5-14-2010

# Cell Wall/Surface Proteome of Candida albicans: its Application in Rapid Identification of Yeast Species by Mass Signature and Characterization by in vitro and in vivo Chemical Labelings

Jiang Qian
*University of New Orleans*

# Cell Wall/Surface Proteome of *Candida albicans*: its Application in Rapid Identification of Yeast Species by Mass Signature and Characterization by *in vitro* and *in vivo* Chemical Labelings

A Dissertation

Submitted to the Graduate Faculty of the

University of New Orleans

in the partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in

The Department of Chemistry

By

Jiang Qian

B.S. Henan University of Technology, China 2000

M.S. Henan University of Technology, China 2004

May 2010

**To my family for all their love**

Father: Guanfu Qian

Mother: Linfen Xia

Wife: YanXia Zhou

Daughter: Trinity Qian

Daughter: Teresa Qian

Son: Jason Qian

# Acknowledgement

Completing this dissertation would not have been possible without supports from a group of people who have continuously helped me during my study at University of New Orleans and the Research Institute for Children, Children's Hospital, New Orleans. First and foremost, my advisor Dr. Yang Cai deserves my deepest appreciation for his guidance throughout the research journey, his instruction in proteomics and the independence he gave to me to pursue my interest in multidispilinary research area. Besides I want to express my appreciation to Dr. Jim Cutler from Louisiana State University Health Science Center for his edifying lessons on how to be become a decent scientist and for elucidating puzzels related to microbiology and biochemistry. I am also thankful to other committee members Dr. Bruce Gibb, Dr. Branko Jursic, Dr. Steven Rick for their constructive judgment in the general exam and the dissertation defense.

Additionally I wish to express my gratitude to Dr. Richard Cole for his generosity and trust in allowing me to learn and operate mass spectrometers in his lab.

I need to acknowledge Qiang Zhang from Xavier University at New Orleans for his helping me with interpreting NMR spectra of the synthesized compounds.

Finally I want to extend my thanks to Ping Wang's lab, Jim Cutler's lab and Seth Pincus' lab for all their support and help throughout my research study in the Research Institute for Children, Children's Hospital, New Orleans.

# TABLE OF CONTENTS

# LIST OF SCHEMES

## CHAPTER 3

# LIST OF FIGURES

## CHAPTER 1



## CHAPTER 2

## CHAPTER 3

## CHAPTER 4

# LIST OF TABLES

## CHAPTER 3



## CHAPTER 4



## APPENDIX

# ABBREVIATIONS

| | |
|---|---|
| 1D | One-dimentional |
| 2D | Two-dimentional |
| 2D-DIGE | Two-dimentional differential in-gel electrophoresis |
| BCA | Bicinchoninic acid |
| BSA | Bovine serum albumin |
| C | Cysteine |
| CGD | Candida genome database |
| CHCA | α-cyano-4-hydroxycinnamic acid |
| CI | Cluster ionization |
| CID | Collision iduced dissociation |
| CWP | Cell wall protein |
| Da | Dalten |
| DCM | Dichloromethane |
| DHB | 2,5-dihydroxybenzoic acid |
| DIGE | Two-dimentional differential in-gel electrophoresis |
| DMEM | Dulbecco's Modified Eagle's Medium |
| DMF | Dimethylformamide |
| DTT | Dithiothreitol |
| EDTA | Ethylene diamine tetraacetic acid |
| ESI | ElectroSpray Ionization |
| FA | Formic acid |
| FBPNHS | Boc-pentanoic-NHS-perfluoroctane (see A in figure 3.1) |
| FP1CMSNHS | sulfo-NHS-$(OEG)_1$-perfluoroctane (see B in figure 3.1) |
| FP3CMSNHS | Sulfo-NHS-$(OEG)_3$-perfluoroctane (see C in figure 3.1) |
| FT-IR | Fourier transform-infrared |
| GO | Gene Ontology |
| GPI | Glycosylphosphatidylinositol |
| GYEP | Glucose yeast extract peptone |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HPLC | High pressure liquid chromatography |
| IA | Iodoacetamide |
| ICAT | Isotope coded affinity tag |
| Iodoacetyl-PEO2-biotin | (N-Iodoacetyl-N-biotinylhexylenediamine |
| iTRAQ | Isobaric tag for relative and absolute quantitation |
| K | Lysine |
| LC | Liquid chromatography |
| LC-MS/MS | Liquid chromatography tandem mass spectrometry |

| | |
|---|---|
| MALDI-TOF MS | Matrix assisted laser desorption ionization-time of flight mass spectrometry |
| MS | Mass spectrometry |
| MS | Mass spectrometry |
| MS/MS | Tandem mass spectrometry |
| NHS | N-hydroxysuccinimidyl |
| NMR | Nuclear magnetic resonance |
| OEG | Oligo ethylene glycol |
| ORF | Open reading frame |
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phospate buffered saline |
| PCA | Principle component analysis |
| PCR | Polymerase chain reaction |
| PEG | Polyethylene glycol |
| PI | Photochemical ionization |
| Pir | Proteins with internal repeats |
| PMSF | Phenylmethanesulphonylfluoride |
| PNA-FISH | Peptide nucleic acid-fluorescence in situ hybridization |
| Py-MS | Pyrolysis-mass spectrometry |
| rpm | Revolutions per minute |
| SA | Sinapinic acid |
| SDS | Sodium dodecyl sulfate |
| SILAC | Stable isotope labeling by amino acids in cell culture |
| Sulfo-NHS-LC-biotin | Sulfosuccinimidyl-6-(biotinamido)hexanoate |
| Sulfo-NHS-SS-biotin | Sulfosuccinimidyl-2-[biotinamido]ethyl-1,3-dithiolpropionate |
| TCEP | Tris(2-carboxyethyl) phosphine hydrochloride |
| TFA | Trifluoroacetic acid |
| THF | Tetrahydrofuran |
| TLC | Thin layer chromatography |
| Tris | Tris(hydroxymethyl)aminomethane |
| UV | Ultroviolet |
| YNB | Yeast nitrogen base |
| YPD | Yeast extract peptone dextrose |

Note: Abbreviations of protein names are listed in Table S1

# ABSTRACT

*Candida albicans* is an opportunistic fungal pathogen that may cause mucutaneous infection and/or disseminated candidasis if the host defense system is impaired (such as those in HIV patients). Cell surface of *C. albicans* is the frontier where initial interplay between host-pathogen takes place and therefore is of great importance in understanding the mechanism of host-pathogen interaction. MALDI-TOF-MS analysis of intact fungal cells yielded mass signatures for rapid species differentiation, strain grouping and yeast morphogenesis monitoring.

Cell surface biotinylations at low temperature (4°C), enzymatic digestion of the intact fungal cell surface proteins ("whole cell shaving"), biotin-avidin affinity enrichment of biotinylated peptides, liquid chromatography mass spectrometry (LC-MS) based proteomic approach were employed for unambiguous identification of cell wall/cell wall associated proteins and the exposed peptide segments of these proteins. SILAC (Stable Isotope Labeling by Amino acids in Cell Culture) based CWP quantification analyses were performed to monitor CWP accumulation level change in response to hyphae induction. Information on surface exposed peptide segments and regulation of cell wall/surface protein during morphogenesis provided new candidates to the pool of potential peptide targets for protective vaccine development.

A New type of "fluorous" (fluorinated alkane) affinity gained popularity due to its low level non-specific protein/peptide binding. Fluorous labeling reagents that target primary amine groups in proteins/peptides were synthesized and characterized. The acid labile linker in the labeling reagents allows cleavage of the bulky fluorous tag moiety and the long oligo ethylene glycol (OEG) spacer after fluorous affinity purification. Upon collision induced decomposition, the

labeled peptide ion yielded a characteristic fragment that could be retrieved from the residual portion of fluorous affinity tag, and serve as a marker to indicate that the relevant peptide had been successfully labeled. Results showed that both the protein/peptide labeling and affinity enrichment/separation process were highly efficient.

**Keywords**

*Candida albicans*, Yeast differentiation, MALDI-TOF MS, Fluorous labeling reagent, LC-MS, Cell surface proteome, Surface biotinylation, SILAC

# CHAPTER 1

# Introduction

## 1.1 Overview of *Candida albicans*

*C. albicans* is a member of yeast kingdom that possesses basic biological features for the yeast species such as a singular cell with an oval shape, reproduction by budding or fission. However, its opportunistic pathogenicity distinguished *C. albican* from its benign relative *Saccharomyces cerevisiae* that is commonly used as baker's yeast in our daily lives. *C. albicans* commensally exists in most healthy human populations and can usually be observed in the oral cavity, gastro-intenstinal and urogential tracts. The tipping point of the balance relies on both the host immuno-system and the fungal traits. This facultative pathogen may turn into a pathogenic agent of infection if the host defense system is impaired. The severity of the fungal infection, depending on the nature and extent of the impairment of the host immune-system, can vary from superficial mucocutaneous infections to life-threatening, invasive infection (124). Over the last few decades, reports of human fungal infection have been increasing at an alarming rate worldwide (132), especially in immuno-compromised patients like HIV infected individuals, patients undergoing transplanting surgery, chemotherapy and other immuno-suppressed treatments (25, 45). Candidiasis is now the fourth leading cause of nosocomial bloodstream infections in the US (83), and for many years the overwhelming leading cause of this disease has been *C. albicans.*

*C. albicans* is a dimorphic organism, which has the ability to transform singular budding yeast to filamentous hyphae cells under environmental factors such as growth temperature at 37°C, serum medium and near neutral pH (9). This morphogenic development is recognized as a key

virulence factor of the fungal species during infection as it enables attacks on host cells either through epithelium invasion or endothelium barrier breaching (29) (82). As a matter of fact, it was shown that non-filamentous mutants were avirulent in murine models (88). One of the key features during the morphological change from yeast to hyphae is the cell wall remodeling, which is companied not only by composition change of cell wall but also by differential regulations of cell wall proteins. Previous study showed that it was easier for *C. albicans* hyphae cells to escape from the capture of mammalian macrophages than yeast cells (88). Information on the quantitative change of cell wall protein expression level during the conversion may shed lights on the path to understand what renders the hyphae more virulent compared to the yeast form morphogenesis.

The cell surface is of critical consideration in the attempts to understand morphogenesis and immunologic target identification, as the cell surface is the forefront where the initial contact between host-fungus occurs and such contact may induce immune response from the host. Understanding of the pathogen-host interaction may provide insight into the functional mechanism of fungal invasion. Mechanisms by which *C. albicans* cause disease are still largely unknown. Although progress has been made on identification of immunologic targets that can be recognized by either antibodies or cell-mediated immune defenses (30) more information is needed before immuno-approaches to prevention and treatment of candidiasis become a reality.

### 1.2 MALDI -TOF Mass Spectrometry

Matrix-assisted laser-desorption/ionization (MALDI) was first introduced in late 1980s by Karas, Hillenkamp, (76) (78) and Tanaka (128). This revolutionary method for ionizing large molecules

blazed new trails in the mass spectrometry field along with another parallel ionization technique, electrospray ionization (ESI), ushered by Fenn (42). Since then both of these two techniques have been the most successful ionizing methods for large bio-molecules such as proteins, nucleotides, and carbohydrates. In MALDI, analytes are usually co-crystalized with certain types of matrix, which transfer the absorbed large excess of laser energy (UV or IR wavelengths) to the neighboring analyte molecules during the excitation process. A plum of ions was released upon laser irradiation, accelerated by electrical potential and analyzed by a mass spectrometer as illustrated in Figure 1.1.

Photochemical ionization (PI) (39) and cluster ionization (CI) (77) have been proposed as the two major models for explaining the mechanism of MALDI process. Matrix molecules were initially excited through energy pooling and multiphoton extraction from laser irradiation in the PI model. Analyte ions were consequently generated after protons were transferred to the analyte molecules. The PI model has been commonly accepted as the primary mechanism for MALDI , although it can not be applied to explain some experimental phenomena such as matrix suppression effect (80) and "sweet spot" (87). Contrary to the PI model (39), the CI model (77) redefined the origin of analyte ions by proposing a large protonated analyte cluster clouded with excessive amount of matrix molecules. Singly charged analyte ions were released from the large analyte-matrix cloud after "shaking off" neutral matrix molecules in gas phase. The ionization process in the CI model is analogous to the solvent desolvation stepin ESI. However, the CI model failed to explain why large biomolecules were singly charged in the gas phase while the same type of molecules were multiply charged in ESI. Recently, Chen and co-workers (20) proposed an energy transfer induced disproportionation (ETID) model which enabled them to

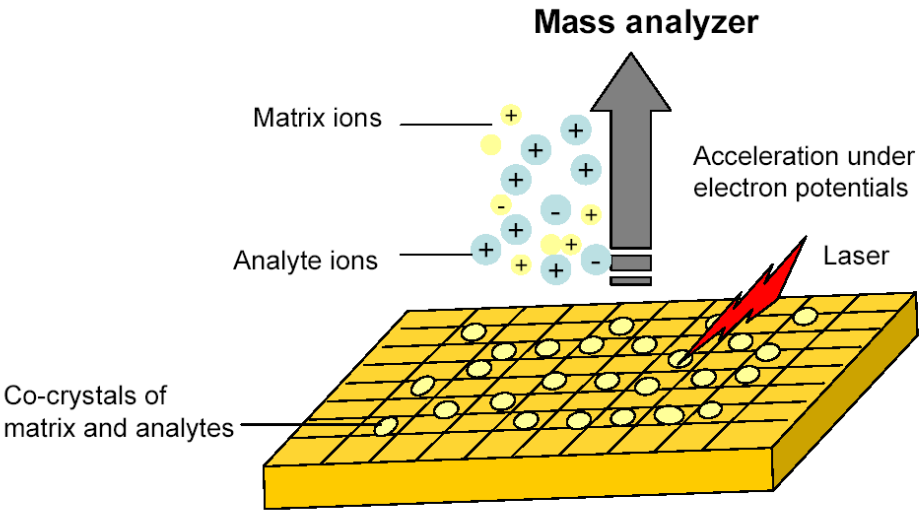



Figure 1.1. Schematic overview of matrix-assisted laser desorption/ionization (MALDI).

explain the observation of equal number of positive and negative analyte ions in MALDI.

Matrix selection plays pivotal roles in the success of MALDI analysis. Matrices are usually aromatic compounds with low vapor pressure. Nicotinic acid (NA) was the first matrix applied in ionizing proteins and peptides (78). A broad range of matrices have been surveyed to determine if a compound is suitable for the ionization of a particular type of analyte molecule. For example, picolinic acid (PA) and 3-hydroxypicolinic acid (HPA) have been used for oligonucleotides and DNA analyses, (129) (140) 2,5-dihydroxybenzoic acid (DHB) for oligosaccharides, (127) α-cyano-4-hydroxycinnamic acid (CHCA) for peptides (8), and sinapinic acid (SA) for proteins (7). In addition to matrix selection, sample preparation is also an important factor for a successful MALDI analysis. The typical molar ratio of matrix-to-analyte is kept between 500 : 1 and 5000 : 1, (62) in order to ensure a good signal-to-noise ratio while minimizing the "matrix effect". Also, co-crystallization of a mixture of analyte and matrix is critical in analysis. Unevenly formed co-crystals can cause poor reproducibility due to the presence of "sweet spots", only from which good quality mass spectrum may be obtained. "Thin-layer" depositing technique can improve the co-crystallization process by slow evaporation of low volatile solvents such as acetone and acetonitrile (134). Upon irradiation by a focused laser beam, the matrix and the analyte molecules embedded in the co-crystals leave the MALDI target plate and enter into the gas phase in the form of a plume of ions and neutral molecules. Those singly charged ions are then accelerated through an electronic potential field, separated on the basis of their molecular weight and finally recorded at a mass detector.

### 1.3 Advantages and challenges in microbe identification by MALDI-TOF MS

World-wide infections caused by various microorganisms have been constant threats to human health, even though progress had been made in combating those infections by new inventions in medicine. Application of mass spectrometry based methods in microorganism detection has been fueled by the advances in mass spectrometry instrumentation recently. Short analysis time (a few minutes) is a major advantage of MS-based characterization of microorganism when compared with other methods such as the classical polymerase chain reaction (PCR) based approaches. Moreover, the MS-based methods usually do not pre-specify targets, whereas DNA primer designs in PCR require predetermined targets. MS spectral data can be readily processed and incorporated into a reference library, which can subsequently be searched to identify unknown samples (67).

The first mass spectrometry based application for bacteria identification was reported in 1975 by Anhalt et al (3), who demonstrated the feasibility of differentiating pathogenic bacterial species via pyrolysis-mass spectrometry (Py-MS). Basile et al (6) reduced the analysis time of bacterial samples to minutes by combining thermal hydrolysis and methylation step with Py-MS. Characteristic mass spectra were also obtained through fast atom bombardment mass spectrometry analysis of lysed bacteria (60). It was not until the late 1990s that MALDI and ESI were introduced to generate mass spectral data for bacterial species (63) (24) and viral species (34). Since then, MALDI-MS analyses of toxigenic and pathogenic bacteria were extensively reported in the literature(58) (139) (64, 92, 121) (133, 141). Experimental factors that potentially affect the quality and reproducibility of the mass spectra obtained from intact whole bacteria cells were defined for identification and species differentiation (64). Bacterial species readily yield abundant peaks in mass spectra without prior sample treatment. Several automated MALDI

mass spectra comparison programs were also developed to facilitate rapid identification and speciation of intact bacteria. Mathematical algorithms in automated spectra comparison programs were usually based on a cross correlation between two mass spectra (5) or derived from a statistical test of significance with reliance on specific bio-makers (38, 46, 67, 116). Cluster analysis such as Phylogenetic Analysis Using Parsimony (PAUP*) had been employed to compare the MALDI mass spectra of *Aeromonas* species and determine the relationships among the analyzed species (35). Principle component analysis (PCA) is another popular multivariate analysis which can be employed to extract features in dataset generated from mass spectra and project each spectrum to a single spot on a PCA plot (57). Compared to ESI-MS, rapid MALDI-MS analysis has several advantageous features: 1. there is no clogging problem in the ion source; 2. mostly singly charged ions simplify the acquired spectrum (32).

Compared to the growing number of bacterial analysis by MALDI, analysis of fungi was rarely reported. A few reports in the literature showed varying degree of success (2, 21, 66, 86). If the fungal cell wall is untreated, very few peaks appear on MALDI spectra intact fungal cells as a result of low yield of medium size ions (5000 – 15000 Da). To address this problem, partial disruption of cell wall matrix may be necessary. Cell wall digestion is one of the methods that facilitate MALDI-MS analyses of fungal species (43). Other treatments such as strong acid extraction and detergent solubilization have also been employed (43). These fungal cell wall pretreatments were usually time consuming and not entirely compatible to MALDI-MS analysis due to the suppression effect of detergents. Development of simple MS-friendly sample treatment method is an important consideration in fungal analysis by MALDI-MS.

### 1.4 Organization of the cell wall of *C .albicans*

The *C. albicans* cell wall is a rigid matrix that mainly consists of polysaccharides (about 80-90%) represented by β-1,3 and β-1,6 glucans, chitin and mannan. Proteins (6-25%) and a small amount of lipid (1 to 7%) are also integrated into the cell wall (19). The cell wall is partitioned into several layers distinguished by different densities when studied by electron microscopy. Cell wall of *C. albicans* is usually composed of three to eight layers (112). An electron-dense outer layer space is mainly composed of mannoproteins, which can be detected by concana-valin A (ConA) labeling (90) or specific antibody (69). The electron transparent inner layer is largely composed of a net work of highly branched polysaccharides and a small number of proteins. Yeast and hyphae forms share similar percent composition of the cell wall, although the relative amounts of individual component differ significantly (15). The cell wall is highly dynamic and is involved in almost all aspects of the biology and the interactions of this fungus with the host (18) (19) while it maintains a balance between rigidity and plasticity. The rigidity of the cell wall provides strong protection for the intracellular components against environmental stress, while the plasticity allows the cell wall to adapt to different stages of development during morphogenesis of yeast to hyphal forms or yeast to pseudohyphal forms. The rigidity of the cell wall is largely attributed to β-glucans and chitin near the plasma membrane (74) (93, 110).

### 1.5 Cell wall proteins and their functions in *C. albicans*

Besides the role in maintaining the architecture of the wall, polysaccharides, the major component of the cell wall, may also serve as ligands for receptor proteins from host immune cells (33) (44, 52, 100). However, the protein population of the cell wall is more likely to distinguish *C. albicans* as pathogenic yeast from non-pathogenic *S.cerevisiae* (124). The outer

layer of the cell wall is replete with various mannoproteins, which are either covalently bound to the inner layer of glucan/chitin networks through β-1,3 and/or β-1,6 linkage or loosely associated with cell wall matrix (74, 110). There are two classes of covalently bound cell wall proteins (CWPs): GPI (Glycosylphosphatidylinositol)-CWPs and Pir (Proteins with internal repeats)-CWPs. GPI-CWPs are those proteins with C-termini linked to the reducing end of β-1,6 glucan chain via a phosphodiester bond, which is then integrated into a inner flexible network of β-1,3 glucan chains or chitin chains (72, 81). In fungi, some of those GPI proteins are released from the membrane and incorporated with the cell wall matrix (48, 56). The GPI-CWPs can be extracted through hydrolysis of the phosphodiester bond using HF-pyridine. Pir-CWPs (130) are linked directly to the network of β-1,3 glucan chains without any bridging from of β-1,6 glucan and can be cleaved from the glucan chains by NaOH treatment(75, 99). Despite significant amount of studies on GPI-CWPs and Pir-CWPs, information is very fragmentary on other CWPs covalently bound either to β-1,3 glucan or chitin via uncharacterized linkages. Those proteins are usually extracted through enzymatic treatment like Quantazyme (β-1,3 glucanase) and/or exochitinase digestion in which by the polysaccharide moieties of the glycoproteins were destroyed (110). Additionally, a large set of CWPs do not have strong attachment to the cell wall matrix (or connected to the covalently bound CWPs merely through disulfide bond). Some of the proteins related to the cell wall are viewed as secreted proteins (95) (84) because they are usually also observed in culture medium. Hot sodium dodecyl sulfate (SDS) buffer containing reducing reagents (such as dithiolthreitol) is commonly used to strip those noncovalently associated CWPs from the cell wall matrix.

Bona fide cell wall proteins are accepted as those that meet the following criteria (115): 1. Proteins have a secretory motif (N-terminal secretion signal peptides). 2. Proteins have other characteristic features such as GPI-binding motif or inner repeats in addition to their glycosylation modifications. Based on these standards, hundreds of cell wall proteins in *C. albicans* were predicted from ORFs analysis (31). This paradigm, however, has never been challenged with the accumulating evidence of detected nonglycosylated proteins (37, 110) and/or numerous proteins without signal motif on the surface (103). Therefore, the concept of cell wall protein should be redefined based on experimental findings in order to accommodate those non-canonical wall proteins.

Even though the fungal cell wall is comprised of highly diverse proteins, five main functions have been assigned to those CWPs. 1). Some proteins are involved in enzymatic degradation of large molecules which could serve as a nutritional source. Other protein enzymes play roles in maintaining the architecture of cell wall polymers by their involvement in the pathways of polysaccharide degradation/synthesis;(19) 2). Some proteins are involved in initiating "cross-talk" with host cells. Adhesins are typical fimbriae mannoproteins that facilitate the attachment of the fungal cell to host cell (epithelial, endothelial or platelets) through four major recognition systems; (115) 3). Antigenicity is another important function shared by different cell wall proteins from both yeast and hyphal cells (17). Those proteins with immunogenic motifs are good candidates for vaccine development; 4). Certain cell wall proteins may contribute to virulence factors on cell surface as evidenced by their involvement in pathogenesis (47).

### 1.6 ESI-MS based proteomic approach to studying cell surface of *C. albicans* and current challenges

In this post-genomic era, advancement in proteomics has been fueled by the development of versatile mass spectrometry instrumentation. The term "proteomics" was coined by Marc Wilkins in 1994 (36, 138) with a notion of 2D gel based analysis of total protein expressed by a genome. The extension of this concept, however, is far more comprehensive then its original definition and still rapidly expanding driven by the demand for answering new questions arising from biology studies. Proteomics is now related not only to protein sequencing and proteome profiling in a certain compartment of an organism, but also to the study on the dynamic change of protein accumulation levels under various conditions, post modifications of proteins and protein-protein interactions. MS-oriented proteomics is undoubtedly the mainstream in studying proteins in biological field that is benefited from discovery and development of protein ionization techniques.

Two "soft" ionization techniques, ESI (42) and MALDI, (76) are most commonly employed mass spectrometry based analyses of proteins and peptides. In ESI, high electric potential is applied to analyte solution to form charged droplets. Attracted by a potential gradient and a pressure gradient along the ion passage, the droplet undergoes a desolvation process, which leads to downsize the droplet. The process continues by constant evaporation of the solvent (and ions in some cases) until fully desolvated ions are formed. ESI-MS is usually coupled with liquid chromatography (LC) to form an on-line LC-MS system that has become the most powerful instrumentation for peptide sequencing in proteomics field (Figure 1.2.).

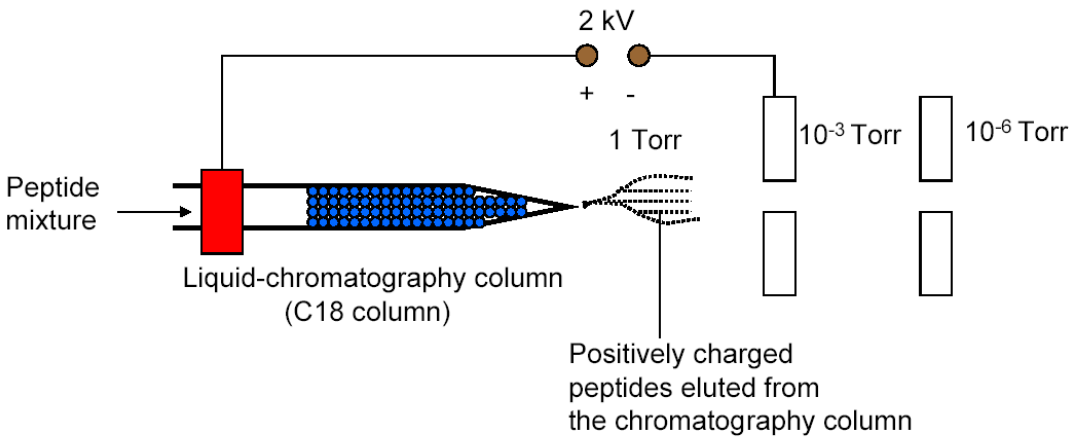



Figure 1.2. Schematic presentation of liquid chromatography (LC)-mass spectrometry (MS) analysis of a peptide mixture in positive mode.

In most tandem mass spectrometry-based proteomic approaches, the proteins are chemically or enzymatically digested to yield smaller peptides and these peptides are subsequently analyzed by tandem mass spectrometry. Database searches are performed on the tandem mass spectra of peptides to determine the peptides' amino acid sequences, and protein identification is achieved by the determination (via database search) of the unique peptide(s) derived from this protein (135). The *C. albicans* protein sequence database (available from http://www.candidagenome.org/) facilitates the identification of *C. albicans* proteins via database search.

The advancement of MS techniques and availability of the full genome sequence of *C. albicans* seem to make it theoretically straightforward to catalogue cell wall proteins of the fungus. However, little progress has been made. Part of the reason is that the heterogeneity of the cell wall proteins (covalently linked to cell wall) makes sample preparation a formidable task. Complicated sequential fractionation is required to isolate those proteins from the polysaccharides network under various chemical and/or enzymatic treatments either from intact cells (4, 16, 79, 89, 123) or cell wall pellet after cell lyses (70, 110, 132) (40, 71, 98). The fractionated CWPs were further separated by two-dimensional gel electrophoresis (2D-PAGE) before LC-MS analysis. This fractionation based sample preparation method may provide some additional information about how CWPs are linked to the network of polysaccharides. Unfortunately, the detection of the identical wall proteins in different fractions adds ambiguity in protein localization. Also, this conventional approach has several limitations: First, the cell wall fractions collected after cell lyses were subject to potential cytosolic contamination, even after extensive washing by NaCl solution of various concentrations. While some researchers (132) had

shown that cytosolic components released from cell lyses didn't interact with cell wall fractions in their approach, others reported that the possibility of cytosolic contamination in the cell wall fractions could not be unequivocally ruled out (41) (94). Second, fractionating steps are time consuming. Moreover, the extraction process is often incomplete and strain dependent (73). Third, 2D-gel electrophoresis may not resolve glycoproteins extracted from yeast cell walls (142) because the same protein may have different glycoforms with different isoelectric points (51) and molecular weights (144), which may complicate the comparative gel based quantitative analysis of CWPs because of their multiple locations in the gel.

### 1.7 Biotinylation technique for cell surface proteome determination

Labeling intact fungal cells by membrane impermeable biotinylating reagents is an attractive approach toward minimizing the influence of cytosolic contamination. The advantage of this method is that the biotin reagents react only with cell wall proteins (or their carbohydrate moiety) on or proteins adjacent to the cell surface. The labeling strategies and reagents share these common features including: 1. the biotin tag for affinity chromatography; 2. chemical groups reactive with proteins at lysine (K) or cysteine (C) residues or carbohydrate moieties of glycoproteins; 3. a spacer of variable length that links the biotin moiety to the protein. The strong affinity between avidin and biotin was first recognized in 1941 (55), but it was not until 1990 when a biotin-avidin affinity based purification technique was introduced by Wilchek (137). Due to the very low dissociation constant ($K_d \approx 10^{-15}$ ) of the biotin-avidin complex (136), the captured biotinylate proteins are stable and released from avidin affinity column only under harsh elution conditions such as 8 M guanidine HCl. The recovery of sample is therefore often incomplete (117) and irreversible protein damage may occur during the elution process. This

problem may be partly solved either by choosing modified avidins with lower affinity such as nitrated avidin derivatives (96), CaptAvidin (97) and monomeric avidin (61) or using cleavable biotinylating reagents (e.g. Sulfo-NHS-SS-Biotin, Pierce, Rockford, IL) which allows the dissociation of labeled protein from the biotin-avidin complex through chemical cleavage of the linker (e.g. disulfide bond). The lower affinity avidins increase the recovery of biotinylated molecules as a result of reduced biotin-avidin binding strength, but they are also less tolerant to detergents, which are necessary in solubilization of membrane proteins co-precipitated with the cell walls. A disulfide bond linker may also be problematic because it could be susceptible to endogenous reduction prior to affinity purification (11). In order to increase the efficiency of binding between large biotinylated biomolecules and immobilized avidin, efforts were made to incorporate a long spacer into the labeling reagents, which would reduce the steric hindrance imposed by the bulky biomolecule and facilitate the docking of the biotin moiety into the avidin binding pocket. The most popular spacer available on the market is a monodisperse PEG (polyethylene glycol) based linker, which not only allows the efficient capture of the labeled molecule, but also prevents the aggregation of labeled protein due to the hydrophilic nature of the linker. It should be noted that mono-PEG spacers have an advantage over their aliphatic counterparts because of their lower immunogenicity *in vivo* (59).

The most important part of the labeling reagent is the protein/peptide reactive moiety that enables biotin conjugation to proteins. An active N-hydroxysuccinimidyl (NHS) ester is the most commonly used functional group to modify free amine groups in proteins/peptides at pH 7.5-8.5. Sulfo-NHS ester is water soluble and its biotin derivatives are widely used in labeling cell surface proteins *in vivo* due to membrane impermeability of the labeling reagents (negative charged sulfonate group on NHS ring prevents the penetration of sulfo-NHS-esters of biotin

through the plasma membrane). Other important functional groups that target cysteine residues are malemide and iodoacetyl moieties, which react with the free sulfhydryl group in cysteine residue at pH 6.5-7.5 and 7.5-8.5 respectively. It is noteworthy that the pH of the labeling reaction mixture needs be strictly controlled to minimize side reactions such as transformation of malemide by primary amines at pH above 7.5 and iodoacetyl by imidazoles at pH 5.0-7.0.

Sulfosuccinimidyl-6-(biotinamido)hexanoate (sulfo-NHS-LC-biotin) and sulfosuccinimidyl-2-(biotinamido)ethyl-1,3-dithiolpropionate (sulfo-NHS-SS-biotin) (see Figure 1.3. for their structures) are the most commonly used protein biotinylating reagents that react with free amines from the side chain of lysine residues or N-termini of proteins/peptides, and these labeling reagents have recently gained much popularity in studying membrane subproteome of mammalian cells (22, 105, 108, 119, 122, 131), cell surface proteome of bacteria (118) (113) and fungi species (38, 46, 116). In the field of *C. abicans* research, investigators have also benefited from this technique as well as the advancement in mass spectrometry instrumentation (4, 16) (132) (68) (89). For example, Urban and co-workers (132) employed sulfo-NHS-biotin labeling technique to monitor the expression of Tsa1p protein in different cell localizations. As mentioned earlier, the biotinylation labeling reagents are membrane impermeable due to the sulfonate group in sulfo-NHS-biotin molecule(125). Thus it is surprising that classic cytosolic and nuclear proteins were also found in of sulfo-NHS-biotin labeled protein fractions (103, 104). Although it is possible that the labeling reagent may penetrates the membranes through certain ion channels (e.g $Na^+$) (143) or enters freely into the aging or dead cells whose membrane integrity has been damaged, it is more likely that the cytosolic and nuclear proteins were non-specifically bound to avidin during biotin-avidin based affinity chromatography (94).

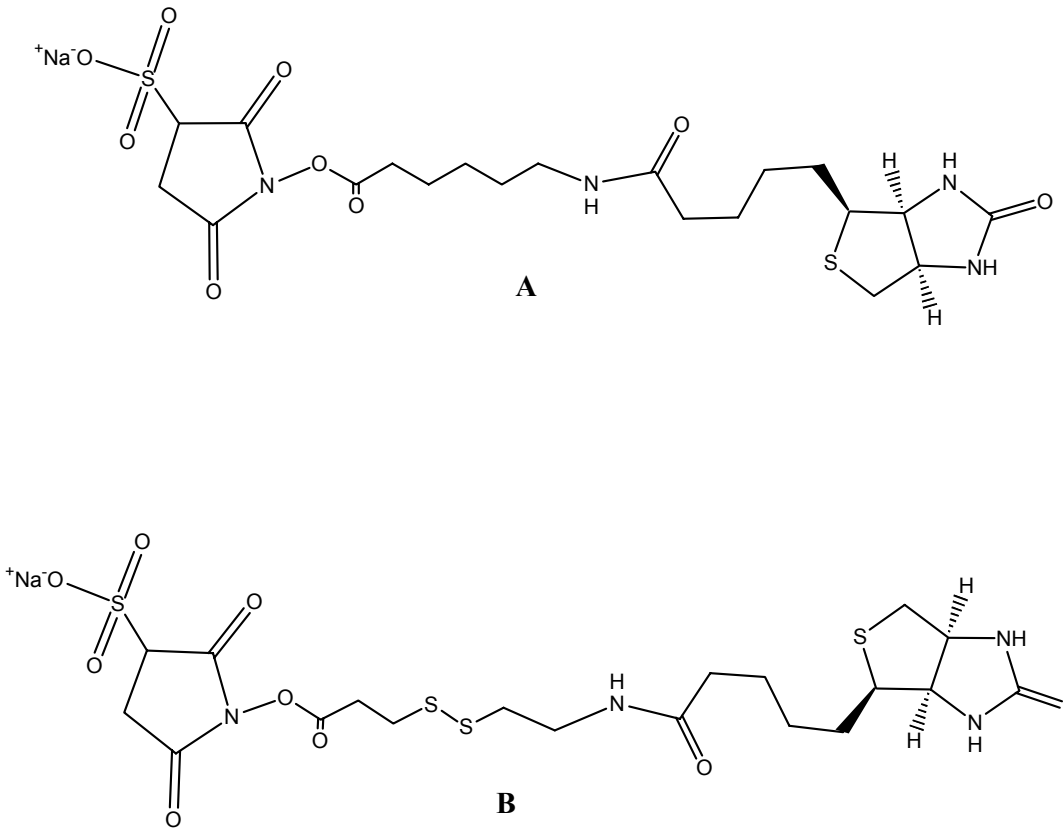



Figure 1.3. Structures of sulfosuccinimidyl-6-(biotinamido)hexanoate (sulfo-NHS-LC-biotin) (A) and sulfosuccinimidyl-2-(biotinamido)ethyl-1,3-dithiolpropionate (sulfo-NHS-SS-biotin) (B)

Fortunately, such problem can be addressed by coupling the surface labeling technique with tandem mass spectrometry-based proteomic tools to determine the biotinylation site(s) of proteins.

### 1.8 Study on dynamic change of cell surface proteome of *C. albicans* during the morphogenesis

Conversion from the yeast form to the hyphal form has become generally recognized as a key virulence factor when *C. albicans* cells attack host cells either through epithelium invasion or endothelium barrier breaching (29) (82). Compared to the CWP profile of *C. albicans*, the expression change of CWPs during morphogenesis is largely unexplored. Only a few reports (37) (110) related to systematic comparison of cell wall proteomes of yeast and hyphal forms were reported. Recently, Gil and co-workers (91) conducted a comparative proteomic analysis on the yeast, hyphae and biofims of *C. albicans*. 2D-gel based approaches were employed in those studies but yielded limited information due to drawbacks in 2D-gel based quantification (e.g. low reproducibility and inaccurate measurement of gel spot volume). A two-dimentional differential in-gel electrophoresis (2D-DIGE ) technique (1) could be employed to improve the quantification of gel spot, but such approach still has the limitations inherited by all 2D-gel/MS based quantitative proteomic methods as discussed above(109).

Stable isotopic labeling techniques, especially ICAT® (54), iTRAQ® (114), and SILAC (107) methods gained considerable popularity and were recently reviewed (23, 49, 102, 120). These labeling techniques involve both in vitro labeling via attaching isotopic tags to cysteine residues in protein (ICAT®) or to free amino group in peptide (iTRAQ®) through chemical reactions,

and *in vivo* metabolic incorporation of stable isotopes into the proteins expressed by cells cultured in media supplemented by different isotopes (SILAC). Relative quantification is achieved by comparative analysis of ion intensities of peptide isotope pairs in the mass spectrum (ICAT®, SILAC) or by measuring the intensities of report ions derived from the labeled peptides in the tandem mass spectrum (iTRAQ®). The stable isotope labeling technology offers accurate quantification of protein accumulation levels from various sources with good reproducibility. A *C. albicans* lysine auxotrophic mutant was generated (10) that, when used along the SILAC technique, will allow us to monitor the CWP accumulation level changes during a yeast to hypae transformation of this pathogenic fungus.

**1.9 Development of "fluorous" affinity tag based labeling reagents**

Although biotin-avidin affinity based approaches are widely employed in the enrichment/separation of biological molecules, several limitations have to be considered when commercially available biotinylating reagents are employed to label proteins/peptides(50): 1. Some unlabeled proteins/peptides may nonspecifically bind to avidin (94), resulting in false identification; 2. The recovery of the biotinylated molecules is often low (26) (126) because of the high stability of biotin-avidin complex (136); 3. The biotin moiety tends to average the overall hydrophobicity of the tagged-peptides and compromise the separation of these peptides on reverse-phase liquid chromatography column(85); 4. Fragments originated from the biotin moiety can complicate the mass spectral interpretation of biotinylated peptides (106). 5. Biotin moiety may be cleaved under biotinidase catalyzed enzymatic reactions or through other even nonenzymatic mechanisms *in vivo* (12, 13). All these limitations call for the development of new

labeling reagents with non-biotin affinity tags for enrichment/separation of biological molecules including proteins/peptides.

The strong non-covalent interaction between fluorinated alkyl group (fluorine-fluorine interactions) makes the fluorous tag one of the alternative approach to biotin labeling. (27) The term “fluorous” was initially coined in 1990s to denote molecules with high content of fluorine atoms (perfluorinated species). The perfluorinated moiety enables fluorous organic compounds (molecules attached by perfluorinated alkyl chain, like $C_8F_{17}$) to be separated from non-fluorous compounds by fluorous-functionalized silica gel through solid phase extraction. Non-fluorous compounds are first washed away by an anti-fluorous isocratic solvent, such as 60% methanol in water. While the affinity between fluorous moieties is strong enough to allow fluorous molecules to be retained on immobilized fluorous carriers after the washing, the fluorous interaction can be disrupted under very mild conditions. Fluorphilic solvents such as methanol and acetonitrile may serve to elute the fluorous fraction. Interestingly, fluorous compounds can be further fractionated based on the different number of fluorous tags attached to a molecule by tuning the fluorophilicity of the elution solvent. Molecules with higher fluorine content would be eluted later with more fluorophilic solvent (14) (28).

The fluorous affinity technique was first introduced in fluorous biphasic catalysis technique (65) and soon utilized in a wide spectrum of organic syntheses (145). Recently, its application in enriching tagged peptides from highly complex biological mixtures was also demonstrated (14). The technique started to gain popularity in metabolomics (53), and microarrays for immobilizing small-biomolecules (101, 111).

## References:


1. **Alban, A., S. O. David, L. Bjorkesten, C. Andersson, E. Sloge, S. Lewis, and I. Currie.** 2003. A novel experimental design for comparative two-dimensional gel analysis: Two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics **3:**36-44.
2. **Amiri-Eliasi, B. J., and C. Fenselau.** 2001. Characterization of protein biomarkers desorbed by MALDI from whole fungal cells. Analytical Chemistry **73:**5228-5231.
3. **Anhalt, J. P., and C. Fenselau.** 1975. Identification of Bacteria Using Mass-Spectrometry. Analytical Chemistry **47:**219-225.
4. **Apaire-Marchais, V., J. Cottin, A. Marot-Leblond, C. Lefrancois, G. Tronchin, and R. Robert.** 2005. *In vitro* and *in vivo* cell surface expression of a thiol-specific antioxidant-like protein in *Candida albicans*. Journal de Mycologie Medicale **15:**1-12.
5. **Arnold, R. J., and J. P. Reilly.** 1998. Fingerprint matching of E-coli strains with matrix-assisted laser desorption ionization time-of-flight mass spectrometry of whole cells using a modified correlation approach. Rapid Communications in Mass Spectrometry **12:**630-636.
6. **Basile, F., M. B. Beverly, K. J. Voorhees, and T. L. Hadfield.** 1998. Pathogenic bacteria: their detection and differentiation by rapid lipid profiling with pyrolysis mass spectrometry. Trac-Trends in Analytical Chemistry **17:**95-109.
7. **Beavis, R. C., and B. T. Chait.** 1989. Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. Rapid Communications in Mass Spectrometry **3:**432-435.
8. **Beavis, R. C., T. Chaudhary, and B. T. Chait.** 1992. Alpha-cyano-4-hydroxycinnamic acid as a matrix for matrix-assisted laser desorption mass-spectrometry. Organic Mass Spectrometry **27:**156-158.
9. **Berman, J., and P. E. Sudbery.** 2002. *Candida albicans*: A molecular revolution built on lessons from budding yeast. Nature Reviews Genetics **3:**918-930.
10. **Bhattacherjee, V., and J. K. Bhattacharjee.** 1999. Characterization of a double gene disruption in the LYS2 locus of the pathogenic yeast, *Candida albicans*. Medical Mycology **37:**411-417.
11. **Bickel, U., Y. S. Kang, and W. M. Pardridge.** 1995. *In-vivo* cleavability of A disulfide-based chimeric opioid peptide in rat-brain. Bioconjugate Chemistry **6:**211-218.
12. **Bogusiewcz, A., N. I. Mock, and D. M. Mock.** 2004. Release of biotin from biotinylated proteins occurs enzymatically and nonenzymatically in human plasma. Analytical Biochemistry **331:**260-266.
13. **Bogusiewicz, A., N. I. Mock, and D. M. Mock.** 2004. Instability of the biotin-protein bond in human plasma. Analytical Biochemistry **327:**156-161.
14. **Brittain, S. M., S. B. Ficarro, A. Brock, and E. C. Peters.** 2005. Enrichment and analysis of peptide subsets using fluorous affinity tags and mass spectrometry. Nature Biotechnology **23:**463-468.
15. **Calderone, R. A., and P. C. Braun.** 1991. Adherence and receptor relationships of *Candida albicans*. Microbiological Reviews **55:**1-20.
16. **Casanova, M., J. L. Lopezribot, J. P. Martinez, and R. Sentandreu.** 1992. Characterization of cell-wall proteins from yeast and mycelial cells of *Candida albicans*

by labeling with biotin - comparison with other techniques. Infection and Immunity **60:**4898-4906.

17. **Casanova, M., J. P. Martinez, and W. L. Chaffin.** 1990. Fab fragments from a monoclonal-antibody against a germ tube mannoprotein block the yeast-to-mycelium transition in *Candida albicans*. Infection and Immunity **58:**3810-3812.
18. **Cassone, A.** 1989. Cell wall of *Candida albicans*: its functions and its impact on the host. Current Topics in Medical Mycology **3:**248-314.
19. **Chaffin, W. L., J. L. Lopez-Ribot, M. Casanova, D. Gozalbo, and J. P. Martinez.** 1998. Cell wall and secreted proteins of *Candida albicans*: Identification, function, and expression. Microbiology and Molecular Biology Reviews **62:**130-180.
20. **Chang, W. C., L. C. L. Huang, Y. S. Wang, W. P. Peng, H. C. Chang, N. Y. Hsu, W. B. Yang, and C. H. Chen.** 2007. Matrix-assisted laser desorption/ionization (MALDI) mechanism revisited. Analytica Chimica Acta **582:**1-9.
21. **Chen, H. Y., and Y. C. Chen.** 2005. Characterization of intact *Penicillium* spores by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Communications in Mass Spectrometry **19:**3564-3568.
22. **Chen, W. N. U., L. R. Yu, E. F. Strittmatter, B. D. Thrall, D. G. Camp, and R. D. Smith.** 2003. Detection of in situ labeled cell surface proteins by mass spectrometry: Application to the membrane subproteome of human mammary epithelial cells. Proteomics **3:**1647-1651.
23. **Chien, K. Y., and M. B. Goshe.** 2009. Advances in quantitative mass spectrometry analysis: Weighing in on isotope coding and label-free approaches for expression and functional proteomics. Current Analytical Chemistry **5:**166-185.
24. **Claydon, M. A., S. N. Davey, V. EdwardsJones, and D. B. Gordon.** 1996. The rapid identification of intact microorganisms using mass spectrometry. Nature Biotechnology **14:**1584-1586.
25. **Corner, B. E., and P. T. Magee.** 1997. *Candida* pathogenesis: Unravelling the threads of infection. Current Biology **7:**R691-R694.
26. **Cronan, J. E.** 1990. Biotination of proteins *in vivo* - A posttranslational modification to label, purify, and study proteins. Journal of Biological Chemistry **265:**10327-10333.
27. **Curran, D. P.** 2008. Chemistry - Fluorous tags unstick messy chemical biology problems. Science **321:**1645-1646.
28. **Curran, D. P., and Z. Y. Luo.** 1999. Fluorous synthesis with fewer fluorines (light fluorous synthesis): separation of tagged from untagged products by solid-phase extraction with fluorous reverse-phase silica gel. Journal of the American Chemical Society **121:**9069-9072.
29. **Cutler, J. E.** 1991. Putative virulence factors of *Candida albicans*. Annual Review of Microbiology **45:**187-218.
30. **Cutler, J. E., G. S. Deepe, and B. S. Klein.** 2007. Advances in combating fungal diseases: vaccines on the threshold. Nature Reviews Microbiology **5:**13-28.
31. **De Groot, P. W. J., H. L. Dekker, A. D. De Boer, K. J. Hellingwerf, C. G. De Koster, and F. M. Klis.** 2003. Identification of cell wall proteins of the fungal pathogen *Candida albicans* and other fungi using mass-spectrometric and genome-wide computational approaches. Yeast **20:**S57-S57.
32. **Demirev, P. A., and C. Fenselau.** 2008. Mass spectrometry for rapid characterization of microorganisms. Annual Review of Analytical Chemistry **1:**71-93.

