Magee, P.T. (1990). The use of genetic analysis in the study of the biology of *Candida albicans*. In: *The Genetics of Candida*. Eds. Kirsch, D.R., Kely, R. and Kurtz, M.B. CRC Press. Boca Raton.
- Magee, P.T. (1997). Which came first, the hypha or the yeast? *Science* **277**, 52-53.
- Marquis, G., Garzon, S., Montplaisir, S., Strykowski, H. and Benhamou, N. (1991). Histochemical and immunochemical study of the fate of *Candida albicans* inside human neutrophil phagolysosomes. *Journal of Leukocyte Biology* **50**, 587-599
- Matthews, R. and Burnie, J. (1996). Antibodies against *Candida*: potential therapeutics? *Trends in Microbiology* **4**, 354-358.
- McEachern, M.J. and Hicks, J.B. (1993). Unusually large telomeric repeats in the yeast *Candida albicans*. *Molecular and Cellular Biology* **13**, 551-560.
- Meunier-Carpentier, F., Kiehn, T.E. and Armstrong, D. (1981). Fungemia in the immunocompromised host. Changing patterns, antigenemia, high mortality. *American Journal of Medicine* **71**, 363-370.
- Milewski, S., Andruskiewicz, R. and Borowski, E. (1988). Substrate specificity of peptide permeases in *Candida albicans*. *FEMS Microbiology Letters* **50**, 73-78.
- Miller, M., Jaskolski, M., Rao, J.K.M., Leis, J. and Wlodawer, A. (1989). Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337**, 576-579.
- Miyada, C.G. and Wallace, R.B. (1987). Oligonucleotide hybridisation techniques. *Methods in Enzymology* **154**, 94-107.

- Miyasaki, S.H., White, T.C. and Agabian, N. (1994). A fourth secreted aspartyl proteinase gene (*SAP4*) and a *CARE2* element are located upstream of the *SAP1* gene in *Candida albicans*. *Journal of Bacteriology* **176**, 1702-1710.
- Monod, M., Togni, G., Hube, B. and Sanglard, D. (1994). Multiplicity of genes encoding secreted aspartic proteinases in *Candida* species. *Molecular Microbiology* **13**, 357-368.
- Monod, M., Hube, B., Hess, D. and Sanglard, D. (1998). Differential regulation of *SAP8* and *SAP9*, which encode two new members of the secreted aspartic proteinase family in *Candida albicans*. *Microbiology* **144**, 2731-2737.
- Moore, S.A., Sielecki, A.R., Chernaia, M.M., Tarasova, N.I. and James, M.N.G. (1995). Crystal and molecular structures of human progastricsin at 1.62 Å resolution. *Journal of Molecular Biology* **247**, 466-485.
- a) Morrison, C.J., Hurst, S.F., Bragg, S.L., Kuykendall, R.J., Diaz, H., Pohl, J. and Reiss, E. (1993). Heterogeneity of the purified extracellular aspartyl proteinase from *Candida albicans*: characterisation with monoclonal antibodies and N-terminal amino acid sequence analysis. *Infection and Immunity* **61**, 2030-2036.
- b) Morrison, C.J., Hurst, S.F., Bragg, S.L., Kuykendall, R.J., Diaz, H., McLaughlin, D.W. and Reiss, E. (1993). Purification and characterization of the extracellular aspartyl proteinase of *Candida albicans*: Removal of extraneous proteins and cell wall mannoprotein and evidence for lack of glycosylation. *Journal of General Microbiology* **139**, 1177-1186.
- Morrow, B., Anderson, J., Wilson, E. and Soll, D.R. (1989). Bidirectional stimulation of the white-opaque transition of *Candida albicans* by ultraviolet irradiation. *Journal of General Microbiology* **135**, 1201-1208.

Morrow, B., Ramsey, H. and Soll, D.R. (1994). Regulation of phase-specific genes in the more general switching system of *Candida albicans* strain 3153A. *Journal of Medical and Veterinary Mycology* **32**, 287-294.

