



The
University
Of
Sheffield.

Quantitative Proteomic Analysis of Kinase and Phosphatase Interactions in *Candida albicans*

By:

Iliyana Nancheva Kaneva

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Science
Department of Molecular Biology and Biotechnology

And

Faculty of Engineering
Department of Chemical and Biological Engineering

September 2016

Summary

The ability of the fungus *Candida albicans* to switch interchangeably between yeast and hyphal forms of growth contributes significantly to its pathogenesis. This morphogenetic shift occurs in response to environmental changes and it is accomplished by a complex network of signal transduction pathways. Protein kinases and phosphatases are important messengers in these pathways and several of them have been directly implicated in controlling *C. albicans* morphogenesis. Kinases and phosphatases (KP) are enzymes that modulate the function of their substrates via reversible phosphorylation and dephosphorylation, respectively. While the specificity of KP is tightly controlled, some enzymes can target a huge number of proteins and have a master regulatory role over various cell processes. The function of KP in *C. albicans* is poorly understood and methods for global analysis of KP interactions have not been adapted to this organism. This study developed a protocol for large scale analysis of protein interactions in *C. albicans* using immunoprecipitation and SILAC in conjunction with quantitative mass spectrometry analysis. The protocol was successfully applied for identification of Cdc14 interactors using the substrate-trapping mutant Cdc14C275S. Cdc14 is a phosphatase required for proper hyphal formation, cytoskeletal organisation and cell separation at the end of mitosis. This study reveals over 100 potential substrates of Cdc14 and new roles of the phosphatase in DNA damage repair, DNA replication, chromosome segregation and transcription regulation. In addition, experiments were performed separately with both yeast and hyphae allowing for direct comparison of Cdc14 interactome between both forms. Many of the identified proteins have unknown function and the significance of these putative interactions remains to be found.

Acknowledgments

I am very grateful to my supervisors Professor Pete Sudbery and Professor Mark Dickman for their guidance and support during my PhD project and for giving me the opportunity to finish my work after thesis submission.

I am thankful to Dr Stefan Milson for giving me valuable advice with protein purification techniques at the start of my PhD when I most needed it.

I also wish to thank Beata Florczak for useful discussion of SILAC and mass spectrometry while I preparing this thesis.

I would like to express my gratitude to Dr Phil Jackson, who thought me how to use a mass spectrometer and gave me assistance with troubleshooting various software problems.

I would like to thank Laura Hunt for ordering all lab consumables and for her assistance as a lab technician at the start of this project.

I also thank Dr Joseph Longworth for providing me with an R script for formatting of Distiller tables in Excel and for writing another R script for correcting arginine-to-proline conversion.

Further, I would like to thank the Life Sciences Team of Technology Transfer Managers at Oxford University Innovation for allowing me to do a placement at their company and for showing me the business side of science.

Finally, I would like to acknowledge my sponsor, BBSRC and The University of Sheffield for providing the funding and facilities for this project.

Abbreviations

5-FOA	5-fluoroorotic acid
ABC	ammonium bicarbonate
ACN	acetonitrile
AP	affinity purification
APC/C	anaphase promoting complex/cyclosome
APS	ammonium persulphate
Arg	arginine
AS	analogue-sensitive
ATP	adenosine triphosphate
AUX	auxotroph
bp	base pairs
BSA	bovine serum albumin
C	control
<i>C. albicans</i>	<i>Candida albicans</i>
<i>C. glabrata</i>	<i>Candida glabrata</i>
<i>C. krusei</i>	<i>Candida krusei</i>
<i>C. kyfer</i>	<i>Candida kyfer</i>
<i>C. parapsilosis</i>	<i>Candida parapsilosis</i>
<i>C. stellatoidea</i>	<i>Candida stellatoidea</i>
<i>C. tropicalis</i>	<i>Candida tropicalis</i>
CaCl ₂	calcium chloride
cAMP	cyclic adenosine monophosphate
CDK	cyclin-dependent kinase
cDNA	complementary DNA
Da	Daltons
DAPI	4',6-diamidino-2-phenylindole
ddH ₂ O	distilled deionised water
DIC	Differential interference contrast
DMSO	dimethyl sulfoxide

DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	exemplī grātiā (for example)
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionisation
et al.	et alia (and others)
EtBr	ethidium bromide
FDR	false discovery rate
FEAR	fourteen early anaphase release
fig.	figure
FTICR	fourier transform ion cyclotron resonance
g	gram
GAP	GTPase-activating protein
GEF	guanosine exchange factor
GFP	green fluorescent protein
GTP	guanosine triphosphate
GTPase	guanosine triphosphatase
H	heavy
hr	hour
i.e.	id est (that is)
ID	identification
IP	immunoprecipitation
KAc	potassium acetate
KP	kinase and phosphatase
KPI	kinase and phosphatase interactions
L	light
L	litre
LB	Luria-Bertani
LC	liquid chromatography
LiAc	lithium acetate
Lys	lysine

m	mass
MALDI	matrix-assisted laser desorption ionisation
MAPK	Mitogen-activated protein kinase
MEN	mitotic exit network
mg	milligram
min	minute
mM	millimolar
MnCl ₂	manganese chloride
mRNA	messenger ribonucleic acid
MS	mass spectrometry
NaAc	sodium acetate
NaN	not a number
NDR	nuclear Dbf2-related
NEB	New England Biolabs
nSILAC	native SILAC
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
PD	phosphatase-dead
PEG	polyethylene glycol
PKA	protein kinase A
PMF	peptide mass fingerprints
PMSF	phenylmethane sulfonyl fluoride
Pro	proline
PRO	prototroph
QO	quadrupole orbitrap
RAM	regulation of Ace2p activity and cellular morphogenesis
RbCl ₂	rubidium chloride
RENT	regulator of nucleolar silencing and telophase
RIA	rate of isotope abundance
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>

SAINT	Significance Analysis of INTERactome
sdH ₂ O	sterile water
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SILAC	stable isotope labelling of amino acids in cell culture
SPB	spindle pole bodies
TAE	tris base, acetic acid, EDTA
TAP	tandem affinity purification
TBS	tris-buffered saline
TE	Tris-EDTA
TOF	time-of-flight
v	volume
w	weight
WB	western blot
WT	wild type
YFP	yellow fluorescent protein
YNB	yeast nitrogen base
YPD	yeast extract, peptone and dextrose
z	charge

Contents

	Page
Summary	i
Acknowledgments	ii
Abbreviations	iii
Contents	vii
Figures	xi
Tables	xiii
1. Introduction	1
1.1. Introduction to <i>Candida albicans</i>	1
1.2. The role of protein phosphorylation and dephosphorylation in the cell	6
1.3. The role of protein kinases in <i>C. albicans</i> morphogenesis	7
1.3.1. Cdc28	7
1.3.2. Cbk1	8
1.3.3. Dbf2	9
1.3.4. Crk1	11
1.4. Protein phosphatases in <i>C. albicans</i>	13
1.4.1. General overview	13
1.4.2. Cdc14	13
1.5. Methods of identifying novel kinase and phosphatase interactions	19
1.5.1. Methods currently used in <i>C. albicans</i>	19
1.5.2. High-throughput techniques used in other species	19
1.6. Introduction to mass spectrometry	22
1.6.1. Sample preparation	22
1.6.2. Principles of MS instruments	22
1.6.3. Processing of MS data	24
1.6.4. Stable isotope labelling of amino acids in cell culture	24
1.7. Aims of this study	28
2. Materials and methods	29
2.1. Cell culture techniques	29

2.1.1. Growth media	29
2.1.2. Growth conditions	34
2.1.3. Cell transformations	35
2.1.4. Strain storage conditions	37
2.1.5. <i>C. albicans</i> strains used in this study	37
2.2. DNA techniques	40
2.2.1. PCR of plasmid DNA	40
2.2.2. Colony PCR	41
2.2.3. DNA precipitation	43
2.2.4. Agarose gel electrophoresis	43
2.2.5. Restriction digest	43
2.2.6. DNA gel extraction	44
2.2.7. DNA ligation	44
2.2.8. Plasmid DNA miniprep	44
2.2.9. DNA sequencing	44
2.2.10. Oligonucleotides used in this study	45
2.2.11. Plasmids used in this study	49
2.3. Protein techniques	50
2.3.1. Soluble protein extraction	50
2.3.2. SDS-PAGE	51
2.3.3. Western blot	52
2.3.4. Stripping of antibodies from a nitrocellulose membrane	53
2.3.5. In vitro protein dephosphorylation	53
2.3.6. Protein purification	54
2.3.7. Coomassie staining of polyacrylamide gels	55
2.3.8. Trypsin in-gel digestion of proteins and peptides gel extraction	55
2.4. Mass spectrometry	57
2.5. Microscopy	60
3. Optimisation strategies for protein purification and mass spectrometry	61
3.1. Introduction	61
3.1.1. Dbf2 and Mob1	61
3.1.2. Cdc14	62

3.1.3. Crk1	62
3.2. General overview of experimental workflow	63
3.3. Affinity purification	66
3.3.1. Generation of epitope-tagged strains	66
3.3.2. Selection of affinity matrix and tag	66
3.3.3. Optimisation of protein elution from beads	68
3.3.4. Optimisation of protein concentration in a cell lysate for IP	68
3.3.5. Optimisation of total protein amount in a cell lysate for IP	71
3.3.6. Pre-incubation of affinity beads with BSA does not reduce background	71
3.3.7. Optimisation of beads-washing steps after IP	71
3.4. Mass spectrometry analysis	76
3.4.1. Fractionation of eluted proteins followed by MS	76
3.4.2. Analysis of MS results using ProHits software	78
3.4.3. SAINT analysis of MS results	96
3.5. Discussion	97
4. Characterization of the substrate-trapping mutant Cdc14C275S	99
4.1. Introduction	99
4.2. Generation of phosphatase-dead strains	101
4.2.1. Identification of the catalytic residues in the active pocket of Cdc14	101
4.2.2. Generation of <i>cdc14C275S</i>	101
4.2.3. Generation of a regulatable Cdc14PD	101
4.3. Characterisation of Cdc14PD by Western blot	106
4.3.1. Phosphorylation status of Cdc14PD	106
4.3.2. Expression of Cdc14PD in yeast and hyphae	106
4.3.3. Co-IP of Cdc14 and Cdc14PD	108
4.4. Phenotypic analysis of Cdc14PD using microscopy	110
4.4.1. Morphology of PD mutants	110
4.4.2. Localisation of Cdc14PD	110
4.4.3. Localisation of Mlc1 in the presence of Cdc14 ^{PD}	112
4.4.4. IP of Cdc14 ^{PD}	112
4.5. Discussion	114
5. SILAC labelling in <i>Candida albicans</i>	117

5.1. Introduction	117
5.2. Media formulation and growth conditions used in SILAC	120
5.2.1. Lysine	120
5.2.2. Arginine	120
5.2.3. Other constituents of the media	120
5.2.4. Growth conditions in SILAC media	121
5.3. Labelling efficiency in yeast	122
5.4. Labelling efficiency in hyphae	126
5.5. Examination of Arg10 to Pro6 conversion	129
5.6. Discussion	132
6. A Screen for Cdc14PD interacting partners using quantitative SILAC-MS	135
6.1. Introduction	135
6.2. Cdc14PD interactors in yeast	136
6.2.1. SILAC experiments in yeast using QTOF-MS	136
6.2.2. SILAC experiments in yeast using QO-MS	144
6.3. Cdc14PD interactors in hyphae	153
6.4. Correction of H/L ratios of proline-containing peptides	166
6.5. Measuring protein levels in the cell lysate	171
6.6. Gene ontology analysis of potential Cdc14 interactors	173
6.7. Discussion	183
7. Discussion	185
7.1. Quantitative MS methods for studying kinase and phosphatase interactions in <i>C. albicans</i>	185
7.2. Strengths and limitations of using an overexpressed mutant version of Cdc14 in interaction studies	188
7.3. Future work	190
References	191

Figures

	Page
Figure 1.1: Morphology of <i>C. albicans</i>	3
Figure 1.2: Graphical view of protein coding genes in <i>C. albicans</i>	5
Figure 1.3: Morphology of Dbf2-depleted cells	10
Figure 1.4: Phenotype of Crk1 mutants	12
Figure 1.5: Morphological defects of <i>cdc14Δ/cdc14Δ</i> cells	18
Figure 1.6: Identifying protein interactions by co-IP-MS using SILAC	26
Figure 3.1: Tagging a protein with an epitope via PCR-based homologous recombination	64
Figure 3.2: Immunoprecipitation of Cdc14 and Dbf2	67
Figure 3.3: Elution of protein from beads following immunoprecipitation	69
Figure 3.4: Effect of protein concentration in cell lysate on IP output	70
Figure 3.5: Effect of total protein amount in cell lysate on IP output	72
Figure 3.6: Incubation of affinity beads with BSA prior to IP	73
Figure 3.7: Immunoprecipitation of Mob1 after one and three bead-washing steps	75
Figure 4.1: Alignment and CaCdc14 and ScCdc14 amino acid sequence using BLAST	102
Figure 4.2: Cloning steps for the generation of <i>cdc14^{PD}</i>	103
Figure 4.3: Colony screen for integration of <i>MET3</i> promoter in front of either <i>CDC14</i> or <i>cdc14^{PD}-MYC</i>	105
Figure 4.4: Phosphatase treatment of Cdc14 ^{PD}	107
Figure 4.5: Comparison of expression levels of Cdc14 and Cdc14 ^{PD}	107
Figure 4.6: Co-immunoprecipitation of Cdc14 ^{PD} and Cdc14	109
Figure 4.7: Phenotype of Cdc14 ^{PD}	111
Figure 4.8: Localisation of Mlc1-GFP in <i>cdc14^{PD}/CDC14</i> background	113
Figure 5.1: Metabolic pathway for conversion of arginine to proline in <i>C. albicans</i>	119
Figure 5.2: Determining labelling efficiency using an ion chromatogram	125
Figure 5.3: Incorporation of Lys8 in hyphae	127
Figure 5.4: Incorporation of Arg10 in hyphae	128
Figure 5.5: Effect of Pro6 on H/L ratio	130
Figure 5.6: Relative Lys8 abundance in <i>S. cerevisiae</i>	134
Figure 6.1: SILAC results from QTOF-MS experiments in yeast	139

Figure 6.2: SILAC results from QO-MS experiments in yeast	146
Figure 6.3: SILAC results from QTOF-MS experiments in hyphae	155
Figure 6.4: SILAC results from QO-MS experiments in yeast	159
Figure 6.5: Change in peptide and protein isotopic ratios after arginine-to-proline correction	167

Tables

	Page
Table 1.1: Experimentally characterised protein phosphatases in <i>C. albicans</i>	14
Table 3.1: General results from label-free MS	77
Table 3.2: Proteins identified by MS	95
Table 5.1: Incorporation efficiency of Arg10 and Lys8	123
Table 6.1: Summary of SILAC data from all experiments performed by QTOF-MS	137
Table 6.2: Summary of SILAC data from all experiments performed by QO-MS	137
Table 6.3: Proteins identified as hits in each data set (yeast QTOF-MS)	141
Table 6.4: Proteins identified as hits in each dataset from yeast QO-MS	149
Table 6.5: Non-quantified proteins enriched in heavy peptides (experiments from yeast QO-MS)	152
Table 6.6: Proteins identified as hits in each data set (hyphae QTOF-MS)	156
Table 6.7: Proteins identified as hits in each dataset from hyphae QO-MS	162
Table 6.8: Non-quantified proteins enriched in heavy peptides (experiments from hyphae QO-MS)	165
Table 6.9: Final list of Cdc14 hits composed after correcting for arginine-to-proline conversion	170
Table 6.10: Ambiguous proteins that were enriched in heavy isotopes in both the IP and the cell lysate	172
Table 6.11: GO analysis of Cdc14 hits based on cellular process	179
Table 6.12: GO analysis of Cdc14 hits based on protein function	180
Table 6.13: GO analysis of Cdc14 hits based on cellular components	182

Chapter 1

Introduction

1.1. Introduction to *Candida albicans*

Candida albicans is a commensal fungus that is normally found on the skin and mucosal surfaces of healthy people. However, a wide variety of factors can contribute to abnormal overgrowth of this fungus and lead to candidiasis, which is the most prevalent opportunistic yeast infection in humans (Martins et al., 2014). In more serious cases, *C. albicans* can disseminate to the bloodstream and internal organs of patients causing life-threatening systematic infections (Antinori et al., 2016). Between 50-90 % of cases are caused by *C. albicans*, while the rest are attributed to other species of the *Candida* genus, such as *C. tropicalis*, *C. glabrata*, *C. parapsilosis*, *C. stellatoidea*, *C. krusei* and *C. kyfer*. Most susceptible to infections are immunocompromised individuals, new-born babies and patients with implanted medical devices. Additionally, up to 75 % of women suffer from vulvovaginal candidiasis at least once in their life (Kabir et al., 2012).

The transition from commensal to pathogenic state is a multifactorial event. Changes in host environment, such as weakened immune response, suppressed microbiota due to antibiotic treatment, malnutrition or variations in pH can all play a part (Nobile and Johnson, 2015). In turn, *C. albicans* has evolved a range of adaptation strategies that allow it to survive in a changing environment and invade the host. Of significant importance is the ability of *C. albicans* to form biofilms on soft tissues and medical devices, which are resistant to the host immune system or conventional antifungal treatments. Expression of adhesins molecules enable the fungus to adhere to a wide variety of surfaces, while secreted hydrolytic enzymes degrade host surface molecules (Schaller et al., 2005). Strains with impaired production of adhesins or hydrolases are less virulent than wild type strains (Finkel et al., 2012). Another striking feature that contributes to *C. albicans* pathogenicity is its

ability to interchangeably switch between unicellular and filamentous growth in response to environmental cues. At <30 °C and acidic pH the fungus forms individual yeast cells that divide by budding and fully separate from each other at the end of mitosis (fig 1.1). Conditions commonly found inside the host, such as 37 °C, neutral pH and presence of blood serum and N-acetylglucosamine, induce the formation of hyphae – long filaments with parallel walls and multiple nuclei separated in compartments by a septum (Sudbery et al, 2004). Conditions between these two extremes may favour the production of pseudohyphae, which are wider than hyphae and have constrictions at the septation sites. The ability of *C. albicans* to grow as filaments is crucial during biofilm formation and host tissue penetration (Whiteway and Oberholzer, 2004). In particular, the yeast-to-hyphae transition has been implicated to play a role during infection, because mutant strains that cannot form hyphae are less virulent than wild type cells (Lo et al., 1997). However, many of the molecular pathways that control morphogenesis also control the expression of various virulence factors, and so the effect may be pleiotropic (Trevijano-Contador et al., 2016). Interestingly, some hyperfilamentous strains have shown decreased infection rate, while some non-filamentous strains have retained their virulence (Alonso-Monge et al., 1999, Noble et al., 2010). An additional complication comes from the fact that many studies have used *URA3* as a selectable marker to create their strains, but adequate expression of this gene is required for virulence and for morphological transition (Lay et al., 1998). The expression of *URA3*, however, is dependent on its chromosome position and thus may vary in mutants. Altogether, the relationship between morphogenesis and virulence in *C. albicans* is very complex and requires further study to be fully understood.

Our understanding of the complex nature of *C. albicans* has greatly advanced in the past two decades. This fungus has been employed as a model organism for studying human pathogens and polarised cell growth. Many molecular biology techniques have been adapted from the *Saccharomyces cerevisiae* research field, since both fungi are genetically related. However, manipulation of the *C. albicans* genome is more difficult due to its obligate diploid nature, genomic plasticity (possible aneuploidy) and the lack of clearly defined sexual cycle (Noble and Johnson, 2007). *Candida* species also have a codon bias for the CTG codon, which they translate as serine rather than leucine (Santos and Tuite, 1995). Because of this bias, all exogenous genes have to be modified before they can be introduced

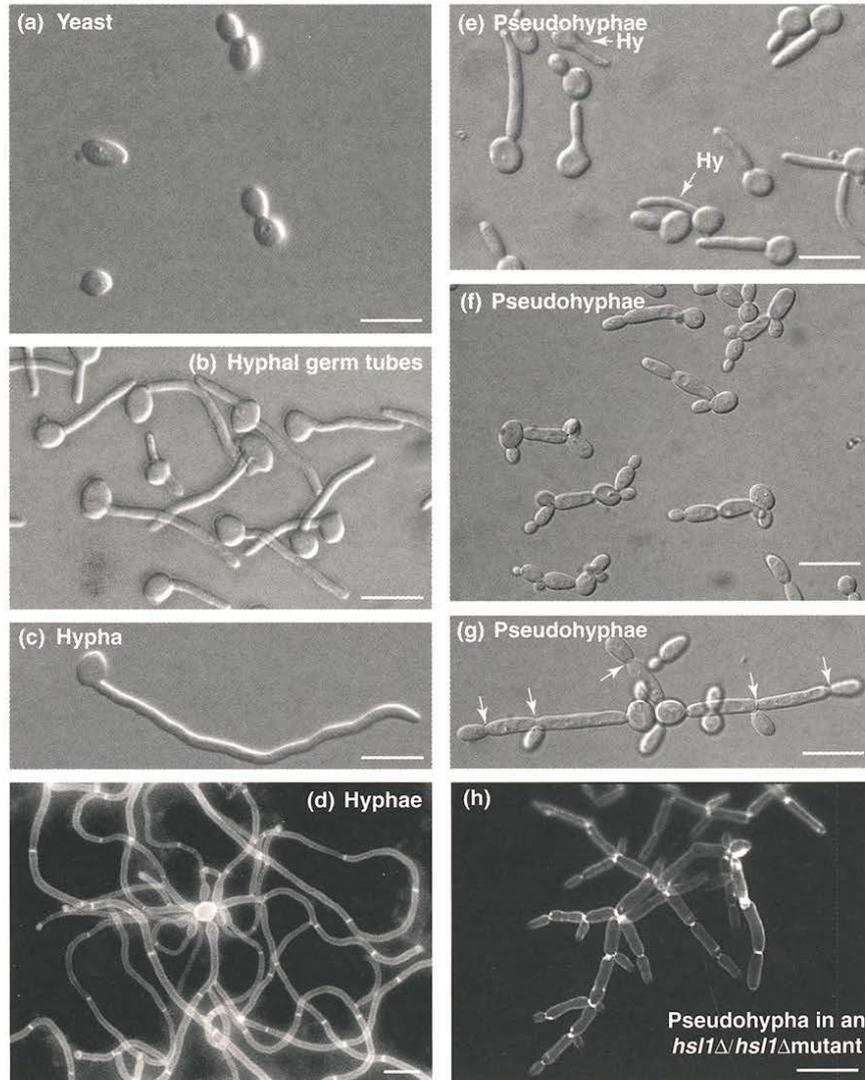


Fig. 1.1: Morphology of *C. albicans*. Cells can grow as either yeast, pseudohyphae or “true” hyphae depending on their environment. This figure is taken from Sudbery *et al.*, 2004.

in *C. albicans*. The complete genome sequence of the *C. albicans* strain SC5314 is available on the Candida Genome Database (candidagenome.org) since 2004 and it has made genetic manipulation of this fungus much easier than before. It has 6218 protein coding genes, only a quarter of which have been experimentally verified (fig. 1.2). This means that much more remains to be discovered about the biology of *C. albicans*.

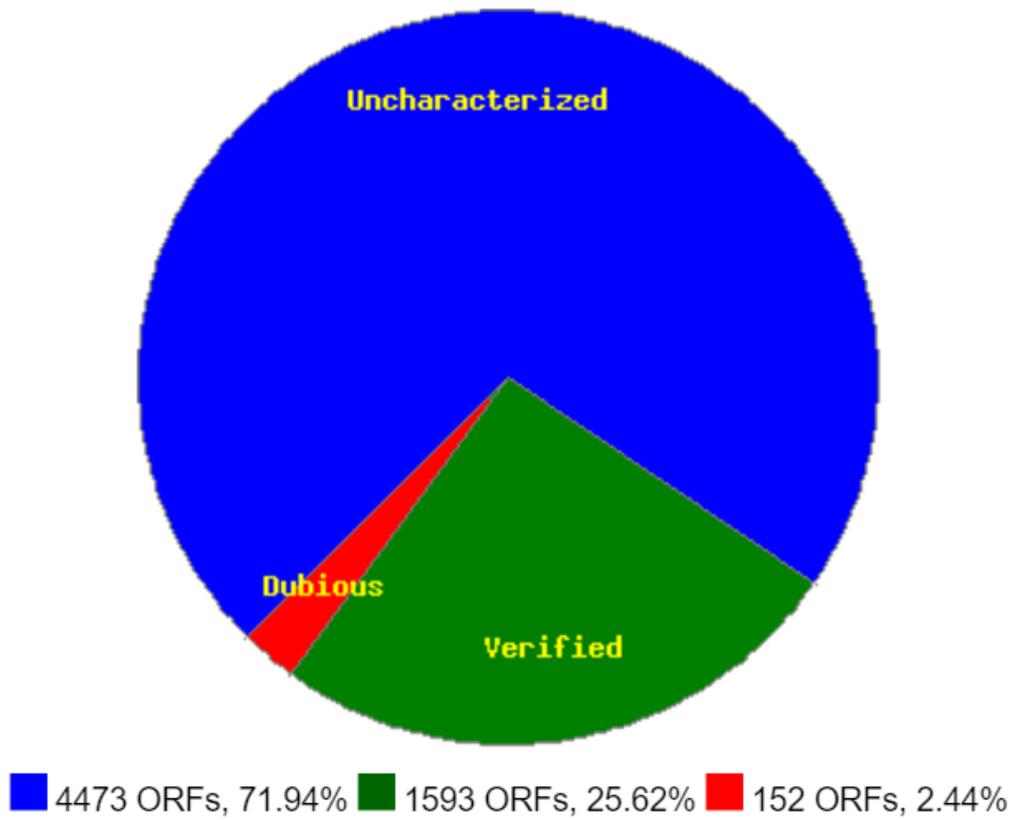


Fig. 1.2: Graphical view of protein coding genes in *C. albicans* (as of 26/06/2016). This figure has been adopted from the Candida Genome Database.

1.2. The role of protein phosphorylation and dephosphorylation in the cell

Posttranslational modifications, such as protein phosphorylation and dephosphorylation, have evolved as mechanisms to control proteins function through reversible alteration of their folding, stability and activity. Protein phosphorylation is done by kinases, which transfer a phosphoryl group from ATP (or more rarely from GTP) to another protein that is their substrate. The addition of a phosphate can change the activity of a protein in two ways. First, it could induce a conformational change that prepares the protein for a downstream action. Second, it can disrupt a protein's surfaces in a way that either creates or blocks another protein binding site (Cheng et al., 2011). Phosphatases reverse this action by catalysing the transfer of a phosphate from a protein to a water molecule, a process known as dephosphorylation.

Kinases and phosphatases (KPs) modulate a huge number of molecular pathways implemented in every basic cellular process in the cell. A single enzyme can have up to several hundred substrates and interaction with each of them is tightly controlled in space and time. KPs have exquisite specificity for their substrates, achieved through recognition of selected amino acid sequence surrounding the phosphoacceptor site (Hutti et al., 2004). In addition, the structure of the catalytic site, the formation of complexes with regulatory subunits, interaction with docking sites on the substrate, localisation of both enzyme and substrate, competition of substrates at any given time and various error correction mechanisms all affect the specificity of an enzyme (Ubersax and Ferrell Jr, 2007).

Since KPs have antagonistic functions, their action must be balanced at any time. The main mechanisms for achieving this are compartmentalisation of the enzymes and modulation of their activity (Bononi et al., 2011). The spatial organisation of KPs creates a gradient of phosphorylated substrates across the cell units. Additionally, KPs often regulate each other by positive and negative feedback loops (Kamioka et al., 2010). Disruption of this complex interplay has been implicated in variety of human disease and continues to be extensively studied. Given their importance, it is not surprising that kinases have become one of the most researched classes of drug targets (Zhang et al., 2009).

1.3. The role of protein kinases in *C. albicans* morphogenesis

1.3.1. Cdc28

Several kinases have been identified to be important for hyphal formation and development in *C. albicans*, although most of their direct targets remain to be found. The cyclin-dependent kinase Cdc28 (also known as Cdk1) is one of the key regulators of cell cycle progression and morphogenesis by controlling several distinct pathways. In association with two G₁ cyclins Ccn1-Cdc28 and Cln3-Cdc28 initiate cell budding through polarised growth, which is maintained during hyphal development by a complex with a third G₁ cyclin Hgc1-Cdc28 (Zheng et al., 2004; Wang, 2016). One of the known targets of Ccn1-Cdc28 is the septin Cdc11, phosphorylated upon hyphal induction first by the kinase Gin4 and then by Ccn1-Cdc28 (Sinha et al., 2007). This phosphorylation is then sustained in hyphae by Hgc1-Cdc28. Abolishing these phosphorylation events impairs hyphal development after initial tube evagination. Phosphorylation of another septin, Sep7, is also dependent on Hgc1, although the involvement of Cdc28 has not been shown (Gonzalez-Novo et al., 2008). However, Sep7 is also phosphorylated by Gin4, which is activated by Clb2-Cdc28 phosphorylation (Li et al., 2012).

Hgc1-Cdc28 further support polarised growth by phosphorylation of the GTPase-activating protein (GAP) Rga2 (Zheng et al., 2007). This event ensures that Rga2 will not go to the hyphal tip and inactivate the GTPase Cdc42, which is required for hyphal development (Court and Sudbery, 2007).

Sec2 and Exo84, which are both involved in the transport of secretory vesicles to the tip, are also substrates of Hgc1-Cdc28 (Bishop et al., 2010; Caballero-Lima and Sudbery, 2014). Exo84 is part of the exocyst, a multiprotein complex at the hyphal tip that tethers secretory vesicles before they fuse with the membrane. Sec2 is a guanosine exchange factor (GEF) for the Rab GTPase Sec4 involved in the docking of secretory vesicles to the exocyst (Guo et al., 1999). The polarisome member Spa2 is yet another protein at the hyphal tip phosphorylated by Cdc28 (Wang et al., 2016). Clb2-Cdc28 targets Spa2 in both yeast and hyphae, while Hgc1-Cdc28 phosphorylate the protein only in hyphae. Abolishing CDK sites in

Spa2 leads to translocation of the protein from the tip to the septum and disrupts hyphal morphology.

Finally, Cdc28 is known to phosphorylate two morphology-related transcription factors. Hgc1-Cdc28 regulates Efg1, which blocks the expression of genes involved in septum degradation after cytokinesis (Wang et al., 2009). Ccn1/Cln3-Cdc28 together with the kinase Cbk1-Mob2 phosphorylate Fkh2 which promotes the expression of genes supporting hyphal development (Greig et al., 2015). Given the wide range of substrates that Cdc28 phosphorylates, it is not surprising that blocking the kinase activity with the ATP analogue 1NM-PP1 completely disrupted the formation of true hyphae (Sinha et al., 2007). It is suspected that many more interactors of Cdc28 remain unknown.

1.3.2. Cbk1

The nuclear Dbf2-related (NDR) kinases have a highly conserved role in regulation of cell cycle and polarised growth. *C. albicans* has two NDR kinases, Cbk1 and Dbf2, which are activated by their regulatory subunits, Mob2 and Mob1 respectively. Cdc28 controls Cbk1 activity through phosphorylation of Mob2 shortly after hyphal induction, which is required for sustaining polarised growth (Gutiérrez-Escribano et al., 2011). Cbk1 and Mob2 are members of the regulation of Ace2p activity and cellular morphogenesis (RAM) network that maintains cell polarity during yeast budding and hyphal development. Both proteins are essential for hyphal growth, as deletion mutants remained permanently locked in the yeast form (Song et al., 2008). The importance of Cbk1 for polarised growth has prompted several studies to investigate its downstream targets. Bharucha et al. (2011) did a haploinsufficiency-based genetic interaction screen for targets of Cbk1 using *CBK1/cbk1Δ* strain. The screen specifically focused on strains displaying morphological defects on Spider medium and identified 41 genes that show synthetic interaction with *CBK1*, half of which are under the transcriptional control of Ace2. A third of those genes were also controlled by the cAMP-protein kinase A (PKA) pathway, suggesting that the interplay between the RAM and cAMP-PKA pathways largely determines cell morphology. Recently, Saputo et al. (2016) did a similar screen but solely looking for strains with decreased filamentation on serum medium. They found 36 genetic interactions with *CBK1*, indicating distinct set of substrates

during yeast and hyphal growth. The study also identified Rgd3 a second GAP for Cdc42 and showed that Cbk1 phosphorylation is required for localisation of Rgd3 to sites of polarised growth. Other targets of Cbk1 include the transcription factors Bcr1, which play a role in biofilm formation, and the mRNA-binding protein Ssd1 (Gutiérrez-Escribano et al., 2011). On hyphal induction, Cbk1-dependent phosphorylation of Ssd1 downregulates the levels of the transcription factor Nrg1, which represses the expression of several hyphae-specific genes (HSGs) (Lee et al., 2015).

1.3.3. Dbf2

Dbf2-Mob1 are part of the mitotic exit network (MEN, also including Lte1, Tem1, Cdc5, Cdc15, Bfa1, Bub2 and Cdc14) in *S. cerevisiae*, which regulates late mitotic events and M-G₁ transition. Although the presence of this network in *C. albicans* has not been established, one study looked specifically at the role of Dbf2 in this fungus and found a great degree of similarity between both orthologues. González-Novo et al. (2009) showed that while *DBF2* is an essential gene in *C. albicans*, conditional mutants, where the gene is downregulated, display severe defects in cell separation due to failure to form the primary septum and contract the actomyosin ring during mitosis (fig. 1.3). In addition, dividing cells were unable to form the mitotic spindle correctly and divide their DNA content equally between the mother and daughter cells. The authors also demonstrated that during mitosis Dbf2 sequentially moves from the nucleus to the mitotic spindle, the spindle pole bodies (SPB) and to the bud neck at the end of the cycle. Co-immunoprecipitation experiments proved a direct physical interaction with tubulin, which is still the only known CaDbf2-interacting protein to date. Finally, the study revealed that the kinase is also important for hyphal morphogenesis. Although the cellular localisation of Dbf2 in hyphae was not investigated, depletion mutants formed very swollen tubes with no septum between the nuclei and with constrictions reminiscing dividing yeast cells. It is unclear whether the morphological defects of hyphae are solely due to disruptions in cell cycle progression or the kinase has a separate role in polarised growth.

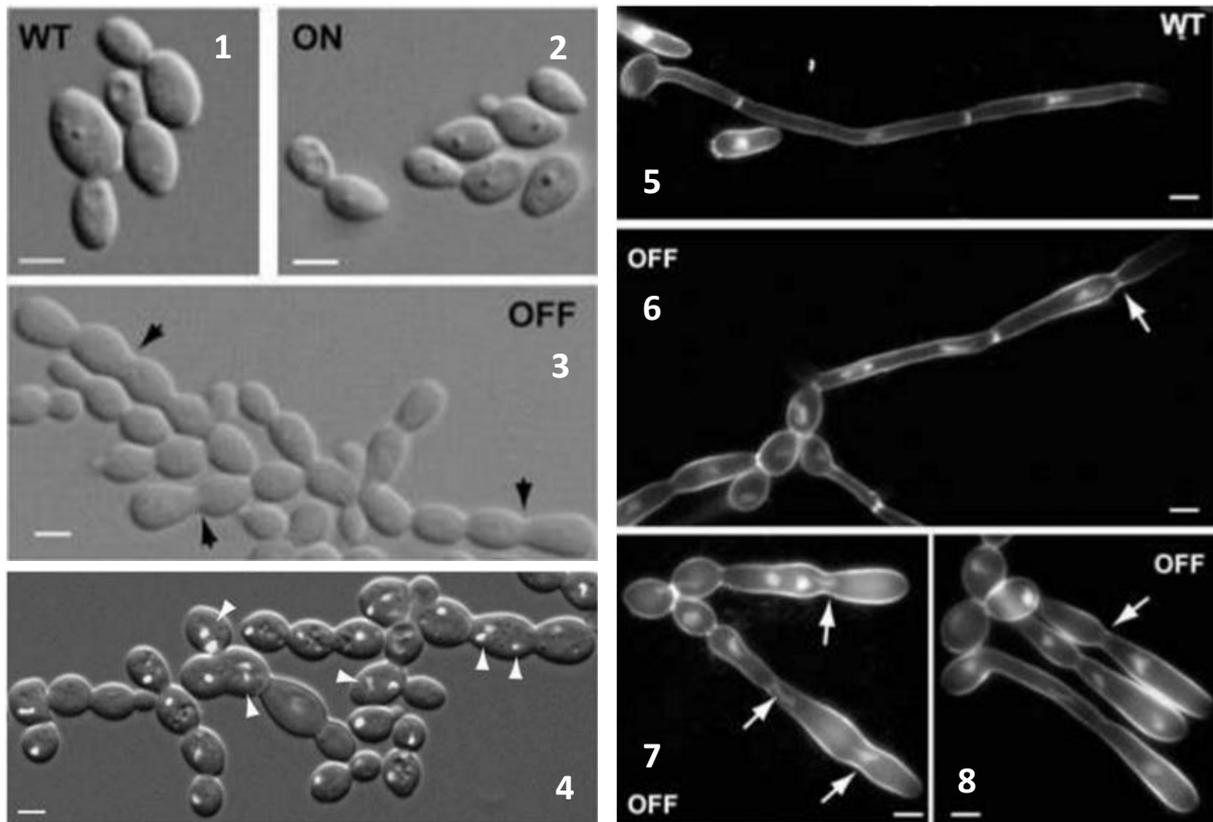


Fig. 1.3: Morphology of Dbf2-depleted cells. Mutants have a single Dbf2 allele that is under the control of MET3 promoter. When the promoter is induced (2), yeast cells grow as normal wild type cells (1). In MET3 repressing conditions, yeast fail to complete cytokinesis and form chains of connected cells (3-4). Black arrows in image 3 indicate abnormally wide bud necks. The white arrows in image 4 point towards cells with more than one nuclei. Images 1-3 are taken by DIC microscopy, and image 4 is a merged picture of DIC and DAPI channels. Images 5-8 show hyphae stained with DAPI and Calcofluor White. Comparing to wild type hyphae (5), depleted mutants (6-8) form wider tubes with constrictions but no septum (indicated by white arrows). All bars are 5 μ m. All images were taken from González-Novo et al., 2009.

1.3.4. Crk1

Unlike Dbf2, the Cdc2-related kinase 1, Crk1 is not essential for viability and construction of the deletion strain *crk1Δ/crk1Δ* has revealed a key role in *C. albicans* filamentation and virulence. In the absence of Crk1 cells are swollen and permanently locked in the yeast form unable to complete cytokinesis (fig. 1.4) (Chen et al., 2000). On the other hand, ectopic expression of Crk1 or its catalytic domain, Crk1N, induces polarised growth on solid YPD (but not in liquid YPD) at 30 °C. The authors have shown that *crk1Δ/crk1Δ* mutants fail to express the HSGs *ECE1* and *HWP1*, which are abnormally expressed in yeast when Crk1 or Crk1N are ectopically introduced. Introduction of Crk1N in the non-filamentous *efg1Δ/efg1Δ*, *cph1Δ/cph1Δ* and the double deletion mutants induces polarised growth, indicating that Crk1 acts independently from the transcription factors Efg1 and Cph1. The study concludes through a series of gene deletion experiments that Crk1 may act in the same pathway with Ras1/cAMP to promote hyphal development. Using a yeast-two hybrid system, Ni et al. (2004) identified Cdc37 and Sti1 as interacting partners of Crk1, which likely assist with protein folding of the kinase. Despite the prominent role of Crk1 in polarised growth its substrates remain unknown. In contrast to the other kinases discussed so far, it is unlikely that Crk1 phosphorylates proteins involved in cytoskeleton organisation (Dhillon et al., 2003). The closest homologues of Crk1 are the *S. cerevisiae* Bur1 (also known as Sgv1) and the human Cdk9, both of which are CDKs controlled by a single cyclin Bur2 and CycT respectively (Malumbres, 2014). Bur1 and Cdk9 are transcriptional regulators of several genes, and it likely that Crk1 has a similar role in *C. albicans*. Although a Bur2 homologue with a cyclin domain exists in the Candida Database, the protein (also called Bur2) is still uncharacterised and no association with Crk1 has been shown (Yao et al., 2000).

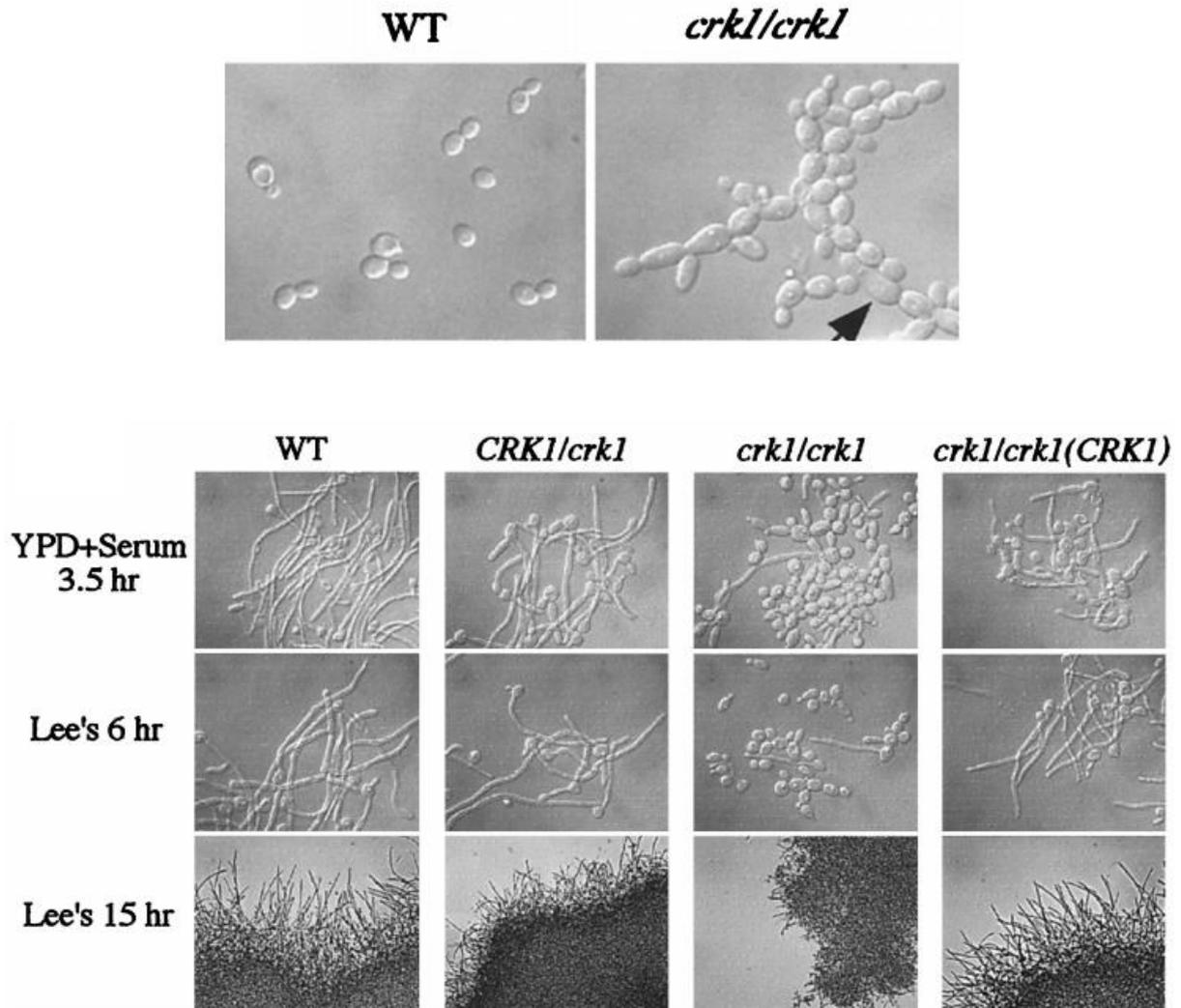


Fig. 1.4: Phenotype of Crk1 mutants. Deletion of both copies of *CRK1* produced chain of swollen cells under yeast conditions (top). Hyphal formation was also impaired in Lee's medium, but a single allele of *CRK1* is sufficient to rescue this phenotype (bottom). All images were taken from Chen et al., 2000.

1.4. Protein phosphatases in *C. albicans*

1.4.1. General overview

Protein kinases greatly outnumber protein phosphatases in all studied organisms. In budding yeast, there are 124 protein kinases and 37 protein phosphatases. The exact number in *C. albicans* has not been determined yet, but one comparative study of fungal kinomes found about 20% less kinases in this fungus compared to *S. cerevisiae*, and another one reported 28 putative phosphatase genes (Kosti et al., 2010; Hanaoka et al., 2008). Analysis of the *C. albicans* phosphoproteome in hyphae has identified 15,906 unique phosphosites on 2,896 proteins (Willger et al., 2015). Overall, it seems that a very small group of phosphatases is responsible for the dephosphorylation of thousands of proteins, many of which have multiple phosphosites. Studying the phosphatase-substrate network in *C. albicans* is therefore a huge task, which is still at the very beginning. At present, only 17 genes have been experimentally verified to express a protein phosphatase (summarised in table 1.1), 9 of which have been implicated in hyphal development. Very few of the substrates of these phosphatases are known. The MAPK phosphatase Cpp1 suppresses hyphal formation by dephosphorylation of the kinase Cek1 and by repressing HSGs expression (Schroppe et al., 2000). Tpd3-Pph21 controls septin ring disassembly by dephosphorylation of Sep7 and *tpdΔ/tpdΔ* mutants grow constitutively as pseudohyphae with multiple septin rings (Liu et al., 2016). Ppg1 promotes filament extension through downregulation of Nrg1 and induction of several HSGs (Albataineh et al., 2014). Since Cdc14 is the main target of this study, its role is reviewed in more detail below.

1.4.2. Cdc14

Cdc14 is a dual specificity protein phosphatase, which means that it can dephosphorylate both phosphoserine/phosphothreonine and phosphotyrosine residues. The Cdc14 family is highly conserved and it is found in all eukaryotes except plants. Cdc14 orthologues are among the most extensively studied phosphatases owing to the central role of the founding member ScCdc14 in controlling mitotic exit in budding yeast (Mocciaro and Schiebel, 2010).

Phosphatase	Description	Reference
Cmp1	Calcineurin, a Ca ²⁺ -calmodulin-activated, serine/threonine-specific protein phosphatase, essential for virulence	Bader et al., 2003
Cpp1	VH1 family dual specificity MAPK phosphatase that represses yeast-hyphal transition and is required for virulence	Csank et al., 1997
Cdc14	Dual specificity phosphatase involved in exit from mitosis and morphogenesis	Clemente-Blanco et al., 2006
Glc7	PP1 serine/threonine phosphatase involved in cell morphogenesis, cell cycle progression and DNA damage response	Hu et al., 2012
Pph21	Type PP2A phosphatase that dephosphorylates the septin Sep7 and regulates morphogenesis and cytokinesis	Liu et al., 2016
Ptc1	Type PP2C phosphatase involved in hyphal growth and virulence	Hanaoka et al., 2008
Ptc2	PP2C family member involved in regulation of mitochondrial physiology and DNA damage checkpoints	Feng et al., 2010
Ptc4	Type PP2C serine/threonine phosphatase involved in ion homeostasis	Zhao et al., 2010
Ptc5 Ptc6 Ptc7	Mitochondrial protein phosphatases of the PP2C family involved in antifungal drug sensitivity	Zhao et al., 2012
Ptc8	Type PP2C serine/threonine phosphatase required for hyphal growth	Fan et al., 2009
Ppz1	Protein phosphatase Z, serine/threonine specific protein phosphatase involved in cation homeostasis, cell wall integrity and virulence	Adam et al., 2012
Ppg1	A PP2A-type protein phosphatase controls filament extension and virulence	Albataineh et al., 2014
Psy2-Pph3	Phosphatase with a role in filamentous growth induced by genotoxic stress and recovery from the DNA damage checkpoint	Sun et al., 2011
Sit4	PP2A phosphatase with a role in cell wall maintenance, hyphal growth, and virulence	Lee et al., 2004
Yvh1	Dual specificity phosphatase that controls growth, cell cycle progression and virulence	Hanaoka et al., 2005

Table 1.1: Experimentally characterised protein phosphatases in *C. albicans* (as of 1/08/2016).

Due to the high attention that these phosphatases have received, they are now hold responsible for targeting hundreds of substrates in a variety of cellular processes.

In *S. cerevisiae*, the function of Cdc14 is largely controlled by its subcellular localisation. From G1 to metaphase, when CDK activity is high, the phosphatase is sequestered in the nucleolus by its inhibitor Net1 (also known as Cfi1) as part of the regulator of nucleolar silencing and telophase (RENT) complex (Visintin, et al., 1999). A recent study suggests that in S phase Clb5-Cdc28 inhibits the phosphatase activity of Cdc14 by phosphorylating it at S429 (Li et al., 2014). During anaphase the FEAR (fourteen early anaphase release) and MEN pathways ensure sequential release of Cdc14 first to the nucleoplasm and later to the cytoplasm (Faust et al., 2013; Yellman and Roeder, 2015). At the start of anaphase Cdc28 activates the anaphase promoting complex/cyclosome (APC/C) that together with Cdc20 degrades securin and activates separase, an enzyme that drives separation of sister chromatids (Rudner and Murray, 2000). Separase also promotes Net1 phosphorylation by Cdc28 and Cdc5, which leads to the initial release of Cdc14 to the nucleoplasm (Sullivan and Uhlmann, 2002; Queralt et al., 2006). At this stage, Cdc14 promotes mitotic spindle elongation and ribosomal DNA segregation (Higuchi and Uhlmann, 2005). Correct spindle orientation is a prerequisite for Tem1-dependent activation of the kinases Cdc15 and Dbf2-Mob1, both of which further phosphorylate and thus inhibit Net1 (Visintin and Amon, 2001). Dbf2-Mob1 also directly phosphorylates Cdc14, an event that drives export of the phosphatase from the nucleus to the cytoplasm (Mohl et al., 2009). At the end of mitosis Cdc14 returns to the nucleolus until the next division.

The main role of Cdc14 in budding yeast is to orchestrate late mitotic events by transiently inhibiting CDK activity and reversing CDK-dependent events. Cdc14 activates Sic1, which inhibits CDKs by direct association with them (Visintin et al., 1998). Cdc14 further stimulates Sic1 expression by dephosphorylating its transcription factor, Swi5, thus enabling it to enter the nucleus. Cdc14 also induces degradation of mitotic cyclins by dephosphorylating Cdh1. APC/C-Cdh1 targets cyclins for ubiquitin-dependent proteolysis (Visintin et al., 1997). Finally, analysis of known Cdc14 substrates has revealed that the phosphatase has a preference CDK consensus sites with one study suggesting a strong bias towards phosphoserine over phosphothreonine CDK sites (Gray et al., 2003; Bremmer et al., 2012; Sanchez-Diaz et al., 2012). In addition to CDK inactivation, Cdc14 has a recognised role

in cytoskeleton organisation, septum formation and actomyosin ring contraction consistent with its localisation at the bud neck at the end of mitosis (Bloom et al., 2011).

The essential role of Cdc14 in mitotic exit is not conserved in all species. The *S. pombe* orthologue Clp1 (also known as Flp1) actively participates in septum formation, nuclear division, chromosome segregation and cytokinesis but is not required for cell viability (Chen et al., 2013). Clp1 regulates G₂-M transition and overexpression blocks the cells in G₂, which is in contrast to budding yeast Cdc14 that arrests the cells in G₁ (Visintin et al., 1998; Cueille et al., 2001; Trautmann et al., 2001). In the plant pathogen *Fusarium graminearum* Cdc14 is important for morphogenesis, pathogenesis and cytokinesis (Li et al., 2015). Similarly, in the fungal entomopathogen *Beauveria bassiana* the phosphatase controls conidiation, virulence and stress response (Wang et al., 2013). In mammals several Cdc14 homologues regulate DNA damage repair but are dispensable for mitotic exit (Mocciaro and Schiebel, 2010; Lin et al., 2015). Several high-throughput studies of Cdc14 interactors in fungi have identified proteins involved in DNA repair, but the significance of these findings has not been investigated in detail (Bloom et al., 2011; Breitskreutz et al., 2010; Chen et al., 2013).

The *C. albicans* orthologue displays a cell cycle-controlled localisation pattern that is different from that of ScCdc14 (Clemente-Blanco et al., 2006). The phosphatase is completely absent from G₁ cells and gradually start accumulating in the nucleus and nucleolus from S phase onwards. At the start of mitosis, Cdc14 concentrates at the SPB and later moves to the bud neck during cytokinesis, after which it is degraded. While the protein is not essential for vegetative growth or mitotic progression, deletion mutants fail to separate at the end of the cycle due to incomplete septum degradation. Cdc14-dependent dephosphorylation of septin regulator Nap1 at the end of mitosis is required for translocation of the protein from the septum to the cytoplasm (Huang et al., 2014). Additionally, *cdc14Δ/cdc14Δ* cells did not accumulate the master regulator of cell separation Ace2 in daughter nuclei and showed decreased expression of genes controlled by it. Cdc14 has probably retained its function to counteract CDK activity in *C. albicans*, since it is involved in degradation of the cyclins Clb2 and Clb4 during mitosis, but it is not clear whether Cdc14 inhibits CDK activity in any other way. However, in a recent study Yong et al. (2016) have proposed a model whereby the kinase Gin4 regulates septin ring assembly at

the beginning of mitosis and is dephosphorylated by Cdc14 at the end of the cycle to allow disassembly of the ring. As mentioned earlier, Gin4 is also a substrate of Cdc28, which raises the possibility that Cdc14 might also target CDK sites in *C. albicans*.