33. **Dennehy, K. M., and G. D. Brown.** 2007. The role of the beta-glucan receptor dectin-1 in control of fungal infection. Journal of Leukocyte Biology **82:**253-258.
34. **Despeyroux, D., R. Phillpotts, and P. Watts.** 1996. Electrospray mass spectrometry for detection and characterization of purified cricket paralysis virus (CrPV). Rapid communications in mass spectrometry **10:**937-941.
35. **Donohue, M. J., A. W. Smallwood, S. Pfaller, M. Rodgers, and J. A. Shoemaker.** 2006. The development of a matrix-assisted laser desorption/ionization mass spectrometry-based method for the protein fingerprinting and identification of *Aeromonas* species using whole cells. Journal of Microbiological Methods **65:**380-389.
36. **Dunn, M. J.** 1995. 2d Electrophoresis - From protein maps to genomes - proceedings of the internationalmeeting - Siena, September 5-7, 1994. Electrophoresis **16:**U3-U4.
37. **Ebanks, R. O., K. Chisholm, S. McKinnon, M. Whiteway, and D. M. Pinto.** 2006. Proteomic analysis of *Candida albicans* yeast and hyphal cell wall and associated proteins. Proteomics **6:**2147-2156.
38. **Edwards, S. R., R. Braley, and W. L. Chaffin.** 1999. Enolase is present in the cell wall of *Saccharomyces cerevisiae*. Fems Microbiology Letters **177:**211-216.
39. **Ehring, H., M. Karas, and F. Hillenkamp.** 1992. Role of photoionization and photochemistry in ionization processes of organic-molecules and relevance for matrix-assisted laser desorption ionization mass-spectrometry. Organic Mass Spectrometry **27:**472-480.
40. **Elorza, M. V., A. Murgui, and R. Sentandreu.** 1985. Dimorphism in *Candida albicans* - contribution of mannoproteins to the architecture of yeast and mycelial cell-walls. Journal of General Microbiology **131:**2209-2216.
41. **Eroles, P., M. Sentandreu, M. V. Elorza, and R. Sentandreu.** 1997. The highly immunogenic enolase and Hsp70p are adventitious *Candida albicans* cell wall proteins. Microbiology-Uk **143:**313-320.
42. **Fenn, J. B., M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse.** 1989. Electrospray ionization for mass-spectrometry of large biomolecules. Science **246:**64-71.
43. **Fenselau, C., and P. A. Demirev.** 2001. Characterization of intact microorganisms by MALDI mass spectrometry. Mass Spectrometry Reviews **20:**157-171.
44. **Filler, S. G.** 2006. Candida-host cell receptor-ligand interactions. Current Opinion In Microbiology **9:**333-339.
45. **Fisherhoch, S. P., and L. Hutwagner.** 1995. Opportunistic Candidiasis - An epidemic of The 1980s. Clinical Infectious Diseases **21:**897-904.
46. **Foster, A. J., R. A. Bird, and S. N. Smith.** 2007. Biotinylation and characterization of *Cryptococcus neoformans* cell surface proteins. Journal of Applied Microbiology **103:**390-399.
47. **Fradin, C., P. De Groot, D. MacCallum, M. Schaller, F. Klis, F. C. Odds, and B. Hube.** 2005. Granulocytes govern the transcriptional response, morphology and proliferation of *Candida albicans* in human blood. Molecular Microbiology **56:**397-415.
48. **Frieman, M. B., and B. P. Cormack.** 2003. The omega-site sequence of glycosylphosphatidylinositol-anchored proteins in Saccharomyces cerevisiae can determine distribution between the membrane and the cell wall. Molecular Microbiology **50:**883-896.
49. **Gant-Branum, R. L., T. J. Kerr, and J. A. McLean.** 2009. Labeling strategies in mass spectrometry-based protein quantitation. Analyst **134:**1525-1530.

50. **Gauthier, D. J., B. F. Gibbs, N. Rabah, and C. Lazure.** 2004. Utilization of a new biotinylation reagent in the development of a nondiscriminatory investigative approach for the study of cell surface proteins. Proteomics **4:**3783-3790.
51. **Gemmill, T. R., and R. B. Trimble.** 1999. Overview of N- and O-linked oligosaccharide structures found in various yeast species. Biochimica Et Biophysica Acta-General Subjects **1426:**227-237.
52. **Gil, M. L., and D. Gozalbo.** 2006. TLR2, but not TLR4, triggers cytokine production by murine cells in response to Candida albicans yeasts and hyphae. Microbes and Infection **8:**2299-2304.
53. **Go, E. P., W. Uritboonthai, J. V. Apon, S. A. Trauger, A. Nordstrom, G. O'Maille, S. M. Brittain, E. C. Peters, and G. Siuzdak.** 2007. Selective metabolite and peptide capture/mass detection using fluorous affinity tags. Journal of Proteome Research **6:**1492-1499.
54. **Gygi, S. P., B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold.** 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnology **17:**994-999.
55. **Gyorgy, P., C. S. Rose, R. E. Eakin, E. E. Snell, and R. J. Williams.** 1941. Egg-white injury as the result of nonabsorption or inactivation of biotin. Science **93:**477-478.
56. **Hamada, K., H. Terashima, M. Arisawa, N. Yabuki, and K. Kitada.** 1999. Amino acid residues in the omega-minus region participate in cellular localization of yeast glycosylphosphatidylinositol-attached proteins. Journal of Bacteriology **181:**3886-3889.
57. **Harrington, P. D., T. E. Street, K. J. Voorhees, F. R. Dibrozolo, and R. W. Odom.** 1989. Rule-building expert system for classification of mass-spectra. Analytical Chemistry **61:**715-719.
58. **Hathout, Y., P. A. Demirev, Y. P. Ho, J. L. Bundy, V. Ryzhov, L. Sapp, J. Stutler, J. Jackman, and C. Fenselau.** 1999. Identification of *Bacillus* spores by matrix-assisted laser desorption ionization mass spectrometry. Applied and Environmental Microbiology **65:**4313-4319.
59. **He, X. H., P. C. Shaw, and S. C. Tam.** 1999. Reducing the immunogenicity and improving the in vivo activity of trichosanthin by site-directed PEGylation. Life Sciences **65:**355-368.
60. **Heller, D. N., R. J. Cotter, and C. Fenselau.** 1987. Profiling of bacteria by fast-atom-bombardment mass-spectrometry. Analytical Chemistry **59:**2806-2809.
61. **Henrikson, K. P., S. H. G. Allen, and W. L. Maloy.** 1979. Avidin monomer affinity column for the purification of biotin-containing enzymes. Analytical Biochemistry **94:**366-370.
62. **Hillenkamp, F., M. Karas, R. C. Beavis, and B. T. Chait.** 1991. Matrix-assisted laser desorption ionization mass-spectrometry of biopolymers. Analytical Chemistry **63:**A1193-A1202.
63. **Holland, R. D., J. G. Wilkes, F. Rafii, J. B. Sutherland, C. C. Persons, K. J. Voorhees, and J. O. Lay.** 1996. Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry **10:**1227-1232.
64. **Horneffer, V., J. Haverkamp, H. G. Janssen, and R. Notz.** 2004. MALDI-TOF-MS analysis of bacterial spores: Wet heat-treatment as a new releasing technique for

biomarkers and the influence of different experimental parameters and microbiological handling. Journal of the American Society for Mass Spectrometry **15:**1444-1454.
65. **Horvath, I. T., and J. Rabai.** 1994. Facile catalyst separation without water - Fluorous biphase hydroformylation of olefins. Science **266:**72-75.
66. **Jackson, K. A., V. Edwards-Jones, C. W. Sutton, and A. J. Fox.** 2005. Optimisation of intact cell MALDI method for fingerprinting of methicillin-resistant *Staphylococcus aureus*. Journal of Microbiological Methods **62:**273-284.
67. **Jarman, K. H., S. T. Cebula, A. J. Saenz, C. E. Petersen, N. B. Valentine, M. T. Kingsley, and K. L. Wahl.** 2000. An algorithm for automated bacterial identification using matrix-assisted laser desorption/ionization mass spectrometry. Analytical Chemistry **72:**1217-1223.
68. **Kandasamy, R., G. Vediyappan, and W. L. Chaffin.** 2000. Evidence for the presence of Pir-like proteins in *Candida albicans*. Fems Microbiology Letters **186:**239-243.
69. **Kapteyn, J. C., L. L. Hoyer, J. E. Hecht, W. H. Muller, A. Andel, A. J. Verkleij, M. Makarow, H. Van Den Ende, and F. M. Klis.** 2000. The cell wall architecture of *Candida albicans* wild-type cells and cell wall-defective mutants. Molecular Microbiology **35:**601-611.
70. **Kapteyn, J. C., R. C. Montijn, G. J. P. Dijkgraaf, and F. M. Klis.** 1994. Identification of beta-1,6-glucosylated cell-wall proteins in yeast and hyphal forms of *Candida albicans*. European Journal of Cell Biology **65:**402-407.
71. **Kapteyn, J. C., R. C. Montijn, G. J. P. Dijkgraaf, H. Vandenende, and F. M. Klis.** 1995. Covalent association of beta-1,3-glucan with beta-1,6-glucosylated mannoproteins in cell-walls of *Candida albicans*. Journal of Bacteriology **177:**3788-3792.
72. **Kapteyn, J. C., R. C. Montijn, E. Vink, J. delaCruz, A. Llobell, J. E. Douwes, H. Shimoi, P. N. Lipke, and F. M. Klis.** 1996. Retention of *Saccharomyces cerevisiae* cell wall proteins through a phosphodiester-linked beta-1,3-/beta-1,6-glucan heteropolymer. Glycobiology **6:**337-345.
73. **Kapteyn, J. C., A. F. J. Ram, E. M. Groos, R. Kollar, R. C. Montijn, H. VandenEnde, A. Llobell, E. Cabib, and F. M. Klis.** 1997. Altered extent of cross-linking of beta 1,6-glucosylated mannoproteins to chitin in *Saccharomyces cerevisiae* mutants with reduced cell wall beta 1,3-glucan content. Journal of Bacteriology **179:**6279-6284.
74. **Kapteyn, J. C., H. Van Den Ende, and F. M. Klis.** 1999. The contribution of cell wall proteins to the organization of the yeast cell wall. Biochimica Et Biophysica Acta-General Subjects **1426:**373-383.
75. **Kapteyn, J. C., P. Van Egmond, E. Sievi, H. Van Den Ende, M. Makarow, and F. M. Klis.** 1999. The contribution of the O-glycosylated protein Pir2p/Hsp150 to the construction of the yeast cell wall in wild-type cells and beta 1,6-glucan-deficient mutants. Molecular Microbiology **31:**1835-1844.
76. **Karas, M., D. Bachmann, U. Bahr, and F. Hillenkamp.** 1987. Matrix-assisted ultraviolet-laser desorption of nonvolatile compounds. International Journal of Mass Spectrometry and Ion Processes **78:**53-68.
77. **Karas, M., M. Gluckmann, and J. Schafer.** 2000. Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. Journal of Mass Spectrometry **35:**1-12.

78. **Karas, M., and F. Hillenkamp.** 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Analytical Chemistry **60:**2299-2301.
79. **Klis, F. M., M. de Jong, S. Brul, and P. W. J. de Groot.** 2007. Extraction of cell surface-associated proteins from living yeast cells. Yeast **24:**253-258.
80. **Knochenmuss, R., F. Dubois, M. J. Dale, and R. Zenobi.** 1996. The matrix suppression effect and ionization mechanisms in matrix-assisted laser desorption/ionization. Rapid Communications in Mass Spectrometry **10:**871-877.
81. **Kollar, R., B. B. Reinhold, E. Petrakova, H. J. C. Yeh, G. Ashwell, J. Drgonova, J. C. Kapteyn, F. M. Klis, and E. Cabib.** 1997. Architecture of the yeast cell wall - beta(1->6)-glucan interconnects mannoprotein, beta(1-3)-glucan, and chitin. Journal of Biological Chemistry **272:**17762-17775.
82. **Kumamoto, C. A., and M. D. Vinces.** 2005. Contributions of hyphae and hypha-co-regulated genes to *Candida albicans* virulence. Cellular Microbiology **7:**1546-1554.
83. **Lazzell, A. L., A. K. Chaturvedi, C. G. Pierce, D. Prasad, P. Uppuluri, and J. L. Lopez-Ribot.** 2009. *Treatment and prevention of* Candida albicans biofilms with caspofungin in a novel central venous catheter murine model of candidiasis. Journal of Antimicrobial Chemotherapy **63:**567-570.
84. **Lee, S. A., S. Wormsley, S. Kamoun, A. F. S. Lee, K. Joiner, and B. Wong.** 2003. An analysis of the *Candida albicans* genome database for soluble secreted proteins using computer-based prediction algorithms. Yeast **20:**595-610.
85. **Li, J. X., H. Steen, and S. P. Gygi.** 2003. Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents - The yeast salinity stress response. Molecular & Cellular Proteomics **2:**1198-1204.
86. **Li, T. Y., B. H. Liu, and Y. C. Chen.** 2000. Characterization of *Aspergillus* spores by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry **14:**2393-2400.
87. **Liu, Y. D., X. Y. Sun, and B. C. Guo.** 2003. Matrix-assisted laser desorption/ionization time-of-flight analysis of low-concentration oligonucleotides and mini-sequencing products. Rapid Communications in Mass Spectrometry **17:**2354-2360.
88. **Lo, H. J., J. R. Kohler, B. DiDomenico, D. Loebenberg, A. Cacciapuoti, and G. R. Fink.** 1997. Nonfilamentous *C. albicans* mutants are avirulent. Cell **90:**939-949.
89. **LopezRibot, J. L., H. M. Alloush, B. J. Masten, and W. L. Chaffin.** 1996. Evidence for presence in the cell wall of *Candida albicans* of a protein related to the hsp70 family. Infection and Immunity **64:**3333-3340.
90. **Marcilla, A., M. V. Elorza, S. Mormeneo, H. Rico, and R. Sentandreu.** 1991. Candida albicans mycelial wall structure - Supramolecular complexes released by zymolyase, chitinase and beta-mercaptoethanol. Archives of Microbiology **155:**312-319.
91. **Martinez-Gomariz, M., P. Perumal, S. Mekala, C. Nombela, W. L. Chaffin, and C. Gil.** 2009. Proteomic analysis of cytoplasmic and surface proteins from yeast cells, hyphae, and biofilms of *Candida albicans*. Proteomics **9:**2230-2252.
92. **Marvin-Guy, L. F., S. Parche, S. Wagniere, J. Moulin, R. Zink, M. Kussmann, and L. B. Fay.** 2004. Rapid identification of stress-related fingerprint from whole bacterial cells of Bifidobacterium lactis using matrix assisted laser desorption/ionization mass spectrometry. Journal of the American Society for Mass Spectrometry **15:**1222-1227.

93. **Masuoka, J.** 2004. Surface glycans of *Candida albicans* and other pathogenic fungi: Physiological roles, clinical uses, and experimental challenges. Clinical Microbiology Reviews **17:**281-+.
94. **Masuoka, J., L. N. Guthrie, and K. C. Hazen.** 2002. Complications in cell-surface labelling by biotinylation of *Candida albicans* due to avidin conjugate binding to cell-wall proteins. Microbiology-Sgm **148:**1073-1079.
95. **Monteoliva, L., M. L. Matas, C. Gil, C. Nombela, and J. Pla.** 2002. Large-scale identification of putative exported proteins in *Candida albicans* by genetic selection. Eukaryotic Cell **1:**514-525.
96. **Morag, E., E. A. Bayer, and M. Wilchek.** 1996. Immobilized nitro-avidin and nitro-streptavidin as reusable affinity matrices for application in avidin-biotin technology. Analytical Biochemistry **243:**257-263.
97. **Morag, E., E. A. Bayer, and M. Wilchek.** 1996. Reversibility of biotin-binding by selective modification of tyrosine in avidin. Biochemical Journal **316:**193-199.
98. **Mormeneo, S., H. Rico, M. Iranzo, C. Aguado, and R. Sentandreu.** 1996. Study of supramolecular structures released from the cell wall of *Candida albicans* by ethylenediamine treatment. Archives of Microbiology **166:**327-335.
99. **Mrsa, V., T. Seidl, M. Gentzsch, and W. Tanner.** 1997. Specific labelling of cell wall proteins by biotinylation. Identification of four covalently linked O-mannosylated proteins of *Saccharomyces cerevisiae*. Yeast **13:**1145-1154.
100. **Netea, M. G., J. W. M. Van der Meer, and B. J. Kullberg.** 2006. Role of the dual interaction of fungal pathogens with pattern recognition receptors in the activation and modulation of host defense. Clinical Microbiology and Infection **12:**404-409.
101. **Nicholson, R. L., M. L. Ladlow, and D. R. Spring.** 2007. Fluorous tagged small molecule microarrays. Chemical Communications**:**3906-3908.
102. **Nita-Lazar, A., H. Saito-Benz, and F. M. White.** 2008. Quantitative phosphoproteomics by mass spectrometry: Past, present, and future. Proteomics **8:**4433-4443.
103. **Nombela, C., C. Gil, and W. L. Chaffin.** 2006. Non-conventional protein secretion in yeast. Trends in Microbiology **14:**15-21.
104. **Nosanchuk, J. D., J. N. Steenbergen, L. Shi, G. S. Deepe, and A. Casadevall.** 2003. Antibodies to a cell surface histon-like protein protect against *Histoplasma capsulatum*. Journal of Clinical Investigation **112:**1164-1175.
105. **Nunomura, K., K. Nagano, C. Itagaki, M. Taoka, N. Okamura, Y. Yamauchi, S. Sugano, N. Takahashi, T. Izumi, and T. Isobe.** 2005. Cell surface labeling and mass spectrometry reveal diversity of cell surface markers and signaling molecules expressed in undifferentiated mouse embryonic stem cells. Molecular & Cellular Proteomics **4:**1968-1976.
106. **Oda, Y., T. Nagasu, and B. T. Chait.** 2001. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nature Biotechnology **19:**379-382.
107. **Ong, S. E., B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, and M. Mann.** 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular & Cellular Proteomics **1:**376-386.

108. **Peirce, M. J., R. Wait, S. Begum, J. Saklatvala, and A. P. Cope.** 2004. Expression profiling of lymphocyte plasma membrane proteins. Molecular & Cellular Proteomics **3:**56-65.
109. **Peng, J. M., and S. P. Gygi.** 2001. Proteomics: the move to mixtures. Journal of Mass Spectrometry **36:**1083-1091.
110. **Pitarch, A., M. Sanchez, C. Nombela, and C. Gil.** 2002. Sequential fractionation and two-dimensional gel analysis unravels the complexity of the dimorphic fungus *Candida albicans* cell wall proteome. Molecular & Cellular Proteomics **1:**967-982.
111. **Pohl, N. L.** 2006. Fluorous-based microarrays for probing carbohydrate-protein recognition. Abstracts of Papers of the American Chemical Society **231**.
112. **Poulain, D., G. Tronchin, J. F. Dubremetz, and J. Biguet.** 1978. Ultrastructure of cell-wall of *Candida albicans* blastospores - study of its constitutive layers by use of a cytochemical technique revealing polysaccharides. Annales de Microbiologie **A129:**141-&.
113. **Rodriguez-Ortega, M. J., N. Norais, G. Bensi, S. Liberatori, S. Capo, M. Mora, M. Scarselli, F. Doro, G. Ferrari, I. Garaguso, T. Maggi, A. Neumann, A. Covre, J. L. Telford, and G. Grandi.** 2006. Characterization and identification of vaccine candidate proteins through analysis of the group *A Streptococcus* surface proteome. Nature Biotechnology **24:**191-197.
114. **Ross, P. L., Y. L. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, and D. J. Pappin.** 2004. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular & Cellular Proteomics **3:**1154-1169.
115. **Ruiz-Herrera, J., M. V. Elorza, E. Valentin, and R. Sentandreu.** 2006. Molecular organization of the cell wall of *Candida albicans* and its relation to pathogenicity. Fems Yeast Research **6:**14-29.
116. **Ruiz-Herrera, J., A. I. Martinez, and R. Sentandreu.** 2002. Determination of the stability of protein pools from the cell wall of fungi. Research in Microbiology **153:**373-378.
117. **Rybak, J. N., S. B. Scheurer, D. Neri, and G. Elia.** 2004. Purification of biotinylated proteins on streptavidin resin: A protocol for quantitative elution. Proteomics **4:**2296-2299.
118. **Sabarth, N., S. Lamer, U. Zimny-Arndt, P. R. Jungblut, T. F. Meyer, and D. Bumann.** 2002. Identification of surface proteins of *Helicobacter pylori* by selective biotinylation, affinity purification, and two-dimensional gel electrophoresis. Journal of Biological Chemistry **277:**27896-27902.
119. **Scheurer, S. B., C. Roesli, D. Neri, and G. Elia.** 2005. A comparison of different biotinylation reagents, tryptic digestion procedures, and mass spectrometric techniques for 2-D peptide mapping of membrane proteins. Proteomics **5:**3035-3039.
120. **Schuchardt, S., and J. Borlak.** 2008. Quantitative mass spectrometry to investigate epidermal growth factor receptor phosphorylation dynamics. Mass Spectrometry Reviews **27:**51-65.
121. **Shaw, E. I., H. Moura, A. R. Woolfitt, M. Ospina, H. A. Thompson, and J. R. Barr.** 2004. Identification of biomarkers of whole *Coxiella burnetii* phase I by MALDI-TOF mass spectrometry. Analytical Chemistry **76:**4017-4022.

122. **Shin, B. K., H. Wang, A. M. Yim, F. Le Naour, F. Brichory, J. H. Jang, R. Zhao, E. Puravs, J. Tra, C. W. Michael, D. E. Misek, and S. M. Hanash.** 2003. Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function. Journal of Biological Chemistry **278:**7607-7616.
123. **Singleton, D. R., and K. C. Hazen.** 2004. Differential surface localization and temperature-dependent expression of the *Candida albicans* CSH1 protein. Microbiology-Sgm **150:**285-292.
124. **Sohn, K., J. Schwenk, C. Urban, J. Lechner, M. Schweikert, and S. Rupp.** 2006. Getting in touch with *Candida albicans*: The cell wall of a fungal pathogen. Current Drug Targets **7:**505-512.
125. **Staros, J. V.** 1982. N-hydroxysulfosuccinimide active esters - Bis(N-Hydroxysulfosuccinimide) esters of 2 dicarboxylic-acids are Hydrophilic, membrane-impermeant, protein cross-linkers. Biochemistry **21:**3950-3955.
126. **Steen, H., and M. Mann.** 2002. A new derivatization strategy for the analysis of phosphopeptides by precursor ion scanning in positive ion mode. Journal of the American Society for Mass Spectrometry **13:**996-1003.
127. **Strupat, K., M. Karas, and F. Hillenkamp.** 1991. 2,5-Dihydroxybenzoic acid - A new matrix for laser desorption ionization mass-spectrometry. International Journal of Mass Spectrometry and Ion Processes **111:**89-102.
128. **Tanaka, K., H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, and T. Matsuo.** 1988. Protein and polymer analyses up to *m/z* 100000 by laser ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry **2:**151-153.
129. **Tang, K., N. I. Taranenko, S. L. Allman, C. H. Chen, L. Y. Chang, and K. B. Jacobson.** 1994. Picolinic-acid as a matrix for laser mass-spectrometry of nucleic-acids and proteins. Rapid Communications in Mass Spectrometry **8:**673-677.
130. **Tohe, A., S. Yasunaga, H. Nisogi, K. Tanaka, T. Oguchi, and Y. Matsui.** 1993. 3 yeast genes, Pir1, Pir2 And Pir3, containing internal tandem repeats, are related to each other, and Pir1 and Pir2 are required for tolerance to heat-shock. Yeast **9:**481-494.
131. **Trotti, D., M. Aoki, P. Pasinelli, U. V. Berger, N. C. Danbolt, R. H. Brown, and M. A. Hediger.** 2001. Amyotrophic lateral sclerosis-linked glutamate transporter mutant has impaired glutamate clearance capacity. Journal of Biological Chemistry **276:**576-582.
132. **Urban, C., K. Sohn, F. Lottspeich, H. Brunner, and S. Rupp.** 2003. Identification of cell surface determinants in *Candida albicans* reveals Tsa1p, a protein differentially localized in the cell. Febs Letters **544:**228-235.
133. **Vargha, M., Z. Takats, A. Konopka, and C. H. Nakatsu.** 2006. Optimization of MALDI-TOF MS for strain level differentiation of *Arthrobacter* isolates. Journal of Microbiological Methods **66:**399-409.
134. **Vorm, O., P. Roepstorff, and M. Mann.** 1994. Improved resolution and very high-sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. Analytical Chemistry **66:**3281-3287.
135. **Washburn, M. P., D. Wolters, and J. R. Yates.** 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nature Biotechnology **19:**242-247.
136. **Wilchek, M., and E. A. Bayer.** 1999. Foreword and introduction to the book (strept)avidin-biotin system. Biomolecular Engineering **16:**1-4.

137. **Wilchek, M., and E. A. Bayer.** 1990. Introduction to avidin-biotin technology. Methods in Enzymology **184:**5-13.
138. **Wilkins, M. R., C. Pasquali, R. D. Appel, K. Ou, O. Golaz, J. C. Sanchez, J. X. Yan, A. A. Gooley, G. Hughes, I. HumpherySmith, K. L. Williams, and D. F. Hochstrasser.** 1996. From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Bio-Technology **14:**61-65.
139. **Williams, T. L., D. Andrzejewski, J. O. Lay, and S. M. Musser.** 2003. Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. Journal of the American Society for Mass Spectrometry **14:**342-351.
140. **Wu, K. J., A. Steding, and C. H. Becker.** 1993. Matrix-assisted laser desorption time-of-flight mass-spectrometry of oligonucleotides using 3-hydroxypicolinic acid as an ultraviolet-sensitive matrix. Rapid Communications in Mass Spectrometry **7:**142-146.
141. **Wunschel, S. C., K. H. Jarman, C. E. Petersen, N. B. Valentine, K. L. Wahl, D. Schauki, J. Jackman, C. P. Nelson, and E. White.** 2005. Bacterial analysis by MALDI-TOF mass spectrometry: An inter-lab oratory comparison. Journal of the American Society for Mass Spectrometry **16:**456-462.
142. **Yin, Q. Y., P. W. J. de Groot, C. G. de Koster, and F. M. Klis.** 2008. Mass spectrometry-based proteomics of fungal wall glycoproteins. Trends in Microbiology **16:**20-26.
143. **Yu, M. J., T. Pisitkun, G. H. Wang, R. F. Shen, and M. A. Knepper.** 2006. LC-MS/MS analysis of apical and basolateral plasma membranes of rat renal collecting duct cells. Molecular & Cellular Proteomics **5:**2131-2145.
144. **Zeng, C. H., and K. Biemann.** 1999. Determination of N-linked glycosylation of yeast external invertase by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Journal of Mass Spectrometry **34:**311-329.
145. **Zhang, W., and D. P. Curran.** 2006. Synthetic applications of fluorous solid-phase extraction (F-SPE). Tetrahedron **62:**11837-11865.

# CHAPTER 2

# Rapid Identification of Yeast Species and Differentiation of *C. albicans* Strains by Mass Signatures Acquired by Matrix-assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-TOF-MS)

## 2.1 Abstract

Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS) was used for rapid identification of yeasts, including strains of *Candida albicans*. Intact yeasts were collected from surface colonies on agar-based media, washed in deionized water followed by 50% methanol, suspended in deionized water to a concentration of approximately $10^6$ cells/ μl and 1 μl of each suspension was deposited into a well of a MALDI target plate. Among several matrix compounds investigated, sinapinic acid crystallization worked especially well. Under these conditions, a sufficient number of medium size ions (4 to 15 kDa) were detectable in the MALDI mass spectra that could serve as "mass signatures" for unambiguous identification of the various yeast species and even among strains of *C. albicans*. Resulting mass signatures were influenced by many factors, including growth medium, length of incubation and cell preparation. Under conditions that were optimized for the greatest number of ions with highest intensities, *C. albicans* was distinguishable from other *Candida* species (*C. lusitaniae, C. dubliniensis, C.glabrata, C. rusei, C. kefyr*) and from other yeast genera (*Cryptococcus neoformans , Saccharomyces cerevisiae and a Rhodotorula sp*.). Within the species *C. albicans* species, several medium sized ions in the 5 to 10 kDa range appeared on the mass spectra of all the investigated strains. In addition to those shared ions, the mass spectrum of individual strains

contained unique ions that could be used to distinguish strains of this species. Finally, *C. albicans* germ tubes produced MALDI-TOF MS signatures that differed from yeast forms of this species. This is a simple, rapid and sensitive method of identifying yeasts. Identification of fungal molecules yielding signature mass spectra is under investigation.

**Keywords**

Mass signature, MALDI-TOF mass spectrometry, Alcohol fixation, Yeast differentiation, Principal component analysis

## 2.2 Introduction

Several *Candida* species, and most especially *C. albicans,* are opportunistic pathogens that are among the most common causes of fungal infections. Since the advent and widespread use of the azole (such as fluconazole) group of antifungals which are especially inhibitory for *C. albicans*, other *Candida spp*. resistant to these drugs have emerged with greater frequency, notably like *Candida glabrata* (13). Although *C. albicans* remains the cause of the majority of cases of disseminated candidiasis, and with resulting high mortality rates, the increase in occurrence due to several other *Candida spp*. and their tendency for resistance against standard antifungal drugs, necessitates identification to the species level. Moreover other yeast genera are also important etiologic agents of disease and some are common contaminants in clinical and research laboratories (9, 20, 36, 39). Rapid identification of yeasts to the species level for clinical purposes, distinguishing strains within a given species for epidemiological investigations and for strain verification purposes in clinical and research laboratories, and monitoring yeast surface development in pathogenesis studies are compelling demands that continuously drive technical advances in the field. There are a plethora of rapid identification methods, ranging from species-specific chromophore development on specialized media to molecular genetic approaches (37). Some of the methods are simple and especially applicable to the presumptive identification of *C. albicans* from other yeasts, whereas others provide at least rapid presumptive identification of most species and can even distinguish strains of a given species, but may be either time consuming and/or require considerable expertise. The approach of peptide nucleic acid-fluorescence in situ hybridization (PNA-FISH) shows potential for rapid identification of yeast directly in clinical specimens (42). A notable advantage of the PNA FISH is rapid determination since prior culturing of the etiologic agent is not required. Despite this impressive development

and other approaches cited above, a single method has not been described that is capable of handling all of the above applications, including strain grouping and monitoring developmentally expressed fungal surface characteristics.

Spectroscopic approaches such as pyrolysis-mass spectrometry (Py-MS), Fourier transform-infrared (FT-IR) spectroscopy (45), nuclear magnetic resonance (NMR) (21, 22), and Raman spectroscopy (25, 32) have been used for *Candida* species differentiation, but with mixed results and significant limitations. Either the spectra showed few qualitative differences between species (45), or the methods and analyses were too complex for practical application (46) (21, 33). In addition, none have shown utility for either strain grouping or for detection of developmental cell surface moieties.

Matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is an emerging tool for microbe characterization and differentiation at the species and strain levels (12). Identification of microbes by MALDI-TOF MS requires a relatively small number of intact cells ($10^6$ -$10^7$) along with simple sample preparation procedures. Rapid data acquisition (in seconds) by MALDI-TOF MS and the high specificity of mass spectrometric analysis relative to the above mentioned techniques allows fast and accurate identification under well-defined conditions. Early reports focused on rapid identification of intact whole bacteria (23, 28), since bacteria naturally yield abundant mass spectral signals under standard conditions. These studies were followed by MALDI-MS identification of toxigenic and pathogenic bacteria (10) (15, 19) (29, 31, 34) (41) (48, 51, 52), factors affecting spectral quality and reproducibility were defined

(24, 51), and automated MALDI mass spectra comparison programs were developed to facilitate rapid identification and speciation of intact bacteria.

Studies of fungal species by MALDI-TOF mass spectrometry (1) (8, 26) (30, 47) are more recent, but unlike simple preparation of bacteria for analysis, investigators report the need for multiple sample preparation steps, such as cell wall lysis (1) in order to increase the intensities of diagnostic signals for species differentiation. Thus, fungi do appear to produce characteristic MALDI-MS signals, but the signals may be more difficult to obtain as compared to those of bacteria. Our first aim in this study was to optimize the MALDI-TOF MS analysis of intact yeast cells and then show the potential application of this technique toward identifying yeast species, grouping strains within a species, following developmental cell surface changes. Mathematical analysis of mass spectra from intact cells was also applied, which should be a useful first step in the development of an automated yeast identification system.

## 2.3 Materials and Methods

### 2.3.1 Yeast strains

*C. albicans* strain 3153A (serotype A) was obtained from the American Type Culture Collection (ATCC 28367). Strain A9 (serotype B) was originally isolated from the oral cavity of an AIDS patient (50) and has been used extensively in mouse-*C. albicans* interaction studies in our laboratory (16, 38). Strain 6284 (or YJB 6284) was the result of repairing strain BWP17 to prototrophy (5); it has all the characteristics of a typical *C. albicans,* including virulence and the ability to germinate, thus, for the purposes of these studies, it is considered as a wild-type strain. Strain GPH1 is a vacuolar mutant (*vps11*Δ) derived from YJB6284 (35). The 6284 and GPH1

strains were gracious gifts from Dr. G. Palmer (Louisiana State University Health Sciences Center, New Orleans). Other yeasts in this study included *C. albicans* ATCC 60193, *C. glabrata* (ATCC 15126), *C. krusei* (ATCC 6258), *C. kefyr* (ATCC 8553), *Cryptococcus neoformans* (ATCC 32045 and ATCC 66031), and *Saccharomyces cerevisiae* (ATCC 9763 and ATCC 24903). In addition, strains of *Rhodotorula spp.* (y2.929 and y2.196) and *Candida dubliniensis* were graciously provided by Dr. K.C. Hazen (University of Virginia Health Sciences Center, Charlottesville, VR).

### 2.3.2 Culture conditions

Unless otherwise specified, the surveyed yeast cells were stored as water stocks, and cultured on an agar medium containing YNB (yeast nitrogen base) (Becton, Dickinson and Company) supplemented with 2% glucose in Petri dishes at 22-23˚C for 4 days. In some cases, yeast cells were grown at 37˚C for two days or at 22-23˚C for 4 days on glucose (2%)-yeast extract (0.3%)-peptone (1%) agar (GYEP agar). After incubation, isolated colonies were picked to make fresh water suspensions for MALDI-TOF mass spectrometry analysis.

The yeast form of *C. albicans* strain 6284 was induced to germinate (or produce hyphae) by nutritional and temperature shifts. After growth on GYEP agar plates at 22-23ºC for 4 days, the yeast cells were carefully removed, washed twice in sterile deioinized water, and inoculated at 2 x $10^6$ yeast forms/mL in DMEM (Dulbecco's Modified Eagle's Medium HEPES Modification, Sigma) with 2.5% fetal bovine serum followed by incubation at 37 ˚C for 2 h or for 4 h under aeration by rotation at 180 rpm. Under these conditions, over 95% of the yeast cells germinated to become hyphal elements as determined microscopically by the formation of elongated

structures without evidence of constrictions at the point of emergence from the mother cell. The germ tubes were washed in deionized water at least twice before further treatment.

### 2.3.3 Methanol (alcohol) fixation of fungi.

Unless otherwise specified, the washed yeast cells were fixed by suspension in 50% methanol/water (v/v) (EMD Chemicals, San Diego, CA) and either analyzed immediately or stored at 4-6°C for up to 45 days for subsequent comparison analysis. The germ tubes were subjected to the same wash and fixation steps as the yeast cells. In some cases, the fungal cells were fixed in 80% or 100% methanol, and in 50% or 100% isopropanol (Sigma, St Louis, MO) using the same approach as above.

### 2.3.4 Matrix/sample spotting method.

Each alcohol-fixed preparation of yeast cells was suspended by vortexing and 0.5 µl of each was immediately transferred to wells of a 100-well gold-coated MALDI target plate (Applied Biosystems). Each well was immediately overlaid with 0.5 µl matrix solution. The sample-matrix mixtures were allowed to dry at room temperature in ambient air for about 20 min. All sample suspensions were deposited in duplicate on the MALDI target plate. A two-layer method of sample preparation (49, 53{Vorm, 1994 #181), which entailed drying the sample before covering it by the matrix solution, was also tested, but it provided no advantage over the previously described dried-droplet method (27) (data not shown). Three commonly used matrices were compared: α-cyano-4-hydroxycinnamic acid (CHCA, Fluka), sinapinic acid (SA, Fluka) and 2,5-dihydroxybenzoic acid (DHB, Fluka). Each was freshly prepared as a saturated solution in 2:1 acetonitrile (EMD Chemicals) : water containing 0.1% trifluoroacetic acid (TFA, Sigma).

### 2.3.5 MALDI-TOF mass spectra acquisition

Mass spectra were acquired on a Voyager-Elite MALDI rTOF mass spectrometer (Applied Biosystems, Framingham, MA) operated in the positive ion linear mode. The intensity of the nitrogen laser (at 337 nm) was set significantly above the threshold for desorption/ionization (between 2000 and 2400). The accelerating voltage was 25 kV with an extraction delay time of 250 ns. For each sample spot (approximately 5 mm in diameter on the MALDI target plate), an average of 75 laser pulses were delivered at one point, and the final spectrum was the accumulation/average results obtained from at least 5 different points on a given same sample spot. Due to performance limitations of the linear mode of the MALDI-TOF mass spectrometer, mass assignments of ions with a mass to charge ratio (m/z) larger than 5,000 Daltons (Da) may bear an error of 10-20 Da. Cytochrome C (12,362 Da), myoglobin (16,700 Da) and ubiquitin (8,565 Da) (Sigma, St. Louis, MO) were used as standards for external calibration.

### 2.3.6 Principal component analysis (PCA)

MALDI mass spectra of intact yeast cells were calibrated and processed using "Data Explorer" (Applied Biosystems). Some of the mass spectra were evaluated by PCA (17, 18). After the spectrum baseline was corrected to zero, up to 25 of the most abundant ions in the range of m/z 4000 – 15,000 were first selected. Among the possible 25 peaks, those with a relative intensity less than 5% and those with a signal/noise ratio less than 3 were discarded. The most intense peaks from each mass spectrum were compiled into a peak list with no redundancy. Each mass spectrum was compared with the compiled peak list to generate a data matrix in which the matched peaks (relative intensity higher than 5%, S/N larger than 3 and no more than 25 peaks) were denoted as 1 (100%) and the rest of the unmatched peaks were denoted as zero. PCA was

performed on the data matrix by using the MatLab program (The Math Works, Inc. version 7.0). Plotting of each spectrum was limited to the first and second principal component scores, PC1 and PC2, respectively.

## 2.4 Results

### 2.4.1 Concentration of yeast cell suspension

*C. albicans* 6284 was used as the prototypic strain to evaluate the influence of yeast cell concentration on the quality of mass spectra. In preliminary experiments, water-washed yeast cells mixed directly with the matrix sinapinic acid (SA) matrix provided evidence that approximately $0.5x10^6$ yeast cells/0.5µl was optimal for MALDI mass spectral analysis. Fewer satisfactory mass spectral peaks were observed at both higher and lower cell concentrations (data not shown).

### 2.4.2 Effect of alcohol fixation

Fixing the yeast cells in 50% methanol was initially employed as a killing step, but it surprisingly led to a significant increase in the number of ions in the MALDI mass spectrum. For instance, non-methanol-treated *C. albicans* A9 yielded only two significant peaks between m/z 3000 and 15000 (Figure 2.1a) as compared to a large number of peaks (more than ten) following methanol fixation (Figure 2.1b). Fixation treatments of 80% or 100% methanol or isopropanol (50% and 100%) offered no advantage over the 50% methanol fixation. Fixation had the additional advantage of minimizing yeast cell aggregation and, therefore, more reproducible cell number transfer to the target plate, as compared to the tendency for increased aggregation of

some strains suspended in water alone. For the remainder of this study, 50% methanol fixation was used unless otherwise specified.

### 2.4.3 MALDI matrix selection

MALDI analyses of *C. glabrata* ATCC 15126 and *C. albicans* A9 produced different spectral qualities depending on the choice of matrix (Figure 2.2). When DHB was employed as the matrix, both *C. glabrata* ATCC 15126 (Figure 2.2a) and *C. albicans* A9 (Figure 2.2d) yielded poor quality mass spectra with only a few significant signals in the high mass range (m/z 5000 – 15,000). Matrix CHCA (Figure 2.2b and 2e) worked somewhat better than DHB, but not as well as SA (Figure 2.2c, and 2f). SA was chosen as the most suitable matrix because it produced the most peaks with acceptable signal/noise ratios, and it promoted signal generation in the high-mass range.

### 2.4.4 Effect of incubation conditions on *C. albicans* mass spectra

Examination of four *C. albicans* strains grown on GYEP medium under the two conditions of 37 °C for 2 days and 22-23 °C for 4 days showed that the influence of incubation conditions on MALDI mass spectra of intact yeast cells was strain dependent. For example, the mass spectra of *C. albicans* strain A9 under the two growth conditions (Figure 2.3a and 3b) did not change dramatically as noted by the presence of high abundant ions on both spectra, such as m/z at 5217, 5910, 6207, 6904, 8552, and 9762. As for *C. albicans* 3153A, however, significant changes could be observed when comparing the MALDI spectra acquired under the two conditions (Figure 2.3c, at 22-23 °C for 4 days, and Figure 2.3d, at 37°C for 2 days). Some ions in the high mass range (m/z at 11341, 12279 and 13161) were evident from cells grown at 37ºC,

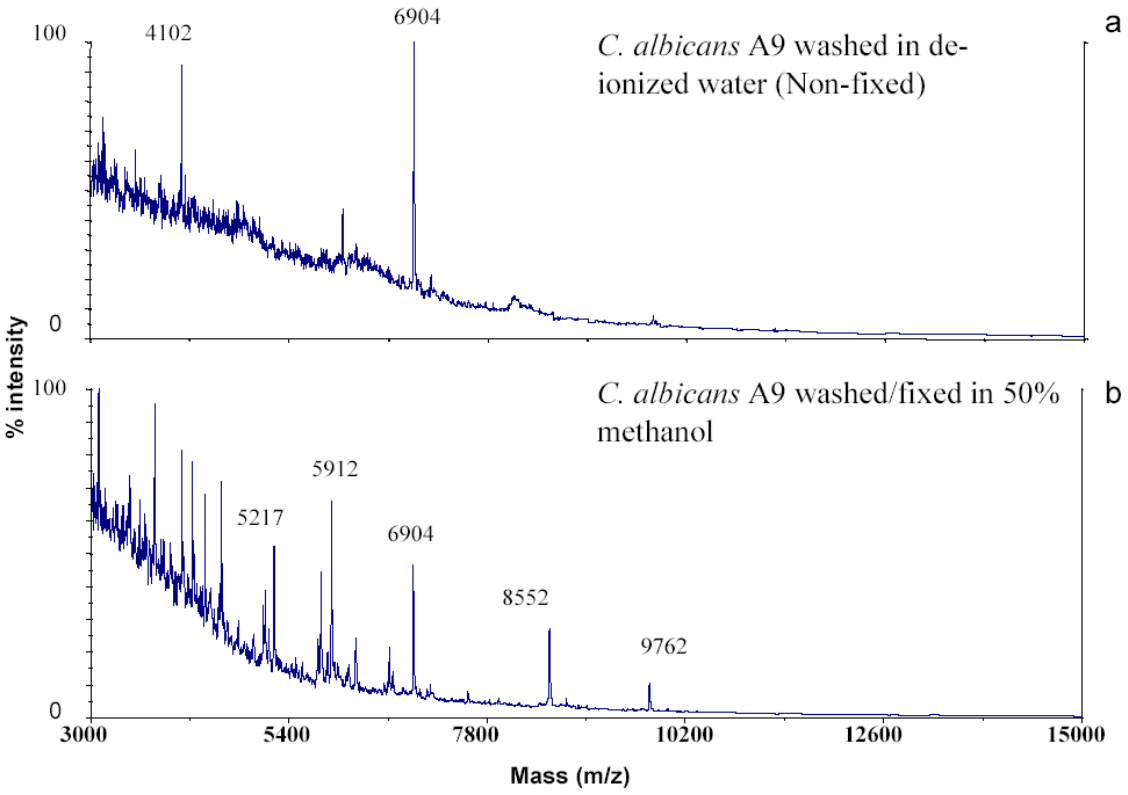



**Figure 2.1.** Alcohol fixation improves the quality of the MALDI mass spectrum. *Candida albicans* A9 yeast cells were prepared by either washing in deionized water (a), or washing in deionizied water followed by fixation in 50% methanol:water (b). Sinapinic acid was employed as the MALDI matrix and the concentration of the yeast cell suspensions was $10^6$ cells/μL.

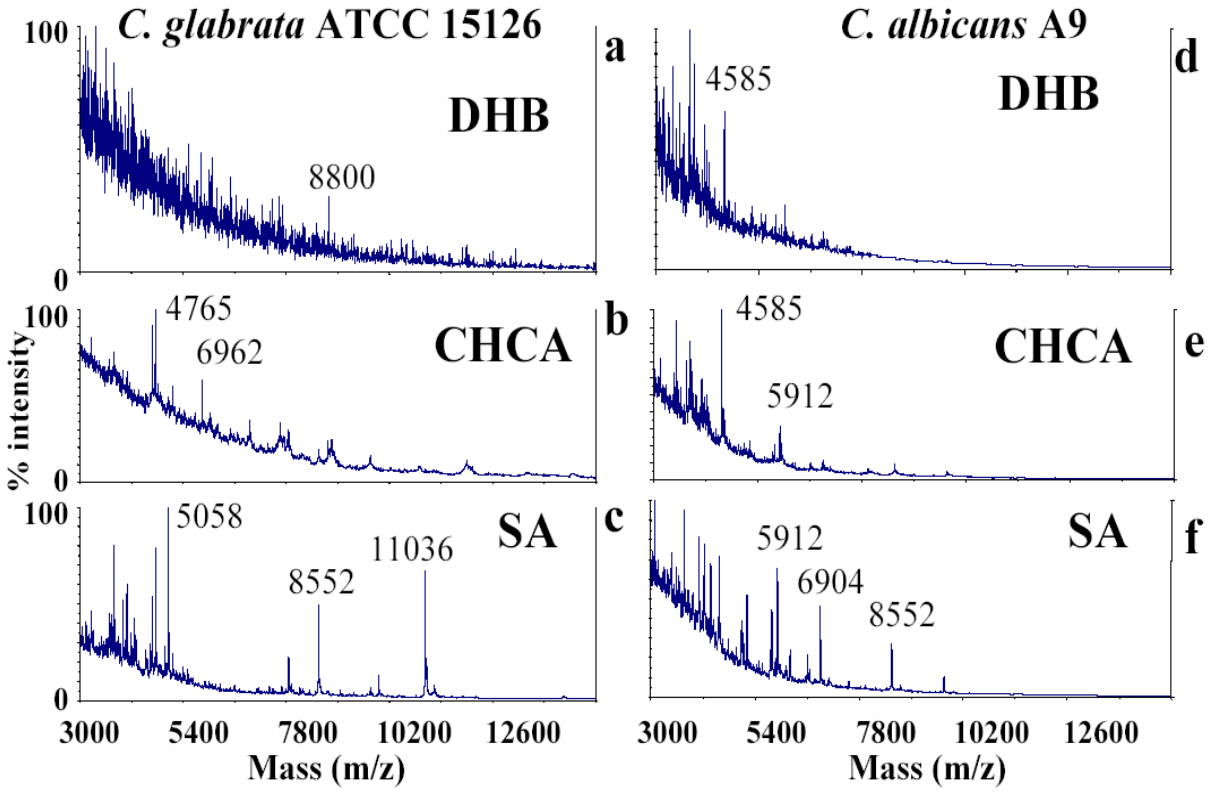



**Figure 2.2**. Three commonly used matrix compounds: 2,5-dihydroxybenzoic acid (DHB), α-cyano-4-hydroxycinnamic acid (CHCA) and sinapinic acid (SA) were compared. MALDI spectra of *C. glabrata* ATCC 15126 were acquired in the presence of DHB (a), CHCA (b) and SA (c). In parallel and under the same conditions, MALDI mass spectra of *C. albicans* A9 were acquired in the presence of DHB (d), CHCA (e), and SA (f). When SA was employed as the matrix, MALDI-MS produced more useful mass signatures for both *C. glabrata* ATCC 15126 (c) and *C. albicans* A9 (f). SA was selected as the optimal matrix in the mass signature acquisitions of all other yeast species and strains in this study.

but not 22-23ºC. *C. albicans* strains GPH1 and 6284 also showed similar patterns of mass spectral changes when the spectra (not shown) of the same strain grown at the two temperatures and durations were compared.

### 2.4.5 Characteristic mass spectral ions generated from *C. albicans*.

The four *C. albicans* strains grown at 22-23˚C on GYEP agar for 4 days provided highly reproducible mass spectra, not only for the same strain tested at different locations on the target plate, but also for replicates the same strain grown under the same conditions on different days. Mass spectral comparisons showed a high degree of similarity between strains 3153A (Figure 2.3a) and A9 (Figure 2.3b), and both closely matched strain 6284 (Figure 2.3e). Mutant strain GPH1 however, showed significant differences from the other strains (Figure 2.3f), ions in the range of m/z 7000 – 8000 and above 11,000 were

apparently not shared by the other three strains. Eight common ions at m/z 5110, 5217, 5785, 5910, 6207, 6904, 8552 and 9762 appeared on the spectra of all four strains, which provided evidence that these signals might serve as signature markers for *C. albicans* under the employed conditions.