Morrow, B., Srikantha, T. and Soll, D.R. (1992). Transcription of the gene for a pepsinogen, *PEP1*, is regulated by white-opaque switching in *Candida albicans*. *Molecular and Cellular Biology* **12**, 2997-3005.

Morrow, B., Srikantha, T., Anderson, J. and Soll, D.R. (1993). Coordinate regulation of two opaque-phase specific genes during white-opaque switching in *Candida albicans*. *Infection and Immunity* **61**, 1823-1828.

Navia, M.A., Fitzgerald, P.M.D., McKeever, B.M., Leu, C-T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L. and Springer, J.P. (1990). Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615-620.

Negi, M., Tsuboi, R., Matsui, T. and Ogawa, H. (1984). Isolation and characterisation of proteinase from *Candida albicans*: substrate specificity. *Journal of Investigative Dermatology* **83**, 32-36.

Nesterenko, M.V., Tilley, M. and Upton, S.J. (1994). A simple modification of Blum's silver stain method allows for 30 min detection of proteins in polyacrylamide gels. *Journal of Biochemical and Biophysical Methods* **28**, 239-242.

Newport, G. and Agabian, N. (1997). *KEX2* influences *Candida albicans* proteinase secretion and hyphal formation. *Journal of Biological Chemistry* **272**, 28954-28961.

Nguyen, M.H., Peacock, J.E., Morris, A.J., Tanner, D.C., Nguyen, M.L., Snyderman, D.R., Wagener, M.M., Rinaldi, M.G. and Yu, V.L. (1996). The changing face of candidemia: emergence of non-*Candida albicans* species and antifungal resistance. *The American Journal of Medicine* **100**, 617-623.

- Niimi, M., Shepherd, M.G., Monk, B.C. (1996). Differential profiles of soluble proteins during the initiation of morphogenesis in *Candida albicans*. *Archives of Microbiology* **166**, 260-268.
- North, M.J. (1982). Comparative biochemistry of the proteinases of eukaryotic microorganisms. *Microbiological Reviews* **46**, 308-340.
- Nothwehr, S.F. and Gordon, J.I. (1990). Targeting of proteins into the eukaryotic secretory pathway: signal peptide structure/function relationships. *BioEssays* **12**, 479-484.
- Odds, F.C. (1985). Morphogenesis in *Candida albicans*. *Critical Reviews in Microbiology* **12**, 45-93.
- Odds, F.C. (1988). *Candida and Candidosis: a review and bibliography*, 2nd edn. London: Balliere and Tindall.
- Oehlen, L., McKinney, J. and Cross, F. (1996). Ste12 and Mcm1 regulate cell cycle-dependent transcription of *FAR1*. *Molecular and Cellular Biology* **16**, 2830-2837.
- Ogrydziak, D.M. (1993). Yeast extracellular proteases. *Critical Reviews in Biotechnology* **13**, 1-55.
- Olaiya, A.F. and Sogin, S.J. (1979). Ploidy determination of *Candida albicans*. *Journal of Bacteriology* **141**, 1284-1290.
- Ollert, M.W., Kadlec, J.V., David, K., Petrella, E.C. Bredehorst, R. and Vogel, C.W. (1994). Antibody-mediated complement activation on nucleated cells. A quantitative analysis of the individual reaction steps. *Journal of Immunology* **153**, 2213-2221.