In hyphae, Cdc14 localises only to the nucleus of the apical compartment, but not to the septum. This pattern is dependent on Hgc1 and Sep7 and if either of them is deleted Cdc14 goes to the septum of germ tubes and induces cell separation (Gonzalez-Novo et al., 2008). Deletion of *CDC14* impairs invasive and filamentous growth since in the presence of serum, cells form much shorter tubes than wild type hyphae. While in the absence of Cdc14 yeast cells are able to progress through the cell cycle in time, hyphae exhibit delay in G2-M transition, suggesting that mitosis may be regulated differently in these forms.

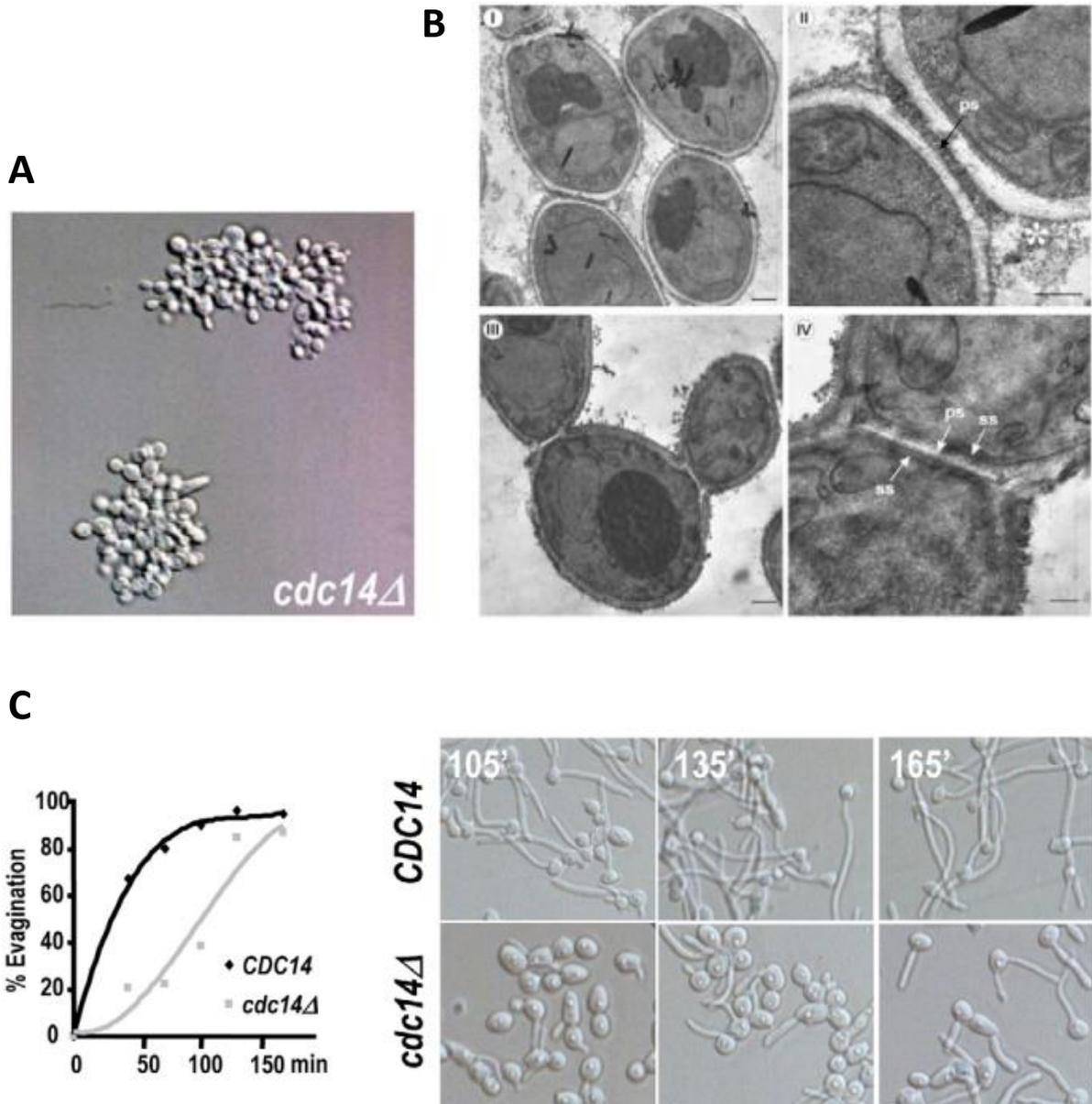


Fig. 1.5: Morphological defects of *cdc14Δ/cdc14Δ* cells. (A) Yeast cells lacking Cdc14 form clumps due to incomplete separation after each cell cycle. (B) Comparing to wild type cells (I-II), deletion mutants (III-IV) failed to degrade the septum at the end of mitosis. (C) Hyphae also exhibit severe delay in evagination and tube elongation in the presence of serum. All images are taken from Clemente-Blanco et al., 2006.

1.5. Methods of identifying novel kinase and phosphatase interactions

1.5.1. Methods currently used in *C. albicans*

So far, studies of *C. albicans* KPs have mostly focused on investigating individual interactions. Putative interacting partners are usually found by either: 1) analogy with KPs in other organisms; or 2) co-localisation of fluorescently-tagged enzyme and another protein; or 3) implication of an enzyme and a protein in a common pathway (e. g. if deletion mutants have similar phenotype). Suspected interactions are then confirmed by co-immunoprecipitation (co-IP)/affinity purification (AP), KP assays, changes in electrophoretic mobility of a substrate in the absence of enzyme activity, co-localisation experiments or mutation of putative phosphosites. Although these methods are very useful for understanding the role of KPs, a more comprehensive analysis of KP interactions (KPI) is required for better interpretation of their function. The complex haploinsufficiency-based screen for genetic interactors of Cbk1 is the first (and so far the only) large-scale analysis of a kinase substrates in *C. albicans* (Bharucha et al., 2011; Saputo et al., 2016). These studies presented a bigger picture of how a single kinase regulate different pathways in response to varying conditions. Although *C. albicans*-adapted yeast two hybrid technology exist, it has not been used for KPI screens (Stynen et al., 2010).

1.5.2. High-throughput techniques used in other species

In other organisms, proteome-wide screens for KPI have been carried out for over a decade. The yeast two-hybrid system is the most applied method for studying physical protein interactions but it is becoming less popular with the advancements of more recent technologies (Bruckner et al., 2009). High-throughput studies commonly involve affinity-based purification of a bait (e. g. a kinase, a phosphatase or a regulatory protein) coupled to mass spectrometry (MS) analysis for identification of co-eluted prey proteins (Gavin et al., 2006). In the simplest scenario, the bait and prey proteins are all expressed in the same

system and allowed to interact in their natural environment prior to purification (Gavin et al., 2002). More commonly however, capturing KPI requires an intervention from the researchers that may disrupt the physiological environment of an enzyme. For example, baits often have to be overexpressed with the use of exogenous promoters in order to provide enough material for an MS analysis (Breitkreutz et al., 2010). Although this might create some false positive results, it often allows the capture of many true interactions that would not be detected otherwise. Alternatively, a bait of interest may be expressed in large amount and purified from a different organism, such as bacteria, before being incubated with cell lysates to allow binding to interacting partners (Knebel et al., 2001). A purified bait can also be incubated with a phage display library or a protein/peptide array chip. In the first method, proteins encoded by cDNA are displayed on the surface of a phage and interaction is detected by immunological assays (Zhou et al., 2003). On the other hand, incubation of a kinase with a microarray chip array is followed by detection of substrate phosphorylation with the use of a phosphorimager (Fasolo et al., 2011). A major caveat of this technique is that immobilised proteins do not always fold correctly, which may prevent interaction with the enzyme. Several studies have combined *in vitro* kinase assays with MS for identification of kinase targets (Li et al., 2014; Muller et al., 2016). A purified kinase is incubated with cell extracts and ATP and phosphorylated proteins are detected by MS. Background phosphorylation can be minimised by several strategies, for example, by using an analogue-sensitive (AS) enzyme as a bait that are designed to accommodate an ATP analogue that other kinases cannot use (Xue and Tao et al., 2013). ATP analogues transfer a non-conventional phosphoryl group that labels the substrates and can be identified by MS. The use of AS kinases is one of the most reliable methods for studying kinase substrates and can also be applied *in vivo* (Shah et al., 1997). However, not every kinase can be engineered to be AS (Koch and Hauf, 2010).

Phosphatases can also be modified for improved detection of their targets. Substrate-trapping enzymes, constructed by a single amino acid substitution, lose their catalytic activity but bind their substrates with higher affinity (Blanchetot et al., 2005). Normally, enzyme-substrate complexes are very short-lived but have to remain intact for the course of purification experiments. Mutants are more reliable in this regard are therefore often used in interaction studies (Bloom et al., 2011).

An alternative method for studying KPs is to look at phosphoproteome dynamics rather than physical associations between proteins (Ficarro et al., 2002; Ptacek et al., 2005). Proteolytically digested whole cell extracts contain a mixture of phosphorylated and non-phosphorylated peptides. The former can be isolated by affinity chromatography and analysed by MS to create a map of all phosphorylated proteins in the cell, i.e. the phosphoproteome. Comparing the phosphoproteome of wild type cells to those of knockout mutants reveals downstream pathways controlled by the missing enzyme as well as many of its direct phosphosites (Kao et al., 2014).

Finally, it is worth noting that several software platforms have been developed to predict phosphorylation sites by scanning protein sequences for a known target motif (Xue et al., 2005; Hornbeck et al., 2015). This can be a good starting point for identifying new KPI and may complement an experimental approach.

1.6. Introduction to mass spectrometry

Mass spectrometry is a century old technique for qualitative and quantitative analysis of compounds and molecules. The main principle of MS is separation of charged species from a sample based on their mass (m) to charge (z) ratio. The resulting spectrum of masses is used to identify the composition of the original sample. MS has many applications in the fields of chemistry, physics, biology and others but for the purpose of this study, the discussion here is limited to the context of proteomics.

1.6.1. Sample preparation

A typical MS-based experiment starts with isolation of a protein mixture by cell extraction, sometimes followed by enrichment of a desired protein by various methods of purification. Complex samples may be separated by gel electrophoreses in one or two dimensions. Whole proteins can be analysed by MS, but more commonly, they are digested to shorter peptides with enzymes, such as trypsin. Peptides may be further fractionated by chromatography or an enrichment method may be used to select for peptides with desired characteristics, e.g. phosphorylation.

1.6.2. Principles of MS instruments

A mass spectrometer generally consists of an ion source, a mass analyser and a detector. Protein and peptides are non-volatile and thermally unstable, so ionisation is necessary to prevent degradation in the gas phase. Most commonly, analytes in solution are ionised by electrospray ionisation (ESI), while dry samples are ionised by matrix-assisted laser desorption ionisation (MALDI) (Fenn et al., 1989; Karas and Hillenkamp, 1988). ESI instruments are usually coupled to a liquid chromatography (LC) that separates the molecules prior to ionisation (Pitt, 2009). There are several other types of ion sources used in exceptional cases, but most proteomics studies use ESI-LC-MS instruments.

The resulting ions then enter the mass analyser, where they are separated based on their m/z ratio. The main types of mass analysers are:

- Time-of-flight (TOF) – ions are accelerated through an electric field and their m/z is deduced from the time it takes them to reach the detector (Guilhaus, 1995). TOF analysers are most often coupled to MALDI ionizers.
- Quadrupole – ions travel in spiral trajectories between four parallel metal rods in an electric field created by static direct current and alternating radio frequency current voltage (Finnigan, 1994). Some quadrupoles also include a magnetic field. At any given time, only ions of certain m/z reach the detector depending on the applied voltage.
- Ion trap – this analyser works on the same principles as quadrupoles, except that, as the name suggests, ions are trapped in confined spaced and sequentially ejected (Hager, 2002). Ions can also be fragmented inside the trap, which generates a tandem mass spectrum. Variations of this technology include linear ion trap, quadrupole ion trap and orbitrap.
- Fourier transform ion cyclotron resonance (FTICR-MS) – in this analyser electric and magnetic fields accelerate the ions of particular m/z around a cyclotron (Marshall et al., 1998). Ion frequency and intensity is determined by Fourier transform mathematical operation, which is used to calculate the corresponding m/z .

Some MS instruments may combine two or more mass analysers to achieve higher throughput and resolution. Common combinations include quadrupole-TOF, quadrupole-ion trap, linear ion trap-orbitrap, triple quadrupole, TOF-TOF and others (Medzihradszky et al., 2000; Chernushevich et al., 2001). All mass analysers have certain advantages and disadvantages. Orbitrap and FTICR-MS instruments offer the highest mass accuracy and resolution (Scigelova et al., 2011).

After passing through the mass analyser, ions hit a detector that passes the signal to a recording system. The most common type of detectors is the electron multiplier (Neetu et al., 2012). Some instruments, such as FTICR-MS and orbitrap, have detector plates within the mass analyser.

1.6.3. Processing of MS data

The output of an MS experiment is a mass spectrum of ions m/z plotted against their intensity. The intensity of ions shows their relative abundance in the sample. This spectrum is fed into a software that contains a database of all proteins in the experimental species. In the database, proteins have been *in silico* “digested” with the same enzyme used in the experiment creating a library of peptide mass fingerprints (PMF). The real mass spectrum is compared to the theoretical one and the software creates a list of peptides that best matches the available data. In other words, the computer is sorting the m/z of peptides in a way that gives the least probability of identifying proteins that are not in the sample (i.e. false positives). The probability of having false positive hits in the final list of proteins is presented as false discovery rate (FDR).

Deducing the presence of a peptide from a given m/z is not an error-free process. In a typical database of thousands of proteins, digestion yields hundreds of thousands of peptides, many of which will have the same (or very close) m/z . This means that one experimental peptide can be matched to several theoretical peptides. A better way to confirm the presence of a peptide is by knowing its amino acid sequence. This information is obtained by tandem MS (denoted as MS/MS or MSⁿ) when ions in the mass analyser are fragmented following the initial MS scan (Nesvizhskii et al., 2003). The produced ions are analysed in the same way like the precursor ions to create an MS² spectrum, and if they are further fragmented, it creates an MS³ spectrum and so on. Sequence information obtained by MS/MS significantly improves the quality of the data, but not all mass analysers are capable of doing tandem MS. The probability of peptides and proteins of being true hits is calculated based on MS/MS and is reported as their respective scores.

1.6.4. Stable isotope labelling of amino acids in cell culture

One of the challenges of MS-based interaction experiments is distinguishing between prey and background proteins. Noise is generated by proteins that stick non-specifically to the purification matrix and do not interact with the bait. Housekeeping proteins, that are generally found at high abundance in cell extracts, are the most common contaminants, but

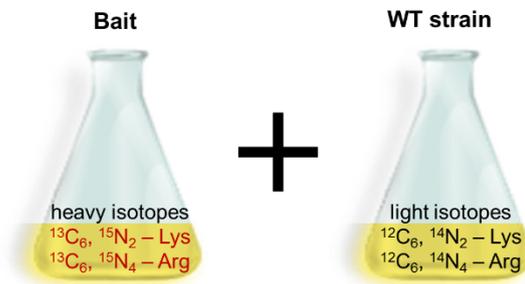
in some cases they can also be preys. Since there is no rule of thumb to identify bait-specific hits, one way to circumvent this problem is to use stable isotope labelling of amino acids in cell culture (SILAC) (fig. 1.6). In SILAC, two cell cultures are grown in differentially labelled media (Ong and Mann, 2006). One culture grows as normal in the presence of naturally occurring light amino acids, while the other grows in medium supplied with one or more amino acids containing heavy isotopes, such as ^{13}C or ^{15}N . After a few cycles, the amino acids are incorporated into newly synthesised proteins and the proteome becomes fully “labelled”. The heavy amino acids do not affect cell growth, nor interfere with any cellular processes.

In MS, heavy and light peptides with identical sequence produce ions with slightly different m/z . In the mass spectrum chromatogram, they appear as peak pairs just a few daltons (Da) apart. The intensity of the peaks from MS1 is used to compare the relative abundance of proteins in the original samples (Trinkle-Mulcahy, 2012). Thus SILAC-MS provides additional quantitative information about the analysed samples.

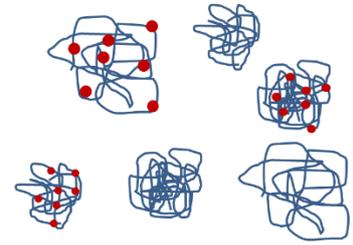
In a typical AP-MS experiment only one of the cultures expresses a tagged bait of interest, while the other one expresses either the tag alone or no tagged proteins. The bait is purified via the tag from a combined cell extract containing equal amount of proteins from both cultures. Following MS analysis, contaminating proteins show similar intensity of light and heavy peptides, while the bait and preys are enriched in the isotopic versions of the culture that they came from.

A number of other labelling strategies for quantitative MS have also been developed, including isobaric tags for relative and absolute quantitation (iTRAQ), tandem mass tags (TMT), isotope-coded affinity tag (ICAT), ^{18}O labelling and ^{15}N labelling. The most popular of these techniques is iTRAQ, where two or more protein samples are prepared separately and digested to peptides, which are then labelled at the N-terminus with different isobaric tags. The samples are then mixed together and analysed by MS. All isobaric tags have the same mass but can be distinguished from each other through the release of a different reporter ion during collision-induced dissociation. ITRAQ is commonly used to compare four or eight multiplexed samples, which makes it efficient in terms of reducing mass spectrometry time and data analysis. Both, iTRAQ and SILAC, have the advantages of being relatively easy to

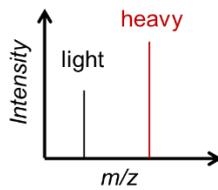
Step 1: Cells are grown in differentially labelled media



Step 2: Cells are mixed together in 1:1 ratio and lysed, and the bait is pulled down



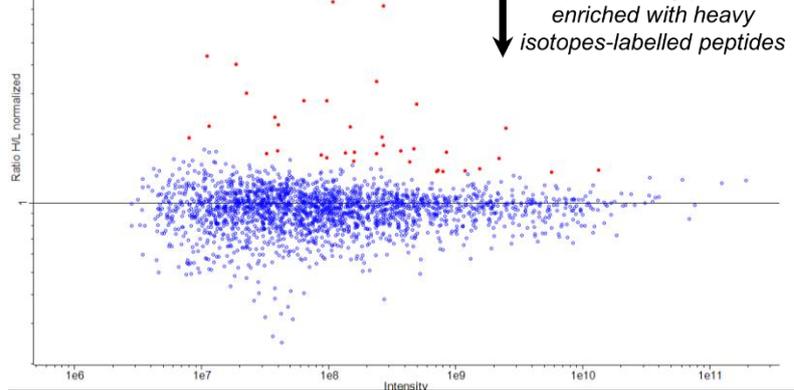
all proteins from the bait strain are labelled with heavy Lys and Arg



EELFTK
EELFTK
VDHVSDVQEISLLK
GINRTQQVIVAIR
GINRTQQVIVAIR
MAPYQGGMR

Step 3: Separate proteins on a gel a digest with trypsin

Step 4: Run peptides through LC-MS/MS and identify proteins



Step 5: Identify proteins enriched with heavy isotopes-labelled peptides

Fig. 1.6: Identifying protein interactions by co-IP-MS using SILAC. In SILAC, two cell cultures are grown in the presence of either heavy or light isotopes. Any amino acid could be labelled, but lysine and arginine are the most common choice. This is because downstream in the experimental procedure, proteins are often digested with trypsin, a protease that cleaves the backbone after each lysine or arginine residue. Thus, all of the resulting peptides will have one of these two amino acids at the C terminus. Consequently, in the heavy labelled proteome, all peptides will carry a heavy isotope and can be used for quantitation.

use and efficient at labelling. However, in SILAC, samples are mixed early in the experimental procedure which minimises handling-induced variation. In contrast, iTRAQ samples are processed separately and therefore this method is generally considered less accurate. SILAC has also been applied to compare up to five samples simultaneously, although the use of multiple isotopes significantly complicates data analysis (Molina et al., 2009). Using SILAC can be disadvantageous in some eukaryotes that can convert arginine to proline, and hence heavy arginine to heavy proline (Marcilla et al., 2011). This can reduce the accuracy of data analysis, since the computer algorithms are not programmed to account for the additional variant of proline. Various experimental and computational methods can be employed to correct this issue (Lossner et al., 2011). This is not a problem in iTRAQ, because the samples are labelled *ex vivo*. For the same reason, iTRAQ can be used to compare samples prepared *in vitro*, whereas SILAC is only used for *in vivo* labelling.

SILAC experiments have traditionally involved generation of auxotrophic strains that must take up the heavy amino acids to survive. Recent studies have demonstrated that *S. cerevisiae*, *S. pombe* and *E. coli* all downregulate lysine biosynthesis in the presence of exogenous lysine and use the amino acids from the media instead (Frohlich et al., 2013). The authors concluded that prototrophs are a viable option in what they termed native SILAC (nSILAC). However, subsequent analysis of protein turnover in *S. cerevisiae* revealed that the yeast achieves full metabolic labelling during exponential growth but it restarts the production of endogenous lysine during stationary phase (Martin-Perez and Villeñ, 2015). Nevertheless, nSILAC presents a promising new approach for studying proteome dynamics in actively growing cells.

1.7. Aims of this study

This study aimed to first of all, develop an unbiased screen for kinase and phosphatase interacting partners in *C. albicans* using co-IP and MS as core techniques and second, to look at differences in KPI between yeast and hyphae. Since KPI screens by MS have not been previously performed in *C. albicans*, the project was designed to test established methods in other fungi and optimise them for use in *Candida*. Initially, experiments were guided by protocols published by Breitkeutz et al. (2011), who developed a global KPI network screen in *S. cerevisiae* using label-free MS. Two kinases (Dbf2 and Crk1), one phosphatase (Cdc14) and one kinase regulatory subunit (Mob1) were selected for preliminary experiments based on their established role in *C. albicans* morphogenesis. These four protein candidates were used as baits in a series of co-IP experiments with the aim to optimise AP-MS for each individual protein. Following optimisation of co-IP experiments in conjunction with label-free MS analysis of potential substrates, the project focused on investigating Cdc14 interactions further with the use of substrate-trapping technology and quantitative MS analysis. Strategies were developed and optimised in *C. albicans* in conjunction with SILAC-MS for the analysis of interacting partners. This thesis describes the first application of SILAC in *C. albicans* and demonstrates that it is a viable method for studying protein interaction in this organism. Furthermore, proteins were labelled with heavy arginine and lysine in a strain that is auxotrophic for lysine, but prototrophic for arginine, suggesting that *C. albicans* may also be used for nSILAC. Finally, this study identified over 100 potential Cdc14 substrates in yeast and hyphae, suggesting a role of the phosphatase in DNA damage repair, chromosome segregation, cytoskeleton organisation and mitotic progression.

Chapter 2

Materials and Methods

2.1. Cell culture techniques

2.1.1. Growth media

All media was sterilised by autoclaving at 121 °C, 15 psi for 20 min after preparation. All media was stored at room temperature and once plated solid media was stored at 4 °C.

YPD

YPD was prepared using either of the following recipes:

Ingredients	Amount
Difco Bacto-yeast extract	10 g
Difco Bacto-peptone	20 g
D-glucose (Fisher Scientific)	20 g
Uridine	80 mg
Distilled water	Up to 1 L

Ingredients	Amount
Formedium™ YPD broth	50 g
Uridine	80 mg
Distilled water	Up to 1 L

YNB

YNB was prepared using either of the following recipes:

Ingredients	Amount
Difco yeast nitrogen base without amino acids	6.7 g
D-glucose	20 g
Uridine*	80 mg
Arginine*	80 mg
Lysine*	80 mg
Histidine*	80 mg
Distilled water	Up to 1 L

Ingredients	Amount
Formedium™ yeast nitrogen base without amino acids	6.9 g
D-glucose	20 g
Uridine*	80 mg
Arginine*	80 mg
Lysine*	80 mg
Histidine*	80 mg
Distilled water	Up to 1 L

*Amino acids were added selectively depending on cells' requirements.

MET3 promoter-inducing media

Ingredients	Amount
Difco yeast nitrogen base without amino acids	6.7 g
D-glucose	20 g
Formedium™ complete supplement mixture drop-out: -arginine -lysine -methionine	670 mg
Uridine	80 mg
Arginine	80 mg
Lysine	80 mg
Distilled water	Up to 1 L

MET3 promoter-repressing media

Ingredients	Amount
Difco yeast nitrogen base without amino acids	6.7 g
D-glucose	20 g
Formedium™ complete supplement mixture drop-out: -arginine -lysine -methionine	670 mg
Methionine	60.6 mg
Cysteine	373 mg
Uridine	80 mg
Arginine	80 mg
Lysine	80 mg
Distilled water	Up to 1 L

5-FOA media

Ingredients	Amount
Difco yeast nitrogen base without amino acids	6.7 g
D-glucose	20 g
5-fluoroorotic acid*	1 g
Uridine	80 mg
Arginine	80 mg
Lysine	80 mg
Distilled water	Up to 1 L

*5-fluoroorotic acid powder was resuspended in water and sterilised with a 22µm filter. It was added to solution after the media was autoclaved and cooled down to 55 °C in a water bath.

Heavy isotopes enriched media

Ingredients	Amount
Difco yeast nitrogen base without amino acids	6.7 g
D-glucose	20 g
Uridine	80 mg
Arginine (¹³ C ₆ , 99%; ¹⁵ N ₄ , 99%)*	100 mg
Lysine (¹³ C ₆ , 99%; ¹⁵ N ₂ ,99%)*	100 mg
Distilled water	Up to 1 L

*Heavy isotopes of arginine and lysine were purchased from Cambridge Isotope Laboratories, Inc.

2TY

Ingredients	Amount
Difco Bacto-yeast extract	10 g
Difco Bacto-tryptone	11 g
NaCl	5 g
5 M NaOH	adjust pH to 7.4
Ampicillin*	100 mg
Distilled water	Up to 1 L

*Ampicillin was filter sterilised and added after the media was autoclaved.

LB

Ingredients	Amount
Difco Bacto-yeast extract	5 g
Difco Bacto-tryptone	10 g
NaCl	10 g
5 M NaOH	adjust pH to 7.4
Ampicillin*	100 mg
Distilled water	Up to 1 L

*Ampicillin was filter sterilised and added after the media was autoclaved.

Solid media

Solid media was prepared by adding 20 g/L Difco Bacto-agar to any of the liquid broth prior to autoclaving.

Hyphae-inducing media

Hyphae-inducing media was prepared by adding 20% v/v fetal new-born calf serum to the appropriate liquid broth immediately before the media was used. The media was then pre-warmed to 37 °C before any cells were added.

2.1.2. Growth conditions

C. albicans strains were revived by inoculating frozen stock of cells on solid media and leaving the plates at 30 °C overnight. Plates were then stored at 4 °C for up to 1 week.

***C. albicans* yeast**

C. albicans yeast was routinely grown in appropriately selected liquid broth in conical flasks shaking at 200 rpm at 30 °C. Stationary phase culture was grown overnight. Log phase culture was prepared by re-inoculating overnight culture into fresh medium at $OD_{595} = 0.25 \pm 0.02$ and letting the cells grow until the culture reaches $OD_{595} = 0.8 \pm 0.02$.

***C. albicans* hyphae**

C. albicans hyphae were induced by inoculating a stationary phase yeast culture into selected pre-warmed hyphae-inducing media. Cells were then grown by shaking at 200rpm at 37 °C until hyphae reached the size of interest.

E.coli

E. coli cells were grown in LB or 2TY broth in conical flasks shaking at 200 rpm at 37 °C for the required length of time.

2.1.3. Cell transformations

All reagents were sterilised prior to use. Transformations were done near a Bunsen flame to minimise risk of contamination.

C. albicans

An overnight yeast culture was inoculated in 50 ml YPD at $OD_{595} = 0.25$ and incubated at 30 °C to $OD_{595} = 0.8$. Cells were pelleted by centrifugation at 3000 rpm for 1 min and transferred to an eppendorf tube where they were washed with 1 ml of wash buffer (TE [10 mM Tris-HCl, 1mM EDTA, pH 7.4], 100 mM LiAc, distilled water) and pelleted at 7000 rpm for 15 sec. Cells were then resuspended in 200 µl wash buffer by gentle pipetting. Each transformation reaction contained the following reagents:

Ingredients	Amount
1 M LiAc	36 µl
10 x TE	30 µl
Single stranded DNA from salmon testes (10 g/L)	10 µl
60 % w/v PEG	300 µl
Cell suspension	100 µl
Precipitated DNA to be inserted*	Combined amount of 5 PCR reactions
DMSO**	40 µl

*A transformation reaction containing 100 µl of distilled water instead of DNA was used as a negative control.

**All ingredients except DMSO were mixed together in an eppendorf tube, which was then incubated at 30 °C overnight. DMSO was added to the mixture on the following morning and the reaction was incubated at 42 °C for 15 min. The cells were pelleted by centrifugation at

8000 rpm for 15 sec, resuspended in 200 µl distilled water and plated on minimal media supplemented with the appropriate amino acids. Plates were incubated at 30 °C for 2-4 days and single colonies were tested for successful integration of the DNA by colony PCR and western blot. Successful transformants were grown on a separate plate and stored in the collection as described in section 2.1.4.

E. coli

All *E. coli* transformations were done using the strain *DH5α* from Delta Biotechnology.

In order to make the cells competent, 10 µl of an overnight culture was inoculated in 1 L fresh 2TY broth and left to grow at 37 °C with shaking at 200 rpm until it reached an OD₅₅₀ = 0.5. Cells were pelleted by centrifugation at 4000 rpm for 5 min and resuspended in 40 ml freshly prepared ice-cold transformation buffer I. Cells were left for 10 min on ice and pelleted again in the same manner. Cells were resuspended in 5 ml freshly prepared ice-cold transformation buffer II and incubated on ice for 15 min. Cells were then divided into 50 ml aliquots, frozen in liquid nitrogen and stored at -80 °C.

Transformation buffer I:

30 mM KAc

10 mM RbCl₂

10 mM CaCl₂

50 mM MnCl₂

15 % v/v glycerol

pH to 5.8 using acetic acid

Transformation buffer II:

10 mM MOPS

75 mM CaCl₂

10 mM RbCl₂

15 % v/v glycerol

pH to 6.5 using KOH

In order to transform *E. coli*, competent cells were defrosted on ice for 5 min and 25 µl of cells were mixed with either 1 µl of plasmid DNA or 1 µl of water (negative control). Cells were incubated for 1 min at 42 °C and for 3 min on ice. A hundred microliters of 2TY was added to each reaction before cells were plated on 2TY solid media containing ampicillin and incubated at 37 °C overnight. Individual colonies were screened for positive results by inoculating them in 5 ml 2TY containing ampicillin and isolating plasmid DNA by doing a miniprep as described further bellow.

2.1.4. Strain storage conditions

Single colony of transformed cells was inoculated in liquid broth and grown to stationary phase overnight. Cells were then re-inoculated in 50 ml fresh medium and grown to OD₅₉₅ = 0.8. Cells were harvested by centrifugation, washed twice in 1 ml distilled water and resuspended in 2 ml 20 % glycerol. The mixture was split into two eppendorf tubes, which were stored in two separate freezers at -80 °C.

2.1.5. *C. albicans* strains used in this study

Strains constructed in this study

Number in lab collection	Strain	Parental strain	Genotype
1822	Dbf2-HA	MDL04	<i>DBF2/DBF2-HA::URA3</i>
1823	Dbf2-MYC	MDL04	<i>DBF2/DBF2-MYC::ARG4</i>
1824	Cdc14-HA	MDL04	<i>CDC14/CDC14-HA::URA3</i>
1825	Cdc14-MYC	MDL04	<i>CDC14/CDC14-MYC::URA3</i>
1826	Cdc14-2xMYC	MDL04	<i>CDC14-MYC::URA3/CDC14-MYC::ARG4</i>
1827	Cdc14-2xHA	MDL04	<i>CDC14-HA::URA3/CDC14-HA::ARG4</i>

Number in lab collection	Strain	Parental strain	Genotype
1828	Dbf2-2xHA	MDL04	<i>DBF2-HA::URA3/DBF2-HA::ARG4</i>
1829	Dbf2-2xMYC	MDL04	<i>DBF2-MYC::ARG4/DBF2-MYC::URA3</i>
1830	Mob1-MYC	MDL04	<i>MOB1/MOB1-MYC::URA3</i>
1831	Mob1-2xMYC	MDL04	<i>MOB1-MYC::URA3/MOB1-MYC::ARG4</i>
1832	Crk1-MYC	MDL04	<i>CRK1/CRK1-MYC::ARG4</i>
1833	Crk1-2xMYC	MDL04	<i>CRK1-MYC::ARG4/CRK1-MYC::URA3</i>
1834	Cdc14/cdc14Δ	MDL04	<i>CDC14/cdc14::frt</i>
1835	Dbf2-MYC in Cdc14- HA	MDL04	<i>DBF2/DBF2-MYC::ARG4 CDC14/CDC14-HA::URA3</i>
1836	Dbf2-TAP	MDL04	<i>DBF2/DBF2-TAP::ARG4</i>
1837	Cdc5-MYC	MDL04	<i>CDC5/CDC5-MYC::ARG4</i>
1838	Tup1-HA	BWP17	<i>TUP1/TUP1-HA::URA3</i>
1839	Mob1-TAP	MDL04	<i>MOB1/MOB1-TAP::ARG4</i>
1840	Lys2/lys2Δ	BWP17	<i>LYS2/lys2::frt</i>
1841	Mob1-2xTAP	MDL04	<i>MOB1-TAP::ARG4/MOB1-TAP::URA3</i>
1842	MET3-Cdc14	MDL04	<i>CDC14/ARG4::MET3-CDC14</i>
1843	Cdc14-TAP	MDL04	<i>CDC14/CDC14-TAP::URA3</i>
1844	Cdc14-2xTAP	MDL04	<i>CDC14-TAP::URA3/CDC14-TAP::ARG4</i>
1845	cdc14C275S-MYC	MDL04	<i>CDC14/cdc14C275S-MYC::URA3</i>
1846	cdc14C275S-GFP	MDL04	<i>CDC14/cdc14C275S-GFP::ARG4</i>
1847	cdc14C275S-TAP	MDL04	<i>CDC14/cdc14C275S-TAP::ARG4</i>
1848	Mlc1-GFP in cdc14C275S-MYC	MDL04	<i>CDC14/cdc14C275S-MYC::URA3 MLC1/MLC1-GFP::ARG4</i>
1849	MET3-cdc14C275S- MYC	MDL04	<i>CDC14/ARG4::MET3-cdc14C275S- MYC::URA3</i>
1850	MET3-Cdc14/ cdc14C275S-MYC	MDL04	<i>ARG4::MET3-CDC14/cdc14C275S- MYC::URA3</i>

Number in lab collection	Strain	Parental strain	Genotype
1851	Cdc14-GFP/ cdc14C275S-MYC	MDL04	<i>CDC14-GFP::ARG4/cdc14C275S-MYC::URA3</i>
1852	Cdc14-GFP	BWP17	<i>CDC14/CDC14-GFP::ARG4</i>

Other strains used in this study

Strain	Genotype	Reference
BWP17	<i>ura3::imm434/ura3::imm434 his1::his1G/his1::his1G arg4::hisG/arg4::hisG</i>	Wilson <i>et al.</i> , 1999
MDL04	<i>lys2::CmLEU2/lys2::CdHIS1 arg4Δ/arg4Δ leu2Δ/leu2Δ his1Δ/his1Δ ura3Δ::imm434/ura3Δ::imm434 iro1Δ::imm434/iro1Δ::imm434</i>	Gift from Munro Lab, University of Aberdeen

2.2. DNA techniques

2.2.1. PCR of plasmid DNA

DNA sequences containing an epitope tag and a selectable marker of interest were amplified by polymerase chain reaction (PCR). The conditions of the reaction were as follows:

PCR stage	Temperature	Time	Number of cycles
Initial denaturation	95 °C	5 min	1
DNA denaturation	95 °C	30 sec	} 30
Primers annealing	50 - 55 °C*	30 sec	
DNA extension	72 °C	2-4 min**	
Final DNA extension	72 °C	10 min	1

*The annealing temperature of each reaction was primer pair-specific and was determined by trial and error.

**The extension time of each reaction was determined based on the length of the DNA template. One minute extension time was allowed for every 1kbp of DNA template.

Each PCR contained the following ingredients mixed together in nuclease-free sterile tubes:

Ingredients	Amount
ddH ₂ O	25 µl
5x HiFi buffer (Bioline)	10 µl
10 mM dNTP (Bioline)	5 µl
5 µM forward and reverse primers	4 µl each
DNA template at approx. 20ng/ µl	1 µl
Velocity polymerase (Bioline)	1 µl
MgCl ₂ *	4 µl

*MgCl₂ was added to some reactions if it was found to give better product yield. When MgCl₂ was added, the amount of ddH₂O was reduced to 21 µl.

Vector plasmids were typically used as templates for PCR. Primers were designed to contain homologous sequences to the vector that will anneal to it and amplify a cassette, and up to 80 bp flanking sequences that are homologous to a region of genomic DNA where the cassette will be inserted. Successful amplification of DNA was confirmed by agarose gel electrophoreses. When PCR was carried out for the purpose of cell transformation, the products of five reactions were mixed together and the DNA was precipitated as described in section 2.2.3.

2.2.2. Colony PCR

After cell transformation, insertion of a DNA cassette was tested by isolating single colony cells and re-suspending them in 5 µl ddH₂O. Cells were boiled for 5 min at 95 °C and frozen

at -80 immediately afterwards. This step aims to break the cells open, so the DNA is released in solution. To set up a PCR, each tube contained the following reagents:

Ingredients	Amount
Cell suspension	5 μ l
2x Biomix Red	12.5 μ l
5 μ M forward and reverse primers	2 μ l each
ddH ₂ O	3.5 μ l

The conditions of each reaction were as follows:

PCR stage	Temperature	Time	Number of cycles
Initial denaturation	95 °C	5 min	1
DNA denaturation	95 °C	30 sec	} 35
Primers annealing	50 – 55 °C*	30 sec	
DNA extension	68 °C	1-2 min*	
Final DNA extension	68 °C	10 min	1

*As already mentioned above, primers annealing temperature and DNA extension time were reaction-specific and determined as previously described.

Primers used in colony PCR were designed to amplify a DNA product containing a few hundred base pairs from the 3' end of the tagged gene and a few hundred base pairs from the 5' end of the epitope tag. Amplification of the desired PCR product was confirmed by agarose gel electrophoresis.

2.2.3. DNA precipitation

In order to precipitate DNA products of PCR, the contents of five reactions were pooled together to a total volume of 250 μ l and mixed with 25 μ l of 3.5 M NaAc at pH 5.2 and 675 μ l of 100% ethanol. The solution was incubated at -20 °C overnight and then spun at maximum speed in a microfuge at 4 °C for 30 min. The supernatant was discarded and the pellet was resuspended in 500 μ l of 70% ethanol. The solution was centrifuged again and the supernatant was discarded. The pellet was air-dried in a sterile hood for 15 min and then resuspended in 100 μ l distilled water. Successful DNA precipitation was confirmed by agarose gel electrophoresis.

2.2.4. Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA molecules according to their size and charge in order to check for the presence of DNA or estimate its amount in solution, e.g. after PCR or DNA precipitation. Commonly, agarose gel was prepared using 1% w/v agarose dissolved in 100ml TAE buffer (400mM Tris-HCl, 200 mM acetate and 10 mM EDTA) and 1 μ l EtBr. Settled gel was fully submerged in TAE buffer and DNA aliquots diluted with DNA loading buffer (Bioline) were loaded in wells. HyperLadder I (Bioline) was routinely used as a DNA marker in order to visually estimate the size of DNA molecules. Electric current of 80 V was applied for 40 min to stimulate migration of DNA along the gel. DNA was visualised by exposing the gel to UV transillumination.

2.2.5. Restriction digest

Restriction digest of DNA with endonucleases was performed prior to ligating two molecules in order to generate compatible ends. All enzymes and buffers were purchased from New England Biolabs (NEB). Each reaction contained 1 μ l of two different nucleases, 2 μ l of DNA, 2 μ l of the NEB-recommended buffer and 14 μ l of ddH₂O. All ingredients were mixed together in a tube and incubated at 37 °C for 2 hrs followed by 65 °C for 5 min. Correct size of the cut DNA fragments was confirmed by agarose gel electrophoresis.

2.2.6. DNA gel extraction

Following endonuclease digestion, DNA fragments were separated on agarose gel and the bands were cut out of the gel with a scalpel while being illuminated by UV light. DNA was purified from the gel using Qiagen's QIAquick Gel Extraction Kit and following the manufacturer's protocol.

2.2.7. DNA ligation

DNA ligation reactions contained a linearised vector and an insert in a molar ratio of 1:2, 1 μ l of T4 DNA ligase (NEB) and 1 μ l of 10x T4 buffer (NEB) in a total volume of 10 μ l. The reaction was incubated at 16 °C for 15 hrs and at 65 °C for 10 min.

When linker oligonucleotides were used, 1 μ l of both oligos was mixed with 23 μ l of ddH₂O and incubated at 95 °C for 2 min and then at 45 °C for 10 min to allow for hybridisation. One microliter of this solution was added to the ligation reaction before incubation at 16 °C.

One microliter of the ligation reaction was used to transform *E. coli* cells as described above.

2.2.8. Plasmid DNA miniprep

Plasmid DNA was prepared using QIAprep Spin Miniprep Kit from Qiagen following the manufacturer's instructions. DNA was eluted in 100 μ l of ddH₂O and stored at -20 °C.

2.2.9. DNA sequencing

All DNA sequencing was done by the University of Sheffield Core Genomic Facility. Samples were prepared by following the facility's guidelines.

2.2.10. Oligonucleotides used in this study

All oligonucleotides were produced by Sigma Aldrich.

Primer	Sequence in direction 5'->3'
S1-Cdc14-XFP	GGATTGCTTCTGGAAACTCACAACATCAAGAGCACACTCTGGTGGTGTGA GAAAGTTAAGTGGAAAGAAACATggtgctggcgcaggtgcttc
S2-Cdc14*- XFP	CCGACTTGGCCAAGCCTAGATCCCGACTAATAGGAATTGATTTGGATGGTAT AAACGGAAACAAAAAAGAGCTGGTACTACTctgatatcatcgatgaattcgag
Cdc14*- TAP/HA/MYC_ R	CCGACTTGGCCAAGCCTAGATCCCGACTAATAGGAATTGATTTGGATGGTAT AAACGGAAACAAAAAAGAGCTGGTACTACTctgatgaattcgagctcgtt
S2-Cdc14-XFP	GGATTTTCGATATATTGGCTTTTGCATATGGTTCGGAAGAACAAATTGAAATT GTTGAACCAGCTTATGAAGAAGACTAATTTAGtctgatatcatcgatgaattcgag
Cdc14-TAP- MYC_F	CTCACAACATCAAGAGCACACTCTGGTGGTGTGAGAAAGTTAAGTGGAAA GAAACATggtcgcagggatccccgggttagaacagaagcttatatccgaa
Cdc14-TAP- HA-MYC_R	CGAGTGGCCTATCCAAAAGATTCAACTCAGCCTTATTCCAATAACTGGATTG AATTGAGTGAAGATAGTGATATTGCTGtctgatgaattcgagctcgtt
Cdc14-chk_F	GACATCTCCACTTGCTGATTCTTCTG
Cdc14-chk- myc_R	GGCTTTTGCATATGGTTCGGAAGAAC
Cdc14- URAF_F	GTCTTTCATTCAAAAACACGTTTGTCTTCTACCATACCGTTCTAAACTACCTAA TAATCACAACACCTTTTCCGctcgaggaagttcctatactttc
Cdc14- URAF_R	CGAGTGGCCTATCCAAAAGATTCAACTCAGCCTTATTCCAATAACTGGATTG AATTGAGTGAAGATAGTGATATTGCTGctctagaactagtgatctgaagtt
NotI-Cdc14- 3'_F	GGCGGG GCGGCCGC GTCTTCTTCATAAGCTGGTTCAAC

Primer	Sequence in direction 5'->3'
SacI-Cdc14-3'_R	GGCGGG GAGCTC CGGAGAATACAAGTACCATTCTCAAG
Cdc14-S1-MET3_F	AAATGTATATAACGAAGATGACTATCATCAATGGTCCGGTTAGTAAAGCGA ACAAGCTTTATAAAAATAGTTATGCTGAACGTACCATgaagcttcgtacgctgcag gtc
Cdc14-S2-MET3_R	AAAGGTAGAACAATCAATTTGAAGTAGATTTTCCCAACATACTTTTAAGAAA CTCTATAAGAGGCACATGAACCAGTGAACATGcatgttttctggggagggtatttac
MET3-chk_F	GCGCCCTCTAAAACAATACCC
Cdc14-5'chk_R	GGTAATGCGTCTTCAACTGTG
Cdc14-TAP-MYC_F	CTCACAACATCAAGAGCACACTCTGGTGGTGTGAGAAAGTTAAGTGGAAA GAAACATggtcgcacggatccccgggttagaacagaagcttatatccgaa
Cdc14-TAP-HA-MYC_R	CGAGTGGCCTATCCAAAAGATTCAACTCAGCCTTATTCCAATAACTGGATTG AATTGAGTGAAGATAGTGATATTGCTGtcgatgaattcgaactcgtt
Cdc14-chk_F	GACATCTCCACTTGCTGATTCTTCTG
Cdc14-chk-myc_R	GGCTTTTGCATATGGTTCGGAAGAAC
NotI-Cdc14-3'_F	GGCGGG GCGGCCGC GTCTTCTTCATAAGCTGGTTCAAC
SacI-Cdc14-3'_R	GGCGGG GAGCTC CGGAGAATACAAGTACCATTCTCAAG
Cdc14-S1-MET3_F	AAATGTATATAACGAAGATGACTATCATCAATGGTCCGGTTAGTAAAGCGA ACAAGCTTTATAAAAATAGTTATGCTGAACGTACCATgaagcttcgtacgctgcag gtc
Cdc14-S2-MET3_R	AAAGGTAGAACAATCAATTTGAAGTAGATTTTCCCAACATACTTTTAAGAAA CTCTATAAGAGGCACATGAACCAGTGAACATGcatgttttctggggagggtatttac

Primer	Sequence in direction 5'->3'
S2-Cdc14-XFP	GGATTTTCGATATATTGGCTTTTGCATATGGTTCGGAAGAACAAATTGAAATT GTTGAACCAGCTTATGAAGAAGACTAATTTAGtctgatatcatcgatgaattcgag
BamHI linker	GAT CCC TCC CAG AAC
XhoI linker	TCG AGT TCT GGG AGG
<u>XhoI-Cdc14</u> F	GGC GGG CTC GAG GGC TTT CCT TTC CTT TGC TAT G
XbaI-Cdc14- MYC_R	GGC GGG TCT AGA CTA ATT TGT GAG TTT AGT ATA CAT GC
Cdc14-seq1_R	GAGGCACATGAACCAGTGAAC
Cdc14-seq2_F	GATGGAAGAGATCTTTTTGGAATTC
Cdc14-seq3_F	CCAGAATTGGGCTCCTCATCAAG
Cdc14-seq4_F	GGTTGTTTGATTGGAGCCCATC
Cdc14-seq5_F	GCTCACCAGCAAGGTATGACTC
Cdc14 C275S F	GCAGTACATTCTAAAGCAGGGTTAGG
Cdc14 C275S R	CCTAACCTGCTTTAGAATGTACTGC
Cdc14C275S_c hk_R	CCGGTTCTTCCTAACCTGCTTTAG
URA-F	GGAAGAGATCCAGATATTGAAGG
URA-R	TGTGCTACTGGTGAGGCATG
ARG-F1	CTGCTAAAAGTGCCGTTTTAAAACAATT
ARG-R1	ACCGGTGAAACGACCACCCATAATTT
myc-R	GTATACATGCATTTACTTATAATGGCGCGC
Crk1-TAP- MYC_F	CAGTTTATAAAAGAATTATAAATGAGAAAATGAGGTTTGAAAAGTTATCGG GAGGACACAAATCTATGggtcgacggatccccgggtagaacagaagcttatatccgaa

Primer	Sequence in direction 5'->3'
Crk1-TAP-HA-MYC_R	GCTCAGTTGCAAGAATGGTTTAGTGGTAAAATCCAACGTTGCCATCGTTGG GCCCCGGGTTTCGATTCCCGGTTCTTGcTcgatgaattcgagctcgtt
Crk1-chk_F	GTGCTGTTGCTGTCTAGATCG
Crk1-chk_R	CGTGACTTGATGGACCTAAGG
Dbf2-TAP-MYC_F	GGAAATGGAATTGGAAATGGAAATCTCGATCAAGTAGATTAAATCCATTA GCTACATTGTATggtcgcgacggatccccgggttagaacagaagcttatatccgaa
Dbf2-TAP-HA-MYC_R	GATAAAATTAAGAATGATTATATTTGGAAACAAGAAAGGGAAGATGAATAA GAAGAAGAAGAAGAATAGTGGGGAGTGGTcgatgaattcgagctcgtt
Dbf2-chk_F	CTCCCAATTGGATAATGAAGAAGATGC
Dbf2-chk-myc_R	GCCGGATCTCTACGAGTTTACAAGTC
S1-Mlc1-XFP	GTTGATGAGTTATTA AAAAGGGTCAATGTA ACTTCTGATGGAAATGTGGAT TATGTTGAATTTGTCAAATCAATTTTAGACCAAggtgctggcgcaggtgcttc
S2-Mlc1-XFP	GGGAACGAGATGGAATCTTTCGTTACGCCTCACATCTGTTTCAGGGTTATCT ATGCTATTAGCTGTTATCGTTATGCTTTCACTCtctgatatcatcgatgaattcgag
Mlc1-chk_F	CATCAACAGACCAGACGGTTTC
Mob1-TAP-MYC_F	CAATTAATTAGCAGGAAAGACTACGGTCCATTAGAGGACTTGGTAGACACG ATGCTTCAAAGAggtcgcgacggatccccgggttagaacagaagcttatatccgaa
Mob1-TAP-HA_F	CAATTAATTAGCAGGAAAGACTACGGTCCATTAGAGGACTTGGTAGACACG ATGCTTCAAAGAggtcgcgacggatccccgggttataccatac gatgttcctgac
Mob1-TAP-HA-MYC_R	CAATATAAAATCAA ACTAACAAGCTACTTAGATTGCCTACACCAGAAGAAA TGGGGTCACCACCGTCAGGtcgatgaattcgagctcgtt
Mob1-chk_F	CCAGTATCATTACCTGCTTGTG
Mob1-chk_R	CGAAGAGTTAGAGCAAGAAAG

2.2.11. Plasmids used in this study

Plasmids constructed in this study:

Plasmid	Description
pINK1	<i>GFP</i> gene was cut out of pRSC3 vector and replaced with <i>CDC14-MYC</i> sequence, including 400 bp upstream sequence of <i>CDC14</i>
pINK2	pINK1 vector containing mutation <i>cdc14C275S</i>
pINK3	pINK2 vector, where 400 bp downstream sequence of <i>CDC14</i> was cloned between <i>SacI</i> and <i>NotI</i> restriction sites

Other plasmids used in this study:

Plasmid	Application	Reference
pFA-MYC-URA3	Amplification of MYC-URA3 cassette	Lavoie <i>et al.</i> , 2008
pFA-MYC-ARG4	Amplification of MYC-ARG4 cassette	Lavoie <i>et al.</i> , 2008
pFA-HA-URA3	Amplification of HA-URA3 cassette	Lavoie <i>et al.</i> , 2008
pFA-HA-ARG4	Amplification of HA-ARG4 cassette	Lavoie <i>et al.</i> , 2008
pFA-TAP-URA3	Amplification of TAP-URA3 cassette	Lavoie <i>et al.</i> , 2008
pFA-TAP-ARG4	Amplification of TAP-ARG4 cassette	Lavoie <i>et al.</i> , 2008
pFA-ARG4-MET3	Amplification of ARG4-MET3 cassette	Gola <i>et al.</i> , 2003
pFA-GFP-URA3	Amplification of GFP-URA3 cassette	Gola <i>et al.</i> , 2003
pFA-GFP-ARG4	Amplification of GFP-ARG4 cassette	Gola <i>et al.</i> , 2003
pFA-URA3	Amplification of URA3 cassette	Gola <i>et al.</i> , 2003

2.3. Protein techniques

2.3.1. Soluble protein extraction

Protein extraction was achieved by breaking the cells open to release their protein contents in solution and removing the cell debris by centrifugation afterwards. This method captures only the soluble proteins of the cell and omits the insoluble fractions such as membrane-embedded proteins. Cell pellets were commonly resuspended in ice-cold lysis buffer (20 mM HEPES, 150 mM NaCl, pH 7.4, EDTA-free protease inhibitors (Roche), 1mM PMSF) and depending on the total volume, cells were broken by one of the following methods:

- Small scale extraction

When cells were resuspended in less than 1 ml of lysis buffer, they were mixed with the same volume of acid-washed glass beads in a tube and violently agitated in a Mini-Beadbeater-16 (Biospec) for 3x 30 sec. Tubes were chilled on ice for 1 min between beatings. Cell debris was pelleted by centrifugation in a top bench centrifuge at maximum speed for 10 min at 4 °C. The supernatant was transferred to a clean eppendorf tube before being used in downstream applications.

- Large scale extraction

When large volume of cell lysate was required, pelleted cells were resuspended in 10 ml of lysis buffer and broken in a high pressure cell disrupter (Constant Systems Ltd.) at 35 psi, 4°C. Cell debris was pelleted in a Beckman Coulter Avanti™ J-20 centrifuge using rotor JA-20 at 20 000 rpm, 4 °C for 15 min. The supernatant was passed through a 44 µm syringe filter to get rid of small residual debris. This method gives approximately 10 ml of cell lysate.