### 2.4.6 PCA analysis of *C. albicans* mass spectra.

PCA analysis was performed on the spectra of the four *C. albicans* strains grown under identical conditions except that the length and the temperature of growth were varied for comparison purposes. *C. albicans* strain 3153A incubated under two different conditions (at 22-23 °C for 4 days and at 37 °C for 2 days) showed variation in the mass signatures (Figure 2.3c and 3d), which became more obvious by PCA analysis (Figure 2.4, the solid and hollow rectangle

symbols, respectively). The mass signatures of *C. albicans* strain 6284 grown under the two growth conditions also had noticeable differences (Figure 2.4, the solid and hollow triangle symbols). The mass signature of *C. albicans* GPH1 (Figure 2.3f) was quite different from those of *C. albicans* strains A9 (Figure 2.3a), 3153A (Figure 2.3b), and 6284 (Figure 2.3e). PCA analysis of these mass signatures indicated that GPH1 (the hollow circle in Figure 2.4) can be distinguished from the other three strains (the hollow rectangle, triangle and plus symbol in Figure 2.4) based on the first principle component (the x-axis in Figure 2.4). Similar mass signatures were grouped together on the PCA plot (Figure 2.4). GPH1 incubated under two conditions formed one group, whereas the other three strains (3153A, A9 and 6284) grown 22-23 °C for 4 days formed another cluster. The latter cluster (around coordinates -1, 0 in Figure 2.4) also included the spots (small # and + symbols that represented the mass signatures of strains 3153A and A9 incubated at 22-23 °C for 4 days followed by fixation in 50% methanol and storage at 4 °C for 45 days before mass spectrometry analysis). Storage at 4 °C for 45 days after fixation in 50% methanol appeared to have little effect on the mass signatures of *C. albicans* 3153A and A9. These data demonstrate that the mass signature of the same strain is reproducible and stable even upon relatively long-term storage.

### 2.4.7 Differentiation of yeast species.

A mass spectral survey on strains of various *Candida* species cultured on YNB-glucose agar medium at 22-23 ˚C for 4 days indicated that each species could be distinguished (Figure 2.5). The three non-*Candida* yeasts*, C. neoformans, Rhodotorula sp.* and *S. cerevisiae* also produced distinct mass spectral signatures that were distinguishable from the *Candida* species and from each other (Figure 2.6). *C. neoformans* produced ions evident at m/z 5736, 6722, 8539, 8539,

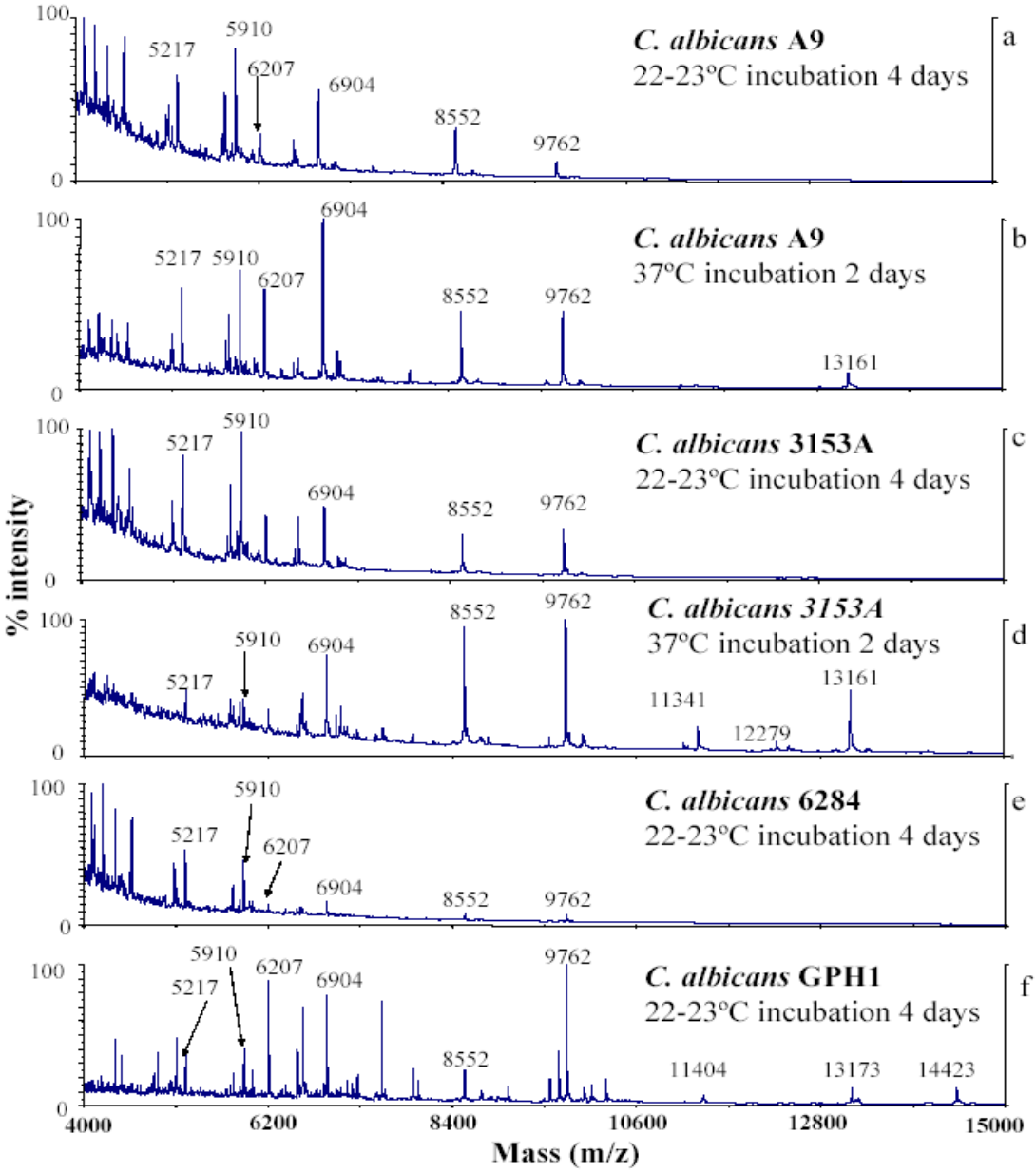



**Figure 2.3**. MALDI mass spectra were acquired for the four *C. albicans* strains grown on GYEP agar medium. Effect of growth conditions on the MALDI mass spectra is strain dependent. Similar MALDI mass spectra were acquired for *C. albicans* A9 grown at 22-23ºC for 4 days (a) and 37ºC for 2 days (b). Differences, especially in the high mass range, are quite obvious when comparing the mass spectra of *C. albicans* 3153A grown at 22-23ºC for 4 days (c) and grown at 37ºC for 2 days (d). Under constant growth conditions at 22-23ºC for 4 days *C. albicans* strain GPH1 (f) could be distinguished by its MALDI mass signature from (a) *C. albicans* A9, (c) *C. albicans* 3153A, and (e) *C. albicans* 6284.

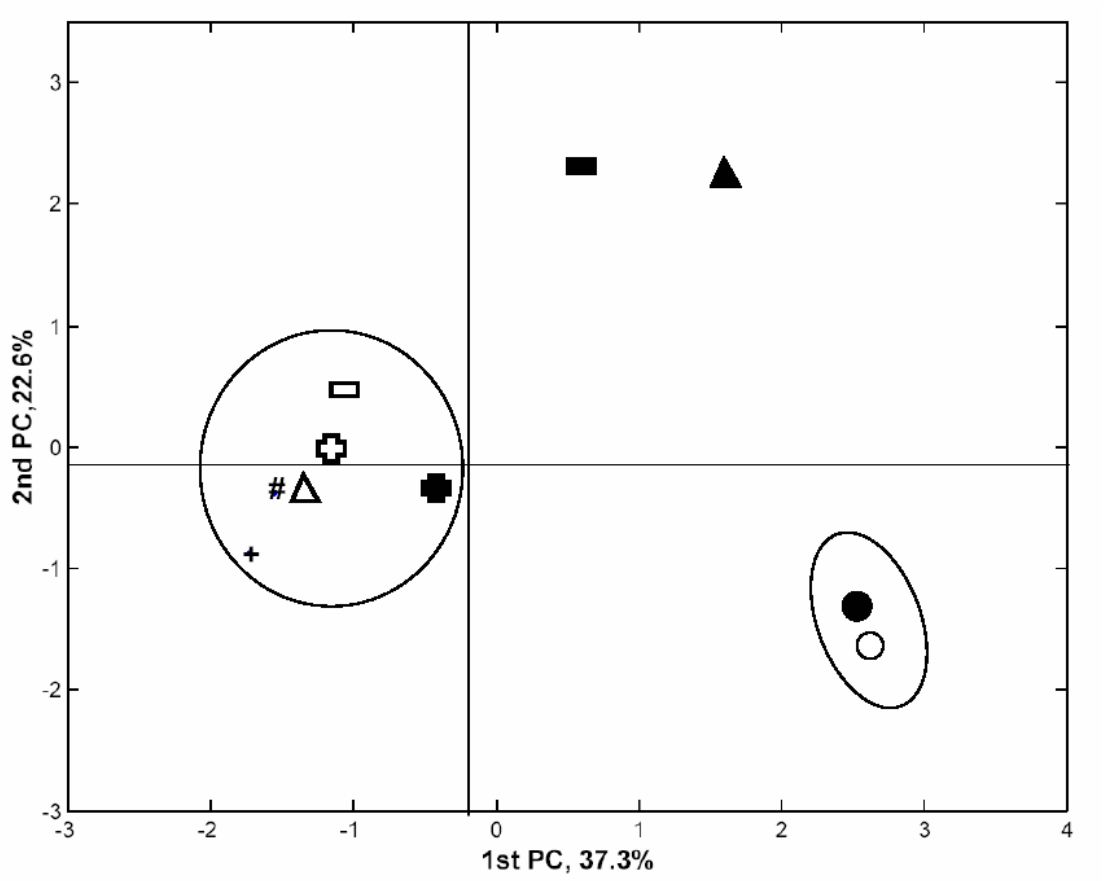



**Figure 2.4**. Principal component analysis (PCA) of the mass signatures of *C. albicans* strains A9, 3153A, 6284, and GPH1. Each spot on the plot represents a mass spectrum. For each mass spectrum used in the PCA analysis, no more than 25 peaks with relative intensity higher than 5%, S/N larger than 3 were denoted as 1 (100%) and the rest of the peaks were denoted as zero. PCA analysis reduced the multi-dimensional features (i.e., mass peaks denoted as 1) of each signature spectrum down to two principal components (PC1 and PC2). Hollow symbols: cells cultured at 22-23ºC for 4 days; solid symbols: cells cultured at 37ºC for 2 days. All strains were cultured on GYEP agar medium. The small cluster on the right includes the mass signatures of *C. albicans* GPH1 grown under two different conditions: ○ represents the mass spectrum of Figure 3f. The large cluster on the left includes the mass signatures: ✚ *C. albicans* A9 (spectrum see Figure 3a); ▭ *C. albicans* 3153A (spectrum see Figure 3c): ⬢ *C. albicans* A9 (spectrum see Figure 3b); △ *C. albicans* 6284 (spectrum see Figure 3e); # (*C. albicans* 3153A) and + (*C. albicans* A9): cell cultured at 22-23ºC for 4 days, treated by 50% methanol and stored at 4ºC for 45 days before MALDI analysis.

8640 and 9405, all of which are shared by the two strains of this species (Figure 2.6a and 6b). The two strains of *S. cerevisiae* shared ions at m/z: 4053, 4681, 4752 and 5500 (Figure 2.6c and 6d). Compared to other species in the survey, the two *Rhodotorula sp.* strains produced fewer peaks in their spectra (Figure 2.6e and 6f), with two shared ions at m/z 4501 and 8536. For all the MALDI-MS profiled yeast strains, no two strains within a given species produced identical spectra, and the mass spectral differences between different species were even more evident.

PCA analysis of the mass signatures of *C. albicans* ATCC60193, *C. kefyr* (ATCC 8553), and *C. neoformans* (ATCC 32045 and ATCC 66031) showed that these species could be clearly distinguished on a PCA plot (Figure 2.7). Different strains within the same species, such as *C. neoformans* ATCC 32045 (Figure 2.6a) and ATCC 66031(Figure 2.6b), could also be distinguished by their mass signatures, however, usually there were fewer differences in mass signatures between strains within a species as compared to differences between species. Examples of mass signature reproducibility are offered by the close proximity of the two *C. glabrata* ATCC 15126 spots and the three spots of *C. albicans* ATCC 60193 on the PCA plot (Figure 2.7). The two mass signatures of *C. glabrata* ATCC 15126 were from replicate analyses of the same sample, while the three mass signatures of *C. albicans* ATCC 60193 were from three independent colonies. These data along with stability and reproducibility of spectra shown in the previous section provide evidence that MALDI-TOF mass spectral analyses are predictable under prescribed conditions of growth and sample preparation.

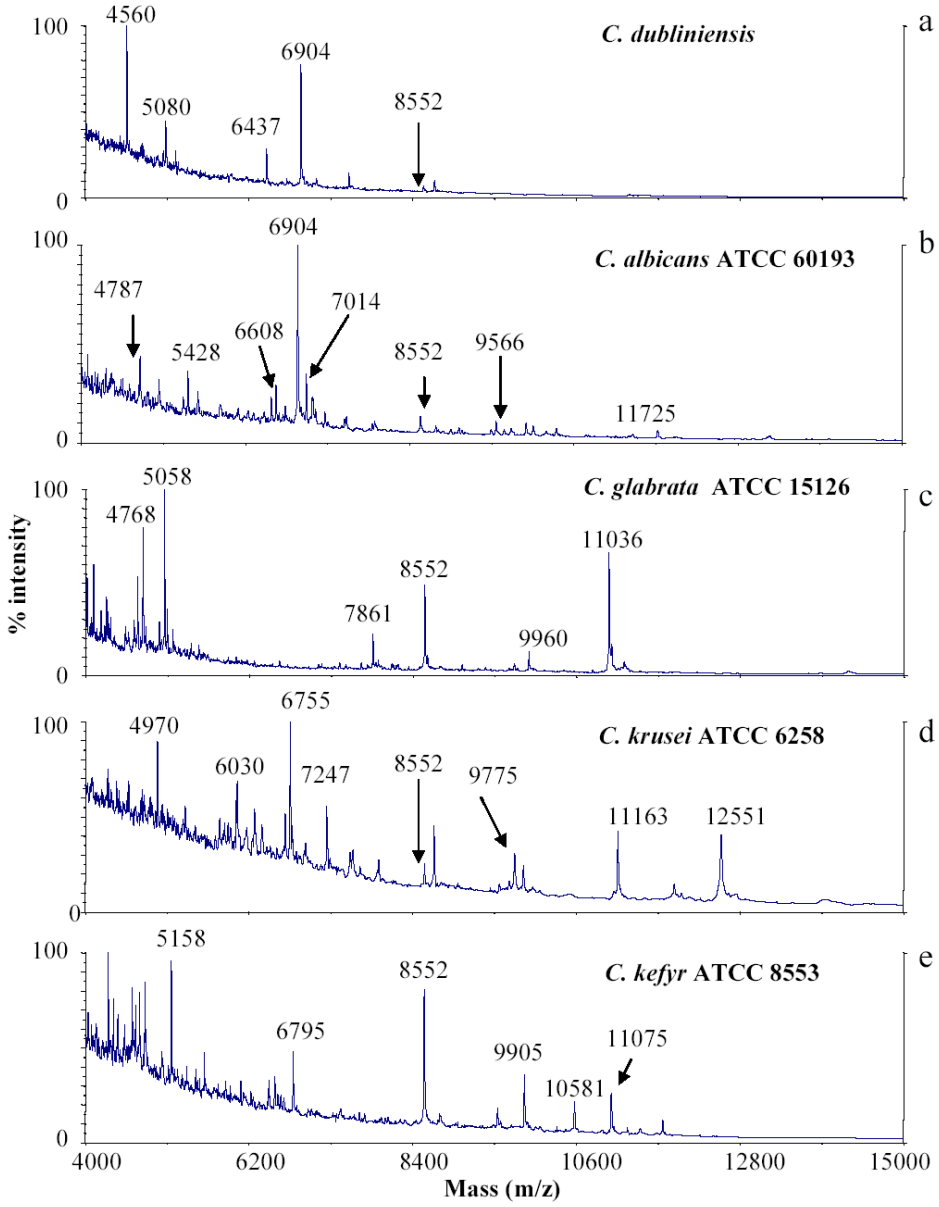



**Figure 2.5**. The MALDI mass signatures distinguish different species of *Candida.* When cultured under the same conditions (YNB with 2% glucose media at 22-23ºC for 4 days), *C. dubliniensis* (a), *C. albicans* ATCC 60193 (b), *C. glabrata* ATCC 15126 (c), *C. krusei* ATCC 6258 (d) and *C. kefyr* ATCC 8553*(*e*)* could be distinguished by their mass signatures.

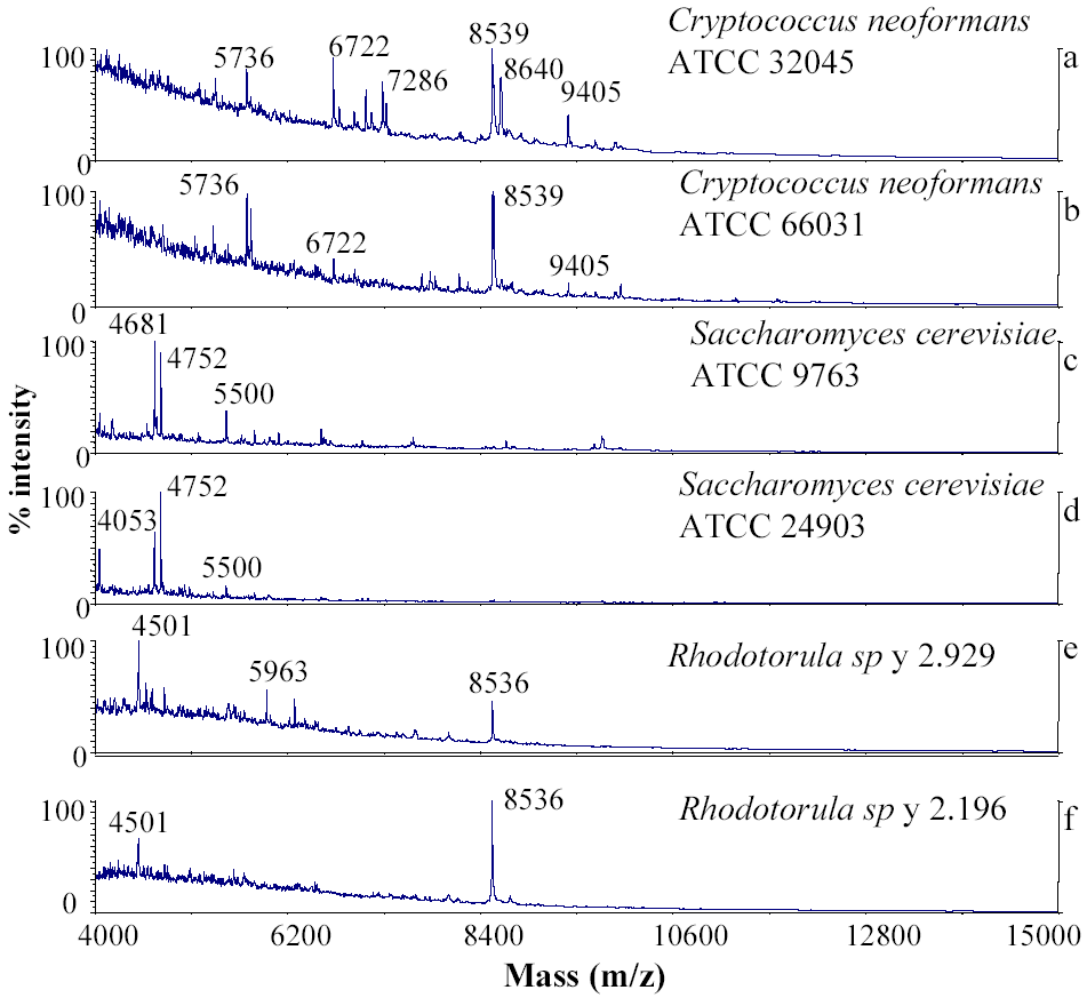



**Figure 2.6**. MALDI mass spectra can be used to differentiate between different fungal genera and strains within each genus. When grown under conditions similar to those used for data obtained in Figure 5, unique mass spectra were obtained from *Cryptococcus neoformans* strains ATCC 32045 (a) and ATCC 66031 (b), *Saccharomyces cerevisice* strains ATCC 9763 (c) and ATCC 24903 (d), *Rhodotorula sp.* strains y 2.929 (e) and y 2.196 (f).

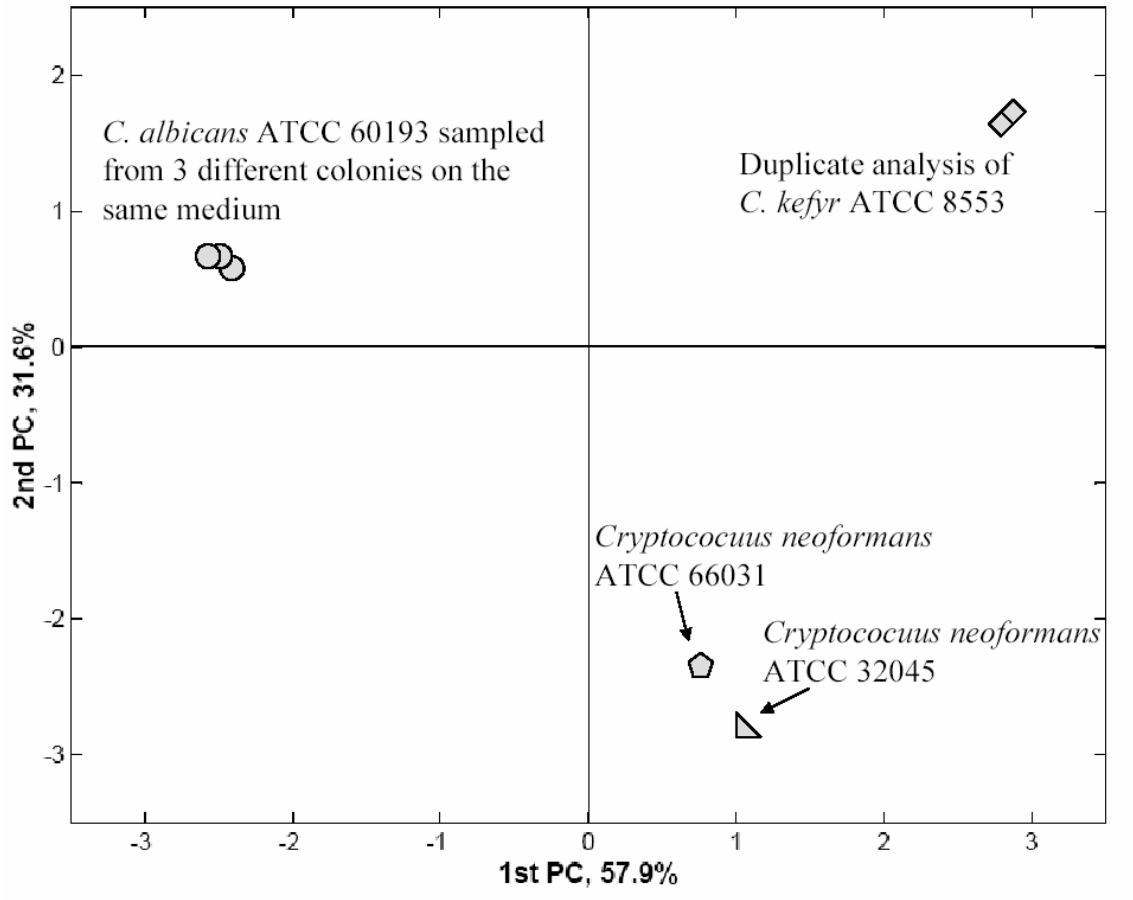



**Figure 2.7**. The reproducibility of mass signatures was confirmed by PCA analysis. Acquired under identical conditions, the mass spectra (three circles) of *C. albicans* ATCC 60193 sampled from 3 different colonies on the same Petri dish were similar; nearly identical mass spectra (two diamonds) were also acquired for the same *C. kefyr* ATCC 8553 sample spotted on different locations of a MALDI target plate. It should also be noted that the mass spectral differences between yeast species (*C. albicans* ATCC 60193, *C. kefyr* ATCC 8553 and *Cryptococcus neoformans*) were greater than the mass spectral differences between strains (ATCC 32045 and ATCC 66031) of the same species (*Cryptococcus neoformans*).

### 2.4.8 Comparison of yeast and hyphal (germ tubes) forms of *C. albicans*

The MALDI mass spectrum of *C. albicans* strain 6284 yeast forms (Figure 2.8a) was compared to that of the same strain germinated for 2 h (Figure 2.8b) and for 4 h (Figure 2.8c). The surprising finding was that the spectrum of the 4 h germination sample was rather similar to that of the yeast form, whereas the spectrum obtained from germination at 2 h differed significantly. More ions, especially in the high mass region, were generated by the 2 h germination sample during the MALDI process. A possible explanation is that at 4 h almost all germ tubes had aggregated and were difficult to suspend in 50% methanol, which caused variability of sample delivery to the target plate. To test this possibility, the amount of sample applied to the target plate was varied, but the high mass ions were consistently absent in the mass spectra (data not shown). The same experiment was repeated on another germination sample preparation of *C. albicans* strain 6284, and the results were consistent. That is, several high mass ions (e.g., above m/z 7500) appeared on the mass spectra of 2 h germinated sample, but the higher mass ions were absent on the mass spectra of the 4 h old germ tubes.

## 2.5 Discussion

The work presented here shows the potential power of the use of MALDI-TOF mass spectrometry as a rapid means for yeast identification, grouping strains within a species and for following surface changes during a yeast-to-hyphal transition of *C. albicans.* Many parameters (e.g. the concentration of yeast suspension, type of MALDI matrix employed) were considered during optimization of the experiments as well as data analysis and presentation. The approach developed in this study should be applicable to all fungi, including molds.

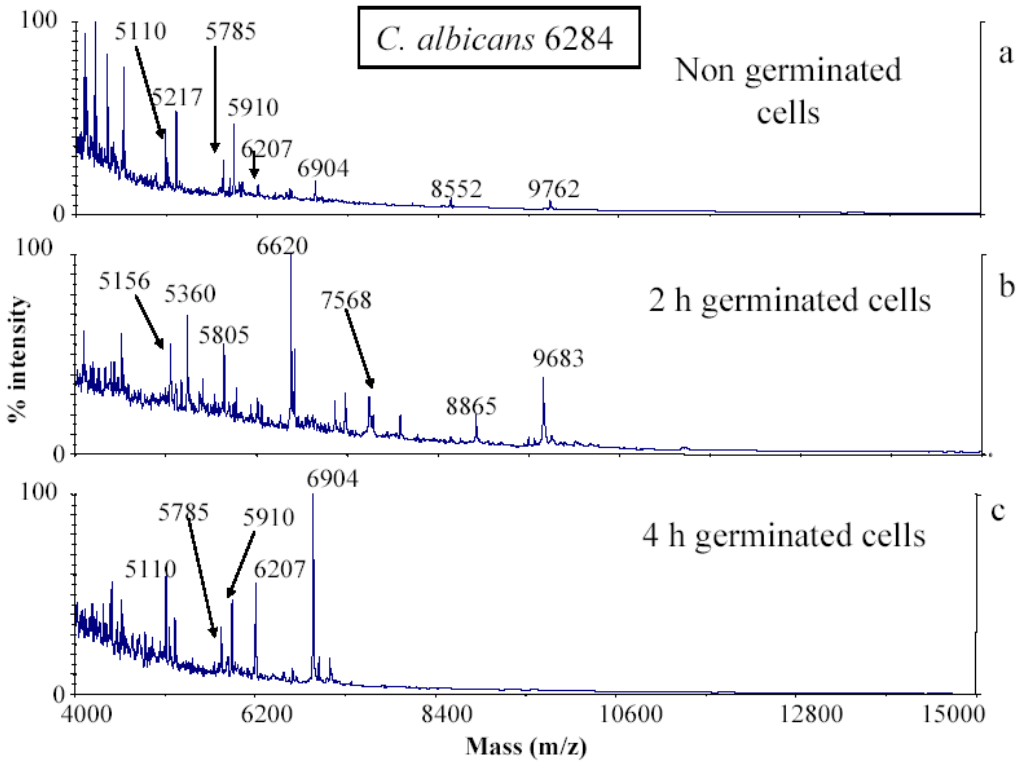



**Figure 2.8**. *C. albicans* strain 6284 was analyzed as yeast forms (a), as 2 h old germ tubes (b), and as 4 h old germ tubes (c). The mass spectra of yeast form (a) and 4 h old germ tubes (c) were similar, but the mass spectrum of 2 h old germ tubes (b) differed from (a) and (c).

In MALDI analyses, sample preparation and selection of an appropriate matrix compound are essential considerations in protocol development. The tight cell wall construction of fungi apparently limits intimate contact of the surface components with the MALDI matrix crystals, which results in low count ion signals and poor quality of mass signatures. Enhancement of contact of surface fungal molecules with the matrix requires an optimal ratio between the matrix and fungal molecules (analyte); unbalanced matrix/analyte ratios lead to reduced or even total supression of analyte signals. Since the volume and concentration of the yeast cell suspension deposited onto each sample spot of the MALDI target plate was usually constant, the ratio of matrix:number of intact yeast cell was optimized by varying the concentration of the cell suspension solution. Analysis of approximately $0.5x10^6$ yeast cells provided the best mass spectral result for *C. albicans* 6284 and other fungal cells under the conditions employed in this study.

**2.5.1 Methanol fixation of fungal cells** – When analyzed by MALDI mass spectrometry, intact fungal cells that were washed in water prior to mixing with a matrix solution and spotting did not produce sufficient diagnostic ions in the range of m/z 4000 to 15,000 (Figure 2.1a), which is in agreement with a previous report (1). Several sample treatment methods, such as sonication (47) and boiling (24) (51), increased MALDI-MS signals of bacteria, however, these methods failed to give satisfactory results in our investigation on yeast cells (unpublished data). Since one group of investigators obtained improved ionization of medium sized molecules following ethanol fixation of whole bacterial cells (51), and as a measure of safety in the handling of potentially infectious organisms, we examined the effects of alcohol fixation on yeast cells. In addition to improving the mass spectral signals, inclusion of alcohol, such as methanol or isopropanol, in the

fixation buffer also facilitated suspension of the yeast cells and reproducibility in sample spotting onto the MALDI target plate. The novel mass spectral signals obtained following alcohol fixation were also reproducible as discussed below, leading to mass signature assignments. Compared to the cell wall lysis procedures used by others (43) (1) which may yield a larger number of mass spectral peaks, the alcohol fixation approach described in our work is more rapid, simple and it also inactivates potentially hazardous organisms.

**2.5.2 Yeast genera and species differentiation by mass signatures** – Although the number of strains was small for each species studies, the results indicate that MALDI mass signatures can be employed to differentiate yeast cells at the genus level, as is clearly demonstrated in Figure 2.6. The mass spectral differences between *C. neoformans* (Figure 2.6a and 6b), *S. cerevisiae* (Figure 2.6c and 6d), and *Rhodotorula sp* (Figure 2.6e and 6f) are evident by simple visual inspection. Differences in mass signatures can also be determined by visual examination for different species within the genus *Candida* (e.g. *Candida spp* in Figure 2.5). The reproducibility of mass spectra collected under a prescribed set of conditions allows comparisons to be made on spectra acquired at different times. For example, the mass signatures of *Candida spp* in Figure 2.5 can be compared with mass signatures shown in Figure 2.6, whereas the mass spectra shown in Figure 2.3 should not be compared to those displayed in Figure 2.5 and 6 because the yeast cells used for data shown in Figure 2.3 were cultured under different conditions.

Mass signatures may be useful for the differentiation of strains within a given species, but this is not always the case. For example, in Figure 2.6 different strains of the same species showed differences in the mass signatures, and in Figure 2.3 the mass spectrum of *C. albicans* GPH1 (Figure 2.3f) was quite different from the mass spectra of *C. albicans* A9 (Figure 2.3a), 3153A

(Figure 2.3c) and 6284 (Figure 2.3e). The departure of the mutant GPH1 from other wild-type strains of *C. albicans* is of interest, because the vacuolar deficiency of GPH1 may well result in cell wall surface changes which may be reflected in the mass spectral signature for this strain. Further work is needed to determine the usefulness of mass spectral analysis of yeast mutants. The mass spectra of *C. albicans* strains A9, 3153A and 6284 were not sufficiently distinct for differentiation. The signatures of these strains shared a series of ions, such as m/z 5217, 5910, 6904, 8552 and 9762 (Figure 2.3); these ions can be used to group the three strains. A greater number of strains need to be surveyed in order to gain a more thorough appreciation of the power of the mass signatures approach in grouping strains, and possibly mutants, within a given species.

**2.5.3 Reproducibility of mass signatures under well defined conditions -** Rapid yeast identification by MALDI-MS is an attractive approach because of its simplicity and potential for high throughput analyses. Although the production of diagnostic signals in MALDI mass spectra is the key to the approach, reproducibility of the signals is critical. In our work, the spectra were highly reproducible provided that the organisms were grown and fixed under a common set of conditions. In general, alcohol fixation worked well, but 50% methanol fixation consistently yielded a sufficient number of diagnostic signals that could be used to constitute a mass signature for a given species. Other important parameters that could affect mass signature reproducibility included the number of yeast cells placed in the MALDI sample spot, choice of matrix compound, sample spotting technique, and cell incubation conditions. For example, yeast cells, such as *C. albicans* strain A9, grown at 37 ºC for 2 days (Figure 2.3a) yielded essentially the same mass signature as when grown at 22-23 ºC for 4 days (Figure 2.3b); but for other yeast

cells, such for *C. albicans* strain 3153A, the mass signatures (Figure 2.3c and 3d) varied somewhat depending on the temperature and length of growth. In the latter example, even though there were obvious changes in some of the characteristic signals, several diagnostic peaks were recorded regardless of growth temperature (e.g. m/z 6904, 8552, 9762). For identification and differentiation purposes, the MALDI mass signatures of yeast cells should be acquired from fungal cells grown and prepared under the same conditions so as to minimize variations in the mass signature of a given strain.

Under the same conditions of growth and sample preparation, our findings concerning the reproducibility of mass signatures is consistent with other results reported on MALDI-TOF analysis of bacterial cells (2, 40). In addition, we found that fungal cells could be stored prior to mass analysis without significantly affecting the results. For example, the mass signature acquired for *C. albicans* strain A9 incubated at 22-23 ˚C in GYEP for 4 days and immediately treated with 50% methanol followed promptly by MALDI-TOF mass spectrometry analysis provided essentially identical signals as when the same strain was grown under identical conditions except that it was stored for 45 days at 4 ˚C before methanol fixation and mass analysis. The two mass signatures (spots: + and ✠ in Figure 2.4) can be clustered together as revealed by PCA analysis. Mass signatures were also consistent for the same yeast strain deposited on different sample spots on the same MALDI target plate, as well as on different sample spots on different MALDI target plates (See Figure 2.7, duplicate analysis of *C. kefyr* ATCC 8553), and for separate colonies grown in different areas of the same Petri dish (See Figure 2.7, the three *C. albicans* ATCC 60193 spots).

**2.5.4 PCA analysis of mass signatures** – As indicated above, diagnostic signals are consistent under a prescribed set of conditions, but the relative peak intensities of a given ion signal may vary from spectrum to spectrum. To minimize this variability for PCA analysis we degenerated the relative intensities of diagnostic peaks to zero and 100%, thus leaving the presence of diagnostic peaks as the only characteristic feature. This simplified approach was effective in clustering the mass signatures of the same yeast strain, even when the data were acquired from yeast prepared under slightly different conditions. As shown in Figure 2.4, the spot that represents the mass signature of *C. albicans* 3153A grown at 22 ºC for 4 days ( ▭ ) can be clustered with the same strain grown under the same conditions but stored for 45 days (#). Importantly, this PCA analysis approach is an effective tool in clustering intact yeast cell mass signatures even without prior information, such as the genotype and phenotype of a given yeast of interest. Thus, PCA employed as a mass signature clustering tool has the potential of playing an important role in both small and large scale yeast identification.

**2.5.5 Origin of the diagnostic signals** - Analyte molecules that are in physical contact with the matrix molecules are those that can be desorbed/ionized leading to their detection and display in mass spectra. Since the matrix molecules are not capable of penetrating the cell wall, the medium sized molecules (4000-15000 Da) detected in our work most likely originate from the yeast cell surface. One possible explanation for the increased signals following alcohol fixation is that cell surface macromolecules unfold, thus exposing medium size molecules (4000 to 15000 Da) to the matrix. A second possibility is that the fixation may have promoted release of cell membrane or cytosolic materials to the cell surface and contact with the matrix. The yeast cell wall biopolymers are comprised of proteins and polysaccharides (7). Since MALDI-MS ionization

conditions in the linear mode preserve even weak non-covalent hydrogen bonds (11) (6, 44), it is unlikely that covalent bonds of the large biopolymers would be cleaved to yield fragments smaller than 15 kDa. Given the flux and energy of the UV laser employed in our experiments, we cannot, however, totally rule out the possibility that certain weaker covalently bound cell surface biopolymers might be cleaved and yield ions in the range of m/z 4000 – 15,000.

Sinapinic acid is a well described MALDI matrix compound that is especially useful for the analyses of peptides and small proteins (3, 4). In our study, this matrix worked the best among all those tested for analyses of the intact yeast cells (Figure 2.2), which is consistent with another report (19), and this suggests that the detected ionized yeast cell surface molecules may well be proteins/peptides, or molecules with amino acid residues. In addition, in this study we employed the "dried droplet" method in MALDI sample spotting, which favors ionization of proteins and peptides (14). The polysaccharides that are abundant on the yeast cell surface would require a different matrix compound and sample spotting method for efficient ionization (6), (14).

**2.5.6 Monitoring the mass signatures of *C. albicans* at different morphological states** - *C. albicans* typically undergoes morphogenesis upon tissue invasion, resulting in the formation of several fungal forms, including yeast (blastospores or blastoconidia) , hyphae (germ tubes or mycelia) and pseudohyphae. We found that each morphogenetic form of *C. albicans* (wild type strain 6284) is characterized by a unique mass signature (Figure 2.8). It should be noted that the 2 h germinated *C. albicans* (pseudohyphae) yielded the most ions in the range of m/z 4000 – 15,000, which implies that changes in the pseudohyphal cell wall may promote the release of more mid-size molecules (Figure 8b). The results for the 4 h germ tubes were surprising in that

they differed from the 2 h germ tubes, and, instead, appeared more like yeast forms. The noted difficulty in spotting due to aggregation of 4 h germ tubes may have resulted in clumps of hyphae directed away from contact with the matrix, exposing mostly the mother yeast cells. Alternatively, by 4 h the hyphal cells may be in transition back to the yeast form. These explanations are currently under investigation. Regardless of the explanation, unique signals were obtainable during the transition of yeast to hyphal forms. Identification of these unique signals is the subject of current studies.

## 2.6 Concluding remarks

In this study, we demonstrated that the MALDI-TOF mass spectra of intact yeast cells can provide mass signatures for identification of fungal cells at the genus and species levels. If appropriately fixed before analysis, intact yeast cells will produce mass signatures just as readily as those produced by bacteria without prior any treatment. Mass signatures could also be employed to differentiate strains within a genus, or at least group together yeast strains in case where difference between individuals was not remarkable. The identification process is rapid, and requires a tiny amount of sample ($10^6$ -$10^7$ intact cells). Fixation of the pathogenic yeast with 50% methanol in water greatly improves the quality of the MALDI mass spectra in terms of increasing the number of characteristic ions with better signal-to-noise ratios. Aided by mathematical tools such as principal component analysis, the creation of MALDI mass signature libraries will serve as an effective tool for fast identification of intact yeast cells in clinical and research laboratories.

## References:


1. **Amiri-Eliasi, B. J., and C. Fenselau.** 2001. Characterization of protein biomarkers desorbed by MALDI from whole fungal cells. Analytical Chemistry **73:**5228-5231.
2. **Arnold, R. J., and J. P. Reilly.** 1998. Fingerprint matching of E-coli strains with matrix-assisted laser desorption ionization time-of-flight mass spectrometry of whole cells using a modified correlation approach. Rapid Communications in Mass Spectrometry **12:**630-636.
3. **Beavis, R. C., and B. T. Chait.** 1989. Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. Rapid Communications in Mass Spectrometry **3:**432-435.
4. **Beavis, R. C., and B. T. Chait.** 1990. Rapid, sensitive analysis of protein mixtures by mass-spectrometry. Proceedings of The National Academy of Sciences of the United States of America **87:**6873-6877.
5. **Benson, E. S., S. C. Filler, and J. Berman.** 2002. A forkhead transcription factor is important for true hyphal as well as yeast morphogenesis in *Candida albicans*. Eukaryotic Cell **1:**787-798.
6. **Cai, Y., Y. J. Jiang, and R. B. Cole.** 2003. Anionic adducts of oligosaccharides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Analytical Chemistry **75:**1638-1644.
7. **Chaffin, W. L., J. L. Lopez-Ribot, M. Casanova, D. Gozalbo, and J. P. Martinez.** 1998. Cell wall and secreted proteins of *Candida albicans*: Identification, function, and expression. Microbiology and Molecular Biology Reviews **62:**130-180.
8. **Chen, H. Y., and Y. C. Chen.** 2005. Characterization of intact *Penicillium* spores by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Communications in Mass Spectrometry **19:**3564-3568.
9. **Enache-Angoulvant, A., and C. Hennequin.** 2005. Invasive *Saccharomyces* infection: A comprehensive review. Clinical Infectious Diseases **41:**1559-1568.
10. **Evason, D. J., M. A. Claydon, and D. B. Gordon.** 2001. Exploring the limits of bacterial identification by intact cell-mass spectrometry. Journal of The American Society For Mass Spectrometry **12:**49-54.
11. **Farmer, T. B., and R. M. Caprioli.** 1998. Determination of protein-protein interactions by matrix-assisted laser desorption/ionization mass spectrometry. Journal of Mass Spectrometry **33:**697-704.
12. **Fenselau, C., and P. A. Demirev.** 2001. Characterization of intact microorganisms by MALDI mass spectrometry. Mass Spectrometry Reviews **20:**157-171.
13. **Fridkin, S. K., and W. R. Jarvis.** 1996. Epidemiology of nosocomial fungal infections. Clinical Microbiology Reviews **9:**499-&.
14. **Gluckmann, M., A. Pfenninger, R. Kruger, M. Thierolf, M. Karas, V. Horneffer, F. Hillenkamp, and K. Strupat.** 2001. Mechanisms in MALDI analysis: surface interaction or incorporation of analytes? International Journal of Mass Spectrometry **210:**121-132.
15. **Haag, A. M., S. N. Taylor, K. H. Johnston, and R. B. Cole.** 1998. Rapid identification and speciation of *Haemophilus* bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Journal of Mass Spectrometry **33:**750-756.

16. **Han, Y. M., and J. E. Cutler.** 1995. Antibody response that protects against disseminated candidiasis. Infection and Immunity **63:**2714-2719.
17. **Harrington, P. D., T. E. Street, K. J. Voorhees, F. R. Dibrozolo, and R. W. Odom.** 1989. Rule-building expert system for classification of mass-spectra. Analytical Chemistry **61:**715-719.
18. **Harrington, P. D., and K. J. Voorhees.** 1990. Multivariate rule building expert system. Analytical Chemistry **62:**729-734.
19. **Hathout, Y., P. A. Demirev, Y. P. Ho, J. L. Bundy, V. Ryzhov, L. Sapp, J. Stutler, J. Jackman, and C. Fenselau.** 1999. Identification of *Bacillus* spores by matrix-assisted laser desorption ionization mass spectrometry. Applied and Environmental Microbiology **65:**4313-4319.
20. **Hazen, K. C.** 1995. New and Emerging Yeast Pathogens. Clinical Microbiology Reviews **8:**462-478.
21. **Himmelreich, U., R. L. Somorjai, B. Dolenko, H. M. Daniel, and T. C. Sorrell.** 2005. A rapid screening test to distinguish between *Candida albicans* and *Candida dubliniensis* using NMR spectroscopy. FEMS Microbiology Letters **251:**327-332.
22. **Himmelreich, U., R. L. Somorjai, B. Dolenko, O. C. Lee, H. M. Daniel, R. Murray, C. E. Mountford, and T. C. Sorrell.** 2003. Rapid identification of *Candida* species by using nuclear magnetic resonance spectroscopy and a statistical classification strategy. Applied and Environmental Microbiology **69:**4566-4574.
23. **Holland, R. D., J. G. Wilkes, F. Rafii, J. B. Sutherland, C. C. Persons, K. J. Voorhees, and J. O. Lay.** 1996. Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry **10:**1227-1232.
24. **Horneffer, V., J. Haverkamp, H. G. Janssen, and R. Notz.** 2004. MALDI-TOF-MS analysis of bacterial spores: Wet heat-treatment as a new releasing technique for biomarkers and the influence of different experimental parameters and microbiological handling. Journal of the American Society for Mass Spectrometry **15:**1444-1454.
25. **Ibelings, M. S., K. Maquelin, H. P. Endtz, H. A. Bruining, and G. J. Puppels.** 2005. Rapid identification of *Candida* spp. in peritonitis patients by Raman spectroscopy. Clinical Microbiology and Infection **11:**353-358.
26. **Jackson, K. A., V. Edwards-Jones, C. W. Sutton, and A. J. Fox.** 2005. Optimisation of intact cell MALDI method for fingerprinting of methicillin-resistant *Staphylococcus aureus*. Journal of Microbiological Methods **62:**273-284.
27. **Karas, M., and F. Hillenkamp.** 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Analytical Chemistry **60:**2299-2301.
28. **Krishnamurthy, T., P. L. Ross, and U. Rajamani.** 1996. Detection of pathogenic and non-pathogenic bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry **10:**883-888.
29. **Lay, J. O., and R. D. Holland.** 2000. Rapid identification of bacteria based on spectral patterns using MALDI-TOFMS. Methods in Molecular Biology **146:**461-487.
30. **Li, T. Y., B. H. Liu, and Y. C. Chen.** 2000. Characterization of *Aspergillus* spores by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry **14:**2393-2400.
31. **Madonna, A. J., F. Basile, E. Furlong, and K. J. Voorhees.** 2001. Detection of bacteria from biological mixtures using immunomagnetic separation combined with matrix-

assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry **15:**1068-1074.

32. **Maquelin, K., L. P. Choo-Smith, H. P. Endtz, H. A. Bruining, and G. J. Puppels.** 2002. Rapid identification of *Candida* species by confocal Raman micro spectroscopy. Journal of Clinical Microbiology **40:**594-600.
33. **Maquelin, K., L. P. Choo-Smith, T. van Vreeswijk, H. P. Endtz, B. Smith, R. Bennett, H. A. Bruining, and G. J. Puppels.** 2000. Raman spectroscopic method for identification of clinically relevant microorganisms growing on solid culture medium. Analytical Chemistry **72:**12-19.
34. **Marvin-Guy, L. F., S. Parche, S. Wagniere, J. Moulin, R. Zink, M. Kussmann, and L. B. Fay.** 2004. Rapid identification of stress-related fingerprint from whole bacterial cells of Bifidobacterium lactis using matrix assisted laser desorption/ionization mass spectrometry. Journal of the American Society for Mass Spectrometry **15:**1222-1227.
35. **Palmer, G. E., M. N. Kelly, and J. E. Sturtevant.** 2005. The *Candida albicans* vacuole is required for differentiation and efficient macrophage killing. Eukaryotic Cell **4:**1677-1686.
36. **Perfect, J. R., and A. Casadevall.** 2002. Cryptococcosis. Infectious Disease Clinics of North America **16:**837-874.
37. **Pincus, D. H., S. Orenga, and S. Chatellier.** 2007. Yeast identification - past, present, and future methods. Medical Mycology **45:**97-121.
38. **Qian, Q. F., M. A. Jutila, N. Vanrooijen, and J. E. Cutler.** 1994. Elimination of mouse splenic macrophages correlates with increased susceptibility to experimental disseminated candidiasis. Journal of Immunology **152:**5000-5008.
39. **Ruhnke, M.** 2006. Epidemiology of *Candida albicans* infections and role of non-*Candida-albicans* yeasts. Current Drug Targets **7:**495-504.
40. **Saenz, A. J., C. E. Petersen, N. B. Valentine, S. L. Gantt, K. H. Jarman, M. T. Kingsley, and K. L. Wahl.** 1999. Reproducibility of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for replicate bacterial culture analysis. Rapid Communications in Mass Spectrometry **13:**1580-1585.
41. **Shaw, E. I., H. Moura, A. R. Woolfitt, M. Ospina, H. A. Thompson, and J. R. Barr.** 2004. Identification of biomarkers of whole *Coxiella burnetii* phase I by MALDI-TOF mass spectrometry. Analytical Chemistry **76:**4017-4022.
42. **Sheppard, D. C., E. DeSouza, Z. Hashmi, H. G. Robson, and P. Rene.** 1998. Evaluation of the Auxacolor system for biochemical identification of medically important yeasts. Journal of Clinical Microbiology **36:**3726-3727.
43. **Sherburn, R. E., and R. O. Jenkins.** 2003. A novel and rapid approach to yeast differentiation using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Spectroscopy **17:**31-38.
44. **Strupat, K., D. Sagi, H. Bonisch, G. Schafer, and J. Peter-Katalinic.** 2000. Oligomerization and substrate binding studies of the adenylate kinase from *Sulfolobus acidocaldarius* by matrix-assisted laser desorption/ionization mass spectrometry. Analyst **125:**563-567.
45. **Timmins, E. M., S. A. Howell, B. K. Alsberg, W. C. Noble, and R. Goodacre.** 1998. Rapid differentiation of closely related *Candida* species and strains by pyrolysis mass spectrometry and Fourier transform-infrared spectroscopy. Journal of Clinical Microbiology **36:**367-374.