- Ollert, M.W., Wende, C., Gorlich, M., McMullan-Vogel, C.G., Borg-von Zepelin, M., Vogel, C-W. and Korting, H.C. (1995). Increased expression of *Candida albicans* secretory proteinase, a putative virulence factor, in isolates from human immunodeficiency virus-positive patients. *Journal of Clinical Microbiology* **33**, 2543-2549.
- Panagoda, G.J., Ellepola, A.N. and Samaranayake, L.P. (1998). Adhesion to denture acrylic surfaces and relative cell-surface hydrophobicity of *Candida parapsilosis* and *Candida albicans*. *APMIS* **106**, 736-742.
- Pearl, L.H. (1987). The catalytic mechanism of aspartic proteinases. *FEBS Letters* **214**, 8-12.
- Pfaller, M.A., Jones, R.N., Doern, G.V., Sader, H.S., Hollis, R.J. and Messer, S.A. (1998). International surveillance of bloodstream infections due to *Candida* species: frequency of occurrence and antifungal susceptibilities of isolates collected in 1997 in the United States, Canada and South America for the SENTRY program. The SENTRY participant group. *Journal of Clinical Microbiology* **36**, 1886-1889
- Piller, F. and Piller, V. (1993). Structural characterization of mucin-type *O*-linked oligosaccharides. In *Glycobiology: A practical approach*. Edited by M. Fukuda and A. Kobata. Oxford University Press.
- Pilus, L. and Rine, J. (1989). Epigenetic inheritance of transcription states of *S. cerevisiae*. *Cell* **59**, 637-647.
- Pomes, R., Gil, C. and Nombela, C. (1985). Genetic analysis of *Candida albicans* morphological mutants. *Journal of General Microbiology* **131**, 2107-2113.
- Ram, S.P., Romana, L.K., Shepherd, M.G. and Sullivan, P.A. (1984). Exo-(1-3)-beta-glucanase, autolysin and trehalase activities during yeast growth and germ-tube formation in *Candida albicans*. *Journal of General Microbiology* **130**, 1227-1236.

Ramsey, H., Morrow, B. and Soll, D.R. (1994). An increase in switching frequency correlates with an increase in recombination of the ribosomal chromosomes of *Candida albicans* strain 3153A. *Microbiology* **140**, 1525-1531.

Ray, T.L. and Payne, C.D. (1988). Scanning electron microscopy of epidermal adherence and cavitation in murine candidiasis: a role for *Candida* acid proteinase. *Infection and Immunity* **56**, 1942-1949.

Remold, H., Fasold, H. and Staib, F. (1968). Purification and characterisation of a proteolytic enzyme from *Candida albicans*. *Biochemica et Biophysica Acta*. **167**, 399-406.

Ridky, T.W., Cameron, C.E., Cameron, J., Leis, J., Copeland, T., Wlodawer, A., Weber, I.T. and Harrison, R.W. (1996). Human immunodeficiency virus, type 1 protease substrate specificity is limited by interactions between substrate amino acids bound in adjacent enzyme subsites. *Journal of Biological Chemistry* **271**, 4709-4717.

Rikkerink, E.H., Magee, B.B. and Magee, P.T. (1988). Opaque-white phenotype transition: a programmed morphological transition in *Candida albicans*. *Journal of Bacteriology* **170**, 895-899.

Roberts, N.A., Martin, J.A., Kinchington, D., Broadhurst, A.V., Craig, J.C., Duncan, I.B., Galpin, S.A., Handa, B.K., Kay, J., Krohn, A., Lambert, R.W., Merrett, J.H., Mills, J.S., Parkes, K.E.B., Redshaw, S., Ritchie, A.J., Taylor, D.L., Thomas, G.J., Machin, P.J. (1990). Rational design of peptide-based HIV proteinase inhibitors. *Science* **248**, 358-361.

Ross, I.K., De Bernardis, F., Emerson, G.W., Cassone, A. and Sullivan, P.A. (1990). The secreted aspartate proteinase of *Candida albicans*: Physiology of secretion and virulence of a proteinase-deficient mutant. *Journal of General Microbiology* **136**, 687-694.