2.3.2. SDS-PAGE

Proteins were separated according to their size and charge using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Mini-PROTEAN® TGX™ Precast Gels 4-15% (Bio-Rad) were used in all mass spectrometry experiments. For all other applications, gels were prepared using the following recipe:

Resolving gel

Reagents	6%	8%	10%	12%
ProtoGel (30%)	2 ml	2.67 ml	3.33 ml	4 ml
ProtoGel Resolving Buffer (4x)	2.5 ml	2.5 ml	2.5 ml	2.5 ml
ddH ₂ O	5.39 ml	4.72 ml	4.06 ml	3.39 ml
10% w/v APS	100 µl	100 µl	100 µl	100 µl
TEMED	10 µl	10 µl	10 µl	10 µl

Stacking Gel

Reagents	Amount
ProtoGel (30%)	0.52 ml
ProtoGel Stacking Buffer	1 ml
ddH ₂ O	2.44 ml
10% w/v APS	20 µl
TEMED	2 µl

Polymerised gels loaded with protein samples were run in running buffer (25mM Tris-Base, 190 mM glycine, 0.1 % w/v SDS) at 180 V for as long as necessary. Protein samples were diluted with 5x protein loading dye (62 mM Tris-HCl (pH 6.8), 25% v/v glycerol, 25%

w/v SDS, 0.01% w/v bromophenol blue, 0.05% β -mercaptoethanol) and boiled for 5 min at 95 °C prior to loading on the gel. In order to estimate protein size Prestained Protein Marker 7-175 kDa (NEB) was loaded on each gel.

2.3.3. Western blot

Proteins were transferred from the polyacrylamide gel onto nitrocellulose membranes (GE Healthcare) using Bio-Rad's Wet/Tank Blotting System Mini Trans-Blot® Cell. Protein transfer was done in transfer buffer (25 mM Tris-HCl (pH 7.6), 190 mM glycine, 20 % v/v methanol) at 150 mA for 100 min. Successful transfer was confirmed by staining membranes with 0.2% w/v Ponceau S. Membranes were then washed with water to remove the dye and placed in SNAP i.d.® 2.0 Protein Detection System (Merk Millipore) where all subsequent steps were done. Membranes were blocked with 1 % w/v BSA dissolved in TBST buffer (20 mM Tris pH 7.5, 150 mM NaCl, 0.1% Tween 20, pH 7.6) for 30 min. Both primary and secondary antibodies diluted in 1 % BSA were applied for 15 min, and membranes were washed with 3x 15 ml TBST buffer after each application. Membranes were then transferred to falcon tubes containing 1 ml of each enhanced chemiluminescence (ECL) solutions I and II and incubated at 4 °C in the dark for 5 min with gentle rolling. Protein bands were visualised using GeneGnomeXRQ gel doc system and GeneTools analysis software, both from Syngene.

ECL I:

1 ml 1M Tris-HCl pH 8.5

100 μ l 250 mM luminol in DMSO

44 μ l 90 mM p-coumaric acid in DMSO

8.856 ml sdH₂O

ECL II:

1 ml 1M Tris-HCl pH 8.5

6.1 μ l 30 % H₂O₂

8.994 ml sdH₂O

Antibodies used in this study:

Antibody	Dilution	Type	Supplier
Mouse anti-HA	1:3 000	Primary monoclonal	Bioserv
Mouse anti-MYC	1:1 000	Primary monoclonal	Bioserv
Mouse anti-GFP	1:1 000	Primary monoclonal	Roche
Goat anti-mouse	1:10 000	Secondary polyclonal	Roche
Anti-PSTAIRE	1:10 000	Primary monoclonal	Roche

2.3.4. Stripping of antibodies from a nitrocellulose membrane

Some membranes were probed with two primary antibodies. In this case after proteins were detected with the first antibody, membranes were washed with stripping buffer (20 ml 10% SDS, 12.5 ml 0.5 M Tris-HCl, 67.5 ml ddH₂O, 0.8 ml β -mercaptoethanol) for 2 hrs at 4 °C. Membranes were then washed five times with 10 ml TBST for 10 min and probed with the second primary antibody as described above.

2.3.5. *In vitro* protein dephosphorylation

Dephosphorylation of proteins *in vitro* was achieved by treating 100 μ l protein sample with 1 μ l (400 units/ μ l) λ phosphatase (NEB) in the presence of 1x Buffer for Protein Metallophosphatases (NEB) and 1mM MnCl₂. The reaction was allowed to proceed at 30 °C for 1 hr. A control reaction containing no phosphatase was set up in parallel. The reaction was stopped by adding protein loading dye to the samples and boiling them for 5 min at 95 °C.

2.3.6. Protein purification

Proteins fused to an epitope tag were purified from cell lysates with the use of beads coupled to an antibody against the tag. All steps were carried while keeping the tubes on ice as much as possible. All buffers and solution were pre-cooled to 4 °C prior to use. Cells were lysed as previously described and depending on the beads used as an affinity matrix one of the following protocols was followed:

- Dynabeads® Protein G (Thermo Fisher Scientific)

Magnetic Dynabeads® were used according the manufacturer's instructions with the following modifications. A hundred microliters of bead slurry were transferred to a tube and beads were pelleted using a magnet. After removing the supernatant, 20 µg of either anti-MYC or anti-HA antibody diluted in 400 µl of lysis buffer was added to the beads and they were incubated at 4 °C for 2 hrs with shaking. The beads were pelleted and washed three times with lysis buffer. They were then added to the cell lysate and incubated at 4 °C for 1 hr with inversion. Beads were pelleted once again and washed once with lysis buffer. Beads were resuspended in 50 µl protein loading buffer and boiled at 95 °C for 5 min to elute the antigen.

- EZview™ Red Anti-c-Myc Affinity Gel (Sigma Aldrich)

EZview™ agarose beads were used according to the manufacturer's instructions with the following modifications. Between 20-200 µl of bead slurry were transferred to a tube and beads were pelleted in a bench centrifuge at 1500 rpm for 30 sec. Beads were washed three times with lysis buffer and incubated with the cell lysate at 4 °C for 1-2 hrs with rotation. Beads were then pelleted again and washed 1-3 times with lysis buffer. Antigen was eluted from the beads by adding 1x beads volume of protein loading dye and boiling the suspension at 95 °C for 5 min.

After proteins were eluted in solution, a small aliquot was analysed by Western blot and, if required, a larger aliquot was separated electrophoretically and the gel was stained with Coomassie.

2.3.7. Coomassie staining of polyacrylamide gels

Coomassie staining was used to visualise proteins on polyacrylamide gels. Typically, gels were transferred to a sterile and keratin-free Petri dish and were soaked in approximately 30 ml of Coomassie InstantBlue (Expedeon) for at least 1 hr with gentle shaking. After this, gels were washed with ddH₂O until water was clear. If gels were not used in the same day, they were stored in 1% acetic acid at 4 °C.

2.3.8. Trypsin in-gel digestion of proteins and peptides gel extraction

Peptide samples prepared by trypsin in-gel digestion were commonly analysed by MS. Therefore, great care was taken to minimise keratin contamination during preparation. All working surfaces were cleaned with decontamination solvent (0.1% acetic acid in 10% isopropanol) prior to work. Proteins were separated on keratin-free precast gels. Gels were kept in keratin-free petri dishes and samples and solutions were handled in keratin-free falcon or eppendorf tubes.

Each IP sample was divided into 15-20 bands. Protein bands were excised from a Coomassie-stained gel and cut into ca. 1x 1 mm pieces with the aid of sterile scalpel. Gel pieces were submerged in 100 µl 50 mM ammonium bicarbonate (ABC) for 5 min before adding 100 µl acetonitrile (ACN) for further 10 min. After that the supernatant was removed. Gel pieces were repeatedly washed in this manner with ABC and ACN until the Coomassie stain was no longer visible (usually 2-4 times). Gel pieces were then dried in a vacuum centrifuge at room temperature until they shrink. Dry pieces were soaked in 30 µl digestion buffer (12.5 ng/µl Trypsin in 50 mM ABC). Additional 50 mM ABC was added after 1 hr until it fully covers the gel pieces. Samples were incubated at 37 °C overnight to allow for enzymatic digestion of proteins into peptides. On the next day, supernatants were

transferred to clean lo-bind collection tubes and gel pieces were soaked in 15 μ l 25 mM ABC for 10 min at room temperature. Supernatants were transferred to the collection tubes again and gel pieces were soaked in 50 μ l ACN for 15 min at 37 °C with shaking. Supernatants were transferred to the collection tubes and gel pieces were incubated with 50 μ l 5% v/v formic acid (FA) for 15 min at 37 °C with shaking. Supernatants were transferred to the collection tubes and gel pieces were submerged in 50 μ l ACN for 15 min at 37 °C with shaking. The supernatant was transferred to the collection tubes and the gel pieces were discarded. At this point all digested peptides are found in solution in the collection tubes. Peptides were dried out in a vacuum centrifuge at 30 °C until completely dry. Peptides were stored at -20 until further use.

2.4. Mass spectrometry

Dry peptides were resuspended in ca. 10 µl MS buffer C (0.1% Trifluoroacetic Acid (TFA), 3% ACN prepared with LC-grade H₂O) and sonicated in a water bath for 5 min. An aliquot of this solution was transferred into MS tubes and loaded on an LC-MS instrument for further analysis. The remaining samples were stored at -20 °C.

Depending on the requirements of the MS analysis, samples were processed in either of three MS instruments:

- Amazon ESI-ion trap (Bruker Daltonics)

Label-free samples were loaded on ESI-ion trap MS coupled to an Ultimate 3000 online capillary liquid chromatography system with a 30 cm x 5mm Acclaim PepMap300 C18 trapping column with 300 Å pore size and 5 µm particle size (Thermo Fisher Scientific, Hemel Hempstead, UK). Peptides were separated by linear gradient in Buffer A (0.1% FA, 3% ACN) and Buffer B (0.1% FA, 97% ACN) at 3-36% over 60 min at a flow rate of 300nL/min using 15 cm x 75 µm onto Acclaim PepMap C18 column with 100 Å pore size and 5 µm particle size (Thermo Fisher Scientific, Hemel Hempstead, UK). MS1 profile scans were acquired in positive mode at a range of m/z=300-1500 and a speed of 8100 m/z/s. Collision-induced dissociation (CID) of 8 ions was performed with active exclusion after 2 spectra and release after 2 min. MS2 scans were acquired at a range of m/z=50-3000. During fragmentation, the target value of the trap was 250 000 for maximum of 50 ms.

Raw data was used to generate Mascot Generic Files (MGF) in Data Analysis v 4.1 with the following settings: signal to noise threshold of 1, base peak intensity of 0.1, absolute intensity threshold of 100, peak width at half maximum m/z 0.1 Charge state deconvolution from fragment spectra was allowed. MGF files were submitted to Mascot Daemon v 2.5.1 (Matrix Science) where searches were performed against the *C. albicans* database, taxonomy all entries, with maximum of 2 missed cleavages allowed, peptide charge +2, +3 and +4, MS and MS/MS peptide tolerance of 0.1 Da, methionine oxidation as a variable modification.

- MaXis ESI-QTOF (Bruker Daltonics)

For the purpose of SILAC experiments, samples were loaded on ESI-QTOF coupled to the same LC system with the same settings as described for Amazon ESI-ion trap. Peptides were eluted on the PepMap C18 column (as above) by liner gradient in Buffer A and then Buffer B 4-40% over 90 min at a flow rate of 300 nL/min. MS1 scans were performed in positive profile mode with an m/z range of 100-1800. CID fragmentation was performed with maximum of three precursor ions per cycle with absolute threshold of 3000, active exclusion after 2 spectra and release after 0.25 min. Captive Spray capillary was set to 4500 V and end plate offset was set to 500 V.

Raw data (.baf files) was submitted to Mascot Distiller v. 2.5.1.0 where peak picking and peptide quantitation was done by default parameters for maXis ESI-TOF. Peptides were searched against the *C. albicans* database, with maximum of 2 missed cleavages allowed, MS and MS/MS tolerance of 0.1 Da, methionine oxidation as a variable modification, peptide charge +2, +3 and +4 and SILAC quantitation (K+8, R+10). True proteins require at least 2 peptides, and quantified proteins require at least two peptides with L/H ratios. Data was exported to Excel and contaminants were filtered out. The median L/H ratio of all proteins was calculated and proteins with $L/H < (\text{median } L/H)/2$ were selected as hits.

- Q Exactive HF ESI-Quadrupole Orbitrap (Thermo Scientific)

SILAC samples were also analysed by LC-ESI-Quadrupole Orbitrap coupled to the same LC system with the same settings as described for Amazon ESI-ion trap. Peptides were eluted onto 15cm x 75 μm Easy-spray PepMap C18 column (2 μm particle size and 100 \AA pore size) in Buffer A and then Buffer B 5-35% over 75 min at a constant flow rate of 300 nL/min. Peptide ionisation was performed by Q-Exactive HF NSI. MS1 scans were performed in positive profile mode with an m/a range of 375-1500. CID fragmentation was performed with the 10 most intense ions from the first scan after accumulation of $5e4$ ions in m/z range of 200-2000. Data was acquired with XCalibur software.

Raw data was processed by MaxQuant v. 1.5.2.8 using Andromeda search engine and *C. albicans* database. False discovery rate (FDR) of peptides and proteins was set to 1%. MS

and MS/MS tolerance was 4.5 ppm, 2 missed cleavages were allowed, methionine oxidation was set as a variable modification.

Data from ProteinGroups.txt files was further processed in Perseus v. 1.2.5.6. Contaminants, reversed and identified by site only proteins were filtered out. Hits were selected by Benjamini-Hochberg procedure (shown as significance B in Perseus) using the normalised H/L ratio and protein intensity, with FDR either 1% or 5%.

The use of each MS instrument is indicated in the results chapters, when data is presented.

2.5. Microscopy

All microscopy was performed using live cells. Differential interference contrast (DIC) images were taken on a Leica microscope, model 0202-519-508L coupled to a HC Image Live software and a Hamamatsu digital camera. Fluorescence microscopy was carried out on an Olympus Delta vision Spectris 4.0 microscope (Applied Precision). Images were captured and deconvolved using Softworx™ 3.2.2.

Chapter 3

Optimisation of AP-MS Methods

3.1. Introduction

The phosphorylation status of a protein commonly affects its function and subcellular localisation. This study sought to investigate the molecular targets of protein kinases or phosphatases that are known to be important for hyphal growth in *C. albicans*. Since the experimental workflow includes purification of the enzymes, in order to have better chances at capturing sufficient amount of them, only proteins expressed at relatively high levels were considered. These criteria were fulfilled by the kinases Dbf2 and Crk1, and by the phosphatase Cdc14.

3.1.1. Dbf2 and Mob1

The NDR family kinase DBF2 is essential for *C. albicans* growth (Gonzalez-Novo *et al.*, 2009). In mutants where its expression is inhibited, cells display pronounced defects in actomyosin ring contraction, cytokinesis, mitotic spindle organisation and nuclear segregation. Hyphal growth is also impaired in DBF2-depleted cells, which form wider than normal hyphal tubes that look swollen and fail to form septa.

Dbf2 is highly conserved in higher eukaryotes and it is known to act together with its activating subunit Mob1. In *S. cerevisiae*, binding of Mob1 to Dbf2 is essential for kinase activation (Mah *et al.*, 2001). Although Mob1 has not been characterised in *C. albicans*, it is likely that it acts in a similar manner as activator of Dbf2 on the basis of functional conservation. Additionally, Mob1 likely physically interacts with the substrates that Dbf2 phosphorylates and therefore it was selected as bait on its own in this study.

3.1.2. Cdc14

The phosphatase Cdc14 has been extensively studied in animals and fungi, and it has been found to interact with hundreds of other proteins. Mitotic exit, chromosome segregation and completion of cytokinesis are universally controlled by Cdc14. The phosphatase is also involved in DNA replication in budding yeast, DNA repair in humans, and morphogenesis in the fungi *Fusarium graminearum* (Li *et al.*, 2015) and *C. albicans* (Clemente-Blanco *et al.*, 2006).

CaCdc14 is not essential for cell viability, but deletion mutants display severe defects in cell separation, cell cycle progression and hyphal formation. It is not clear whether Cdc14 regulates cell morphogenesis through distinct pathways, or the observed phenotype is merely a consequence of the cell cycle disruption.

3.1.3. Crk1

The Cdc2-related protein kinase Crk1 is crucial for hyphal development and virulence of *C. albicans* (Chen *et al.*, 2000). Deletion mutants are unable to form hyphae, while ectopic expression of the Crk1 kinase domain (Crk1N) promotes hyphal formation even under yeast-inducing conditions. Additionally, *crk1/crk1* cells exhibit reduced expression of the hyphae-specific genes *ECE1* and *HWP1*. Crk1 stimulates polarized growth through an unknown pathway but independently of hyphal inducers Cph1 and Egf1. The protein sequence of Crk1N is most similar to that of *S. cerevisiae* Bur1 and the human Pkl1/Cdk9. Indeed, ectopic expression of either Crk1 or Crk1N in *S. cerevisiae* can partially complement *bur1* mutants suggesting some functional homology between the kinases Crk1 and Bur1. Crk1N also promotes filamentation of *S. cerevisiae* cells through the transcriptional regulator Flo8. Collectively, these results portray Crk1 as a strong inducer of filamentous growth and thus make it an interesting candidate of this study.

It is clear that Dbf2, Cdc14 and Crk1 all play a role in *C. albicans* morphogenesis, but none of their direct targets were known prior to this study. The homologues of these proteins have been well researched in other organisms and it is expected that some common interactors will be found.

3.2. General overview of experimental workflow

Identification of protein interactions by mass spectrometry is a well-established technique. A single experiment takes several weeks to be completed and many steps have to be optimised to reflect the strength of the protein interaction, the biological properties and the structure of the whole protein complex as well as the individual proteins, the affinity of the purified protein to the pull-down matrix, the stability of the complex *in vitro* and *in vivo*, the concentration and the total amount of protein required, the labelling of the protein, the sensitivity of the MS instruments available and the error-inducing steps in data processing and analysis. Tandem affinity purification (TAP) is the most common method of protein purification, but it is too stringent to preserve weak protein interactions. Complexes were thus purified by co-IP using various combinations of tags and antibodies.

In general, each experiment followed the following steps:

- 1. Both copies of a gene of interest were fused at their C termini to an epitope which was used as an affinity tag to purify the protein coded by the gene** (fig. 3.1). Cells were transformed with a PCR-generated cassette containing a tag and a selectable marker. Expression of the tagged protein was confirmed by Western blot. The same strain was then transformed again in order to tag the second allele with the same epitope but a different selectable marker. Integration of the cassette was confirmed by PCR by amplifying a DNA starting from the 3' of the marker and ending within a genomic sequence adjacent to the marker. Additionally, protein levels were quantified by Western blot, using Cdc28 levels as a normalisation control. A strain having both alleles tagged produced a protein band that is approximately twice brighter than a single tag band relative to Cdc28 protein brightness. The phenotype of the cells was assessed by microscopy to confirm that the tags do not cause any visible defects.

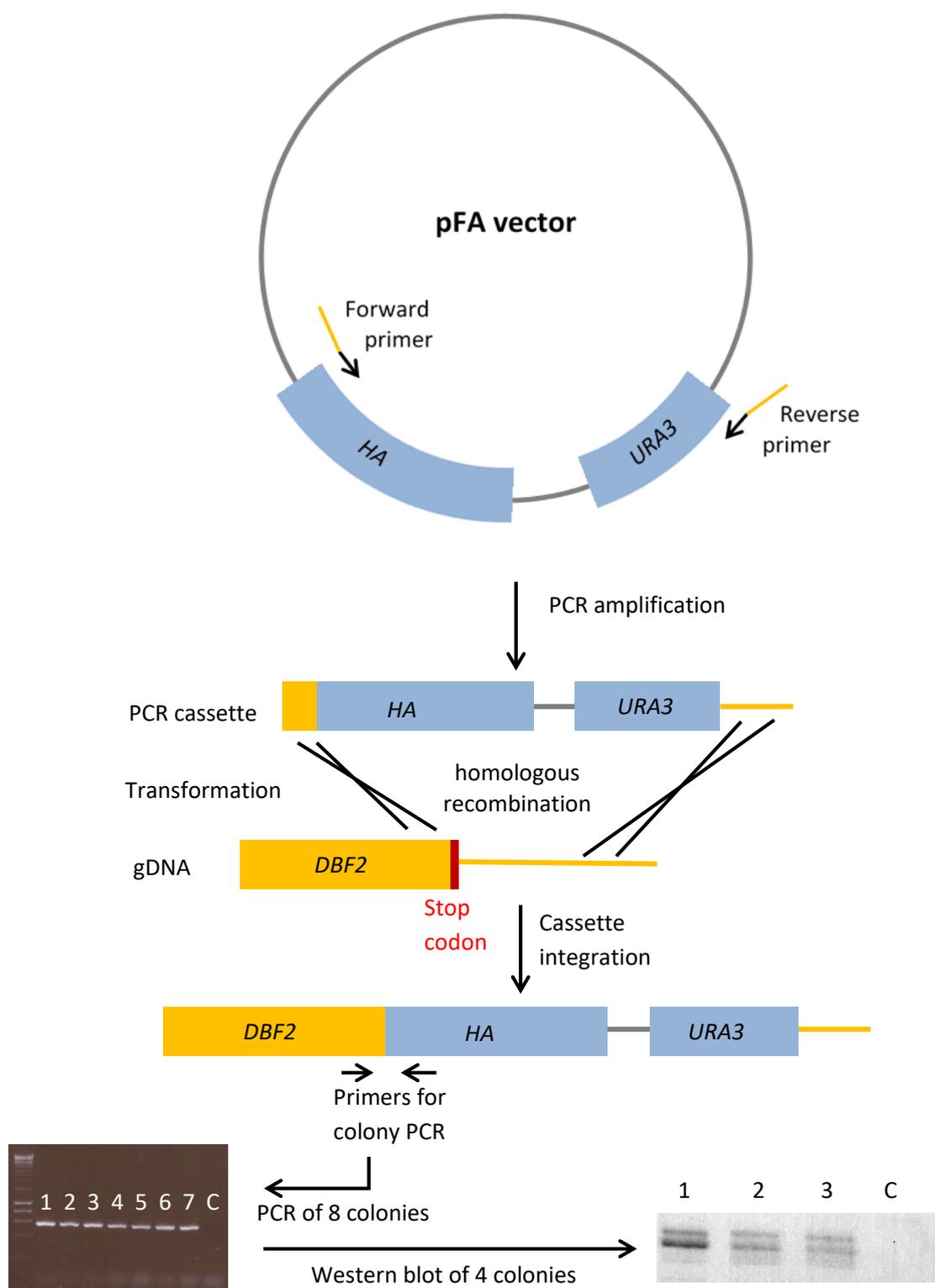


Fig. 3.1: Tagging a protein with an epitope via PCR-based homologous recombination. An epitope of interest (HA) and a selectable marker (URA3) were amplified from pFA vectors using primers with homologous sequences to genomic DNA. The forward primer contains the 3' sequence of the gene to be tagged (*DBF2*) excluding the stop codon. The reverse primer contains a homologous sequence about 100bp downstream of the gene. PCR amplification generates a cassette, which when transformed into *C. albicans* cells integrates at the 3' end of the desired gene. Correct integration is confirmed by colony PCR and protein expression bearing the tag is confirmed by Western blot. Dbf2 produces two bands due to phosphorylation. C – negative control.

- 2. Wild type and epitope-tagged cells were grown in YPD liquid media to an $OD_{595}=0.8$ and lysed, and the bait was precipitated from the cell lysate.** The lysate was incubated with an affinity matrix that interacts with the epitope for 1 hr at 4 °C.
- 3. Different affinity matrixes were tested as described below in order to determine which one gives the highest yield of bait protein.** The bait was eluted from the matrix by boiling in protein loading buffer. At this stage the bait and all associated proteins were released in solution.
- 4. Eluted proteins were separated by SDS-PAGE on a 4-12% gradient gel which was stained with Coomassie in order to visualise protein bands.** A gradient gel allows for optimal separation of all proteins of different sizes. This step is necessary because the MS instrument would be overwhelmed if all proteins were loaded on it simultaneously. For highest precision, the proteins were separated into fractions according to their size.
- 5. All protein bands were cut out of the gel and digested with the protease trypsin.** This step generated 15-20 fractions of proteins. Each fraction was digested overnight with trypsin which has a high specificity for cleaving proteins at the carboxyl sites of the amino acids lysine and arginine, except when either is followed by proline. The generated peptides were extracted from the gel and dried in eppendorf tubes.
- 6. Peptide fractions were analysed by MS and the data was processed** as described later in this chapter.

3.3. Affinity purification

3.3.1. Generation of epitope-tagged strains

The genomic sequences of Dbf2, Mob1, Cdc14 and Crk1 were retrieved from the Candida Genome Database (candidagenome.org). A DNA cassette containing either *MYC* or *HA* tag and either *URA3* or *ARG4* gene sequence as a selectable marker was amplified by PCR using primers containing flanking sequences homologous to a region in the 3' end of the gene to be tagged (fig. 3.1). Cells were transformed with the PCR product and individual colonies were screened by colony PCR and by Western blot.

Initially Dbf2 and Cdc14 were each consecutively tagged with HA-URA3 and HA-ARG4 to create Dbf2-2xHA and Cdc14-2xHA. In a similar manner Dbf2-2xMYC and Cdc14-2xMYC were created.

3.3.2. Selection of affinity matrix and tag

All optimisation experiments were carried out with yeast culture, because it is easier to handle than hyphae. Dbf2 and Cdc14 were each immunoprecipitated using either anti-HA-conjugated magnetic Dynabeads Protein G or anti-MYC-conjugated EZview Red Agarose Affinity Gel. Both affinity matrixes were tested under the same conditions. One litre of log phase culture ($OD_{595}=0.8$) was used to produce 10 ml of cell lysate, which was incubated with affinity beads for 1hr at 4°C. The beads were then washed once and boiled to elute all bound proteins. Comparison of protein yield by Western blot showed that the EZview beads have higher affinity for their antigen than Dynabeads (fig. 3.2). Additionally, only the EZview IP produced a visible band of the bait on a Coomassie stained gel, which is considered a good yield for MS analysis. Therefore, all subsequent IPs were done using EZview anti-MYC beads. Mob1-2xMYC and Crk1-2xMYC strains were generated as previously described.

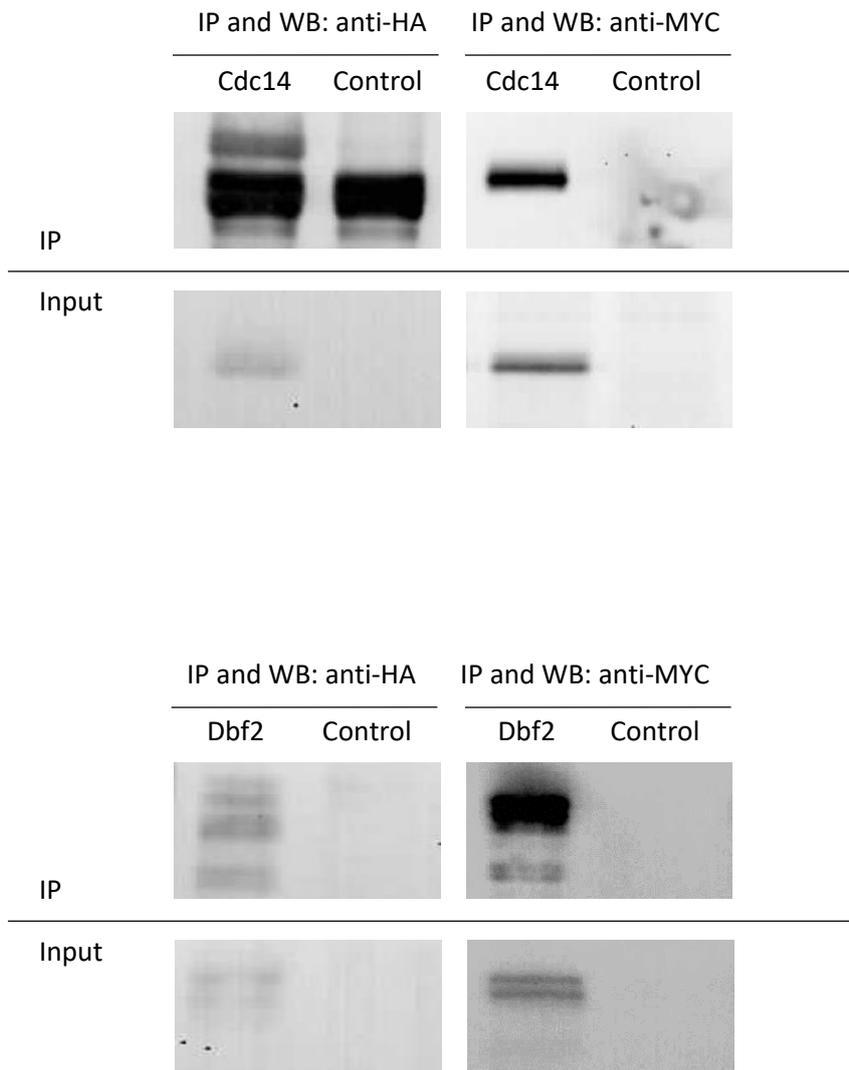


Fig. 3.2: Immunoprecipitation of Cdc14 and Dbf2 using either Dynabeads conjugated to anti-HA (left) or EZview beads conjugated to anti-MYC (right). Cdc14 is very close to the heavy chain of the anti-HA antibody, which produces very thick band on a Western blot (top, left panel). The anti-MYC antibody has fragments of different size, which do not migrate near the baits. Some protein degradation is seen during Dbf2 precipitation, which produces multiple bands on a Western blot. The EZview anti-MYC beads generated higher yield and lower background than Dynabeads. 2% of the IPs and 0.1% of the inputs were loaded on the gels for Western blot.

3.3.3. Optimisation of protein elution from beads

Proteins can be eluted from the beads either by heating at high temperature or by incubation with low pH glycine. The latter method has the potential of eluting less non-specific proteins bound to the beads. Both elution protocols were tested. Cdc14-2xMYC was precipitated as previously described and the protein-bound beads were incubated with 50 mM glycine pH 2.8 for 10 min at room temperature with shaking. The glycine was then removed and the beads were boiled to elute residual proteins. Western blot analysis of both elutions showed that most of the bait remained on the beads after the glycine elution (fig 3.3). Coomassie staining showed that the glycine elution had significantly lower background of contaminating proteins (data not shown). However, due to the low yield, glycine elution was not a viable option for an MS experiment. All subsequent IPs were therefore performed using heat elution.

3.3.4. Optimisation of protein concentration in a cell lysate for IP

The protein concentration of a cell lysate often correlates with the amount of background proteins sticking to the beads. In the initial IP, 1 L of cell culture produced 10 ml of cell lysate. This lysate has the highest concentration of proteins that could be achieved by breaking the cells in the available cell disrupter. Diluting the cell lysate would supposedly reduce the background, but it is only worth doing so if protein yield remains the same.

To test the effect of protein concentration of the cell lysate on total yield, Dbf2-2xMYC cells were lysed as previously described and 3 ml of cell lysate were diluted with 3 ml of lysis buffer. Dbf2 was precipitated from either 6 ml of diluted lysate or from 6 ml of concentrated lysate. The protein yield was assessed by Western blot and the background contamination was assessed by staining a gel with Coomassie (fig.3.4). The experiments confirmed that reducing the protein concentration produced visibly less background, but unfortunately, it also reduced significantly the yield of Dbf2. It would be difficult to objectively measure the optimal balance between bait yield and background. Therefore, the protein concentration was kept high in all further experiments, although the exact value was not measured.

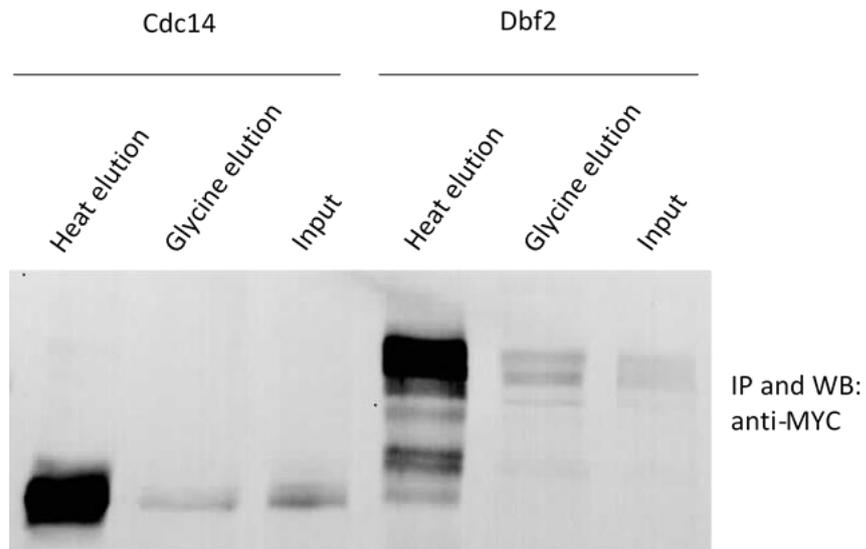


Fig. 3.3: Elution of protein from beads following immunoprecipitation. Cdc14 and Dbf2 were immunoprecipitated using EZview anti-MYC beads and the beads were incubated with 50 mM glycine pH 2.8 for 10 min at room temperature. The supernatant was then transferred to a fresh tube and the beads were resuspended in protein loading dye and boiled for 5 min. As shown here the glycine did not elute much of the bait proteins compared to the heat elution. 2% of the eluted IP volume and 0.1% of the inputs were loaded on the gels for Western blot.

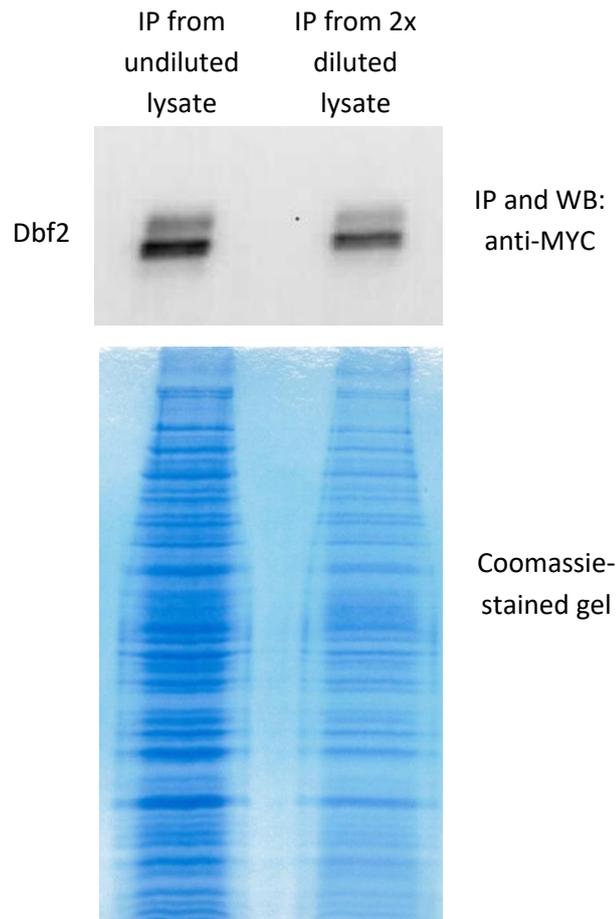


Fig. 3.4: Effect of protein concentration in cell lysate on IP output. Dbf2 was precipitated from undiluted and 2x diluted lysate. Western blot (top) shows that diluting the lysate reduces total protein yield although it also reduces background contamination as shown on the Coomassie-stained gel. 2% of the total IP volume was loaded on the gel for Western blot, and 80% of the IP was loaded on the Coomassie-stained gel.

3.3.5. Optimisation of total protein amount in a cell lysate for IP

The total amount of protein in a cell lysate correlates with the amount of the bait protein in the lysate. The amount of the bait should be sufficient to saturate the beads to their maximum capacity. Unsaturated beads would unnecessarily increase the background without proportional increase in bait yield.

In order to determine whether more cell lysate would improve protein yield, IP of Dbf2-2xMYC was carried out using either 5 ml or 15 ml of cell lysate. Both experiments produced the same yield of Dbf2 as well as the same amount of background as determined by Western blot and Coomassie staining respectively (fig 3.5). This means that 5 ml of lysate is sufficient to fully saturate the beads and more total protein does not result in increased background. It is also evident from figure 3.5 that the beads did not fully deplete the pool of Dbf2 available in the cell lysate. Significant amount of protein was still seen in the flowthrough after immunoprecipitation, although the exact recovery was not formally measured. In all further IP experiment 10 ml of cell lysate and 50 μ l of beads were used.

3.3.6. Pre-incubation of affinity beads with BSA does not reduce background

In an attempt to reduce non-specific proteins sticking to the affinity matrix, beads were incubated with 3% w/v BSA in lysis buffer for 4 hrs at 4 °C prior to immunoprecipitating Cdc14-2xMYC. The beads were then used as previously described. Coomassie staining showed no difference in background between blocked and non-blocked beads (fig 3.6). BSA blocking was not used in further experiments.

3.3.7. Optimisation of beads-washing steps after IP

In order to assess the impact of washing the beads with lysis buffer after IP on protein yield, Mob1-2xMYC was immunoprecipitated as previously described and the beads were separated into two tubes. One aliquot was washed once and the other was washed 3 times with 1 ml lysis buffer. Western blot of the eluted proteins showed that washing the beads 3 times reduced the yield of Mob1 approximately two times comparing to washing them once

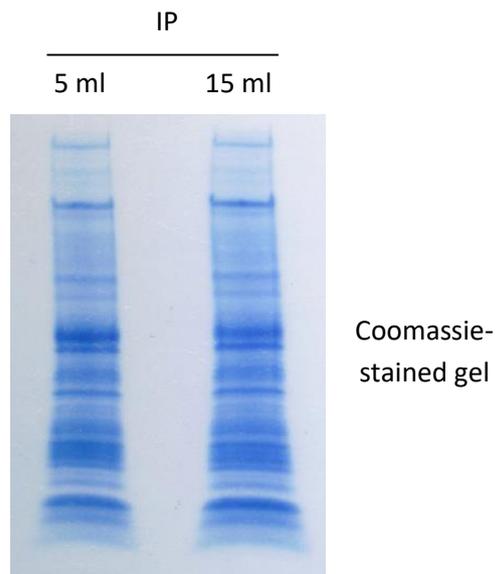
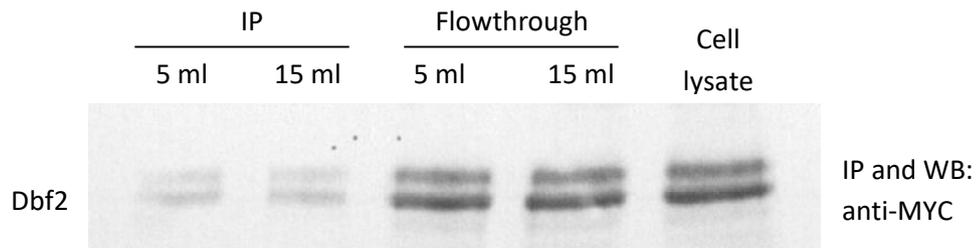


Fig. 3.5: Effect of total protein amount in cell lysate on IP output. Dbf2 was precipitated from either 5 ml or 15 ml cell lysate. Western blot (top) shows the same protein yield from both experiments and considerable amount of Dbf2 remain unbound in the flowthrough. This indicates that the affinity beads are saturated with the bait even with 5 ml of lysate. No difference was seen in the amount of contaminants on a Coomassie-stained gel (bottom). 2% of the eluted IP and 0.1% of the flowthrough and input volumes were loaded on the gel for Western blot. 80% of the IP was loaded on the Coomassie-stained gel.

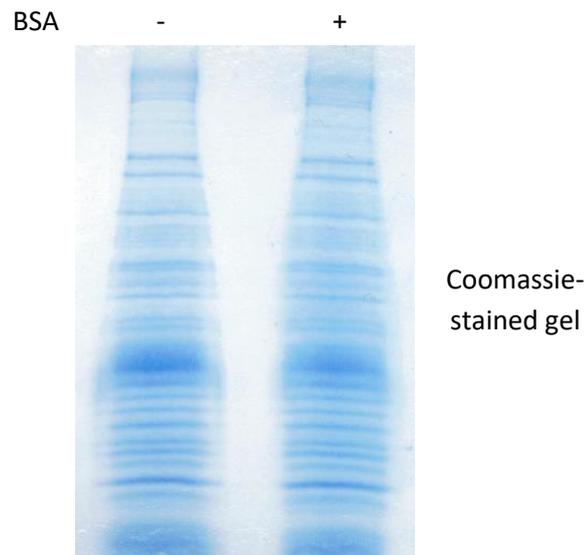


Fig. 3.6: Incubation of affinity beads with BSA prior to IP. Blocking the beads with BSA before the IP does not reduce the amount of contaminants sticking to the them as shown on this gel. 80% of the IP was loaded on the Coomassie-stained gel.

(fig. 3.7). This means that the antibody-antigen bond is not strong enough to overcome multiple washes. Although the background was also significantly lower in experiment with 3 washes, in all further experiments beads were washed only once.

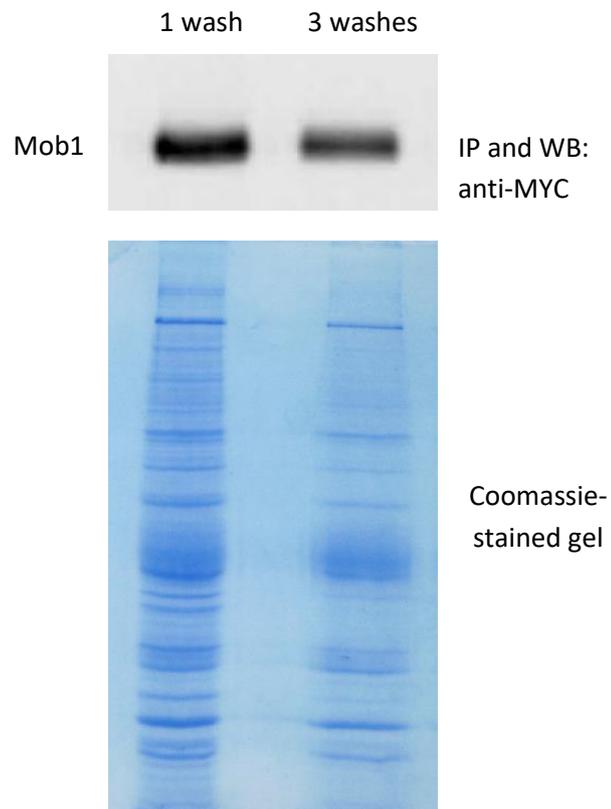


Fig. 3.7: Immunoprecipitation of Mob1 after one and three bead-washing steps. Mob1 yield and contamination were both reduced in IP with 3 washes compared to IP with only 1 wash. 2% of the IP volume was loaded on the gel for Western blot, 80% of it was loaded on the Coomassie-stained gel.

3.4. Mass spectrometry analysis

Once the affinity purification of all four proteins was optimised, mass spectrometry of the eluted fractions was performed to identify the proteins. The simplest way to do that is label-free MS since it does not require the additional step of protein labelling. However, as discussed later in this chapter, quantitative data analysis of label-free MS is challenging especially for the identification of protein interaction and a large amount of data is required in order to obtain statistically significant results.

3.4.1. Fractionation of eluted proteins followed by MS

Affinity purification of a single bait protein results in the elution of several hundred to several thousand other proteins in the mixture. Protein abundance varies highly and reflects the chances of a protein to be detected by MS. If the MS instrument is overwhelmed by the most abundant proteins, it is likely to miss some of the least abundant. Therefore, separating the eluted protein mixture into fractions is optimal for detecting the greatest number of proteins.

Following affinity purification, proteins were separated according to their size (also influenced by charge) by SDS-PAGE and the gel was stained with Coomassie. No proteins were allowed to run out of the gel. All protein bands were cut out of the gel into 15-20 fractions, so that each fraction contains proteins of similar size. Each fraction was individually digested with trypsin and the resulting peptides were analysed by ESI-MS using an amaZon ion trap instrument coupled to an online capillary liquid chromatography system.

Six IPs were analysed by MS in total: Dbf2, Mob1, Cdc14 and Crk1 were each used as bait separately in four of them and the remaining two were mock IPs of the wild type strain where no bait was present. The mock IP was duplicated in order to capture all contaminating protein. All experiments were done in yeast.

The results of each MS are summarised in table 3.1. Although great care was taken to keep the conditions of all experiments the same, there is a striking variability in the

IP bait	Dbf2	Mob1	Crk1	Cdc14	Control 1	Control 2
Total number of proteins identified by MS	1403	785	832	747	1155	710
FDR	1.39%	1.26%	1.28%	1.37%	1.57%	1.49%
Number of proteins after applying ProHits filters	648	388	345	313	514	325

Table 3.1: General results from label-free MS. Untagged wild type cells were used as controls. FDR – False discovery rate.

number of proteins identified from each IP. These proteins include bait-specific interactors as well as contaminants. In order to separate both groups all six lists of identified proteins were compared in parallel.

3.4.2. Analysis of MS results using ProHits software

Defining protein-protein interactions from label-free MS data is a challenging bid. Each set of data presents a list of proteins identified by MS but on its own this list does not entail any information about protein interactions within it. Comparison of all six data sets is likely to reveal such information in two ways. First, the mock IPs contain only non-specific proteins sticking to the beads and thus those proteins can be regarded as contaminants in all other experiments. Second, bait specific interactors would be missing from IPs using different bait. However, as previously discussed, Dbf2 and Mob1 are likely to interact with many common proteins and Cdc14 may also dephosphorylate some of Dbf2's substrates since they act in the same pathway in other species of yeast.