46. **Tintelnot, K., G. Haase, M. Seibold, F. Bergmann, M. Staemmler, T. Franz, and D. Naumann.** 2000. Evaluation of phenotypic markers for selection and identification of *Candida dubliniensis*. Journal of Clinical Microbiology **38:**1599-1608.
47. **Valentine, N. B., J. H. Wahl, M. T. Kingsley, and K. L. Wahl.** 2002. Direct surface analysis of fungal species by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Communications in Mass Spectrometry **16:**1352-1357.
48. **Vargha, M., Z. Takats, A. Konopka, and C. H. Nakatsu.** 2006. Optimization of MALDI-TOF MS for strain level differentiation of *Arthrobacter* isolates. Journal of Microbiological Methods **66:**399-409.
49. **Vorm, O., P. Roepstorff, and M. Mann.** 1994. Improved resolution and very high-sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. Analytical Chemistry **66:**3281-3287.
50. **Whelan, W. L., J. M. Delga, E. Wadsworth, T. J. Walsh, K. J. Kwonchung, R. Calderone, and P. N. Lipke.** 1990. Isolation and characterization of cell-surface mutants of *Candida albicans*. Infection and Immunity **58:**1552-1557.
51. **Williams, T. L., D. Andrzejewski, J. O. Lay, and S. M. Musser.** 2003. Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. Journal of the American Society for Mass Spectrometry **14:**342-351.
52. **Wunschel, S. C., K. H. Jarman, C. E. Petersen, N. B. Valentine, K. L. Wahl, D. Schauki, J. Jackman, C. P. Nelson, and E. White.** 2005. Bacterial analysis by MALDI-TOF mass spectrometry: An inter-lab oratory comparison. Journal of the American Society for Mass Spectrometry **16:**456-462.
53. **Xiang, F., and R. C. Beavis.** 1994. A method to increase contaminant tolerance in protein matrix-assisted laser-desorption ionization by the fabrication of thin protein-doped polycrystalline films. Rapid Communications in Mass Spectrometry **8:**199-204.

# CHAPTER 3

## Synthesis and Applications of "Fluorous" (Fluorinated Alkyl) Affinity Reagents That Label Primary Amine Groups in Proteins/Peptides

## 3.1 Abstract

Strong non-covalent interactions such as biotin-avidin affinity are commonly utilized in protein/peptide enrichment/separation. A new type of "fluorous" (fluorinated alkane) affinity gained popularity due to its low level non-specific binding to proteins/peptides. We developed two generations of fluorous labeling reagents which were reactive toward primary amine groups in proteins/peptides. The 2$^{nd}$ generation reagents were highly water soluble and the labeling reactions with proteins/peptides could proceed in an aqueous environment. The acid labile linker in the labeling reagent allows cleavage of the bulky fluorous tag moiety and the long oligo ethylene glycol (OEG) spacer after fluorous affinity purification. Upon collision induced decomposition, the labeled peptide ion yielded a characteristic fragment that could be retrieved from the residual portion of fluorous affinity tag, and served as a marker to indicate that the relevant peptide had been successfully labeled. Analysis of results showed that both the labeling and affinity enrichment/separation process were highly efficient.

## 3.2 Introduction

Proteins and peptides in living organisms accumulate at different levels with wide dynamic range. It is usually necessary to enrich and/or purify native proteins and peptides for further characterization. Given the complexity of biological samples, extraction and fractionation of proteins/peptides of interest can be challenging.

Conjugating an affinity tag to protein/peptide is an attractive approach in protein/peptide enrichment and purification. Certain functional groups, such as a thiol group from cysteine(7), an amine group from lysine (20) (12) and a phenol group from tyrosine (10) could serve as direct targets for conjugation. A few other functional groups from post modifications, such as phosphoserine and phosphothreonine were suitable to be converted to thiol first, and then conjugated with an affinity tag (1, 14).Biotin is almost exclusively chosen as the affinity tag moiety in conjugating reagents due to the biotin-avidin interaction, which is one of the strongest non-covalent interactions in nature with a dissociation constant ($K_d$) of $10^{-15}$ (21). For affinity purification purpose, this well-known biotin-avidin interaction also has several limitations: 1. Some proteins/peptides may nonspecifically bind to avidin (11); 2. Low recovery of the biotin-tagged molecules in affinity purification is common because of the incomplete elution of biotinylated molecules from avidin resins (4) (19); 3. The biotin moiety tends to average the overall hydrophobicity of the tagged-peptides and compromise the separation of these peptides on reverse-phased liquid chromatography column (9); 4. Fragments originated from biotin moiety can complicate the mass spectral interpretation of biotin tagged peptides (14). Incorporation of cleavable linker between the biotin moiety and the peptide reactive group allows the removal of biotin moiety, and this approach has been employed in cleavable ICAT reagents (9, 22) that targets the free thiol group in proteins/peptides.

The fluorine-fluorine interaction ("fluorous affinity") based binding and separation was first demonstrated in fluorous biphasic catalysis technique (8). Since then the fluorous affinity method has been successfully utilized in a wide spectrum of organic syntheses (23). Recently, its application was introduced to proteomics (3), metabolomics (6), and microarrays for immobilizing small-biomolecules (17) (13, 16). The flourous affinity is characterized by the strong non-covalent interactions between the fluorous affinity tag (which can be conjugated to molecules of interest) and fluorous resin. The interactions between fluorous resin and proteins/peptides are generally weak, the non-specific binding of protein/peptides to the fluorous resin, if any, can be easily disrupted. Moreover, fluorous affinity based separation technique allows fractionation of the same molecules with different number of fluorous tags (3). Similar phenomenon can not be found in biotin-avidin based separation approaches.

Unlike the biotin group, which is quite small (244 Da), the fluorous tags are usually quite large with 448 Da for a commonly used tag $C_8F_{17}CH_2CH_3$. The strength of fluorous affinity increases with the number of fluorine in the fluorous tag (5). Moreover, fluorous compounds are general water insoluble due to the very strong hydrophobicity of fluorous tag. This main caveat limits its application in biological system where physiological environment need be maintained (e.g. in PBS buffer) and may cause protein aggregation from multiple "post-modifications" by fluorous tags. Also, the bulky fluorous affinity tag complicates the fragmentation pattern of tagged peptides and makes it more difficult to interpret the tandem mass spectra of tagged peptides. Incorporation of a polar, hydrophilic spacer between the fluorinated alkyl moiety and the reactive group will increase the solubility of the reagent and labeled protein, however, the fragmentation pattern of the tagged peptides will be complicated by the fragment ions from the spacer.

Therefore, it is necessary to remove the fluorous tag and spacer from the labeled molecule before LC-MS analysis.

In this study, we developed two generations of fluorous reagents (1st generation fluorous labeling compound: Boc-pentanoic-NHS-perfluoroctane (FBPNHS) (Figure 3.1 A ) and 2nd generation compounds including two analogues: sulfo-NHS-$(OEG)_3$-perfluoroctane (FP3CMSNHS), sulfo-NHS-$(OEG)_1$-perfluoroctane (FP1CMSNHS) (Figure 3.1 B and C) which both could label free amines from peptides and proteins. Compared to our 1st generation labeling reagent (FBPNHS), which is not water soluble, the 2nd generation molecules are novel water soluble fluorous labeling reagents that included a protein/peptide reactive group (Sulfo-N-Hydroxylsuccinimidyl ester, Sulfo-NHS) (18), an acid labile linker (a tertiary carbomate group), an oligoethylene glycol $(OEG)_3$ spacer and a perfluorinated alkane moiety ($C_8F_{17}$). Labeling protein/peptides with our synthesized flourous labeling reagents (A and C in Figure 3.1) was demonstrated along with the purification/separation and LC-MS analysis of the tagged protein/peptide.

## 3.3 Materials and Methods

Fluorenylmethoxycarbonyloxy)succinimide (Fmoc-OSu), $LiALH_4$, pyridine, N,N-tertrabutylammonium fluoride (TBAF), trifluoroacetic acid (TFA), ammonium formate, bovine serum albumin (BSA), DL-dithiothreitol (DTT), iodoacetamide (IA) and 5-aminovaleric acid were purchased from Sigma (St. Louis, MO). 2-{2-[2-(2-Amino-ethoxy)-ethoxy]-ethoxy}-ethanol and [N- (ε –maleimidocaproyloxy) sulfosuccinimide ester] (Sulfo-EMCS) were purchased from Molecular Bioscience, Inc (Boulder, CO).

N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDCI) (Advanced ChemTech, Louisville, KY). 1-hydroxybenzotriazole, anhydrous (HOBt) was purchased from Chem-Impex International, Inc.(Wood Dale, IL). N,N′-disuccinimidyl carbonate (DSC), 1-Boc-piperazine (Oakwood Products, Inc., West Columbia, SC), tetrabutylammonium fluoride (TBAF,) and Tris(2-carboxyethyl) phosphine hydrochloride (TCEP hydrochloride) were purchased from EMD Chemicals Inc. (Gibbstown, NJ). Diisopropylethylamine (DiPEA), and p-nitrophenylchloromate were purchased from Acros (Raritan, New Jersey). ACTH (4-11) was purchased from American Peptide Company, Inc. (Sunnyvale, CA). Sequencing grade modified trypsin was purchased from Promega, Inc (Madison, WI). Fluorous silica gel (5 μm in diameter), 2H,2H,3H,3H-perfluoroundecanoic acid and 2-[2-(1H,1H,2H,2H-perfluorohexyl) isopropoxycarbonyloxyimino]-2-phenylacetonitrile (F-BOC-ON) were purchased from Fluorous Technologies Inc (Pittsburg, PA).

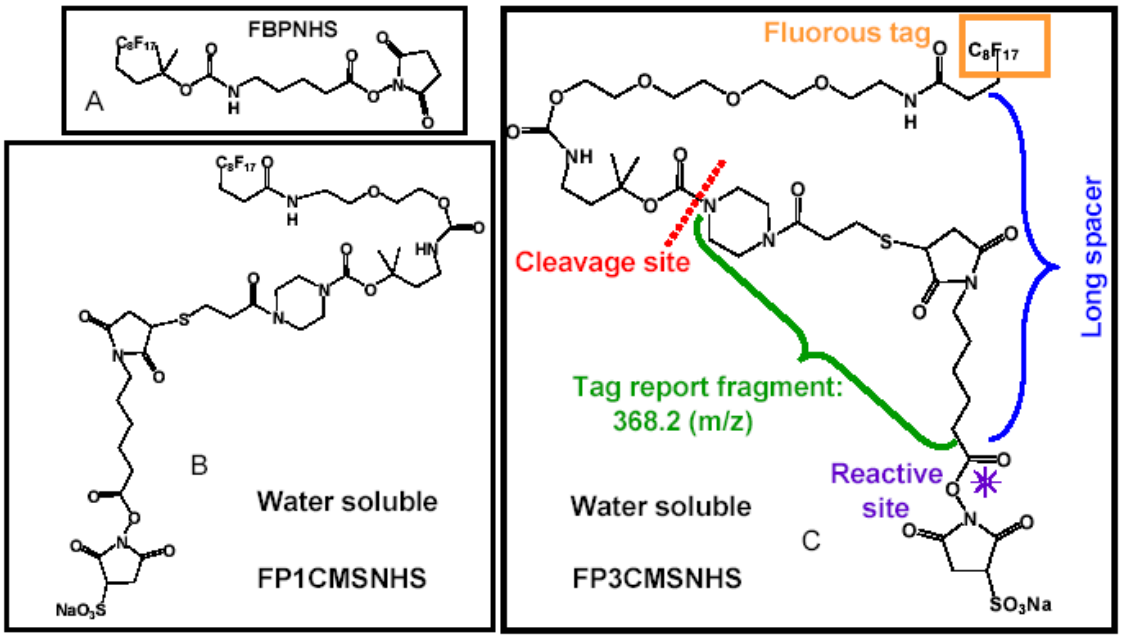



**Figure 3.1.** Structures of synthesized fluorous labeling reagents A, B and C.

## 3.4 Experimentals

**3.4.1 Synthesis of fluorous labeing reagent (FBPNHS)** (Scheme 3.1)

**5-(4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-heptadecafluoro-1,1-dimethyl-undecyloxycarbonylamino)-pentanoic acid (1')**

90.7 mg (0.77 mM, 1.05 eq) of 5-amino-pentanoic acid was taken up in 1.5 mL of dichloromethane and 315 μL of diisopropylethylamine (2.5 eq) was added.

The stirred mixture was then cooled to 0° C in an ice bath. To this mixture was added 110 μL of trimethylsilylchloride (0.87 mM, 1.2 eq) under nitrogen and the reaction was allowed to proceed for 10 minutes. Then a solution of F-BOC-ON (500 mg, 0.73 mM, 1.0 eq) in 2 mL THF was slowly added to the flask. The ice bath was removed to let the reaction continue for 10 hours at room temperature. The reaction was quenched with 2mL of 2N HCl and extracted into ether 2-3 times. The organic layers are combined, dried over magnesium sulfate, followed by solvents evaporation. The extracted product was purified by silica gel column chromatography with ethyl acetate-hexane 1:3 as the elution solvent. Purified product **2** (464 mg, 98%) was obtained as white solid **1'**. $^{1}$H NMR (400 MHz, DMSO-$d_6$) δ 1.34 (t, J=7.0 Hz, 2H), δ 1.38 (s, 6H), 1.44 (t, J=7.0 Hz, 2H), 1.94 (t, J=7.0 Hz, 2H), 2.16 (t, J= 7.0 Hz, 4H), 2.88 (q, J=6.2 Hz ,2H), 6.94 (t, J=5.1 Hz, 1H), 11.96 (s, 1H)**.** ESI-MS: $(M+H)^+$ =650.1.

**1.2 5-(4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-Heptadecafluoro-1,1-dimethyl-undecyloxycarbonylamino)-pentanoic acid 2,6-dioxo-piperidin-1-yl ester (FBPNHS) (2'')**

To a stirred solution of **2** (200 mg, 0.31 mM, 1.0 eq) was dissolved in 2mL anhydrous DMF, to which DSC ( 158 mg, 0.62 mM, 2.0 eq) and pyridine ( 73 mg, 0.93 mM, 3.0 eq) were added. The mixture was stirred overnight at room temperature and solvent was

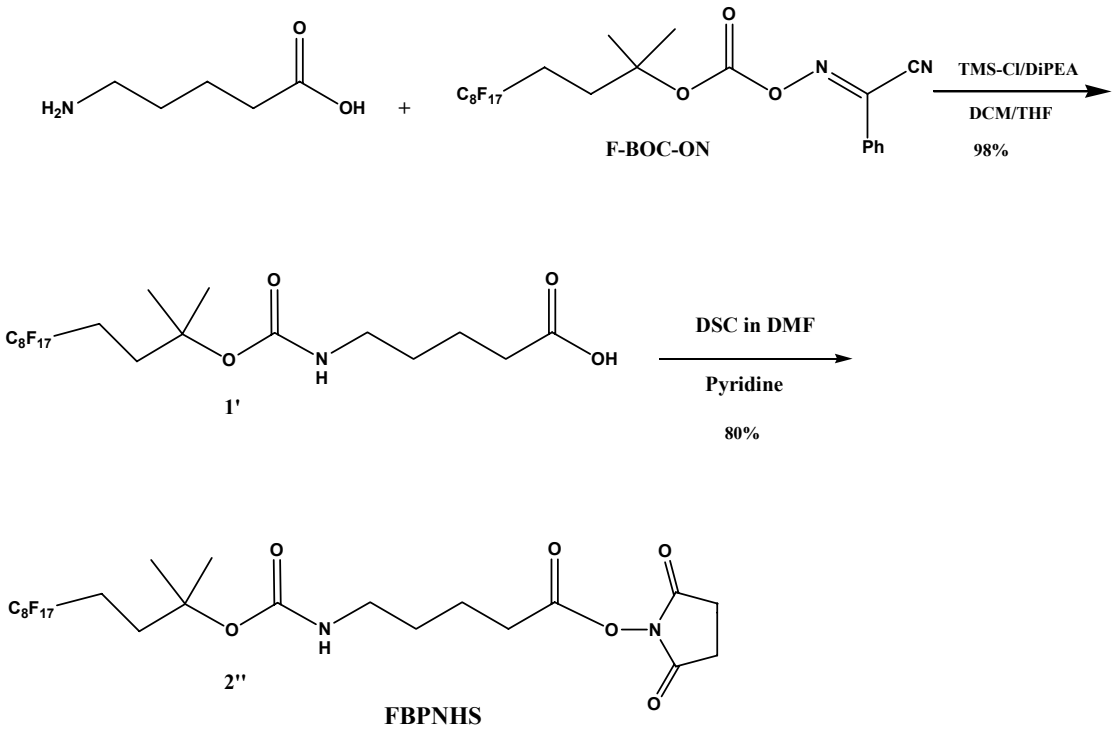
removed under vacuum. Crude product was purified by silica gel column with ethyl acetate-hexane 2:1 as the elution solvent to give the pure product **3** (188 mg, 80%) as white solid **2''**. $^{1}$H NMR (400 MHz, $CDCl_3$) δ 1.48 (s, 6H), 1.59 (q, J=7.0 Hz, 2H), 1.75 (q, J=7.0 Hz, 2H), 2.12 (m, 2H), 2.75 (m, 6H), 3.15 (t, J=7.0 Hz, 2H), 4.80 (s, 1H). ESI-MS: $(M+H)^+$ =761.1.

### 3.4.2 Labeling small peptides MRFA and ACTH (4-11) using FBPNHS

To avoid rapid hydrolysis of the compound, excessive amount of peptide was used in this study. 10μL of 153 mM MRFA aqueous solution was mixed with 50 μL of 2.5 mM FBPNHS THF solution. The reaction proceeded at room temperature for one hour. 0.5 μL of mixture was diluted in 60 μL of ethanol containing 5% FA and analyzed on the mass spectrometer through direct infusion.

Longer peptides ACTH (4-11) were dissolved in DMF : $NaHCO_3$ (50 mM in water) = 3: 1 (v/v, pH=7.8). 9 μL of 50 mM ACTH (4-11) was mixed 20 μL of 50 mM FBPNHS in THF solution. The mixture was set at room temperature for 1hour, followed by addition of 2 μL of 100 mM Tris (in THF/$H_2O$ ,vol/vol=1/1) and mixing for 30 minutes at room temperature. The products were dried under vacuum, and the fluorous tag was cleaved in 100 μL of ICAT cleavage reagents according to the procedures described in the manufacture's protocol (Applied Biosystems). After 1 hour, an aliquot of 1 μL solution was diluted to 100 μL of 50% MeOH for immediate ESI-MS analysis.

### 3.4.3 Labeling a standard protein bovine serum albumin (BSA) using FBPNHS




**Scheme 3.1.** Synthesis of 1st generatation fluorous labeling reagent Boc-pentanoic-NHS-perfluoroctane (FBPNHS)

BSA (~0.7 nM, 20 μL) (Sigma-Aldrich, St Louis, MO) in the same solvents used in the peptide labeling above was reacted with 15 μL of FBPNHS (1.5 μM) in THF solution with mixing at room temperature. The reaction was allowed to proceed for 1 hour and quenched by adding 5μL of 5 M Tris and mixing for 20 minutes. After solvents evaporation, 200 μL of the ICAT cleavage reagents was added. The cleavage reaction was allowed to proceed for 2 hours at 37°C. After the completion of the cleavage, solvents were evaporated, and the labeled BSA was re-dissolved in 50% MeOH and separated from excessive labeling reagents, which became insoluble in 50% MeOH. The labeled BSA was vacuum dried and dissolved in 50 mM ammonium bicarbonate for alkylation and trypsin digestion. Digested tryptic peptides were analyzed by LC-MS and by database search as described below.

**3.4.4 Synthesis of Sulfo-NHS-$(OEG)_3$-perfluoroctane** ( Scheme 3.2 to Scheme 3.5 )

**4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-heptadecafluoro-N-(2-(2-(2-(2-hydroxyethoxy)ethoxy)ethoxy)ethyl)undecanamide (1)**

EDCI (0.64 g, 3.3 mM, 1.1 eq) and HOBt (0.68 g, 4.5 mM, 1.5 eq) was added to a stirred mixture of 2-(2-(2-(2-aminoethoxy)ethoxy)ethoxy)ethanol (0.62 g, 3.2 mM, 1.05 eq) and 4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-heptadecafluoroundecanoic acid (1.50 g, 3.0 mM, 1.0 eq.) in 20 mL anhydrous dichloromethane ($CH_2Cl_2$) at room temperature. After 16 hr stirring, the solution was diluted by $CH_2Cl_2$, washed by saturated sodium bicarbonate, extracted by $CH_2Cl_2$. Combined organic layer was washed by brine, dried over $MgSO_4$ and concentrated under reduced pressure. The crude product was purified by flash column chromatography ($SiO_2$, eluent: 9:1 $CH_2Cl_2$/MeOH) as pale brown oil **1** (1.56 g, yield: 78%). $^1H$ NMR (400 MHz, $CDCl_3$) δ 2.49

(m, 2H), 2.58 (m, 2H), 3.49 (t, J=5.0 Hz, 2H), 3.56 (t, J=5.1 Hz, 2H), 3.67 (m, 4H), 3.71 (m, 4H), 3.81 (t, J=6.0 Hz, 2H), 4.45 (t, J=6.0 Hz, 2H), 6.26 (s, 1H). ESI-MS: $(M+H)^+$ =668.1.

**16,16,17,17,18,18,19,19,20,20,21,21,22,22,23,23,23-heptadecafluoro-13-oxo-3,6,9-trioxa-12-azatricosyl 4-nitrobenzoate (2)**

4-nitrophenyl chlorofomate (0.72 g, 3.6 mM, 1.5 eq.) in 5mL anhydrous $CH_2Cl_2$ was slowly added to a solution of **c** (1.4 g, 2.4 mM, 1.0 eq.) and DiPEA (0.93g, 7.2 mM, 3.0 eq.) in 15 mL anhydrous $CH_2Cl_2$ at 4°C. The mixture was stirred for 16 hr while gradually warmed to the room temperature. The reaction was quenched by adding ice-cold 0.5N HCl, extracted by $CH_2Cl_2$. Organic layer was washed by $H_2O$, brine, dried over $MgSO_4$ and concentrated in vacuo. The crude product was purified by flash column chromatography ($SiO_2$, eluent: 1:1 hexane/ethylacetate) as pale yellow waxy solid **2** (1.80 g, yield: 90%). $^1H$ NMR (400 MHz, $CDCl_3$) δ 2.46 (m, 2H), 2.49 (m, 2H), 3.47 (t, J=5.0 Hz, 2H), 3.57 (t, J=5.0 Hz, 2H), 3.65 (m, 4H), 3.70 (m, 4H), 3.81 (t, J=6.0 Hz, 2H), 4.44 (t, J=6.0 Hz, 2H), 6.25 (s, 1H), 7.38 (d, J=9.2 Hz, 2H), 8.28 (d, J=9.2 Hz, 2H). ESI-MS: $(M+H)^+$ =817.1

**4-amino-2-methyl-butan-2-ol (3)**

To a vigorously stirred solution of $LiAlH_4$ (2.8 g, 74 mM, 5.0 eq.) in 140 mL ice-cold dry ether was slowly added 2-methylpent-4-yn-2-ol (1.45 g, 14.8 mM, 1.0 eqiv.) in 50 mL dry ether. After stirring for 1 hour at room temperature, the mixture was refluxed for 10 hours. After cooling, 15 mL ice cold 10% NaOH was slowly added to the reaction under stirring to quench the reaction. After additional 2 hour stirring, precipitates were filtered and washed thoroughly by 70 mL THF three times. After removal of THF in vacuo, 80 mL of 15% NaOH was added to the residue

followed by extraction of 90 mL DCM three times. Organic phase was combined, washed by brine and dried over $MgSO_4$ overnight. Solvents were evaporated to afford product as pale yellow oil **3** (1.40g, yield: 92%), which was used for the next step without further purification.

**(3-Hydroxy-3-methyl-butyl)-carbamic acid 9H-fluoren-9-ylmethyl ester (4)**

Compound **3** (0.60 g, 5.8 mM, 1.0 eq ) was taken up in 20 mL 5% $Na_2CO_3$ and cooled off in an ice-water slurry bath. Fmoc-Osu (2.16 g, 6.4 mM, 1.1 eq) in 25 mL THF was slowly added to the mixture. After stirring for 16 hr, THF was removed *in vacuo* and the aqueous phase was extracted by 70 mL DCM three times. The extracts were combined, washed by brine, dried over $MgSO_4$ and evaporated to dryness. Crude product was purified by flash column chromatography ($SiO_2$, eluent: 6:1 DCM/ethylacetate) as white solid **4** (1.75g, yield: 93 %). $^1$H NMR (400 MHz, $CDCl_3$) δ 1.27 (s, 6H), 1.71 (t, J=6.8 Hz, 2H), 3.64 (t, J=6.8 Hz, 2H), 4.22 (t, J=6.8 Hz, 1H), 4.38 (d, J=6.8 Hz, 2H), 5.37 (s, 1H), 7.31 (t, J=7.2 Hz, 2H), 7.40 (t, J=7.2 Hz, 2H), 7.60 (d, J=7.2 Hz, 2H), 7.76(d, J=7.2 Hz, 2H). ESI-MS: $(M+H)^+$ =326.2

**Carbonic acid 3-(9H-fluoren-9-ylmethoxycarbonylamino)-1,1-dimethyl-propyl ester 4-nitro-phenyl ester (5)**

To a solution of **4** (0.76 g, 2.3 mM, 1.0 eq) and pyridine (5 mL) in 25 mL DCM was slowly added 4-nitrophenyl chlorofomate (0.94 g, 4.6 mM, 2.0 eq) in 9 mL DCM at 0°C. The mixture was stirred for 8 hours at room temperature and poured into 90 mL 0.5 N cold HCl. 300 mL DCM was used for 3 times of extraction and the extracts were washed by brine, dried over $MgSO_4$ before evaporated to dryness. Crude product was purified by flash column chromatography ($SiO_2$, eluent: 5:1 hexane/ethylacetate) as colorless to light yellow oil **5** (1.10 g,

yield: 96%). $^{1}$H NMR (400 MHz, $CDCl_3$) δ 1.60 (s, 6H), 2.07 (t, J=7.6 Hz, 2H), 3.40 (t, J=7.6 Hz, 2H), 4.22 (t, J=6.8 Hz, 1H), 4.42 (d, J=6.8 Hz, 2H), 4.93 (s, 1H), 7.28 (t, J=7.2 Hz, 2H), 7.23 (t, J=7.2 Hz, 2H), 7.35 (d, J=8.4 Hz, 2H), 7.40 (t, J=7.2 Hz, 2H), 7.58 (d, J=7.2 Hz, 2H), 7.76 (d, J=7.2 Hz, 2H), 8.24 (d, J=8.4 Hz, 2H). ESI-MS: $(M+H)^+$ =491.2

**Bis(2,5-dioxopyrrolidin-1-yl) 3,3'-disulfanediyldipropanoate (6)**

To a solution of 3,3'-disulfanediyldipropanoic acid (50 mg, 0.24 mM, 1.0 eq) and pyridine (0.11 mL, 6.0 eq) in 3mL DCM was added DSC (182 mg, 0.72 mM, 3.0 eq). The mixture was stirred for at room temperature overnight then diluted by 10 mL DCM before poured into 10 mL 0.2 N cold HCl. Additional 50 mL DCM was used for extractions and the extracts were washed by brine, dried over $MgSO_4$ before evaporated to dryness. Crude product was purified by flash column chromatography ($SiO_2$, eluent: 2:3 hexane/ethylacetate) as white solid **6** (89 mg, yield 93%). ESI-MS: $(M+H)^+$ =405.53

**Di-tert-butyl 4,4'-(3,3'-disulfanediylbis(propanoyl))bis(piperazine-1-carboxylate) (7)**

To a solution of 6 (0.5 g, 1.23 mM, 1.0 eq) and DiEA (0.47 mL) in 11 mL DCM was slowly added tert-butyl piperazine-1-carboxylate (0.51 g, 2.72 mM, 2.2 eq) in 5 mL DCM at room temperature. After 3 hours stirring, solvents were evaporated and crude product was recrystalized by hexane/ethylacetate as white solid **7** (0.62 g, yield: 94%). $^{1}$H NMR (400 MHz, CDCl3) δ 1.47 (s, 18H), 2.76 (t, J=6.8 Hz, 4H), 2.97 (t, J=6.8 Hz, 4H), 3.46 (m, 8H), 3.48 (t, J=4.9 Hz, 4H), 3.60 (t, J=4.9 Hz, 4H). ESI-MS: $(M+H)^+$ =546.4

**Bis(4-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-2-methylbutan-2-yl) 4,4'-(3,3'-disulfanediylbis(propanoyl))bis(piperazine-1-carboxylate) (8)**

Compound **7** (0.5 g, 0.91 mM, 1.0 eq.) was taken up by 2.5 mL of 50% TFA in DCM under stirring at room temperature. After 1 hour cleavage reaction, solvents were evaporated under reduced pressure and the resulting residue was mixed with 6 mL DCM, 1.7 mL DiEA and HOBt (27.6 mg, 0.20 eq). To this mixture was added compound **5** (0.98 g, 2.0 mM, 2.2 eq) in 5 mL DCM. After stirring for 16 hours, 30 mL of 0.5 N cold HCl was added to quench the reaction and 200 mL DCM was then used for 3 times of extractions. The extracts were washed by brine, dried by $MgSO_4$ and evaporated to dryness. Crude product was purified by flash column chromatography ($SiO_2$, eluent: ethylacetate) as white solid **8** (0.7 g, yield: 73%). $^1$H NMR (400 MHz, $CDCl_3$) δ 1.49 (s, 12H), 2.30 (t, J=6.8 Hz, 4H), 2.74 (t, J=7.0 Hz, 4H), 2.95 (t, J=7.0 Hz, 4H), 3.28 (t, J=6.8 Hz, 4H), 3.41 (m, 4H), 3.44 (m, 8H), 3.57 (m, 4H), 4.20 (t, J=6.4 Hz, 2H), 4.40 (d, J=6.4 Hz, 4H), 4.80 (s, 2H), 7.31 (t, J=7.6 Hz, 4H), 7.40 (t, J=7.6 Hz, 4H), 7.58 (d, J=7.6 Hz, 4H), 7.62 (d, J=7.6 Hz, 4H). ESI-MS: $(M+H)^+$ =1049.4

**23,23,24,24,25,25,26,26,27,27,28,28,29,29,30,30,30-heptadecafluoro-2-methyl-6,20-dioxo-7,10,13,16-tetraoxa-5,19-diazatriacontan-2-yl 4-(3-((3-(4-(24,24,25,25,26,26,27,27,28,28,29,29,30,30,31,31,31-heptadecafluoro-3,3-dimethyl-7,21-dioxo-2,8,11,14,17-pentaoxa-6,20-diazahentriacontan-1-oyl)piperazin-1-yl)-3-oxopropyl)disulfanyl)propanoyl)piperazine-1-carboxylate (9 )**

To a solution of **8** in 0.7 mL anhydrous DMF was added 66 mg TBAF. After 2 hours, TLC showed complete de-protection and compound **2** (135 mg, 0.16 mM, 2.0 eqiv.) in 4 mL anhydrous DCM was slowly added at room temperature, followed by addition of DiEA (0.03

mL). The mixture was stirred overnight, and then the reaction was quenched by adding 10 mL ice cold 0.2 N HCl. The separated aqueous phase was further extracted by 60 mL DCM 3 times. The extracts were washed by brine, dried by $MgSO_4$ and evaporated to dryness. Crude product was purified by flash column chromatography ($SiO_2$, eluent: 13:1 DCM: MeOH) as pale oil **9** (135 mg, yield: 84 %). $^1$H NMR (400 MHz, $CDCl_3$) δ 1.47 (s, 12H), 2.00 (t, J=7.2 Hz, 4H), 2.47 (m, 4H), 2.51 (m, 4H), 2.76 (t, J=7.1 Hz, 4H), 2.95 (t, J=7.0 Hz, 4H), 3.25 (q, J=7.2 Hz, 4H), 3.40 (m, 16H), 3.56 (t, J=5.0 Hz, 4H), 3.60 (m, 24H), 4.19 (t, J=5.0 Hz, 4H), 4.99 (s, 2H), 6.63 (s, 2H). ESI-MS: $(M+H)^+$ =1991.5

**23,23,24,24,25,25,26,26,27,27,28,28,29,29,30,30,30-heptadecafluoro-2-methyl-6,20-dioxo-7,10,13,16-tetraoxa-5,19-diazatriacontan-2-yl 4-(3- mercaptopropanoyl)piperazine -1-carboxylate (10)**

0.6 mL of TCEP·HCl (65 mg, 0.23 mM, 4.0 eq ) in saturated $NaHCO_3$ was added to 3 mL THF solution of compound **9** (113 mg, 0.057 mM, 1.0 eq). After stirring for 9 hours, solvents were evaporated under reduced pressure. The crude product was purified by flash column chromatography ($SiO_2$, eluent: 20:1 DCM: MeOH) as pale brown oil **10** (52 mg, yield: 93%). $^1$H NMR (400 MHz, $CDCl_3$) δ 1.47 (s, 6H), 1.73 (t, J=8.0 Hz, 1H), 2.00 (t, J=7.2 Hz, 2H), 2.49 (m, 2H), 2.52 (m, 2H), 2.65 (t, J=6.8 Hz, 2H), 2.81 (q, J=7.2 Hz, 2H), 3.24 (q, J=7.2 Hz, 2H), 3.40 (m, 8H), 3.56 (t, J=5.2 Hz, 2H), 3.63 (m, 12H), 4.19 (t, J=5.2 Hz, 2H), 4.99 (s, 1H), 6.63 (s, 1H). ESI-MS: $(M+H)^+$ =997.3

**Sodium 1-((6-(3-((3-(4-(24,24,25,25,26,26,27,27,28,28,29,29,30,30,31,31,31-heptadecafluoro-3,3-dimethyl-7,21-dioxo-2,8,11,14,17-pentaoxa-6,20-diazahentriacontan-1-oyl)piperazin-1-**

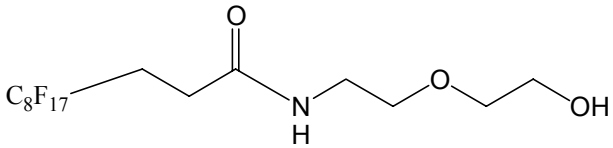
**yl)-3-oxopropyl)thio)-2,5-dioxopyrrolidin-1-yl)hexanoyl)oxy)-2,5-dioxopyrrolidine-3-sulfonate (Sulfo-NHS-$(OEG)_3$-perfluoroctane) (11)**

To a solution of **10** (10 mg, 24 mM, 1.0 eq.) in anhydrous DMF was added Sulfo-EMCS (25.5 mg, 25.6 mM, 1.05 eq) and the mixture was stirred at room temperature for 1hour and then for additional 3 hours at 40°C. Solvents were removed under reduced pressure and final product **11** was obtained as pale yellow waxy solid. $^1$H NMR (400 MHz, DMSO-$d_6$) δ 1.30 (q, J=7.2 Hz, 2H), 1.38 (s, 6H), 1.49 (q, J=7.2 Hz, 2H), 1.60 (q, J=7.2 Hz, 2H), 1.84 (t, J=7.6 Hz, 2H), 2.41 (m, 4H), 2.44 (t, J=7.2 Hz, 2H), 2.62 (t, J=6.8 Hz, 2H), 2.70 (m, 2H), 2.84 (m, 2H), 3.02 (m, 2H), 3.20 (q, J=6.0 Hz, 2H), 3.30~3.51 (m, 24H), 4.02 (m, 4H), 7.18 (t, J=5.6 Hz, 1H), 8.12 (t, J=5.6 Hz, 1H). Negative ion ESI-MS: $(M-H)^-$=1383.2.

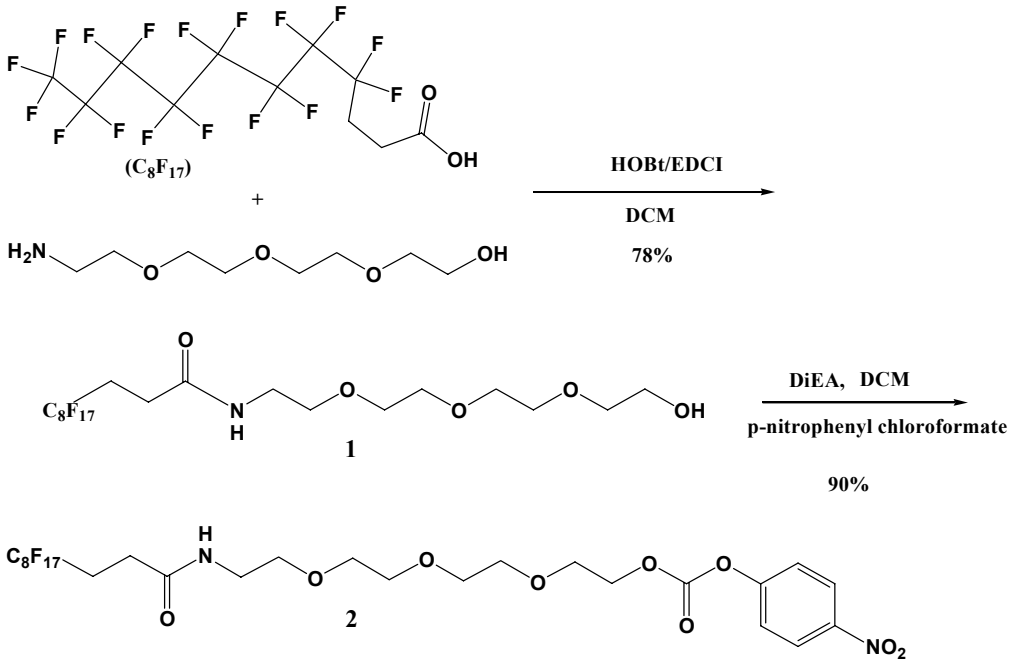
**3.4.5 Synthesis of sulfo-NHS-$(OEG)_1$-perfluoroctane** (B in Figure 3.1)

The compound sulfo-NHS-$(OEG)_1$-perfluoroctane was synthesized according to the procedures above for synthesizing its analogue except that $(OEG)_1$ was used instead of $(OEG)_3$. The chareacterizations of all intermediates from synthesizing the final product sulfo-NHS-$(OEG)_1$-perfluoroctane by 1HNMR and ESI-MS are listed below:

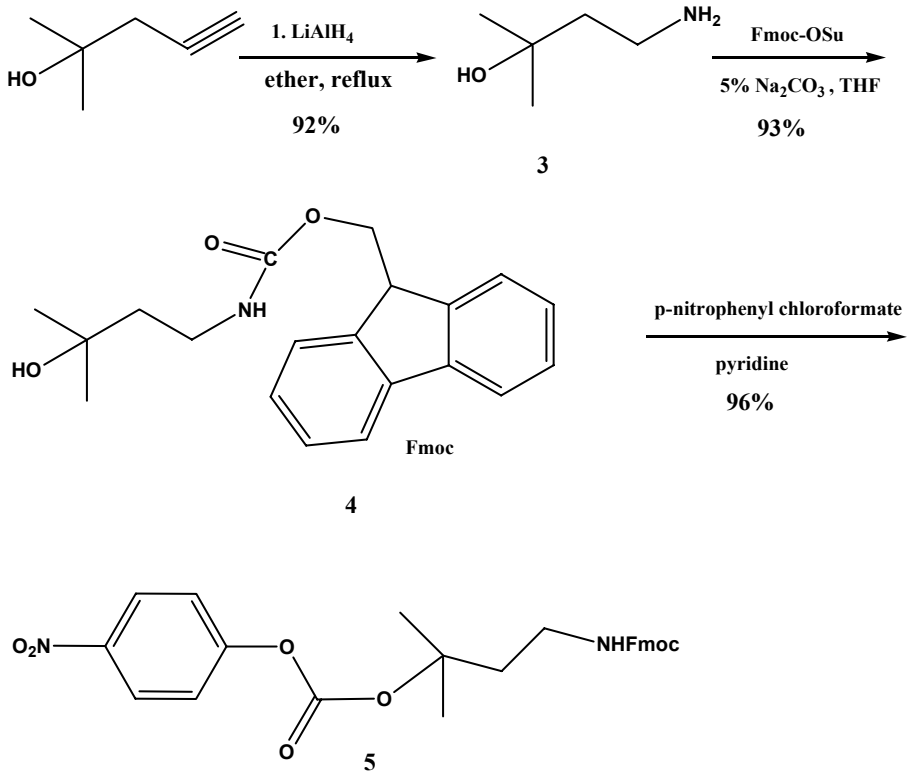
**4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-heptadecafluoro-N-(2-(2-hydroxyethoxy)ethyl)undecanamide**



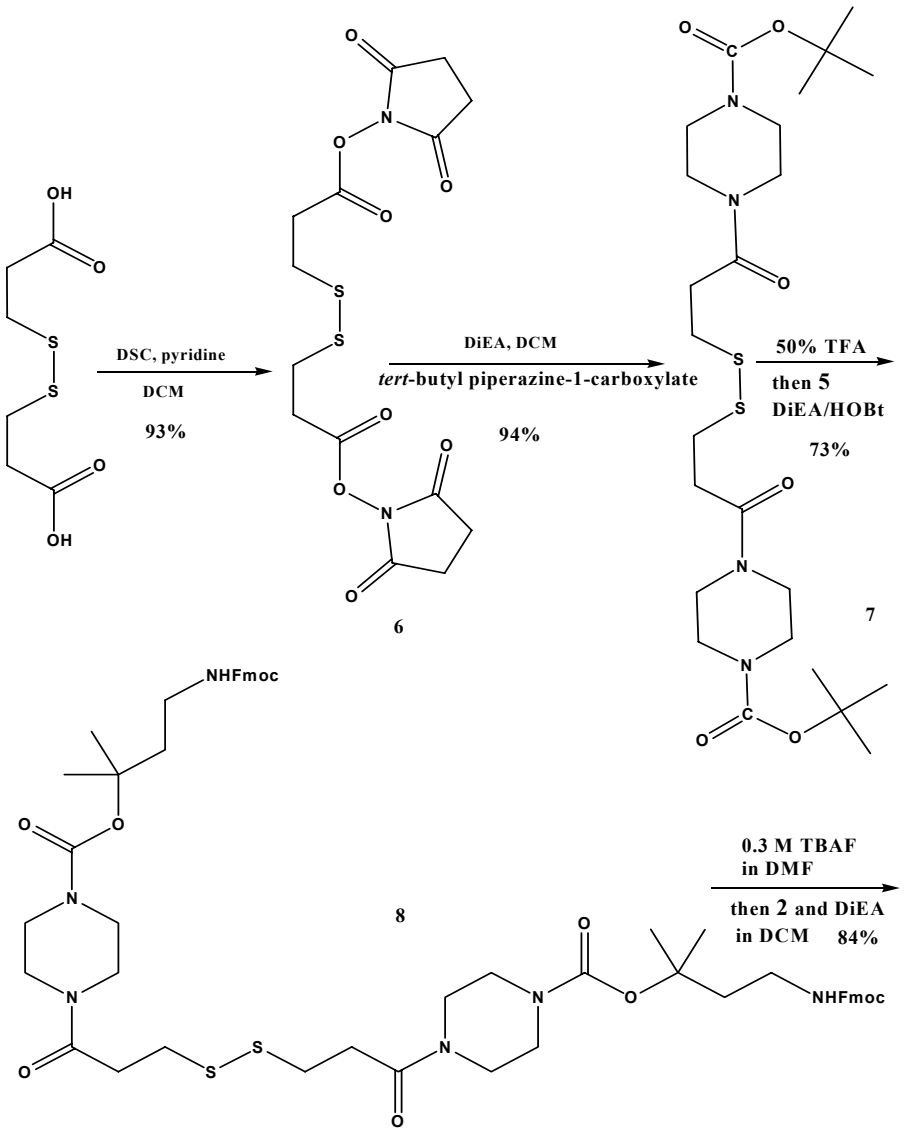

$^1$H NMR (400 MHz, $CDCl_3$) δ 2.49 (m, 2H), 2.56 (m, 2H), 3.50 (t, J=4.8 Hz, 2H), 3.58 (t, J=4.8 Hz, 2H), 3.59 (t, J=4.4 Hz, 2H), 3.76 (t, J=4.4 Hz, 2H), 6.21 (s, 1H). ESI-MS: $(M+H)^+$ =580.1.



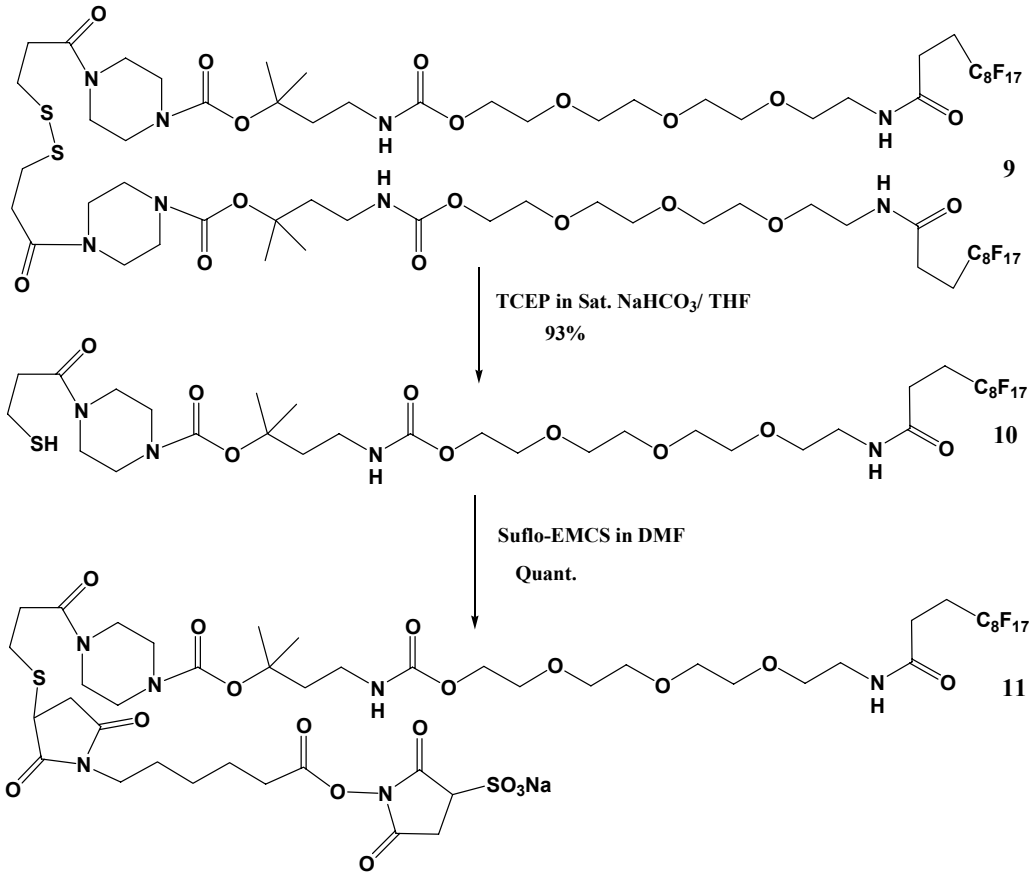

**Scheme 3.2.** Part 1: Synthesis of intermediates **1-2** for sulfo-NHS-$(OEG)_3$-perfluoroctane




**Scheme 3.3.** Part 2: Synthesis of intermediates **3-5** for sulfo-NHS-$(OEG)_3$-perfluoroctane

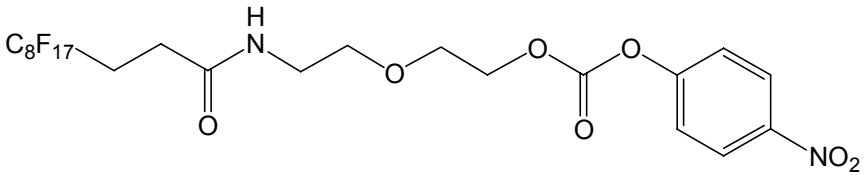



**Scheme 3.4.** Part 3: Synthesis of intermediates **6-8** for sulfo-NHS-$(OEG)_3$-perfluoroctane



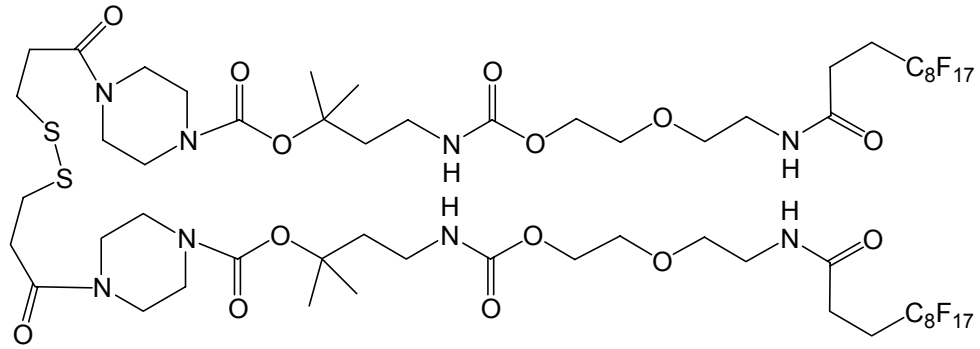

**Scheme 3.5.** Part 4: Synthesis of intermediates **9-11** for sulfo-NHS-$(OEG)_3$-perfluoroctane

**2-(2-(4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,11-heptadecafluoroundecanamido)ethoxy)ethyl (4-nitrophenyl) carbonate**



$^1$H NMR (400 MHz, $CDCl_3$) δ 2.49 (m, 2H), 2.56 (m, 2H), 3.55 (t, J=4.8 Hz, 2H), 3.62 (t, J=4.8 Hz, 2H), 3.78 (t, J=4.4 Hz, 2H), 4.45 (t, J=4.4 Hz, 2H), 5.99 (s, 1H), 7.38 (d, J=9.2 Hz, 2H), 8.28 (d, J=9.2 Hz, 2H). ESI-MS: $(M+H)^+$ =745.08.