- Rotrosen, D., Edwards, J.E.Jr., Gibson, T.R., Moore, J.C., Cohen, A.H. and Green, I. (1985). Adherence of *Candida* to cultured vascular endothelial cells: mechanism of attachment and endothelial cell penetration. *Journal of Infectious Diseases* **152**, 1264-1274.
- Ruchel, R. (1981). Properties of a purified proteinase from the yeast *Candida albicans*. *Biochimica et Biophysica Acta* **659**, 99-113.
- Ruchel, R. (1986). Cleavage of immunoglobulins by pathogenic yeasts of the genus *Candida*. *Microbiological Sciences* **3**, 316-319.
- Ruchel, R., De Bernardis, F., Ray, T.L., Sullivan, P.A. and Cole, G.T. (1992). *Candida* acid proteinases. *Journal of Medical and Veterinary Mycology* **30**, 123-132.
- Samaranayake, L.P., Hughes, A. and MacFarlane, T.W. (1984). The proteolytic potential of *Candida albicans* in human saliva supplemented with glucose. *Journal of Medical Microbiology* **17**, 13-22
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: a laboratory manual*. Cold Spring Harbour Laboratory.
- Sanglard, D., Hube, B., Monod, M., Odds, F.C. and Gow, N.A.R. (1997). A triple deletion of the secreted aspartyl proteinase genes *SAP4*, *SAP5*, and *SAP6* of *Candida albicans* causes attenuated virulence. *Infection and Immunity* **65**, 3539-3546.
- Santos, M.A., Keith, G. and Tuite, M.F. (1993). Non-standard translational events in *Candida albicans* mediated by an unusual seryl-tRNA with a 5'-CAG-3' (leucine) anticodon. *EMBO Journal* **12**, 607-616.
- Scarborough, P.E., Richo, G.R., Kay, J., Connor, G.E. and Dunn, B.M. (1991). In *Structure and Function of the aspartic proteinases - genetics, structures and mechanisms*, pp 343-347. Dunn, B.M. (ed.), Plenum press, New York.

Schaller, M., Hube, B., Ollert, M.W., Schafer, W., Borg-von Zepelin, M., Thoma-Greber, E. and Korting, H.C. (1999). *In vivo* expression and localization of *Candida albicans* secreted aspartyl proteinases during oral candidosis in HIV-infected patients. *Journal of Investigative Dermatology* **112**, 383-386.

Schreiber, B., Lyman, C.A., Gurevich, J. and Needham, C.A. (1985). Proteolytic activity of *Candida albicans* and other yeasts. *Diagnostic Microbiology Infectious Disease* **3**, 1-5

Sham, H.L., Betebenner, D.A., Wideburg, N.E., Saldivar, A.C., Kohlbrenner, W.E., Vasavanonda, S., Kempf, D.J., Norbeck, D.W., Zhao, C., Clement, J.J., Erickson, J.E., Plattner, J.J. (1991). Potent HIV-1 protease inhibitors with antiviral activities *in vitro*. *Biochemical and Biophysical Research Communications* **175**, 914-919.

Shearer, G.M. and Clerici, M. (1992). T helper cell immune dysfunction in asymptomatic, HIV-1-seropositive individuals: the role of TH1-TH2 cross-regulation. *Chemical Immunology* **54**, 21-43.

Shuster, J.R. (1991). Gene expression in yeast: protein secretion. *Current Opinion in Biotechnology* **2**, 685-690.

Sielecki, A.R., Hayakawa, K., Fujinaga, M., Murphy, M.E.P., Fraser, M., Muir, A.K., Carilli, C.T., Lewicki, J.A., Baxter, J.D., James, M.N.G. (1989). Structure of recombinant human renin, a target for cardiovascular-active drugs, at 2.5 Å resolution. *Science* **243**, 1346-1351.

Sikorski, J.A., Devadas, B., Zupec, M.E., Freeman, S.K., Brown, D.L., Lu, H-F., Nagarajan, S., Mahta, P.P., Wade, A.C., Kishore, N.S., Bryant, M.L., Getman, D.P., McWherter, C.A. and Gordon, J.I. (1997). Selective peptidic and peptidomimetic inhibitors of *Candida albicans* MyristoylCoA:Protein N-Myristoyltransferase: a new approach to antifungal therapy. *Biopolymers* **43**, 43-71.