The software ProHits provides means of comparing label-free MS data sets with the aim of evaluating bait specific hits (Liu *et al.*, 2010). In particular, ProHits was used to compare the total number of peptides representing a given protein across all 6 experiments (table 3.2). In order to reduce false positive result, an arbitrary filter for protein score <50 and unique peptides <2 was applied. All baits were successfully recovered, but some of them do not stand out in the noise of contaminants. For example, Dbf2 was represented by only 11 peptides when used as a bait, and there were 284 proteins with more peptides from the same IP. Its activating partner, Mob1 was not found at all in this experiment. When Mob1 was the bait, it was represented by 33 peptides, ranking at position 93. Interestingly, Dbf2 ranked 14th in the Mob1 IP with 102 peptides. This means that more of the Dbf2 protein was pulled when Mob1 was the bait, rather than the kinase itself. Nevertheless, the overwhelming number of background proteins makes it difficult to identify true preys in any of the experiments. The most abundant contaminants include ribosomal proteins (e.g. RPL10, RPP0, RPL6, RPL3, RPL8B, RPS16A, RPL20B etc.), glycolytic enzymes (e.g. TDH3, PDC11, ENO1, CDC19, FBA1, PGI1 etc.), translation elongation factors (e.g. CEF3, TEF2,

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
P02768-1	P02768-1	216					
P00761	P00761	178		192	433	241	269
orf19.6814	TDH3	171	135	301	156	177	130
orf19.4152	CEF3	143	137	228	121	197	119
orf19.1065	SSA2	131	129	119	145	157	123
orf19.2877	PDC11	129	93	226	126	149	66
orf19.979	FAS1	127	135	209	91	284	101
orf19.382	TEF2	119	186	301	135	237	109
orf19.395	ENO1	112	116	270	155	159	104
orf19.5788	EFT2	99	76	153	96	131	105
orf19.6515	HSP90	93	101	133	121	136	101
orf19.7392	DED1	92	105	63	108	76	77
orf19.5949	FAS2	89	97	147	93	191	80
orf19.4980	HSP70	89	73	84	79	95	65
orf19.2803	HEM13	86	135	144	180	82	51
orf19.6367	SSB1	86	101	152	75	114	77
orf19.6312	RPS3	82	63	96	78	60	50
orf19.4477	CSH1	77	86	127	136	71	89
orf19.2360	URA2	76	54	105	80	300	117
orf19.3575	CDC19	72	98	158	105	96	88
orf19.2478.1		72	73	60	85	53	40
orf19.3523	CRK1	69					
orf19.6873	RPS8A	68	83	99	106	30	45
orf19.717	HSP60	67	49	73	53	69	43
orf19.2935	RPL10	66	70	54	51	46	45
orf19.7015	RPP0	63	72	56	56	41	38
orf19.3003.1	RPL6	60	47	84	65	46	53
orf19.1601	RPL3	57	52	59	56	40	43
orf19.6002	RPL8B	56	69	92	68	57	71
orf19.5653	ATP2	56	49	103	69	112	57
orf19.5341	RPS4A	55	58	92	60	68	59
orf19.2551	MET6	55	40	61	58	60	41
orf19.236	RPL9B	55	37	56	38	34	30
orf19.4618	FBA1	54	60	114	66	56	54
orf19.6854	ATP1	54	52	59	80	99	54
orf19.2994.1	RPS16A	53	55	68	48	28	40
orf19.4632	RPL20B	52	65	70	43	43	35
orf19.6265.1	RPS14B	50	45	27	23	20	33
orf19.4660	RPS6A	49	37	22	56	26	32
orf19.171	DBP2	48	50	39	39	49	28
orf19.6906	ASC1	48	28	57	43	55	27
orf19.7018	RPS18	46	53	59	41	32	31
orf19.3888	PGI1	46	24	55	49	47	24
orf19.930	PET9	45	42	86	47	58	39
orf19.2651	CAM1-1	45	20	53	41	31	33
orf19.2309.2	RPL2	44	45	66	43	34	37

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.1896	SSC1	44	36	45	45	56	34
orf19.5982	RPL18	44	33	61	46	31	43
orf19.1700	RPS7A	43	35	57	45	30	33
orf19.4371	TAL1	43	13	64	37	25	22
orf19.5904	RPL19A	42	54	63	25	42	44
orf19.3334	RPS21	42	46	57	49	25	28
orf19.2232	RPL11	42	41	31	23	20	42
orf19.4622		42	30	32	25	30	28
orf19.4336	RPS5	42	29	77	42	47	20
orf19.840	RPL21A	42	25	15	15	12	20
orf19.5197	APE2	42	21	38	46	30	33
orf19.493	RPL15A	41	35	41	36	36	47
orf19.4931.1	RPL14	40	51	49	38	44	36
orf19.2262		40	28	29	25	8	16
orf19.6385	ACO1	40	20	44	48	45	33
orf19.771	LPG20	39	42	42	60	22	13
orf19.903	GPM1	39	25	76	37	25	21
orf19.2013	KAR2	39	22	50	29	44	25
orf19.3465	RPL10A	38	45	56	29	27	39
orf19.2340	CDC48	38	42	54	63	19	28
orf19.838.1	RPS9B	37	31	47	36	26	38
orf19.5996.1	RPS19A	36	36	40	27	15	30
orf19.7382	CAM1	36	30	39	38	19	30
orf19.6375	RPS20	36	29	30	18	13	27
orf19.2435	MSI3	36	25	74	45	76	27
orf19.5466	RPS24	35	49	39	33	19	27
orf19.1378	SUP35	35	32	38	41	58	30
orf19.6975	YST1	34	42	56	53	45	54
orf19.3002	RPS1	34	40	54	46	42	34
orf19.1839	RPA190	34	23	31	6	28	35
orf19.3138	NOP1	34	20	27	29	18	17
orf19.3911	SAH1	34	11	45	28	26	27
orf19.7466	ACC1	33	60	49	75	76	44
orf19.1635	RPL12	33	43	40	13	26	9
orf19.827.1	RPL39	33	29	19	17	13	22
orf19.5024	GND1	33	22	80	40	45	25
orf19.4149.1		32	33	21	30	13	32
orf19.4193.1	RPS13	31	40	24	25	18	26
orf19.1854	HHF22	31	32	14	23	12	22
orf19.3690.2		31	26	26	30	11	17
orf19.5225.2	RPL27A	30	53	47	17	25	41
orf19.1064	ACS2	30	34	43	43	73	33
orf19.4885	MIR1	30	23	40	40	36	21
orf19.4602	MDH1-1	30	17	45	20	29	28
orf19.7238	NPL3	29	47	31	35	26	45
orf19.2364	MIS11	28	30	58	52	25	25

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.2994	RPL13	28	26	36	40	28	36
orf19.2111.2	RPL38	28	26		12	5	10
orf19.51		28	19	63	38	59	28
orf19.3788.1	RPL30	28	14	33	11	21	12
orf19.3812	SSZ1	28	10	24	15	30	9
orf19.6540	PFK2	27	48	67	60	38	19
orf19.542	HXK2	27	23	52	40	50	12
orf19.7217	RPL4B	26	96	103	76	51	79
orf19.3149	LSP1	26	26	39	29	21	19
orf19.6265	RPS22A	26	20	28	15	19	20
orf19.4393	CIT1	26	7	45	28	41	29
orf19.6541	RPL5	25	34	58	20	35	27
orf19.6561	LAT1	25	18	18	20	32	21
orf19.5112	TKL1	25	7	20	23	56	12
orf19.3311	IFD3	24	33	22	38	12	22
orf19.2864.1	RPL28	24	33	19	29	14	18
orf19.3037		24	25	30	41	64	32
orf19.3415.1	RPL32	24	25	15	17	7	19
orf19.7569	SIK1	24	24	42	27	31	12
orf19.5964.2	RPL35	24	19	13	17	7	16
orf19.2489		24	16	54	32	51	32
orf19.7332	ELF1	24	13	42	32	40	16
orf19.3572.3		23	33	23	24	26	30
orf19.6286.2	RPS27	23	19		9		14
orf19.3504	RPL23A	23	17	9	8	13	12
orf19.778	PIL1	23	15	45	18	10	18
orf19.2560	CDC60	23	12	56	24	38	26
orf19.2709	ZUO1	23	9	21	16	25	10
orf19.3325.3	RPS21B	23	8		6	8	4
orf19.6085	RPL16A	22	35	34	28	21	36
orf19.7417	TSA1	22	34	48	23	33	30
orf19.5943.1		22	15	21	12	20	7
orf19.2179.2	RPS10	22	14	8	9	14	18
orf19.1288	FOX2	22			5		
orf19.687.1	RPL25	21	26	38	13	16	23
orf19.4490	RPL17B	21	16	19	12	12	24
orf19.6702	DED81	21	15	23	15	22	12
orf19.6127	LPD1	20	22	33	29	36	14
orf19.6663	RPS25B	20	22	7	17	13	9
orf19.3789	RPL24A	20	21	16	22	14	17
orf19.5927	RPS15	20	20	31	17	19	13
orf19.6472	CYP1	20	15	25	16	14	18
orf19.1578		20	15	4	5	15	8
orf19.4635	NIP1	20	13	37	45	24	16
orf19.4506	LYS22	20		33	24	30	9
orf19.5746	ALA1	19	13	39	38	26	22

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.5858	EGD2	19	11	24	31	22	22
orf19.6403.1	RPP2A	19	8	12	12	11	7
orf19.1652	POX1-3	19		3			
orf19.4311	YNK1	18	21	22	14	21	22
orf19.5383	PMA1	18	18	43	47	131	28
orf19.2329.1	RPS17B	18	18	18	8	12	7
orf19.6090		18	16	10	15	17	17
orf19.18	IMH3	18	15	23	9	26	12
orf19.3541	ERF1	18	11	38	13	43	14
orf19.6701		18	11	45	11	39	9
orf19.6584	PRT1	18	10	15	31	33	7
orf19.3942.1	RPL43A	18	10	7	12	11	7
orf19.7239	MDG1	18	7	22	8	23	9
orf19.6749	KRS1	17	13	24	19	41	16
orf19.3997	ADH1	16	63	156	79	65	58
orf19.1199	NOP5	16	27	31	17	37	14
orf19.1154	EGD1	16	12	13	6	12	11
orf19.4560	BFR1	16	6	16	10	12	6
orf19.3014	BMH1	15	25	36	23	32	12
orf19.5007	ACT1	15	23	29	24	8	14
orf19.7534	MIS12	15	13	17	22	17	9
orf19.6745	TPI1	15	11	32	18	15	11
orf19.6160		15	8	11	13	10	
orf19.3010.1	ECM33	15	6	31		30	
orf19.6925	HTB1	14	18	25	17	10	16
orf19.6665		14	14	11	18	8	9
orf19.1750	SLR1	14	8	6	10	8	9
orf19.2138	ILS1	13	17	60	28	53	24
orf19.6253	RPS23A	13	17	6	7	5	5
orf19.5294	PDB1	13	16	22	16	12	8
orf19.1470	RPS26A	13	16	13	9	6	
orf19.1295	VAS1	13	14	49	23	40	14
orf19.946	MET14	13	13	21	12	6	7
orf19.2407	DPS1-1	13	12	16	16	28	12
orf19.4623.3	NHP6A	13	12	5		6	10
orf19.198	ASN1	13	8	31	15	30	9
orf19.5806	ALD5	13	3	20	7	9	7
orf19.3967	PFK1	12	40	69	75	35	15
orf19.2183	KRE30	12	13	16	20	27	6
orf19.1833		12	12	18	20	12	7
orf19.1042	POR1	12	12	20	15	13	5
orf19.5493	GSP1	12	11	26	6	11	7
orf19.1738	UGP1	12	6	17	10	12	4
orf19.4718	TRP5	12	6	10	4	20	4
orf19.4223	GCD11	12	5	11	12	12	5
orf19.269	SES1	12	4	28	16	21	9

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.4427	SKP1	12		14	11	10	8
orf19.7188	RPP1B	12		14		7	
orf19.4284	BUR2	12					
orf19.6987	DNM1	11	28		38	2	18
orf19.5928	RPP2B	11	14	13	9	9	8
orf19.6165	KGD1	11	11	19	23	8	18
orf19.5137.1	HHO1	11	11		6		9
orf19.5081	FUN12	11	2	31	20	22	5
orf19.5281		11		50	29	30	17
orf19.3426	ANB1	11		10		12	6
orf19.1051	HTA2	11		5		6	
orf19.3324	TIF	10	30	60	35	36	25
orf19.6763	SLK19	10	22		5	6	5
orf19.1853	HHT2	10	17	4	4	3	7
orf19.5177		10	13	19	15	20	8
orf19.6785	RPS12	10	9	19	10	9	12
orf19.3599	TIF4631	10	7	24	10	19	9
orf19.6213	SUI2	10	6	18	10	13	4
orf19.2960	FRS2	10	3	15	14	8	7
orf19.5641	CAR2	10		13	14	4	
orf19.2573	FRS1	10		18	9	13	5
orf19.1149	MRF1	9	54	46	48	25	23
orf19.5682		9	15	25	9	23	14
orf19.492	ADE17	9	14	36	12	29	10
orf19.6257	GLT1	9	11	40	16	49	10
orf19.6882.1		9	11	6	8	6	12
orf19.6387	HSP104	9	7	17	7		
orf19.7161	SUI3	9	5	8	7	3	9
orf19.3358	LSC1	9	4	11	3	6	
orf19.4777	DAK2	9		34	12	15	
orf19.3428		9					
orf19.5779	RNR1	8	19	26	28	11	13
orf19.4716	GDH3	8	9	49	32	41	18
orf19.6345	RPG1A	8	9	19	14	5	20
orf19.2929	GSC1	8	9	42	5	30	3
orf19.7048.1	RPS28B	8	9		3		7
orf19.1409.1		8	7	11	4	8	8
orf19.3591	APE3	8	5	8	7	14	8
orf19.2843	RHO1	8	5	10		5	5
orf19.437	GRS1	8	4	11	12	12	4
orf19.759	SEC21	8	4	18	10		3
orf19.3423	TIF3	8	4	14	7	10	5
orf19.7312	ERG13	8	2	14	10	11	7
orf19.3349		8		30	15	13	10
orf19.5801	RNR21	8		19	11	7	4
orf19.5263	SER33	8		34	10	19	6

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.5964	ARF2	8		21	10	15	4
orf19.1030		8		5	9	14	
orf19.2098	ARO8	8		17	7	20	4
orf19.2992	RPP1A	8		12	5	15	
orf19.4704	ARO1	7	11	52	19	37	10
orf19.2871	SDH12	7	10	18	16	31	13
orf19.3391	ADK1	7	10	32	11	19	19
orf19.3590	IPP1	7	9	23	7	16	5
orf19.3838	EFB1	7	8	22	19	14	6
orf19.2884	CDC68	7	8				4
orf19.5437	RHR2	7	4	17	13	11	
orf19.5130	PDI1	7	4	20	10	21	6
orf19.7308	TUB1	7	4	10	10	7	
orf19.4813	GUA1	7	4	13	8	22	7
orf19.2937	PMM1	7	4	28	7	15	6
orf19.5854	SBP1	7		24	15	11	7
orf19.1613	ILV2	7		18	7	17	3
orf19.6034	TUB2	7		13	7	14	2
orf19.6109	TUP1	7		21	6	13	6
orf19.6507		7		13	5	6	
orf19.822	HSP21	7					
orf19.3097	PDA1	6	10	16	17	8	5
orf19.3579	ATP4	6	9	19	10	17	18
orf19.5750	SHM2	6	8	14	14	25	7
orf19.3268	TMA19	6	7	15	5	11	
orf19.7062	RPA135	6	7	12	4	20	4
orf19.754	YBN5	6	6	23	9	12	9
orf19.1672		6	5	15	7	9	11
orf19.5550	MRT4	6	5	7	5	3	4
orf19.2168.3		6	5	4		3	4
orf19.6634	VMA2	6	2	17	10	16	
orf19.6081	PHR2	6	2	14	6		5
orf19.4879.2	NTF2	6	2				
orf19.387	GCR3	6			10		
orf19.4309	GRP2	6		24	9	4	
orf19.4956	RPN1	6		15	6	13	9
orf19.3015	ARX1	6		11	6	19	
orf19.3915		6		10		2	
orf19.3034	RLI1	6		9		13	3
orf19.2416.1	MLC1	6		4		2	3
orf19.4317	GRE3	6		2		5	
orf19.3087	UBI3	5	9	7	6	3	
orf19.5885		5	6		5	5	
orf19.339	NDE1	5	6	10		4	
orf19.685	YHM1	5	5	6	12	7	
orf19.7509.1	ATP17	5	4		4		

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.6717		5	4	5		4	8
orf19.7611	TRX1	5	4				6
orf19.3700	TOM70	5	3	8	2	16	
orf19.6197	DHH1	5	3	11		8	4
orf19.7057	GUS1	5		19	9	18	8
orf19.4261	TIF5	5		13	7	6	
orf19.92		5		16	6	10	
orf19.6126	KGD2	5		10	5	10	
orf19.7136	SPT6	5		9	4		
orf19.3430		5		6	4		
orf19.4536	CYS4	5		7	3	7	
orf19.3129		5			3	3	
orf19.7421	CYP5	5		9		7	
orf19.6645	HMO1	5		8		6	
orf19.5917.3		5		6		4	
orf19.5351	TIF11	5		6			3
orf19.6229	CAT1	5		4			
orf19.6010.1	RPB11	5		3		3	4
orf19.4021		5		2			
orf19.2241	PST1	5					
orf19.2288	CCT5	5					
orf19.4833	MLS1	5					
orf19.789	PYC2	4	9	48	31	70	19
orf19.3496	CHC1	4	8	36	18	26	5
orf19.3955	MES1	4	5	21	10	15	4
orf19.4099	ECM17	4	5	41	9	23	6
orf19.1770	CYC1	4	4	4		4	
orf19.4959		4	4				5
orf19.3171	ACH1	4	3	11	3	7	4
orf19.7011		4		25	9	17	
orf19.3799		4		10	8	10	
orf19.6724	FUM12	4		11	7	11	
orf19.5285	PST3	4		15	5	8	10
orf19.3507	MCR1	4		7	4	7	6
orf19.2066.1	ATP18	4			4		
orf19.6327	HET1	4		12		9	2
orf19.2014	BCY1	4		11		9	
orf19.3974	PUT2	4		11		6	
orf19.2694	TYS1	4		7		5	
orf19.3294	MBF1	4		5			7
orf19.7384	NOG1	4		3		6	
orf19.3150	GRE2	4		2			
orf19.4577.3	TIM10	4					
orf19.5597.1		4					2
orf19.3651	PGK1	3	62	128	65	49	48
orf19.953.1	COF1	3	7	13		6	

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.526	NHP2	3	4	4	5		5
orf19.6994	BAT22	3		22	13	5	
orf19.7438	UBA1	3		26	12	16	
orf19.6317	ADE6	3		47	5	25	14
orf19.4754	ZWF1	3		11	4		2
orf19.3959	SSD1	3		6	3		
orf19.5627		3		5	2	5	
orf19.2531	CSP37	3		4	2	2	
orf19.5006	GCV3	3		11			3
orf19.2598	VMA4	3		10		6	
orf19.7187	MAM33	3		7			
orf19.4650	ILV6	3		6		4	
orf19.1402	CCT2	3		4			
orf19.4870	DBP3	3				6	3
orf19.873.1	COX6	3				5	
orf19.5198	NOP4	3					
orf19.7459		3					
orf19.5850	NOC2	2	3	2		15	5
orf19.3467	SEC27	2		11	12	2	3
orf19.3053		2		13		5	
orf19.7592	FAA4	2		6		5	
orf19.5968	RDI1	2					
orf19.7288		2					
280717	ALB		54	5	52		15
orf19.1048	IFD6		42	58	48	26	26
orf19.4476			35	50	34	24	
orf19.1880	HEM15		34	42	45	30	30
orf19.4192	CDC14		32				
orf19.6190	SRB1		31	50	30	12	18
3848	KRT1		21	121	76	45	61
orf19.4826	IDH1		18	32	21	15	23
orf19.657	SAM2		12	21	10	2	6
orf19.88	ILV5		11	38	17	21	10
orf19.6047	TUF1		11	16	11	4	8
orf19.1591	ERG10		10	24	10	7	11
orf19.3223	ATP3		9	4	6	5	
orf19.4827	ADE12		8	8	7		6
orf19.6415.1			8		6		8
orf19.3356	ESP1		8				
orf19.5260	RPN2		7	17	11	3	2
orf19.385	GCV2		7	9	10	4	3
orf19.1986	ARO2		7	12	8	6	5
orf19.646	GLN1		7	30	3	10	12
16687	Krt6a		7				
orf19.1860	LSC2		6	20	7	7	7
orf19.5791	IDH2		6	25	5	13	12

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.327	HTA3		6			5	
3858	KRT10		5	73	36	12	8
orf19.6250			5	3	9	11	5
orf19.3442			5		4		3
3868	KRT16		5				
orf19.2310.1	RPL29		5				
orf19.5328	GCN1		4	4	22	3	3
3857	KRT9		4		9		17
orf19.2352			4		5		
orf19.5015	MYO2		4	22	3	4	
orf19.1254	SEC23		4	10	3	4	
orf19.2119	NDT80		4		3		
orf19.2507	ARP9		4	9		13	
orf19.2644	QCR2		4	9		5	6
3872	KRT17		4				
orf19.4437			4				
orf19.5685	THS1		3	16	6	21	4
orf19.4060	ARO4		3	16	4	4	
orf19.5591	ADO1		3	10	4	6	4
orf19.3126	CCT6		3	14		2	
orf19.5450	ETR1		3	9		10	
orf19.5793	PR26		3	9		8	
orf19.5025	MET3		3	9		2	
orf19.2511.1	MRPL33		3				
orf19.2873	TOP2		3				
orf19.3276	PWP2		3				
orf19.4831	MTS1		2	12	9	5	5
orf19.3527	CYT1		2	6	8	7	
orf19.2422	ARC1		2	8	5	4	4
orf19.2967	TIF34		2	10		5	
orf19.691	GPD2		2	8			
orf19.2917			2	7		5	
orf19.2533.1			2	4			
orf19.2699	ABP1		2	2		5	4
orf19.2017			2			6	
orf19.6447	ARF1		2				
orf19.1223	DBF2			11	102		
orf19.5528	MOB1				33		
3849	KRT2			87	23	12	17
orf19.4732	SEC24			2	14		
orf19.1342	SHM1			10	11	8	4
orf19.528	SEC26			6	11	4	
orf19.3870	ADE13			13	10	22	5
orf19.1680	TFP1			11	9	20	7
orf19.1618	GFA1			8	8	9	
orf19.7626	EIF4E			19	7	9	5

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.6539				12	7	5	
orf19.316	SEC13			4	7	6	
orf19.518				2	7	7	
orf19.7178	PRE5				7	3	
orf19.1047	ERB1			9	6	11	
orf19.506	YDJ1			9	6	3	
orf19.2250	SPE3			8	6	4	
orf19.5639	HIS4			8	6		
orf19.3335				6	6	4	2
orf19.1661	DBP5			2	6	6	
orf19.7424	NSA2				6		
orf19.6217	PGA63			14	5	6	5
orf19.6882	OSM1			8	5		
orf19.2275				5	5	3	
orf19.1756	GPD1				5	4	
orf19.3278	GSY1				5	4	
orf19.585					5		
orf19.6402	CYS3			19	4	3	
orf19.2525	LYS12			12	4	5	
orf19.922	ERG11			9	4	5	
orf19.5645	MET15			9	4		
orf19.406	ERG1			8	4		
orf19.1559	HOM2			7	4	7	
orf19.2785	ATP7			7	4	2	8
orf19.1789.1	LYS1			6	4		
orf19.338				5	4	4	
orf19.4931				3	4	10	
orf19.1569	UTP22			3	4	3	
orf19.5991				2	4		
orf19.3681					4	7	3
orf19.4375.1	RPS30				4		
orf19.4413	CMD1				4		
orf19.5607					4		
orf19.6696	TIM9				4		
orf19.6220.3	MMD1			15	3	7	
orf19.7327	PHO88			10	3	2	
orf19.3941	URA7			8	3	12	
orf19.5893	RIP1			7	3	3	
orf19.2720				6	3	6	
orf19.6757	GCY1			6	3		
orf19.5061	ADE5,7			5	3	7	
orf19.7483	CRM1			4	3		
orf19.1233	ADE4				3	3	
orf19.1966	BUD23				3		
orf19.2500					3		
orf19.522	PIM1				3		

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.96	TOP1				3		
orf19.734	GLK1			8	2		
orf19.5083	DRG1			7	2	4	
orf19.3348				5	2		
orf19.6524	TOM40			5	2		
orf19.441	RPT1			3	2	10	
orf19.5420				3	2		
orf19.5505	HIS7			3	2		
orf19.7064	GLN4				2	4	
orf19.3367					2		
orf19.7328					2		
orf19.125	EBP1			41		46	
orf19.5776	TOM1			30		15	
orf19.744	GDB1			26		22	
orf19.1517	ARO3			21		5	4
orf19.2947	SNZ1			17		12	3
orf19.3462	SAR1			17		9	
orf19.1552	CPR3			17		3	
orf19.3341				15		23	
orf19.637	SDH2			15		10	
orf19.6099	CCT8			14		9	
orf19.1631	ERG6			14		2	
orf19.4040	ILV3			13		13	
orf19.2023	HGT7			13			
orf19.2951	HOM6			13			9
orf19.5180	PRX1			12		6	
orf19.2852				12		5	4
orf19.3554	AAT1			12		4	
orf19.512				11		6	
orf19.544.1	PRE6			11		5	2
orf19.2762	AHP1			11		2	
orf19.4517				11			3
orf19.1336	PUP3			10		6	
orf19.3013	CDC12			10		6	
orf19.850				10		6	3
orf19.251	GLX3			10		4	
orf19.5073	DPM1			10		4	
orf19.1815				10		3	
orf19.1375	LEU42			10			
orf19.1448	APT1			10			
orf19.505	SRV2			10			
orf19.4969	KEM1			9		7	
orf19.3064	MRPL27			9		6	
orf19.4016				9		6	
orf19.4051	HTS1			9		6	
orf19.847	YIM1			9		2	

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.1868	RNR22			9			
orf19.2396	IFR2			9			
orf19.3221	CPA2			9			
orf19.401	TCP1			9			
orf19.5956	PIN3			9			
orf19.7297				9			
orf19.7481	MDH1			9			
orf19.610	EFG1			8		8	
orf19.3206	CCT7			8		6	3
orf19.4076	MET10			8		5	
orf19.5773				8		4	
orf19.5378	SCL1			8		3	
orf19.3168	RPN8			8			
orf19.4491	ERG20			8			
orf19.7269				8			5
orf19.7600	FDH3			8			
orf19.7676	XYL2			8			
orf19.1086				7		6	3
orf19.4759	COX5			7		6	
orf19.2930				7		5	
orf19.6041	RPO41			7		5	8
orf19.4233	THR4			7		4	
orf19.5085				7		3	
orf19.941	SEC14			7		3	
orf19.2640	FUR1			7		2	4
orf19.548	CDC10			7		2	
orf19.2483	RIM1			7			
orf19.3696	TOM22			7			5
orf19.645.1	VMA13			7			
orf19.797	BAT21			7			
orf19.3192	STI1			6		13	2
orf19.7655	RPO21			6		11	5
orf19.680	TIM50			6		8	
orf19.7448	LYS9			6		7	
orf19.300	AIP2			6		6	
orf19.2555	URA5			6		5	
orf19.5021	PDX1			6		5	
orf19.5484	SER1			6		5	
orf19.997	SNL1			6		4	
orf19.1967				6		3	
orf19.1840				6			
orf19.213				6			
orf19.2244				6			
orf19.239				6			
orf19.3103				6			
orf19.424	TRP99			6			

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.4290	TRR1			6			
orf19.4697	MDN1			6			
orf19.5211	IDP1			6			
orf19.6285	GLC7			6			
orf19.828				6			
orf19.4246				5		6	
orf19.3054	RPN3			5		5	
orf19.3052	YPT1			5		4	
orf19.4609				5		4	
orf19.1164	GAR1			5		3	
orf19.2841	PGM2			5		3	
orf19.3350	MRP20			5		3	
orf19.4248				5		3	
orf19.550	PDX3			5		3	
orf19.1390	PMI1			5		2	
orf19.238	CCP1			5		2	
orf19.4909.1	RPL42			5		2	
orf19.5912	MAK21			5		2	
orf19.1115	GUK1			5			
orf19.1229				5			
orf19.1354	UCF1			5			
orf19.1946				5			
orf19.2571	SEC4			5			
orf19.390	CDC42			5			
orf19.4382				5			
orf19.4492				5			
orf19.482	RPT4			5			
orf19.5235				5			
orf19.5480	ILV1			5			
orf19.5525				5			
orf19.5620				5			
orf19.6014	RRS1			5			
orf19.7124	RVS161			5			
orf19.895	HOG1			5			
orf19.978	BDF1			5			
orf19.6632	ACO2			4		13	
orf19.5419	ATP5			4		4	
orf19.6559				4		4	
orf19.7215				4		4	
orf19.763				4		4	
orf19.1299	RPN6			4		3	
orf19.1687				4		3	
orf19.2233	PRE2			4		3	
orf19.4490.2	QCR8			4		3	
orf19.7635	DRS1			4		3	
orf19.1575	PRS1			4		2	

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.2795	LHP1			4		2	
orf19.3300	ZPR1			4		2	
orf19.1153	GAD1			4			
orf19.1340				4			
orf19.1553	ENT3			4			
orf19.1649	RNA1			4			
orf19.1665	MNT1			4			
orf19.1691				4			
orf19.2093	RFA1			4			
orf19.2549	SHP1			4			
orf19.2895	VMA8			4			
orf19.3038	TPS2			4			
orf19.3106	MET16			4			
orf19.3340	SOD2			4			
orf19.3478	NIP7			4			
orf19.3707	YHB1			4			
orf19.4591	CAT2			4			3
orf19.4796				4			
orf19.4848	SKI3			4			
orf19.5228	RIB3			4			
orf19.5369				4			
orf19.5517				4			
orf19.5622	GLC3			4			
orf19.5832	HPT1			4			
orf19.6151	ARC15			4			
orf19.6176	SEC61			4			
orf19.6729	TIP120			4			
orf19.6809				4			
orf19.7019	YML6			4			
orf19.7021	GPH1			4			
orf19.7322				4			
orf19.798	TAF14			4			
orf19.886	PAN1			4			
orf19.989				4			
orf19.1453	SPT5			3		5	
orf19.5647	SUB2			3		5	
orf19.7215.3				3		5	
orf19.1055	CDC3			3		4	
orf19.1214				3		4	
orf19.6810				3		4	
orf19.4236	RET2			3		3	
orf19.5440	RPT2			3		3	
orf19.6136				3		3	
orf19.6837	FMA1			3		3	
orf19.5100	MLT1			3		2	
orf19.6417	TSR1			3		2	

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.1085				3			
orf19.1394				3			
orf19.1630				3			
orf19.1646				3			
orf19.1891	Apr-01			3			
orf19.2283	DQD1			3			
orf19.3003				3			
orf19.3123	RPT5			3			
orf19.3251	ARC19			3			
orf19.3297				3			
orf19.3322	DUT1			3			
orf19.3846	LYS4			3			
orf19.3962	HAS1			3			
orf19.4204				3			
orf19.4451	RIA1			3			
orf19.4640	PWP1			3			
orf19.4751				3			
orf19.5078				3			
orf19.5126				3			
orf19.5178	ERG5			3			
orf19.5230	MRPS9			3			
orf19.5293				3			
orf19.5698				3			
orf19.581				3			
orf19.5834				3			
orf19.5870	CTP1			3			
orf19.6293	EMP24			3			
orf19.6322	ARD			3			
orf19.6503				3			
orf19.6798	SSN6			3			
orf19.6804				3			
orf19.688				3			
orf19.6948	CCC1			3			
orf19.7264				3			
orf19.7335	PRE8			3			
orf19.7409	ERV25			3			
orf19.7613	HCR1			3			
orf19.7654	CPR6			3			
orf19.809				3			
orf19.863				3			
orf19.1235	HOM3			2		4	
orf19.3133	GUT2			2		3	
orf19.2755				2		2	
orf19.7236	TIF35			2		2	3
orf19.976	BRE1			2		2	
orf19.1108	HAM1			2			

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.1166	CTA3			2			
orf19.1628	LAP41			2			
orf19.1662				2			
orf19.2214				2			
orf19.3938				2			
orf19.4024	RIB5			2			
orf19.4032	RPN5			2			
orf19.4230				2			
orf19.4669	AAT22			2			
orf19.4686				2			
orf19.4898				2			
orf19.5104	LTP1			2			
orf19.5226	WRS1			2			
orf19.5597	POL5			2			
orf19.5629	QCR7			2			
orf19.5747				2			
orf19.5958	CDR2			2			
orf19.6236	NOP6			2			
orf19.6264.3				2			
orf19.667.1	RPL37B			2			
orf19.6752				2			
orf19.6844	ICL1			2			
orf19.7035	RFC2			2			
orf19.7086				2			
orf19.7153				2			
orf19.7261	GDI1			2			
orf19.810				2			
orf19.882	HSP78			2			
orf19.7076	GBP2					8	3
orf19.231	APL2					5	
orf19.1949	VPS1					4	4
orf19.2150						4	
orf19.4683	MLP1					4	
orf19.5516						4	
orf19.5544	SAC6					4	
orf19.6812	PMT2					4	
orf19.7201	SLA2					4	
orf19.7501						4	
orf19.7678	ATP16					4	
orf19.1494	RAD23					3	
orf19.3547						3	
orf19.3755						3	
orf19.4089	SGT1					3	
orf19.4594	CLC1					3	
orf19.6063	UBP6					3	
orf19.2028	MXR1					2	

Gene ID	Gene Name	Crk1	Cdc14	Dbf2	Mob1	Control 1	Control 2
orf19.2095						2	
orf19.2286						2	
orf19.2601	HEM1					2	
orf19.2672	NCP1					2	
orf19.2688	NAN1					2	
orf19.3205	MPRL36					2	
orf19.3333						2	
orf19.3480						2	
orf19.4093	PES1					2	
orf19.4102	RPN10					2	
orf19.4147	GLR1					2	
orf19.498						2	
orf19.5989						2	
orf19.6582	PRE10					2	
orf19.6612						2	
orf19.6967	USO6					2	
orf19.7081	SPL1					2	
orf19.7234						2	
orf19.7552						2	
orf19.6924	HTA1						7

Table 3.2: Proteins identified by MS. Six IP-MS experiments were carried out where the bait protein was either Crk1, Cdc14, Dbf2, Mob1 or none. In the control experiments untagged cells were used in the same manner as the other four tagged strains. The results of the MS experiments were processed using the software ProHits, where this table was exported from. The Gene ID is a unique number assigned to each gene in the Candida Genome Database. If the genes have been characterised, they also have names corresponding to the protein names. Blank spaces in the second column indicate that the genes are uncharacterised. The numbers in the remaining six columns indicate the number of total peptides of the corresponding protein, that have been identified in each experiment. Results are filtered, so that only proteins with at least 2 unique peptides are present, and only peptides with a score >50 are counted. Cell colours indicate peptide abundance in a decreasing order from red to blue.

CAM1-1, CAM1, SUP35 etc.) and chaperones (e.g. SSA2, HSP90, HSP70, KAR2, SSZ1, HSP104 etc.).

3.4.3. SAINT analysis of MS results

The Significance Analysis of INTERactome (SAINT) is another computational method to assess interaction probabilities in a set of MS data (Choi, *et al.*, 2011). SAINT is an integrated tool of ProHits and provides a more sophisticated analysis for quantifying the probability of an interaction between two proteins by using spectral counts normalised to the length of the proteins and to the total number of spectra in the purification.

Unfortunately, after the MS data was processed by SAINT, the software was not able to find any interactions, because it would require more data sets to generate statistically significant scores.

3.5. Discussion

Affinity purification coupled to mass spectrometry is the main tool of this study for identifying interacting partners of Dbf2, Mob1, Cdc14 and Crk1. All four proteins were immunoprecipitated using agarose beads coupled to antibodies against a MYC tag fused to the proteins. This combination of a tag and affinity matrix was found to capture the most amount of bait, while also producing less background than the other tested variations. A series of optimisation experiments were used to derive the optimal conditions for maximal yield of bait proteins, which is a major limiting factor in the pipeline. Although a large amount of purified bait proteins was achieved, Coomassie-stained gels showed an overwhelming background of contaminants which likely obstruct the identification of low abundant specific interactors. Experimental methods that reduce the non-specific binding of proteins to the affinity matrix, inevitably compromise bait recovery and are likely to disturb protein interactions. Increasing signal-to-noise ratio was a major goal in following experiments.

Label-free MS can be a powerful tool for identifying unknown proteins in a mixture but it is likely not the easiest way to map interactions. Previously Breitkreutz, *et al.* (2011) have used label-free MS followed by SAINT to construct a global kinase and phosphatase interaction network in *S. cerevisiae*. Their success was in part due to the fact that bait proteins were very strongly overexpressed in order to force a maximum number of *in vivo* interactions. Another hallmark of their study is the use of 276 different baits, which allowed them to devise a sophisticated statistical method for analysis of the MS results. This study attempted to use their approach on a smaller scale in *C. albicans*. However, due to insufficient data size no statistical significance was reached by SAINT.

Although ProHits provides a rather loose method for identifying interactions, very few proteins stand out as likely hits. Proteins present in all six experiments can be excluded as interactors with high confidence, but proteins present in a single IP are very few. An example of such protein is Bur2, which was represented by 12 peptides in the Crk1 IP, but it was not found in any of the other IPs. Bur2 is also a known cyclin of Bur1 (the Crk1

homologue in *S. cerevisiae*) and it was even identified as a prey in the Breitzkreutz study. Thus Bur2 is most likely to be a real hit, which raises the possibility that Crk1 is a cyclin-dependent kinase, although further experiments would be needed to confirm it.

The vast majority of proteins are found in 2-5 IPs. Some of those proteins have very different representation in each experiment. For example, Sec27 has 12 peptides in Mob1 IP, 11 peptides in Dbf2 IP, 3 peptides in one of the mock IPs, and 2 peptides in the other mock IP and Crk1 IP. As previously discussed, Dbf2 and Mob1 are likely to pull the same interactors. Sec27 is visibly overrepresented in these 2 IPs comparing to the rest, but the presence of a few peptides in the other three IPs creates the possibility that it may be a contaminant. In fact, Sec27 was identified as interactor of Dbf2 in *S. cerevisiae* by MS (Ho et al., 2002), but the ProHits analysis is not sufficient to confidently call it a hit in *C. albicans*. This example illustrates the difficulty in using a label-free MS data for assigning protein interactions.

It is noteworthy that Mob1 has pulled significantly more of the Dbf2 protein than the direct IP of the kinase. On the other hand, Mob1 was not recovered at all in the Dbf2 IP. Thus, it is likely that significant proportion of Mob1 in the cell is bound to Dbf2, whereas very little of the Dbf2 pool is bound to Mob1. Considering that Dbf2 is only active when bound to Mob1, the direct IP of Dbf2 most likely recovered predominantly an inactive kinase. The Dbf2-Mob1 complex recovered in the Mob1 IP suggests that this dataset is likely to contain more clients of the kinase. In conclusion, Mob1 is better suited as a bait in IP experiments aiming to identify Dbf2 targets.

Expanding the current data by doing more IPs would likely create more meaningful results. However, rather than that, a strategic decision was made to take a different approach for identifying interactions, namely using SILAC in conjunction with MS analysis.

Chapter 4

Characterization of the Substrate-Trapping Mutant Cdc14C275S

4.1. Introduction

Protein phosphorylation and dephosphorylation are transient reactions in which kinases or phosphatases remain bound to their substrates for a very short time until a phosphate group is added to or removed from the proteins. Such interactions are often too short-lived to withstand co-IP experiments and thus are very difficult to detect by AP-MS.

Previous research has shown that the interaction between protein tyrosine phosphatases (PTP) and their targets can be artificially enhanced if the catalytic residues of the enzymes are changed (Tonks and Neel, 1996). Phosphatase-dead (PD) mutants are engineered by substitution either of two essential catalytic amino acids within the active pocket of the enzyme – Cys → Ser/Ala or Asp → Ala (reviewed in Blanchetot *et al.*, 2005). Substrate-trapping mutants have significantly lower or completely absent catalytic activity and thus may interrupt downstream events governed by their substrates. Nevertheless, since PD phosphatases have higher affinity for their substrates, they have been a valuable tool for identifying physiological interactions.

Structural and kinetic analyses of hCdc14B have demonstrated striking similarities between the active pocket of this phosphatase and other PTPs (Gray *et al.*, 2003). The study identified the active Cys and Asp residues in the signature motif of hCdc14B, which were later found to be conserved in Cdc14 homologues in all species. PD mutants of Cdc14 have been engineered in *H. sapiens* (Lanzetti *et al.*, 2007), *S. cerevisiae* (Bloom *et al.*, 2011), *S. pombe* (Wolfe and Gould, 2004) and recently in *Fusarium graminearum* (Li *et al.*, 2015).

Importantly, the use of these mutants in AP-MS experiments has revealed many substrates of the phosphatase which were not detected by the same experiments, but using the wild type enzyme. This illustrates the power of using PD Cdc14 when looking for substrates of this phosphatase.

The substrate-trapping approach was recognised as a possible solution of the problems encountered in the initial MS experiments, namely the difficulty of obtaining sufficient amount of bound substrates to be detected by MS. Therefore, this approach was employed in this study too.

4.2. Generation of phosphatase-dead strains

4.2.1. Identification of the catalytic residues in the active pocket of Cdc14

The protein sequences of CaCdc14 and ScCdc14 were aligned using the Basic Local Alignment Search Tool (BLAST). The catalytic residues of ScCdc14 are D253 and C283. The corresponding amino acids in CaCdc14 are D244 and C275 respectively (fig. 4.1). C275 was arbitrarily chosen to be substituted with a serine to create a PD Cdc14.

4.2.2. Generation of *cdc14C275S*

In order to engineer PD Cdc14 in *C. albicans*, one of the endogenous alleles of *CDC14* was mutated to *cdc14C275S* using the cloning strategy illustrated in figure 4.2. A cassette containing *cdc14C275S-MYC::URA3* was cut out of the vector and transformed into MDL04 strain. The insert replaces one of the wild type alleles of *CDC14* creating a strain expressing one wild type allele of *CDC14* and one PD allele of *cdc14C275S* fused to a *MYC* epitope. A wild type allele of *CDC14* was purposely left in the genome, because cell expressing only a PD allele have a specific phenotype as discussed later in this chapter.

A separate strain expressing the PD allele *cdc14C275S* fused to GFP was also created in order to visualise the localisation of the protein. *MYC::URA3* was replaced by *GFP::ARG4* in the genome by homologous recombination.

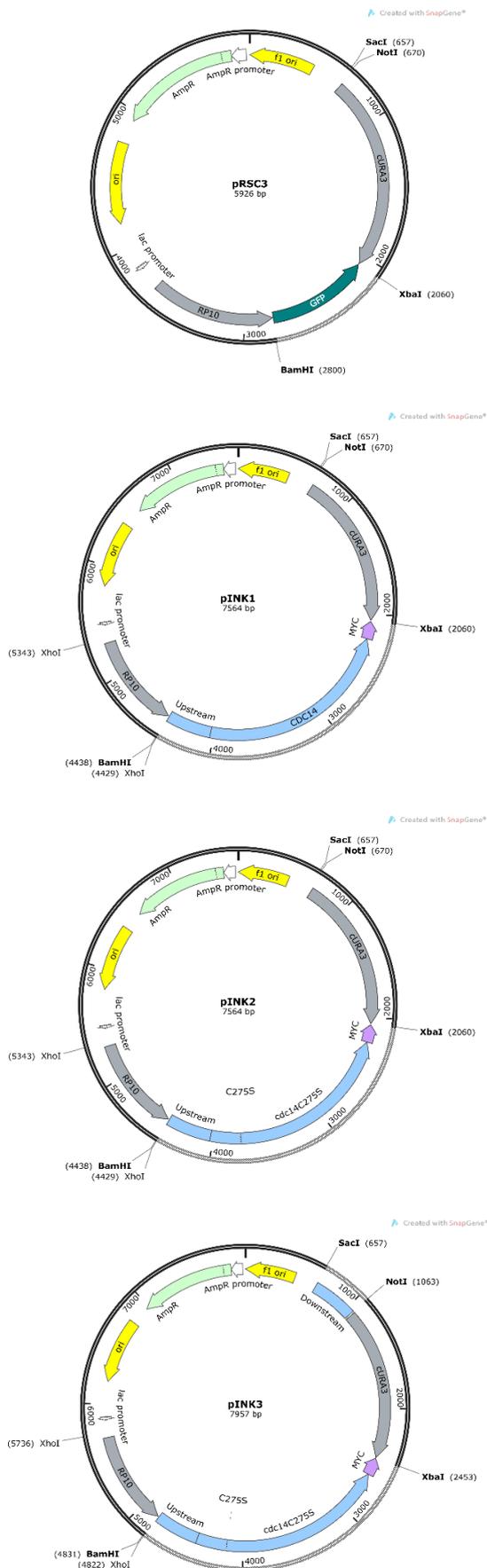
Strains expressing *cdc14C275S* will from now on be written as Cdc14^{PD}.

4.2.3. Generation of a regulatable Cdc14^{PD}

The inactive Cdc14^{PD} is expected to affinity purify more targets than the wild type enzyme. In order to further optimise the AP-MS and enrich protein binding partners, *CDC14^{PD}* was put under a regulatable *MET3* promoter, which allows the gene to be overexpressed several

Score	Expect	Method	Identities	Positives	Gaps
543 bits(1398)	0.0	Compositional matrix adjust.	281/563(50%)	365/563(64%)	60/563(10%)
Query 11	IEFLKNRIYLGAYDHHKRDTE	DLAYFTVEDALPYNAFYMDFGPLNIGHLYRFAVLLHKKL			70
Sbjct 11	IEFLRGRVYLGAYDYTPEDTDELVFFTVEDAIFYNSFHLD	DFGPMNIGHLYRFAVIFHEIL			70
Query 71	NEDSTQGKGLVIYSSTSPKERANLACLLCCYMILLQNWAPHQVLOPIAQITPPLQAFRDA				130
Sbjct 71	NDPENANKAVVYSSASTRQRANAACMLCCYMILVQAWTPHQVLQPLAQVDPFMPFRDA				130
Query 131	GYSSADYEITIQDVVYAMWRAKERGMIDLAKFDLDEYEQYERVDQGDVFNVISKDFIAFAS				190
Sbjct 131	GYSNADFEITIQDVVYGVWRAKEKGLIDLHSFNLESYEKYEHVDFGDFNVLTPDFIAFAS				190
Query 191	PQQ-----SKRGLNEPFQKVFVYVENNVQLVRLNSHLYDAKEFTKRNKIHIDMI				242
Sbjct 191	PQ+ +K LN+PF+ VL +F NNVQLVRLNSHLY+ K F I+H+D+I				250
Query 243	FDDGTCPTLEYVQKFIGAAECIINKGGKIAVHC	AGLGRGTGCLIGAHLIYTHGFTANECI			302
Sbjct 251	FDDGTCPLSIVKNFVGAETIIRKGGKIAVHC	AGLGRGTGCLIGAHLIYTYGFTANECI			310
Query 303	AYMRMIRPGMVVGPQQHWHLYLHDDFRSWRHTMIVDNRDPDLIGNLFPLCSYEDYKQRLK				362
Sbjct 311	GFLRFIRPGMVVGPQQHWHLYLHQNDFREWKYTRISLKPSEATGGLYPLISLEEYRLQKK				370
Query 363	EAKRKERLQLQQQLTSPLADSSVINTPIRRRKVSGALASKIQT-----VVP	IES			411
Sbjct 371	KLKDDKRVA-QNNIEGELRDLTMTPPSNHGALSARNSSQPSTANNGSNSFKSSAVPQTS				429
Query 412	PGQPRKYFEDSEIDEVE-----MVNNSDDENTMQDIIQSSSPARYDSVTP				456
Sbjct 430	PGQPRK S I+++ NNSDDE +MQD +S Y V+				487
Query 457	KTKDNSDWRVLRISISTNNVSSQQSIHIKTTTTKTVNETLSSPPGTSPTNVLRVSKARSK				516
Sbjct 488	KKND-----ISSASSRMEDNEPSATNINNAADD-----ILRQLLPKNR				527
Query 517	NRIASGNSQTSRAHSGGVRKLSG				539
Sbjct 528	-RVTSGRRRTSAA--GGIRKISG				547

Fig. 4.1: Alignment and CaCdc14 (Query) and ScCdc14 (Sbjct) amino acid sequence using BLAST. The catalytic residues in ScCdc14 D253 and C283 correspond to D244 and C275 in CaCdc14.



START: The backbone vector used in the cloning procedure is pRSC3.

STEP 1: The *GFP* gene was cut out of the vector using the endonucleases BamHI and XbaI generating a linearized plasmid.

STEP 2: *CDC14-MYC* was amplified from gDNA starting from 400 bp upstream of *CDC14* (insert 1).

STEP 3: The insert contained the restriction sites for XhoI (5' end) and XbaI (3' end) and was digested with these two enzymes to create sticky ends.

NOTE: The insert could not be designed with a BamHI end, because it has this restriction site internally.

STEP 4: The insert and the linear vector were ligated together with the use of a short linker sequence containing BamHI and XhoI restriction sites (pINK1).

STEP 5: The pINK1 plasmid was amplified by PCR using a primer pair that mutates a TGT (Cys275) codon to a TCT (Ser) codon (pINK2).

STEP 6: A 400 bp sequence downstream of *CDC14* was amplified by PCR from a gDNA (insert 2). This sequence contained restriction sites for NotI (5' end) and SacI (3' end).

STEP 7: The pINK2 vector and insert 2 were both digested with NotI and SacI and ligated together to create the final vector pINK3.

STEP 8: pINK3 was digested with XhoI and SacI. This created 3 fragments, the biggest of which contains *CDC14-MYC::URA3* flanked by 400 bp upstream and downstream sequence of *CDC14*. This fragment was transformed into *C. albicans* cells to create the mutant strain *cdc14^{PD}*.

Fig. 4.2: Cloning steps for the generation of *cdc14^{PD}*.

folds above physiological levels. For this purpose, cells expressing Cdc14^{PD}-Myc were transformed with *ARG4::MET3* cassette containing flanking sequences of the 5' region of *CDC14*. The cassette may integrate in the 5' region of either *CDC14* or *cdc14^{PD}*. To separate these two outcomes apart, PCR-positive colonies were grown individually in *MET3*-inducing media and *MET3*-repressing media. Using Western blot, two colonies were identified to have upregulated levels of Cdc14^{PD} in a *MET3*-on culture and completely absent Cdc14^{PD} in a *MET3*-off environment (fig. 4.3). This shows that in these colonies, the *MET3* promoter is controlling the expression of Cdc14^{PD}.

The expression of *CDC14* cannot be measured by Western blot because the gene is not fused to an epitope. Thus, integration of the cassette in front of *CDC14* was confirmed by DNA sequencing. However, the Western blot pattern of Cdc14^{PD} of one of the colonies suggested this to be the case as explained in figure 4.3. In a *MET3*-on culture Cdc14^{PD} was expressed at its natural levels but without the characteristic hyperphosphorylation seen in the *MET*-off environment (as discussed in section 4.3.1). Clp1 is known to transautodephosphorylate in *S. pombe* (Wolfe *et al.*, 2006). This is the first evidence suggesting that CaCdc14 may be doing the same. Overexpression of the catalytically active Cdc14 produced fully dephosphorylated Cdc14^{PD}, while downregulating Cdc14 results in hyperphosphorylated Cdc14^{PD}. Further evidence that in this strain the *MET3* promoter controls *CDC14* expression came from the observation that cell with induced *MET3* display the same phenotype as *cdc14Δ/Δ* (see section 4.4.1), while cells with repressed *MET3* have the phenotype of wild type cells.

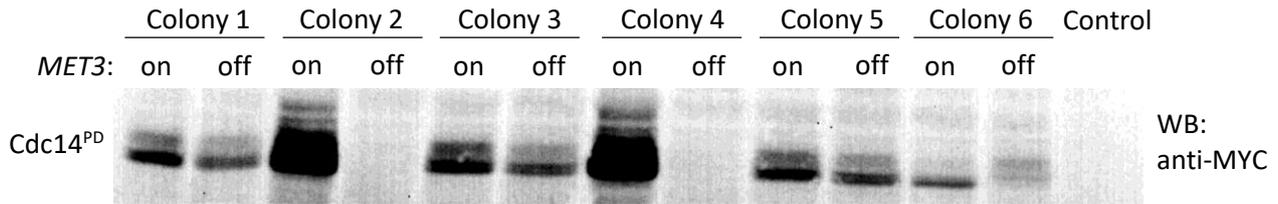


Fig. 4.3: Colony screen for integration of *MET3* promoter in front of either *CDC14* or *cdc14^{PD}-MYC*. Six colonies transformed with *MET3* were grown in liquid broth that renders the promoter either on or off. Colonies 2 and 4 showed highly increased expression of *cdc14^{PD}* when the promoter is switched on and complete absence of the protein when the promoter is off. This indicates that the *MET3* promoter was integrated in front of the mutant allele in these two colonies. The wild type *CDC14* allele is not fused to any tagged so its expression cannot be tested by Western blot. However, the expression pattern of colony 6 suggests that *MET3* is controlling *CDC14*. When the promoter is on, *cdc14^{PD}* loses its characteristic hyperphosphorylation pattern. Cdc14 is known to autodephosphorylate in *S. pombe*, so overexpression of the wild type protein likely dephosphorylates the mutant protein. When the promoter is off, i.e. no wild type Cdc14 is expressed, the majority of *cdc14^{PD}* is phosphorylated. Colonies 1, 3 and 5 were unsuccessful transformants. Control is lysate from untagged cells.

4.3. Characterisation of Cdc14^{PD} by Western blot

4.3.1. Phosphorylation status of Cdc14^{PD}

Cdc14 expressed from asynchronous cells produces a single band on a Western blot, although the protein is phosphorylated in a part of the cell cycle. In contrast to this, Cdc14^{PD} extracted from *cdc14^{PD}/CDC14* cells, produced a clearly smeared band, characteristic of phosphorylated proteins. To confirm that the protein is indeed phosphorylated, cell lysates were incubated in the presence or absence of lambda phosphatase at 37 °C. The phosphatase-treated sample produced a single band on a Western blot (fig. 4.4). Hence, the smear of Cdc14^{PD} is due to phosphorylation of the protein.

4.3.2. Expression of Cdc14^{PD} in yeast and hyphae

Protein levels of Cdc14 vary widely throughout the cell cycle from being completely absent in G1 to being highly expressed in anaphase in both yeast and hyphae (Clemente-Blanco *et al.*, 2006). To follow the expression of Cdc14^{PD}, an overnight culture of *cdc14^{PD}/CDC14* cells was left in water at room temperature for 4 hours to induce all cells into entering stationary phase. Cells were then released into fresh medium and allowed to grow as either yeast or hyphae for 90 min. Although this method of synchronisation is not as efficient as elutriation, it is much easier to carry out in the lab and the vast majority of cells are in the same phase of the cycle. Samples were taken every 15 minutes and cells were immediately lysed to extract soluble proteins. As shown by Western blot on fig. 4.5, the Cdc14^{PD} was not detectable in the initial stages and its levels increased steadily during the course of the experiment in both yeast and hyphae. This shows that Cdc14^{PD} follows the same expression pattern as the active Cdc14.



Fig. 4.4: Phosphatase treatment of Cdc14^{PD}. Cdc14^{PD} produces a band shift on a Western blot which collapses to a sharp band when a cell lysate is treated with λ phosphatase. This shows that the protein is phosphorylated and most likely at multiple sites, since no clear distinction can be made between the phosphorylated and non-phosphorylated forms on a Western blot. Note: the white mark in the untreated sample is a defect of the membrane.

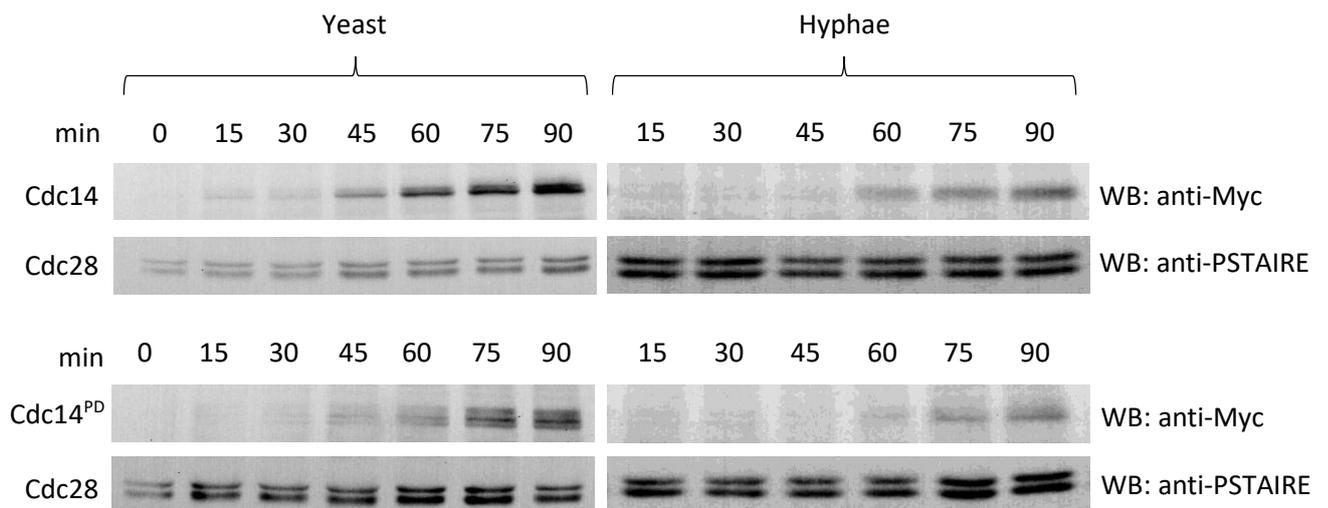


Fig. 4.5: Comparison of expression levels of Cdc14 and Cdc14^{PD}. A time course experiment showed very similar expression pattern of the wild type and the mutant proteins. Cells were grown overnight and then starved in water for 4 hours to induce transition into G₀. They were then released into fresh medium and left to grow in either yeast- or hyphae-promoting conditions. Cells are semi-synchronised. Cells were taken every 15 min and lysed. The phosphatase is not present in stationary phase cells and it starts appearing after about 45 min in yeast and 60 min in hyphae. More importantly, Cdc14^{PD} does not show any signs of untimely expression. The budding index after starvation in water was not measured.

4.3.3. Co-IP of Cdc14 and Cdc14^{PD}

As already mentioned in section 4.2.3, overexpression of the wild type Cdc14 completely diminished the phosphorylation of Cdc14^{PD}. The possibility of interaction between both forms of the phosphatase was further investigated by co-IP. The substrate-trapping Cdc14^{PD}-Myc was co-expressed with Cdc14-GFP and cell lysates were incubated with anti-Myc affinity beads. Western blot of the immunoprecipitated proteins detected both forms of Cdc14 (fig. 4.6), adding further evidence for direct physical interaction.

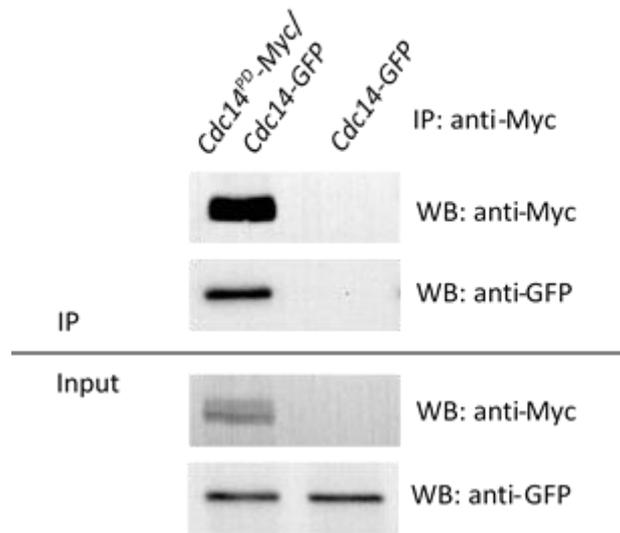


Fig. 4.6: Co-immunoprecipitation of Cdc14^{PD} and Cdc14. IP of Cdc14^{PD} pulls down the wild type phosphatase Cdc14, suggesting physical interaction between both proteins.

4.4. Phenotypic analysis of Cdc14^{PD} using microscopy

4.4.1. Morphology of PD mutants

CDC14/cdc14^{PD} cells grew completely normal in both yeast and hyphae-inducing conditions. Cells displayed no visible differences to *CDC14/CDC14* when examined by brightfield microscopy. Hyphal formation was also not affected by the presence of the mutant allele (fig. 4.7). This suggests that *C. albicans* is haplosufficient for *CDC14* and that *cdc14^{PD}* is not a dominant negative allele.

When the wild type allele was repressed by MET3 promoter, *MET3-CDC14/cdc14^{PD}* cells exhibited the same phenotype as *cdc14Δ/Δ*, namely defects in cell separation resulting in chains of yeast cells and inability to form proper hyphae (fig 4.7). This is an evidence that Cdc14^{PD} has indeed lost its phosphatase activity.