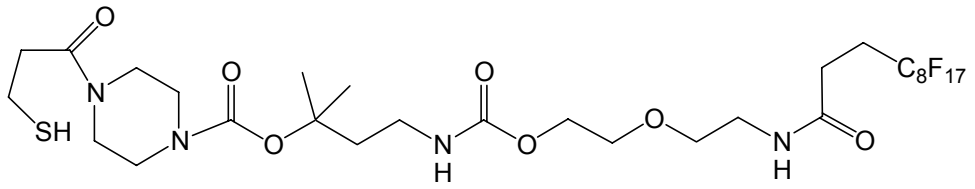
**17,17,18,18,19,19,20,20,21,21,22,22,23,23,24,24,24-heptadecafluoro-2-methyl-6,14-dioxo-7,10-dioxa-5,13-diazatetracosan-2-yl 4-(3-((3-(4-(18,18,19,19,20,20,21,21,22,22,23,23,24,24,25,25,25-heptadecafluoro-3,3-dimethyl-7,15-dioxo-2,8,11-trioxa-6,14-diazapentacosan-1-oyl)piperazin-1-yl)-3-oxopropyl)disulfanyl)propanoyl)piperazine-1-carboxylate**



$^1$H NMR (400 MHz, $CDCl_3$) δ 1.47 (s, 12H), 2.02 (t, J=7.2 Hz, 4H), 2.47 (m, 4H), 2.51 (m, 4H), 2.76(t, J=7.0 Hz, 4H), 2.95 (t, J=7.0 Hz, 4H), 3.27 (q, J=7.2 Hz, 4H), 3.45 (m, 16H), 3.56 (t,

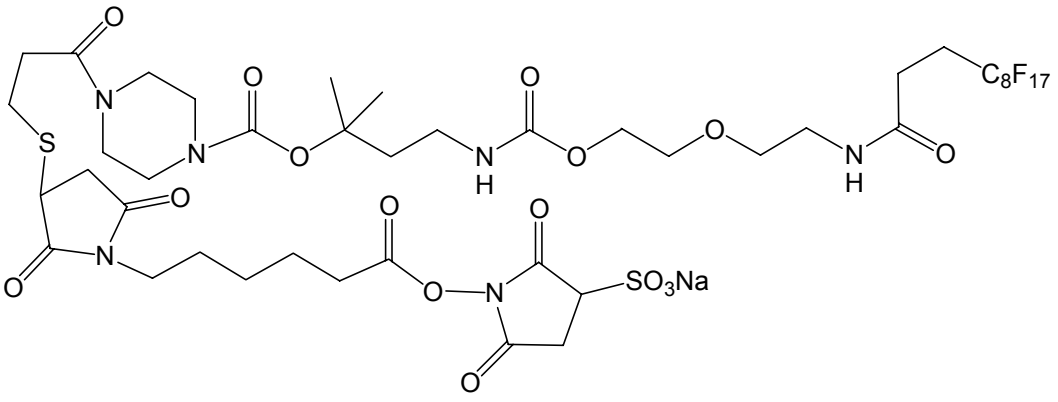
J=5.0 Hz, 4H), 3.60 (m, 8H), 4.20 (t, J=5.0 Hz, 4H), 4.88 (s, 2H), 6.50 (s, 2H). ESI-MS: $(M+H)^+$ =1815.4

**17,17,18,18,19,19,20,20,21,21,22,22,23,23,24,24,24-heptadecafluoro-2-methyl-6,14-dioxo-7,10-dioxa-5,13-diazatetracosan-2-yl 4-(3-mercaptopropanoyl)piperazine-1-carboxylate**



$^1$H NMR (400 MHz, $CDCl_3$) δ 1.47 (s, 6H), 1.73 (t, J=8.0 Hz, 1H), 2.02 (t, J=7.2 Hz, 2H), 2.49 (m, 2H), 2.52 (m, 2H), 2.66 (t, J=6.7 Hz, 2H), 2.82 (q, J=7.2 Hz, 2H), 3.29 (q, J=7.2 Hz, 2H), 3.43 (m, 8H), 3.56 (t, J=5.2 Hz, 2H), 3.63 (m, 12H), 4.21 (t, J=5.2 Hz, 2H), 4.83 (s, 1H), 6.50 (s, 1H). ESI-MS: $(M+H)^+$ =909.2

**sodium 1-((6-(3-((3-(4-(18,18,19,19,20,20,21,21,22,22,23,23,24,24,25,25,25-heptadecafluoro-3,3-dimethyl-7,15-dioxo-2,8,11-trioxa-6,14-diazapentacosan-1-oyl)piperazin-1-yl)-3-oxopropyl)thio)-2,5-dioxopyrrolidin-1-yl)hexanoyl)oxy)-2,5-dioxopyrrolidine-3-sulfonate**



Negative ion ESI-MS: $(M-H)^-$ =1295.3

### 3.4.6 Labeling ACTH (4-11) using sulfo-NHS-$(OEG)_3$-perfluoroctane

10 μL of 1mg/mL ACTH (4-11) was diluted in 120 μL PBS 8.0 and the solution was added to a vial containing 0.038 mg labeling reagent (Sulfo-NHS-$(OEG)_3$-perfluoroctane, structure see C in Figure 3.1). The labeling was proceeded at room temperature for 2 hours and then excessive labeling reagent was consumed by reacting with about 0.5 mg polymer-bound tris(2-aminoethyl)-amine for additional 1hour at room temperature. After centrifugation, the polymer beads were washed by 100 μL 80% MeOH and 100% MeOH twice, supernatants were combined. 1 μL of the collected supernatant was resuspended in 100 μL 2.5% FA in MeOH for direct ESI-MS analysis. The rest was vacuum dried and resuspended in 50 μL TFA for cleavage. The cleavage reaction was completed after 1 hour at 37° C. After removal of the TFA under reduced pressure, residues were dissolved in 100 μL 50% MeOH containing 2.5% FA for ESI-MS analysis.

### 3.4.7 Labeling BSA using sulfo-NHS-$(OEG)_3$-perfluoroctane and affinity purification of the labeled tryptic BSA peptides

10 μL of 10 mg/mL BSA stock solution was diluted in 35 μL PBS 8.0 buffer and the solution was transferred to a vial containing 0.26 mg fluorous reagent, followed by a brief vortex. The mixture was incubated for 2 hours at room temperature while shaking. The excessive labeling reagent was consumed by adding 15 μL of 50 mM tris-HCl for additional 1 hour mixing. The labeled BSA was then subjected to in solution trypsin digestion. Briefly, 80 μL of 8.0 M urea was added to the labeling mixture, followed by addition of 1.7 μL of 200 mM DTT in 50 mM

ammonium bicarbonate for 1hour reduction at 45°C. 6 μL of 200 mM iodoacetamidein 50 mM ammonium bicarbonate was added for 1hour alkylation at room temperature in dark, and the excessive iodoacetamide was consumed by adding 6 μL DTT for an additional 1 hour incubation at room temperature. The mixture was diluted by adding 1 mL of 50 mM ammonium biocarbonate to final before trypsin (6 μg) digestion at 37° C for 15 hours. The digestion was stopped by adding 5% FA till colorless precipitates occurred ( pH≈3.0, hydrolysis by-product of the labeling reagent.), which was then separated by centrifugation and further were washed by 200 μL 5% FA. Combined supernatant were evaporated to dryness and 1/6 sample was resuspended in 450 μL buffer A (20 mM ammonium formate, 0.1 M Acetic acid in 60% MeOH) for fluorous affinity purification. Tryptic BSA mixture was loaded onto cartridge packed by fluorous silica gel beads at a flow rate 10 μL/min using a syringe pump. After washing by buffer A (3 mL) and retained modified peptides were eluted by 2 mL buffer B (20 mM ammonium formate, 0.1 M acetic acid in 100% MeOH) at the flow rate of 10 μL/min. The eluate was dried under vacuum and resuspended in 95% TFA with 5% anisole as a scavenger for incubation at 37°C for 2 hours. After removal of solvents, cleaved peptides were resuspended in 50 μL of 5% FA, 2% ACN for ESI-LC-MS analysis.

### 3.4.8 LC-MS/MS analysis and database search

LC-ESI MS/MS was performed on a Finnigan LTQ$^{TM}$-ion trap mass spectrometer (Thermo Electron, San Jose, CA). Peptide sample (8 μl) was first loaded to a $C_{18}$ trap column and washed with 3% acetonitrile and 0.1% formic acid for 60 minutes for desalting, then the peptides were eluted to a reverse-phase $C_{18}$ analytical column by a 60 minutes gradient made of A buffer (0.1% formic acid: 97% water:3% acetonitrile v/v/v) and B buffer (0.1% formic acid: 3% acetonitrile:

97% water v/v/v) at a flow rate of 200 ~ 500 nL/min. Separated peptides were analyzed by the data-dependent acquisition mode set by Xcalibur, 2.2 version (Thermo Electron, San Jose, CA). After a MS survey scan over the m/z range of 300-2000, the seven most intensive precursors were selected and subjected to fragmentation by CID (collision induced dissociation). The normalized collision energy was set at 35% with activation Q value being 0.25 and dynamic exclusion of 100 seconds. The acquired raw data were processed by BioWorks software, version 3.3(Thermo Electron). Parameters for SEQUEST database search were set as follows: differential mass increase of 57.02 Da on cysteinyl, 15.99 Da on methionine residue, 367.02 Da (for Sulfo-NHS-$(OEG)_3$-perfluoroctane labeling) and 99.06 Da (for FBPNHS labeling) on lysine residue respectively. The number of missed cleavage sites was set to three. The search results were filtered by cross-correlation score (XCorr) which must be $\geq 2.0$ for singly charged ions, 2.5 for the doubly charged, 3.0 for the triply charged and by delta correlation score ($\Delta Cn$) $\geq 0.1$. MS/MS spectra for detected modified peptides were manually examined to ensure quality of identification.

## 3.5 Results and discussion

### 3.5.1 Synthesis of fluorous labeling reagent Boc-pentanoic-NHS-perfluoroctane (FBPNHS)

The design of our 1st fluorous labeling reagent included fluorous tag ($C_8F_{17}$) moiety for affinity enrichment, a reactive group (HOSu derived active ester, NHS ester) targeting free amine groups and a TFA labile carbamate linker. The NHS ester was chosen as the reactive group because of its well known ability to react with $NH_2$ groups in peptides and proteins in pH=7-9 buffer. The carbamate group was selected as the cleavable linker based on robust chemistry and easy incorporation into the reagent. Moreover, a relatively long spacer (5 carbon backbone) would

reduce the steric hindrance around the NHS ester. The condition that favors the linker cleavage was similar to TFA cleavage of the Boc protected amino acids, which has been routinely employed in peptide synthesis. The compound was synthesized following the two steps shown in Scheme 3.1 with an overall 78% yield.

The first step – production of fluorous tagged Boc group, was relatively easy with high yield. The second step however, was not as straightforward. A variety of reaction conditions such as using coupling reagents (e.g. DCC, EDC) in the presence of organic bases (e.g. pyridine, DMAP and DiPEA) had been explored, but none of those approaches resulted in acceptable condensation efficiency as monitored by TLC and mass spectrometry. The efficiency was significantly enhanced when NHS was replaced by DSC, an activated derivative of of HOSu, which had been utilized in the preparation of active ester (15) and peptide synthesis. Presumably, the NHS anions are more easily generated from the attack of carboxylic groups to the carbonyl in DSC, which was proven to be more reactive toward deprotonated carboxyl group in the presence of pyridine.

### 3.5.2 Labeling peptides and BSA using FBPNHS

As proof of principle experiments, the newly synthesized reagent was tested in the labeling of some small peptides. After about 30 minutes labeling reaction, a small peptide MRFA gave a new peak at m/z: 1155.3, which corresponded to the N-terminally labeled MRFA, the unconsumed MRFA at m/z: 524.4 also appeared in the same mass spectrum (Figure 3.2). Increasing the amount of the labeling reagent resulted in complete conversion of MRFA to its labeled form (mass spectral data not shown). It is interesting that the cleavage of carbamate

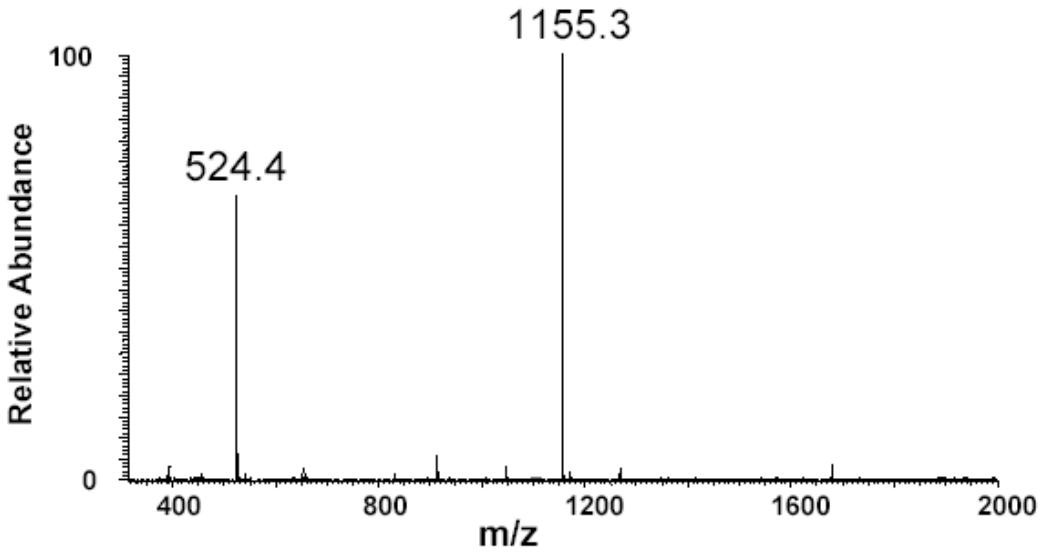



**Figure 3.2.** MS spectrum of a modified small peptide MRFA by Boc-pentanoic-NHS-perfluoroctane after one hour reaction at room temperature. The ion at m/z 524.4 corresponds to the unconsumed MRFA and the ion at m/z 1155.3 represents N-terminal modified MRFA.

linker was preferred to peptide bond cleavage upon CID (collision induced dissociation). It is also noteworthy that no fragmentation of fluorous tag was observed under the CID conditions that were optimized for peptide fragmentation; this phenomenon confirmed the inertness of the fluorous tag in MS analysis (2).

Longer peptide ACTH (4-11) was also tested and trifluoroacetic acid (TFA) was employed to cleave the cabamate linker in the labeled peptides. Results indicated that both the N-terminal and the lysine side chain of the peptides were modified as evidenced
by the appearance of two ions after the cleavage at m/z: 1189.6 (singly tagged) and 1288.7 (doubly tagged) (Figure 3.3). Its MS/MS analysis (Figure 3.4) revealed that
C-terminal lysine residue was preferentially modified over the N-terminus in ACTH (4-11) possibly because of free amines from the lysine side chain were more accessible to the labeling reagent than the ones at N-terminus.

The labeled BSA was digested by trypsin, and the tryptic peptide mixture was treated by ICAT cleavage reagents to remove the bulky fluorous label. HPLC-MS analysis followed by database search against the amino acid sequence of BSA resulted in the identification of 33 peptides among which 11 peptides were labeled by FBPNHS. The FBPNHS labeled peptides covered nearly 26% of the BSA sequence (highlighted in blue in Figure 3.5).
Moreover, a small moiety after the cleavage does not appear to interfere with the peptide backbone fragmentation pattern without producing its visible fragments in MS/MS spectrum as exemplified in Figure 3.6.

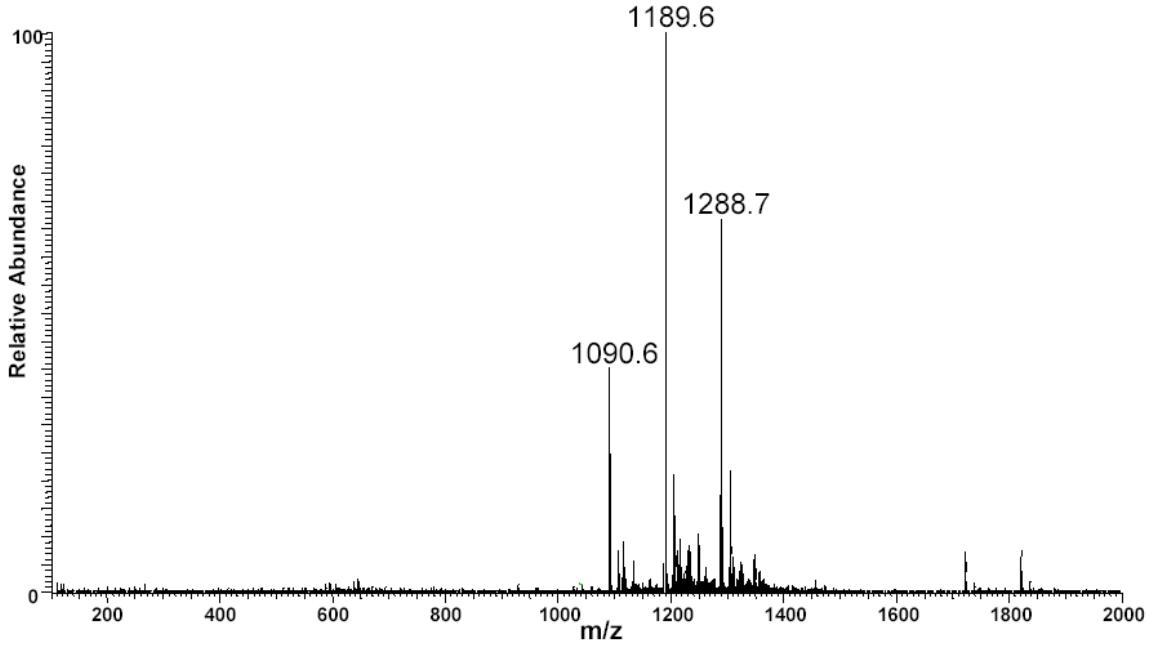



**Figure 3.3.** MS analysis a modified peptide ACTH (4-11) by Boc-pentanoic-NHS-perfluoroctane after cleaving the fluorous tag. 1090.6 (m/z) indicates the unconsumed peptide ACTH (4-11), 1189.6 (m/z) represents the singly tagged ACTH (4-11) and 1288.7524.4 (m/z) corresponds to the doubly tagged ACTH (4-11).

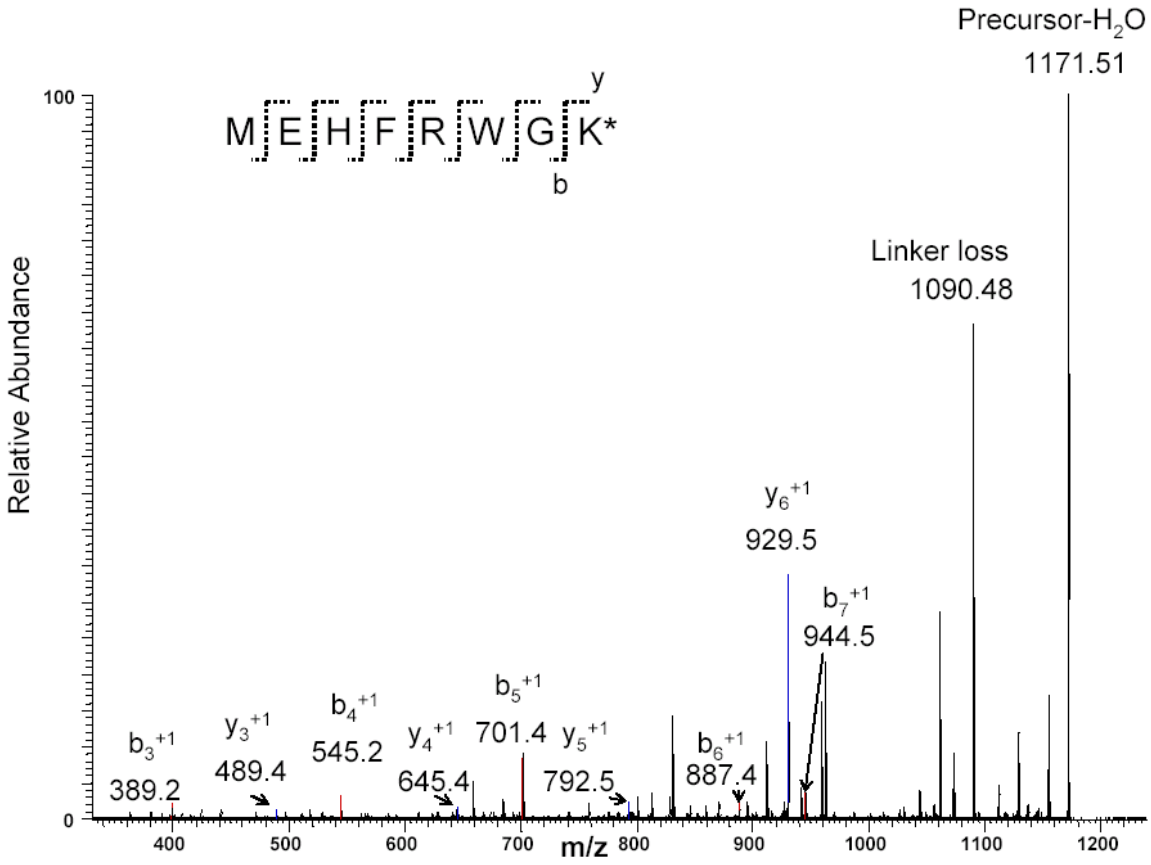



**Figure 3.4.** $MS^2$ analysis of singly tagged ACTH (4-11) peptide precursor ion at m/z 1189.6. Series of b ions and y ions could be used to confirm the peptide's sequence.

mkwvtfisllllfssaysrgvfrrdthkseiahrfkdlgeehfkglvliaf
sqylqqcpfdehvklvneltefaktcvadeshagcekslhtlfgdelckva
slretygdmadccekqepernecflshkddspdlpklkpdpntlcdefkad
ekkfwgkylyeiarrhpyfyapellyyankyngvfqeccqaedkgacllpk
ietmrekvlassarqrlrcasiqkfgeralkawsvarlsqkfpkaefvevt
klvtdltkvhkecchgdllecaddradlakyicdnqdtissklkeccdkpl
lekshciaevekdaipenlppltadfaedkdvcknyqeakdaflgsflyey
srrhpeyavsvllrlakeyeatleeccakddphacystvfdklkhlvdepq
nlikqncdqfeklgeygfqnalivrytrkvpqvstptlvevsrslgkvgtr
cctkpesermpctedylslilnrlcvlhektpvsekvtkcckteslvnrrpc
fsaltpdetyvpkafdeklftfhadictlpdtekqikkqtalvellkhkpk
ateeqlktvmenfvafvdkccaaddkeacfavegpklvvstqtala

**Figure 3.5.** Peptides (highlighted in blue) modified by Boc-pentanoic-NHS-perfluoroctane in BSA. The fluorous tag moiety was cleaved before LC-MS analysis.

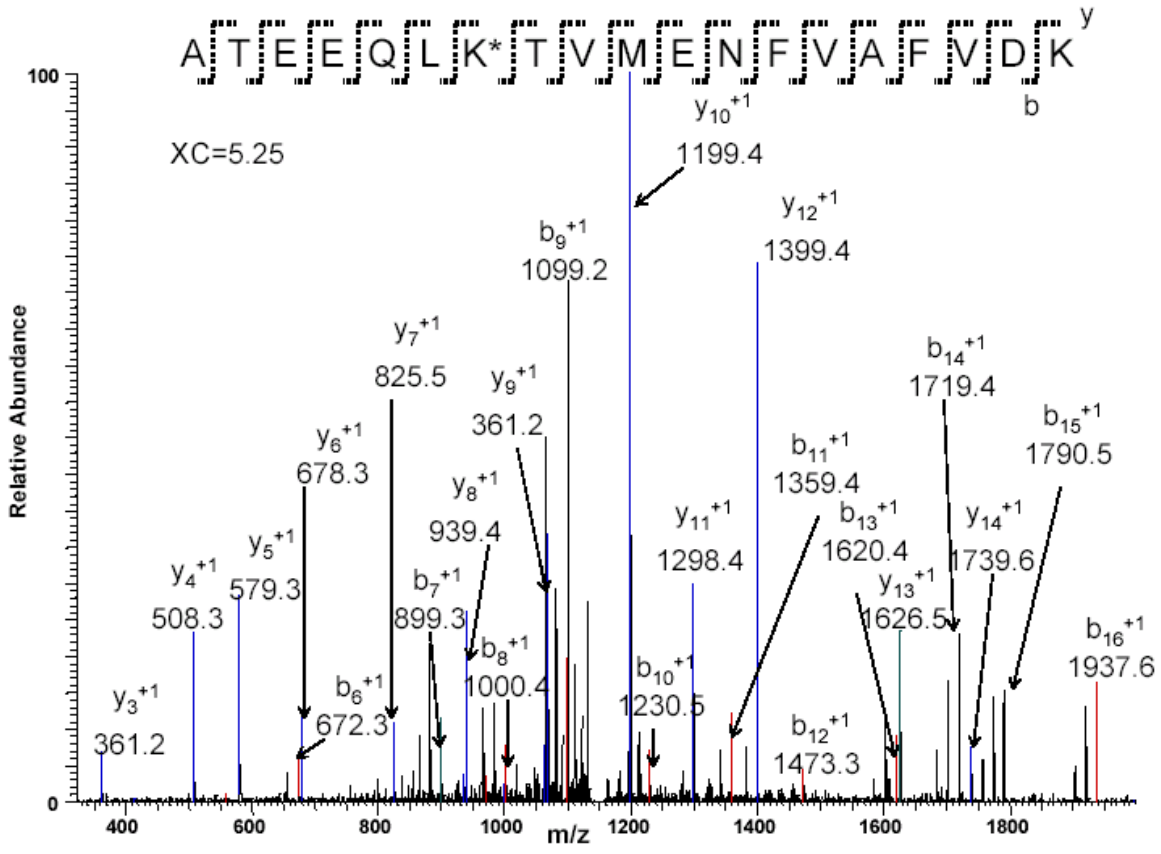



**Figure 3.6.** $MS^2$ spectrum of a BSA tryptic peptide modified by Boc-pentanoic-NHS-perfluoroctane. The bulky fluorous tag moiety was cleaved before LC-MS analysis.

### 3.5.3 Retrosynthetic analysis of 2nd generation fluorous labeling reagent sulfo-NHS-$(OEG)_3$-perfluoroctane

The novel fluorous labeling reagent, sulfo-NHS-$(OEG)_3$-perfluoroctane can be divided to five building blocks (a in Figure 3.7). The synthesis of this labeling reagent was based on a retrosynthetic analysis (b in Figure 3.7). The target product, sulfo-NHS-$(OEG)_3$-perfluoroctane, can be obtained through a highly efficient coupling reaction between 1 (Sulfo-EMCS, block E in Figure 3.7 a) and a thiol compound **2** (block A to D) (in Figure 3.7 a), which could be derived from the cleavage of a disulfide bond in its dimmer, compound **3**. Compound **3** could be formed through a carbamate formation between 4

and carbonate compound **5** after removal of the Fmoc group with TBAF in DMF. The acid labile linker in compound **4** could be constructed through carbonate compound **7** (10) and piperazine derivative **6**, which be easily prepared from dithiobis[succinimidylpropionate] **9** and 1-Boc-piperazine **10**. Compound **5** could be prepared from hydroxyl compound **8** by reacting with p-nitrophenylchloroformate. The $(OEG)_3$ spacer in compound **8** could be introduced by the incorporation of $(OEG)_3$-mono-amine **12** to fluorous compound **11** via an amide bond formation. The fluorous labeling reagent was synthesized via a 10-step reaction with an overall yield of 25%.

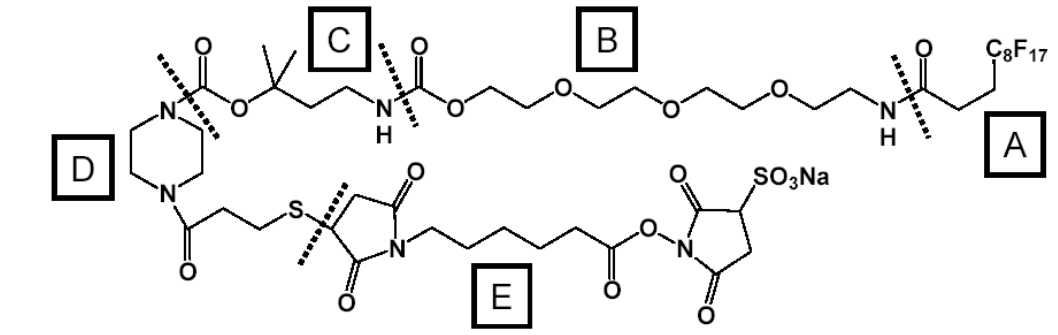
### 3.5.4 Characterization of sulfo-NHS-$(OEG)_3$-perfluoroctane by Negative ESI-MS

The newly synthesized sulfo-NHS-$(OEG)_3$-perfluoroctane was characterized by negative electrospray mass spectrometry and $^1H$ NMR (see detailed data in experimental), Dissolved in 50% MeOH, the labeling reagent could be detected in negative ion mass spectrometry as shown in Figure 3.8. Upon collision induced dissociation (CID), the molecular ion [Sulfo-NHS-$(OEG)_3$-perfluoroctane]$^-$ at m/z 1383.30 yielded to two major product ions at m/z 605.0 and m/z 560.90 (Figure 3.8) corresponding to the product ion (Figure 3.8) from the cleavage at the carbamate

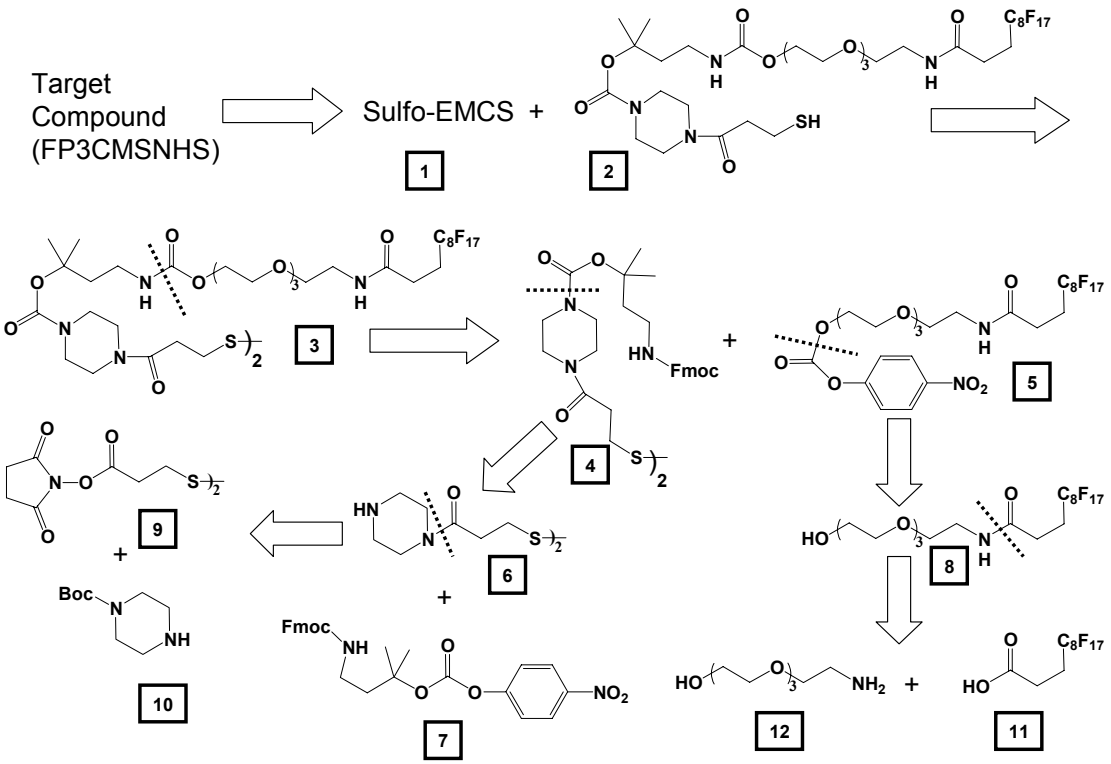
bond in block C (Figure 3.7) and subsequent loss of a $CO_2$ (block D and E in a Figure 3.7) respectively. Further fragmentation of product ion at m/z 605.0 yielded ion at m/z 560.90 (Figure 3.8) after the neutral loss of a $CO_2$. The product ion at m/z 560.90 in the $MS^2$ spectrum of m/z 1383.30 and the m/z 560.90 in the $MS^3$ spectrum of m/z 605.0 are the same, both ions yielded m/z 387.01 (spectrum not shown) upon CID, corresponding to the formation of Sulfo-EMCS as the result of thiol ether bond cleavage (block E in Figure 3.7 a). These mass spectral evidences support the structure of newly synthesized sulfo-NHS-$(OEG)_3$-perfluoroctane as shown in Figure 3.1 C.

### 3.5.5 $MS^2$ analysis of ACTH (4-11) labeled by sulfo-NHS-$(OEG)_3$-perfluoroctane

Sulfo-NHS-$(OEG)_3$-perfluoroctane was employed to label a small peptide ACTH(4-11) (NH2-M-E-H-F-R-W-G-K) in PBS buffer at room temperature. In a 1 hour reaction, the free amine groups (N-terminal and lysine side chain) in ACTH (4-11) were almost completely tagged by the labeling reagent (the molecular ion of intact ACTH disappeared and doubly labeled ACTH became a major ion on the mass spectrum of the reaction mixture, data not shown). The bulky part (correspond to building blocks A + B + C in Figure 3.7 a) of the labeling reagent was removed from the labeled ACTH peptide upon treatment with TFA as expected, leaving a small tag of 367.17 Da (net mass increase to the amine group in the peptide by building blocks D + E after losing the sulfo-NHS group in Figure 3.7 a) remained covalently linked to the N-terminal and lysine side chain amine of the peptide. The positive ion ESI MS/MS spectrum of the labeled




**a**




**b**

**Figure 3.7.** Retrosynthetic analysis of sulfo-NHS-$(OEG)_3$-perfluoroctane. a. Structure of synthesized the fluorous labeling reagent with indications of its building blocks (A to E). b. Retrosynthetic steps for the construction of sulfo-NHS-$(OEG)_3$-perfluoroctane.

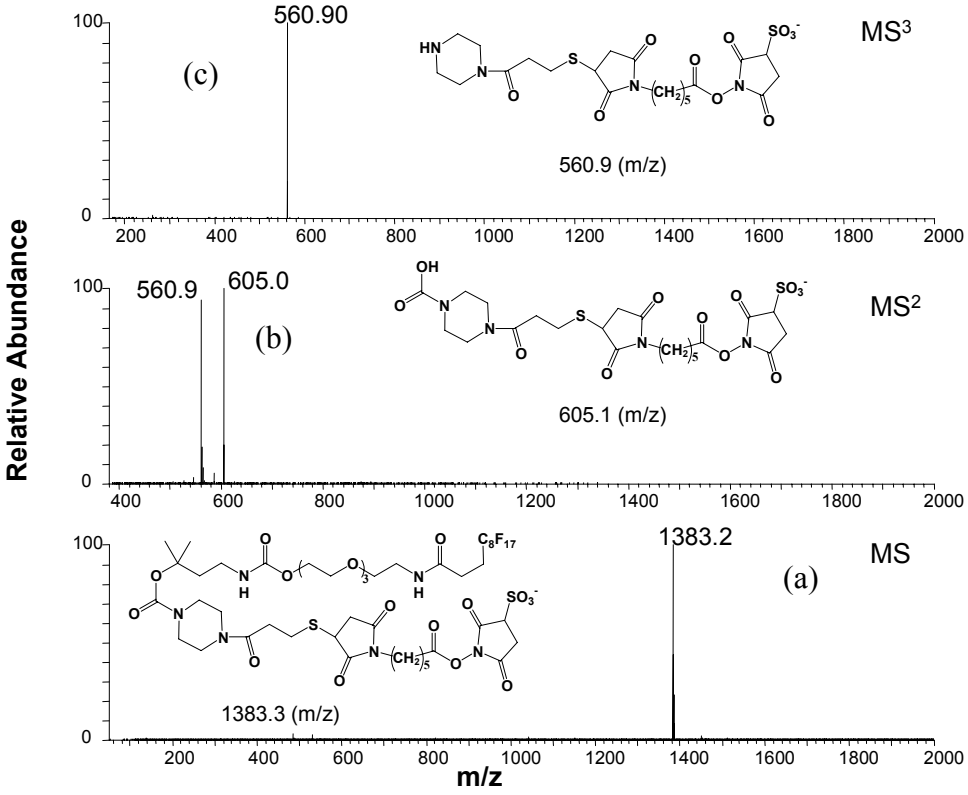



**Figure 3.8.** Negative ion ESI-MS of sulfo-NHS-$(OEG)_3$-perfluoroctane and its fragmentation pattern. (a) MS spectrum of the labeling reagent showing expected molecular weight after deprotonated molecular ion. (b) $MS^2$ spectrum of precursor ion at m/z 1383.3. The product ions at m/z 605.0 corresponds to the cleavage of precursor ion at the tertiary carbamate linkage (TFA labile site as well), and ion at m/z 560.9 corresponds to the neutral loss of $CO_2$ from m/z 605.0 (c) $MS^3$ spectrum of precursor ion at m/z 605.0 showing its further loss of the $CO_2$ to yield ion at m/z 560.9.

ACTH peptide (after TFA cleavage) is shown in Figure 3.9. Most of the b ions and y ions can be observed. Interestingly, a fragment ion at m/z 368.1 which corresponds to the residual of sulfo-NHS-$(OEG)_3$-perfluoroctane cleaved from the amide bonds formed during the labeling, also appeared on the $MS^2$ spectrum and could serve as a marker of the labeling and fluorous tag moiety removal processes. The newly synthesized fluorous labeling reagent indeed efficiently modified the free amine groups in peptide. The residual of sulfo-NHS-$(OEG)_3$-perfluoroctane after acid cleavage did not interfere with the fragmentation of the labeled peptide.

### 3.5.6 $MS^2$ analysis of BSA peptides labeled by sulfo-NHS-$(OEG)_3$-perfluoroctane

To further demonstrate the potential of the Sulfo-NHS-$(OEG)_3$-perfluoroctane in protein/peptide modification, it was employed to label BSA in PBS buffer. No precipitation was observed throughout the labeling reaction, which indicated that the labeling can be performed in aqueous solution without introducing any organic solvent to the reaction mixture. After the trypsin digestion, the tagged peptides were easily enriched through fluorous affinity purification. After TFA cleavage as mentioned above, the tagged peptides were analyzed by LC-MS/MS. About 95% of the identified peptides were tagged (see Table 3.1). This indicates that affinity purification based on fluorine-fluorine interaction was indeed highly effective in enriching labeled peptides with low nonspecific binding. The sulfo-NHS-$(OEG)_3$-perfluoroctane tagged peptides had longer retention time on reversed-phase $C_{18}$ column due to the increased hydrophobicity introduced by the fluorous alkyl chain. Once the fluorous tag moiety was removed, the tagged peptide behaved in the same way as its unmodified counterpart in reversed-

phased LC separation. Fifteen out of sixteen tryptic peptides (see Table 3.1) were identified as tagged peptides in a single LC-MS/MS experiment followed by database search against BSA sequence using SEQUEST algorithm. As shown in Figure 3.10, the residual of fluorous tag had little influence on the fragmentation of the labeled BSA peptide. In addition to the b ions and y ions from peptides fragmentation, a characteristic ion at m/z 368.1 appeared on the spectrum confirmed that this peptide was labeled by the fluorous reagent.

## 3.6 Conclusion

A serial of fluorous labeling reagents were synthesized with aim to effectively react with free amine groups on peptides and proteins. The first labeling reagent Boc-pentanoic-NHS-perfluoroctane (FBPNHS) was synthesized through 2 steps with an overall 78% yield. The reactivity of FBPNHS with small peptides (MRFA and ACTH) was tested in water/organic mixtures under slight basic conditions. ESI-MS analysis of the product mixture confirmed the labeling reaction and $MS^2$ spectrum of the labeled peptide showed that it was normally fragmented after removal of the fluorous tag. As one of our novel $2^{nd}$ generation fluorous labeling reagents, sulfo-NHS-$(OEG)_3$-perfluoroctane was also synthesized through 11 steps of simple chemical reactions with a high overall yield of 23%. Sulfo-NHS-$(OEG)_3$-perfluoroctane also reacts with free amine groups from peptide/protein through an active succinimidyl ester group. Notably, the labeling reaction proceeded efficiently in PBS buffer without the need of any extra organic solvent in the reaction mixture. More importantly, a homogeneous solution of reaction mixture was consistently observed throughout the labeling process, which indicated that the fluorous labeling reagent was not only soluble in water, but also didn't cause any aggregation of the labeled proteins due to the incorporation of a hydrophilic $(OEG)_3$ linker. Fluorous affinity

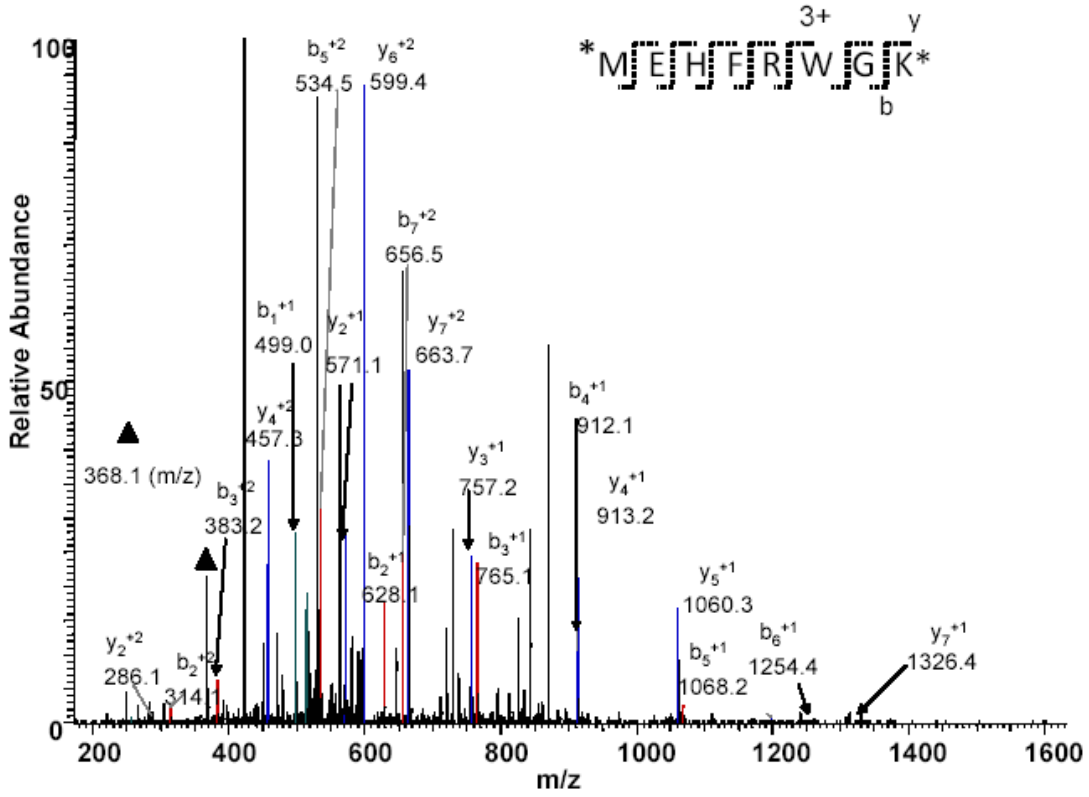



**Figure 3.9.** Positive ion $MS^2$ of peptide ACTH(4-11) tagged by the labeling reagent sulfo-NHS-$(OEG)_3$-perfluoroctane after TFA cleavage. The peptide was labeled at the N-terminal and the lysine side chain. ▲ unique fragment ion at 368.2 m/z from the residual of the cleaved tag.

**Table 3.1.** Identified tryptic peptides derived from sulfo-NHS-$(OEG)_3$-perfluoroctane labeled BSA*

| Detected peptides from elution | Xc |
|---|---|
| K.NYQEAK*DAFLGSFLYEYSR.R | 6.22 |
| R.ADLAK*YIC@DNQDTISSK.L | 6.14 |
| K.GLVLIAFSQYLQQC@PFDEHVK.L | 5.80 |
| R.ADLAK*YIC@DNQDTISSK*LK.E | 5.55 |
| R.LAK*EYEATLEEC@C@AK.D | 5.16 |
| K.EAC@FAVEGPK*LVVSTQTALA | 4.86 |
| K.FWGK*YLYEIAR.R | 4.43 |
| R.K*VPQVSTPTLVEVSR.S | 4.32 |
| K.SLHTLFGDELC@K*VASLR.E | 4.28 |
| R.ETYGDMADC@C@EK*QEPER.N | 4.17 |
| K.VTK*C@C@TESLVNR.R | 3.79 |
| R.ETYGDMADC@CEK*QEPER.N | 3.67 |
| R.ETYGDM#ADC@CEK*QEPER.N | 3.56 |
| R.ALK*AWSVAR.L | 3.39 |
| R.C@ASIQK*FGER.A | 3.02 |
| R.C@C@TK*PESER.M | 2.64 |

*Bulky perfluoroctane moiety and OEG linker of the labeling reagent was removed by acid cleavage, which resulted in a net mass increase of 367.12 Da on lysine residue. * lysine residue modified by fluorous affinity tag, #methionine was oxidized, and @ cysteine was modified by iodoacetamide.