- Slutsky, B., Buffo, J. and Soll, D.R. (1985). High-frequency switching of colony morphology in *Candida albicans*. *Science* **230**, 666-669.
- Slutsky, B., Staebell, M., Anderson, J., Risen, L., Pfaller, M. and Soll, D.R. (1987). "White-opaque transition": a second high frequency switching system in *Candida albicans*. *Journal of Bacteriology* **169**, 189-197.
- Smolenski, G., Sullivan, P.A., Cutfield, S.M. and Cutfield, J.F. (1997). Analysis of secreted aspartic proteinases from *Candida albicans*: purification and characterisation of individual Sap1, Sap2 and Sap3 isoenzymes. *Microbiology* **143**, 349-356.
- Sobel, J.D., Muller, G. and Buckley, H.R. (1984). Critical role of germ-tube formation in the pathogenesis of candidal vaginitis. *Infection and Immunity* **44**, 576-580.
- Soll, D.R. (1997). Gene regulation during high-frequency switching in *Candida albicans*. *Microbiology* **143**, 255-263.
- Soll, D.R. (1992). High frequency switching in *Candida albicans*. *Clinical Microbiology Reviews* **5**, 183-203.
- Soll, D.R., Anderson, J. and Bergen, M. (1991). The developmental biology of the white-opaque transition in *Candida albicans*. In *The molecular biology of Candida albicans*. pp20-45. R. Prasad (ed). Springer-Verlag, Berlin.
- Soll, D.R., Galask, R., Isly, S., Rao, T.V., Store, D., Hicks, J., Schmid, J., Mac, K. and Hanna, C. (1989). Switching of *Candida albicans* during successive episodes of recurrent vaginitis. *Journal of Clinical Microbiology* **27**, 681-690.
- Soll, D.R., Langtimm, C.J., McDowell, J., Hicks, J. and Galask, R. (1987). High-frequency switching in *Candida* strains isolated from vaginitis patients. *Journal of Clinical Microbiology* **25**, 1611-1622.

- Soll, D.R., Srikantha, T., Morrow, B., Chandrasekhar, A., Schroppel, K. and Lockhart, S. (1995). Gene regulation in the white-opaque transition of *Candida albicans*. *Canadian Journal of Botany* **73** (Suppl. 1), S1049-S1057.
- Sprague, G. (1990). Combinatorial associations of regulatory proteins and the control of cell type in yeast. *Advances in Genetics* **27**, 33-62.
- Srikantha, T. and Soll, D.R. (1993). A white-specific gene in the white-opaque switching system of *Candida albicans*. *Gene* **31**, 53-60.
- Srikantha, T., Chanrasekhar, A. and Soll, D.R. (1995). Functional analysis of the promoter of the phase-specific *WH11* gene of *Candida albicans*. *Molecular and Cellular Biology* **15**, 1797-1805.
- Staib, F. (1965). Serum-proteins as nitrogen source for yeastlike fungi. *Sabouraudia*, **4**:187-193.
- Stoldt, V.R., Sonneborn, A., Leuker, C.E. and Ernst, J.F. (1997). Efg1p, an essential regulator of morphogenesis of the human pathogen *Candida albicans*, is a member of a conserved class of bHLH proteins regulating morphogenetic processes in fungi. *EMBO Journal* **16**, 1982-1991.
- Strahl-Bolsinger, S., Gentsch, M. and Tanner, W. (1999). Protein O-mannosylation. *Biochimica et Biophysica Acta* **1426**, 297-307.
- Suguna, K., Padlan, E.A., Smith, C.W., Carlson, W.D. and Davies, D.R. (1987). Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis*: implication for a mechanism of action. *Proceedings of the National Academy of Sciences* **84**, 7009-7013.
- Tang, J. and Lin, X. (1994). Engineering aspartic proteases to probe structure and function relationships. *Current Opinion in Biotechnology* **5**, 422-427.