When Cdc14^{PD} was overexpressed with the use of MET3, cells grown in the absence of methionine remained with normal morphology, but grew slightly slower than wild type cells. Overexpression of *cdc14^{PD}* likely depletes the large pool of Cdc14 substrates. Considering the role of Cdc14 in mitotic progression through anaphase, this phenotype is not surprising. However, it is evident that as long as one wild type allele of the phosphatase is expressed, cells are able to grow without morphological defects.

4.4.2. Localisation of Cdc14^{PD}

The localisation of Cdc14^{PD}-GFP was examined by fluorescent microscopy. The mutant showed difference to wild type protein localisation in neither yeast, nor hyphae. In yeast, Cdc14^{PD} was seen in the nucleus during interphase and at the spindle pole bodies during cell separation (fig. 4.7). In hyphae, Cdc14^{PD} was only detected in the nucleus. This is also the localisation pattern of wild type Cdc14-GFP in the presence of Cdc14^{PD}. All of these results show that the catalytic inactivation of Cdc14^{PD} does not affect its localisation in the cell.

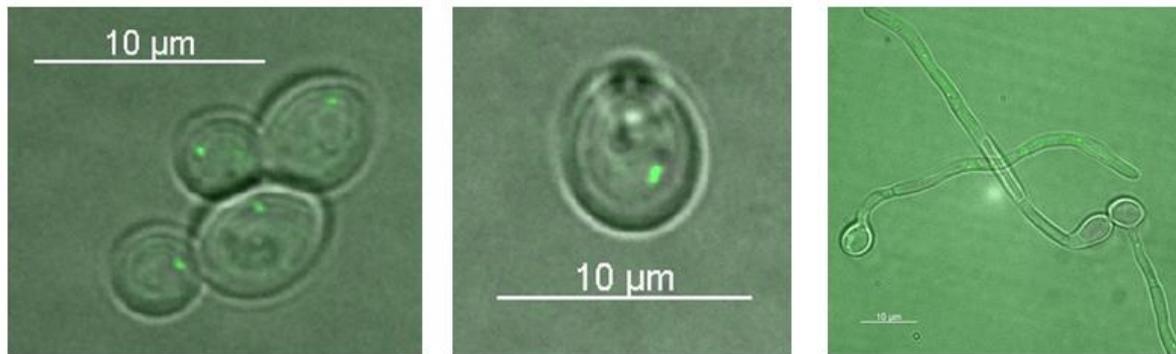
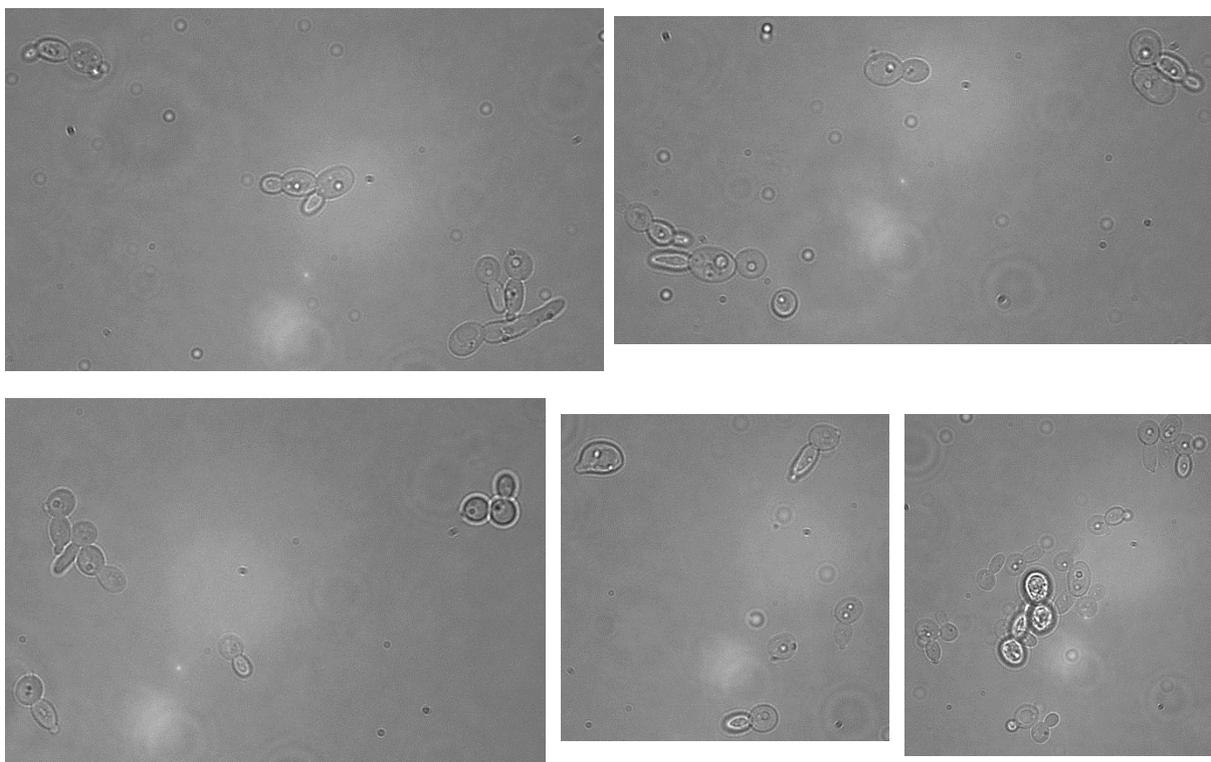
A**B**

Fig. 4.7: Phenotype of Cdc14^{PD}. (A) In *cdc14^{PD}-GFP/CDC14* cells, *CDC14* is dominant to *cdc14^{PD}*, since in the presence of both alleles cells did not display any visible defects. Just like the wild type phosphatase, Cdc14^{PD}-GFP localised to two spots likely corresponding to the spindle pole bodies in dividing yeast and to one spot, most likely the nucleolus, in non-dividing yeast and hyphae. (B) In *cdc14^{PD}/MET3-CDC14* the wild type allele was turned off with the use of *MET3* promoter, and only Cdc14^{PD} was expressed. As a result, hyphal formation was severely impaired even 2 hours after induction. The budding index was not measured.

4.4.3. Localisation of Mlc1 in the presence of Cdc14^{PD}

Cdc14 Δ/Δ cells have severe defects in septum formation (Clemente-Blanco et al., 2006). To examine whether this process is affected by cdc14^{PD}, the septum marker protein Mlc1 was fused to GFP in *CDC14/cdc14^{PD}* background. As in wild type cells, Mlc1 localised to the septum of both yeast and hyphae and demonstrated dynamic contraction of the septum ring at the end of mitosis (fig. 4.8).

4.4.4. IP of Cdc14^{PD}

IP of overexpressed Cdc14^{PD} was performed as previously described, except that the affinity beads were washed 3 times instead of 1 time after incubation with cell lysate. It was hypothesised that the stronger bait-prey interaction will withstand the washes, while the level of contaminants will be lower than previously seen. Indeed, Coomassie stained gel showed much brighter Cdc14^{PD} bands on a much clearer background (data not shown).

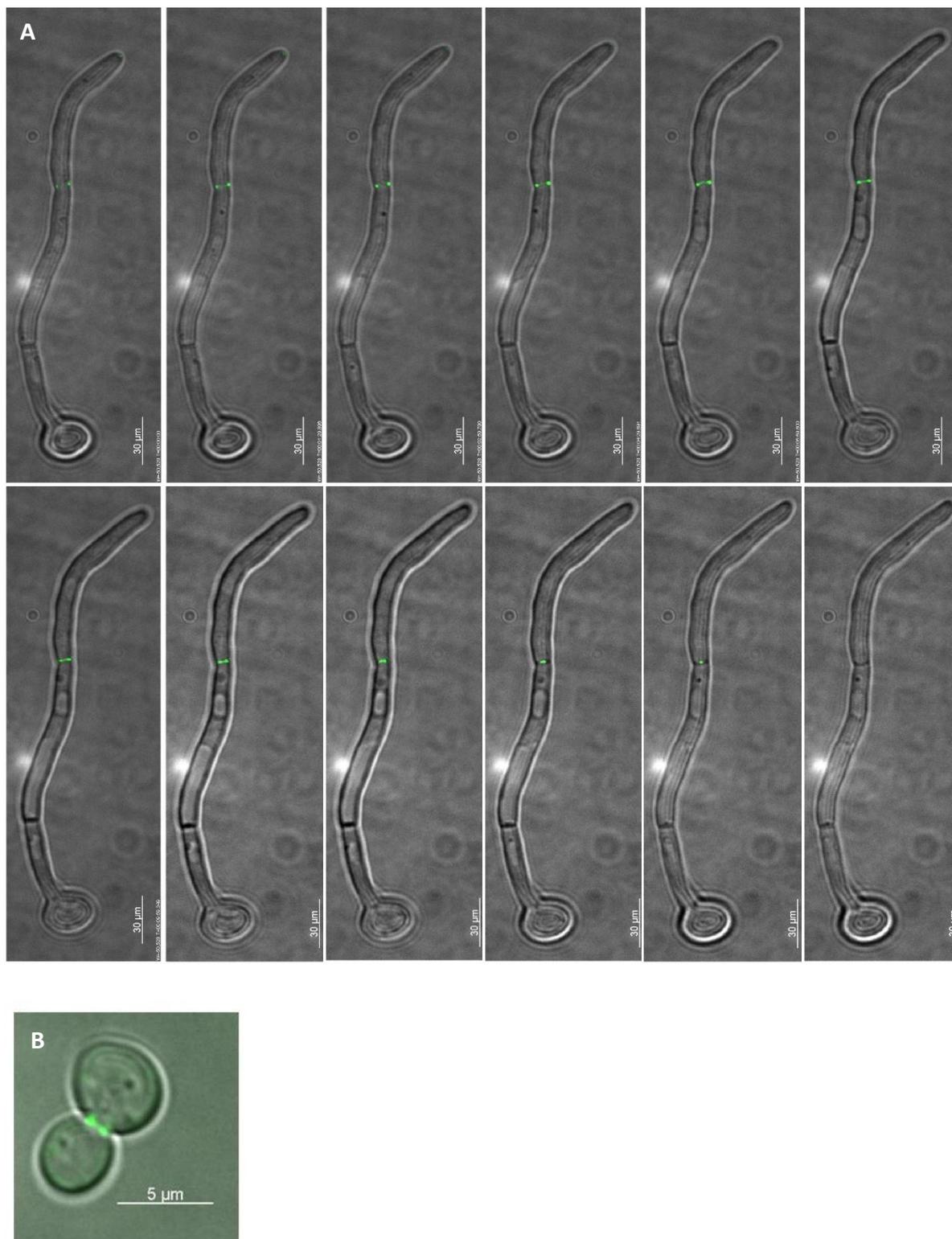


Fig. 4.8: Localisation of Mlc1-GFP in *cdc14^{PD}/CDC14* background. A) Time lapse microscopy shows that Mlc1 localises to the cytokinetic ring in hyphae and its contraction is not affected by the presence of *Cdc14^{PD}*. **B)** In yeast, Mlc1 has a normal localisation at the bud neck. Bright field and fluorescence images are overlaid.

4.5. Discussion

The initial experiments of this study, where label-free MS was used, showed ambiguous results and did not allow for clear identification of interacting partners of the baits. Such potential partners were often underrepresented in comparison to the contaminants, which reflects the fact that interactions are easily lost in the experimental procedure due to their weak and transient nature. At this stage it became evident that in order to produce meaningful results, the ratio of prey-to-background proteins has to be significantly higher.

This study made use of a substrate-trapping method to stabilise the interaction of the phosphatase Cdc14 with its partners. The method employing a single amino acid substitution has already aided in identification of Cdc14 clients in other organisms. The catalytic Cys and Asp residues of Cdc14 are highly conserved across species and their position was easily identified in *C. albicans* by sequence alignment. The phosphatase dead Cdc14C275S was created by series of cloning steps. The catalytic activity of the mutant was not directly measured by *in vitro* assay, but the phenotype of *MET3-CDC14/cdc14^{PD}* when MET3 is repressed reminisces that of *cdc14Δ/Δ*, which strongly suggested that Cdc14^{PD} is not functional. It was confirmed that this phenotype is not due to haploinsufficiency since cells with a single copy of functional Cdc14 display wild type characteristics. These findings were taken into consideration when further MS experiments were performed, namely all strains used in these experiments were expressing one active and one inactive allele of Cdc14. The active phosphatase is required to maintain the cells in a healthy physiological state. The disadvantage of keeping the wild type Cdc14 is that it competes for the same substrates as Cdc14^{PD}. However, the mutant phosphatase is overexpressed and largely outnumbers the wild type protein. Thus, competition for substrates is not a concern, as most of them will be bound to Cdc14^{PD}.

A conserved feature of all Cdc14 phosphatases is that they become constitutively hyperphosphorylated when mutated to inactive enzymes and CaCdc14 is no exception. *In vitro* phosphatase treatment abolished the gel shift seen in non-treated samples of Cdc14^{PD}. Wolfe *et al.* have shown that in fission yeast Clp1 regulates its own activity by

dephosphorylating itself at six Cdk1 sites. This is likely the case with Cdc14 since interaction between Cdc14 and Cdc14^{PD} was shown by co-immunoprecipitation. Additionally, overexpression of Cdc14 reduced the phosphorylation state of Cdc14^{PD} to a single gel band. Although this is not a direct evidence for autodephosphorylation, it makes a strong case towards this hypothesis. Additional experiments would be required to prove beyond doubt that Cdc14 is acting on itself, such as *in vitro* dephosphorylation reaction. However, it was not the aim of this study to investigate Cdc14 regulation, so no additional experiments were performed.

The localisation of Cdc14^{PD} was further studied by fluorescent microscopy. It is important that Cdc14^{PD} localised in the same manner as Cdc14, as mislocalisation of the protein could cause non-physiological interactions. In addition, Cdc14^{PD} disrupted neither the morphology nor the growth of the cells, as long as the wild type protein is also expressed. The completion of cytokinesis was specifically examined, because Cdc14 is known to be involved in this process. In the presence of Cdc14^{PD}, Mlc1 showed proper localisation at the bud neck of dividing cells and time lapse movies demonstrated successful and undisrupted contraction of the septum ring in hyphae.

Overexpression of Cdc14^{PD} is another tactic that was employed with the aim to capture maximum number of protein partners. Cdc14 is not an abundant protein, which makes AP-MS very difficult without overexpressing it. Furthermore, it is not universally expressed throughout the cell cycle as shown by Clemente-Blanco *et al.* (and the same pattern was observed here for Cdc14^{PD}). Instead, its levels vary widely from completely missing in G₁ to peaking in anaphase. Although the phosphatase is present for the most of the cell cycle, it is thought to be active only for a brief period during mitosis. This makes a screen for interactors very difficult, because capturing sufficient amount of the bait and preys would be very challenging. On the other hand, overexpression of the phosphatase creates the risk of identifying false positive hits, i.e. interaction that would not normally occur at physiological levels of Cdc14. Nevertheless, taking everything into account, a decision was made that an overexpressed protein is likely to produce a better quality data. It is important to mention that purifying Cdc14 from synchronised cells in mitosis (when the protein is most abundant) was not a viable option due to technical constraints. Cell elutriation is a rather laborious technique that can be used to synchronise a small amount of

cells in the same phase of the cycle. An MS experiment requires far more cells than could possibly be obtained by elutriation. Other methods for obtaining mitotic cells include generating mutants that arrest the cells in mitosis. However, perturbing molecular pathways creates an artificial environment for Cdc14 and may also lead to identifying false interactions. Thus, for the purpose of this study, using non-synchronised cells was the best option.

All of the findings presented here cumulatively indicate that Cdc14^{PD} is a viable candidate for an MS screen of Cdc14 interactions. The rest of this study focussed solely on identifying Cdc14 targets.

Chapter 5

SILAC Labelling in *Candida albicans*

5.1. Introduction

Stable isotope labelling by amino acids in cell culture (SILAC) is a metabolic labelling technique that produces mass difference between two proteomes read out by MS. It was originally developed in 2002 for use in mammalian cell lines, but since then it has been successfully applied to various organisms, including bacteria, yeast, plants, worms and flies (Ong et al., 2002; Kerner et al., 2005; Jiang and English, 2002; Gruhler et al., 2005). SILAC depends on *in vivo* incorporation of stable isotope-labelled amino acids present in the growth medium of cell cultures. Thus, any organism that feeds on amino acids can be used in SILAC experiments (Ong and Mann, 2006). This study describes the first application of SILAC labelling in *C. albicans*.

The first important consideration in SILAC is the choice of labelled amino acids. This is largely dependent on the protease used to digest the samples prior to MS analysis. Ideally, every peptide coming from a heavy-labelled proteome should carry at least one heavy amino acid. Heavy and light peptide analogues form SILAC pairs in the mass spectrum that show their relative abundance in each sample. Trypsin cleaves at the C-terminus of arginine (Arg) and lysine (Lys), so it is commonly used in combination with heavy Arg and Lys. In cases where only one of those amino acids is used, e.g. only Lys, peptides cleaved after Arg are not used in quantitation but are still used for protein identification. Alternatively, other protease may be used (e.g. LysC cleaves only after Lys) or other heavy amino acids may be added. Amino acid labelling is achieved with the use of the heavy isotopes ^{13}C and ^{15}N , while naturally occurring forms contain ^{12}C and ^{14}N . Deuterium (^2H) is rarely used because it may affect the retention time of the peptides in the chromatography phase (Zhang and Regnier, 2001).

Many organisms have metabolic pathways for conversion of Arg into other amino acids, most notably proline (Pro). This includes some yeast species, such as *S. cerevisiae* and *S. pombe* (Gruhler et al., 2005; Bicho et al., 2010). Arginine is utilised as a nitrogen source, when other sources are not available. The arginine degradation pathway has not been studied in *C. albicans*, but orthologues of all *S. cerevisiae* genes involved in it have been found (fig. 5.1). In experiments using heavy Arg, conversion to Pro creates additional satellite peaks in the mass spectrum that reduce the intensity of heavy ions and change the ratio of SILAC pairs. Therefore, the presence of heavy Pro must be carefully examined in all experiments, because it can significantly compromise the quality of the data.

SILAC experiments are inherently reliant on full incorporation of the heavy amino acids into the proteome, so it is important to allow the cells sufficient time to grow in heavy medium. Most cultures require at least five doublings for that, but the exact number may vary between species. Labelling efficiency must be determined by MS analysis of cell extracts derived solely from a heavy-labelled population before proceeding further with experiments. Incorporation efficiency of >95% is considered good, because anything less than that will decrease the maximum observable ratio of the SILAC pairs. For example, at 95% incorporation the ratio of heavy to light peptides will be 1:20, whereas at 90% it can only go up to 1:10.

SILAC typically involves using essential amino acids for proteomic labelling, which cells cannot synthesise on their own. The use of prototrophic organisms has only recently been investigated, but seems to be a valid approach in some yeast and bacteria (Frohlich et al., 2013). In the presence of readily available Lys sources, growing cells downregulate endogenous Lys production and achieve full incorporation of the heavy variants. This method, termed native SILAC (nSILAC), conveniently bypasses the requirement of generating auxotrophic strains. However, it is important to note that cells can switch back to producing Lys if they sense the need to do so, for example if they reach a stationary phase of growth (Martin-Perez and Villeñ, 2015).

This chapter explores the opportunity of using the SILAC method in *C. albicans* by carefully examining labelling strategies and cell behaviour. The use of this fungus for nSILAC is also discussed.

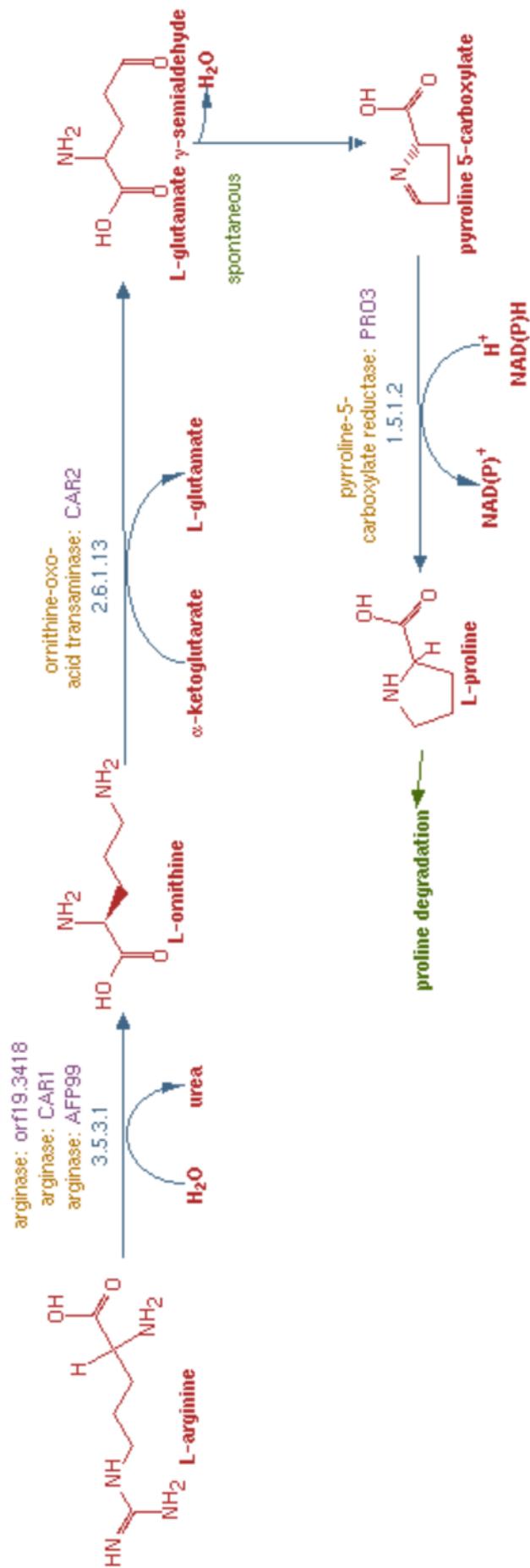


Fig. 5.1: Metabolic pathway for conversion of arginine to proline in *C. albicans*. This figure is taken from the Candida Genome Database.

5.2. Media formulation and growth conditions used in SILAC

5.2.1. Lysine

All SILAC experiments were done using cells expressing Cdc14^{PD} under the *MET3* promoter in MDL04 background strain. Cdc14^{PD} was used as bait and untagged MDL04 cells as control. Both strains were grown in *MET3* inducing media. *MET3*-Cdc14^{PD} cells were always grown in “heavy” medium, while wild type cells were grown in “light” medium. Both strains are lysine auxotrophs, so they were supplemented with either 100 mg/L unlabelled Lys or 80 mg/L Lys8 (¹³C₆, ¹⁵N₂). The molar concentration of lysine in both media is the same because Lys8 has higher molecular weight (MW) than Lys (227.05 g/mol and 182.6 g/mol respectively).

5.2.2. Arginine

The *MET3*-Cdc14^{PD} strain does not require supplementation with any amino acids other than Lys. However, in order to label Arg-containing tryptic peptides, *MET3*-Cdc14^{PD} was grown in the presence of 80 mg/L Arg10 (¹³C₆, ¹⁵N₄). MDL04 requires supplementation with arginine, so 100 mg/L of unlabelled Arg was added to the media of wild type cells. Again, both media were formulated with equal molar concentration of arginine, which explains the difference in mass concentration (Arg MW=174.2 g/mol; Arg10 MW=220.59 g/mol). When cells were cultured in media with equal mass concentration, they grew at different rate (data not shown).

5.2.3. Other constituents of the media

Both heavy and light *MET3*-inducing media contained glucose, yeast nitrogen base, complete supplement media lacking methionine, lysine and arginine, and 100 mg/L uridine. Although *MET3*-Cdc14^{PD} cells did not require uridine, it was added to both media to keep the conditions as equal as possible. When hyphae were induced, both media were mixed

with 10% v/v serum. *MET3*-repressing media contained methionine and cysteine in addition to all ingredients of *MET3*-inducing media.

5.2.4. Growth conditions in SILAC media

To prepare a yeast culture, *MET3*-*Cdc14*^{PD} cells were allowed to grow in heavy *MET3*-repressing medium overnight and were then subcultured into fresh heavy *MET3*-inducing medium at OD₅₉₅ 0.25 for 4 hrs. The control culture was grown in the same conditions but in “light” medium. Both strains grew at the same rate. After four hours the light and heavy cultures reached and OD₅₉₅ 0.70 and 0.69 respectively. Cell morphology was inspected by bright field microscopy and the heavy isotopes did not cause any visible phenotype (data not shown). The conclusion is that the presence of heavy isotopes in the growth medium of *C. albicans* does not affect cell development.

When hyphae were induced, an overnight yeast culture of both strains was prepared as described above. Each strain was then inoculated into six flasks containing SILAC media plus serum at OD₅₉₅ 0.4. Cells were allowed to grow at 37 °C for 60, 75, 90, 105, 120 and 135 min. All cell pellets were mixed together and processed further as one sample. This means that all hyphal experiments examine the interactions of *Cdc14*^{PD} that occur in hyphae between 60-135 min after induction. This is the time when the phosphatase is most abundant, so it is likely to be most active as well (Clemente-Blanco et al., 2006).

5.3. Labelling efficiency in yeast

The incorporation efficiency of Lys8 and Arg10 into the proteome of MET3-Cdc14^{PD} was measured by MS. Heavy-labelled cell extract were prepared as described in chapter 3 and proteins were separated by SDS-PAGE. The gel was stained with Coomassie to make protein bands visible and a small piece of the gel was excised. Proteins contained within this piece were digested with trypsin and the resulting peptides were extracted from the gel and analysed by MS.

The data analysis yielded in total 1777 peptides (after filtering out contaminants) assigned to 184 proteins. In a SILAC data each peptide is represented by a heavy and a light form and a software program measures the intensity of each form. This is used to calculate the heavy to light ratio (H/L). Incorporation of the heavy amino acids in each peptide was calculated by the following equation:

$$\text{Incorporation} = \frac{H/L}{H/L + 1} \times 100$$

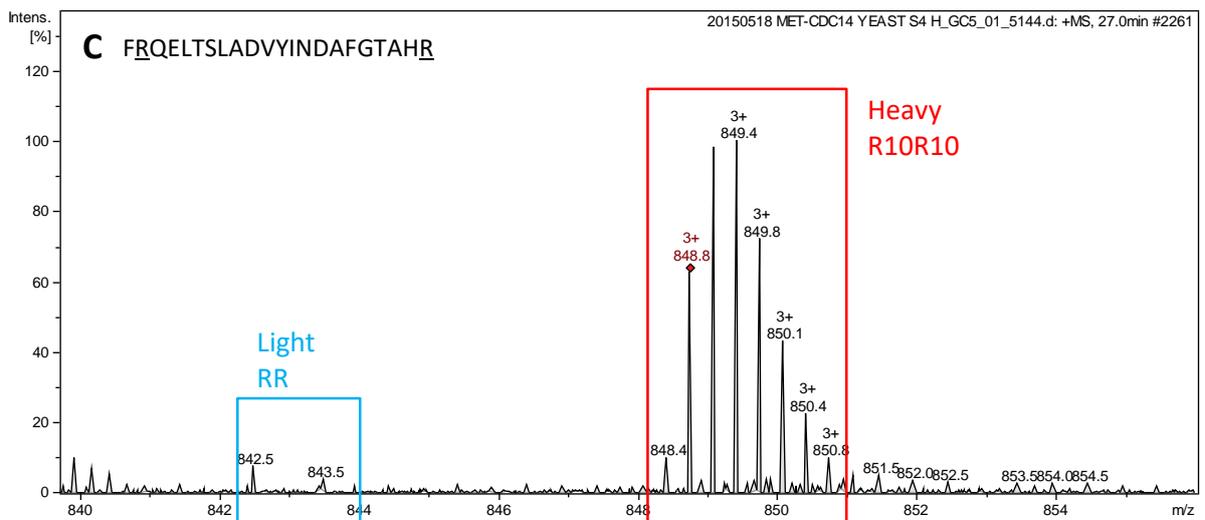
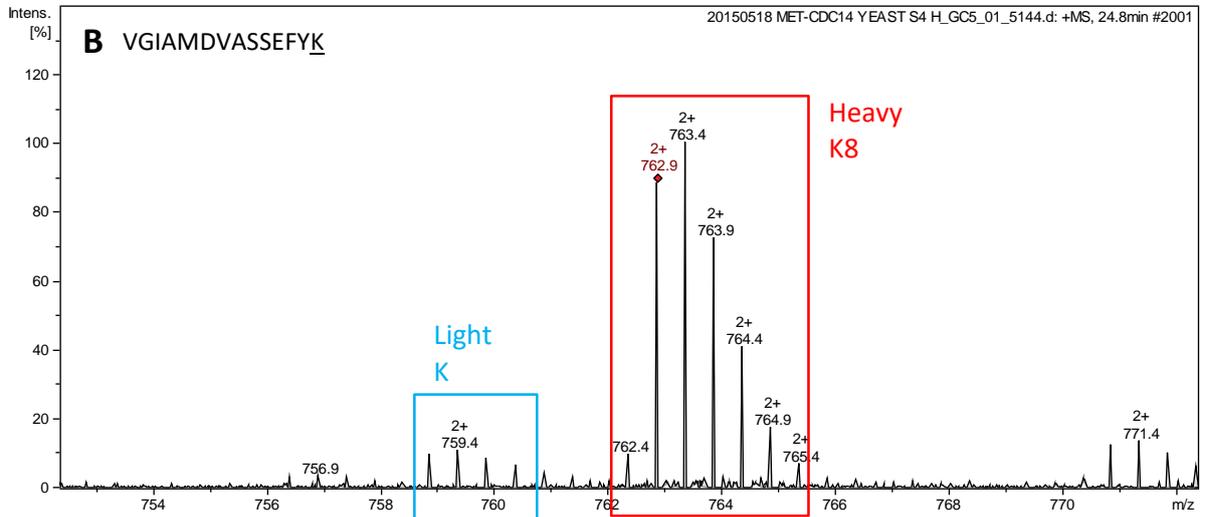
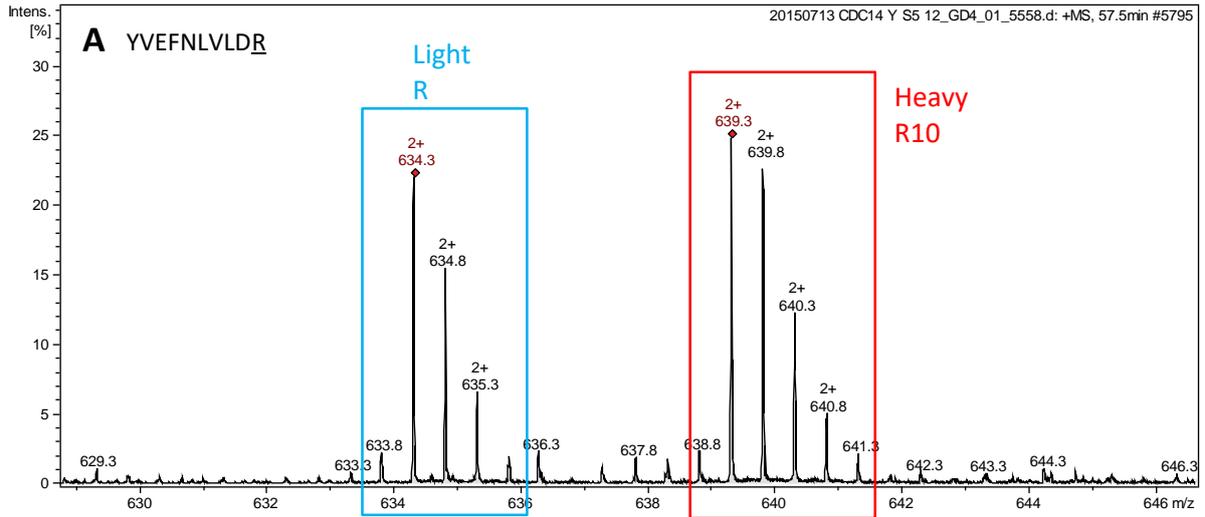
Peptides that are found only in the heavy form are fully labelled, so the incorporation is 100%. Peptides that are only light have incorporation of 0%. An example of selected peptides is given in table 5.1. The labelling efficiency in the whole sample was calculated by taking the average incorporation of all 1776 peptides, which is 93.87%.

The incorporation of Lys8 and Arg10 was also calculated separately by averaging peptides containing only one of these amino acids. Altogether 1129 peptides contained only Lys8 with incorporation of 95.43%, while 499 Arg10 peptides had 89.54% incorporation. The incorporation of both amino acids is good enough to proceed with SILAC experiments.

Large scale analysis was performed by quadrupole orbitrap and data was processed by MaxQuant. However, this procedure was repeated routinely on a smaller scale before each SILAC experiment as a quality control. Representative mass spectra from QTOF-MS are shown in figure 5.2.

Peptide sequence	Lys8	Arg10	Missed cleavages	Ratio H/L	Intensity	Intensity L	Intensity H	Incorporation %
SINPNYTPVPVP ET <u>K</u>	1	0	0	NaN	2.08E+08	0	207870000	100
PTIPVD <u>K</u> EDL <u>F</u> <u>K</u>	2	0	1	NaN	6368400	0	6368400	100
DAYVYQR <u>P</u> VYIG LPSNLVDM <u>K</u>	1	1	0	NaN	6928900	0	6928900	100
AL <u>K</u> EDDD <u>F</u> KS <u>N</u> L NDPVYTLG <u>K</u>	3	0	2	NaN	8113800	0	8113800	100
<u>R</u> YEDPEVQ <u>R</u>	0	2	1	NaN	8820500	0	8820500	100
NPENTIVN <u>F</u> <u>R</u>	0	1	0	9.2414	6747600	1177400	5570200	90
VTPSFVAFTSEE <u>R</u>	0	1	0	7.4277	80531000	9744800	70787000	88
AFNMFI <u>L</u> DPI <u>F</u> <u>R</u>	0	1	0	NaN	1982800	0	1982800	100
KIDLSLHPNDPE SQTEVIETVEK	2	0	1	NaN	17580000	17580000	0	0
RLETINEEDLQK	1	1	1	NaN	73065000	73065000	0	0
VVAIVESTSGDK VPPNTPSDEQSR	1	1	1	NaN	5949800	5949800	0	0
EVVFGMS <u>K</u>	1	0	0	52.238	15727000	309240	15417000	98
SLDSIMAVG <u>E</u> <u>K</u>	1	0	0	47.863	53559000	1117400	52442000	98
YVEDV <u>L</u> <u>K</u>	1	0	0	35.625	21063000	552740	20510000	97

Table 5.1: Incorporation efficiency of Arg10 and Lys8. The MaxQuant software reports the intensity of light and heavy version of each peptide. The sum of both is the total peptide intensity (column 6). The H/L ratio in column 5 was used to calculate the incorporation efficiency of the heavy amino acids (last column). Heavy amino acids are underlined. NaN – not a number



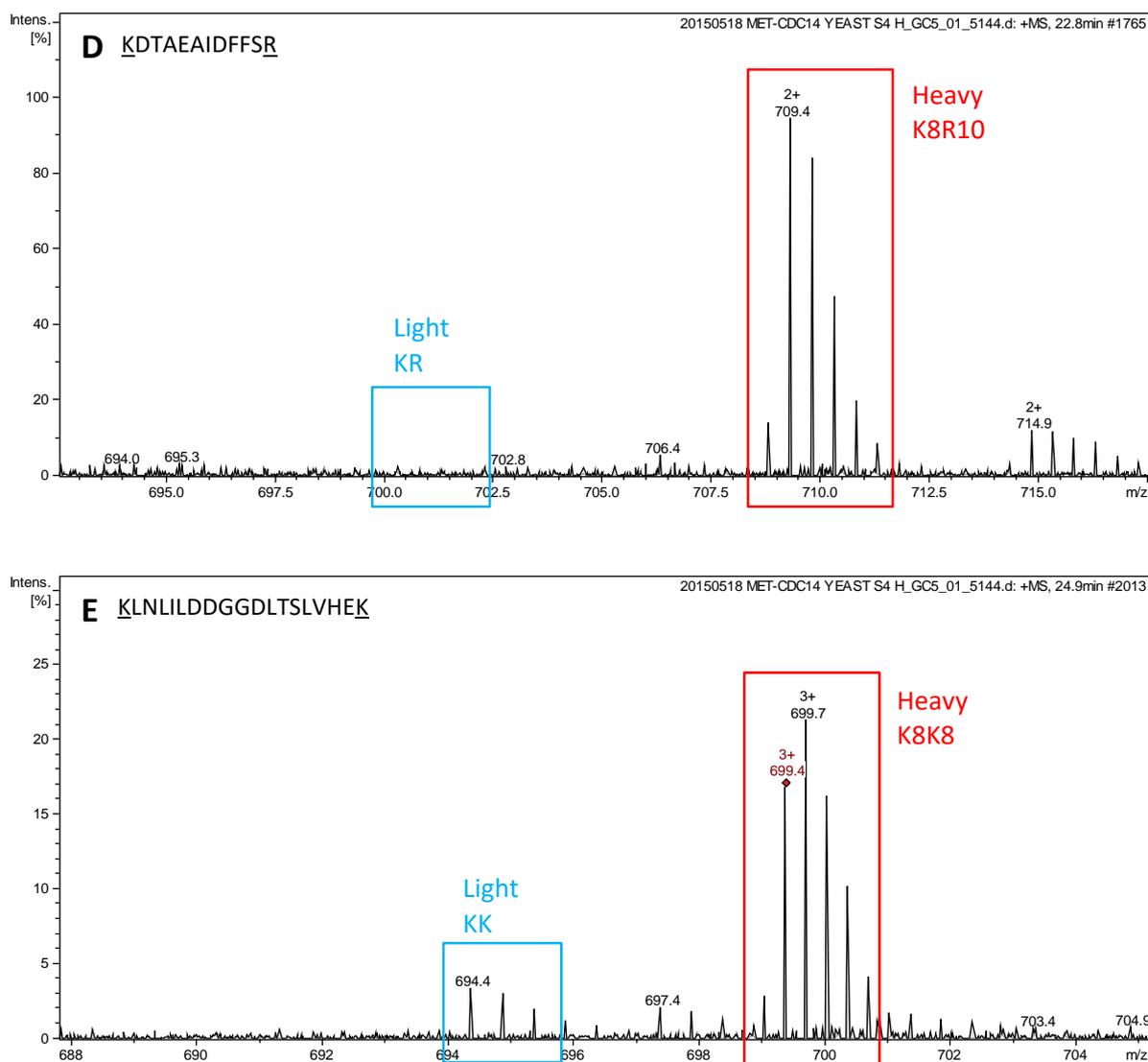


Fig. 5.2: Determining labelling efficiency using an ion chromatogram. An ion chromatogram displays the m/z of identified ions (x axis) plotted against their intensity (y axis). Part A shows a SILAC pair of the peptide YVEFNLVLDLDR which is equally enriched in light (blue) and heavy (red) isotopes, because both ions have the same intensity. Parts B and E show example of ions that are predominantly labelled, but have a small amount of unlabelled species. Parts C and D show ions that are fully labelled (the blue squares indicate the area where the light ions would be seen if they were present). Part A comes from a mixed sample (i.e. heavy and light culture in 1:1 ratio) and parts B-E come from a heavy-labelled sample. In this case the sample would be considered successfully labelled. All samples were analysed by QTOF-MS.

5.4. Labelling efficiency in hyphae

Hyphal formation is commonly induced by addition of serum to the growth medium. The amino acid composition of the serum is not known. If the serum contains light Lys and Arg, it could potentially reduce the labelling efficiency of cells. In order to test if this is the case, labelled cell extracts from hyphae were prepared as described in section 5.3 and analysed by QTOF-MS. The QTOF data is processed using Mascot Distiller, which does not produce a comprehensive table of peptide intensities like MaxQuant does. Therefore incorporation cannot be calculated on a large scale. Instead, isotope incorporation was visually assessed using the ion chromatogram.

As shown in figure 5.3 Lys8 achieved high percent of incorporation (>98%). The vast majority of peptides examined contained no light peaks at all. However, the percent of Arg10 incorporation was lower (fig. 5.4). Residual unlabelled peptides were a common place, but their intensity was typically below 5% (100% is the intensity of the most abundant ion hitting the detector in a given time point). In contrast, heavy ions produced significantly greater intensities. Thus, it is clear that Arg10 incorporation was very high, but not as high as Lys8 incorporation.

The data presented here show that *C. albicans* can successfully incorporate the heavy amino acids Arg10 and Lys8 when cells are grown as either yeast or hyphae. Therefore, SILAC-MS is a viable method for investigating protein interactions in this organism.

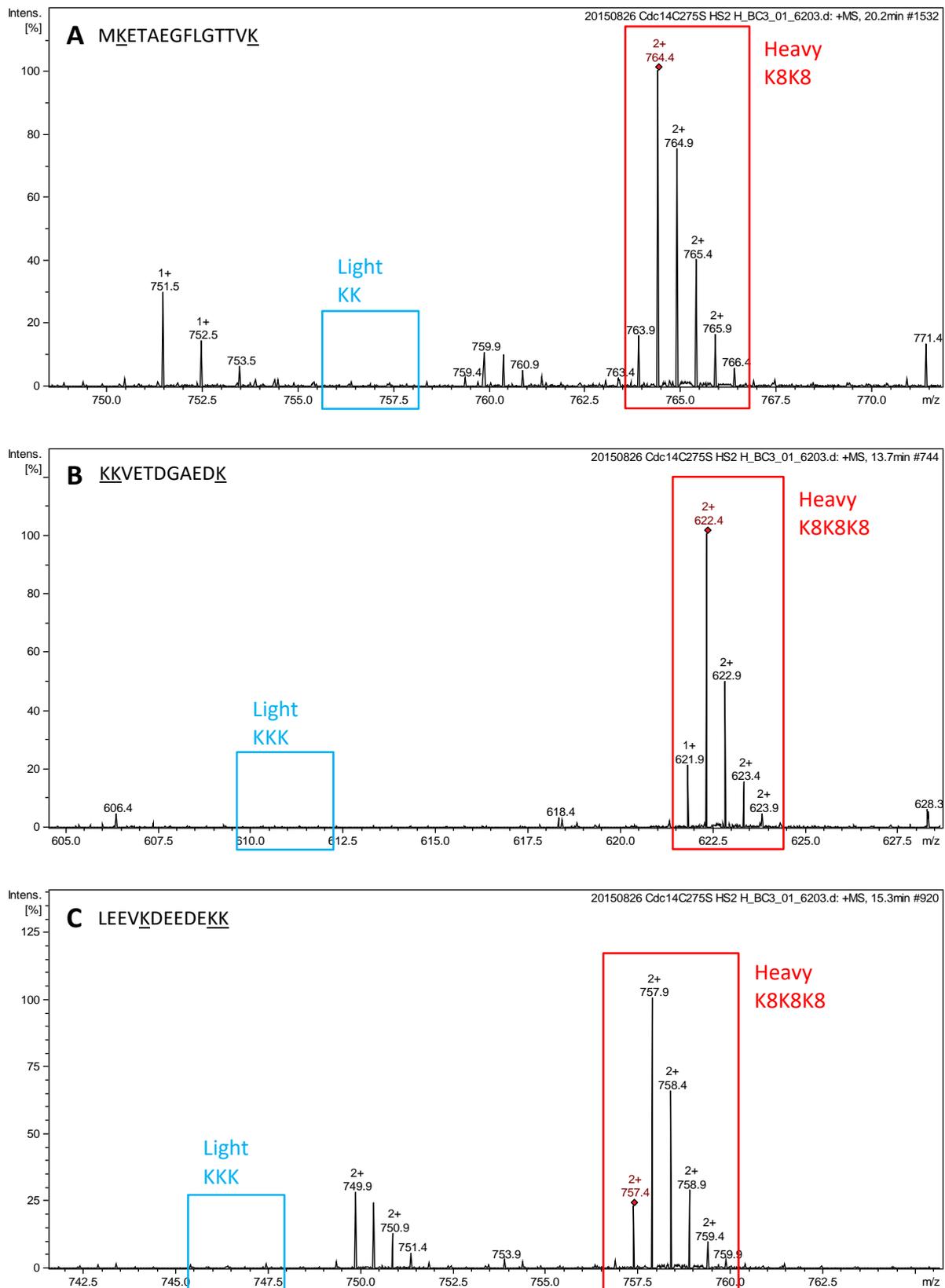


Fig. 5.3: Incorporation of Lys8 in hyphae. All three peptides here show full incorporation of heavy amino acids as no light analogues are present in the blue squares. Figure annotation is the same as fig 5.2.

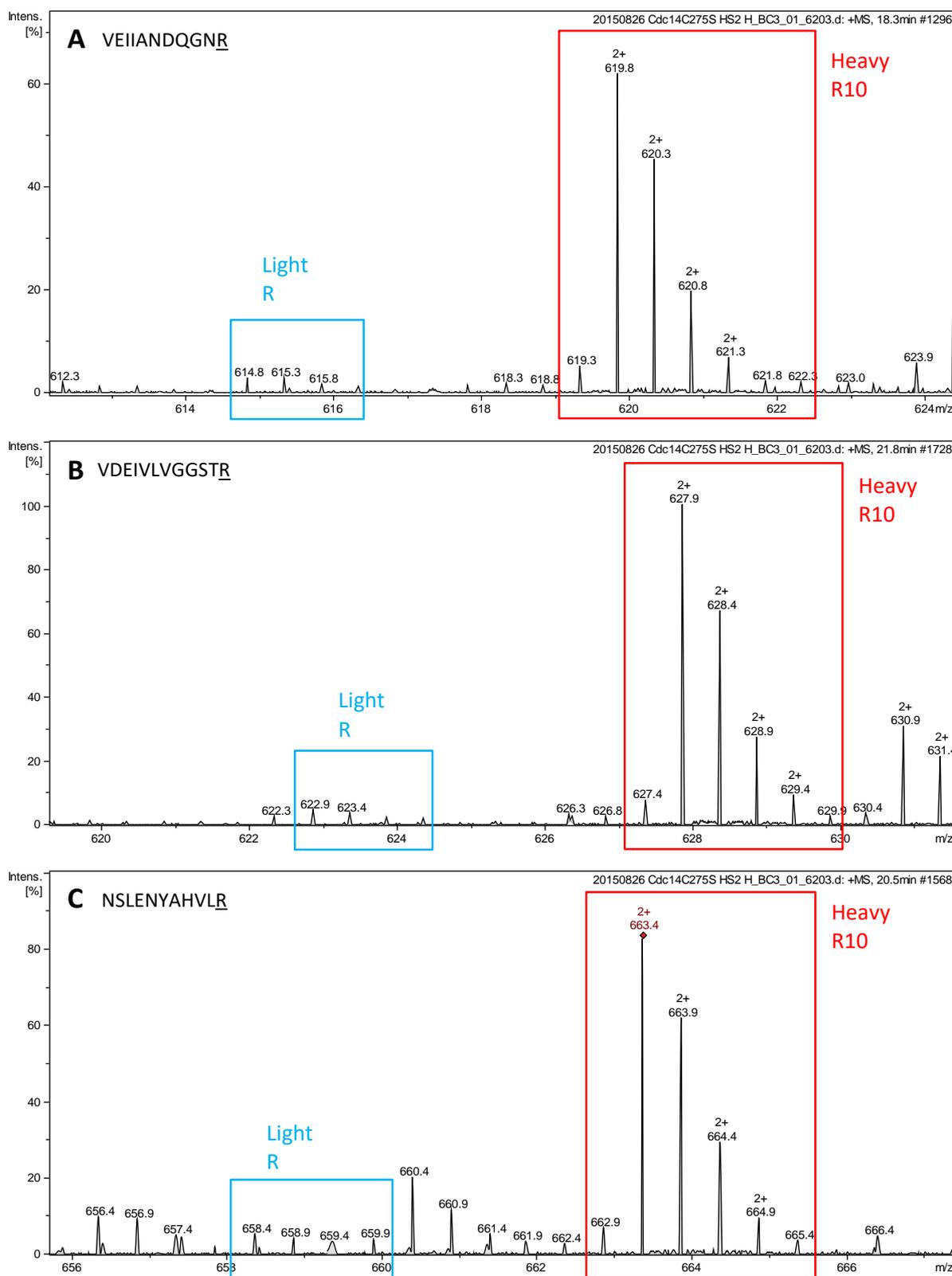


Fig 5.4: Incorporation of Arg10 in hyphae. Labelling with Arg10 was less efficient than that with Lys8. Nevertheless, labelled peptides showed significantly higher intensity than the light ones. Here the intensities of the light peaks are 2.89% (A), 4.80% (B) and 5.18% (C), while the intensities of the heavy peaks are 61.79% (A), 100% (B) and 82.49% (C) respectively. Peptides incorporation rates are 95.5% (A), 95.4% (B) and 94.1% (C), which is 95% on average. Figure annotation is the same as fig. 5.2.

5.5. Examination of Arg10 to Pro6 conversion

Cellular metabolism of the labelled amino acid Arg10 may lead to production of labelled proline, i.e. Pro6 (¹³C5, ¹²N). To test if this is the case, the heavy-labelled protein sample described in section 5.3 was processed in MaxQuant with Pro6 as a variable modification. The results were as follows:

- Proline:

Total number	1147
Light Pro	930 (81.1%)
Heavy Pro6	217 (18.9%)

- Peptides:

Total number	1777
With Pro and/or Pro6	819 (46.1%)
With Pro6 only	186 (10.5%)
Without Pro6	1591 (89.5%)

- Proteins:

Total number	184
With Pro6	62 (33.7%)

The data shows that Arg10->Pro6 conversion has indeed occurred, which has caused almost a fifth of all proline residues to be labelled with heavy isotopes. Almost half of all tryptic peptides in the sample contained a proline amino acid with 10.5% containing Pro6. At the protein level, a third of all identified proteins have a Pro6-containing peptide.

So why does all of this matter? When using a heavy form of arginine, heavy isotope labels can be inserted into proline through arginine catabolism. If the stable isotope incorporated into proline is not considered, ratios of proline-containing light and heavy peptides can be incorrectly calculated, leading to a reduction in intensity of the isotopic labeled heavy peptide. As explained in figure 5.5, the intensity of heavy peptides is divided between several peak clusters, but the software is programmed to match one light ion to

one heavy ion, so it takes only the first cluster into account. This skews the H/L ratio of peptides in favour of the light ones. In this study proteins of interest are enriched in heavy peptides, so the additional Pro6 may create some false negatives, but not false positives.

5.6. Discussion

SILAC labelling is a powerful and accurate method to distinguishing between bait-interacting proteins and contaminants in a complex mixture in conjunction with quantitative MS analysis. In order to use SILAC for identifying Cdc14 interactors, a protocol for labelling of the *C. albicans* proteome was developed.

In the present study all MS samples were prepared by trypsin digestion and tryptic peptides universally contain an arginine or lysine residue at the carboxyl terminus (except the C terminal peptide of every protein). This project used the *C. albicans* strain MDL04, which requires supplementation with arginine, lysine and uridine. However, two of the selectable markers, arginine and uridine, were used to generate the mutant strain MET3-Cdc14^{PD}. This leaves only lysine as an essential amino acid. In order to label every peptide, cells have to grow in the presence of both heavy arginine and heavy lysine. Therefore, we tested the ability of *C. albicans* to incorporate both amino acids in a strain that can synthesise arginine. MET3-Cdc14^{PD} cells were grown in “heavy” medium overnight and re-inoculated into fresh “heavy” medium for either 4hrs (yeast) or 60-135 min (hyphae). This time should be sufficient, because most organisms require about 5-10 cell divisions to fully incorporate labelled amino acids. As expected, MS analysis showed that Lys8 incorporation was very high (>95% in yeast) and it can be used in further SILAC experiments.

Arg10 incorporation was also high (almost 90% in yeast), which shows that cells use amino acids that are readily available to them even if they can make them endogenously. However, about 10% of all arginine was light. There are three possible reasons, why arginine peptides were not fully labelled. In the first scenario, cells had used up all of the Arg10 in the medium (possibly overnight) and had to switch back to endogenous production to meet their needs. A second possibility is that cells never fully turned off arginine biosynthesis, and so they used 90% from the medium and made 10% internally throughout the course of the experiment. However, the most likely explanation is that cells had enough arginine in the medium overnight and they achieved full incorporation, but when they reached stationary phase they restarted synthesis of light arginine. When cells were reinoculated into fresh medium, they started using the heavy arginine again, but could not replace all light isotopes within the given time. This speculation is based on a study by Martin-Perez and Villeñ

(2015), who showed that an *S. cerevisiae* prototroph reach full incorporation of Lys8 after 10 hrs, but after that, cells restarted endogenous lysine synthesis, regardless of the Lys8 availability in the medium (fig. 5.6). The authors suggested that nSILAC should be limited to exponentially growing cells for that reason. Therefore, to improve Arg10 incorporation it is proposed that cells are grown for 10 hours instead of overnight (about 16 hrs) and then reinoculated as described above, resulting in potentially higher isotope incorporation. The obvious caveat of this approach is that some work will have to be done at very inconvenient times during the night.

Ninety percent Arg10 incorporation is still satisfactory, so it was concluded that SILAC experiments can be performed using both Arg10 and Lys8. The incidence of Lys8 peptides was much higher than that of Arg10 peptides, so the average rate of isotope abundance was 93.87%. Thus the presented protocol here was applied in further SILAC experiments with the aim to identify Cdc14^{PD} interactors.

This study also examined the amount of arginine to proline conversion in *C. albicans*, which may be a source of error in quantitative SILAC experiments. Indeed, it was found that cells metabolised Arg10 to Pro6, so nearly a fifth of all proline in the proteome was labelled. There are several ways to prevent Arg10->Pro6 conversion. The easiest one is to add proline to the growth media, because as described above, cells tend to use what is available first. Alternatively, the amount of Arg10 may be reduced instead, because too much arginine stimulates proline production. However, this is not an option here, because low Arg10 availability may prompt the cells to revert to biosynthesis. Some studies report using strains with deleted genes involved in arginine catabolism, while others have taken a bioinformatics approach to account for the additional heavy Pro (Borek et al., 2015; Park et al., 2009). In summary, Arg10->Pro6 conversion may be prevented prior to MS or corrected after that. In this study we have determined the percent of stable isotope incorporated in proline, which may be corrected for computationally in downstream quantitative proteomic analysis.