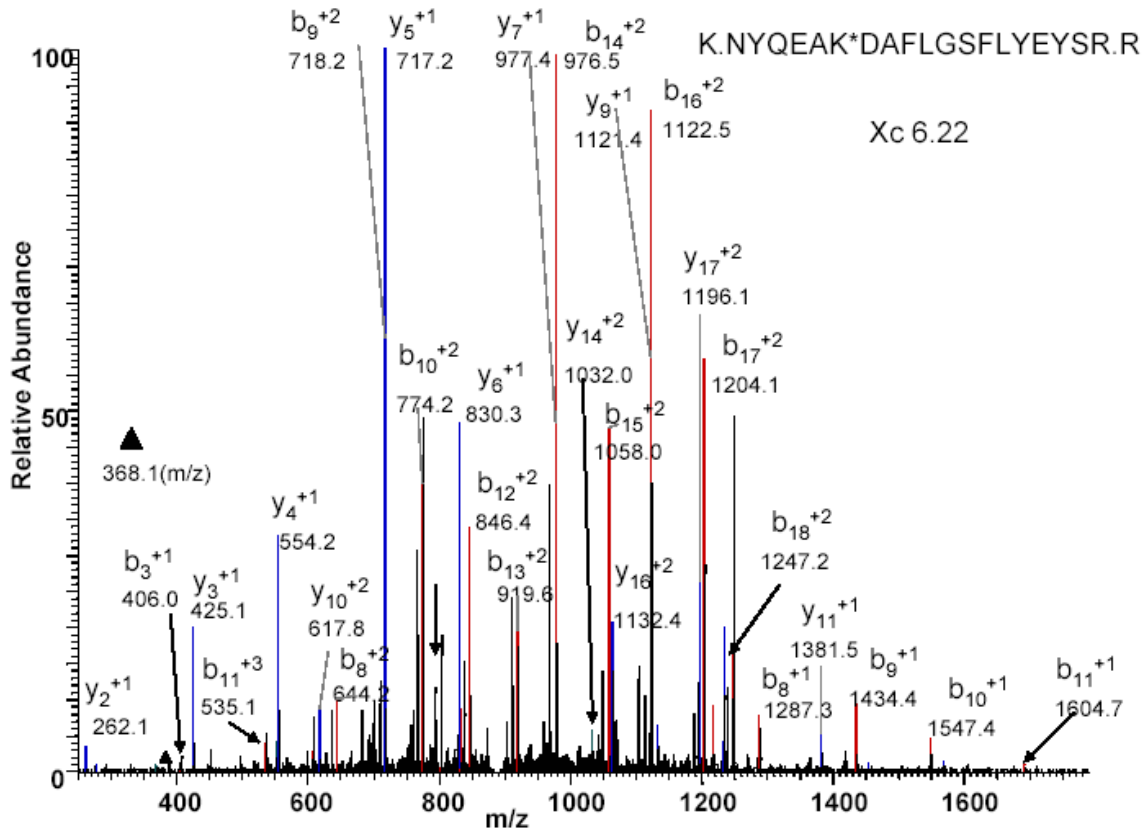



**Figure 3.10.** Positive ion $MS^2$ of a typical BSA tryptic peptide tagged by the labeling reagent sulfo-NHS-$(OEG)_3$-perfluoroctane at the lysine residue

purification was performed to isolate the tagged tryptic peptides of BSA, majority of peptides eluted from the fluorous affinity column were identified by LC-MS analysis as labeled peptides. Once the bulky fluorous tag moiety and the OEG linker were removed by acid cleavage, the labeled peptides (with a small residue from the fluorous labeling reagent) behaved similar to those unlabeled peptide in terms of fragmentation pattern. All these results indicate that labeling proteins/peptides by the newly synthesized sulfo-NHS-$(OEG)_3$-perfluoroctane followed by affinity separation/enrichment can by incorporated into most of LC-MS based proteomic approaches.

## References:


1. **Adamczyk, M., J. C. Gebler, and J. Wu.** 2001. Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry. Rapid Communications in Mass Spectrometry **15:**1481-1488.
2. **Brittain, S. M., A. Brock, D. E. Mason, S. B. Ficarro, and E. C. Peters.** 2005. Proteome characterization using fluorous affinity enrichment and mass spectrometry. Biopolymers **80:**508-508.
3. **Brittain, S. M., S. B. Ficarro, A. Brock, and E. C. Peters.** 2005. Enrichment and analysis of peptide subsets using fluorous affinity tags and mass spectrometry. Nature Biotechnology **23:**463-468.
4. **Cronan, J. E.** 1990. Biotination of proteins *in vivo* - A posttranslational modification to label, purify, and study proteins. Journal of Biological Chemistry **265:**10327-10333.
5. **Curran, D. P., and Z. Y. Luo.** 1999. Fluorous synthesis with fewer fluorines (light fluorous synthesis): separation of tagged from untagged products by solid-phase extraction with fluorous reverse-phase silica gel. Journal of the American Chemical Society **121:**9069-9072.
6. **Go, E. P., W. Uritboonthai, J. V. Apon, S. A. Trauger, A. Nordstrom, G. O'Maille, S. M. Brittain, E. C. Peters, and G. Siuzdak.** 2007. Selective metabolite and peptide capture/mass detection using fluorous affinity tags. Journal of Proteome Research **6:**1492-1499.
7. **Gygi, S. P., B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold.** 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnology **17:**994-999.
8. **Horvath, I. T., and J. Rabai.** 1994. Facile catalyst separation without water - Fluorous biphase hydroformylation of olefins. Science **266:**72-75.

9. **Li, J. X., H. Steen, and S. P. Gygi.** 2003. Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents - The yeast salinity stress response. Molecular & Cellular Proteomics **2:**1198-1204.
10. **Li, S. W., and D. X. Zeng.** 2007. CILAT - a new reagent for quantitative proteomics. Chemical Communications**:**2181-2183.
11. **Masuoka, J., L. N. Guthrie, and K. C. Hazen.** 2002. Complications in cell-surface labelling by biotinylation of *Candida albicans* due to avidin conjugate binding to cell-wall proteins. Microbiology-Sgm **148:**1073-1079.
12. **Mrsa, V., T. Seidl, M. Gentzsch, and W. Tanner.** 1997. Specific labelling of cell wall proteins by biotinylation. Identification of four covalently linked O-mannosylated proteins of *Saccharomyces cerevisiae*. Yeast **13:**1145-1154.
13. **Nicholson, R. L., M. L. Ladlow, and D. R. Spring.** 2007. Fluorous tagged small molecule microarrays. Chemical Communications**:**3906-3908.
14. **Oda, Y., T. Nagasu, and B. T. Chait.** 2001. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nature Biotechnology **19:**379-382.
15. **Ogura, H., T. Kobayashi, K. Shimizu, K. Kawabe, and K. Takeda.** 1979. Novel active ester synthesis reagent (N,N'-Disuccinimidyl Carbonate). Tetrahedron Letters**:**4745-4746.
16. **Pohl, N. L.** 2006. Fluorous-based microarrays for probing carbohydrate-protein recognition. Abstracts of Papers of the American Chemical Society **231**.
17. **Pohl, N. L.** 2008. Fluorous tags catching on microarrays. Angewandte Chemie-International Edition **47:**3868-3870.
18. **Staros, J. V.** 1982. N-hydroxysulfosuccinimide active esters - Bis(N-Hydroxysulfosuccinimide) esters of 2 dicarboxylic-acids are Hydrophilic, membrane-impermeant, protein cross-linkers. Biochemistry **21:**3950-3955.
19. **Steen, H., and M. Mann.** 2002. A new derivatization strategy for the analysis of phosphopeptides by precursor ion scanning in positive ion mode. Journal of the American Society for Mass Spectrometry **13:**996-1003.
20. **Urban, C., K. Sohn, F. Lottspeich, H. Brunner, and S. Rupp.** 2003. Identification of cell surface determinants in *Candida albicans* reveals Tsa1p, a protein differentially localized in the cell. Febs Letters **544:**228-235.
21. **Wilchek, M., and E. A. Bayer.** 1999. Foreword and introduction to the book (strept)avidin-biotin system. Biomolecular Engineering **16:**1-4.
22. **Wu, W. W., G. H. Wang, S. J. Baek, and R. F. Shen.** 2006. Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. Journal of Proteome Research **5:**651-658.
23. **Zhang, W., and D. P. Curran.** 2006. Synthetic applications of fluorous solid-phase extraction (F-SPE). Tetrahedron **62:**11837-11865.

# CHAPTER 4

# Surface Exposed Peptide Segments of *Candida albicans* Cell Wall Proteome Characterized by *in vitro* and *in vivo* Chemical Labelings and LC-MS

## 4.1 Abstract

The fungal pathogen *Candida albicans* is one of the major causes of candidiasis which results in high mortality among various immunocompromised hosts (such as AIDS patients). Cell wall proteins (CWPs) play essential functions in this pathogen's colonization, invasion and induction of immune response from hosts. Information on the dynamic regulations of CWPs during the morphological transition from yeast to hyphae provides better understanding on the mechanism of pathogen-host interaction. Cell surface biotinylations at low temperature (4°C), enzymatic digestion of the intact fungal cell surface proteins ("whole cell shaving"), biotin-avidin affinity enrichment of biotinylated peptides, liquid chromatography mass spectrometry (LC-MS) based proteomic approach were employed for unambiguous identification of cell wall/cell wall associated proteins and the exposed peptide segments of these proteins. SILAC (Stable Isotope Labeling by Amino acids in Cell Culture) based CWP quantification analyses were performed to monitor CWP accumulation level change in response to germination. 148 CWPs of yeast form of *C. albicans* were identified based on stringent criteria through intact cell labeling by three biotinylation labeling reagents. Characteristic fragments originated from biotin tag on tandem mass spectra of labeled peptides were also employed to verify that the peptides were indeed labeled. Notably, fifty one proteins found in our study were not reported in the literature as CWPs or cell wall associated proteins. Metabolic labeling of proteins with isotopic coded lysine

during yeast germination followed by and SDS-DTT (sodium dodecyl sulfate-dithiolthreitol) extraction allowed us to quantify one hundred and eighteen cell wall-associated proteins, among which ninety two proteins were up-regulated and twenty four proteins were down-regulated with one yeast specific protein and one hyphae specific protein. Surface exposed peptide segments of CWPs determined in this study could be new candidates to the pool of potential peptide targets for protective vaccine development.

## 4.2 Introduction

The cell surface of *C. albicans* is the frontier where initial interaction takes place between the fungus and the host. Cell surface proteins are likely those members in cell wall matrix to initiate fungal adhesion, colonization, pathogenic invasion and therefore modulation of immunological response from the host (68). Studying the surface proteome of *C. albicans* is essential for gaining a further understanding on the mechanism of the host-pathogen interactions. A paradigmatic approach was established to profile the surface proteins in fugal species using *Saccharomyces cerevisiae* as a model (93), featuring protein extraction via an array of chemical/enzymatic means, protein separation by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and protein detection by mass spectrometry (MS) (66). With the recent advancement in mass spectrometry based proteomic techniques, especially the coupling of MS and SDS-PAGE and/or liquid chromatography (LC), characterization of fungal cell surface proteins has reached an unprecedented level both qualitatively and quantitatively. For instance, non-gel based methods were complementary to 2D-SDS-PAGE, which was not efficient in glycoprotein separation (31, 92). A "whole cell-shaving" technique using proteolytic enzymes (usually trypsin) was employed to release peptides from surface accessible proteins of *Cryptococcus neoformans* (22) without the need of laborious protein extraction procedures. Moreover, labeling the whole fungal cell surface proteins by membrane impermeable biotinylation reagent (87) offered an alternative approach to differentiating surface proteins from cytosolic contaminants. The labeled proteins can be enriched through avidin affinity chromatography followed by SDS-PAGE separation. A typical cell membrane impermeable biotinylation reagent consists of two moieties: a biotin group and a peptide reactive group. The peptide reactive moiety of biotinylation reagents usually targets either the free amine groups (82)or free thiol groups (55) in

protein/peptides. These labeling methods were successfully employed in the detection of Pir (Protein with internal repeats) proteins in *S. cerevisiae* (51), Pir and Tsa1p in *C. albicans* (35) (87). During avidin-biotin affinity purification of the labeled proteins/peptides, unlabeled proteins/peptides may nonspecifically bind to avidin and eventually result in false identification. This non-specific binding problem could be aggravated especially in the purification of heavily glycosylated cell wall proteins.(49)

The accumulation level of *C. albicans* cell wall/surface proteins is highly dynamic and often correlates with the extracellular environment. The dynamic change of cell wall/surface is especially prominent during the conversion from yeast to hyphae of this pathogenic fungus. Switching to hyphae is believed to be a necessary process for the pathogen's penetration into epithelia, endothelia and tissue as a result of the morphological adaptation of the fungal cell, thereby contributing to its virulence *in vivo* (57). The morphogenesis is orchestrated by a wide spectrum of complex signal network pathways under certain stimulus factors such as elevated culture temperature at 37° C (29). Understanding the surface proteomic change of *C. albicans* during morphogenesis is essential to the elucidation of fungal pathogenesis, thereby shedding lights on the path of new drug target discovery. In previous studies, comparative 2D-PAGE based approaches were employed to quantify the surface proteome change during a yeast to hyphae conversion of *C. albicans* (66) (20). A small number (24) of proteins were found to be differentially regulated before and after germination (20), probably due to the complicated sample preparation procedure and the incompatibility of the gel-based method employed in the cell wall protein separation. More recently, a comprehensive proteomic analysis of cytoplasmic proteins and cell wall associated proteins from yeast cells, hyphae, and biofilms of *C. albicans*

has been reported by Gil and co-workers(47). Compared to the conventional 2D-PAGE based approaches, 2D-DIGE (2-D Fluorescence Difference Gel Electrophoresis) technique significantly enhanced the capacity of protein quantification.(1) In addition to the 2D-PAGE based comparative proteomic analysis, stable isotopic labeling techniques has more attractive advantages such as good reproducibility, high accuracy and high throughput in protein quantification (34). SILAC (61) is one of the important isotopic labeling methods and has been successfully applied in various biological systems (74) (43) (88). The first step in SILAC method is to metabolically label proteins through the incorporation of essential or nonessential amino acids to living cells. Lysine is commonly chosen for its relative high abundance in proteins and presence in the peptides produced by trypsin digestion. Relative quantification of protein can be achieved by comparing the chromatographic peak areas of two paired peptides with the isotope coded amino acid residues. Another advantage of the stable isotopic tagging is that it can be introduced *in vivo* at early stages of cell growth; thereby variations due to sample preparation, which were commonly observed in *in vitro* labeling, can be greatly reduced. One major disadvantage inherent with the SILAC is that it can not be applied to microbes which do not normally need any exogenous amino acids. Fungal cells (including *C. albicans*) are just one such organism. Fortunately, this issue can be addressed through choosing a mutant strain which is auxotrophic for certain amino acids such as lysine.(5)

In this study, the cell surface proteome of *C. albicans* was profiled by intact cell surface biotinylation labeling technique coupled with LC-MS identification of surface exposed peptide segments. Characteristic fragments originated from biotin tag moiety on tandem mass spectra of peptide were employed to confirm that the peptides were labeled. Facilitated by the availability

of a lysine production deficient mutant, the dynamic change of the surface proteome during morphogenesis was quantified by using a SILAC based method which involved in an *in vivo* stable isotope coded lysine labeling technique.

## 4.3 Experimentals

### 4.3.1 Yeast strains

*C. albicans* 6284 was the result of repairing strain BWP17 to prototrophy (69). It was a gracious gift from Dr. G. Palmer (Louisiana State University Health Sciences Center, New Orleans, LA). Strain CLD2 was obtained after sequential disruption of the LYS2 alleles and auxotrophic for lysine (5). It was generously given as a gift from Dr. J. K. Bhattacharjee (Department of Microbiology, Miami University, Oxford, OH).

### 4.3.2 Cell surface labeling with biotinylation reagents

*C. albicans* 6284 were cultured on YPD plate (Becton, Dickinson and Company, Franklin Lakes, NJ) at 37 ° C for 2 days. Cells were slightly scratched off the medium using sterilized pipette into 1.5 mL centrifuge tube and consequently washed by ice-cold phosphate-buffered saline PBS pH 7.4, five times with final $OD_{600}$=2.7-2.8 in 1mL PBS 7.4. Biotin labeling reagent (sulfo-NHS-LC-biotin or sulfo-NHS-SS-biotin or iodoacetyl-$PEO_2$-biotin) (Pierce, Rockford, IL) (14 mM) in 300 ul of ice cold PBS 8.0 was added to the washed cells. After a short vortexing to form a cell suspension, the mixture was incubated at 4°C with 400 rpm mixing in a cold room for 2 hours. When iodoacetyl-$PEO_2$-biotin reagent was used, intact yeast cells were initially treated by 10 mM tris-(2-carboxyethyl) phosphine (TCEP) (Promega, Madison, WI) at 4°C for 20 minutes prior to the incubation step, and the cells were washed again by PBS 7.4 3~5 times

before the labeling. The labeling reaction was quenched by adding ice cold 1 mL of 50 mM Tris-HCl (pH 8.0) and by further incubation at 10°C for 15 minutes. Finally the labeled cells were extensively washed by ice cold PBS buffer (pH 7.4) to remove by-products of the biotinylating labeling reaction.

### 4.3.3 Tryptic digestion of intact cells ("whole cell shaving") and peptides extraction

Yeast cells labeled with sulfo-NHS-LC-biotin or iodoacetyl-$PEO_2$-biotin were first washed by 1 mL of 50% methanol two times and suspended in 500 µL 50 mM ammonium bicarbonate. To this suspension of yeast cells, 15 µl of 200 mM dithiothreitol (DTT),(Sigma, St. Louis, MO) was added followed by 40 minutes incubation at room temperature. 7 µL of 1 M iodoacetamide (IA) (Sigma, St. Louis, MO) was added to the cell suspension followed by an additional 40 minutes incubation at room temperature. After centrifugation, cell pellets were washed by 500 µL of 50 mM ammonium bicarbonate three times. Finally, 8 µg trypsin in 40 µL of 50 mM ammonium bicarbonate was added to the cell pellet followed by a brief vortexing. The trypsin digestion was carried out at 37°C for 18 h. For sulfo-NHS-SS-biotin labeled cells, washing the cells with 60% MeOH was conducted instead of dithiothreitol/iodoacetamide (DTT/IA) treatment before trypsin digestion.

The digestion was stopped by freezing the sample in liquid nitrogen. Supernatant containing soluble peptides were collected. More tryptic peptides were subsequently extracted by acidic water (5% formic acid), 10% ACN, 30%ACN and 50% ACN. Extracts were combined and then evaporated to dryness in vacuum.

### 4.3.4 Affinity enrichment of biotinylated peptides

Extracted peptides after "whole cell shaving" were first purified using ICAT-cartridge-cation exchange column (Applied Biosystems, Foster City, CA) to remove DTT, IA and trypsin. Peptides labeled by non-cleavable reagent were enriched on cartridge-avidin column according to the protocol provided with the ICAT reagent kit (Applied Biosystems, Foster City, CA). After vacuum drying at room temperature, the labeled peptides were dissolved in 3% ACN containing 1% formic acid for LC-MS/MS analysis.

Sulfo-NHS-SS-biotin labeled peptides were incubated in 0.2 mL neutravidin beads (Pierce, Rockford, IL) in 200 µl binding buffer (2×PBS pH 7.2) at room temperature for 1 hour with mixing. After centrifuge, supernatant was discarded and the beads were washed three times by 2 mL of PBS (pH 7.2), 2 mL of PBS ( pH 7.2 ) containing 1% Triton X-100, 2 mL of 50 mM ammonium bicarbonate in 20% MeOH and finally 2 mL of $H_2O$. The beads were incubated in 250µL of 20 mM TCEP at room temperature for 30 minutes to release the labeled peptides. The reduction and extraction process were repeated two times followed by additional four times washing of the bead with 5% ACN. All extracts and washes were combined and vacuum dried before LC MS/MS analysis.

### 4.3.5 In vivo isotopic lysine labeling during germination growth of yeast

Lys2$^-$ auxotroph strain (*C. albicans* CLD2) cells (5) were first streaked on a YPD ((1% yeast extract, 2% bacto peptone, 2% glucose) plate and incubated for 3 days at room temperature. A single colony was then selected and resuspended in 5 mL of liquid YPD medium in a centrifuge tube for pre-culturing at 25°C for 10 hours. These cultures were inoculated in 100 mL of YPD with a dilution of 1:200 and grown under 25°C for 36 hours as yeast cells. Cells were harvested

by centrifugation under 4°C and extensively washed by cold PBS buffer (pH 7.4). Cells were subsequently inoculated at $8x10^6$ /cells in 100 mL SILAC DMEM (minus lysine and arginine) (Pierce, Rockford, IL) which had been supplemented by isotopic labeled lysine (K*.2HCl, 180 mg/L) (Cambridge Isoptope Laboratories, Andover, MA), arginine (R.HCl, not isotopic labeled, 84 mg/L) (Sigma, St. Louis, MO) and HEPES (5.958 g/L) (EMD Chemicals Inc., Gibbstown, NJ). After 2 hours of germination at 37 °C hyphae cells were then harvest by centrifugation under 4°C and repeatedly washed by cold PBS buffer (pH=7.4).

### 4.3.6 Extraction cell wall proteins by SDS-DTT

A cell wall fraction was prepared according to procedures described previously with some modifications (66). Briefly, hyphae cells were washed 3 times by ice cold lysis buffer (10 mM Tris-HCl, pH 7.4, 1 mM phenylmethylsulfonyl fluoride, PMSF), resuspended in cold lysis buffer and transferred in screw capping vials which contained equal volume of glass beads (425-600 μm in diameter, acid washed). The cells were mechanically broken with the glass beads in a dry ice precooled bead mill cell homogenizer at the maximum speed for 45s for 7 cycles with intermittent cooling on ice immediately after each cycle in a cold room (4°C). The complete cell breakage was confirmed by microscopic examination and failure of cell growth in YPD-chloramphenicol (10 μg/mL) plates. Glass beads were filtered out through a sponge packed 10 mL syringe and washed by cold lysis buffer untill the washing solution was clear. Crude cell wall pellet was separated from of soluble cytosolic fraction after centrifugation at 2000 g for 10 minutes at 4°C. The collected cell wall fraction was washed and vortexed in 10 mL of cold 1mM PMSF solution for 3 times. The wall pellet was subjected to extensive washing and vortexing by a gradient of cold 10 mL NaCl solution (5%, 2%, 1% in 1mM PMSF) for 5 times each. The

pellet was finally washed by another 10 mL of cold 1 mM PMSF again. The washed cell wall fraction was boiled in 300 ul of SDS extraction buffer (2% SDS, 0.1 M EDTA, 15 mM freshly made DTT, 50 mM Tris-HCl, pH 8.0) for 4 times of 10 minutes. The supernatants were combined for the first three extractions after centrifugation at 1000 g for 10 minutes and the $4^{th}$ extract was discarded. SDS extracted proteins were immediately carboxyamidomethylated with 500 mM iodoacetamide (4 fold mole excessive to DTT in the extracts) at room temperature for 1 hour in the dark. SDS extracted proteins were precipitated by $CHCl_3$/MeOH (90). The protein pellet (SDS-DTT extracted cell wall protein) was stored at -80°C for future use.

### 4.3.7 SDS-gel electrophoresis (PAGE) and in-gel digestion

The protein pellet from SDS extraction was suspended in 1% SDS and heated at 90°C for 2 minutes. After a brief vortexing to solubilize the pellet, protein concentration was determined by BCA. Protein samples (three biological replicates, 100 μg for each sample) were loaded onto each well of a mini-gel, separated by SDS-PAGE 9% acrylamide gels and stained by Coomassie blue. Stained bands for each sample were excised and in-gel digested as previously described (71).

### 4.3.8 LC-MS/MS analysis and database search

LC-ESI MS/MS was performed on Finnigan LTQ$^{TM}$-ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) coupled with a Finnigan Micro Auto Sampler. Vacuum-concentrated samples (not completely dried) were suspended in 25 μL of 5% acetonitrile and 5% formic acid in a ninety six well micro plate and the volume for each sample injection was 8 μL. The sample mixture was first loaded to a $C_{18}$ trap column and washed with 3% acetonitrile and

0.1% formic acid for 60 minutes for desalting, then the purified peptides were eluted to a reverse-phase C18 analytical column by a 60 minutes gradient made of A buffer (0.1% formic acid: 97% water:3% acetonitrile v/v/v) and B buffer (0.1% formic acid: 3% acetonitrile: 97% water v/v/v) at a flow rate of 200 ~ 500 nL/min. Separated peptides were analyzed under the data-dependent acquisition mode set by Xcalibur, 2.2 version (Thermo Electron). After a MS survey scan over the m/z range of 300-2000, seven (three for samples from SILAC labeling experiments) most intensive precursors were selected and subjected to fragmentation by CID (collision induced dissociation). The normalized collision energy was set at 35% with activation Q value being 0.25 and dynamic exclusion of 100 seconds. The acquired raw data were processed by BioWorks software, version 3.3 (Thermo Electron) and tandem mass spectra were searched against a *C. albicans* protein database (downloaded from NCBI) by using SEQUEST algorithm.

Parameters for SEQUEST database search were set as follows: differential mass increase of 414.19 Da on cysteinyl (for iodoacetyl-$PEO_2$-biotin labeling), 340.1 Da on lysine residue (for sulfo-NHS-LC-biotin labeling), 88.01 Da on lysine residue (for sulfo-NHS-SS-biotin labeling) and 8.00 Da (for SILAC labeling) in addition to 57.02 Da on cysteinyl and 15.99 Da on methionine residue. The number of missed cleavage sites was set to three. The output for the search results was filtered by cross-correlation score (XCorr) which must be $\geq 3.0$ for the triply charged ions, $\geq 2.5$ for the doubly charged ions and $\geq 2.0$ for the singly charged ions. Also, delta correlation score ($\Delta$Cn) was set to be $\geq 0.1$. MS/MS spectra of all the biotinylated modified peptides were manually inspected to ensure the quality of spectra. “Pepquan” tool in BioWork

3.3 was used for SILAC protein quantification. Chromatographic peak areas of peptide pairs were manually inspected or adjusted to ensure accurate quantification.

## 4.4 Results and Discussion

### 4.4.1 Surface protein profiling by three biotinylating reagents and LC-MS

The intact yeast form of *C. albicans* 6284 was labeled with three membrane impermeable biotinylation reagents respectively at low temperature. These labeling reagents were routinely used in cell surface biotinylation method (23) as a procedure to reduce possible cytosolic contamination in the cell membrane protein fraction (21). Sulfo-NHS-LC-biotin and sulfo-NHS-SS-biotin target free amine groups in the side chain of lysine residue and unmodified N-termini of proteins, while iodoacetyl-$PEO_2$-biotin reacts with cysteine residues. To maximize the sequence coverage, confirm the labeling and test the efficiency of those reagents in the bitotinylation of *C. albicans* cell surface proteins, all three labeling reagents were employed. In order to reduce the potential cell response to the labeling reactions, the labeling procedures were conducted at low temperature (4°C) in PBS (8.0) buffer. Urban's group also confirmed the efficacy of the labeling and membrane impermeability of sulfo-NHS-LC-biotin by using colloidal gold immunoassay (87).

Unlike other methods in which cell wall fractions were prepared from cell lysate or biotinylated proteins were enriched after the labeling process, the "whole cell-shaving" method employed in our experiments yielded labeled peptides by trypsin digestion of the intact fungal cells. The "shaving" technique had been previously utilized to identify bacterial surface proteins for vaccine candidates search (70). We employed this technique with an aim to preferentially cleave

surface proteins mostly exposed to the extracellular environment. Also, membrane proteins (such as GPI anchored proteins) are likely to be shielded by outer cell wall proteins and partial trypsin digestion of these proteins was expected. However, this "whole cell shaving" technique could not totally eliminate cytosolic contamination, as the DTT/IA treatment of the whole fungal cell might damage the plasma membrane and lead to leakage of cytosolic proteins (36). Avidin affinity chromatography allowed us to enrich biotinylated peptides from a highly complex mixture of peptides, but the enriched biotinalyted peptide fraction may still contain digested cytosolic proteins. Fortunately LC-MS and bioinformatic analyses of enriched biotinalyted peptide fraction would provide unambiguous information as to which peptide was biotinalyted, and thus determine the cell wall/surface location of the protein from which the biotinayated peptide was derived.

LC-MS analysis yielded tandem mass spectra of peptides in the enriched biotinalyted peptide fraction. Database searches were subsequently performed on the peptide tandem mass spectra by using SEQUEST algorithms. Once a tandem mass spectrum was determined to match a peptide sequence, the search result would also include information on which amino acid residue (s) was (were) modified by the labeling reagent defined prior database search. Moreover, characteristic fragments originated from the labeling reagent further confirmed that the peptide was labeled. Figure 4.1B is a tandem mass spectrum of peptide ion at m/z: (the precursor ion was detected in MS spectrum as shown in Figure 4.1A). The b ions and y ions in the spectrum provided information about the sequence of the peptide as well as which amino acid residue was modified by the labeling reagent (sulfo-NHS-LC-biotin). Two fragment ions at m/z 340.2 and m/z 227.1 were originated from the sulfo-NHS-LC-biotin tag (see Figure 4.1C). As for iodoacetyl-$PEO_2$-

biotin labeled peptides, characteristic fragment ions at (m/z): 449.1, 375.2, 332.2 and 270.1 corresponding to the cleavage sites as shown in Figure 4.2 could also be observed for the biotinylated peptides upon CID. These fragment ions originated from iodoacetyl-$PEO_2$-biotin tag were comprehensively studied previously (7).

Even though sulfo-NHS-SS-biotin labeled peptides do not yield characteristic fragments that can be retrieved to the labeling reagent, the tandem mass spectra of the labeled peptides contains enough information for the sequence and modification site(s) determination (data not shown). It should be noted that biotinylated peptides determined in the above mentioned experiments not only confirmed the cell surface localization of the proteins from which they were derived, but also represented the peptide segments of these surface/surface affiliated proteins that were likely exposed outside to the extracellular environment. For instance, several biotinylated peptides were detected for CaPMA1 plasma membrane H+-transporting ATPase (CA 2300). The detected peptides ( Figure 4.3 (A)) were located in the hydrophilic loops that were likely exposed to the outside as predicted by TopPred (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html) (Figure 4.3 (B) ), rather than in the more hydrophobic parts that might be buried in the cell wall or cell membrane.

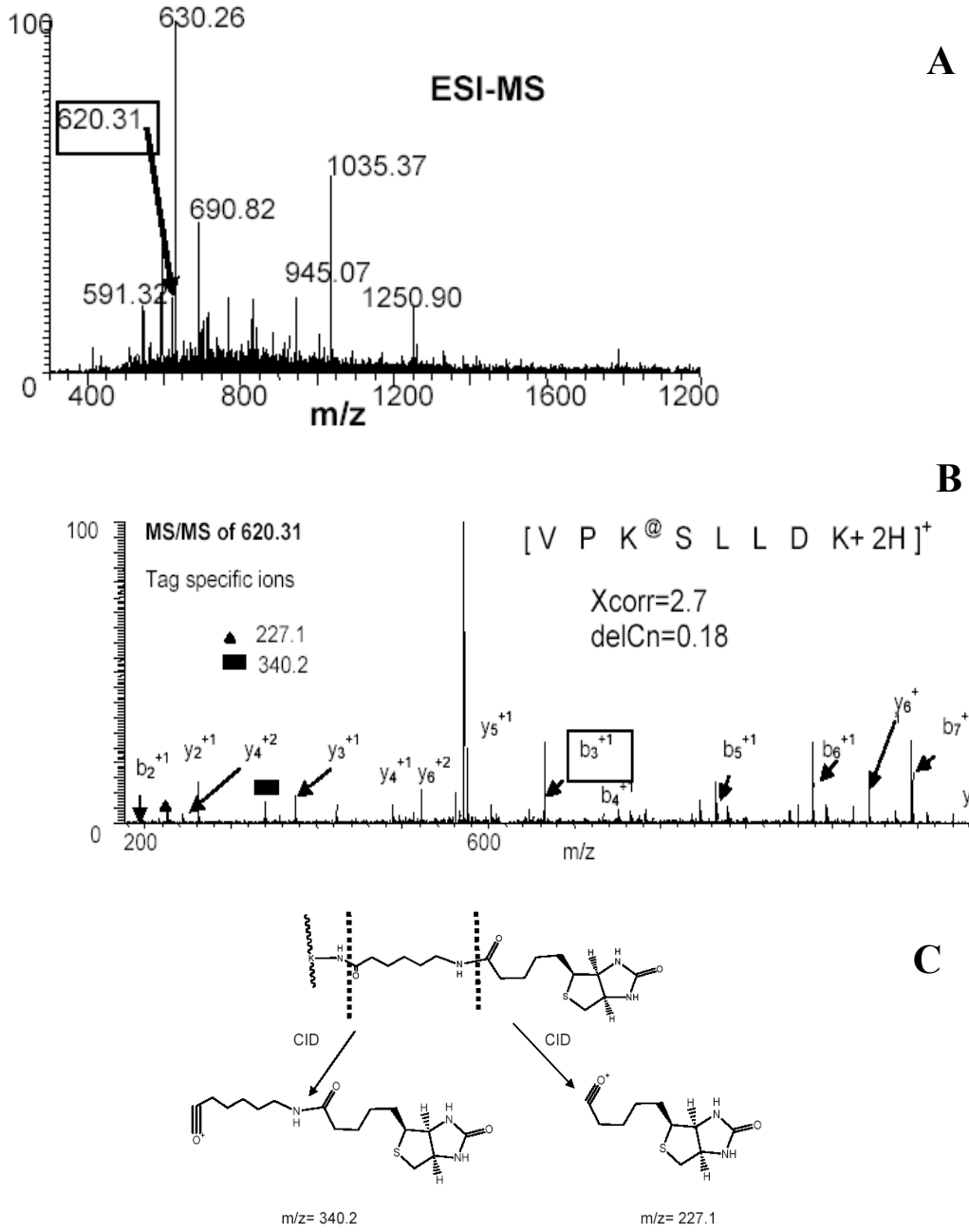



**Figure 4.1.** An example of LC-MS analysis of enriched peptide released from *C. albicans* cells labeled by sulfo-NHS-LC-biotin. (A) Precursor ion 620.31 (m/z) was selected for MS/MS experiment. (B) The MS/MS spectrum for the precursor ion 620.31 (m/z). The "b" and "y" ions are labeled together with the two biotin moiety specific ions of 227.1 (m/z), 340.2 (m/z). The peptide was identified by SEQUEST as VPK@SLLDK (K@, modified lysine residue). (C) Two characteristic fragment ions produced upon CID from the biotin moiety attached to a lysine residue

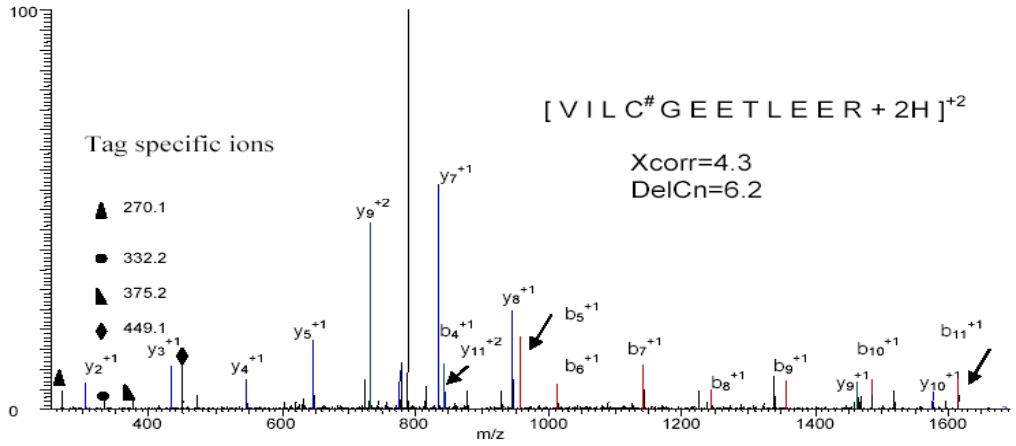



**Figure 4.2.** MS/MS spectrum of an iodoacetyl-$PEO_2$-biotin labeled peptide matched the sequence of VILC#GEETLEER as a result of SEQUEST search. Four tag specific ions at m/z: 270.1, 332.2, 375.2, and 449.1 can be observed.

(A)

MSATEPTNEKVDKIVSDDEDEDIDQLVADLQSNPGAGDEEEEEENDSSFKAVPEELLQTDPRVGLTDDEVTKRRKRY
GLNQMAEEQENLVLKFVMFFVGPIQFVMEAAAVLAAGLEDWVDGVICALLLLNAFVGFIQEYQAGSIVDELKKTLANS
ALVVRNGQLVEIPANEVVPGDILQLEDGTVIPTDGRIVSEDCLLQVDQSAITGESLAVDKRSGDSCYSSSTVKTGEAF
MIVTATDSTFVGRAAALVNKASAGTGHFTEVLNGIGTTLLVFVIVTLLVVWVACFYRTVRIVPILRYTLAITIIGVPVGLPA
VVTTTMAVGAAYLAKKQAIVQKLSAIESLAGVEILCSDKTGTLTKNKLSLHEPYTVEGVEPDDLMLTACLAASRKKKG
LDAIDKAFLKSLINYPRAKAALPKYKVIEFQPFDPVSKKVTAIVESPEGERIICVKGAPLFVLKTVEDDHPIPEDVH

(B)

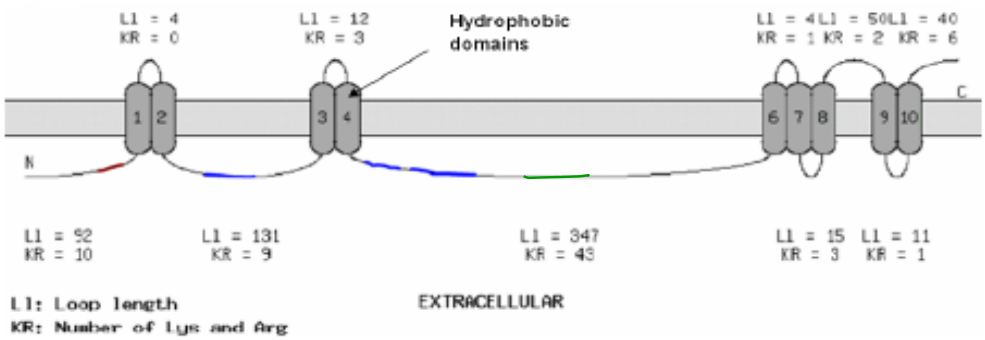



**Figure 4.3.** (A) Part of amino acids sequences of CaPMA1 plasma membrane H+-transporting ATPase and peptides detected by labeling using sulfo-NHS-SS-biotin (highlighted in dark red), iodoacetyl-$PEO_2$-biotin (highlighted in blue) and Sulfo-NHS-LC-biotin (highlighted in green) respectively (B) Topology of CaPMA1 predicted by TopPred from (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html). Lines colored in brown and blue corresponding to the peptides in the same color shown in (A).

### 4.4.2 Cell wall proteins identified based on the detection of biotinylated peptides

Based on this approach, LC-MS analysis of biotin-avidin affinity enriched tryptic peptides resulted in the identification of more than 200 proteins. The experiments included: 1. whole cell labeling by sulfo-NHS-LC-biotin or sulfo-NHS-SS-biotin or iodoacetyl-$PEO_2$-biotin; 2. trypsin digestions of the labeled whole cell to release peptides; 3. biotin-avidin enrichment of the biotinylated peptides; 4. LC-MS/bioinformatic analyses of the biotinylated peptide enriched fraction. Among these 200 plus proteins, one hundred and fourty eight cell wall proteins (supplemental Table S1) were assigned based on the detection of multiple biotinylated peptides (see Table 4.1 for examples). If protein identification was based on three or less peptides, the tandem mass spectra of the peptides were manually inspected for biotin tag specific ions to verify the identification. This study, to our knowledge, is the most accurate profile of cell surface proteome of *C. albicans*, because the identification of proteins was based on the detection of labeled peptides, and in some cases was further confirmed by the characteristic fragment ions originated from the biotinylation labeling reagent. In our experiments, fifty one proteins were identified as cell wall proteins for the first time (highlighted in green in supplemental Table S1). Among these fifty one identified proteins, twenty seven were identified by only one biotinylated peptide, possibly due to the low abundance of these proteins. Even though these twenty seven proteins had never been reported as cell wall proteins, their cell wall/surface localization could be confidently assigned based on the unambiguous detection of biotinylation labeling reagent modified peptide, i.e. detection of the exposed peptide segment of this protein.

**Table 4.1.** A short list of 5 examples of identified cell surface proteins of *C. albicans* strain 6284 based on detection of biotinylated peptide(s) by LC-MS

| Accession number | Protein name | Detected peptides | Xcor | ΔCN | Function |
|---|---|---|---|---|---|
| CA2300 | CaPMA1 plasma membrane $H^+$-transporting ATPase 1 | R.IVSEDC@LLQVDQSAITGESLAVDKR.S<br>K.LSLHEPYTVEGVEPDDLMLTAC@LAASR.K<br>K.LSAIESLAGVEILC@SDK.T<br>R.SGDSC@YSSSTVK.T<br>K.VIEFQPFDPVSKK&VTAIVESPEGER.I<br>R.VGLTDDEVTK#R.R | 6.3<br>6.3<br>4.4<br>2.8<br>3.8<br>3.8 | 0.6<br>0.6<br>0.5<br>0.5<br>0.5<br>0.5 | Hydrogen ion pump |
| CA2582 | CaTAL1 transaldolase (by homology) | R.QFGK#DAVTLLTELESR.F<br>K.EK#AEIALDR.L | 4.3<br>2.6 | 0.6<br>0.3 | An enzyme catalyses the transfer of a dyhydroxy-acetone moiety |
| CA2857 | CaSSA1 Heat shock protein of HSP70 family | K.SK&LDASEIEEVTK.A<br>R.QATK#DAGTIAGLNVMR.I | 4.6<br>4.3 | 0.5<br>0.5 | Member of HSP70 family, highly immunogenic, protein folding, Adhesin |
| CA0855 | CaAMYG1 glucoamylase | R.K&NPFGLLVALDAEGTASGK.L<br>R.K#NPFGLLVALDAEGTASGK.L | 5.6<br>4.5 | 0.6<br>0.5 | Involved in cell wall biosynthesis, An enzyme for carbohydrate metabolism |
| CA3746 | CaARA1 D-arabinose dehydrogenase | R.YIC@DPWGYGIGFR.W | 3.4 | 0.5 | An enzyme of erythroas-corbic acid biosynthesis |

Note: The accession numbers are adopted from Candida DB (http://genodb.pasteur.fr/CandidaDB). K&: lysine residue modified by sulfo-NHS-LC-biotin; K#: lysine residue modified by sulfo-NHS-SS-LC-biotin; C@: cysteine residue modified by iodoacetyl-$PEO_2$-biotin

Three biotinylation reagents, sulfo-NHS-LC-biotin, sulfo-NHS-SS-biotin, and iodoacetyl-$PEO_2$-biotin, were employed in three similar approaches to identify cell wall/surface proteins of *C. albicans*. Each approach has unique protein identifications, and there were also protein identifications shared by different approaches (Figure 4.4) (supplemental Table S1). Twenty proteins (GAP1 TEF1, SSA4, FBA1, PDC11, PGK1, EFT2, TPI1, MET6, PET9, RPL3, SSB1, SAH1, HXK2.3f, ACH1, CHT2, RPL82, ATP2, IPF2431, PMA1) were detected in all the three approaches, which implies that these proteins may have relative high abundance in the cell wall/surface. The repetitive detection of these proteins also confirmed their cell wall localization and the corroborative nature of the three approaches. Surprisingly, ACH1, SAH1 were never reported in literature as cell wall proteins even though they could be regarded as abundant proteins in our study. It seems that these proteins may also have known cytosolic functions such as acetyl-CoA hydrolase activity (for ACH1) and adenosylhomocysteinase activity (for SAH1), but their roles in cell wall remain unknown. These two proteins may be regarded as secreted proteins since they were both detected in the serum from patients under systemic candidiasis (64). SAH1 may also contribute to the biofilm formation of *C. albicans* and act as an adhesion molecule to host (25). ENO1 was identified as one of the most abundant cell wall proteins in our study with the detection of twenty eight peptides labeled by sulfo-NHS-LC-biotin or sulfo-NHS-SS-biotin (but not by iodoacetyl-$PEO_2$-biotin because it has no cysteine residue in its sequence). It was also previously identified as a major cell wall glycolytic enzyme located in the inner layers of the glucan wall matrix via monoclonal antibody probing and immunoelectron microscopy(3), and this protein might be account for 0.7 and 2% of the total protein in both yeast and hyphae cells (83). Edwards group (21) first detected the presence of Enolase in the

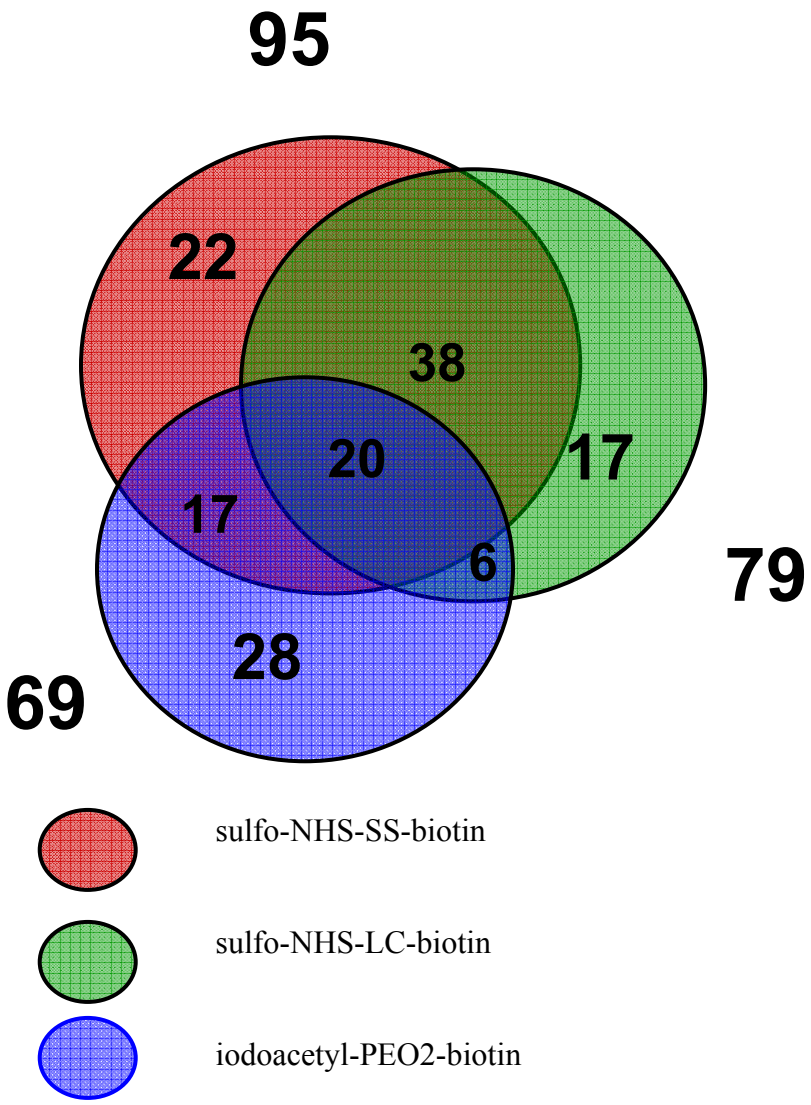



**Figure 4.4.** Pie graphic representation of detected cell wall proteins biotinylated by three labeling reagents. 95 proteins were detected via labeling by sulfo-NHS-SS-biotin (dark red), 69 by iodoacetyl-$PEO_2$-biotin (in blue) and 79 by sulfo-NHS-LC-biotin (in green).

cell wall of *S. cerevisiae*, which prompted them to propose an alternative secretion pathways for proteins without signal sequences. Other proteins (labeled by all three labeling reagents), such as FBA1, MET6, PGK1, SSB1, TEF1, TPI1 and TSA1 are mainly involved in pathogen- host interactions. FBA1, MET6, PGK1, SSB1, and TPI1 showed immunogenecity in a murine model and could be induced by symbiote of host defense response (65). Interestingly, ribosomal proteins (RPL3 and RPL82) and translational elongation factors (EFT2 and TEF1) were also detected as abundant cell wall proteins. Although their functions at cell wall were unknown, they had been shown to be highly antigenic both in mouse and human (48)

The largest overlapping proteins were detected by both of the two approaches that involved in sulfo-NHS-LC-biotin and sulfo-NHS-SS-biotin. These thirty eight cell wall proteins include ENO1, ADH1, ECM33.3, RPL2.3, RPL17B, RPS15.3, CIT1, RPL18, RPL19A.3, GAD1, RPS31, RPL23B.3, RPL7A.3, IPF8762, PRA1, RPL28.3f, IPF14171, RPS4, KAR2, HHF21, RPS3, PGI1, ICL1, POT14, SSC1, HTA1, RPL10, RPL32, HSP70, RPL6.3, IPF9926, RPS22A, HTB1, EFB1, ACT1, ADK1, RPS19A.3, AMYG1. This large overlapped detection was because that both labeling reagents targeted the same functional group (amine) in cell wall proteins. Interestingly, for the same protein detected by both approaches, each approach detected different regions. For example, sulfo-NHS-LC-biotin detected K&NPROFGLLVALDAEGTASGK (& indicates the biotinylation site) while sulfo-NHS-SS-biotin labeled K+NPFGLLVALDAEGTASGK (+ shows the biotinylation site), although both labeling reagents identified the same protein AMYG1. Since the cells were labeled under the same conditions, the difference in the detection of labeled peptides could be due to the peptide elution process during affinity enrichment. The sulfo-NHS-SS-biotin labeled peptides were released

from the avidin column via the reductive cleavage of the disulfide bond, and the released peptide has a smaller tag than sulfo-NHS-LC-biotin labeled peptides.