- Tang, J., James, M.N.G., Hsu, I.N., Jenkins, J.A., Blundell, T.L. (1978). Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* **271**, 618-621.
- Togni, G., Sanglard, D. and Monod, M. (1994). Acid proteinase secreted by *Candida tropicalis*: virulence in mice of a proteinase negative mutant. *Journal of Medical and Veterinary Mycology* **32**, 257-265
- Togni, G., Sanglard, D., Quadroni, M., Foundling, S.I. and Monod, M. (1996). Acid proteinase secreted by *Candida tropicalis*: Functional analysis of preproregion cleavages in *C. tropicalis* and *Saccharomyces cerevisiae*. *Microbiology* **142**, 493-503.
- Tsuboi, R., Ogawa, H., Bramono, K., Richardson, M.D., Shankland, G.S., Crozier, W.J., Sei, Y., Ninomiya, J., Nakabayashi, A., Takaiuchi, I., Payne, C.D. and Ray, T.L. (1994). Pathogenesis of superficial mycoses. *Journal of Medical and Veterinary Mycology* **32**, 91-104.
- Umezawa, H., Aoyagi, T., Morishima, H., Matsuzaki, M. and Hamada, M. (1970). Pepstatin, a new pepsin inhibitor produced by Actinomycetes. *Journal of Antibiotics* **23**, 259-262.
- Vanden Bossche, H. (1991). Anti-*Candida* drugs: mechanism of action, pp. 83-95. In: *Candida and Candidamycesis*. Edited by. E. Tumbay, H.P.R. Seeliger and O. Ang.
- Vedvick, T.S. (1991). Gene expression in yeast: *Pichia pastoris*. *Current Opinion in Biotechnology* **2**, 742-745.
- Wagner, T., Borg-von Zepelin, M. and Ruchel, R. (1995). pH-dependent denaturation of extracellular aspartic proteinases from *Candida* species. *Journal of Medical and Veterinary Mycology* **33**, 275-278.

- Walsh, T.J. and Giri, N. (1997). Pradimicins: a novel class of broad-spectrum antifungal compounds. *European Journal of Clinical Microbiological Infectious Diseases* **16**, 93-97.
- Wang, C., Eufemi, M., Turano, C. and Giartosio, A. (1996). Influence of the carbohydrate moiety on the stability of glycoproteins. *Biochemistry* **35**, 7299-307
- Watts, H.J., Cheah, F.S.H., Hube, B., Sanglard, D. and Gow, N.A.R. (1998). Altered adherence in strains of *Candida albicans* harbouring null mutations in secreted aspartic proteinase genes. *FEMS Microbiology Letters* **159**, 129-135
- Weinberg, R.A., McWherter, C.A., Freeman, S.K., Wood, D.C., Gordon, J.I. and Lee, S.C. (1995). Genetic studies reveal that myristoylCoA:protein N-myristoyltransferase is an essential enzymes in *Candida albicans*. *Molecular Microbiology* **16**, 241-250.
- Whelan, W.L., Partridge, R.M. and Magee, P.T. (1980). Heterozygosity and segregation in *Candida albicans*. *Molecular and General Genetics* **180**, 107-113.
- Whelan, W.T. and Magee, P.T. (1981). Natural heterozygosity in *Candida albicans*. *Journal of Bacteriology* **145**, 896-903.
- White, T.C. and Agabian, N. (1995). *Candida albicans* secreted aspartyl proteinases: Isozyme pattern is determined by cell type, and levels are determined by environmental factors. *Journal of Bacteriology* **177**, 5215-5221.
- White, T.C., Andrews, L.E., Maltby, D. and Agabian, N. (1995). The “universal” leucine codon CTG in the secreted aspartyl proteinase 1 (*SAP1*) gene of *Candida albicans* encodes a serine *in vivo*. *Journal of Bacteriology* **177**, 2953-2955.

b) White, T.C., Kohler, G.A., Miyasaki, S.H. and Agabian, N. (1995). Expression of virulence factors in *Candida albicans*. *Canadian Journal of Botany* **73**(Suppl. 1), S1058-1064.

White, T.C., Marr, K.A. and Bowden, R.A. (1998). Clinical, cellular, and molecular factors that contribute to antifungal drug resistance. *Clinical Microbiology Reviews* **11**, 382-402

White, T.C., Miyasaki, S.H. and Agabian, N. (1993). Three distinct aspartyl proteinases in *Candida albicans*. *Journal of Bacteriology* **175**, 6126-6133.

Wright, R.J. (1994). *SAP2 of Candida albicans*. A thesis submitted for the degree of Doctor of Philosophy at the University of Otago, Dunedin, New Zealand.

Wright, R.J., Carne, A.J., Hieber, A.D., Lamont, I.L., Emerson, G.W. and Sullivan, P.A. (1992). A second gene for a secreted aspartate proteinase in *Candida albicans*. *Journal of Bacteriology* **174**, 7848-7853.