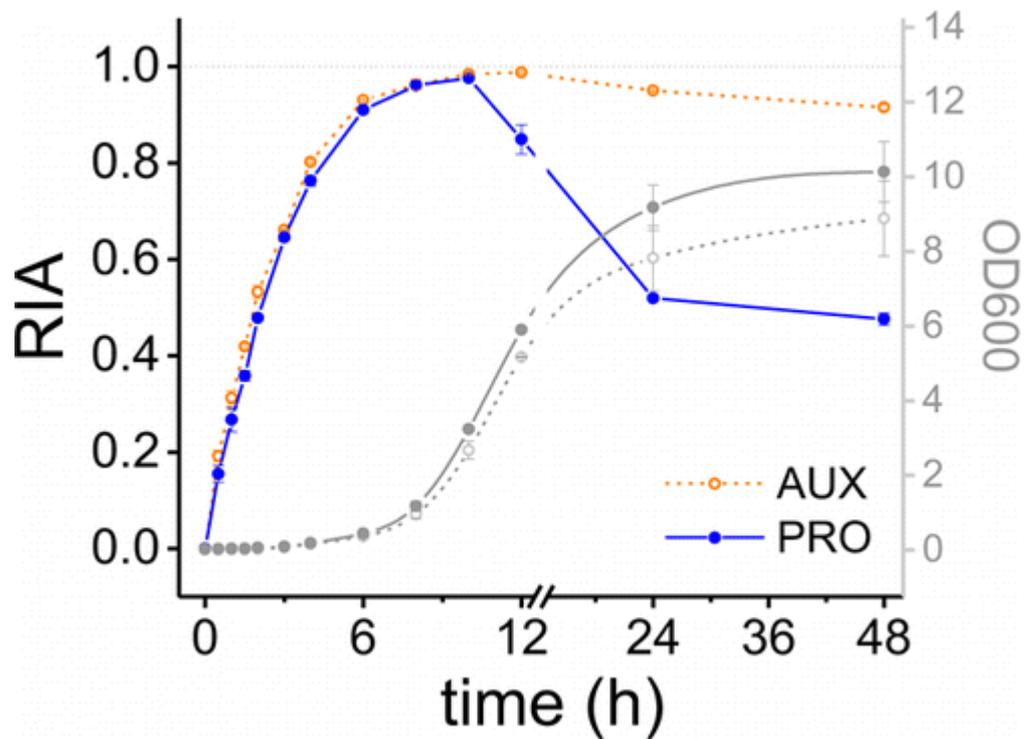


Fig. 5.6: Relative Lys8 abundance in *S. cerevisiae*. An auxotroph (AUX) and a prototroph (PRO) strain of *S. cerevisiae* was grown in the presence of Lys8 and the rate of isotope abundance (RIA) was followed for 48 hrs. Both strains achieved full incorporation at the same rate after 10 hrs of growth. However, longer incubation time resulted in a decay of RIA, mostly in the prototrophic strain. This figure is taken from Martin-Perez and Villeñ (2015).

Chapter 6

A Screen for Cdc14^{PD} Interacting Partners Using Quantitative SILAC-MS

6.1. Introduction

Following the extensive optimising and preparatory work done so far, this chapter returns to the main objective of this study – identifying protein interactions by MS. After taking on board the limitations of the initial preliminary screen described in chapter 3, several improvements to the experimental strategy were made. First, the interactions between Cdc14 and its substrates were stabilised by engineering the phosphatase-dead mutant Cdc14^{PD}. Second, Cdc14^{PD} was overexpressed in order to improve the signal-to-noise ratio in MS data. And third, a SILAC protocol in *C. albicans* was developed with the aim to make a better discrimination between interactors and contaminants.

Initially SILAC experiments were performed using a QTOF-MS, because it was the most sensitive instrument in the facility, where this project was carried out. However, at a later stage, an even better quadrupole orbitrap-MS (QO-MS) was purchased, and subsequent experiments were done with the new equipment. The data from both instruments was analysed using different software, so experimental results cannot be combined together. The QO-MS is superior in terms of sensitivity and resolution, so the data produced by it is of a higher quality.

It is important to note that the results presented here are not final. As shown in chapter 5, *C. albicans* has used the labelled arginine to synthesise labelled proline. The data shown here has not accounted for the heavy proline. An additional correction to the intensity of proline-containing peptides will be applied, but due to time constraints, it was not completed prior to submitting this thesis. Nevertheless, it is expected that the final results will have only minor differences from the results described in this chapter.

6.2. Cdc14^{PD} interactors in yeast

The general workflow described in this chapter is shown in figure 1.6. Cdc14^{PD} was overexpressed with the use of *MET3* promoter and purified separately from yeast and from hyphae. The growth conditions and SILAC media are described in section 5.3. In all experiments, *MET3*-Cdc14^{PD} was grown in heavy medium, and MDL04 was grown in light medium. This means that proteins of interest are enriched in heavy isotopes. Cell pellets from both cultures were weighted and mixed in 1/1 ratio prior to IP and MS analysis. Therefore non-specific background proteins will be present in approximately a 1:1 ratio. Cell extracts and IP of Cdc14^{PD} were prepared as described in chapter 3. MS sample preparation is described in section 5.3.

6.2.1. SILAC experiments in yeast using QTOF-MS

Cdc14^{PD} was immunoprecipitated from yeast three times and each experiment (termed Y1, Y2 and Y3) was run separately on QTOF-MS. The data from each run was analysed individually and then data from all three experiments was combined and analysed together using Mascot Distiller. This software performs peptide identification and quantitation, and then matches the identified peptides to a database of theoretical peptides in order to assign them to proteins. The false discovery rate (FDR) in all experiments was below 2%. Results are summarised in table 6.1 and presented in detail in figure 6.1 and table 6.3. The number of identified proteins varied between 448-1101. Over 70% of proteins were quantified and all data sets. Protein quantitation is reported as the L/H ratio. The L/H ratio of a protein is derived from the L/H ratios of all peptides assigned to it. If equal amounts of cells were mixed at the start of the experiment, contaminating proteins should have an L/H ratio close to 1. Cdc14^{PD}-bound proteins should be enriched in heavy peptides, so their L/H ratio should be <1. How much less than one is difficult to determine in empirical manner, so a cut off value of over two times enrichment in heavy peptides was taken as an arbitrary criterion. In other words, proteins that have two times more heavy peptides than light peptides are considered likely Cdc14^{PD} interactors, i.e. hits. Due to handling errors, cells could never be mixed in exactly 1/1 ratio, but a ratio between 0.5-2 was considered acceptable. In order to

	Y1	Y2	Y3	Y1+Y2+Y3	H1	H2	H1+H2
Identified proteins	717	448	804	1101	515	568	616
Quantified proteins	513	370	641	856	409	428	407
Median L/H ratio	1.22	1.1	1.15	1.19	0.64	0.77	0.69
L/H ratio of hits	<0.61	<0.55	<0.58	<0.59	<0.32	<0.39	<0.35
Hits	26	17	21	39	19	17	21

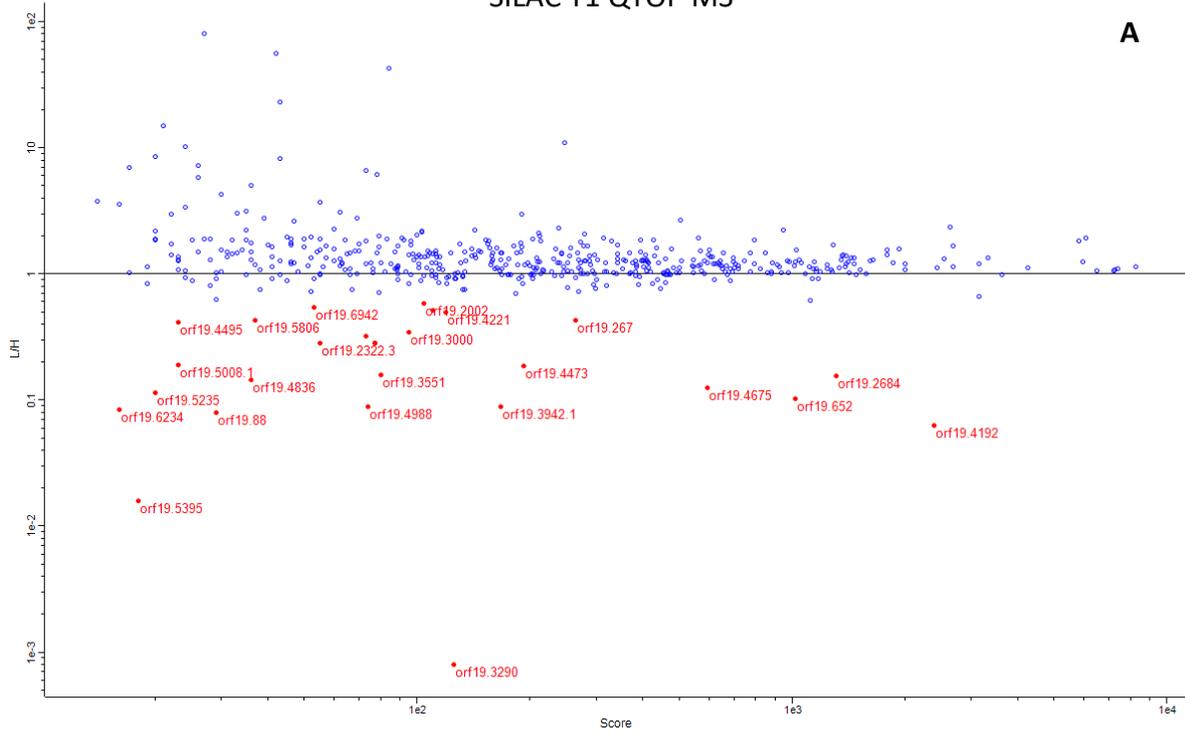
Table 6.1: Summary of SILAC data from all experiments performed by QTOF-MS.

	Y1	Y3	Y1+Y3	H2	H3	H2+H3
Identified proteins	2541	2228	2947	1831	1543	2085
Quantified proteins	2275	1844	2539	1512	1195	1795
Hits	77	44	82	51	47	55
Very low confidence hits	31	27	28	63	8	59

Table 6.2: Summary of SILAC data from all experiments performed by QO-MS.

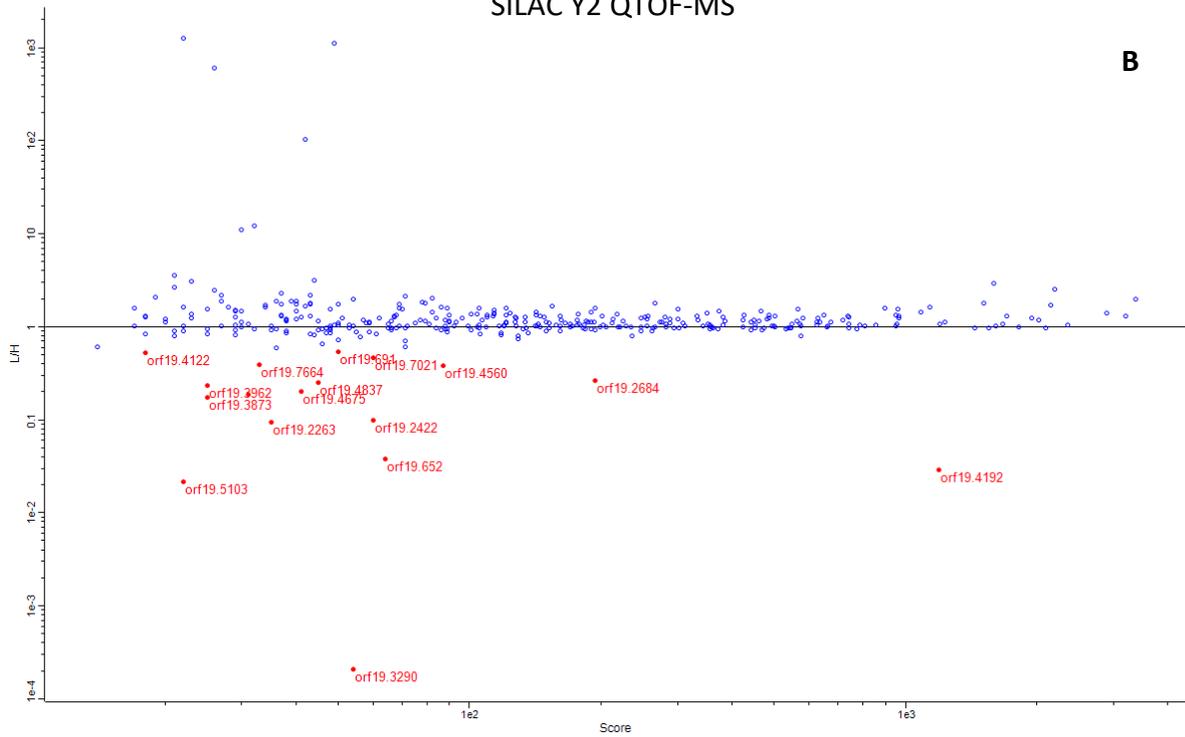
SILAC Y1 QTOF-MS

A



SILAC Y2 QTOF-MS

B



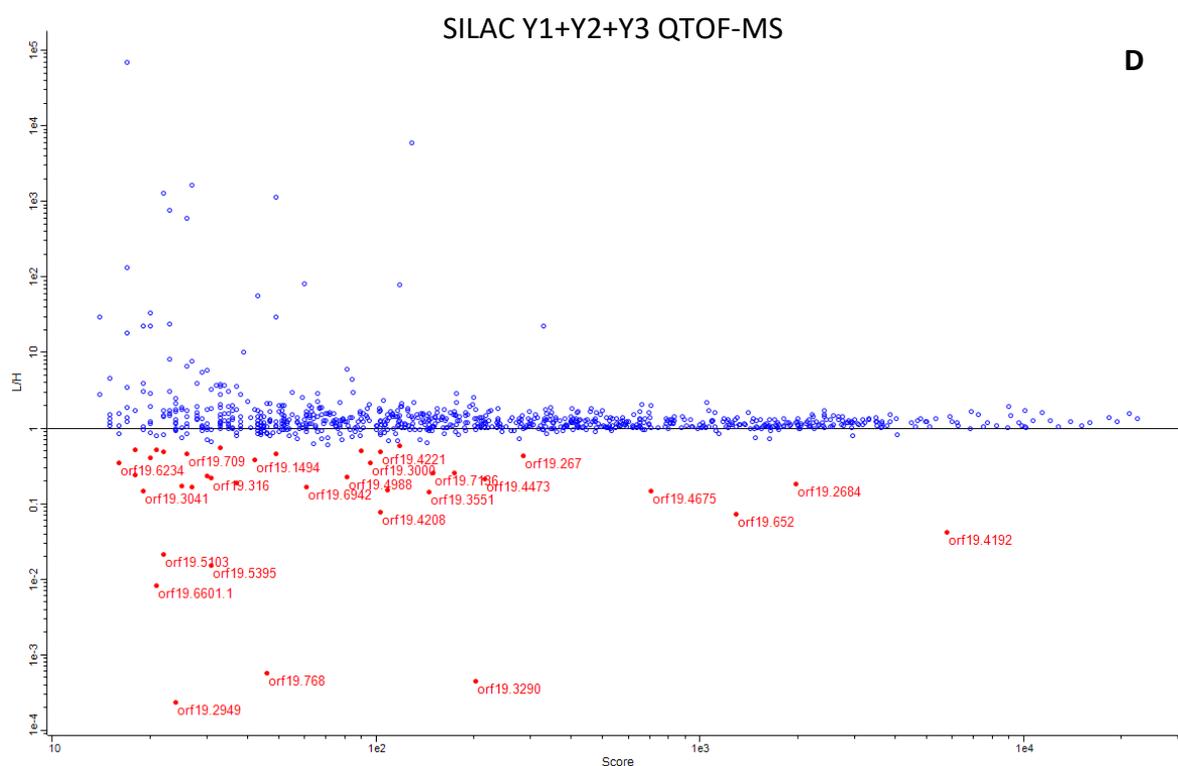
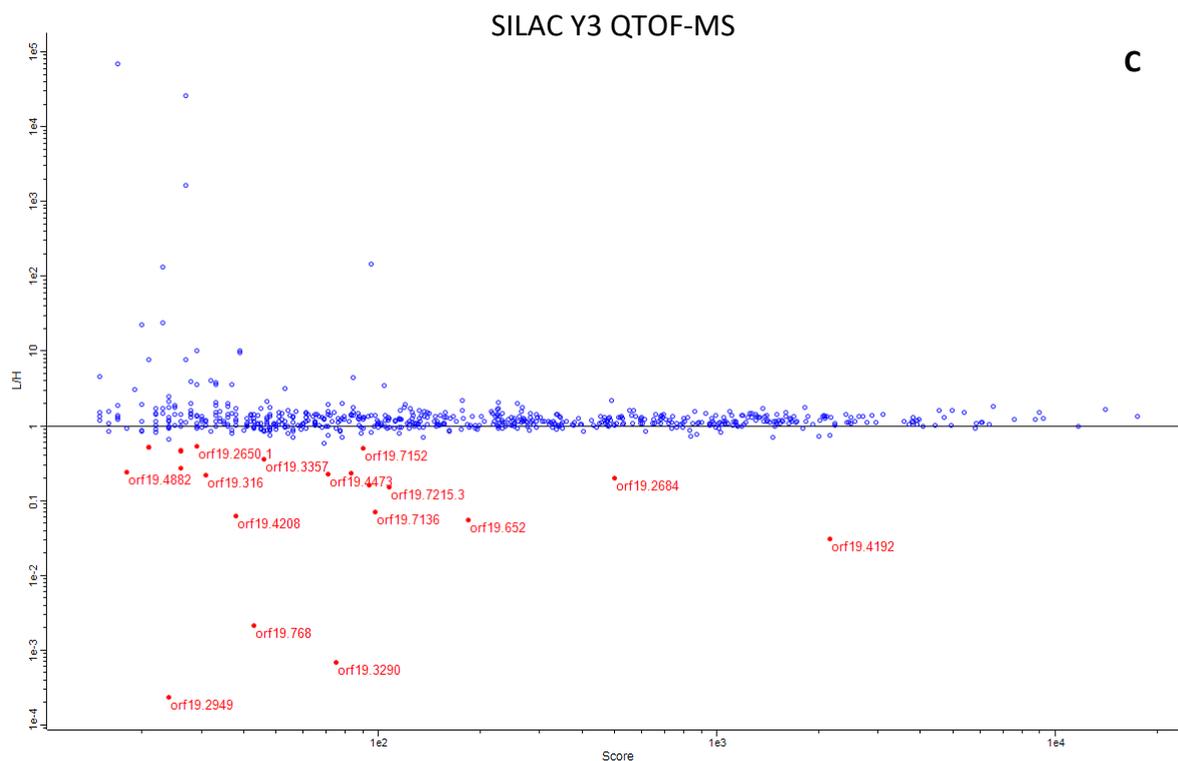


Fig. 6.1: SILAC results from QTOF-MS experiments in yeast. Each dot represents a protein identified by MS. The L/H ratio of proteins is plotted against their score. Proteins with over two times enrichment in heavy peptides are shown in red. Proteins that are considered to be contaminants are marked in blue. Note that only some of the red dots are labelled due to insufficient space on the graph. The full list of enriched proteins is shown in table 6.2.

Protein IDs	Gene	L/H ratio			
		Y1 (<0.61)	Y2 (<0.55)	Y3 (0.58)	Y1+Y2+Y3 (<0.59)
orf19.1340			0.19		
orf19.1494	RAD23				0.38
orf19.1515	CHT4	0.51			
orf19.2002		0.58			0.58
orf19.2263			0.10		
orf19.2322.3		0.28			
orf19.2422	ARC1		0.10		
orf19.2650.1				0.54	
orf19.267	NET1	0.43			0.43
orf19.2684		0.16	0.27	0.20	0.18
orf19.2949				0.00	0.00
orf19.3000	ORC1	0.35			0.35
orf19.3041					0.15
orf19.316	SEC13			0.22	0.22
orf19.3290		0.00	0.00	0.00	0.00
orf19.3357				0.36	
orf19.3551	DAD2	0.16			0.14
orf19.3873	ARC40		0.17		0.17
orf19.3942.1	RPL43A	0.09			
orf19.3962	HAS1		0.23		
orf19.4122			0.52		0.52
orf19.4192	CDC14	0.06	0.03	0.03	0.04
orf19.4208	RAD52	0.32		0.06	0.08
orf19.4221	ORC4	0.49			0.48
orf19.4473	SPC19	0.19		0.23	0.21
orf19.4495	NDH51	0.41			0.49
orf19.4560	BFR1		0.38		
orf19.4675	ASK1	0.13	0.20	0.16	0.15
orf19.4836	URA1	0.14			0.55
orf19.4837	DAM1		0.25	0.23	0.25
orf19.4882				0.24	0.24
orf19.4988		0.09			0.22
orf19.5008.1	DAD1	0.19			0.19
orf19.5103			0.02		0.02
orf19.5235		0.12			0.17
orf19.5358		0.28			
orf19.5395		0.02			0.02
orf19.5806	ALD5	0.43			
orf19.5958	CDR2				0.46
orf19.6000				0.47	
orf19.6234		0.08			0.35
orf19.6417	TSR1				0.40
orf19.652		0.10	0.04	0.05	0.07

Protein IDs	Gene	L/H ratio			
		Y1 (<0.61)	Y2 (<0.55)	Y3 (0.58)	Y1+Y2+Y3 (<0.59)
orf19.6601.1	YKE2				0.01
orf19.691	GPD2		0.54		
orf19.6942	ORC3	0.54			0.17
orf19.7021	GPH1		0.47		
orf19.709	PUP2			0.27	0.46
orf19.7136	SPT6			0.07	0.26
orf19.7152				0.50	0.50
orf19.7215.3				0.15	0.15
orf19.7477	YRB1			0.53	0.53
orf19.7652	CKA1			0.46	
orf19.7664			0.39		
orf19.7672					0.23
orf19.768	SYG1			0.00	0.00
orf19.88	ILV5	0.08			

Table 6.3: Proteins identified as hits in each data set (yeast QTOF-MS). Values show the L/H ratio of each protein. Empty cells indicate that the protein either did not reach the minimum threshold (shown in brackets) or it was not identified in that experiment. Empty spaces in the Gene column mean that the gene has not been named in the Candida Genome Database.

account for errors in mixing, the median L/H ratio of all proteins was calculated (table 6.1). The cut off value for hits is the median L/H ratio divided by 2. Table 6.3 shows all protein hits and their respective L/H ratio. Proteins are shown with their ORF number instead of their names, because such is the format of the output files. The scatter plots in figure 6.1 show the L/H of all quantified proteins plotted against protein score.

In total, 56 proteins were identified as potential hits (excluding the bait). Cdc14 was recovered with high score and low L/H ratio, which is a positive sign that the experiment is working. However, apart from the L/H ratio of these proteins, other criteria can also be used to indicate the likelihood of proteins being true hits. It is evident from looking at the graphs, that low-scoring proteins have higher variation in their L/H ratio than high-scoring proteins. Thus, the former are less likely to be true hits than the latter, even though they are all enriched in heavy isotopes. Low-scoring proteins are generally represented by very few peptides, so they are more likely to have deviations in the L/H ratio just by chance.

Further indication of a protein being a true hit comes from comparing the results in table 6.3. Proteins that were repeatedly identified as hits in more than one experiment have the highest probability of being such (e.g. orf19.2684, orf19.652, Rad52, etc.). However, this does not mean that proteins identified only once are not true hits (e.g. Orc4, Pup2, Spt6, etc.). There are many reasons why a protein would not be identified by MS: it can be lost at any stage during the purification process, the sample preparation or during the data-dependent MS analysis. While reproducibility is certainly a good confirmation of results, the lack of it should not be regarded as a reason to discard hits. High reproducibility was not expected in only three experiments.

The information contained in the last column of table 6.3 can be used as a criterion for assigning a confidence tag to non-reproducible hits. Combining the data from all three experiments together allows the processing software to make more accurate calculations. For example, Cht4 was a hit in Y1, but not in any of the other experiments. When the data was combined, this protein no longer appeared as a hit, because the overall L/H ratio is higher than the threshold. Hence, this protein was most likely a false positive in Y1. On the other hand hits like Dad2, which is also only a hit in Y1, remain with low L/H ratio even after the combined analysis. Such proteins are high confidence hits.

In each experiment, over 95% of identified proteins could be clearly distinguished as contaminants based on their quantitative ratios. Moreover, the quantitative SILAC analysis enabled the identification of some strongly enriched proteins in the Cdc14^{PD} based on their H/L ratios. Between these two extremes, proteins have a narrow difference in their isotope contents and a clear line separating hits and contaminants does not exist. In QTOF-MS experiments, proteins with two times higher abundance of heavy versus light peptides were selected as hits. However, this is not to say that every one of those hits was bound to Cdc14^{PD} during the pull down procedure. The list of hits is highly enriched in Cdc14^{PD} interactors, but it will inevitably contain some non-specific proteins. The latter are false positives that may be enriched in heavy peptides for two reasons. First, their abundance may have been higher in the labelled strain either by chance, or as a consequence of genetic differences between both strains. Second, even if a protein is present at equal amounts in both strains, unequal peptide intensity may be detected by chance. This is more likely to happen if a protein has low peptide intensity. Peptide quantitation is less accurate when the signal of the peptide is low. This means that proteins with low intensity would be less confident hits than proteins with high intensity. Unfortunately, Mascot Distiller (software processing QTOF data) does not report protein intensities, so protein score would be the next best parameter to look at. The protein score is based on the probability of a random match between the theoretical and experimental data bases. While score and intensity are correlated, they are not directly proportional to one another. In this regard, low scoring proteins have lower probability of being true hits.

6.2.2. SILAC experiments in yeast using QO-MS

Tryptic peptides from Y1 and Y3 were also analysed by a QO-MS instrument. Again, data from each experiment was processed separately, as well as in combination together, this time using the software MaxQuant. This software reports the H/L ratio of proteins (not the L/H ratio like seen above), so now hits have values >1 . MaxQuant also have inbuilt algorithms for normalising the H/L ratio, so median value was not calculated here. The software also reports protein intensity, which was used to calculate the probability of proteins being hits. This was done in Perseus, a software that is designed to perform bioinformatic analysis of output from MaxQuant. In Perseus, hits were selected based on the H/L ratio relative to intensity using the Benjamini-Hochberg procedure with false discovery rate (FDR) below 5% and below 1% (as described by Cox and Mann, 2008).

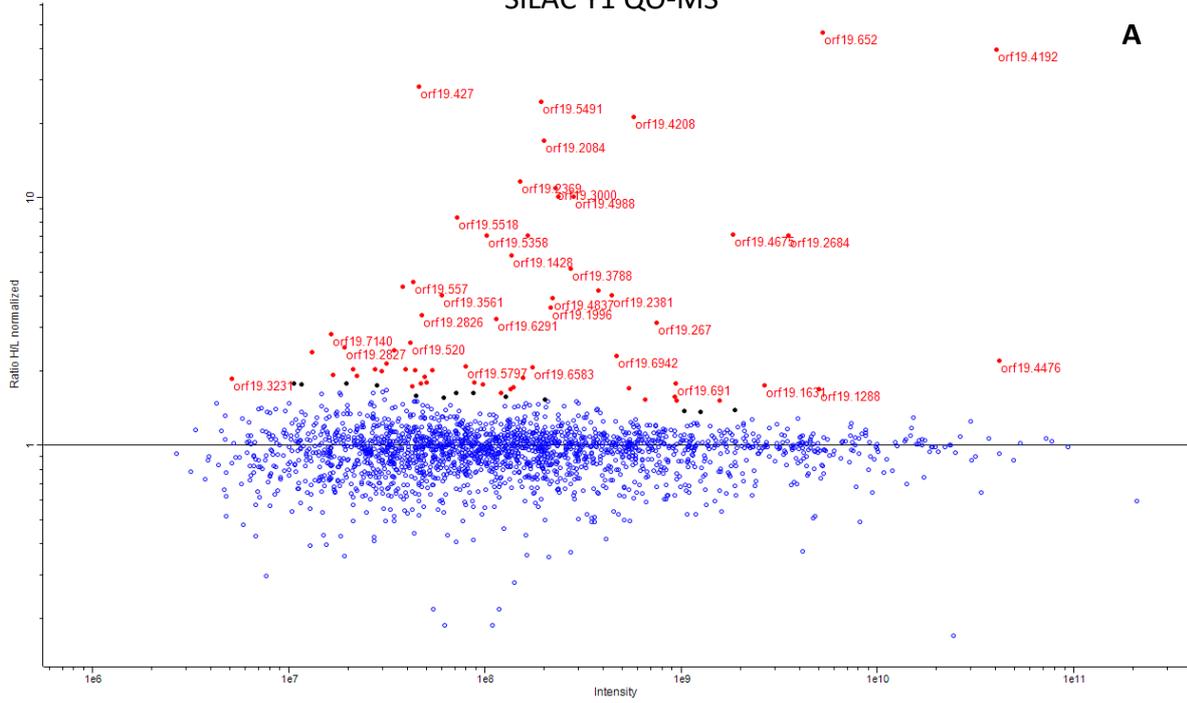
The QO instrument used in this study is one of the latest and most advanced mass spectrometers produced in recent times and it is also coupled to improved data analysis software, namely MaxQuant. Data derived by QO-MS is therefore higher quality than QTOF data. Results from QO-MS are summarised in table 6.2. QO-MS experiments in yeast resulted in the identification of over 2000 proteins. In comparison, QTOF data contained significantly less proteins per experiment (see table 6.1) than QO data. These numbers illustrate the huge difference in performance between the QTOF and QO instruments. Hits from the QO instrument are shown in figure 6.2 and table 6.4.

As expected, bait recovery was good and most proteins have H/L ratio close to 1. The QO data yielded two times more hits than the QTOF data, 115 in total (table 6.4). As discussed above, proteins found in all columns of table 6.4 have the highest probability of being true hits, while non-reproducible hits are questionable.

There is certainly a significant overlap between QTOF and QO data, but in the case of discrepancies, the QO data is regarded valid. For example, if a protein was identified as a hit by QTOF but not by QO in the same experiment, it is not a hit. Hits identified by both instruments have the highest probability of being true.

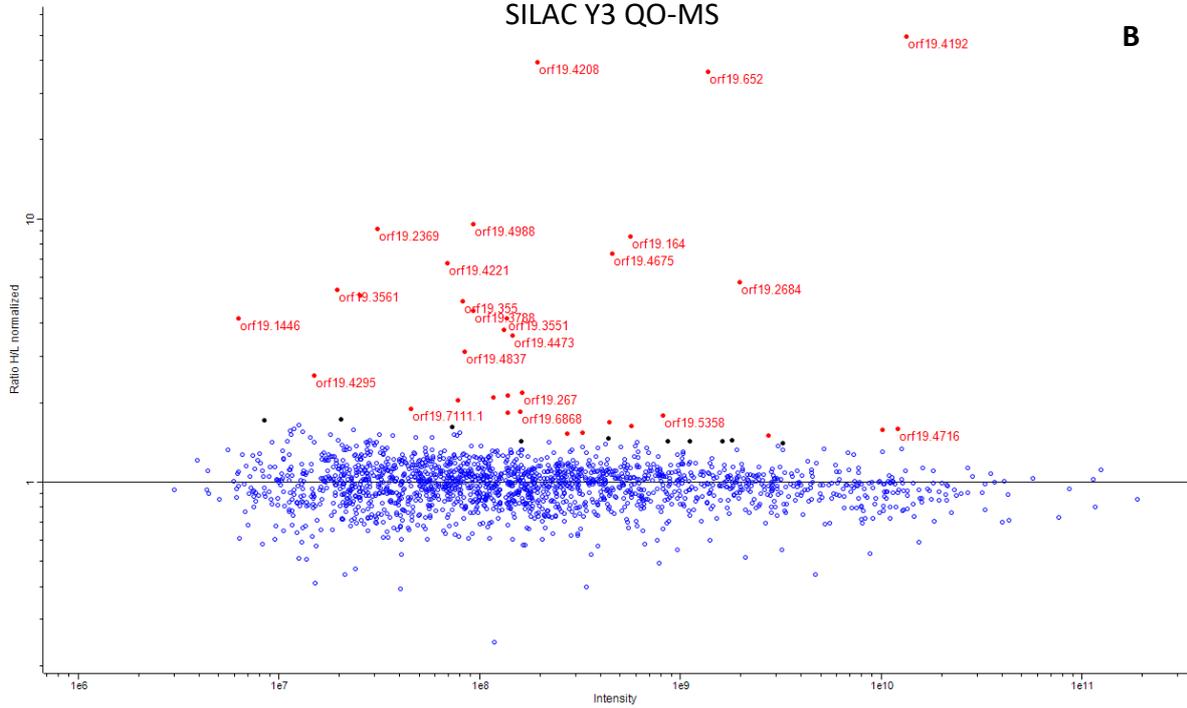
The selection of hits here was based on protein isotopes ratio as well as protein intensity. As shown in tables 6.4 and 6.7, the H/L ratio of some hits was as low as 1.2 (after

SILAC Y1 QO-MS



A

SILAC Y3 QO-MS



B

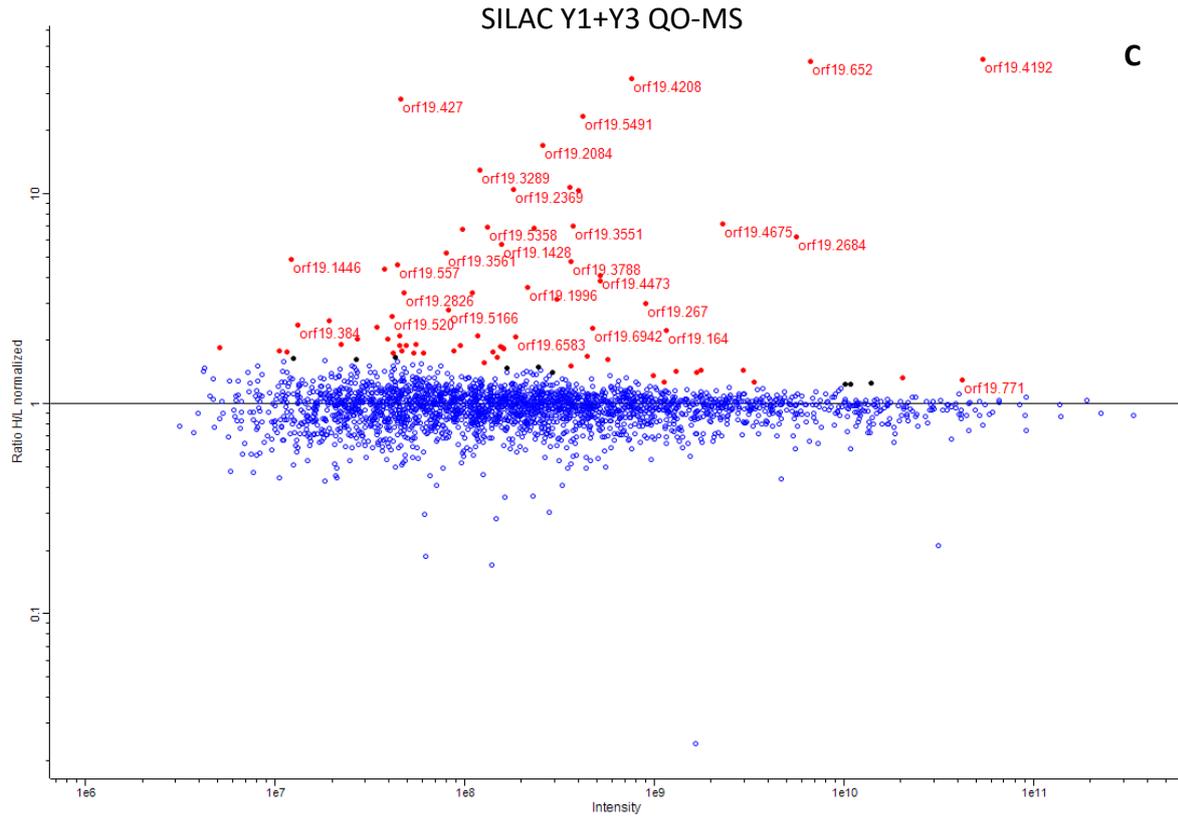


Fig. 6.2: SILAC results from QO-MS experiments in yeast. Blue proteins are contaminants. Red proteins are hits with FDR<1%. Black proteins are hits with FDR<5%. Parts A and B show data from two single experiments, while part C show the combined data from A and B analysed together.

Protein IDs	Gene	Y1		Y3		Y1+Y3	
		Ratio H/L (normalized)	FDR	Ratio H/L (normalized)	FDR	Ratio H/L (normalized)	FDR
orf19.1091						1.6254	<5%
orf19.113	CIP1			1.4971	<1%		
orf19.1153	GAD1			1.4303	<5%		
orf19.1223	DBF2	1.5703	<5%			1.5703	<1%
orf19.1267		1.5345	<5%				
orf19.1288	FOX2	1.6928	<1%				
orf19.1428	DUO1	5.8654	<1%			5.7418	<1%
orf19.1442	PLB4	2.0096	<1%			1.8839	<1%
orf19.1446	CLB2			4.197	<1%	4.8501	<1%
orf19.147	YAK1	1.734	<1%			1.734	<1%
orf19.1515	CHT4	1.5648	<1%			1.4188	<1%
orf19.1591	ERG10	1.6956	<1%				
orf19.1598	ERG24	1.5823	<5%				
orf19.1608				1.4301	<5%	1.3586	<1%
orf19.1631	ERG6	1.7492	<1%			1.4508	<1%
orf19.164				8.5791	<1%	2.2292	<1%
orf19.1652	POX1-3	1.5327	<1%				
orf19.1996	CHA1	3.5988	<1%			3.5988	<1%
orf19.2084	CDH1	16.977	<1%			16.977	<1%
orf19.2125				1.6317	<1%	1.6317	<1%
orf19.2369		11.63	<1%	9.1691	<1%	10.521	<1%
orf19.2381		4.0396	<1%	2.0437	<1%	3.8213	<1%
orf19.2389		1.6238	<5%				
orf19.2397		1.8769	<1%			1.8769	<1%
orf19.2400				1.7204	<5%		
orf19.2416.1	MLC1	1.9855	<1%				
orf19.2488	FAL1			1.6188	<5%		
orf19.2644	QCR2			1.43	<5%		
orf19.267	NET1	3.1376	<1%	2.1859	<1%	2.9978	<1%
orf19.2672	NCP1	1.5197	<1%			1.4135	<1%
orf19.2684		7.0405	<1%	5.7543	<1%	6.1973	<1%
orf19.2826		3.3565	<1%			3.3565	<1%
orf19.2827		2.4765	<1%			2.4765	<1%
orf19.3000	ORC1	10.971	<1%	3.7814	<1%	10.667	<1%
orf19.3014	BMH1	1.3864	<5%				
orf19.3139		1.6234	<5%				
orf19.3231	CDC27	1.8517	<1%			1.8517	<1%
orf19.3240	ERG27	1.6935	<1%			1.6645	<1%
orf19.3289						12.987	<1%
orf19.3311	IFD3			1.5724	<1%		
orf19.3356	ESP1	2.014	<1%			2.1129	<1%
orf19.3477		2.1341	<1%				
orf19.355				4.8789	<1%	3.3597	<1%

Protein IDs	Gene	Y1		Y3		Y1+Y3	
		Ratio H/L (normalized)	FDR	Ratio H/L (normalized)	FDR	Ratio H/L (normalized)	FDR
orf19.3551	DAD2	10.087	<1%	4.1752	<1%	7.043	<1%
orf19.3561	CDC7	4.0357	<1%	5.3986	<1%	5.2344	<1%
orf19.3684		1.6343	<1%				
orf19.3689						1.9167	<1%
orf19.3707	YHB1			2.1273	<1%	1.8295	<1%
orf19.3733	IDP2			1.5442	<1%	1.5049	<1%
orf19.3737		1.7414	<5%				
orf19.3788	SPC34	5.1598	<1%	4.4694	<1%	4.7585	<1%
orf19.384		2.3802	<1%			2.3802	<1%
orf19.4013		1.7812	<1%			1.7812	<1%
orf19.4025	PRE1			1.4293	<5%	1.4822	<5%
orf19.4157	SPS20	1.3707	<5%				
orf19.4192	CDC14	39.523	<1%	49.501	<1%	43.7	<1%
orf19.4208	RAD52	21.25	<1%	39.354	<1%	35.248	<1%
orf19.4221	ORC4	7.0215	<1%	6.7989	<1%	6.817	<1%
orf19.427		28.157	<1%			28.157	<1%
orf19.4295		1.7727	<5%	2.5392	<1%	2.3258	<1%
orf19.4341		1.7645	<5%			1.7645	<1%
orf19.4371	TAL1	1.3622	<5%				
orf19.4435		1.7689	<1%			1.7689	<1%
orf19.4473	SPC19	4.2083	<1%	3.6084	<1%	4.0749	<1%
orf19.4476		2.2008	<1%	7.3655	<1%		
orf19.4675	ASK1	7.1272	<1%			7.1891	<1%
orf19.4716	GDH3			1.5981	<1%	1.2539	<5%
orf19.4837	DAM1	3.9217	<1%	3.1348	<1%	3.1348	<1%
orf19.4960						1.6432	<5%
orf19.4988		10.134	<1%	9.5358	<1%	10.325	<1%
orf19.5005	OSM2	1.7133	<1%				
orf19.5166	DBF4					2.7929	<1%
orf19.5181	NIK1	2.0339	<1%				
orf19.520		2.5985	<1%			2.5985	<1%
orf19.5246		1.7909	<1%			1.7468	<1%
orf19.5276		1.8916	<1%			1.8916	<1%
orf19.5293		1.5601	<5%				
orf19.5358		7.0547	<1%	1.7848	<1%	6.9636	<1%
orf19.5389	FKH2	2.0378	<1%			2.0378	<1%
orf19.539	LAP3			1.5264	<1%	1.4166	<5%
orf19.5437	RHR2			1.4625	<5%		
orf19.5491		24.362	<1%			23.466	<1%
orf19.5518		8.3083	<1%	5.1215	<1%	6.7519	<1%
orf19.557		4.5763	<1%			4.5763	<1%
orf19.5727		1.9117	<1%			1.9117	<1%
orf19.5797	PLC2	2.0764	<1%				

Protein IDs	Gene	Y1		Y3		Y1+Y3	
		Ratio H/L (normalized)	FDR	Ratio H/L (normalized)	FDR	Ratio H/L (normalized)	FDR
orf19.5806	ALD5	1.5208	<1%				
orf19.5825	NCB2	1.9226	<1%				
orf19.5959	NOP14			1.7357	<5%		
orf19.6021	IHD2					1.6664	<5%
orf19.6155						1.7412	<1%
orf19.6254	ANT1	2.035	<1%			2.035	<1%
orf19.6257	GLT1			1.4094	<5%		
orf19.6291		3.2507	<1%			3.3709	<1%
orf19.6294	MYO1					1.2733	<1%
orf19.6385	ACO1					1.2314	<5%
orf19.6443		1.7819	<5%			1.7819	<1%
orf19.652		46.599	<1%	36.301	<1%	42.754	<1%
orf19.6583		2.0663	<1%			2.0803	<1%
orf19.6596				1.69	<1%	1.69	<1%
orf19.6610		1.7923	<1%			1.7923	<1%
orf19.6758				1.8268	<1%	1.5017	<5%
orf19.6868				1.849	<1%	1.849	<1%
orf19.691	GPD2	1.7831	<1%				
orf19.6942	ORC3	2.2915	<1%			2.2915	<1%
orf19.7111.1	SOD3			1.8992	<1%	1.8992	<1%
orf19.7140		2.8008	<1%				
orf19.7288						1.3306	<1%
orf19.7297						1.2742	<1%
orf19.7469	ARG1			2.0986	<1%	2.0986	<1%
orf19.7520	POT1	2.4198	<1%				
orf19.7600	FDH3			1.4373	<5%	1.4373	<1%
orf19.7663		4.3581	<1%			4.3581	<1%
orf19.771	LPG20					1.2961	<1%
orf19.827.1	RPL39					1.2363	<5%

Table 6.4: Proteins identified as hits in each dataset from yeast QO-MS. The table show the normalised H/L ratio, but note that the protein intensity was also factored in when selecting hits. Some proteins have H/L>2, which was the criterion for hits in the QTOF data. The Benjamini-Hochberg method used in Perseus is a more accurate way of selecting outliers in a dataset. If a protein has an H/L ratio that is significantly different from the ratios of other proteins with similar intensity, it is picked as a hit. The FDR next to the ratio indicates the chance of a protein being a false positive. If a protein was not an outlier or not present in the dataset, the space is left blank.

normalisation where 1 is the median intensity of the whole population), which is much lower than the ratio required for QTOF hits to be selected. Nevertheless, hits with low H/L ratio were found at high intensities, so the probability of this ratio occurring by chance is within the set limits (i.e. either 1% or 5%). This data analysis method, although being better, is by no means error-free, so as discussed above, hits should be regarded as only potential interactors. Hits may be sorted into confidence categories (e.g. low, medium, high) based on their probability of being true, when the data analysis is completed. (Readers are reminded that data will be subjected to further corrections to account for the Arg10->Pro6 conversion discussed in chapter 5.)

As already mentioned, the majority of proteins (>80%) identified by QO-MS were also quantified by MaxQuant using sophisticated algorithms. The remaining <20% were not assigned an H/L ratio by MaxQuant, even though they had intensity of light and heavy peptides. This is due to the settings of the software (e.g. minimum ratio count is set to 2, so proteins with a single ratio count will not be quantified, etc.). One could relax the settings, so that all proteins are quantified, but that would compromise the accuracy of quantitation. So, non-quantified proteins were further analysed manually and a number of them were selected as potential hits (table 6.5). These include proteins that had only heavy peptides (i.e. intensity of L=0, intensity of H>0) and proteins with calculated H/L>2. The calculated H/L is derived by simply dividing intensity H (the sum of all heavy peptides intensities) by intensity of L (the sum of all light peptides intensity). This is different from the H/L ratio assigned to proteins by MaxQuant, which is the median intensity of all individual peptide ratios. In addition to the calculated H/L values, proteins were selected only if they were identified by at least 2 peptides (razor+unique). This selection process is likely to refined further by looking at other parameters, before the final list of hits is composed. The most interesting entries in this category are proteins with exclusively or predominantly heavy peptides. In the final list of hits, these proteins may be presented as very low confidence hits (e.g. if they are present in multiple experiments), or they may be completely omitted (e.g. if they only appear once). These proteins are shown as very low confidence hits in table 6.2.

Protein IDs	Gene	H/L calculated (>2)		
		Y1	Y3	Y1+Y3
orf19.1219			2.29	
orf19.1282	CKS1		2.85	2.85
orf19.1331	HSM3		2.11	
orf19.1357	FCY21			3.07
orf19.1428	DUO1		12.89	
orf19.1441			2.05	
orf19.1446	CLB2	2.87		
orf19.1704	FOX3	2.11		
orf19.1738.1			2.75	2.75
orf19.1792		4.13		
orf19.2084	CDH1		36.32	
orf19.2301				2.40
orf19.2384		3.66		2.35
orf19.2630	RAD51			H
orf19.2852		3.16		
orf19.3289		15.80		
orf19.3296		H		73.66
orf19.3615				4.62
orf19.3695		2.24		2.24
orf19.3809	BAS1	H		H
orf19.3871	DAD3	H	H	H
orf19.3901			8.70	
orf19.4101		H		H
orf19.4109	PMT4		3.67	
orf19.4161		H		
orf19.4262			2.17	
orf19.4275	RAD9	H	H	H
orf19.4312			2.92	
orf19.4432	KSP1		4.26	
orf19.4937	CHS3		2.89	
orf19.5008.1	DAD1			6.64
orf19.5166	DBF4	82.25		
orf19.5397		5.31		5.31
orf19.5730			4.14	4.14
orf19.5759	SNQ2		6.20	
orf19.5773		4.23		
orf19.580		12.89		15.77
orf19.6030		H		H
orf19.6049		136.68	H	158.86
orf19.6124	ACE2	3.09		3.09
orf19.6155			5.15	
orf19.6195			8.26	
orf19.6346		2.55		
orf19.6376	PTC5		5.16	

Protein IDs	Gene	H/L calculated (>2)		
		Y1	Y3	Y1+Y3
orf19.638	FDH1		H	H
orf19.6536	IQG1	8.10		8.10
orf19.658	GIN1	H		H
orf19.6662		H		H
orf19.6689	ARG4		H	
orf19.6734	TCC1		7.32	
orf19.6796			H	
orf19.6838		5.75		
orf19.6869		H		H
orf19.6880				5.39
orf19.6884	GWT1	2.60		2.60
orf19.6941			H	
orf19.6990		2.10		
orf19.7060		H		H
orf19.926	EXO1	H		
orf19.927			2.68	
orf19.986	GLY1	4.51		

Table 6.5: Non-quantified proteins enriched in heavy peptides (experiments from yeast QO-MS).

The H/L ratios presented here should not be compared to the H/L ratios in table 6.3, because both values were determined in a different manner. Blank spaces indicate that the protein was either quantified by MaxQuant or not identified in the dataset at all. Note that some proteins in this table are also present in table 6.3, but in different columns. For example, Clb2 was manually selected in Y1 (here), but it was selected by MaxQuant in Y3 and Y1+Y3 (table 6.3). Dbf4 showed an impressive bias towards heavy peptides in Y1 (here), but it was quantified by MaxQuant only in the combined dataset (table 6.3). Dad3 and Rad9 were never quantified by the software, but are very likely to be hits because they were represented exclusively by heavy peptides (H) in all 3 datasets here.

6.3. Cdc14^{PD} interactors in hyphae

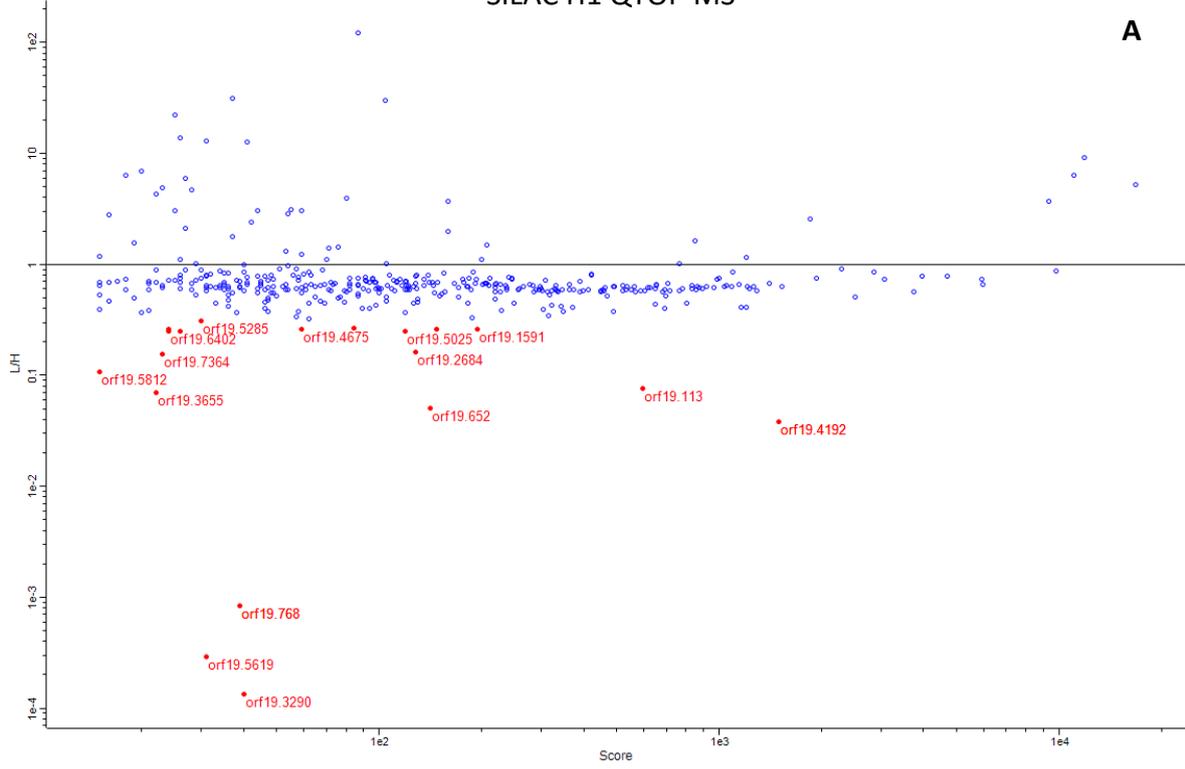
Cdc14^{PD} was immunoprecipitated from hyphae induced for 60-135 min, as described in section 5.2. The experiment was repeated three times (H1, H2 and H3). Samples from H1 and H2 were run through a QTOF-MS, while samples from H2 and H3 were analysed on a QO-MS.

Data from experiments H1 and H2 obtained from the QTOF-MS was processed separately and together. Over 500 proteins were identified and over 400 of them were quantified in each data set. Results are summarised in table 6.1 and shown in detail in figure 6.3 and table 6.6. Data was analysed as described in section 6.2.1, i.e. proteins with $L/H < (\text{median } L/H)/2$ are regarded as hits.