The overlapped detection shared by approaches involved with sulfo-NHS-SS-biotin and iodoacetyl-$PEO_2$-biotin includes seventeen proteins (PCK1, MLS1, CDC19, PDA1, GND1, CEF3, GPH1, CAR2, GCV2, ADH2, RPS23, IPF17186, RPL10A, ZWF1, ACO1, RPL12, IDH2). These two labeling reagents target different functional groups. Similarly, the overlapped detection shared by sulfo-NHS-LC-biotin approach and iodoacetyl-$PEO_2$-biotin approach include six proteins (RPS10, SRB1, RPS6A, NCP1, PHR2, RPL4B). It should be noted that large set of proteins were detected in each approach (ATP1, GPM1, QCR2, IFD4, TAL1, RPL13, LPD1, IPF9803, TKL1, IDH1.3, VMA4, GCV1, RBT5, HHT21, CDC48, AMS1, IPF11299, HSP90, HHT3, KRS1, FDH11.3, IPF3358 in sulfo-NHS-SS-biotin labeling approach; ALD5, IPP1, HSP60, MIS11, MDH1, FAS2.5f, SCW1, SAC6.5f, HEM13, ADO1, CHC1, RIB3, ARO4, IPF11123, YNK1, AAT1, FBP1, ADE6, MNT1, ARA1, CPR3, VMA2, IPF6629, ADE17, IPF9550, BMH2, IPF470, TRX1 in iodoacetyl-$PEO_2$-biotin labeling approach, and ALS4.5f, CYP1, UCF1, IMH3, BGL21, IPF17237, PDI1, ADH5, MCR1, GSP1, ATP4, FUM12.5f, IPF11888, KES1, SOD21, ACC1, NTF2 in sulfo-NHS-LC-biotin labeling). Sulfo-NHS-SS-biotin had been chosen by other investigators (63) to reduce possible intracellular labeling since the reductive cytosolic milieu would potentially cleave the disulfide bond in the labeling reagent, (6) therefore affect the enrichment of  labeled cytosolic proteins. Contrary to this theory, we found that the sulfo-NHS-SS-biotin labeling approach detected the largest number of proteins, and it might imply that proteins detected by biotinylation were not related to the cytosol.

### 4.4.3 Localization of identified putative cell wall proteins

All identified cell wall proteins were analyzed by Gene Ontology (GO) (8) (4) Slim Mapper at CGD (http://www.candidagenome.org/cgi-bin/GO/goTermMapper) with a filter set for "annotations from High-throughput Experiments" and "the Manually Curated". As shown in Figure 4.5, 55.8% of the identified proteins have been fully annotated. Majority of the annotated proteins were cell wall (25.9%), plasma membrane (20.4%), membrane associated (7.5%) and extra cellular region (4.8%) proteins. This result demonstrated that biotinylating reagents were selective in labeling cell wall/membrane proteins. BGL2 was a classical yeast form cell wall protein of *C.albicans* which had been confirmed to have cell surface location (33, 66) (12). It should be noted that many proteins identified in our study might have multi-locations. For example, ENO1 was classified in six locations (cell wall, membrane, plasma membrane, cytoplasm, membrane fraction and extracellular region). ADH1, ENO1, HSP70, HSP90, and PGK1 from cell wall fraction were also assigned as cytosolic proteins (5 out of 19). It is quite common in microbes that proteins may have multiple localizations, and a term "moon-lighting" was coined for those proteins with an emphasis for their second function in the cell wall. Also, a new mechanism (59) of protein exporting had been proposed for cell wall proteins that lacked an N-terminal signal sequence for secretion. An increasing number of proteins without N-terminal signal peptide had been found in both *S. cerevisiae* and *C. albicans* and were compared in a recent review (59). Detection of conventional cytosolic proteins in this study is consistent with previous reports in which the surface presence of highly abundant proteins (GAP1, FBA1, PDC11, TPi1p, ADH1, PGK1, GPM1) involved in glycolysis were confirmed. For instance, GAP1 protein was previously detected as cell surface antigen through indirect immunofluoresence (26). Moreover, work by Cutler's group (91) demonstrated that glycopeptide

vaccines containing peptide moieties from FBA1, ENO1 and GAP1 could induce protection against candidiasis in a murine model. An elegant study had been done to confirm that PGK protein was present in the cell wall by several approaches (2). Similarly, FBA1, PDC11, GPM1 and TPi1p were detected in cell wall fractions by using 2D-gel based proteomic method (47) (20) (66). Heat shock protein like SSA1 and HSP60 had also been previously detected as cell wall proteins both in *C. albicans* (40) and *S. cerevisiae* (40). HSP90 (or its fragments) was isolated as a major antigen from the body fluids of patients

with disseminated candidiasis (50), suggesting that it might be also a secreted protein. For these proteins with multiple cellular localizations, their functions in the cell wall were believed to be unrelated to their well known cytosolic functions. ADH1 was also proven to be critical in biofilm formation on catheter surfaces (53) and recently its existence on the cell surface was confirmed by a biotinylation technique (87). The high abundance of glycolytic enzymes in cell wall may be relevant to their roles in *C. albicans* pathogenesis. Previous studies indicated that those proteins were highly immunogenic in inducing host immune responses. (26) (32) (86) (75). Much attention has been drawn to these glycolytic enzymes in vaccine development regarding their function in stimulating immune response of host cells. ENO1 was viewed as an immuno-dominant antigen as other conventional cytosolic proteins such as elongation factors and glycolytic enzymes in literature. For instance, one of the three cytosolic proteins (ACH1, ICL1, MLS1) identified in our study, ACH1 was well known for converting acetyl-Co A to/from acetate in cytosol , and its possible cell wall location has never been determined previously. However, ACH1(64)was detected in the serum specimens of candidiasis patients and showed high immunogenicity. Interestingly, work by Lorenz group (11) showed that a mutant strain lacking ACH1 is still fully virulent in a murine model.

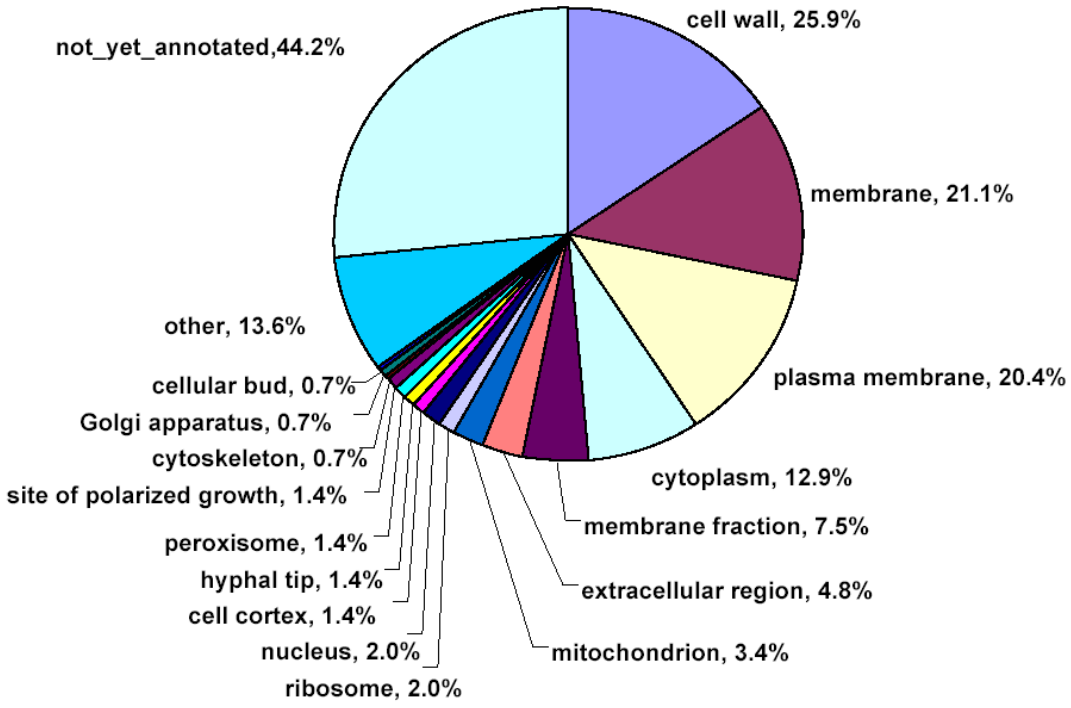



**Figure 4.5.** Component classification of identified cell wall proteins based on GO Slim Mapper at CGD. The classification was based on annotation type selections for *High-throughput Experiments and Manually Curated* (the same for process classification in Figure 4.6). Out called figures represent the percentage of the number of classified proteins at the corresponding component over the total annotated proteins. The same proteins may be classified into two or more components.

MLS is an active enzyme in glyoxylate cycle which is required for wild-type virulence in murine systemic infection (41).

Proteins normally observed in ribosomes were also found in our study. Many of these proteins like RPL10, RPL10A, RPL12, RPL13, RPL18, RPL19A, RPL23A, RPL32, RPL82, RPS10, can not be mapped by GO program at GCD due to no annotation for the specified GO Slim term in the selected annotation set. Therefore the real functions of these "ribosomal" proteins in cell wall need to be studied further in future. Other investigators also detected some ribosomal proteins (RPS4B, RPL3, RP10, RPL3 60S, RPL20B, RPP0) in cell wall pellet fraction (20) of both yeast and hyphal forms; however possible cytosolic contaminations as a result of low speed centrifuging during cell wall pellet fractionation could not be ruled out. On the contrary, results from another group (66) did not include any ribosomal proteins even though the same cell wall preparation method and gel analysis were employed. It is possible that the "ribosomal" proteins may adventitiously bind to the cell wall after they escaped from the broken dead cells and react with the labeling reagents as discussed above. Although this possibility can not be totally ruled out, the extensive labeling of those "ribosomal" proteins suggests their localization on the cell surface. Additionally, RPL20B, RPP0, RPL3 60S, RP10 ribosomal proteins were also detected in the SDS-DTT resistant cell wall fraction(20), which should be free of any non-covalently associated proteins (including cytosolic proteins stuck to the plasma membrane or wall matrix) after extensive hot SDS washings. Therefore, these "ribosomal" proteins may be regarded as cell wall proteins rather than cytosolic contaminations. Study has shown that RP10 was immunogenic in human infection of *C. albicans.* (86) In *Neisseria gonorrhoeae*, RP L12 was believed to play an important role in invasion of human reproductive cells and shown to be

associated with the plasma membrane and exposed to the cell surface (81). Immunochemical assays would be needed to confirm the cell surface localization of any protein of interest identified in this study.

A group of elongation factor proteins (TEF1, EFT2, CEF3,EFB1) was identified in our study as well, which was consistent with reports from other investigators (20) (66). Interestingly, these proteins were reported to be absent on the cell wall of *S. cerevisiae*. Detection of these elongation factors have also been reported as extracellular or cell envelope associated protein in microbes such as *C. albicans* (44), *Trichoderma.reesei* (39) and *Staphylococcus aureus* (76). Moreover, the presence of TEF1 in tobacco cell walls has been confirmed by immunogold SEM analysis (94). Localization of TEF1 on the cell surface of *C. albicans* could also be supported by its ability to bind plasminogen (17). The cytosolic functions of these proteins are related to protein synthesis and interactions with other proteins, but their roles on the cell wall are largely unknown. Although the approaches employed in our study minimized the possibility of cytosolic contamination, these possibilities still could not be ruled out: 1. Labeling reagents might react with those highly abundant cytosolic proteins in aging or non-viable cells whose disintegrating plasma membranes could not serve as barriers for the biotinylating reagents. 2. Certain channels (like $Na^+$ ion channel) in plasma membrane could potentially uptake the sulfo-NHS-LC-biotin and sulfo-NHS-SS-biotin labeling reagents. Some researchers intended to accept the fact that these "ribosomal" proteins could indeed present in the cell wall possibly through the non-canonical secretion pathways as mentioned earlier (59). Also, a series of histone-like proteins (HHT21, HHT3, HHF21, HTA1, HTB1) were detected in our study. Antigenic histone H2B-like

cell surface protein was also confirmed in *Histoplasma capsulatum* and predicted to be a new candidate antigen in vaccine development (60).

### 4.4.4 Identified putative cell wall proteins classified into different cellular processes by GO (8) (4)

The different roles of the identified proteins are summarized in Figure 4.6. The majority of proteins involved in the processes of interspecies interaction between organisms (17.7%), interaction with host (12.9%), carbohydrate metabolic process (8.8%), filamentous growth (8.2%) and pathogenesis (6.8%). Other proteins are mainly associated with other important biological processes such as hyphal growth, biofilm formation, cell wall organization and cell adhesion. Identification of proteins involved in the processes related to pathogen's interaction with the host and other organisms suggests that those proteins are exposed to the extracellular environment and serve as initial points to interact with the host cells (e.g. adhesion, biofilm formation and recognition). Interspecies interaction between organisms generally includes interaction with hosts (HSP70, ADH1, PGK1, GPM1 and FBA) and symbiosis (encompassing mutualism through parasitism), which consists the response to defenses of other organism (through evasion or tolerance of defenses of other organism, MNT1), pathogenesis (CSH1, ECM33, PHR2, HSP90, BGL2, MNT1, ICL1, LPD1, MP65, BMH1), modification of morphology or physiology of other organism ( through induction by symbiont of host defense response, SSB1, HSP70, ACO1, TPI1, ADH1, MET6, CDC19, ENO1, PGK1, MP65, FBA1), adhesion to other organism (through adhesion to host, BGL2, PRA) and movement in environment of other organisms (through entry to host, ENO1). Also 21 (CSH1, SSB1, HSP70, ECM33, PHR2, HSP90, ALS4, TEF1, BGL2, PRA1, TPI1, ADH1, TSA1, MET6, CDC19,

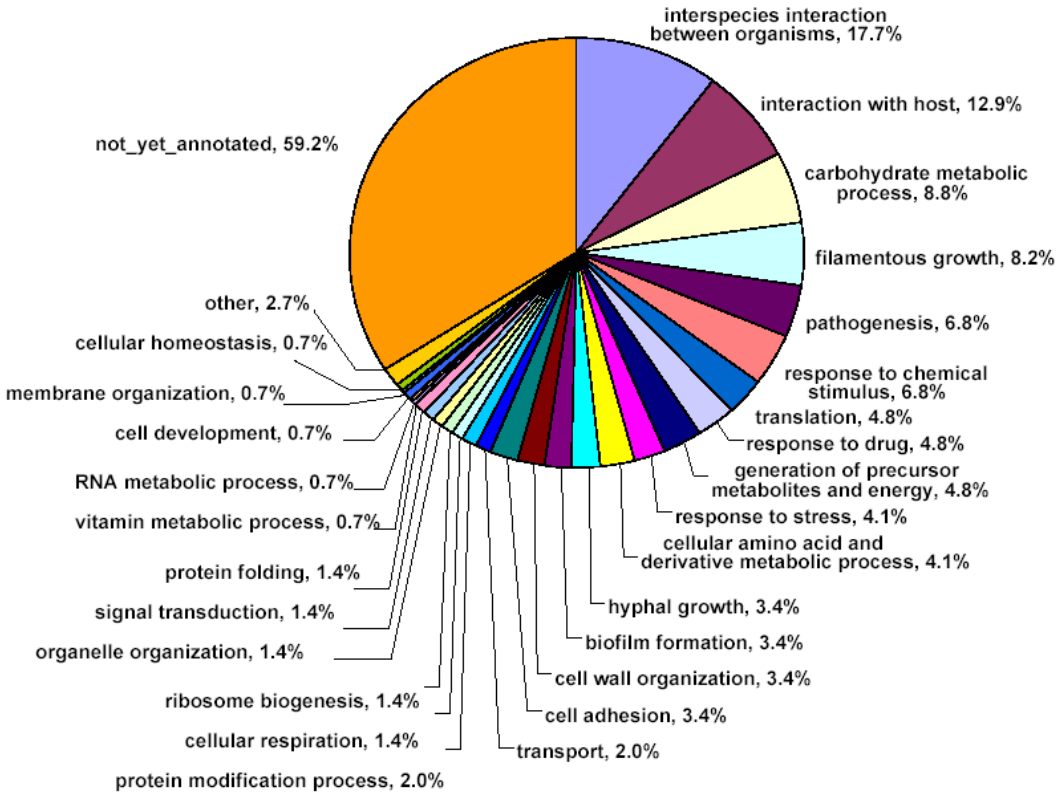



**Figure 4.6.** Process classification of identified cell wall proteins based on GO Slim Mapper at CGD. Outcalled figures represent the percentage of the number of classified proteins involved in the corresponding process among the total annotated proteins. Same proteins may be classified into different processes.

ENO1, PGK1, GPM1, MP65, BMH1, FBA1) out of 26 proteins associated with interactions with organisms have been previously detected in the cell wall. It also should be noted that it is common that one protein can be found in different biological processes. For example, CSH1 plays roles in pathogenesis (77), biofilm formation (58), and cell adhesion (78). ADH1 was classified into processes such as carbohydrate metabolism (85), generation of precursor metabolites and energy and biofilm formation (53) in addition to its involvement in the interaction with the host (17). MNT1 (54) was involved in interacting with the host (9), cell wall organization, filamentous growth and cell wall mannoprotein biosynthetic process. In addition to their roles in various interaction processes, ADH1, CDC19, ENO1, FBA1, GPM1, PGK1, TPI1, TSA1 are also involved in carbohydrate metabolic process (13). BGL2, BMH1, CSH1, ECM33, HSP90, ICL1, LPD1, MNT1, MP65, PHR2 were assigned as cell wall proteins that play important roles in *C. albicans* pathogenesis. MP65 was extensively studied for its capability to induce immune responses in humans (27, 28). ADH, FBA1, PGK1, TEF1, GMP1, TSA1 were among the eight cell wall proteins which had been shown to have the capability of binding to human plasminogen (17, 67). Recent work by Cutler and co-workers (91) demonstrated that peptide segments from FBA1, ENO1, GAP1, MET6, PGK1 could be conjugated to β-mannan trisaccharide [β-$(Man)_3$] to produce novel glycopeptide vaccines against candidiasis in mice. The high effectiveness of protection induced by β-$(Man)_3$-FBA1 and β-$(Man)_3$-MET6 suggested that FBA1 and MET6 may be potent immunological proteins presented at the cell surface of *C. albicans*. Thirteen proteins are associated with the second largest process-carbohydrate metabolic process (based on their cytosolic functions): SRB1 in GDP-mannose biosynthesis; ENO1, PGK1,GPM1 and PCK1 in gluconeogenesis; ADH1,CDC19, ENO1, PGK1, GPM1 in glycolysis; MLS1, ICL1 in glyoxylate cycle; MNT1 in N-glycan processing; SRB1 in protein

amino acid glycosylation; MNT1 in protein amino acid O-linked glycosylation and MP65 in cellular glucan metabolism. Again, new cell wall functions of those well known cytosolic proteins have been being discovered and showed to be different from their primary roles in the carbohydrate metabolic pathway (59).

Heat shock protein and chaperon proteins (HSP90, HSP70, HSP60, SSA1, SSA4, SSB1, SSC1, KAR2) are well known for their function to repair damaged proteins as a result of certain stress to the cell (e.g. drugs, heat and toxins) (37) (89), to maintain the integrity of the cell (42), and for their high immunogenicity in inducing host immune response in *C. albicans* pathogenesis. (46) Classification of those proteins indicated that also play other roles in other complicated biological processes. For example, HSP70, HSP90 and SSB1 were all classified as proteins interacting with host for their roles in response to the ligands released by other organisms. Additionally, HSP70 was identified as a cell envelope binding protein for human salivary histatin 5 (38). Interestingly, HSP70 led to enhanced systemic candidasis in murine model due to its high immunogenicity. (10) Moreover, HSP70 proteins are also involved in many facets of proteins regulation including their folding and trafficking across the membrane (16). Also, Work by Piper (62) and Nombela's (65) group suggested that HSP90 was responsive to pheromone in *C. albicans* infections and SSB1 was an immunoreactive protein in murine model, which indicated that these proteins somehow interacted with the host. Interestingly, the OG didn't classify any heat shock proteins into cell adhesion process while HSP70 family proteins were viewed as major adhesins in a recent review article (13). Notably, for the first time we detected the Dnak-type molecular chaperone (CA 0915), which was encoded by a gene about 50% identical to all eukaryotic HSP70s. However, very little information of this HSP can be found in literature. The

low abundance of this protein in the cell wall might be the main reason that it had never been detected before.

Proteins involved in cell adhesion process also attract much attention from investigators who want to gain better understanding on the mechanism of the host-pathogen interactions. Proteins ALS4 (19), CSH1 (54, 78), MNT1 (54), MP65 (72), PRA1 (80) were classified in this process. As mentioned early, this GO classification does not include ECM33, HSP70 family proteins which, however, were usually considered as adhesins (13).

### 4.4.5 Monitoring putative cell wall proteins during morphogenesis of *C. albicans*

A SILAC based quantification approach was developed to monitor SDS-DTT extractable cell wall associated proteins during the transition of *C. albicans* from yeast and hyphae. Isotopic labeled lysine was incorporated into proteins synthesized during hyphae growth. To avoid potential cytosolic contaminations, it was first tried to label the hyphae cell surface proteins by sulfo-NHS-LC-biotin. However, the amount of biotinylated proteins extracted was too low for MS analysis, and such very low labeling efficiency was probably due to the sever aggregation of hyphae cells in PBS buffer. A classical approach to prepare cell wall pellet (66) (20) was employed in our study. Extensive washing of the cell wall fraction by gradient concentration of NaCl would significantly reduce cytosolic contamination; even though membrane associated proteins would co-precipitated with cell debris as well. Hot SDS-DTT extracted CWPs were first separated by 1D SDS-PAGE to reduce sample complexity and to increase sequence coverage. Fifteen major bands for each biological sample were cut for in-gel digestion followed by LC-MS and bioinformatics analyses of the tryptic peptides. A comparative analysis of those identified

proteins was performed by using Pepquan program (incorporated in BioWorks 3.1, Thermo Electron). The results indicted that ninety two proteins were up-regulated and twenty four proteins were down-regulated during the morphogenesis as shown in Table 4.2. For instance, ATP1 was slightly over-expressed in hyphae (in Figure 4.7) while expression level of ADH1 was significantly suppressed after the germination (Figure 4.8) as evidenced by SILAC quantitative analysis of pairs of individual peptides differentially labeled by isotopic lysine.

However, the degree of the regulation (H/L) was largely between 1.5:1 and 1:1.5 with the highest average ratio (H/L) being 3.26 for ACC1 and the lowest ratio (H/L) being 0.21 for ADH1 as shown in Figure 4.9, indicating that protein expression changes during the germination *C. albicans* was not large. This result is in good agreement with a previous study (47). Among the ninety two up-regulated proteins in response to germination, seven proteins (ACS2, HSP90, PSA1, SSA, SSB1, SSE1 and TUB2.3) were previously reported as being up-regulated in a 2-D gel based comparative analysis on SDS-DTT extracts of hyphae cells wall fraction (20). A recent report from Gil's group (47) also confirmed some of those proteins (ACS2, SSB1) were up-regulated. Notably, five heat shock proteins (HSP90, SSA1, SSB1, SSC1, SSE1) were all moderately up-regulated (from 1.31 to 1.51 fold) during two hour hyphae growth induced by temperature switch from 25°C to 37°. Up-regulation of those proteins may be explained by their primary function in heat shock protection. Similar expression level of SSB1 on yeast and hyphae cells was first predicted based on Northern-blot analysis by Gozalbo's group (45) and its cell surface localization was also demonstrated by immuno-screening experiments. Previous study by Mukai and coworkers (52) also shown that mRNA levels increased by several-fold upon

**Table 4.2** Summary of identified cell wall associated proteins (SDS-DTT extract)*

| Accession No. | Protein Name | Average ratio (H/L) |
|---|---|---|
| CA5816 | CaACC1 acetyl-coenzyme-A carboxylase (by homology) | 3.26 |
| CA2470 | CaSDH12 Succinate dehydrogenase (by homology) | 2.93 |
| CA5857 | CaPCK1 phosphoenolpyruvate carboxykinase | 2.54 |
| CA5062 | CaSEC62 subunit of ER protein-translocation complex (by homology) | 2.40 |
| CA5388 | CaPET9 ADP/ATP carrier protein (by homology) | 2.36 |
| CA1513 | CaMIR1 phosphate transport protein, mitochondrial (MCF) (by homology) | 2.26 |
| CA0290 | CaPOT14 acetyl-CoA acetyltransferase (by homology) | 2.24 |
| CA4897 | CaTUB2.3 Beta-tubulin, 3-prime end | 2.22 |
| CA0959 | CaSAM2 S-adenosylmethionine synthetase 2 | 2.21 |
| CA4159 | CaALD5 aldehyde dehydrogenase (NAD+) (by homology) | 2.14 |
| CA2903 | IPF11725 unknown function | 2.07 |
| CA4502 | IPF10391 Similar to dnaJ proteins | 2.01 |
| CA6107 | CaFAS2.3f fatty-acyl-CoA synthase, alpha chain, 3-prime end | 1.98 |
| CA5992 | CaFAA4 long-chain fatty acid--CoA ligase and synthetas | 1.84 |
| CA6105 | CaFAS2.5f fatty-acyl-CoA synthase, alpha chain, 5-prime end | 1.82 |
| CA0919 | CaPOR1 mitochondrial outer membrane porin (by homology) | 1.80 |
| CA3304 | CaRPL82 60S ribosomal protein L7a.e.B (by homology) | 1.78 |
| CA3115 | CaECM33.3 cell wall biogenesis, 3-prime end (by homology) | 1.73 |
| CA2499 | CaSAP5 secreted aspartyl proteinase 5 | 1.68 |

| | | |
|---|---|---|
| **CA5255** | **CaACT1 actin (by homology)** | **1.65** |
| **CA2939** | **CaTIF1 translation initiation factor** | **1.65** |
| **CA2858** | **CaACS2 acetyl-coenzyme-A synthetase (by homology)** | **1.65** |
| **CA2998** | **CaLPD1 dihydrolipoamide dehydrogenase (by homology)** | **1.59** |
| **CA1564** | **CaGAD1 Glutamate decarboxylase (by homology)** | **1.56** |
| **CA5986** | **CaINO1 myo-inositol-1-phosphate synthase** | **1.56** |
| **CA1598** | **CaSES1 seryl-tRNA synthetase (by homology)** | **1.55** |
| **CA0622** | **IPF14171 unknown function** | **1.55** |
| **CA2065** | **CaQCR2 Ubiquinol--cytochrome-c reductase 40KD chain II** | **1.55** |
| **CA4242** | **CaSEC26 beta chain of secretory vesicles coatomer comp** | **1.55** |
| **CA3453** | **CaSEC23 Component of COPII coat (by homology)** | **1.53** |
| **CA5095** | **CaARF21 GTP-binding protein of the ARF family (by homology)** | **1.52** |
| **CA1911** | **CaSSE1 heat shock protein of HSP70 family (by homology)** | **1.51** |
| **CA5549** | **CaERG13 3-hydroxy-3-methylglutaryl coenzyme A synthase** | **1.51** |
| **CA4959** | **CaHSP90 heat shock protein** | **1.50** |
| **CA3081** | **CaEFT3 translation elongation factor 3** | **1.50** |
| **CA3052** | **CaDPS1 aspartyl-tRNA synthetase (by homology)** | **1.50** |
| **CA2400** | **CaSUP45 Translational release factor (by homology)** | **1.50** |
| **CA0824** | **CaGPD2 Glycerol 3-phosphate dehydrogenase (by homology)** | **1.50** |
| **CA0838** | **CaTFP1 vacuolar ATPase subunit (by homology)** | **1.50** |
| CA4179 | CaTOM20 mitochondrial outer membrane import receptor s | 1.47 |
| CA3208 | CaPSA1 GDP-mannose pyrophosphorylase | 1.45 |

| | | |
|---|---|---|
| CA2764 | CaQCR7 ubiquinol--cytochrome-c reductase subunit 7 (by homology) | 1.44 |
| CA4362 | CaATP2 F1F0-ATPase complex, F1 beta subunit (by homology) | 1.43 |
| CA4301 | IPF2593 amino acid-tRNA ligase homolog (by homology) | 1.42 |
| CA1423 | IPF14247 unknown function | 1.42 |
| CA4259 | CaCDC10 cell division control protein | 1.42 |
| CA0524 | CaCYT12 cytochrome-c1 (by homology) | 1.41 |
| CA0362 | CaTEF1 translation elongation factor eEF1 alpha-A chain | 1.40 |
| CA4060 | CaGRS1 glycine-tRNA ligase (by homology) | 1.40 |
| CA2457 | CaMCR1 NADH-cytochrome-b5 reductase (by homology) | 1.40 |
| CA3261 | CaTYS1 tyrosyl-tRNA synthetase by homology | 1.39 |
| CA2001 | IPF17074 unknown function | 1.38 |
| CA4783 | IPF3358 ubiquinol-cytochrome-c reductase (by homology) | 1.38 |
| CA2857 | CaSSA1 Heat shock protein of HSP70 family | 1.38 |
| CA3760 | CaERG26 C-3 sterol dehydrogenase (C-4 decarboxylase) (by homology) | 1.38 |
| CA4498 | IPF11315 unknown function | 1.38 |
| CA2332 | IPF4792 unknown Function | 1.37 |
| CA4986 | CaSER33 Phosphoglycerate dehydrogenase (by homology) | 1.37 |
| CA5051 | IPF3709 unknown function | 1.37 |
| CA3534 | CaSSB1 heat shock protein 70 | 1.36 |
| CA3978 | CaCPR3 cyclophilin (peptidylprolyl isomerase), mitochondrial (by homology) | 1.36 |
| CA5270 | CaSEC18.5f vesicular fusion protein by homology, 5-prime end | 1.33 |

| | | |
|---|---|---|
| CA1504 | CaARF3 GTP-binding protein of the ARF family (by homology) | 1.33 |
| CA2708 | CaRPS6A ribosomal protein S6 (by homology) | 1.32 |
| CA4474 | CaSSC1 Mitochondrial heat shock protein 70-related pro | 1.31 |
| CA2811 | CaRPS10 Ribosomal protein 10 | 1.30 |
| CA1273 | CaCCT4 Component of chaperonin-containing T-complex | 1.30 |
| CA3499 | CaRPG1 Translation initiation factor eIF3 (by homology) | 1.30 |
| CA2810 | CaEFT2 translation elongation factor 2 | 1.28 |
| CA0408 | CaSAC6.5f actin filament bundling protein, fimbrin, 5- | 1.28 |
| CA1565 | CaEGD1 GAL4 DNA-binding enhancer protein (by homology) | 1.26 |
| CA4412 | CaPDA1 Pyruvate dehydrogenase alpha chain (by homology) | 1.26 |
| CA4457 | CaATP1.exon3 F1F0-ATPase complex, F1 alpha subunit, exon 3 | 1.25 |
| CA3018 | CaSAH1 S-adenosyl-L-homocysteine hydrolase by homology | 1.25 |
| CA4699 | IPF4220 similar to Saccharomyces cerevisiae Pep8p invo | 1.24 |
| CA6093 | CaARF22 GTP-binding protein of the ARF family (by homology) | 1.22 |
| CA1189 | CaRPL24A ribosomal protein L24 (by homology) | 1.22 |
| CA5343 | CaRPL4B Ribosomal protein L4B (by homology) | 1.21 |
| CA4016 | CaGFA1 glutamine:fructose-6-phosphateamidotransferase | 1.20 |
| CA5796 | CaGLO3 zinc finger protein | 1.17 |
| CA3307 | CaRPL2.3 ribosomal protein L8, 3-prime end (by homology) | 1.15 |
| CA4181 | CaHOM6 homoserine dehydrogenase (by homology) | 1.14 |
| CA1751 | CaARO3.exon2 3-deoxy-D-arabinoheptulosonate-7-phosphat | 1.13 |
| CA1945 | CaMSS116 RNA helicase of the DEAD box family (by homology) | 1.13 |

| | | |
|---|---|---|
| CA4088 | CaHMO1 High-mobility protein 1 by homology | 1.13 |
| CA2956 | CaEGD2 Nascent polypeptide associated complex protein a | 1.13 |
| CA4804 | CaARO8 aromatic amino acid aminotransferase I (by homology) | 1.12 |
| CA0697 | CaVTC4 putative polyphosphate synthetase (by homology) | 1.10 |
| CA0632 | CaRPS5 ribosomal protein S5.e (by homology) | 1.08 |
| CA4954 | CaTOM40 mitochondrial import receptor chain TOM40 (by homology) | 1.05 |
| CA1931 | IPF3905 similar to Saccharomyces cerevisiae Sec21p coat | 1.04 |
| CA2720 | IPF7366 Arginyl-tRNA synthetase | 1.03 |
| **CA4765** | **CaADH1 alcohol dehydrogenase (by homology)** | **0.21** |
| **CA1691** | **CaPGK1 Phosphoglycerate kinase** | **0.21** |
| **CA3923** | **CaADH2 alcohol dehydrogenase I (by homology)** | **0.22** |
| **CA5892** | **CaGAP1 Glyceraldehyde-3-phosphate dehydrogenase** | **0.30** |
| **CA0127** | **CaHXK2.3f hexokinase II, 3-prime end (by homology)** | **0.41** |
| **CA3483** | **CaCDC19 pyruvate kinase (by homology)** | **0.43** |
| **CA4671** | **CaGPM1 phosphoglycerate mutase (by homology)** | **0.44** |
| **CA3874** | **CaENO1 Enolase I (2-phosphoglycerate dehydratase)** | **0.45** |
| **CA5613** | **IPF525 unknown function** | **0.48** |
| **CA5991** | **IPF690.3f NADH dehydrogenase (ubiquinone) 78K chain precursor, 3-prime end (by homology)** | **0.48** |
| **CA1927** | **CaYBN5 Putative purine nucleotide-binding protein (by homology)** | **0.55** |
| **CA2474** | **CaPDC11 Pyruvate decarboxylase (by homology)** | **0.64** |

| | | |
|---|---|---|
| **CA5239** | **CaGND1 6-phosphogluconate dehydrogenase** | **0.64** |
| CA1580 | CaTRP5 tryptophan synthase (by homology) | 0.68 |
| CA3112 | CaPFK2 6-phosphofructokinase, beta subunit | 0.75 |
| CA5398 | CaSEC14 phosphatidylinositol(PI)/phosphatidylcholine(PC)transfer | 0.75 |
| CA5206 | CaGPH1 Glycogen phosphorylase (by homology) | 0.81 |
| CA3075 | IPF6105 similar to Saccharomyces cerevisiae Gua1p GMP | 0.83 |
| CA3278 | CaRPS3 Ribosomal protein S3.e (by homology) | 0.88 |
| CA4184 | CaSNZ1 stationary phase protein by homology | 0.89 |
| CA1541 | CaBGL21 endo-beta-1,3-glucanase (by homology) | 0.89 |
| CA0895 | CaSHM2 Serine hydroxymethyltransferase precursor, mito | 0.93 |
| CA3752 | CaTIF51.3 translation initiation factor eIF-5A, 3-prime end | 0.97 |
| CA4001 | CaRPL12 ribosomal protein | 0.99 |
| CA4570 | IPF9550 similar to Saccharomyces cerevisiae Osm1p osmotic growth protein | $-\infty$ |
| CA1019 | IPF13485 unknown function | $\infty$ |

*These identified cell wall associated proteins in SDS-DTT fraction were differentially regulated between hyphae (H) and yeast (L) of *C. albicans*. Protein level up-regulated or down-regulated by at least 1.5 fold were considered significant and displayed in bold.

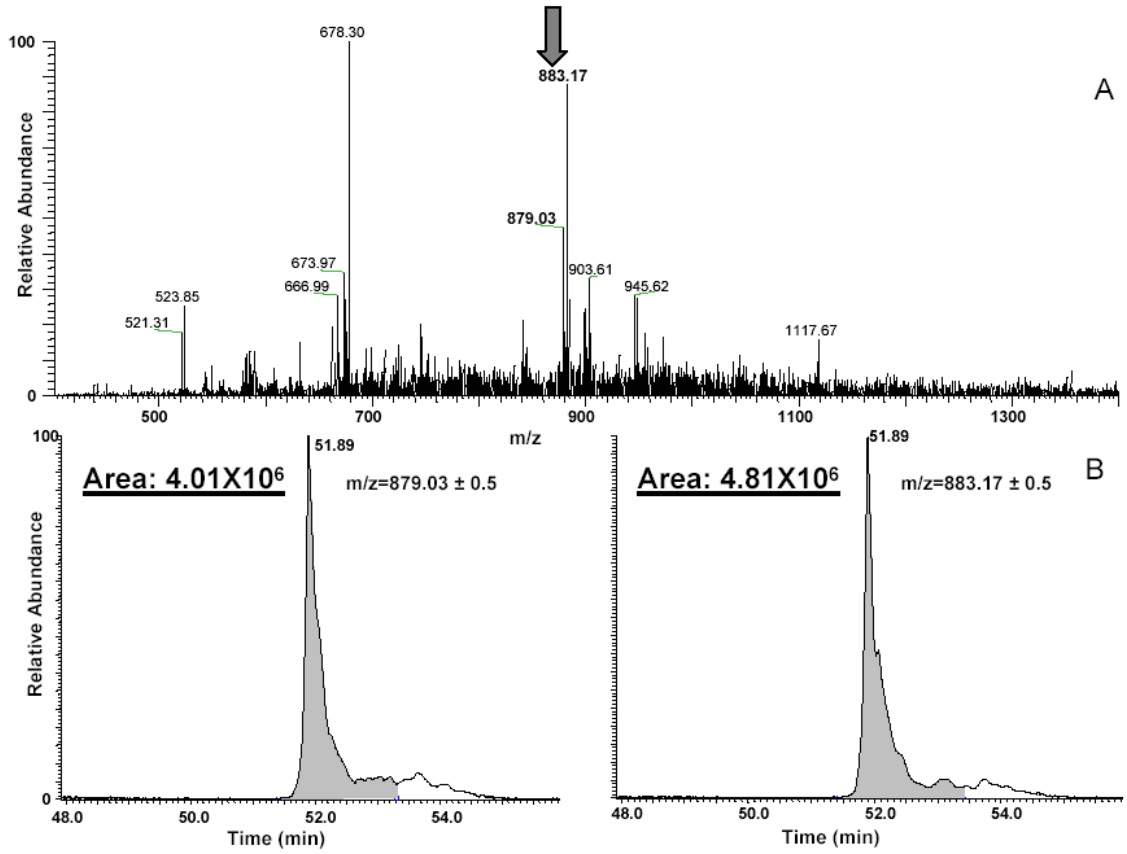



**Figure 4.7** SILAC quantitative analysis of a protein extracted by SDS-DTT from cell wall pellet of hyphae of *C. albicans*. This protein (ATP1) is preferentially accumulated in yeast form (A) Part of full scan ESI mass spectrum acquired at time of 51.89 min. A peptide doublet indicated by a solid arrow represents a pair of peptide differentially labeled by an isotopic lysine residue in their sequence. The mass-to-charge (m/z) ratio of each peptide is 4.14 Da different. (B) Single ion chromatograms of the paired peptides eluted from $C_{18}$ column at the same time window. The ratio (1:1.20) (H/L) calculated from peak area was used for average ratio calculation. Peptide sequence: Peptide precursor labeled by heavy isotopes (H): m/z 883.17; peptide precursor not labeled (L): m/z 879.03

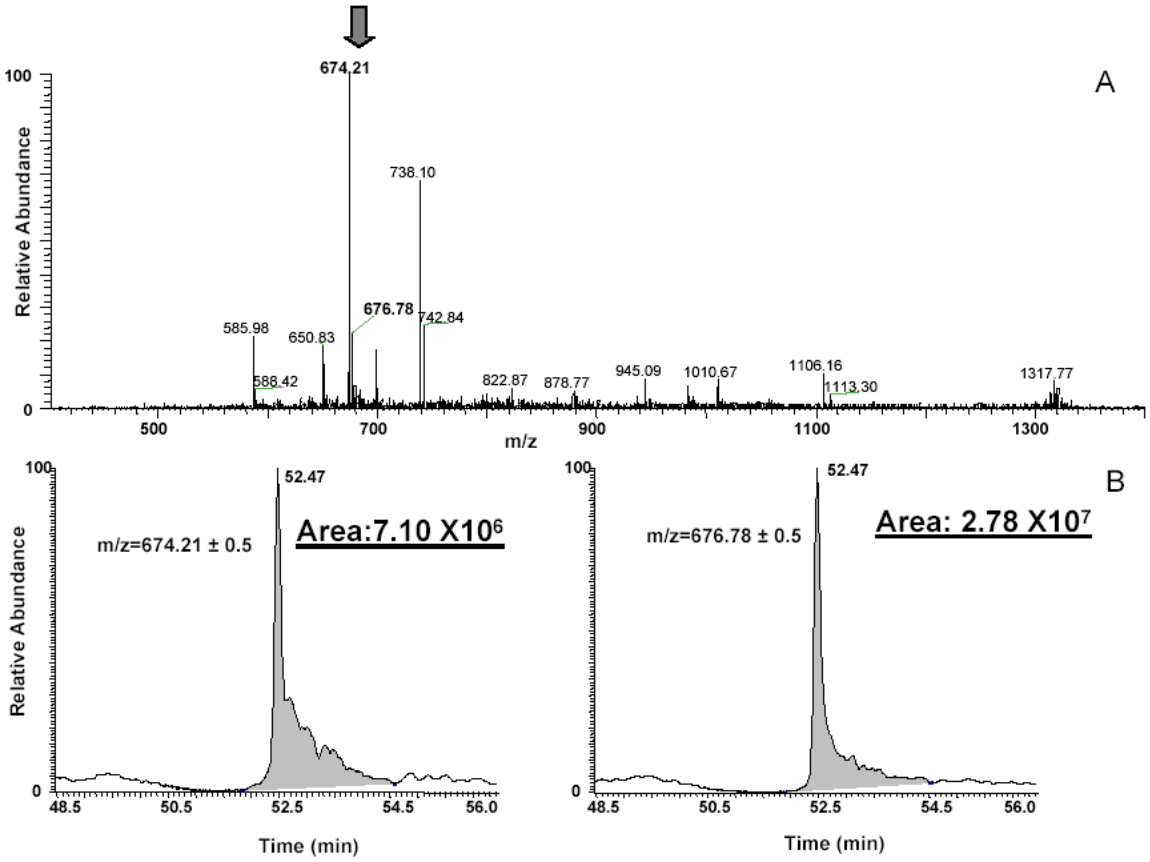



**Figure 4.8** SILAC quantitative analysis of a protein extracted by SDS-DTT from cell wall pellet of hyphae of *C. albicans*. This protein (ADH1) is preferentially accumulated in hyphal form (A) Part of full scan mass spectrum acquired through ESI at time of 52.47 minute. A peptide doublet indicated by a solid arrow represents a pair of peptide differentially labeled by an isotopic lysine residue in their sequence. The mass-to-charge (m/z) ratio of each peptide is separated by 2.57 Da. (B) Single ion chromatograms of the paired peptides eluted from $C_{18}$ column at the same time window. The ratio (4.16:1) (H/L) calculated from peak area was used for average ratio calculation. Peptide sequence: Peptide precursor labeled by heavy isotopes (H): m/z 676.78; peptide precursor not labeled (L): m/z 674.21

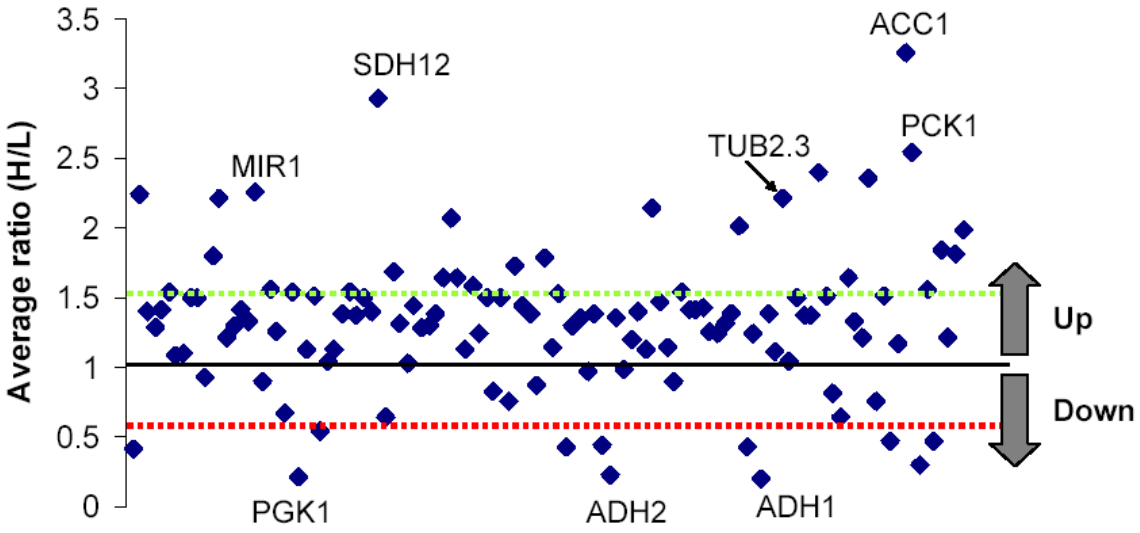



**Figure 4.9.** The overall graphic distribution of regulations of cell wall associated proteins during the yeast cells morphogenesis. The detailed regulation values of proteins (dots) were listed in Table 4.2.

induction to hyphae growth at 37°. However, work by Brwon's group (84) shown that HSP90 mRNA levels had no detectable increase as a result of germination. Our results suggest that there might be an opposite correlation of expression level between mRNA and protein. Moreover, a recent study published by Cowen's group indicated that HSP90 played pivotal roles in regulating temperature-dependent morphogenesis by relieving HSP90-meidated repression of other relevant proteins through Ras1-PKA signaling cascade. (73) Compared to the larger set of up-regualated proteins during the morphogenesis, a small group of downregulated proteins was found to be mainly consisting of glycolysis enzymes (ADH1, PGK1, PFK2, GPM1, ENO1, CDC19). As mentioned above, the real functions of those moonlighting proteins in the cell wall remain enigmatic. Intriguingly, four human plasmin(ogen) binding proteins ( ENO1, ADH1, PGK1, GPM1) were found among those down-regulated proteins, which is consistent with a previous study (17) indicating that *C. albicans* cells bound to plasmin did not increase damages of endothelial cells. Hyphae cells generally efficiently penetrate host cells through a combination of factors that focuses hydraulic pressure at the hyphae tip via the activities of an array of secreted hydrolases (30); and it is further evidenced by our results that some hydrolase activity related proteins (ARF1, CEF3, HSP90, MSI3, SAP5, TIF andTFP1) are significantly up-regulated. These hydrolases activity related proteins account for the highest function percentage (17.9% ) as classified at CGD. Interestingly, work by Cutler's group showed that only moderate protection against candidiasis was achieved while mice were immunolized by synthetic vaccines containing peptide fragments from down-regulated GAP1 and ENO1 (91). Their finding might suggest that proteins highly down-regulated during the morphogenesis are presumably less immunologically relevant in eliciting protective antibodies from the host. Moreover, we tentatively propose IPF13485 as hyphae specific and IPF9550 as yeast specific. It should be

noted that despite some agreement with previous reports for both up and down regulated proteins (20) (47), our results also include proteins that showed opposite trend of regulation changes during the germination. For instance, ATP2, EGD2 were identified to be up-regulated in our study, but were reported as being down-regulated in a previous study (20).