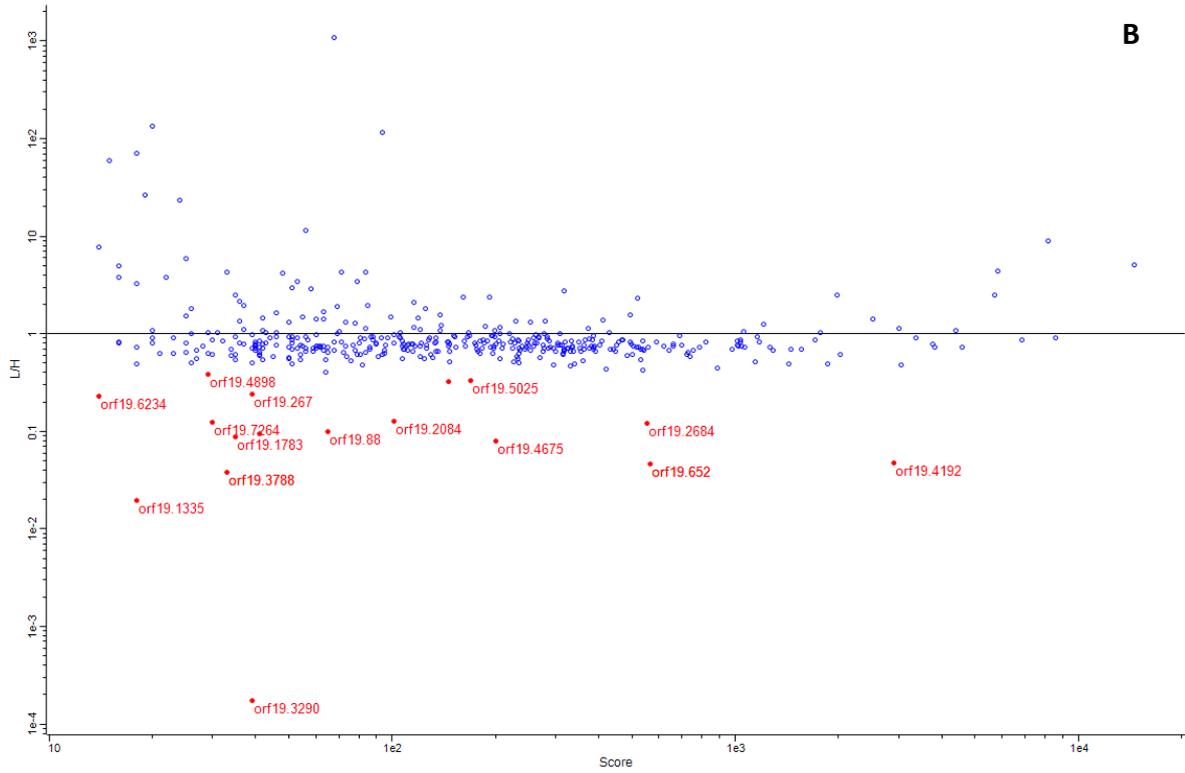
As discussed above, the QO-MS is better suited for SILAC analysis than QTOF-MS, so the data obtained by it is higher quality. QO-MS identified about three times more proteins than QTOF-MS. Results are shown in figure 6.4 and table 6.7. Low confidence hits from non-quantified proteins are listed in table 6.8.

Overall, fewer hits were found in hyphae than in yeast (tables 6.1 and 6.2). In hyphae QTOF data analysis found 40 hits (table 6.6), and QO data produced 97 hits (table 6.7). These include some common as well as some form-specific hits. Since only two experiments were done by QO-MS in each form, very few of the hits are reproducible in both experiments. However, some of the non-reproducible hyphal hits were found in yeast (and vice versa) and this can also be regarded as a reproducibility of results. For example, Cdh1 was a hit in Y1 and H2, but not in Y3 and H3 (it is actually among the low confidence hits in Y3, but these will be ignored in the current discussion). The fact that Cdh1 was selected twice as a hit, although in different morphological forms, suggests a high probability of interaction with Cdc14^{PD}.

SILAC H1 QTOF-MS



SILAC H2 QTOF-MS



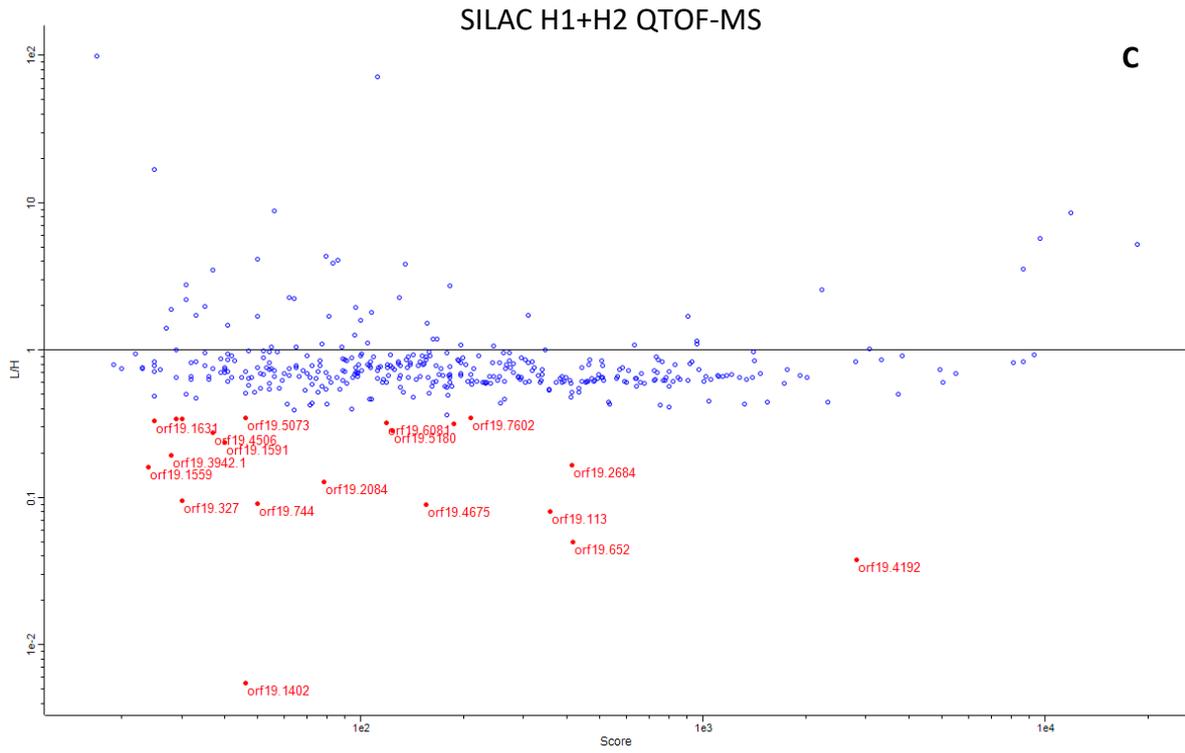


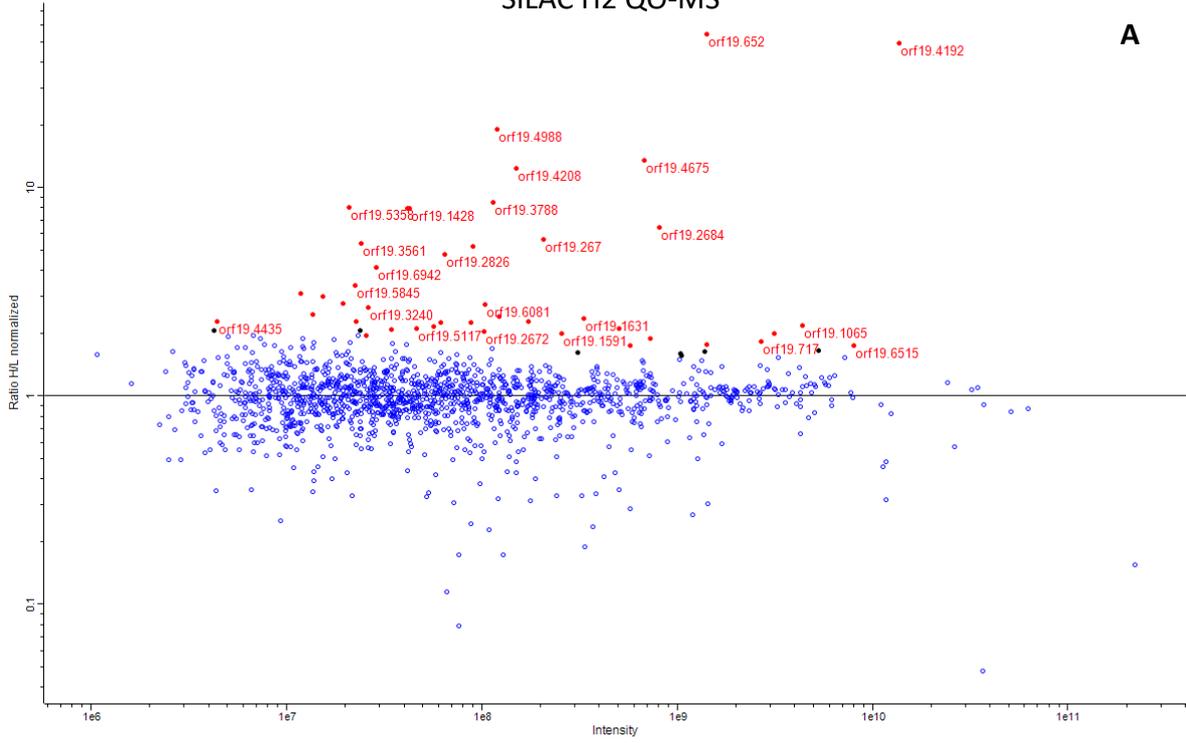
Fig. 6.3: SILAC results from QTOF-MS experiments in hyphae. Red – hits, blue – contaminants.

Protein IDs	Gene	L/H ratio		
		H1 (<0.32)	H2 (<0.39)	H1+H2 (<0.35)
orf19.113	CIP1	0.08		0.08
orf19.1335			0.02	
orf19.1402	CCT2			0.01
orf19.1559	HOM2			0.16
orf19.1591	ERG10	0.26		0.24
orf19.1631	ERG6			0.33
orf19.1783	YOR1		0.09	
orf19.2016				0.34
orf19.2084	CDH1		0.13	0.13
orf19.267			0.24	
orf19.2684		0.16	0.12	0.17
orf19.269	SES1			0.34
orf19.327	HTA3			0.10
orf19.3290		0.00	0.00	
orf19.3655		0.07		
orf19.3788	SPC34		0.04	
orf19.3942.1	RPL43A			0.19
orf19.4076	MET10	0.26		
orf19.4192	CDC14	0.04	0.05	0.04
orf19.4506	LYS22			0.27
orf19.4536	CYS4	0.26		
orf19.4675	ASK1	0.26	0.08	0.09
orf19.4898			0.39	
orf19.5008.1	DAD1		0.10	
orf19.5025	MET3	0.25	0.33	0.32
orf19.5073	DPM1			0.34
orf19.5180	PRX1	0.26		0.28
orf19.5285	PST3	0.31		
orf19.5619		0.00		
orf19.5812		0.11		
orf19.6081	PHR2	0.25	0.32	0.32
orf19.6234			0.23	
orf19.6402	CYS3	0.25		
orf19.652		0.05	0.05	0.05
orf19.7264			0.13	
orf19.7364		0.16		
orf19.744	GDB1			0.09
orf19.7602				0.35
orf19.768	SYG1	0.00		
orf19.88	ILV5		0.10	

Table 6.6: Proteins identified as hits in each data set (hyphae QTOF-MS). Values show the L/H ratio of each protein. Empty cells indicate that the protein either did not reach the minimum threshold

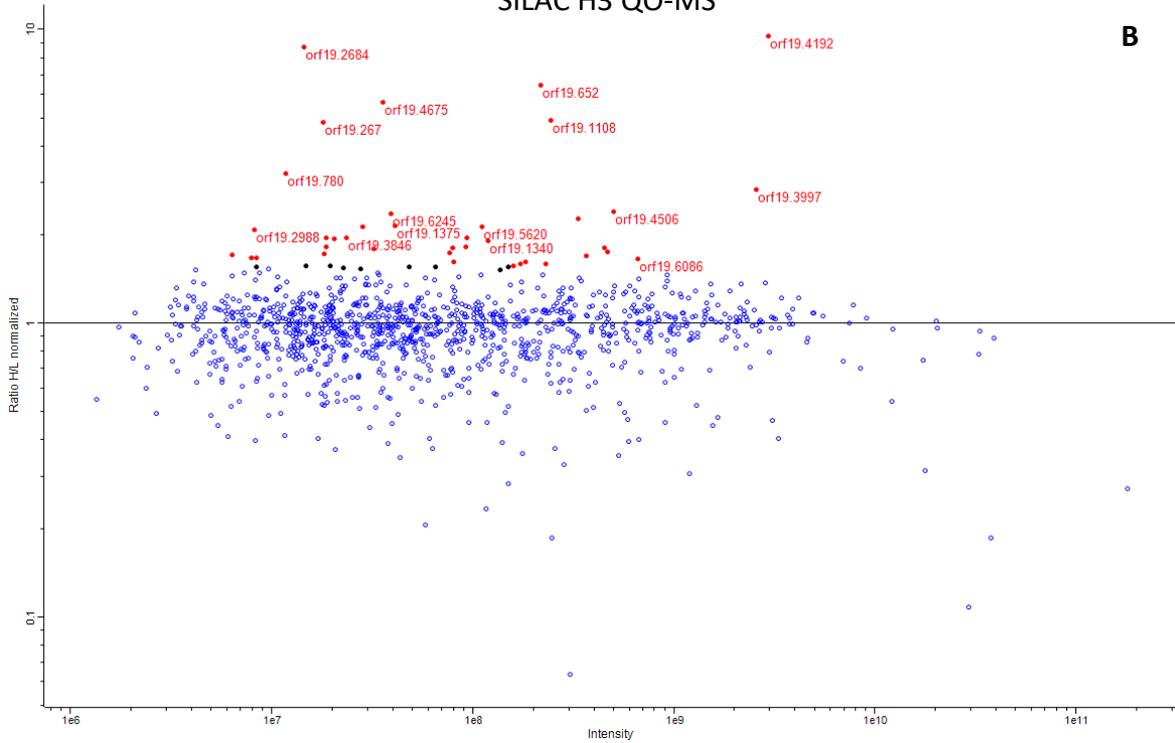
(shown in brackets) or it was not identified in that experiment. Empty spaces in the Gene column mean that the gene has not been named in the Candida Genome Database.

SILAC H2 QO-MS



A

SILAC H3 QO-MS



B

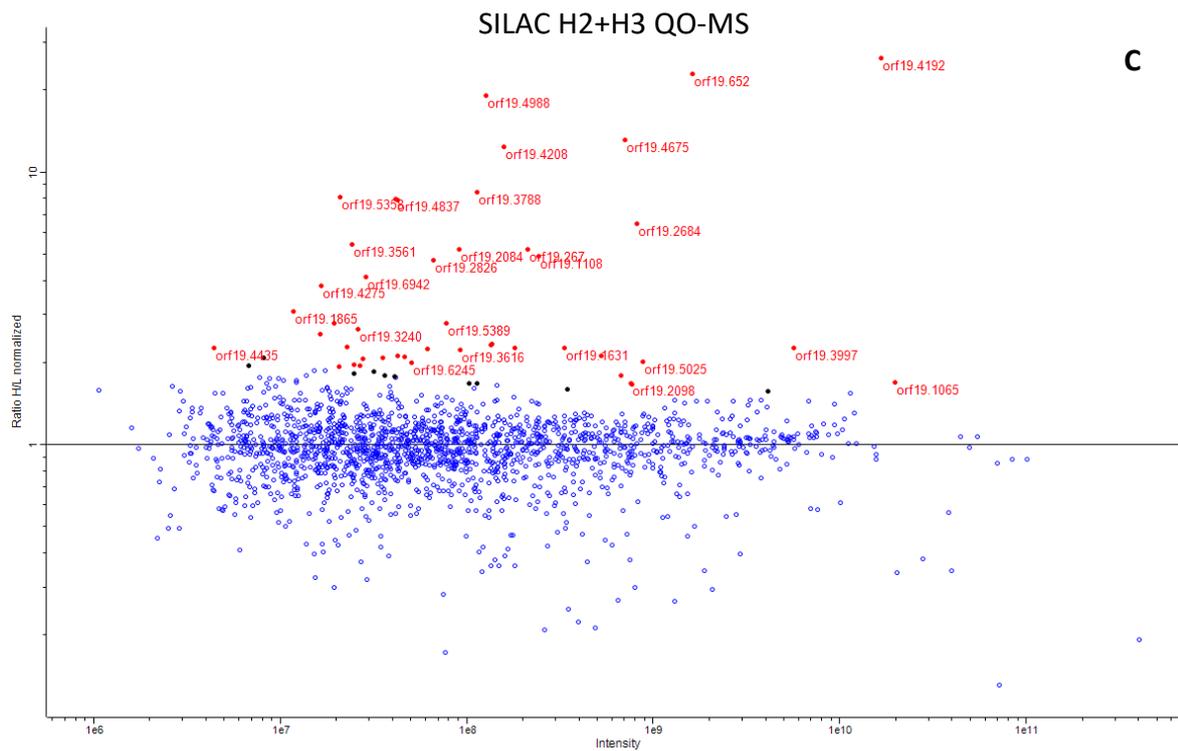


Fig. 6.4: SILAC results from QO-MS experiments in yeast. Red – hits with FDR<1%, black – hits with FDR<5%, blue – contaminants.

Protein IDs	Gene	HS2		HS3		HS2+HS3	
		Ratio H/L normalized	FDR	Ratio H/L normalized	FDR	Ratio H/L normalized	FDR
CaalfMp01	COX2			1.5579	<5%		
orf19.1065	SSA2	2.1844	<1%			1.6903	<1%
orf19.1108	HAM1			4.9168	<1%	4.9168	<1%
orf19.113	CIP1	1.6306	<5%				
orf19.1235	HOM3			1.6181	<1%		
orf19.1340				1.9064	<1%		
orf19.1375	LEU42			2.146	<1%	2.1128	<1%
orf19.1394				1.6711	<1%		
orf19.1415	FRE10	1.9462	<1%			1.7978	<5%
orf19.1428	DUO1	7.9148	<1%			7.9148	<1%
orf19.1467	COX13			1.566	<5%		
orf19.1515	CHT4	2.1621	<1%				
orf19.1564				1.5339	<5%		
orf19.1591	ERG10	1.9886	<1%				
orf19.1631	ERG6	2.347	<1%			2.2693	<1%
orf19.1801	CBR1					1.6803	<5%
orf19.1865		3.0906	<1%			3.0906	<1%
orf19.1986	ARO2			1.6225	<1%		
orf19.2028	MXR1			1.8038	<1%		
orf19.2084	CDH1	5.2064	<1%			5.2064	<1%
orf19.2098	ARO8	1.617	<5%	1.7503	<1%	1.6827	<1%
orf19.2389						1.7737	<5%
orf19.267	NET1	5.6399	<1%	4.8275	<1%	5.2122	<1%
orf19.2672	NCP1	2.0435	<1%				
orf19.2684		6.4516	<1%	8.7403	<1%	6.4516	<1%
orf19.2826		4.7624	<1%			4.7624	<1%
orf19.2847		2.9914	<1%				
orf19.2909	ERG26	2.448	<1%			2.5368	<1%
orf19.2988				2.0777	<1%	2.0777	<5%
orf19.3240	ERG27	2.6505	<1%			2.6505	<1%
orf19.3312						1.8227	<5%
orf19.346				1.5455	<5%		
orf19.3561	CDC7	5.4195	<1%			5.4195	<1%
orf19.3616	ERG9	2.2525	<1%			2.2241	<1%
orf19.3788	SPC34	8.4606	<1%			8.4606	<1%
orf19.3846	LYS4			1.9598	<1%	1.9598	<1%
orf19.3911	SAH1	1.7654	<1%				
orf19.3997	ADH1	1.9864	<1%	2.858	<1%	2.2645	<1%
orf19.4076	MET10	1.7524	<1%	1.9576	<1%	1.7958	<1%
orf19.4099	ECM17			1.8032	<1%		
orf19.4177	HIS5			1.7235	<1%		
orf19.4192	CDC14	49.394	<1%	9.5141	<1%	26.257	<1%
orf19.4208	RAD52	12.346	<1%			12.346	<1%

Protein IDs	Gene	HS2		HS3		HS2+HS3	
		Ratio H/L normalized	FDR	Ratio H/L normalized	FDR	Ratio H/L normalized	FDR
orf19.4220				1.6706	<1%		
orf19.4261	TIF5			1.7914	<1%		
orf19.4275	RAD9					3.8401	<1%
orf19.4435		2.2731	<1%			2.2731	<1%
orf19.449				1.9313	<1%	1.9313	<1%
orf19.4506	LYSS22			2.3886	<1%	2.1107	<1%
orf19.4675	ASK1	13.589	<1%	5.6727	<1%	13.094	<1%
orf19.4837	DAM1	7.9903	<1%			7.9903	<1%
orf19.4988		19.025	<1%			19.025	<1%
orf19.5025	MET3	2.1043	<1%	1.6937	<1%	2.0078	<1%
orf19.506	YDJ1	1.5707	<5%				
orf19.5117	OLE1	2.0981	<1%			2.0981	<1%
orf19.5131				1.7149	<1%		
orf19.5180	PRX1	1.8973	<1%				
orf19.5293				1.8131	<1%	1.6777	<5%
orf19.5358		8.0684	<1%			8.0684	<1%
orf19.5389	FKH2					2.7786	<1%
orf19.5484	SER1			1.5575	<5%		
orf19.5564				1.5577	<5%		
orf19.5614		2.2491	<1%			2.2491	<1%
orf19.5620				2.1288	<1%		
orf19.5811	MET1			1.5516	<5%		
orf19.5845	RNR3	3.3892	<1%				
orf19.5949	FAS2	1.6469	<5%				
orf19.6010	CDC5	2.07	<5%			2.07	<1%
orf19.6081	PHR2	2.7364	<1%			2.3358	<1%
orf19.6086	LEU4			1.6606	<1%	1.6606	<1%
orf19.6155						1.9545	<5%
orf19.6245				2.361	<1%	2.0026	<1%
orf19.6257	GLT1			2.2767	<1%		
orf19.6402	CYS3			1.5939	<1%		
orf19.6515	HSP90	1.7447	<1%				
orf19.652		54.679	<1%	6.4406	<1%	22.934	<1%
orf19.6524	TOM40			1.816	<1%		
orf19.6559				1.5725	<5%		
orf19.6583		2.7916	<1%			2.7916	<1%
orf19.6632	ACO2			2.1247	<1%	1.8576	<5%
orf19.6701		1.5939	<5%				
orf19.6758		2.2841	<1%			2.2841	<1%
orf19.6779	PRO2			1.7307	<1%		
orf19.6837	FMA1	2.2776	<1%			2.267	<1%
orf19.6942	ORC3	4.1333	<1%			4.1333	<1%
orf19.6994	BAT22			1.5155	<5%		

Protein IDs	Gene	HS2		HS3		HS2+HS3	
		Ratio H/L normalized	FDR	Ratio H/L normalized	FDR	Ratio H/L normalized	FDR
orf19.700	SEO1	2.0787	<1%			2.0787	<1%
orf19.717	HSP60	1.818	<1%			1.568	<5%
orf19.7297				1.5952	<1%		
orf19.7325	SCO1					1.9556	<1%
orf19.7498	LEU1			1.9464	<1%		
orf19.76	SPB1	2.0606	<5%				
orf19.7602		1.6089	<5%			1.5963	<5%
orf19.7602							
orf19.780	DUR1,2			3.2289	<1%		
orf19.922	ERG11	2.3954	<1%			2.3269	<1%
orf19.946	MET14			1.5689	<1%		

Table 6.7: Proteins identified as hits in each dataset from hyphae QO-MS.

Protein IDs	Gene	H/L calculated (>2)		
		H2	H3	H2+H3
CaalfMp08.1		3.85		
orf19.1133	MSB1			H
orf19.124	CIC1		2.63	
orf19.125	EBP1	10.25		40.95
orf19.1252	YME1	H		H
orf19.1519				H
orf19.1609		H		H
orf19.183	HIS3			3.36
orf19.2055	NPL6	6.10		
orf19.2115		H		3.14
orf19.2365	POL2			2.37
orf19.2381		H		H
orf19.2549	SHP1	3.92		
orf19.2770.1	SOD1		11.77	14.25
orf19.2782		73.50		
orf19.2827		2.99		2.36
orf19.2847				
orf19.2956	MGM101	H		
orf19.2973		H		2.42
orf19.2985				2.31
orf19.2989	GOR1			2.52
orf19.3000	ORC1	5.64		34.89
orf19.3040	EHT1	5.85		7.54
orf19.3103		3.19		
orf19.3296		H		H
orf19.3309		3.14		2.26
orf19.3474	IPL1	89.21		H
orf19.3540	MAK5	2.06		
orf19.3551	DAD2	H		4.89
orf19.3647	SEC8	3.56		
orf19.3802	PMT6			3.59
orf19.3871	DAD3	H		29.87
orf19.390	CDC42	2.26		
orf19.4013				6.66
orf19.406	ERG1			2.29
orf19.4131		H		6.44
orf19.4208	RAD52		H	
orf19.4221	ORC4	2.13		10.72

		H/L calculated (>2)		
Protein IDs	Gene	H2	H3	H2+H3
orf19.4257	INT1	12.77		
orf19.427		4.94		H
orf19.4275	RAD9	4.15		
orf19.4340		2.71		2.81
orf19.439		2.78		
orf19.4444	PHO15	4.86		6.15
orf19.4449				2.15
orf19.4473	SPC19	H	H	49.43
orf19.4563		H		
orf19.4829	DOA1	3.17		
orf19.489	DAP1	2.43		
orf19.5166	DBF4	3.29		73.50
orf19.5246		6.44		6.10
orf19.5389	FKH2	3.04		
orf19.5395		865.35		
orf19.541			H	
orf19.5491		25.75		H
orf19.5500	MAK16	2.25		2.30
orf19.5517		2.30		10.25
orf19.5518		10.72		H
orf19.5764	SKI8			2.09
orf19.58	RRP6	34.80		5.92
orf19.580		64.87		H
orf19.5852				5.42
orf19.6049		3.44		89.21
orf19.6124	ACE2	8.69		H
orf19.6291		H		9.38
orf19.6408		H		5.85
orf19.6418		4.20		3.67
orf19.643		H		H
orf19.6463		2.42		6.53
orf19.6506		H		2.99
orf19.658	GIN1	9.38		H
orf19.6689	ARG4		12.14	12.14
orf19.6692	MNN7	H		9.77
orf19.6809			H	H
orf19.6886			3.68	
orf19.7080	LEU2	14.39		

Protein IDs	Gene	H/L calculated (>2)		
		H2	H3	H2+H3
orf19.7107		H		2.43
orf19.7119	RAD3			3.45
orf19.7198		H		8.69
orf19.7215.3		5.02		
orf19.7322		7.54		
orf19.7394	GDA1	3.67		
orf19.748		2.68		
orf19.7538		6.15		5.64
orf19.926	EXO1			H
orf19.944	IFG3			21.99

Table 6.8: Non-quantified proteins enriched in heavy peptides (experiments from hyphae QO-MS).

6.4. Correction of H/L ratios of proline-containing peptides

As discussed in chapter 5, significant proportion of the heavy arginine was metabolically converted to heavy proline *in vivo*. Consequently, the H/L ratios of all proline-containing peptides were falsely estimated by the quantitation software. The true ratios of these peptides were calculated using a post-quantitation R script written by Dr Joseph Longworth (University of Sheffield). The script compares the median H/L ratios of peptides with different number of proline residues and estimates the contribution of each proline to the ratios (figure 6.5). In the example shown in figure 6.5A, each proline amino acid reduces the H/L ratio of a peptide by 35.7%, hence the correction factor is 0.357 per proline residue. After calculating the correction factor, the R application normalises the isotopic ratios of all peptides and re-calculates the correct protein ratios in the dataset. As evident in figure 6.5A, the peptide ratios appear very similar after the correction is applied. The change in protein ratios after the correction is not dramatic (figure 6.5B), but has a significant impact on the final list of hits.

All data sets obtained by QO-MS were corrected for arginine-to-proline conversion and outliers in each set were then determined as described in sections 6.2 and 6.3. The updated list of Cdc14 hits is shown in table 6.9. In total, 126 proteins were significantly enriched in labelled peptides, of which 117 were found in yeast and 43 were found in hyphae. Table 6.9 show the final list of Cdc14 hits that can be considered potential Cdc14 interactors. The criteria for selecting these proteins as hits were as follows:

- Proteins passed the FDR threshold of 0.05 in the Significance B test based on the Benjamini-Hochberg multiple hypothesis test.
- Proteins remain outliers when both data sets from each condition (i.e. yeast and hyphae) are combined and analysed together. For example, in tables 6.4 proteins may appear as hits in Y1 or Y2, but are missing when Y1+Y2 are combined. Such proteins will not fulfil this criterion, so they are not present in table 6.9.
- Proteins were identified by at least 2 peptides
- Proteins were enriched in heavy peptides in the cell lysate. For further information about this point, see section 6.5.

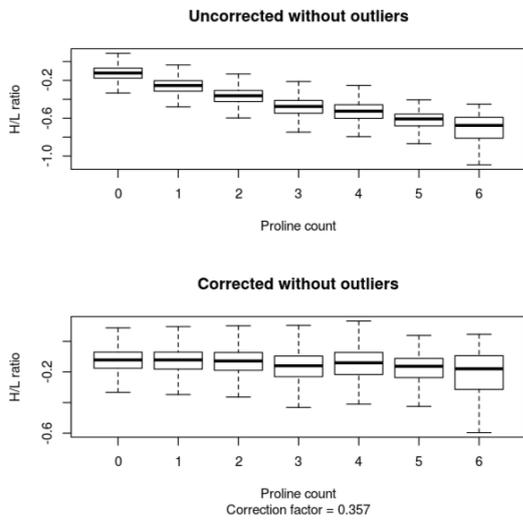
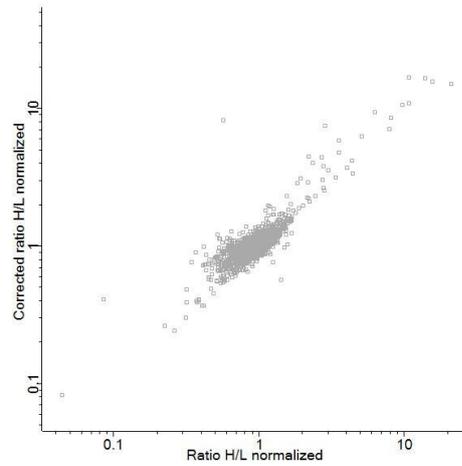
A**B**

Fig. 6.5: Change in peptide and protein isotopic ratios after arginine-to-proline correction. (A) In the example shown here, the peptide ratios decrease as the number of prolines in the peptides go up. The R script has calculated that each proline lowers the peptide ratio by 0.357 and applies this correction factor in order to produce a more accurate set of ratios, where peptides have similar ratios regardless of the proline count. **(B)** A comparison of the protein isotopic ratios before and after the correction show that the change is very subtle, i.e. most dots in the graph lie near a straight diagonal line crossing the 0 intercept.

Protein IDs	Gene	Yeast	Hyphae
orf19.1108	HAM1		+
orf19.1133	MSB1		+
orf19.1223	DBF2	+	
orf19.1357	FCY21	+	
orf19.1428	DUO1	+	+
orf19.1446	CLB2	+	
orf19.1451	SRB9	+	
orf19.147	YAK1	+	
orf19.1515	CHT4	+	+
orf19.1570	ERG27	+	
orf19.1598	ERG24	+	
orf19.1609			+
orf19.1618	GFA1	+	
orf19.1631	ERG6	+	
orf19.1792		+	
orf19.1801	CBR1	+	
orf19.1941	NUF2	+	
orf19.2084	CDH1	+	+
orf19.2369		+	
orf19.2381		+	+
orf19.2389		+	
orf19.2397		+	
orf19.2400		+	
orf19.2416.1	MLC1	+	
orf19.267	NET1	+	+
orf19.2672	NCP1	+	
orf19.2684		+	+
orf19.272	FAA21	+	
orf19.2826		+	+
orf19.2827		+	+
orf19.2909	ERG26	+	
orf19.3000	ORC1	+	+
orf19.3040	EHT1	+	
orf19.3091		+	
orf19.3221	CPA2	+	
orf19.3231	CDC27	+	
orf19.3240	ERG27	+	
orf19.3252	DAL81	+	
orf19.3289		+	
orf19.3296		+	
orf19.3311	IFD3	+	
orf19.3356	ESP1	+	
orf19.3362;orf19.2671		+	
orf19.3474	IPL1		+

Protein IDs	Gene	Yeast	Hyphae
orf19.3477		+	
orf19.3507	MCR1	+	
orf19.3535		+	
orf19.3551	DAD2	+	+
orf19.3561	CDC7	+	+
orf19.3684		+	
orf19.3707	YHB1	+	
orf19.3733	IDP2	+	
orf19.3788	SPC34	+	+
orf19.3809	BAS1	+	
orf19.3823	ZDS1	+	
orf19.384		+	
orf19.3856	CDC28	+	
orf19.3871	DAD3	+	
orf19.3954		+	
orf19.4013		+	
orf19.4043		+	
orf19.406	ERG1	+	
orf19.4101		+	
orf19.4192	CDC14	+	+
orf19.4208	RAD52	+	+
orf19.4221	ORC4	+	+
orf19.427		+	+
orf19.4275	RAD9	+	+
orf19.4311	YNK1	+	
orf19.4435		+	+
orf19.4473	SPC19	+	+
orf19.4675	ASK1	+	+
orf19.4716	GDH3	+	
orf19.4777	DAK2	+	
orf19.4837	DAM1	+	+
orf19.4988		+	+
orf19.5166	DBF4	+	+
orf19.520		+	
orf19.5246		+	
orf19.5276		+	
orf19.5358		+	+
orf19.5389	FKH2	+	+
orf19.5437	RHR2	+	
orf19.5491		+	+
orf19.5518		+	+
orf19.557		+	
orf19.5614			+
orf19.580		+	

Protein IDs	Gene	Yeast	Hyphae
orf19.5825	NCB2		+
orf19.6010	CDC5	+	+
orf19.6011	SIN3	+	
orf19.6026	ERG2	+	
orf19.6030		+	
orf19.6049		+	+
orf19.6124	ACE2	+	+
orf19.6155		+	
orf19.6234			+
orf19.6254	ANT1	+	
orf19.6257	GLT1	+	
orf19.6291		+	
orf19.638	FDH1	+	
orf19.6385	ACO1	+	
orf19.6408			+
orf19.643			+
orf19.6443		+	
orf19.6496	TRS33	+	
orf19.652		+	+
orf19.6536	IQG1	+	
orf19.658	GIN1	+	+
orf19.6583		+	+
orf19.6596		+	
orf19.6610		+	
orf19.6837	FMA1	+	
orf19.6868		+	
orf19.6882	OSM1	+	
orf19.6942	ORC3	+	+
orf19.7060		+	
orf19.7185		+	
orf19.7288		+	
orf19.7306		+	
orf19.7406		+	
orf19.7469	ARG1	+	
orf19.7663		+	
orf19.771	LPG20	+	
orf19.826		+	
orf19.926	EXO1	+	

Table 6.9: Final list of Cdc14 hits composed after correcting for arginine-to-proline conversion. The + indicates the protein was found as a hit in that experiment. If two ORFs are present in the Protein ID column, it means that the software could not determine, which one of them were present in the sample because they belong to the same protein family (i.e. they have the same tryptic peptides).

6.5. Measuring protein levels in the cell lysate

SILAC experiments often rely on the assumption that proteins are expressed equally in both cell cultures that are grown in different media. Thus, proteins enriched in either light or heavy peptides are regarded as outliers in the downstream data analysis. However, it is possible that the difference in growth conditions has caused differential expression of some proteins, which will appear as false positives.

In this study, SILAC experiments were carried out using a Cdc14^{PD} mutant, and a wild type strain as a control. Assuming that all proteins were present in the same amount in both starting cultures, those enriched in heavy peptides would appear so by physical interaction with the bait. Although cells expressing Cdc14^{PD} showed no phenotype suggesting metabolic disturbances, protein levels were formally assessed by mass spectrometry analysis of cell lysates. Cells of both strains were grown in the same condition used in a standard SILAC experiment, i.e. Cdc14^{PD} cells were labelled with heavy amino acids, and wild type cells were grown in light medium. Equal amounts of cells were mixed together and lysed, and protein extracts were separated by SDS-PAGE. Protein samples were then prepared and run through the mass spectrometer as described in chapter 2. Data was analysed as described in this chapter, and all proteins present in the samples were identified and quantified. This experiment, performed in both yeast and hyphae, reveals the protein ratios in the starting material before immunoprecipitating Cdc14^{PD}. In total, 2339 yeast protein and 2265 hyphal proteins were identified and quantified. The vast majority of those proteins had H/L ratios close to one. However, a few exceptions were also found, namely 131 proteins in yeast and 131 proteins in hyphae were significantly enriched in either light or heavy isotopes determined by using the Significance B test FDR=0.05. Proteins that were found to be enriched in the lysate and also in the IP experiments are listed in table 6.10. These proteins cannot be considered as Cdc14 interactors, because they were more abundant in the Cdc14^{PD} cells than the wild type cells. Thus, if they are non-specifically sticking to the matrix, heavy peptides will stick more often than light peptides. On the other side, these proteins cannot be completely excluded, because it is now known whether they associate with Cdc14 or not. Note that a direct comparison of the H/L ratios of these proteins in the IP and in the cell extracts cannot be made because isotope ratios vary greatly between experiments.

Protein IDs	Gene	Yeast	Hyphae
orf19.1065	SSA2		+
orf19.125	EBP1		+
orf19.1442	PLB4.5	+	
orf19.1608		+	
orf19.1631	ERG6		+
orf19.1779	MP65	+	
orf19.1801	CBR1		+
orf19.1865			+
orf19.1996	CHA1	+	
orf19.2124		+	
orf19.2125		+	
orf19.2417	SMC5	+	
orf19.2770.1	SOD1		+
orf19.3040	EHT1		+
orf19.3240	ERG27		+
orf19.355;orf19.6999		+	
orf19.3616	ERG9		+
orf19.3997;orf19.5113	ADH1		+
orf19.4076	MET10		+
orf19.4212;orf19.4213	FET99	+	
orf19.4295		+	
orf19.489	DAP1	+	+
orf19.5025	MET3		+
orf19.5180	PRX1		+
orf19.5517			+
orf19.5730		+	
orf19.5949	FAS2		+
orf19.6081	PHR2		+
orf19.6515	HSP90		+
orf19.6689	ARG4	+	+
orf19.6758		+	+
orf19.6837	FMA1		+
orf19.700;orf19.1855	SEO1		+
orf19.7111.1	SOD3	+	
orf19.717	HSP60		+
orf19.7600	FDH3	+	
orf19.7602			+
orf19.922	ERG11		+
orf19.979	FAS1		+

Table 6.10: Ambiguous proteins that were enriched in heavy isotopes in both the IP and the cell lysates.

6.6. Gene ontology analysis of potential Cdc14 interactors

The Cdc14 hits from table 6.9 were subjected to gene ontology (GO) analysis using the GO tools available on candidagenome.org. The aim is to group these proteins into categories based on their function or localisation in order to have a broader look at Cdc14 role in the cell. The query set for the analysis consisted of 125 putative Cdc14 interactors, while the background was all 2783 proteins identified by MS in all IP experiments. The results of the GO analysis are summarised in tables 6.11, 6.12 and 6.13.

The most common cellular processes identified here were related to cell cycle control, chromosome segregation, DNA metabolism, and organelle and cytoskeleton organisation. The most prevalent protein functions were DNA and cytoskeleton binding, and in agreement with this, the list was enriched in chromosomal and cytoskeletal proteins according to component analysis. These results suggest that in *C. albicans*, Cdc14 is actively involved in DNA maintenance not only during mitosis, but throughout the whole cell cycle. Based on this analysis, Cdc14 does not appear to have unique roles in *C. albicans*. However, there are many uncharacterised genes among the hits, which are not assigned any GO terms. Therefore, the significance of these interactions remain unknown.

GO term	Cluster frequency	Background frequency	FDR
single-organism cellular process	76 out of 125 genes, 60.8%	1231 out of 2783 background genes, 44.2%	0.12%
cell cycle process	42 out of 125 genes, 33.6%	255 out of 2783 background genes, 9.2%	0.00%
cell cycle	42 out of 125 genes, 33.6%	264 out of 2783 background genes, 9.5%	0.00%
mitotic cell cycle process	33 out of 125 genes, 26.4%	163 out of 2783 background genes, 5.9%	0.00%
mitotic cell cycle	33 out of 125 genes, 26.4%	166 out of 2783 background genes, 6.0%	0.00%
chromosome organization	31 out of 125 genes, 24.8%	246 out of 2783 background genes, 8.8%	0.00%
negative regulation of cellular process	30 out of 125 genes, 24.0%	318 out of 2783 background genes, 11.4%	0.02%
regulation of cell cycle	26 out of 125 genes, 20.8%	141 out of 2783 background genes, 5.1%	0.00%
regulation of cell cycle process	25 out of 125 genes, 20.0%	115 out of 2783 background genes, 4.1%	0.00%
regulation of cellular component organization	24 out of 125 genes, 19.2%	235 out of 2783 background genes, 8.4%	0.09%
DNA metabolic process	23 out of 125 genes, 18.4%	172 out of 2783 background genes, 6.2%	0.00%
regulation of organelle organization	21 out of 125 genes, 16.8%	159 out of 2783 background genes, 5.7%	0.00%
negative regulation of nucleobase-containing compound metabolic process	21 out of 125 genes, 16.8%	173 out of 2783 background genes, 6.2%	0.00%
nuclear chromosome segregation	20 out of 125 genes, 16.0%	63 out of 2783 background genes, 2.3%	0.00%
chromosome segregation	20 out of 125 genes, 16.0%	75 out of 2783 background genes, 2.7%	0.00%
regulation of mitotic cell cycle	20 out of 125 genes, 16.0%	92 out of 2783 background genes, 3.3%	0.00%
negative regulation of macromolecule biosynthetic process	20 out of 125 genes, 16.0%	181 out of 2783 background genes, 6.5%	0.12%
negative regulation of cellular macromolecule biosynthetic process	20 out of 125 genes, 16.0%	181 out of 2783 background genes, 6.5%	0.12%
negative regulation of RNA biosynthetic process	19 out of 125 genes, 15.2%	151 out of 2783 background genes, 5.4%	0.02%
negative regulation of nucleic acid-templated transcription	19 out of 125 genes, 15.2%	151 out of 2783 background genes, 5.4%	0.02%
negative regulation of transcription, DNA-templated	19 out of 125 genes, 15.2%	151 out of 2783 background genes, 5.4%	0.02%
negative regulation of RNA metabolic process	19 out of 125 genes, 15.2%	153 out of 2783 background genes, 5.5%	0.02%
cytoskeleton organization	18 out of 125 genes, 14.4%	125 out of 2783 background genes, 4.5%	0.00%
microtubule cytoskeleton organization	16 out of 125 genes, 12.8%	40 out of 2783 background genes, 1.4%	0.00%

GO term	Cluster frequency	Background frequency	FDR
positive regulation of cell cycle process	16 out of 125 genes, 12.8%	42 out of 2783 background genes, 1.5%	0.00%
microtubule-based process	16 out of 125 genes, 12.8%	43 out of 2783 background genes, 1.5%	0.00%
positive regulation of cell cycle	16 out of 125 genes, 12.8%	45 out of 2783 background genes, 1.6%	0.00%
nuclear division	16 out of 125 genes, 12.8%	73 out of 2783 background genes, 2.6%	0.00%
organelle fission	16 out of 125 genes, 12.8%	79 out of 2783 background genes, 2.8%	0.00%
regulation of cell cycle phase transition	15 out of 125 genes, 12.0%	61 out of 2783 background genes, 2.2%	0.00%
regulation of mitotic cell cycle phase transition	15 out of 125 genes, 12.0%	61 out of 2783 background genes, 2.2%	0.00%
regulation of chromosome segregation	14 out of 125 genes, 11.2%	37 out of 2783 background genes, 1.3%	0.00%
DNA-dependent DNA replication	14 out of 125 genes, 11.2%	63 out of 2783 background genes, 2.3%	0.00%
positive regulation of organelle organization	14 out of 125 genes, 11.2%	65 out of 2783 background genes, 2.3%	0.00%
DNA replication	14 out of 125 genes, 11.2%	67 out of 2783 background genes, 2.4%	0.00%
regulation of chromosome organization	14 out of 125 genes, 11.2%	75 out of 2783 background genes, 2.7%	0.00%
negative regulation of cell cycle	14 out of 125 genes, 11.2%	88 out of 2783 background genes, 3.2%	0.02%
positive regulation of cellular component organization	14 out of 125 genes, 11.2%	91 out of 2783 background genes, 3.3%	0.02%
attachment of spindle microtubules to kinetochore	13 out of 125 genes, 10.4%	19 out of 2783 background genes, 0.7%	0.00%
mitotic spindle organization	13 out of 125 genes, 10.4%	20 out of 2783 background genes, 0.7%	0.00%
spindle organization	13 out of 125 genes, 10.4%	23 out of 2783 background genes, 0.8%	0.00%
microtubule cytoskeleton organization involved in mitosis	13 out of 125 genes, 10.4%	24 out of 2783 background genes, 0.9%	0.00%
mitotic nuclear division	13 out of 125 genes, 10.4%	50 out of 2783 background genes, 1.8%	0.00%
sister chromatid segregation	13 out of 125 genes, 10.4%	53 out of 2783 background genes, 1.9%	0.00%
regulation of nuclear division	13 out of 125 genes, 10.4%	60 out of 2783 background genes, 2.2%	0.00%
negative regulation of gene expression, epigenetic	13 out of 125 genes, 10.4%	82 out of 2783 background genes, 2.9%	0.02%
chromatin silencing	13 out of 125 genes, 10.4%	82 out of 2783 background genes, 2.9%	0.02%
gene silencing	13 out of 125 genes, 10.4%	86 out of 2783 background genes, 3.1%	0.10%

GO term	Cluster frequency	Background frequency	FDR
regulation of microtubule-based process	12 out of 125 genes, 9.6%	19 out of 2783 background genes, 0.7%	0.00%
regulation of microtubule cytoskeleton organization	12 out of 125 genes, 9.6%	19 out of 2783 background genes, 0.7%	0.00%
nuclear DNA replication	12 out of 125 genes, 9.6%	33 out of 2783 background genes, 1.2%	0.00%
cell cycle DNA replication	12 out of 125 genes, 9.6%	34 out of 2783 background genes, 1.2%	0.00%
mitotic sister chromatid segregation	12 out of 125 genes, 9.6%	43 out of 2783 background genes, 1.5%	0.00%
regulation of mitotic nuclear division	12 out of 125 genes, 9.6%	43 out of 2783 background genes, 1.5%	0.00%
regulation of cytoskeleton organization	12 out of 125 genes, 9.6%	53 out of 2783 background genes, 1.9%	0.00%
negative regulation of mitotic cell cycle	12 out of 125 genes, 9.6%	56 out of 2783 background genes, 2.0%	0.00%
negative regulation of cell cycle process	12 out of 125 genes, 9.6%	68 out of 2783 background genes, 2.4%	0.02%
organic hydroxy compound metabolic process	12 out of 125 genes, 9.6%	69 out of 2783 background genes, 2.5%	0.02%
organic hydroxy compound biosynthetic process	11 out of 125 genes, 8.8%	45 out of 2783 background genes, 1.6%	0.00%
alcohol metabolic process	11 out of 125 genes, 8.8%	51 out of 2783 background genes, 1.8%	0.00%
DNA replication initiation	10 out of 125 genes, 8.0%	24 out of 2783 background genes, 0.9%	0.00%
regulation of sister chromatid segregation	10 out of 125 genes, 8.0%	30 out of 2783 background genes, 1.1%	0.00%
regulation of mitotic sister chromatid segregation	10 out of 125 genes, 8.0%	30 out of 2783 background genes, 1.1%	0.00%
positive regulation of mitotic cell cycle	10 out of 125 genes, 8.0%	33 out of 2783 background genes, 1.2%	0.00%
alcohol biosynthetic process	10 out of 125 genes, 8.0%	37 out of 2783 background genes, 1.3%	0.00%
double-strand break repair	10 out of 125 genes, 8.0%	51 out of 2783 background genes, 1.8%	0.07%
protein-DNA complex assembly	10 out of 125 genes, 8.0%	55 out of 2783 background genes, 2.0%	0.11%
positive regulation of chromosome segregation	9 out of 125 genes, 7.2%	14 out of 2783 background genes, 0.5%	0.00%
phytosteroid biosynthetic process	9 out of 125 genes, 7.2%	23 out of 2783 background genes, 0.8%	0.00%
secondary alcohol biosynthetic process	9 out of 125 genes, 7.2%	23 out of 2783 background genes, 0.8%	0.00%
cellular alcohol biosynthetic process	9 out of 125 genes, 7.2%	23 out of 2783 background genes, 0.8%	0.00%
ergosterol biosynthetic process	9 out of 125 genes, 7.2%	23 out of 2783 background genes, 0.8%	0.00%

GO term	Cluster frequency	Background frequency	FDR
cellular lipid biosynthetic process	9 out of 125 genes, 7.2%	23 out of 2783 background genes, 0.8%	0.00%
phytosteroid metabolic process	9 out of 125 genes, 7.2%	24 out of 2783 background genes, 0.9%	0.00%
secondary alcohol metabolic process	9 out of 125 genes, 7.2%	24 out of 2783 background genes, 0.9%	0.00%
ergosterol metabolic process	9 out of 125 genes, 7.2%	24 out of 2783 background genes, 0.9%	0.00%
sterol biosynthetic process	9 out of 125 genes, 7.2%	26 out of 2783 background genes, 0.9%	0.00%
steroid biosynthetic process	9 out of 125 genes, 7.2%	26 out of 2783 background genes, 0.9%	0.00%
cellular alcohol metabolic process	9 out of 125 genes, 7.2%	27 out of 2783 background genes, 1.0%	0.00%
sterol metabolic process	9 out of 125 genes, 7.2%	29 out of 2783 background genes, 1.0%	0.00%
chromatin silencing at silent mating-type cassette	9 out of 125 genes, 7.2%	30 out of 2783 background genes, 1.1%	0.00%
steroid metabolic process	9 out of 125 genes, 7.2%	31 out of 2783 background genes, 1.1%	0.00%
mitotic cell cycle checkpoint	9 out of 125 genes, 7.2%	35 out of 2783 background genes, 1.3%	0.00%
negative regulation of cell cycle phase transition	9 out of 125 genes, 7.2%	36 out of 2783 background genes, 1.3%	0.02%
negative regulation of mitotic cell cycle phase transition	9 out of 125 genes, 7.2%	36 out of 2783 background genes, 1.3%	0.02%
cell cycle checkpoint	9 out of 125 genes, 7.2%	41 out of 2783 background genes, 1.5%	0.04%
mitotic spindle organization in nucleus	8 out of 125 genes, 6.4%	8 out of 2783 background genes, 0.3%	0.00%
positive regulation of nuclear division	8 out of 125 genes, 6.4%	15 out of 2783 background genes, 0.5%	0.00%
regulation of G2/M transition of mitotic cell cycle	8 out of 125 genes, 6.4%	23 out of 2783 background genes, 0.8%	0.00%
regulation of cell cycle G2/M phase transition	8 out of 125 genes, 6.4%	23 out of 2783 background genes, 0.8%	0.00%
positive regulation of cytoskeleton organization	8 out of 125 genes, 6.4%	30 out of 2783 background genes, 1.1%	0.02%
regulation of DNA-dependent DNA replication	8 out of 125 genes, 6.4%	30 out of 2783 background genes, 1.1%	0.02%
regulation of DNA replication	8 out of 125 genes, 6.4%	31 out of 2783 background genes, 1.1%	0.02%
pre-replicative complex assembly involved in cell cycle DNA replication	7 out of 125 genes, 5.6%	18 out of 2783 background genes, 0.6%	0.00%
pre-replicative complex assembly	7 out of 125 genes, 5.6%	18 out of 2783 background genes, 0.6%	0.00%
pre-replicative complex assembly involved in nuclear cell cycle DNA	7 out of 125 genes, 5.6%	18 out of 2783 background genes, 0.6%	0.00%

GO term	Cluster frequency	Background frequency	FDR
replication			
mitotic DNA integrity checkpoint	7 out of 125 genes, 5.6%	22 out of 2783 background genes, 0.8%	0.02%
regulation of exit from mitosis	7 out of 125 genes, 5.6%	23 out of 2783 background genes, 0.8%	0.02%
positive regulation of attachment of spindle microtubules to kinetochore	6 out of 125 genes, 4.8%	7 out of 2783 background genes, 0.3%	0.00%
regulation of attachment of spindle microtubules to kinetochore	6 out of 125 genes, 4.8%	7 out of 2783 background genes, 0.3%	0.00%
metaphase plate congression	6 out of 125 genes, 4.8%	9 out of 2783 background genes, 0.3%	0.00%
mitotic metaphase plate congression	6 out of 125 genes, 4.8%	9 out of 2783 background genes, 0.3%	0.00%
establishment of chromosome localization	6 out of 125 genes, 4.8%	10 out of 2783 background genes, 0.4%	0.00%
positive regulation of mitotic nuclear division	6 out of 125 genes, 4.8%	11 out of 2783 background genes, 0.4%	0.00%
mitotic DNA replication	6 out of 125 genes, 4.8%	18 out of 2783 background genes, 0.6%	0.10%
negative regulation of mitotic nuclear division	6 out of 125 genes, 4.8%	19 out of 2783 background genes, 0.7%	0.11%
chromosome localization	6 out of 125 genes, 4.8%	19 out of 2783 background genes, 0.7%	0.11%
DNA double-strand break processing	5 out of 125 genes, 4.0%	5 out of 2783 background genes, 0.2%	0.00%
regulation of microtubule polymerization or depolymerization	5 out of 125 genes, 4.0%	7 out of 2783 background genes, 0.3%	0.00%
regulation of mitotic spindle organization	5 out of 125 genes, 4.0%	7 out of 2783 background genes, 0.3%	0.00%
regulation of spindle organization	5 out of 125 genes, 4.0%	8 out of 2783 background genes, 0.3%	0.00%
DNA replication checkpoint	5 out of 125 genes, 4.0%	10 out of 2783 background genes, 0.4%	0.02%
regulation of nuclear cell cycle DNA replication	5 out of 125 genes, 4.0%	11 out of 2783 background genes, 0.4%	0.07%
regulation of spindle pole body separation	4 out of 125 genes, 3.2%	4 out of 2783 background genes, 0.1%	0.00%
microtubule nucleation	4 out of 125 genes, 3.2%	4 out of 2783 background genes, 0.1%	0.00%
positive regulation of microtubule polymerization or depolymerization	4 out of 125 genes, 3.2%	5 out of 2783 background genes, 0.2%	0.02%
regulation of microtubule polymerization	4 out of 125 genes, 3.2%	5 out of 2783 background genes, 0.2%	0.02%
positive regulation of microtubule polymerization	4 out of 125 genes, 3.2%	5 out of 2783 background genes, 0.2%	0.02%
attachment of mitotic spindle microtubules to kinetochore	4 out of 125 genes, 3.2%	5 out of 2783 background genes, 0.2%	0.02%
regulation of cell cycle checkpoint	4 out of 125 genes, 3.2%	6 out of 2783 background genes, 0.2%	0.02%

GO term	Cluster frequency	Background frequency	FDR
	genes, 3.2%	genes, 0.2%	
regulation of mitotic spindle elongation	4 out of 125 genes, 3.2%	6 out of 2783 background genes, 0.2%	0.02%
microtubule polymerization	4 out of 125 genes, 3.2%	6 out of 2783 background genes, 0.2%	0.02%
microtubule polymerization or depolymerization	4 out of 125 genes, 3.2%	7 out of 2783 background genes, 0.3%	0.14%
regulation of spindle elongation	4 out of 125 genes, 3.2%	7 out of 2783 background genes, 0.3%	0.14%
mitotic DNA replication checkpoint	4 out of 125 genes, 3.2%	7 out of 2783 background genes, 0.3%	0.14%
positive regulation of spindle pole body separation	3 out of 125 genes, 2.4%	3 out of 2783 background genes, 0.1%	0.10%
meiotic DNA double-strand break processing	3 out of 125 genes, 2.4%	3 out of 2783 background genes, 0.1%	0.10%

Table 6.11: GO analysis of Cdc14 hits based on cellular process.

GO term	Cluster frequency	Background frequency	FDR
double-stranded DNA binding	17 out of 125 genes, 13.6%	77 out of 2783 background genes, 2.8%	0.00%
microtubule binding	8 out of 125 genes, 6.4%	14 out of 2783 background genes, 0.5%	0.00%
microtubule plus-end binding	5 out of 125 genes, 4.0%	5 out of 2783 background genes, 0.2%	0.00%
DNA replication origin binding	9 out of 125 genes, 7.2%	25 out of 2783 background genes, 0.9%	0.00%
tubulin binding	8 out of 125 genes, 6.4%	20 out of 2783 background genes, 0.7%	0.00%
DNA binding	25 out of 125 genes, 20.0%	206 out of 2783 background genes, 7.4%	0.00%
sequence-specific double-stranded DNA binding	12 out of 125 genes, 9.6%	54 out of 2783 background genes, 1.9%	0.00%
sequence-specific DNA binding	14 out of 125 genes, 11.2%	87 out of 2783 background genes, 3.1%	0.25%
cytoskeletal protein binding	10 out of 125 genes, 8.0%	46 out of 2783 background genes, 1.7%	0.22%
structural constituent of cytoskeleton	4 out of 125 genes, 3.2%	8 out of 2783 background genes, 0.3%	0.80%

Table 6.12: GO analysis of Cdc14 hits based on protein function.

GO term	Cluster frequency	Background frequency	FDR
microtubule cytoskeleton	27 out of 125 genes, 21.6%	108 out of 2783 background genes, 3.9%	0.00%
chromosome, centromeric region	18 out of 125 genes, 14.4%	47 out of 2783 background genes, 1.7%	0.00%
condensed chromosome outer kinetochore	9 out of 125 genes, 7.2%	9 out of 2783 background genes, 0.3%	0.00%
condensed nuclear chromosome outer kinetochore	9 out of 125 genes, 7.2%	9 out of 2783 background genes, 0.3%	0.00%
condensed chromosome kinetochore	13 out of 125 genes, 10.4%	22 out of 2783 background genes, 0.8%	0.00%
condensed nuclear chromosome kinetochore	12 out of 125 genes, 9.6%	20 out of 2783 background genes, 0.7%	0.00%
condensed chromosome, centromeric region	13 out of 125 genes, 10.4%	25 out of 2783 background genes, 0.9%	0.00%
chromosomal region	19 out of 125 genes, 15.2%	63 out of 2783 background genes, 2.3%	0.00%
kinetochore	13 out of 125 genes, 10.4%	28 out of 2783 background genes, 1.0%	0.00%
condensed nuclear chromosome, centromeric region	12 out of 125 genes, 9.6%	23 out of 2783 background genes, 0.8%	0.00%
chromosomal part	32 out of 125 genes, 25.6%	200 out of 2783 background genes, 7.2%	0.00%
chromosome	32 out of 125 genes, 25.6%	208 out of 2783 background genes, 7.5%	0.00%
spindle	16 out of 125 genes, 12.8%	50 out of 2783 background genes, 1.8%	0.00%
DASH complex	7 out of 125 genes, 5.6%	7 out of 2783 background genes, 0.3%	0.00%
cytoskeletal part	28 out of 125 genes, 22.4%	170 out of 2783 background genes, 6.1%	0.00%
cytoskeleton	28 out of 125 genes, 22.4%	171 out of 2783 background genes, 6.1%	0.00%
nuclear chromosome	27 out of 125 genes, 21.6%	163 out of 2783 background genes, 5.9%	0.00%
condensed nuclear chromosome	13 out of 125 genes, 10.4%	36 out of 2783 background genes, 1.3%	0.00%
nuclear chromosome part	26 out of 125 genes, 20.8%	156 out of 2783 background genes, 5.6%	0.00%
condensed chromosome nuclear origin of replication	14 out of 125 genes, 11.2%	44 out of 2783 background genes, 1.6%	0.00%
recognition complex	6 out of 125 genes, 4.8%	6 out of 2783 background genes, 0.2%	0.00%
origin recognition complex	6 out of 125 genes, 4.8%	6 out of 2783 background genes, 0.2%	0.00%
spindle pole body	16 out of 125 genes, 12.8%	68 out of 2783 background genes, 2.4%	0.00%
microtubule organizing centre	16 out of 125 genes, 12.8%	69 out of 2783 background genes, 2.5%	0.00%
mitotic spindle pole body	13 out of 125 genes, 10.4%	56 out of 2783 background genes, 2.0%	0.00%

GO term	Cluster frequency	Background frequency	FDR
	genes, 10.4%	genes, 2.0%	
microtubule	8 out of 125 genes, 6.4%	19 out of 2783 background genes, 0.7%	0.00%
spindle microtubule	7 out of 125 genes, 5.6%	14 out of 2783 background genes, 0.5%	0.00%
DNA replication preinitiation complex	7 out of 125 genes, 5.6%	15 out of 2783 background genes, 0.5%	0.00%
spindle midzone	7 out of 125 genes, 5.6%	15 out of 2783 background genes, 0.5%	0.00%
pre-replicative complex	7 out of 125 genes, 5.6%	16 out of 2783 background genes, 0.6%	0.00%
nuclear pre-replicative complex	7 out of 125 genes, 5.6%	16 out of 2783 background genes, 0.6%	0.00%
mitotic spindle	10 out of 125 genes, 8.0%	37 out of 2783 background genes, 1.3%	0.00%
supramolecular complex	8 out of 125 genes, 6.4%	28 out of 2783 background genes, 1.0%	0.06%
supramolecular polymer	8 out of 125 genes, 6.4%	28 out of 2783 background genes, 1.0%	0.06%
supramolecular fibre	8 out of 125 genes, 6.4%	28 out of 2783 background genes, 1.0%	0.06%
polymeric cytoskeletal fibre	8 out of 125 genes, 6.4%	28 out of 2783 background genes, 1.0%	0.06%
protein-DNA complex	9 out of 125 genes, 7.2%	37 out of 2783 background genes, 1.3%	0.05%
chromosome passenger complex	3 out of 125 genes, 2.4%	3 out of 2783 background genes, 0.1%	0.32%
anaphase-promoting complex	3 out of 125 genes, 2.4%	4 out of 2783 background genes, 0.1%	0.51%
microtubule associated complex	3 out of 125 genes, 2.4%	4 out of 2783 background genes, 0.1%	0.50%

Table 6.13: GO analysis of Cdc14 hits based on cellular components.

6.7. Discussion

The development and application of a SILAC-based screen for Cdc14^{PD} interactors resulted in the identification of over 100 potential candidates. The SILAC protocol described in chapter 5 was successfully applied here and proved to be a powerful method for studying protein interactions in *C. albicans*. One of the biggest strength of this study is that it used an unbiased approach for identification of unknown Cdc14 interactions. Hits selection was based on clearly defined rules guided by the experimental conditions, and not influenced by previous knowledge of Cdc14 interactions. However, given the conserved roles of Cdc14 homologs, at least some of the interactions found here also occur in other organisms, especially fungi. In the list of hits identified by MS, there are several well-known Cdc14 interactors. In *S. cerevisiae*, Cdc14 is held in the nucleolus by Net1 throughout interphase. The Net1 homologue in *C. albicans* has not been characterised but was found as a hit in all four QO-MS experiments in both yeast and hyphae, suggesting a similar function. The mitotic exit network kinases Dbf2 and Cdc5 were also recovered. Other kinases in the list include Yak1 and both subunits of Cdc7-Dbf4. The mitotic cyclin Clb2 and the APC/C activator Cdh1 are both involved in counteracting CDK activity initiated by Cdc14. Almost the entire Dam1/DASH complex that orchestrates chromosome segregation was present, including Ask1, Dad1, Dad2, Dad3, Dam1, Duo1, Spc19, Spc34 (only Dad4 and Hsk3 were not recovered). It is intriguing whether Cdc14 interacts with all of these proteins or the whole complex was purified via one specific interaction. Another complex amongst the hits is the origin recognition complex (Orc1, Orc3 and Orc4) involved in DNA replication. Other hits with a role in DNA replication are orf19.2369, orf19.3289, orf5358, orf19.2389 and Cdc7-Dbf4, suggesting a prominent role for Cdc14 in controlling this process. Other proteins involved in various DNA processes were recovered too: Rad52, Rad9, orf19.652, orf19.6155 and 19.6291 are all DNA repair proteins; orf19.1865 directs DNA recombination; orf19.7663 takes part in chromosome segregation; Fkh2 and orf19.4295 are transcriptional corepressors; orf19.427 plays a role in chromatin silencing. Full gene ontology analysis of the final list of Cdc14 interactors will be performed and will reveal more information about the role of this phosphatase in *C. albicans*. When Cdc14 was first characterised, it became known as a mitotic exit protein and it was thought to be inactive during interphase while in the nucleus. This notion was later disproved, as in higher eukaryotes, Cdc14 actively controls

DNA dynamics. The results presented here suggest that in *C. albicans* Cdc14 has a prominent role in the nucleus related to DNA maintenance and organisation.

Cdc14 deletion mutants have severe defect in cell separation due to failure to degrade the septum after cytokinesis. Cdc14 may be activating this process by targeting the hydrolytic enzyme Cht4 that was recovered from two experiments. Amongst the low confidence hits are the known target Iqg1 that directs actomyosin ring disassembly, the transcriptional regulator Ace2 and additional DNA-binding proteins.

Many potential Cdc14 interactors were identified from both yeast and hyphae experiments. Cdc14 localises to the nucleus in both of these forms, so it is likely that its role in DNA maintenance is universal. None of the proteins that were form-specific suggest an obvious role of Cdc14 in morphogenesis. However, almost half of the hits are uncharacterised proteins, i.e. those shown with their ORF number and without a name. The role of these hits in the cell and the significance of their potential interaction with Cdc14 remain to be found.

Chapter 7

Discussion

7.1. Quantitative MS methods for studying kinase and phosphatase interactions in *C. albicans*

The main aim of this project was to identify kinase and phosphatase interactions in *C. albicans* using quantitative MS approaches. Experimenting with two different methods of quantitation, label-free and SILAC, allows for a parallel comparison of both. While both techniques have been hugely refined in the past decade and numerous reviews in the literature discuss their merits, this study found SILAC to be a superior approach for the aims of the project even when all other experimental differences are taken into account. These differences include using different MS instruments for both methods, as well as overexpressing the bait and stabilising bait-substrate interactions in SILAC experiments. Label-free techniques strongly rely on precise and accurate replication of each affinity purification procedure. Differences in the end results of bait and control experiments are then attributed to bait interactions. Label-free experiments described in chapter 3 were replicated with meticulous care, yet the number of identified proteins in each of them varied by several hundreds. This is largely due to the gentle beads-washing conditions, which allow large (but varied) amount of contaminants to remain in the sample. As a result, clear discrimination between preys and contaminants cannot be made. For example the experiment using Dbf2 as a bait found over 300 more proteins than one of the control experiments, where no bait was used ("Control 2" in table 3.1) but these are clearly not all preys. This problem would not be solved if a better MS instrument was used, the bait was overexpressed and bait-prey interactions were enhanced. Therefore, the advantages of the SILAC method are not due to the later improvements in the AP-MS protocol.

It is important to note that label-free MS can be a powerful approach for studying protein interactions, when implemented in the right context. Indeed, a label-free proteomic analysis of Clp1 (homologue of Cdc14 in *S. pombe*) interactions found 128 hits from MS experiments performed under 5 different conditions using either a substrate-trapping or a wild type phosphatase as a bait (Chen et al., 2013). The Chen et al. study had several differences to the project described here. The authors used tandem affinity purification (TAP) which strongly reduces the amount of background proteins in the final sample. They performed 10 MS experiments with the bait which allows them to carry out a better statistical analysis on their data. In addition, the study was carried by a group of highly experienced MS scientists, who have a previously composed database of common contaminants. That allows them to compare results from a large number of label-free experiments and achieve low FDR of the final list of hits. It is also clear that the choice of MS instrument and processing software have met the requirements of label-free experiments (the MS instrument used in the study is a linear trap quadrupole from Thermo Electron). So, while the SILAC protocol produced better results in the current project, it is believed that similar results can be achieved with label-free MS methods. However, label-free MS requires rigorous statistical analysis and use of a protein frequency library where known contaminant proteins for a specific set of experimental parameters (cell line, bead matrix, buffer conditions) may be excluded. It is also better suited for stringent purifications such as TAP, where contamination is reduced to minimum.

This study found over 100 potential Cdc14 interactors, but it certainly missed to identify others. It is recognised that some interactions will not be detected by the methods employed here, so further experiments are likely to find more unknown targets. The main advantage of SILAC over label-free quantitation is that bait and control samples are mixed early in the experimental procedure, so differences between samples cannot arise by handling errors. However, this is also a disadvantage for SILAC, because protein interactions in the mixed cell lysate may still occur. Substrates bound by Cdc14 after cell lysis will not be enriched in heavy isotopes. Thus, they will be false negatives. Such dynamic interactions can be detected by MAP-SILAC (mixing after purification) (Wang and Huang, 2014). As the name suggests, in MAP-SILAC samples are mixed after the affinity purification procedure to order

to avoid exchange between differentially labelled protein complexes. The disadvantage of this method is that it may introduce handling error differences between the samples.

Other reasons for not capturing an interaction may be that it occurred below the threshold of AP-MS detection or it did not withstand the course of the experiment. Although some improvements were made to stabilise transient interactions and enhance low abundant interactions, many of them would still be lost. Nevertheless, these improvements certainly played a role in the success of the study. This becomes evident when the results from SILAC and label-free Cdc14 experiments are directly compared to one another. Many of the most confident hits in SILAC experiments that were identified repeatedly by both QTOF and QO instruments were not identified in the label free IP of Cdc14, for example orf19.652, orf19.2684 and Rad52. While label-free experiments were certainly analysed by a less sensitive MS instrument, these three hits had relatively high intensities, so failure of detection cannot be explained with MS sensitivity. It is more likely that these hits (which can be regarded as interacting partners of Cdc14 based on very strong evidence) were either not present in the MS sample (i.e. they were lost, because bait-prey interactions were weak), or they were present in the MS sample at very low abundance because Cdc14 was not overexpressed. From that, it can be concluded that using an overexpressed substrate-trapping mutant of Cdc14 in conjunction with quantitative SILAC-MS was a right decision.

7.2. Strengths and limitations of using an overexpressed mutant version of Cdc14 in interaction studies

Quantitative MS analysis of protein interactions involves detecting subtle changes in protein levels between two samples. Despite the huge advantages of MS technology available today, detecting protein interactions at physiological levels remains a challenging task. This is especially true for proteins with relatively low abundance, such as Cdc14.

In the Chen et al. study described in the previous section, the authors used both wild type Clp1 and Clp1^{PD} as a bait. Out of 128 hits that they found in total, 73 (57%) were enriched three or more times in the Clp1^{PD} compared to Clp1 experiments. As this and many other studies have shown, substrate-trapping mutants are a great tool in protein interaction experiments.

Early AP-MS experiments described in chapter 3 illustrate the difficulty of capturing interacting proteins *in vivo* and later identifying them among the crowd of contaminants. As already discussed above, experiments with wild type Cdc14 failed to identify even the most prominent interactors found with MET3-Cdc14^{PD}. The early experiments were performed on a less sensitive instrument, but this is unlikely to be the sole reason for the lack of hits. A high-sensitivity mass spectrometer would most likely detect more interacting partners, but they would not stand out from the contaminants (even if SILAC is used). The ratio of preys-to-contaminants was increased by using an overexpressed substrate-trapping Cdc14^{PD} while quantitation was improved by using SILAC and high-sensitivity MS.

The non-physiological conditions of the experiment have almost certainly created some aberrant interactions. A study in budding yeast found that a gradual change in the Cdc14-to-CDKs ratio during mitosis is responsible sequential substrate dephosphorylation by Cdc14 (Bouchoux and Uhlmann, 2011). Constitutively high levels of Cdc14 are likely to disrupt the natural sequence of dephosphorylation events. However, since the overexpressed phosphatase was inactive, and an active Cdc14 was present in the cells, abnormalities would not be due to excessive Cdc14 dephosphorylation. Rather than that, high levels of Cdc14^{PD} may deplete the pool of substrates that active Cdc14 dephosphorylates. That may have an effect on downstream events governed by these

substrates, even if no phenotype was seen by microscopic observations. This has two important consequences: 1) some of the hits in this study may be non-physiological substrates of Cdc14 and 2) some of the hits may be false positives due to being more abundant in the Cdc14^{PD} strain than the wild type strain. To clarify the second point: if a contaminating protein is more abundant in the labelled than the non-labelled strain, it will appear with higher H/L ratio relative to the whole population. It will therefore be regarded as a hit, although it did not interact directly with the bait. Thus, hits in this study should be regarded as potential interactors of Cdc14, but further experiments would be required to confirm these interactions.

In conclusion, while the non-physiological conditions of the experiments limit the significance of the findings presented here, it is believed that many physiological interactions would not have been revealed in different experimental conditions. Therefore, this study has an important contribution to understanding the *C. albicans* interactome.

The list of hits presented in this study shows proteins that are likely interacting with Cdc14. These include substrates of the phosphatase, activating subunits, inhibitors, upstream regulators, anchoring proteins and other interactors. Since substrate interactions were artificially enhanced, the list is likely to be enriched in substrates. These cannot be distinguished from the rest without further investigation, but comparisons can be drawn between Cdc14 homologues in other species. For example Dbf2, as part of the mitotic exit network, is known to phosphorylate

Another important consideration when interpreting the results is that physical interactions detected by AP-MS are not necessarily direct. Proteins bound together in a stable complex may be purified via a single member. Two complexes were significantly enriched in the list of potential interactors: the Dam1/Dash complex and the origin recognition complex. Members of these complexes should not be regarded as direct Cdc14 interactors. It is likely that other proteins have also been purified via indirect interaction. Thus, this study presents a global view of the Cdc14 interactome, which include both immediate and distal physical interactions.

7.3. Future work

The data obtained in this project will be subjected to further bioinformatics analysis. A close comparison to known Cdc14 interactors in other species will also be performed. One can further compare the results of this study to the results of similar AP-MS experiments with Cdc14 homologues (e.g. the Chen et al. study described above). This will provide additional validation that the list of final hits is enriched in Cdc14 interactors and that the experimental approach is working.

Finally, the project aims to confirm some of the interactions by further experiments, such as co-IP, and investigate the role of Cdc14 dephosphorylation of a few chosen targets. Unknown proteins will be tagged with GFP in order to examine their localisation. Selected genes will be deleted in order to look for deletion phenotype and investigate the function of these genes.

Investigating every single interaction found by IP-MS is beyond the scope of this project. This study provided a global analysis of Cdc14 interactions in *C. albicans*, but the importance of individual interaction may be researched further by other groups.

The methods of quantitative SILAC-MS used in this study proved to be a valuable tool for large scale analysis of protein interactions in *C. albicans*. The protocols may therefore be applied in further studies of *C. albicans* interactome. SILAC could also be applied in studies of protein dynamics, protein turnover rate, whole proteome analysis and others.

References

- Adam, C., Erdei, E., Casado, C., Kovacs, L., Gonzalez, A., Majoros, L., Petrenyi, K., Bagossi, P., Farkas, I., Molnar, M., Pocsi, I., Arino, J. & Dombradi, V. (2012) Protein phosphatase CaPpz1 is involved in cation homeostasis, cell wall integrity and virulence of *Candida albicans*. *Microbiology*, 158, 1258-67.
- Albataineh, M. T., Lazzell, A., Lopez-Ribot, J. L. & Kadosh, D. (2014) Ppg1, a PP2A-type protein phosphatase, controls filament extension and virulence in *Candida albicans*. *Eukaryot Cell*, 13, 1538-47.
- Alonso-Monge, R., Navarro-Garcia, F., Molero, G., Diez-Orejas, R., Gustin, M., Pla, J., Sanchez, M. & Nombela, C. (1999) Role of the mitogen-activated protein kinase Hog1p in morphogenesis and virulence of *Candida albicans*. *J Bacteriol*, 181, 3058-68.
- Antinori, S., Milazzo, L., Sollima, S., Galli, M. & Corbellino, M. (2016) Candidemia and invasive candidiasis in adults: A narrative review. *Eur J Intern Med*.
- Au Yong, J. Y., Wang, Y. M. & Wang, Y. (2016) The Nim1 kinase Gin4 has distinct domains crucial for septin assembly, phospholipid binding and mitotic exit. *J Cell Sci*, 129, 2744-56.
- Bader, T., Bodendorfer, B., Schroppel, K. & Morschhauser, J. (2003) Calcineurin is essential for virulence in *Candida albicans*. *Infect Immun*, 71, 5344-54.
- Bharucha, N., Chabrier-Rosello, Y., Xu, T., Johnson, C., Sobczynski, S., Song, Q., Dobry, C. J., Eckwahl, M. J., Anderson, C. P., Benjamin, A. J., Kumar, A. & Krysan, D. J. (2011) A large-scale complex haploinsufficiency-based genetic interaction screen in *Candida albicans*: analysis of the RAM network during morphogenesis. *PLoS Genet*, 7, e1002058.

- Bicho, C. C., De Lima Alves, F., Chen, Z. A., Rappsilber, J. & Sawin, K. E. (2010) A genetic engineering solution to the "arginine conversion problem" in stable isotope labeling by amino acids in cell culture (SILAC). *Mol Cell Proteomics*, 9, 1567-77.
- Bishop, A., Lane, R., Beniston, R., Chapa-Y-Lazo, B., Smythe, C. & Sudbery, P. (2010) Hyphal growth in *Candida albicans* requires the phosphorylation of Sec2 by the Cdc28-Ccn1/Hgc1 kinase. *Embo J*, 29, 2930-42.
- Blanchetot, C., Chagnon, M., Dube, N., Halle, M. & Tremblay, M. L. (2005) Substrate-trapping techniques in the identification of cellular PTP targets. *Methods*, 35, 44-53.
- Bloom, J., Cristea, I. M., Procko, A. L., Lubkov, V., Chait, B. T., Snyder, M. & Cross, F. R. (2011) Global analysis of Cdc14 phosphatase reveals diverse roles in mitotic processes. *J Biol Chem*, 286, 5434-45.
- Bononi, A., Agnoletto, C., De Marchi, E., Marchi, S., Patergnani, S., Bonora, M., Giorgi, C., Missiroli, S., Poletti, F., Rimessi, A. & Pinton, P. (2011) Protein kinases and phosphatases in the control of cell fate. *Enzyme Res*, 2011, 329098.
- Borek, W. E., Zou, J., Rappsilber, J. & Sawin, K. E. (2015) Deletion of Genes Encoding Arginase Improves Use of "Heavy" Isotope-Labeled Arginine for Mass Spectrometry in Fission Yeast. *PLoS One*, 10, e0129548.
- Breitkreutz, A., Choi, H., Sharom, J. R., Boucher, L., Neduva, V., Larsen, B., Lin, Z. Y., Breitkreutz, B. J., Stark, C., Liu, G., Ahn, J., Dewar-Darch, D., Reguly, T., Tang, X., Almeida, R., Qin, Z. S., Pawson, T., Gingras, A. C., Nesvizhskii, A. I. & Tyers, M. (2010) A global protein kinase and phosphatase interaction network in yeast. *Science*, 328, 1043-6.
- Bremmer, S. C., Hall, H., Martinez, J. S., Eissler, C. L., Hinrichsen, T. H., Rossie, S., Parker, L. L., Hall, M. C. & Charbonneau, H. (2012) Cdc14 phosphatases preferentially dephosphorylate a subset of cyclin-dependent kinase (Cdk) sites containing phosphoserine. *J Biol Chem*, 287, 1662-9.
- Bruckner, A., Polge, C., Lentze, N., Auerbach, D. & Schlattner, U. (2009) Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci*, 10, 2763-88.

- Caballero-Lima, D. & Sudbery, P. E. (2014) In *Candida albicans*, phosphorylation of Exo84 by Cdk1-Hgc1 is necessary for efficient hyphal extension. *Mol Biol Cell*, 25, 1097-110.
- Chen, J., Zhou, S., Wang, Q., Chen, X., Pan, T. & Liu, H. (2000) Crk1, a novel Cdc2-related protein kinase, is required for hyphal development and virulence in *Candida albicans*. *Mol Cell Biol*, 20, 8696-708.
- Chen, J. S., Broadus, M. R., Mclean, J. R., Feoktistova, A., Ren, L. & Gould, K. L. (2013) Comprehensive proteomics analysis reveals new substrates and regulators of the fission yeast clp1/cdc14 phosphatase. *Mol Cell Proteomics*, 12, 1074-86.
- Cheng, H. C., Qi, R. Z., Paudel, H. & Zhu, H. J. (2011) Regulation and function of protein kinases and phosphatases. *Enzyme Res*, 2011, 794089.
- Chernushevich, I. V., Loboda, A. V. & Thomson, B. A. (2001) An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom*, 36, 849-65.
- Choi, H., Larsen, B., Lin, Z. Y., Breitzkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z. S., Tyers, M., Gingras, A. C. & Nesvizhskii, A. I. (2011) SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat Methods*, 8, 70-3.
- Clemente-Blanco, A., Gonzalez-Novo, A., Machin, F., Caballero-Lima, D., Aragon, L., Sanchez, M., De Aldana, C. R., Jimenez, J. & Correa-Bordes, J. (2006) The Cdc14p phosphatase affects late cell-cycle events and morphogenesis in *Candida albicans*. *J Cell Sci*, 119, 1130-43.
- Court, H. & Sudbery, P. (2007) Regulation of Cdc42 GTPase activity in the formation of hyphae in *Candida albicans*. *Mol Biol Cell*, 18, 265-81.
- Csank, C., Makris, C., Meloche, S., Schroppel, K., Rollinghoff, M., Dignard, D., Thomas, D. Y. & Whiteway, M. (1997) Derepressed hyphal growth and reduced virulence in a VH1 family-related protein phosphatase mutant of the human pathogen *Candida albicans*. *Mol Biol Cell*, 8, 2539-51.
- Cueille, N., Salimova, E., Esteban, V., Blanco, M., Moreno, S., Bueno, A. & Simanis, V. (2001) Flp1, a fission yeast orthologue of the *S. cerevisiae* CDC14 gene, is not required for

- cyclin degradation or rum1p stabilisation at the end of mitosis. *J Cell Sci*, 114, 2649-64.
- Dhillon, N. K., Sharma, S. & Khuller, G. K. (2003) Signaling through protein kinases and transcriptional regulators in *Candida albicans*. *Crit Rev Microbiol*, 29, 259-75.
- Fan, J., Wu, M., Jiang, L. & Shen, S. H. (2009) A serine/threonine protein phosphatase-like protein, CaPTC8, from *Candida albicans* defines a new PPM subfamily. *Gene*, 430, 64-76.
- Fasolo, J., Sboner, A., Sun, M. G., Yu, H., Chen, R., Sharon, D., Kim, P. M., Gerstein, M. & Snyder, M. (2011) Diverse protein kinase interactions identified by protein microarrays reveal novel connections between cellular processes. *Genes Dev*, 25, 767-78.
- Faust, A. M., Wong, C. C., Yates, J. R., 3rd, Drubin, D. G. & Barnes, G. (2013) The FEAR protein Slk19 restricts Cdc14 phosphatase to the nucleus until the end of anaphase, regulating its participation in mitotic exit in *Saccharomyces cerevisiae*. *PLoS One*, 8, e73194.
- Feng, J., Zhao, J., Li, J., Zhang, L. & Jiang, L. (2010) Functional characterization of the PP2C phosphatase CaPtc2p in the human fungal pathogen *Candida albicans*. *Yeast*, 27, 753-64.
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246, 64-71.
- Ficarro, S. B., McClelland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F. & White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol*, 20, 301-5.
- Finkel, J. S., Xu, W., Huang, D., Hill, E. M., Desai, J. V., Woolford, C. A., Nett, J. E., Taff, H., Norice, C. T., Andes, D. R., Lanni, F. & Mitchell, A. P. (2012) Portrait of *Candida albicans* adherence regulators. *PLoS Pathog*, 8, e1002525.

- Frohlich, F., Christiano, R. & Walther, T. C. (2013) Native SILAC: metabolic labeling of proteins in prototroph microorganisms based on lysine synthesis regulation. *Mol Cell Proteomics*, 12, 1995-2005.
- Gavin, A. C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M. A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A. M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J. M., Kuster, B., Bork, P., Russell, R. B. & Superti-Furga, G. (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature*, 440, 631-6.
- Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M. A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. & Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, 415, 141-7.
- Gonzalez-Novo, A., Correa-Bordes, J., Labrador, L., Sanchez, M., Vazquez De Aldana, C. R. & Jimenez, J. (2008) Sep7 is essential to modify septin ring dynamics and inhibit cell separation during *Candida albicans* hyphal growth. *Mol Biol Cell*, 19, 1509-18.
- Gonzalez-Novo, A., Labrador, L., Pablo-Hernando, M. E., Correa-Bordes, J., Sanchez, M., Jimenez, J. & Vazquez De Aldana, C. R. (2009) Dbf2 is essential for cytokinesis and correct mitotic spindle formation in *Candida albicans*. *Mol Microbiol*, 72, 1364-78.
- Gray, C. H., Good, V. M., Tonks, N. K. & Barford, D. (2003) The structure of the cell cycle protein Cdc14 reveals a proline-directed protein phosphatase. *Embo J*, 22, 3524-35.
- Greig, J. A., Sudbery, I. M., Richardson, J. P., Naglik, J. R., Wang, Y. & Sudbery, P. E. (2015) Cell cycle-independent phospho-regulation of Fkh2 during hyphal growth regulates *Candida albicans* pathogenesis. *PLoS Pathog*, 11, e1004630.

- Gruhler, A., Schulze, W. X., Matthiesen, R., Mann, M. & Jensen, O. N. (2005) Stable isotope labeling of *Arabidopsis thaliana* cells and quantitative proteomics by mass spectrometry. *Mol Cell Proteomics*, 4, 1697-709.
- Guo, W., Roth, D., Walch-Solimena, C. & Novick, P. (1999) The exocyst is an effector for Sec4p, targeting secretory vesicles to sites of exocytosis. *Embo J*, 18, 1071-80.
- Gutierrez-Escribano, P., Gonzalez-Novo, A., Suarez, M. B., Li, C. R., Wang, Y., De Aldana, C. R. & Correa-Bordes, J. (2011) CDK-dependent phosphorylation of Mob2 is essential for hyphal development in *Candida albicans*. *Mol Biol Cell*, 22, 2458-69.
- Hanaoka, N., Takano, Y., Shibuya, K., Fugo, H., Uehara, Y. & Niimi, M. (2008) Identification of the putative protein phosphatase gene PTC1 as a virulence-related gene using a silkworm model of *Candida albicans* infection. *Eukaryot Cell*, 7, 1640-8.
- Hanaoka, N., Umeyama, T., Ueno, K., Ueda, K., Beppu, T., Fugo, H., Uehara, Y. & Niimi, M. (2005) A putative dual-specific protein phosphatase encoded by YVH1 controls growth, filamentation and virulence in *Candida albicans*. *Microbiology*, 151, 2223-32.
- Higuchi, T. & Uhlmann, F. (2005) Stabilization of microtubule dynamics at anaphase onset promotes chromosome segregation. *Nature*, 433, 171-6.
- Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B. D., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W., Figgeys, D. & Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*, 415, 180-3.
- Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V. & Skrzypek, E. (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res*, 43, D512-20.

- Hu, K., Li, W., Wang, H., Chen, K., Wang, Y. & Sang, J. (2012) Shp1, a regulator of protein phosphatase 1 Glc7, has important roles in cell morphogenesis, cell cycle progression and DNA damage response in *Candida albicans*. *Fungal Genet Biol*, 49, 433-42.
- Huang, Z. X., Zhao, P., Zeng, G. S., Wang, Y. M., Sudbery, I. & Wang, Y. (2014) Phosphoregulation of Nap1 plays a role in septin ring dynamics and morphogenesis in *Candida albicans*. *MBio*, 5, e00915-13.
- Hutti, J. E., Jarrell, E. T., Chang, J. D., Abbott, D. W., Storz, P., Toker, A., Cantley, L. C. & Turk, B. E. (2004) A rapid method for determining protein kinase phosphorylation specificity. *Nat Methods*, 1, 27-9.
- Jiang, H. & English, A. M. (2002) Quantitative analysis of the yeast proteome by incorporation of isotopically labeled leucine. *J Proteome Res*, 1, 345-50.
- Kabir, M. A., Hussain, M. A. & Ahmad, Z. (2012) *Candida albicans*: A Model Organism for Studying Fungal Pathogens. *ISRN Microbiol*, 2012, 538694.
- Kamioka, Y., Yasuda, S., Fujita, Y., Aoki, K. & Matsuda, M. (2011) Multiple decisive phosphorylation sites for the negative feedback regulation of SOS1 via ERK. *J Biol Chem*, 285, 33540-8.
- Kao, L., Wang, Y. T., Chen, Y. C., Tseng, S. F., Jhang, J. C., Chen, Y. J. & Teng, S. C. (2014) Global analysis of cdc14 dephosphorylation sites reveals essential regulatory role in mitosis and cytokinesis. *Mol Cell Proteomics*, 13, 594-605.
- Karas, M. & Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*, 60, 2299-301.
- Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H. C., Stines, A. P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M. & Hartl, F. U. (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell*, 122, 209-20.

- Knebel, A., Morrice, N. & Cohen, P. (2001) A novel method to identify protein kinase substrates: eEF2 kinase is phosphorylated and inhibited by SAPK4/p38delta. *Embo J*, 20, 4360-9.
- Koch, A. & Hauf, S. (2010) Strategies for the identification of kinase substrates using analog-sensitive kinases. *Eur J Cell Biol*, 89, 184-93.
- Kosti, I., Mandel-Gutfreund, Y., Glaser, F. & Horwitz, B. A. (2010) Comparative analysis of fungal protein kinases and associated domains. *BMC Genomics*, 11, 133.
- Lanzetti, L., Margaria, V., Melander, F., Virgili, L., Lee, M. H., Bartek, J. & Jensen, S. (2007) Regulation of the Rab5 GTPase-activating protein RN-tre by the dual specificity phosphatase Cdc14A in human cells. *J Biol Chem*, 282, 15258-70.
- Lay, J., Henry, L. K., Clifford, J., Koltin, Y., Bulawa, C. E. & Becker, J. M. (1998) Altered expression of selectable marker URA3 in gene-disrupted *Candida albicans* strains complicates interpretation of virulence studies. *Infect Immun*, 66, 5301-6.
- Lee, C. M., Nantel, A., Jiang, L., Whiteway, M. & Shen, S. H. (2004) The serine/threonine protein phosphatase SIT4 modulates yeast-to-hypha morphogenesis and virulence in *Candida albicans*. *Mol Microbiol*, 51, 691-709.
- Lee, H. J., Kim, J. M., Kang, W. K., Yang, H. & Kim, J. Y. (2015) The NDR Kinase Cbk1 Downregulates the Transcriptional Repressor Nrg1 through the mRNA-Binding Protein Ssd1 in *Candida albicans*. *Eukaryot Cell*, 14, 671-83.
- Li, C., Melesse, M., Zhang, S., Hao, C., Wang, C., Zhang, H., Hall, M. C. & Xu, J. R. (2015) FgCDC14 regulates cytokinesis, morphogenesis, and pathogenesis in *Fusarium graminearum*. *Mol Microbiol*, 98, 770-86.
- Li, C. R., Au Yong, J. Y., Wang, Y. M. & Wang, Y. (2012) CDK regulates septin organization through cell-cycle-dependent phosphorylation of the Nim1-related kinase Gin4. *J Cell Sci*, 125, 2533-43.

- Li, Y., Cross, F. R. & Chait, B. T. (2014) Method for identifying phosphorylated substrates of specific cyclin/cyclin-dependent kinase complexes. *Proc Natl Acad Sci U S A*, 111, 11323-8.
- Lin, H., Ha, K., Lu, G., Fang, X., Cheng, R., Zuo, Q. & Zhang, P. (2015) Cdc14A and Cdc14B Redundantly Regulate DNA Double-Strand Break Repair. *Mol Cell Biol*, 35, 3657-68.
- Liu, G., Zhang, J., Larsen, B., Stark, C., Breitkreutz, A., Lin, Z. Y., Breitkreutz, B. J., Ding, Y., Colwill, K., Pasculescu, A., Pawson, T., Wrana, J. L., Nesvizhskii, A. I., Raught, B., Tyers, M. & Gingras, A. C. (2010) ProHits: integrated software for mass spectrometry-based interaction proteomics. *Nat Biotechnol*, 28, 1015-7.
- Liu, Q., Han, Q., Wang, N., Yao, G., Zeng, G., Wang, Y., Huang, Z., Sang, J. & Wang, Y. (2016) Tpd3-Pph21 phosphatase plays a direct role in Sep7 dephosphorylation in *Candida albicans*. *Mol Microbiol*, 101, 109-21.
- Lo, H. J., Kohler, J. R., Didomenico, B., Loebenberg, D., Cacciapuoti, A. & Fink, G. R. (1997) Nonfilamentous *C. albicans* mutants are avirulent. *Cell*, 90, 939-49.
- Lossner, C., Warnken, U., Pscherer, A. & Schnolzer, M. (2011) Preventing arginine-to-proline conversion in a cell-line-independent manner during cell cultivation under stable isotope labeling by amino acids in cell culture (SILAC) conditions. *Anal Biochem*, 412, 123-5.
- Mah, A. S., Jang, J. & Deshaies, R. J. (2001) Protein kinase Cdc15 activates the Dbf2-Mob1 kinase complex. *Proc Natl Acad Sci U S A*, 98, 7325-30.
- Malumbres, M. (2014) Cyclin-dependent kinases. *Genome Biol*, 15, 122.
- Marcilla, M., Alpizar, A., Paradela, A. & Albar, J. P. (2011) A systematic approach to assess amino acid conversions in SILAC experiments. *Talanta*, 84, 430-6.
- Marshall, A. G., Hendrickson, C. L. & Jackson, G. S. (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev*, 17, 1-35.
- Martin-Perez, M. & Villen, J. (2015) Feasibility of protein turnover studies in prototroph *Saccharomyces cerevisiae* strains. *Anal Chem*, 87, 4008-14.

- Martins, N., Ferreira, I. C., Barros, L., Silva, S. & Henriques, M. (2014) Candidiasis: predisposing factors, prevention, diagnosis and alternative treatment. *Mycopathologia*, 177, 223-40.
- Medzihradszky, K. F., Campbell, J. M., Baldwin, M. A., Falick, A. M., Juhasz, P., Vestal, M. L. & Burlingame, A. L. (2000) The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal Chem*, 72, 552-8.
- Mocciaro, A. & Schiebel, E. (2010) Cdc14: a highly conserved family of phosphatases with non-conserved functions? *J Cell Sci*, 123, 2867-76.
- Mohl, D. A., Huddleston, M. J., Collingwood, T. S., Annan, R. S. & Deshaies, R. J. (2009) Dbf2-Mob1 drives relocalization of protein phosphatase Cdc14 to the cytoplasm during exit from mitosis. *J Cell Biol*, 184, 527-39.
- Molina, H., Yang, Y., Ruch, T., Kim, J. W., Mortensen, P., Otto, T., Nalli, A., Tang, Q. Q., Lane, M. D., Chaerkady, R. & Pandey, A. (2009) Temporal profiling of the adipocyte proteome during differentiation using a five-plex SILAC based strategy. *J Proteome Res*, 8, 48-58.
- Muller, A. C., Giambruno, R., Weisser, J., Majek, P., Hofer, A., Bigenzahn, J. W., Superti-Furga, G., Jessen, H. J. & Bennett, K. L. (2016) Identifying Kinase Substrates via a Heavy ATP Kinase Assay and Quantitative Mass Spectrometry. *Sci Rep*, 6, 28107.
- Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem*, 75, 4646-58.
- Ni, J., Gao, Y., Liu, H. & Chen, J. (2004) *Candida albicans* Cdc37 interacts with the Crk1 kinase and is required for Crk1 production. *FEBS Lett*, 561, 223-30.
- Nobile, C. J. & Johnson, A. D. (2015) *Candida albicans* Biofilms and Human Disease. *Annu Rev Microbiol*, 69, 71-92.

- Noble, S. M., French, S., Kohn, L. A., Chen, V. & Johnson, A. D. (2010) Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nat Genet*, 42, 590-8.
- Noble, S. M. & Johnson, A. D. (2007) Genetics of *Candida albicans*, a diploid human fungal pathogen. *Annu Rev Genet*, 41, 193-211.
- Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A. & Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*, 1, 376-86.
- Ong, S. E. & Mann, M. (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc*, 1, 2650-60.
- Park, S. K., Liao, L., Kim, J. Y. & Yates, J. R., 3RD (2009) A computational approach to correct arginine-to-proline conversion in quantitative proteomics. *Nat Methods*, 6, 184-5.
- Pitt, J. J. (2009) Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev*, 30, 19-34.
- Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., Guo, H., Jona, G., Breitkreutz, A., Sopko, R., McCartney, R. R., Schmidt, M. C., Rachidi, N., Lee, S. J., Mah, A. S., Meng, L., Stark, M. J., Stern, D. F., De Virgilio, C., Tyers, M., Andrews, B., Gerstein, M., Schweitzer, B., Predki, P. F. & Snyder, M. (2005) Global analysis of protein phosphorylation in yeast. *Nature*, 438, 679-84.
- Queralt, E., Lehane, C., Novak, B. & Uhlmann, F. (2006) Downregulation of PP2A(Cdc55) phosphatase by separase initiates mitotic exit in budding yeast. *Cell*, 125, 719-32.
- Rudner, A. D. & Murray, A. W. (2000) Phosphorylation by Cdc28 activates the Cdc20-dependent activity of the anaphase-promoting complex. *J Cell Biol*, 149, 1377-90.
- Sanchez-Diaz, A., Nkosi, P. J., Murray, S. & Labib, K. (2012) The Mitotic Exit Network and Cdc14 phosphatase initiate cytokinesis by counteracting CDK phosphorylations and blocking polarised growth. *Embo J*, 31, 3620-34.

- Santos, M. A. & Tuite, M. F. (1995) The CUG codon is decoded in vivo as serine and not leucine in *Candida albicans*. *Nucleic Acids Res*, 23, 1481-6.
- Saputo, S., Norman, K. L., Murante, T., Horton, B. N., Diaz Jde, L., Didone, L., Colquhoun, J., Schroeder, J. W., Simmons, L. A., Kumar, A. & Krysan, D. J. (2016) Complex Haploinsufficiency-Based Genetic Analysis of the NDR/Lats Kinase Cbk1 Provides Insight into Its Multiple Functions in *Candida albicans*. *Genetics*, 203, 1217-33.
- Schaller, M., Borelli, C., Korting, H. C. & Hube, B. (2005) Hydrolytic enzymes as virulence factors of *Candida albicans*. *Mycoses*, 48, 365-77.
- Schroppel, K., Sprosser, K., Whiteway, M., Thomas, D. Y., Rollinghoff, M. & Csank, C. (2000) Repression of hyphal proteinase expression by the mitogen-activated protein (MAP) kinase phosphatase Cpp1p of *Candida albicans* is independent of the MAP kinase Cek1p. *Infect Immun*, 68, 7159-61.
- Scigelova, M., Hornshaw, M., Giannakopoulos, A. & Makarov, A. (2011) Fourier transform mass spectrometry. *Mol Cell Proteomics*, 10, M111 009431.
- Shah, K., Liu, Y., Deirmengian, C. & Shokat, K. M. (1997) Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc Natl Acad Sci U S A*, 94, 3565-70.
- Sinha, I., Wang, Y. M., Philp, R., Li, C. R., Yap, W. H. & Wang, Y. (2007) Cyclin-dependent kinases control septin phosphorylation in *Candida albicans* hyphal development. *Dev Cell*, 13, 421-32.
- Song, Y., Cheon, S. A., Lee, K. E., Lee, S. Y., Lee, B. K., Oh, D. B., Kang, H. A. & Kim, J. Y. (2008) Role of the RAM network in cell polarity and hyphal morphogenesis in *Candida albicans*. *Mol Biol Cell*, 19, 5456-77.
- Stynen, B., Van Dijck, P. & Tournu, H. (2010) A CUG codon adapted two-hybrid system for the pathogenic fungus *Candida albicans*. *Nucleic Acids Res*, 38, e184.
- Sudbery, P., Gow, N. & Berman, J. (2004) The distinct morphogenic states of *Candida albicans*. *Trends Microbiol*, 12, 317-24.

- Sullivan, M. & Uhlmann, F. (2002) A non-proteolytic function of separase links the onset of anaphase to mitotic exit. *Nat Cell Biol*, 5, 249-54.
- Sun, L. L., Li, W. J., Wang, H. T., Chen, J., Deng, P., Wang, Y. & Sang, J. L. (2011) Protein phosphatase Pph3 and its regulatory subunit Psy2 regulate Rad53 dephosphorylation and cell morphogenesis during recovery from DNA damage in *Candida albicans*. *Eukaryot Cell*, 10, 1565-73.
- Tonks, N. K. & Neel, B. G. (1996) From form to function: signaling by protein tyrosine phosphatases. *Cell*, 87, 365-8.
- Trautmann, S., Wolfe, B. A., Jorgensen, P., Tyers, M., Gould, K. L. & Mccollum, D. (2001) Fission yeast Clp1p phosphatase regulates G2/M transition and coordination of cytokinesis with cell cycle progression. *Curr Biol*, 11, 931-40.
- Trevijano-Contador, N., Rueda, C. & Zaragoza, O. (2016) Fungal morphogenetic changes inside the mammalian host. *Semin Cell Dev Biol*, 57, 100-9.
- Trinkle-Mulcahy, L. (2012) Resolving protein interactions and complexes by affinity purification followed by label-based quantitative mass spectrometry. *Proteomics*, 12, 1623-38.
- Ubersax, J. A. & Ferrell, J. E., JR. (2007) Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol*, 8, 530-41.
- Visintin, R. & Amon, A. (2001) Regulation of the mitotic exit protein kinases Cdc15 and Dbf2. *Mol Biol Cell*, 12, 2961-74.
- Visintin, R., Craig, K., Hwang, E. S., Prinz, S., Tyers, M. & Amon, A. (1998) The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. *Mol Cell*, 2, 709-18.
- Visintin, R., Hwang, E. S. & Amon, A. (1999) Cfi1 prevents premature exit from mitosis by anchoring Cdc14 phosphatase in the nucleolus. *Nature*, 398, 818-23.
- Visintin, R., Prinz, S. & Amon, A. (1997) CDC20 and CDH1: a family of substrate-specific activators of APC-dependent proteolysis. *Science*, 278, 460-3.

- Wang, A., Raniga, P. P., Lane, S., Lu, Y. & Liu, H. (2009) Hyphal chain formation in *Candida albicans*: Cdc28-Hgc1 phosphorylation of Efg1 represses cell separation genes. *Mol Cell Biol*, 29, 4406-16.
- Wang, H., Huang, Z. X., Au Yong, J. Y., Zou, H., Zeng, G., Gao, J., Wang, Y., Wong, A. H. & Wang, Y. (2016) CDK phosphorylates the polarisome scaffold Spa2 to maintain its localization at the site of cell growth. *Mol Microbiol*, 101, 250-64.
- Wang, J., Liu, J., Hu, Y., Ying, S. H. & Feng, M. G. (2013) Cytokinesis-required Cdc14 is a signaling hub of asexual development and multi-stress tolerance in *Beauveria bassiana*. *Sci Rep*, 3, 3086.
- Wang, Y. (2016) Hgc1-Cdc28-how much does a single protein kinase do in the regulation of hyphal development in *Candida albicans*? *J Microbiol*, 54, 170-7.
- Whiteway, M. & Oberholzer, U. (2004) *Candida* morphogenesis and host-pathogen interactions. *Curr Opin Microbiol*, 7, 350-7.
- Willger, S. D., Liu, Z., Olarte, R. A., Adamo, M. E., Stajich, J. E., Myers, L. C., Kettenbach, A. N. & Hogan, D. A. (2015) Analysis of the *Candida albicans* Phosphoproteome. *Eukaryot Cell*, 14, 474-85.
- Wolfe, B. A. & Gould, K. L. (2004) Fission yeast Clp1p phosphatase affects G2/M transition and mitotic exit through Cdc25p inactivation. *Embo J*, 23, 919-29.
- Wolfe, B. A., McDonald, W. H., Yates, J. R., 3rd & Gould, K. L. (2006) Phospho-regulation of the Cdc14/Clp1 phosphatase delays late mitotic events in *S. pombe*. *Dev Cell*, 11, 423-30.
- Xue, L. & Tao, W. A. (2013) Current technologies to identify protein kinase substrates in high throughput. *Front Biol (Beijing)*, 8, 216-227.
- Xue, Y., Zhou, F., Zhu, M., Ahmed, K., Chen, G. & Yao, X. (2005) GPS: a comprehensive www server for phosphorylation sites prediction. *Nucleic Acids Res*, 33, W184-7.

- Yao, S., Neiman, A. & Prelich, G. (2000) BUR1 and BUR2 encode a divergent cyclin-dependent kinase-cyclin complex important for transcription in vivo. *Mol Cell Biol*, 20, 7080-7.
- Yellman, C. M. & Roeder, G. S. (2015) Cdc14 Early Anaphase Release, FEAR, Is Limited to the Nucleus and Dispensable for Efficient Mitotic Exit. *PLoS One*, 10, e0128604.
- Zhang, J., Yang, P. L. & Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer*, 9, 28-39.
- Zhang, R., Sioma, C. S., Wang, S. & Regnier, F. E. (2001) Fractionation of isotopically labeled peptides in quantitative proteomics. *Anal Chem*, 73, 5142-9.
- Zhao, J., Sun, X., Fang, J., Liu, W., Feng, C. & Jiang, L. (2010) Identification and characterization of the type 2C protein phosphatase Ptc4p in the human fungal pathogen *Candida albicans*. *Yeast*, 27, 149-57.
- Zhao, Y., Feng, J., Li, J. & Jiang, L. (2012) Mitochondrial type 2C protein phosphatases CaPtc5p, CaPtc6p, and CaPtc7p play vital roles in cellular responses to antifungal drugs and cadmium in *Candida albicans*. *FEMS Yeast Res*, 12, 897-906.
- Zheng, X., Wang, Y. & Wang, Y. (2004) Hgc1, a novel hypha-specific G1 cyclin-related protein regulates *Candida albicans* hyphal morphogenesis. *Embo J*, 23, 1845-56.
- Zheng, X. D., Lee, R. T., Wang, Y. M., Lin, Q. S. & Wang, Y. (2007) Phosphorylation of Rga2, a Cdc42 GAP, by CDK/Hgc1 is crucial for *Candida albicans* hyphal growth. *Embo J*, 26, 3760-9.
- Zhou, T., Aumais, J. P., Liu, X., Yu-Lee, L. Y. & Erikson, R. L. (2003) A role for Plk1 phosphorylation of NudC in cytokinesis. *Dev Cell*, 5, 127-38.