Proteins with significant regulation change (out of the range between 1.5:1 and 1:1.5 as indicated in bold type in Table 4.2 ) were classified according to their biological process by GO at CGD (4). Most of these proteins could not be mapped due to lack of evidence reported in the literature. Our biotinylation method confirmed the localization to cell surface of those proteins: ACC1, ACT1, ADH1, ADH2, ALD5, CDC19, CEF3, ECM33.3, ENO1, FAS2.5f, GAD1, GAP1, GND1, GPM1, HSP90, HXK2.3f, IPF14171, LPD1, PCK1, PDC, PET9, PGK1, POT14, QCR2, RPL82. It was also shown in our study that ENO1 may be highly abundant in yeast cell wall because of its largest sequence coverage (55.4%) among all identified biotinylated proteins. Given that ENO1 was only moderately down-regulated (H/L, 1:2.23) and its low sequence coverage (23.41%) even after 1D SDS-PAGE gel separation, we speculated that our biotinylation labeling approach was biased toward proteins located on cell surface. The largest group for up-regulated proteins (SAP5, FAS2, ECM33, LPD1, HSP90) belong to the process of interspecies interaction between organisms and pathogenesis. Likewise down-regulated proteins (ENO1 ADH1 PGK1 GPM1 CDC19) were also classified in interspecies interaction between organisms (including with the host). Lee's group (14) first detected SAP4-7 proteins in culture supernatants from hyphae cells and monitored their dynamic expression levels during different germination stages under various pH conditions. The pathogenicity of SAPs from various *C. albicans* strains was reviewed by Cassone (18). A recent review paper from Hube's group summarized the roles

that secreted aspartyl proteinases (SAPs) family played in *C. albicans*' interaction with host and pointed to their direct contribution in its pathogenicity (56). Detection of SAP5 in our study indicated that certain secreted proteins may also be non-covalently associated with cell wall and therefore susceptible to SDS-DTT extraction. Contrary to the claim of Pinto's group (20) that FAS2 and LPD1was only detected from SDS-DTT resisted hyphae cell wall fraction, our result showed that both proteins were up-regulated in hyphae cells and both also presented in yeast and were SDS-DTT extractible. This suggests that our approach was quite sensitive in the detection of low abundant proteins. Moreover, LPD1 were identified as labeled protein by sulfo-NHS-SS-biotin and FAS2 were also detected based on one peptide modified by I-biotin. FAS2 has also been found in membrane-coated organelles and known for its essential role in the survival of the yeast in non-exogenous fatty acid environment (79). Work by Nguyen's group (15) suggested that LPD1 could be released from cytosol to the cell surface after the cell being attacked and damaged by the human immune system, and serve as an immunogenic antigen. Its immunogenicity was confirmed in murine model recently by Gil's group (24).

## 4.5 Conclusions

In summary, one hundred and forty eight CWPs from *C. albicans* yeast cell were identified based on the detection of biotinylated peptides and observation of characteristic fragments originated from the biotin moiety on $MS^2$ spectra of labeled peptide. The method developed in our study, to our knowledge, is the most accurate approach for efficient identification of CWPs and cell wall associated proteins of *C. albicans.* Fifty one of one hundred forty eight proteins were never reported before in the literature. Each approach involved with one of the three labeling reagents has its own protein identifications, while some other protein identifications

were shared by at least two labeling approaches even though the detected labeled peptide might not be necessarily the same. Moreover, two hundred and thirty five cell wall-associated proteins were identified from cell wall pellet of hyphae of *C. albicans* after SDS-DTT extraction. Analyzed by this SILAC based comparative proteomic approach, ninety two proteins were upregulated and twenty four proteins were downregulated during the transition from yeast to hyphae growth. IPF 9550 (CA4570) and IPF 13485 (CA1019) were tentatively assigned as yeast and hyphae specific protein respectively. Overall the expression level change ranged from 3.26 fold upregulation and 0.21 fold downregulation after the morphological transition from yeast to hyphae. Thirty nine proteins were significantly upregulated (more than 1.5 fold increase) and thirteen proteins were significantly downregualted (more than 1.5 fold decrease) in hyphae cells. According to the functional classification obtained from CGD, seven upregulated proteins were related to hydrolase enzymatic activities which might be promoted in cell wall during germination, while four major plasmin(ogen) binding proteins were downregulated possibly to facilitate germ tube elongation. This extensive study on cell surface proteome analysis of *C. albicans* will help elucidate the mechanism of host-pathogen interaction and guide in protective vaccine development in future.

## References:


1. **Alban, A., S. O. David, L. Bjorkesten, C. Andersson, E. Sloge, S. Lewis, and I. Currie.** 2003. A novel experimental design for comparative two-dimensional gel analysis: Two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics **3:**36-44.
2. **Alloush, H. M., J. L. LopezRibot, B. J. Masten, and W. L. Chaffin.** 1997. 3-Phosphoglycerate kinase: A glycolytic enzyme protein present in the cell wall of *Candida albicans*. Microbiology-Uk **143:**321-330.

3. **Angiolella, L., M. Facchin, A. Stringaro, B. Maras, N. Simonetti, and A. Cassone.** 1996. Identification of a glucan-associated enolase as a main cell wall protein of *Candida albicans* and an indirect target of lipopeptide antimycotics. Journal of Infectious Diseases **173:**684-690.
4. **Arnaud, M. B., M. C. Costanzo, M. S. Skrzypek, P. Shah, G. Binkley, C. Lane, S. R. Miyasato, and G. Sherlock.** 2007. Sequence resources at the *Candida* genome database. Nucleic Acids Research **35:**D452-D456.
5. **Bhattacherjee, V., and J. K. Bhattacharjee.** 1999. Characterization of a double gene disruption in the LYS2 locus of the pathogenic yeast, *Candida albicans*. Medical Mycology **37:**411-417.
6. **Bickel, U., Y. S. Kang, and W. M. Pardridge.** 1995. *In-vivo* cleavability of A disulfide-based chimeric opioid peptide in rat-brain. Bioconjugate Chemistry **6:**211-218.
7. **Borisov, O. V., M. B. Goshe, T. P. Conrads, V. S. Rakov, T. D. Veenstra, and R. D. Smith.** 2002. Low-energy collision-induced dissociation fragmentation analysis of cysteinyl-modified peptides. Analytical Chemistry **74:**2284-2292.
8. **Boyle, E. I., S. A. Weng, J. Gollub, H. Jin, D. Botstein, J. M. Cherry, and G. Sherlock.** 2004. GO:TermFinder - open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics **20:**3710-3715.
9. **Brand, A., J. D. Barnes, K. S. Mackenzie, F. C. Odds, and N. A. R. Gow.** 2008. Cell wall glycans and soluble factors determine the interactions between the hyphae of *Candida albicans* and *Pseudomonas aeruginosa*. Fems Microbiology Letters **287:**48-55.
10. **Bromuro, C., R. La Valle, S. Sandini, F. Urbani, C. M. Ausiello, L. Morelli, C. F. D'Ostiani, L. Romani, and A. Cassone.** 1998. A 70-kilodalton recombinant heat shock protein of *Candida albicans* is highly immunogenic and enhances systemic murine candidiasis. Infection and Immunity **66:**2154-2162.
11. **Carman, A. J., S. Vylkova, and M. C. Lorenz.** 2008. Role of acetyl coenzyme A synthesis and breakdown in alternative carbon source utilization in *Candida albicans*. Eukaryotic Cell **7:**1733-1741.
12. **Castillo, L., E. Caivo, A. I. Martinez, J. Ruiz-Herrera, E. Valentin, J. A. Lopez, and R. Sentandreu.** 2008. A study of the *Candida albicans* cell wall proteome. Proteomics **8:**3871-3881.
13. **Chaffin, W. L.** 2008. *Candida albicans* cell wall proteins. Microbiology and Molecular Biology Reviews **72:**495-+.
14. **Chen, Y. C., C. C. Wu, W. L. Chung, and F. J. S. Lee.** 2002. Differential secretion of Sap4-6 proteins in Candida albicans during hyphae formation. Microbiology-Sgm **148:**3743-3754.
15. **Cheng, S. J., C. J. Clancy, M. A. Checkley, M. Handfield, J. D. Hillman, A. Progulske-Fox, A. S. Lewin, P. L. Fidel, and M. H. Nguyen.** 2003. Identification of *Candida albicans* genes induced during thrush offers insight into pathogenesis. Molecular Microbiology **48:**1275-1288.
16. **Craig, E., P. J. Kang, and W. Boorstein.** 1990. A review of the role of 70 Kda heat-shock proteins in protein translocation across membranes. Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology **58:**137-146.

17. **Crowe, J. D., I. K. Sievwright, G. C. Auld, N. R. Moore, N. A. R. Gow, and N. A. Booth.** 2003. *Candida albicans* binds human plasminogen: identification of eight plasminogen-binding proteins. Molecular Microbiology **47:**1637-1651.
18. **De Bernardis, F., P. A. Sullivan, and A. Cassone.** 2001. Aspartyl proteinases of *Candida albicans* and their role in pathogenicity. Medical Mycology **39:**303-313.
19. **de Groot, P. W. J., A. D. de Boer, J. Cunningham, H. L. Dekker, L. de Jong, K. J. Hellingwerf, C. de Koster, and F. M. Klis.** 2004. Proteomic analysis of *Candida albicans* cell walls reveals covalently bound carbohydrate-active enzymes and adhesins. Eukaryotic Cell **3:**955-965.
20. **Ebanks, R. O., K. Chisholm, S. McKinnon, M. Whiteway, and D. M. Pinto.** 2006. Proteomic analysis of *Candida albicans* yeast and hyphal cell wall and associated proteins. Proteomics **6:**2147-2156.
21. **Edwards, S. R., R. Braley, and W. L. Chaffin.** 1999. Enolase is present in the cell wall of *Saccharomyces cerevisiae*. Fems Microbiology Letters **177:**211-216.
22. **Eigenheer, R. A., Y. J. Lee, E. Blumwald, B. S. Phinney, and A. Gelli.** 2007. Extracellular glycosylphosphatidylinositol-anchored mannoproteins and proteases of *Cryptococcus neoformans*. Fems Yeast Research **7:**499-510.
23. **Elia, G.** 2008. Biotinylation reagents for the study of cell surface proteins. Proteomics **8:**4012-4024.
24. **Fernandez-Arenas, E., G. Molero, C. Nombela, R. Diez-Orejas, and C. Gil.** 2004. Low virulent strains of *Candida albicans*: Unravelling the antigens for a future vaccine. Proteomics **4:**3007-3020.
25. **Garcia-Sanchez, S., S. Aubert, I. Iraqui, G. Janbon, J. M. Ghigo, and C. d'Enfert.** 2004. Candida albicans biofilms: a developmental state associated with specific and stable gene expression patterns. Eukaryotic Cell **3:**536-545.
26. **GilNavarro, I., M. L. Gil, M. Casanova, J. E. Oconnor, J. P. Martinez, and D. Gozalbo.** 1997. The glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase of *Candida albicans* is a surface antigen. Journal of Bacteriology **179:**4992-4999.
27. **Gomez, M. J., B. Maras, A. Barca, R. La Valle, D. Barra, and A. Cassone.** 2000. Biochemical and immunological characterization of MP65, a major mannoprotein antigen of the opportunistic human pathogen *Candida albicans*. Infection and Immunity **68:**694-701.
28. **Gomez, M. J., A. Torosantucci, S. Arancia, B. Maras, L. Parisi, and A. Cassone.** 1996. Purification and biochemical characterization of a 65-kilodalton mannoprotein (MP65), a main target of Anti-Candida cell-mediated immune responses in humans. Infection and Immunity **64:**2577-2584.
29. **Gow, N. A. R.** 2009. Fungal Morphogenesis: Some Like It Hot. Current Biology **19:**R333-R334.
30. **Gow, N. A. R., A. J. P. Brown, and F. C. Odds.** 2002. Fungal morphogenesis and host invasion. Current Opinion In Microbiology **5:**366-371.
31. **Gygi, S. P., G. L. Corthals, Y. Zhang, Y. Rochon, and R. Aebersold.** 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proceedings of the national academy of sciences of the United States of America **97:**9390-9395.
32. **Ishiguro, A., M. Homma, S. Torii, and K. Tanaka.** 1992. Identification of *candida albicans* antigens reactive with immunoglobulin-E antibody of human sera. Infection and Immunity **60:**1550-1557.

33. **Jeng, H. W., A. R. Holmes, and R. D. Cannon.** 2005. Characterization of two *Candida albicans* surface mannoprotein adhesins that bind immobilized saliva components. Medical Mycology **43:**209-217.
34. **Julka, S., and F. Regnier.** 2004. Quantification in proteomics through stable isotope coding: A review. Journal of Proteome Research **3:**350-363.
35. **Kandasamy, R., G. Vediyappan, and W. L. Chaffin.** 2000. Evidence for the presence of Pir-like proteins in *Candida albicans*. Fems Microbiology Letters **186:**239-243.
36. **Klis, F. M., M. de Jong, S. Brul, and P. W. J. de Groot.** 2007. Extraction of cell surface-associated proteins from living yeast cells. Yeast **24:**253-258.
37. **Lavalle, R., C. Bromuro, L. Ranucci, H. M. Muller, A. Crisanti, and A. Cassone.** 1995. Molecular-cloning and expression of a 70-kilodalton heat-shock protein of *Candida albicans*. Infection and Immunity **63:**4039-4045.
38. **Li, X. W. S., M. S. Reddy, D. Baev, and M. Edgerton.** 2003. *Candida albicans* Ssa1/2p is the cell envelope binding protein for human salivary histatin 5. Journal of Biological Chemistry **278:**28553-28561.
39. **Lim, D. B., P. Hains, B. Walsh, P. Bergquist, and H. Nevalainen.** 2001. Proteins associated with the cell envelope of *Trichoderma reesei*: A proteomic approach. Proteomics **1:**899-909.
40. **LopezRibot, J. L., H. M. Alloush, B. J. Masten, and W. L. Chaffin.** 1996. Evidence for presence in the cell wall of *Candida albicans* of a protein related to the hsp70 family. Infection and Immunity **64:**3333-3340.
41. **Lorenz, M. C., and G. R. Fink.** 2001. The glyoxylate cycle is required for fungal virulence. Nature **412:**83-86.
42. **Mager, W. H., and P. M. Ferreira.** 1993. Stress response of yeast. Biochemical Journal **290:**1-13.
43. **Malik, R., R. Lenobel, A. Santamaria, A. Ries, E. A. Nigg, and R. Korner.** 2009. Quantitative analysis of the human spindle phosphoproteome at distinct mitotic stages. Journal of Proteome Research **8:**4553-4563.
44. **Maneu, V., A. M. Cervera, J. P. Martinez, and D. Gozalbo.** 1996. Molecular cloning and characterization of a Candida albicans gene (EFB1) coding for the elongation factor EF-1 beta. Fems Microbiology Letters **145:**157-162.
45. **Maneu, V., A. M. Cervera, J. P. Martinez, and D. Gozalbo.** 1997. Molecular cloning of a *Candida albicans* gene (SSB1) coding for a protein related to the hsp70 family. Yeast **13:**677-681.
46. **Maresca, B., and G. S. Kobayashi.** 1994. Hsp70 in parasites - As an inducible protective protein and as an antigen. Experientia **50:**1067-1074.
47. **Martinez-Gomariz, M., P. Perumal, S. Mekala, C. Nombela, W. L. Chaffin, and C. Gil.** 2009. Proteomic analysis of cytoplasmic and surface proteins from yeast cells, hyphae, and biofilms of *Candida albicans*. Proteomics **9:**2230-2252.
48. **Martinez-Lopez, R., C. Nombela, R. Diez-Orejas, L. Monteoliva, and C. Gil.** 2008. Immunoproteomic analysis of the protective response obtained from vaccination with *Candida albicans* ecm33 cell wall mutant in mice. Proteomics **8:**2651-2664.
49. **Masuoka, J., L. N. Guthrie, and K. C. Hazen.** 2002. Complications in cell-surface labelling by biotinylation of *Candida albicans* due to avidin conjugate binding to cell-wall proteins. Microbiology-Sgm **148:**1073-1079.

50. **Matthews, R. C., J. P. Burnie, and S. Tabaqchali.** 1987. Isolation of immunodominant antigens from sera of patients with systemic candidiasis and characterization of serological response to *Candida albicans*. Journal of Clinical Microbiology **25:**230-237.
51. **Mrsa, V., T. Seidl, M. Gentzsch, and W. Tanner.** 1997. Specific labelling of cell wall proteins by biotinylation. Identification of four covalently linked O-mannosylated proteins of *Saccharomyces cerevisiae*. Yeast **13:**1145-1154.
52. **Mukai, H., T. Kuno, H. Tanaka, D. Hirata, T. Miyakawa, and C. Tanaka.** 1993. Isolation and characterization of Sse1 And Sse2, new members of the yeast Hsp70 multigene family. Gene **132:**57-66.
53. **Mukherjee, P. K., S. Mohamed, J. Chandra, D. Kuhn, S. Q. Liu, O. S. Antar, R. Munyon, A. P. Mitchell, D. Andes, M. R. Chance, M. Rouabhia, and M. A. Ghannoum.** 2006. Alcohol dehydrogenase restricts the ability of the pathogen *Candida albicans* to form a biofilm on catheter surfaces through an ethanol-based mechanism. Infection and Immunity **74:**3804-3816.
54. **Munro, C. A., S. Bates, E. T. Buurman, H. B. Hughes, D. M. MacCallum, G. Bertram, A. Atrih, M. A. J. Ferguson, J. M. Bain, A. Brand, S. Hamilton, C. Westwater, L. M. Thomson, A. J. P. Brown, F. C. Odds, and N. A. R. Gow.** 2005. Mnt1p and Mnt2p of *Candida albicans* are partially redundant alpha-1,2-mannosyltransferases that participate in O-linked mannosylation and are required for adhesion and virulence. Journal of Biological Chemistry **280:**1051-1060.
55. **Muroi, M., T. Ohnishi, and K. Tanamoto.** 2002. Regions of the mouse CD14 molecule required for Toll-like receptor 2-and 4-mediated activation of NF-kappa B. Journal of Biological Chemistry **277:**42372-42379.
56. **Naglik, J., A. Albrecht, O. Bader, and B. Hube.** 2004. *Candida albicans* proteinases and host/pathogen interactions. Cellular Microbiology **6:**915-926.
57. **Navarro-Garcia, F., M. Sanchez, C. Nombela, and J. Pla.** 2001. Virulence genes in the pathogenic yeast *Candida albicans*. Fems Microbiology Reviews **25:**245-268.
58. **Nobile, C. J., J. E. Nett, A. D. Hernday, O. R. Homann, J. S. Deneault, A. Nantel, D. R. Andes, A. D. Johnson, and A. P. Mitchell.** 2009. Biofilm Matrix Regulation by *Candida albicans* Zap1. Plos Biology **7**.
59. **Nombela, C., C. Gil, and W. L. Chaffin.** 2006. Non-conventional protein secretion in yeast. Trends in Microbiology **14:**15-21.
60. **Nosanchuk, J. D., J. N. Steenbergen, L. Shi, G. S. Deepe, and A. Casadevall.** 2003. Antibodies to a cell surface histon-like protein protect against *Histoplasma capsulatum*. Journal of Clinical Investigation **112:**1164-1175.
61. **Ong, S. E., B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, and M. Mann.** 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular & Cellular Proteomics **1:**376-386.
62. **Panaretou, B., K. Sinclair, C. Prodromou, J. Johal, L. Pearl, and P. W. Piper.** 1999. The Hsp90 of *Candida albicans* can confer Hsp90 functions in *Saccharomyces cerevisiae*: a potential model for the processes that generate immunogenic fragments of this molecular chaperone in *C. albicans* infections. Microbiology-Sgm **145:**3455-3463.
63. **Peirce, M. J., R. Wait, S. Begum, J. Saklatvala, and A. P. Cope.** 2004. Expression profiling of lymphocyte plasma membrane proteins. Molecular & Cellular Proteomics **3:**56-65.

64. **Pitarch, A., J. Abian, M. Carrascal, M. Sanchez, C. Nombela, and C. Gil.** 2004. Proteomics-based identification of novel Candida albicans antigens for diagnosis of systemic candidiasis in patients with underlying hematological malignancies. Proteomics **4:**3084-3106.
65. **Pitarch, A., R. Diez-Orejas, G. Molero, M. Pardo, M. Sanchez, C. Gill, and C. Nombela.** 2001. Analysis of the serologic response to systemic *Candida albicans* infection in a murine model. Proteomics **1:**550-559.
66. **Pitarch, A., M. Sanchez, C. Nombela, and C. Gil.** 2002. Sequential fractionation and two-dimensional gel analysis unravels the complexity of the dimorphic fungus *Candida albicans* cell wall proteome. Molecular & Cellular Proteomics **1:**967-982.
67. **Poltermann, S., A. Kunert, M. von der Heide, R. Eck, A. Hartmann, and P. F. Zipfel.** 2007. Gpm1p is a factor H-, FHL-1-, and plasminogen-binding surface protein of *Candida albicans*. Journal of Biological Chemistry **282:**37537-37544.
68. **Ponton, J., M. J. Omaetxebarria, N. Elguezabal, M. Alvarez, and M. D. Moragues.** 2001. Immunoreactivity of the fungal cell wall. Medical Mycology **39:**101-110.
69. **Qian, Q. F., M. A. Jutila, N. Vanrooijen, and J. E. Cutler.** 1994. Elimination of mouse splenic macrophages correlates with increased susceptibility to experimental disseminated candidiasis. Journal of Immunology **152:**5000-5008.
70. **Rodriguez-Ortega, M. J., N. Norais, G. Bensi, S. Liberatori, S. Capo, M. Mora, M. Scarselli, F. Doro, G. Ferrari, I. Garaguso, T. Maggi, A. Neumann, A. Covre, J. L. Telford, and G. Grandi.** 2006. Characterization and identification of vaccine candidate proteins through analysis of the group *A Streptococcus* surface proteome. Nature Biotechnology **24:**191-197.
71. **Rosenfeld, J., J. Capdevielle, J. C. Guillemot, and P. Ferrara.** 1992. In-gel digestion of proteins for internal sequence-analysis after 1-dimensional or 2-dimensional gel-electrophoresis. Analytical Biochemistry **203:**173-179.
72. **Sandini, S., R. La Valle, F. De Bernardis, C. Macri, and A. Cassone.** 2007. The 65 kDa mannoprotein gene of *Candida albicans* encodes a putative beta-glucanase adhesin required for hyphal morphogenesis and experimental pathogenicity. Cellular Microbiology **9:**1223-1238.
73. **Shapiro, R. S., P. Uppuluri, A. K. Zaas, C. Collins, H. Senn, J. R. Perfect, J. Heitman, and L. E. Cowen.** 2009. Hsp90 orchestrates temperature-dependent *Candida albicans* morphogenesis via Ras1-PKA signaling. Current Biology **19:**621-629.
74. **Sharma, K., C. Weber, M. Bairlein, Z. Greff, G. Keri, J. Cox, J. V. Olsen, and H. Daub.** 2009. Proteomics strategy for quantitative protein interaction profiling in cell extracts. Nature Methods **6:**741-U13.
75. **Shen, H. D., K. B. Choo, H. H. Lee, J. C. Hsieh, W. L. Lin, W. R. Lee, and S. H. Han.** 1991. The 40-kilodalton allergen of *Candida albicans* is an alcohol-dehydrogenase -molecular-cloning and immunological analysis using monoclonal-antibodies. Clinical and Experimental Allergy **21:**675-681.
76. **Singh, V. K., R. K. Jayaswal, and B. J. Wilkinson.** 2001. Cell wall-active antibiotic induced proteins of *Staphylococcus aureus* identified using a proteomic approach. Fems Microbiology Letters **199:**79-84.
77. **Singleton, D. R., P. L. Fidel, K. L. Wozniak, and K. C. Hazen.** 2005. Contribution of cell surface hydrophobicity protein I (Csh1p) to virulence of hydrophobic *Candida albicans* serotype A cells. Fems Microbiology Letters **244:**373-377.

78. **Singleton, D. R., J. Masuoka, and K. C. Hazen.** 2001. Cloning and analysis of a *Candida albicans* gene that affects cell surface hydrophobicity. Journal of Bacteriology **183:**3582-3588.
79. **Smith, S.** 1994. The animal fatty-acid synthase - One gene, one polypeptide, 7 enzymes. Faseb Journal **8:**1248-1259.
80. **Soloviev, D. A., W. A. Fonzi, R. Sentandreu, E. Pluskota, C. B. Forsyth, S. Yadav, and E. F. Plow.** 2007. Identification of pH-regulated antigen 1 released from *Candida albicans* as the major ligand for leukocyte integrin alpha(M)beta(1)(2). Journal of Immunology **178:**2038-2046.
81. **Spence, J. M., and V. L. Clark.** 2000. Role of ribosomal protein L12 in gonococcal invasion of Hec1B cells. Infection and Immunity **68:**5002-5010.
82. **Staros, J. V.** 1982. N-hydroxysulfosuccinimide active esters - Bis(N-Hydroxysulfosuccinimide) esters of 2 dicarboxylic-acids are Hydrophilic, membrane-impermeant, protein cross-linkers. Biochemistry **21:**3950-3955.
83. **Sundstrom, P., and G. R. Aliaga.** 1994. A subset of proteins found in culture supernatants of *Candida albicans* includes the abundant, immunodominant, glycolytic enzyme enolase. Journal of Infectious Diseases **169:**452-456.
84. **Swoboda, R. K., G. Bertram, S. Budge, G. W. Gooday, N. A. R. Gow, and A. J. P. Brown.** 1995. Structure and regulation of the Hsp90 gene from the pathogenic fungus *Candida albicans*. Infection and Immunity **63:**4506-4514.
85. **Swoboda, R. K., G. Bertram, S. Delbruck, J. F. Ernst, N. A. R. Gow, G. W. Gooday, and A. J. P. Brown.** 1994. Fluctuations in glycolytic messenger-Rna levels during morphogenesis in *Candida albicans* reflect underlying changes in growth and are not a response to cellular dimorphism. Molecular Microbiology **13:**663-672.
86. **Swoboda, R. K., G. Bertram, H. Hollander, D. Greenspan, J. S. Greenspan, N. A. R. Gow, G. W. Gooday, and A. J. P. Brown.** 1993. Glycolytic-enzymes of *Candida albicans* are nonubiquitous immunogens during candidiasis. Infection and Immunity **61:**4263-4271.
87. **Urban, C., K. Sohn, F. Lottspeich, H. Brunner, and S. Rupp.** 2003. Identification of cell surface determinants in *Candida albicans* reveals Tsa1p, a protein differentially localized in the cell. Febs Letters **544:**228-235.
88. **Van Hoof, D., J. Munoz, S. R. Braam, M. W. H. Pinkse, R. Linding, A. J. R. Heck, C. L. Mummery, and J. Krijgsveld.** 2009. Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell **5:**214-226.
89. **Vylkova, S., X. S. Li, J. C. Berner, and M. Edgerton.** 2006. Distinct antifungal mechanisms: beta-defensins require *Candida albicans* Ssa1 protein, while Trk1p mediates activity of cysteine-free cationic peptides. Antimicrobial Agents and Chemotherapy **50:**324-331.
90. **Wessel, D., and U. I. Flugge.** 1984. A method for the quantitative recovery of protein in dilute-Solution in the presence of detergents and lipids. Analytical Biochemistry **138:**141-143.
91. **Xin, H., S. Dziadek, D. R. Bundle, and J. E. Cutler.** 2008. Synthetic glycopeptide vaccines combining beta-mannan and peptide epitopes induce protection against candidiasis. Proceedings of the National Academy of Sciences of the United States Of America **105:**13526-13531.

92. **Yin, Q. Y., P. W. J. de Groot, C. G. de Koster, and F. M. Klis.** 2008. Mass spectrometry-based proteomics of fungal wall glycoproteins. Trends in Microbiology **16:**20-26.
93. **Yin, Q. Y., P. W. J. de Groot, H. L. Dekker, L. de Jong, F. M. Klis, and C. G. de Koster.** 2005. Comprehensive proteomic analysis of *Saccharomyces cerevisiae* cell walls. Journal of Biological Chemistry **280:**20894-20901.
94. **Zhu, J. K., B. Damsz, A. K. Kononowicz, R. A. Bressan, and P. M. Hasegawa.** 1994. A higher-plant extracellular vitronectin-like adhesion protein is related to the translational elongation factor-1-alpha. Plant Cell **6:**393-404.

# APPENDIX

Table S1. Summary of CWPs identified through cell surface biotinylation and LC-MS in Chapter 4

| Accession Number | Protein Name and Biotinylated Peptide | Xcorr | ΔCN |
|---|---|---|---|
| CA3874 | CaENO1 Enolase I (2-phosphoglycerate dehydratase) | | |
| | K.AK&IDVVDQAK.I | 3.34 | 0.41 |
| | K.AVANVNDIIAPALIK&AK&IDVVDQAK.I | 5.52 | 0.36 |
| | K.AVANVNDIIAPALIK&AK.I | 5.65 | 0.11 |
| | K.DGK+YDLDFK.N | 3.04 | 0.34 |
| | K.DGK+YDLDFKNPROESDPSK.W | 3.60 | 0.37 |
| | K.EALDLIM*DAIDK&AGYK.G | 4.51 | 0.44 |
| | K.EALDLIMDAIDK&AGYK.G | 5.91 | 0.42 |
| | K.EALDLIMDAIDK+AGYK.G | 5.01 | 0.34 |
| | K.GK&VGIAMDVASSEFYK.D | 5.45 | 0.51 |
| | K.GK+VGIAMDVASSEFYK.D | 5.24 | 0.56 |
| | K.GVLK&AVANVNDIIAPALIK.A | 5.71 | 0.62 |
| | K.HIANISNAK&K.G | 2.99 | 0.38 |
| | K.K&AANALLLK.V | 3.14 | 0.31 |
| | K.SK&LGANAILGVSLAAANAAAAAQGIPLYK.H | 6.42 | 0.63 |
| | K.TPK&EALDLIMDAIDK&AGYK.G | 6.28 | 0.26 |
| | K.TPK&EALDLIMDAIDK.A | 5.38 | 0.49 |
| | K.YGQSAGNVGDEGGVAPDIK&TPK.E | 4.66 | 0.66 |
| | K.YGQSAGNVGDEGGVAPDIK+TPK.E | 5.58 | 0.13 |
| | R.DGDK+SK+WLGK+GVLK.A | 4.28 | 0.38 |
| | R.GNPROTVEVDFTTDK&GLFR.S | 4.29 | 0.59 |
| | R.GNPROTVEVDFTTDK+GLFR.S | 3.71 | 0.54 |
| | R.IEEELGSEAIYAGK&DFQK.A | 3.63 | 0.18 |
| | R.IEEELGSEAIYAGK+DFQK.A | 6.07 | 0.33 |
| | R.IK+TAIEK+K+AANALLLK.V | 5.82 | 0.50 |
| | R.LAK&LNQILR.I | 3.07 | 0.12 |
| | R.LAK+LNQILR.I | 2.81 | 0.23 |
| | R.VGDK&IQIVGDDLTVTNPROTR.I | 4.52 | 0.33 |
| | R.VGDK+IQIVGDDLTVTNPROTR.I | 6.69 | 0.6 |
| CA5892 | CaGAP1 Glyceraldehyde-3-phosphate dehydrogenase | | |
| | K.DIEVVAVNDPFIAPDYAAYMFK+YDSTHGR.Y | 4.17 | 0.55 |
| | K.HIDAGAK&K.V | 2.99 | 0.11 |
| | K.IISNASC#TTNC#LAPLAK.V | 2.72 | 0.42 |
| | K.IISNASC#TTNC@LAPLAK.V | 3.92 | 0.29 |
| | K.IK&VFQER.D | 2.91 | 0.33 |
| | K.K+AASYEEIAQAIK.K | 4.17 | 0.41 |
| | K.K+AASYEEIAQAIK+K+ASEGPLK.G | 4.63 | 0.40 |
| | K.LEGAQK&HIDAGAK.K | 4.24 | 0.4 |
| | K.TVDGPSHK&DWR.G | 2.89 | 0.34 |
| | R.DPANIPWGK+SGVDYVIESTGVFTK.L | 5.74 | 0.61 |
| | R.K+DIEVVAVNDPFIAPDYAAYMFK.Y | 5.02 | 0.51 |
| | R.LK+K+AASYEEIAQAIK+K.A | 5.30 | 0.09 |
| | R.TASGNIIPSSTGAAK&AVGK.V | 4.96 | 0.33 |
| | R.TASGNIIPSSTGAAK+AVGK.V | 3.99 | 0.29 |

| | | | |
|---|---|---|---|
| | R.YK&GEVTASGDDLVIDGHK&IK.V | 5.72 | 0.63 |
| | R.YK+GEVTASGDDLVIDGHK+IK.V | 5.91 | 0.08 |
| | R.YK&GEVTASGDDLVIDGHK.I | 6.71 | 0.62 |
| | R.YK+GEVTASGDDLVIDGHK.I | 6.05 | 0.67 |
| | R.YK+GEVTASGDDLVIDGHK+IK+VFQER.D | 6.81 | 0.58 |
| | R.YKGEVTASGDDLVIDGHK&IK.V | 5.78 | 0.09 |
| **CA0362** | **CaTEF1 translation elongation factor eEF1 alpha-A chain** | | |
| | K.AGK&VTK&AAQK.A | 2.78 | 0.06 |
| | K.FVK&SGDAAIVK.M | 3.86 | 0.36 |
| | K.FVK+SGDAAIVK.M | 2.55 | 0.24 |
| | K.K&LEENPROK.F | 3.09 | 0.16 |
| | K.LEENPROK&FVK&SGDAAIVK.M | 3.83 | 0.47 |
| | K.MVPTK&PMC@VEAFTDYPPLGR.F | 4.46 | 0.62 |
| | K.MVPTKPMC#VEAFTDYPPLGR.F | 4.33 | 0.49 |
| | K.NMITGTSQADC#AILIIAGGTGEFEAGISK.D | 5.60 | 0.51 |
| | K.TLLEAIDAIEPPTRPTDK+PLR.L | 4.32 | 0.54 |
| | K.TVPFVPISGWNGDNMIEPSTNC#PWYK.G | 3.51 | 0.29 |
| | K.VTGK+TLLEAIDAIEPPTRPTDK+PLR.L | 3.97 | 0.50 |
| | K.VTGK&TLLEAIDAIEPPTRPTDK&PLR.L | 4.19 | 0.42 |
| | K.VTK&AAQK.A | 2.84 | 0.28 |
| | R.GNVC@GDSK&NDPPK.G | 3.39 | 0.47 |
| | R.GNVC#GDSK.N | 2.53 | 0.46 |
| | R.LPLQDVYK&IGGIGTVPVGR.V | 4.76 | 0.53 |
| | R.LPLQDVYK+IGGIGTVPVGR.V | 5.41 | 0.63 |
| | R.QTVAVGVIK#SVEK.S | 3.89 | 0.42 |
| | R.QTVAVGVIK+SVEK.S | 4.18 | 0.29 |
| | R.TGK&K&LEENPROK.F | 2.98 | 0.2 |
| | R.TIEK&FEK.E | 2.88 | 0.38 |
| **CA1230** | **CaSSA4 cahsp70 mRNA for heat shock** | | |
| | K.AVGIDLGTTYSC#VAHFANDR.V | 3.33 | 0.44 |
| | K.EIAEGYLGSTVK+DAVVTVPAYFNDSQR.Q | 5.24 | 0.52 |
| | K.FDDPEVINDAK&HFPFK.V | 3.51 | 0.13 |
| | K.GTGK+TQK+ITITNDK+GR.L | 3.88 | 0.46 |
| | K.ITITNDK&GR.L | 2.60 | 0.37 |
| | K.K&DISTNQR.A | 2.93 | 0.18 |
| | K.NQAAMNPROANTVFDAK+R.L | 4.25 | 0.56 |
| | K.TFSPEEISSMVLTK.M | 3.82 | 0.58 |
| | K.TQK&ITITNDK.G | 3.12 | 0.41 |
| | R.FEELC#ADLFR.S | 3.70 | 0.53 |
| | R.IINEPTAAAIAYGLDK+K.G | 5.35 | 0.08 |
| | R.IINEPTAAAIAYGLDK&K.G | 4.44 | 0.07 |
| | R.K&FDDPEVINDAK&HFPFK.V | 6.58 | 0.23 |
| | R.K+NK+K+DISTNQR.A | 4.43 | 0.32 |
| | R.KFDDPEVINDAK&HFPFK&VIDK.A | 4.46 | 0.10 |
| | R.KFDDPEVINDAK&HFPFK.V | 4.71 | 0.19 |
| | R.LIGDAAK+NQAAMNPROANTVFDAK.R | 5.84 | 0.70 |
| | R.LVNFFIQEFK+R.K | 3.58 | 0.45 |
| | R.STLDPVGK&VLADAK.I | 3.13 | 0.41 |

| | | | |
|---|---|---|---|
| **CA5180** | **CaFBA1 fructose-bisphosphate aldolase (by homology)** | | |
| | K.APIILQTSQGGAAYFAGK&GVDNK.D | 6.18 | 0.29 |
| | K.APIILQTSQGGAAYFAGK+GVDNK.D | 3.69 | 0.38 |
| | K.APVGNPROEGADK&PNK&KYFDPR.V | 3.55 | 0.51 |
| | K.APVGNPROEGADK&PNKK.Y | 3.50 | 0.62 |
| | K.APVGNPROEGADK+PNK+K+YFDPR.V | 3.59 | 0.40 |
| | K.APVGNPROEGADKPNK&K&YFDPR.V | 4.18 | 0.10 |
| | K.APVGNPROEGADKPNK&K.Y | 4.38 | 0.22 |
| | K.HPLYLVFHGGSGSTQEEFNTAIK.N | 3.54 | 0.40 |
| | K.K&LLPWFDGM*LK.A | 3.12 | 0.45 |
| | K.K&LLPWFDGMLK.A | 3.49 | 0.35 |
| | K.K+LLPWFDGMLK.A | 2.93 | 0.31 |
| | K.SGVIYGK&DVK.D | 3.10 | 0.39 |
| | K.SGVIYGK+DVK.D | 2.56 | 0.32 |
| | K.TGTPLFSSHMLDLSEETDDENIATC#AK.Y | 5.71 | 0.65 |
| | K.VNLDTDC#QYAYLTGIR.D | 5.11 | 0.64 |
| | R.DYVTNK&IEYLK.A | 4.19 | 0.48 |
| | R.DYVTNK+IEYLK.A | 3.59 | 0.48 |
| | R.IAEALDIFHTK&GQL | 3.99 | 0.49 |
| | R.IAEALDIFHTK+GQL | 4.76 | 0.49 |
| **CA2474** | **CaPDC11 Pyruvate decarboxylase (by homology)** | | |
| | K.EALQK&LLTTVK.K | 2.51 | 0.48 |
| | K.EALQK+LLTTVK.K | 3.01 | 0.31 |
| | K.K&SINPRONYTPVPVPETK.L | 4.65 | 0.62 |
| | K.K+SINPRONYTPVPVPETK.L | 4.91 | 0.47 |
| | K.K+IDLSLHPNDPESQTEVIETVEK.L | 5.34 | 0.49 |
| | K.LIEETQFPVFTTPMGK+SSVDESNPROR.F | 3.66 | 0.41 |
| | K.LISEASNPROVILVDAC#AIR.H | 6.69 | 0.59 |
| | K.VPK&SLLDK.K | 2.69 | 0.18 |
| | K.WEC#NNTYLFVLNNDGYTIER.L | 4.45 | 0.44 |
| | R.FGGVYVGSLSK+PEVK.E | 4.57 | 0.28 |
| | R.HNCK+PEVAK+LIEETQFPVFTTPMGK.S | 5.03 | 0.54 |
| | R.ISTVGELNDLFADK&AFAVPDK.I | 4.39 | 0.32 |
| | R.ISTVGELNDLFADK+AFAVPDK.I | 5.71 | 0.36 |
| | R.LIHGEK+AQYNDIQPWNNLQLLPLFNAK.D | 6.00 | 0.57 |
| | R.MFK&NISQTSAFIADINSAPAEIDR.C | 5.23 | 0.59 |
| | R.MFK+NISQTSAFIADINSAPAEIDR.C | 5.67 | 0.53 |
| | R.PVYIGLPSNLVDMK&VPK.S | 5.59 | 0.16 |
| | R.PVYIGLPSNLVDMK+VPK.S | 5.41 | 0.15 |
| | R.VILFVGDGSLQLTVQEISTMC#K.W | 4.10 | 0.55 |
| **CA4765** | **CaADH1 alcohol dehydrogenase (by homology)** | | |
| | K.C@PIK&IVGLSDLPEVFK.L | 5.76 | 0.62 |
| | K.DK&DIVEAVK.K | 3.31 | 0.21 |
| | K.GDWPLATK&LPLVGGHEGAGVVVGMGENVK.G | 4.04 | 0.41 |
| | K.IVGLSDLPEVFK+LMEEGK.I | 5.28 | 0.40 |
| | K.K&ATDGGPHGAINVSVSEK.A | 3.52 | 0.43 |

| | | | |
|---|---|---|---|
| | K.LPLVGGHEGAGVVVGMGENVK.G | 4.35 | 0.51 |
| | K.SIEIK+GSYVGNR.K | 3.43 | 0.40 |
| | K.SLGAEAYVDFTKDK&DIVEAVK.K | 4.17 | 0.18 |
| | K.TADLAAGQWVAISGAGGGLGSLAVQYAR.A | 4.83 | 0.54 |
| | K.VVLVGLPAHAK&VTAPVFDAVVK.S | 4.06 | 0.57 |
| | R.GLIK&C@PIK&IVGLSDLPEVFK.L | 5.82 | 0.48 |
| | R.GLIK+C@PIK+IVGLSDLPEVFK.L | 5.17 | 0.63 |
| | R.K&DTAEAIDFFSR.G | 4.33 | 0.43 |
| | R.K+DTAEAIDFFSR.G | 3.77 | 0.33 |
| | R.VVAIDGGDEK&GEFVK.S | 4.75 | 0.46 |
| | R.VVAIDGGDEK+GEFVK.S | 5.01 | 0.36 |
| **CA1691** | **CaPGK1 Phosphoglycerate kinase** | | |
| | K.ALENPROERPFLAILGGAK&VSDK.I | 5.11 | 0.34 |
| | K.K+VK+ADPEAVK+K.F | 3.75 | 0.10 |
| | K.LLGQK+VTFLNDCVGPEVTK+AVENAK.D | 4.25 | 0.32 |
| | K.LSVK&DLDVAGK&R.V | 3.12 | 0.4 |
| | K.LSVK&DLDVAGK.R | 3.54 | 0.43 |
| | K.MPIGDSLFDEAGAK&NVEHLVEK.A | 4.97 | 0.09 |
| | K.SAENGNIVIIGGGDTATVAK&K.Y | 5.19 | 0.05 |
| | K.VTFLNDC#VGPEVTK.A | 4.60 | 0.48 |
| | R.NDK&YSLAPVATELEK.L | 4.95 | 0.5 |
| | R.NDK+YSLAPVATELEK.L | 4.25 | 0.52 |
| | R.YHIEEEGSSK&DKDGK.K | 5.57 | 0.22 |
| | R.YHIEEEGSSK+DKDGK.K | 4.01 | 0.08 |
| **CA2810** | **CaEFT2 translation elongation factor 2** | | |
| | K.AGIISAAK&AGDAR.F | 2.54 | 0.42 |
| | R.AGLKPEVPEYTEYYDK+L | 4.41 | 0.45 |
| | R.ALLELQTTK+EDLYQTFAR.T | 4.13 | 0.65 |
| | R.ILADK+HGWDVVDAR.K | 3.33 | 0.49 |
| | R.GLMDK+VTNVR.N | 2.70 | 0.26 |
| | R.AETLYEGPSDDPFC#TAIR.N | 4.99 | 0.57 |
| | K.IWC#FGPDGNGPNLVVDQTK.A | 4.75 | 0.52 |
| | R.NC#DPNADLMLYVSK.M | 4.66 | 0.48 |
| | R.VTDGALVVVDTVEGVC#VQTETVLR.Q | 4.44 | 0.57 |
| | R.SVEQIDDC#PAGNIIGLVGIDQFLLK.S | 4.23 | 0.62 |
| **CA5950** | **CaTPI1 Triose phosphate isomerase** | | |
| | K.ANVDGFLVGGASLK+PEFVDIIK.S | 5.60 | 0.48 |
| | K.DVEVVIC#PPALYLGLAVEQNK.Q | 4.72 | 0.55 |
| | K.GGVTLDVC#AR.Q | 3.15 | 0.41 |
| | K.SC#GAFTGETC@ASQILDVGASWTLTGHSER.R | 5.93 | 0.60 |
| | K.SIGAEQAEK&TR.I | 3.44 | 0.13 |
| | K.VILC#IGETLEER.K | 4.45 | 0.51 |
| | R.AHLAK&SIGAEQAEK.T | 5.77 | 0.47 |
| | R.AHLAK+SIGAEQAEK.T | 4.61 | 0.43 |
| | R.ILYGGSVNGK&NAK.D | 3.64 | 0.34 |
| | R.ILYGGSVNGK+NAK.D | 3.85 | 0.20 |

| | | | |
|---|---|---|---|
| **CA0653** | **CaMET6 BY HOMOLOGY TO S.CEREV.: 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase** | | |
| | K.ADK&DSLDLEPISLLPK.I | 4.18 | 0.40 |
| | K.ASAVVQK+AIEK.V | 3.29 | 0.40 |
| | K.FWVNPRODC#GLK.T | 3.03 | 0.34 |
| | K.FWVNPRODC@GLK&TR.G | 3.96 | 0.39 |
| | K.YNLPLFPTTTIGSFPQTK+DIR.I | 4.19 | 0.51 |
| | K.YNLPLFPTTTIGSFPQTK&DIR.I | 4.36 | 0.6 |
| | R.DDVSGK+IQALQLGLALR.D | 4.16 | 0.42 |
| | R.GWPEVK+ESLTNMVEAAK.E | 3.51 | 0.12 |
| | R.SESSITNDPK+VQER.L | 3.68 | 0.40 |
| | R.YTK&FDLAPIDVLFAMGR.G | 2.80 | 0.44 |
| **CA5857** | **CaPCK1 phosphoenolpyruvate carboxykinase** | | |
| | K.ISEHNANAWLLNTGWVGSSVAQGGK+R.C | 4.50 | 0.46 |
| | K.LIGDDEHC#WSDNGVFNIEGGC#YAK.C | 6.74 | 0.48 |
| | K.C#LDLSAEKEPEIFNSIK.F | 4.60 | 0.45 |
| | K.IPC@LADTHPTNIILLTC#DASGVLPPVSK.L | 4.29 | 0.16 |
| | K.IPC#LADTHPTNIILLTC@DASGVLPPVSK.L | 3.62 | 0.54 |
| | K.LIGDDEHC#WSDNGVFNIEGGC@YAK.C | 3.51 | 0.42 |
| | R.C#AYPIDFIPSAK.I | 3.21 | 0.45 |
| **CA5388** | **CaPET9 ADP/ATP carrier protein (by homology)** | | |
| | R.YFPTQALNFAFK&DK.F | 3.85 | 0.18 |
| | R.LANDAK+SSK+GDGK+R.E | 4.61 | 0.48 |
| | R.VK+LLIQNQDEMIK.Q | 4.31 | 0.42 |
| | R.EFNGLVDVYK+K.T | 3.50 | 0.06 |
| | R.K+VVAAEGVGSLFK.G | 3.22 | 0.43 |
| | R.YTGIVDC#FK.R | 3.38 | 0.53 |
| | K.YDGALDC#FR.K | 3.29 | 0.36 |
| **CA3115** | **CaECM33.3 cell wall biogenesis, 3-prime end (by homology)** | | |
| | K.LK&TIGGALQISDNSELR.S | 5.47 | 0.58 |
| | K.LK+TIGGALQISDNSELR.S | 3.47 | 0.55 |
| | K.LSC@SAFNK&LNTNGDIK.G | 5.36 | 0.48 |
| | K.TIGGALQISDNSELR.S | 4.53 | 0.58 |
| | K.TDGK&LSC@SAFNK&LNTNGDIKGDK.F | 4.27 | 0.39 |
| | K.TDGK&LSC@SAFNK.L | 4.06 | 0.41 |
| | K.VDVILK&QLSSAK.N | 3.23 | 0.46 |
| **CA3307** | **CaRPL2.3 ribosomal protein L8, 3-prime end (by homology)** | | |
| | R.TSGNYVIIIGHNPRODENK&TR.V | 4.48 | 0.48 |
| | R.TSGNYVIIIGHNPRODENK+TR.V | 4.51 | 0.46 |
| | R.GAVSGQK&AGLIAAR.R | 4.17 | 0.48 |
| | R.GVVK&QIIHDPGR.G | 3.35 | 0.42 |

| | | | |
|---|---|---|---|
| | R.GVVK+QIIHDPGR.G | 3.34 | 0.45 |
| | R.K+GAGSIFTSHTR.L | 3.71 | 0.40 |
| **CA4457** | **CaATP1.exon3 F1F0-ATPase complex, F1 alpha subunit, exon 3** | | |
| | R.IGEFEEAFLGHLK+SNETGILDAIK.T | 6.25 | 0.35 |
| | K.SNETGILDAIK+TK.G | 4.35 | 0.19 |
| | K.TKGELSK+DELEK.L | 3.85 | 0.22 |
| | K.GELSK+DELEK.L | 3.35 | 0.41 |
| | R.VGSAAQVK+AMK.Q | 2.88 | 0.27 |
| | R.LTQLLK+QK.Q | 2.83 | 0.23 |
| **CA0615** | **CaRPL3 60S large subunit ribosomal protein L3.e (by homology)** | | |
| | K.K&TINPROMGGFVR.Y | 3.27 | 0.28 |
| | R.DLDRPGSK&MHK.R | 2.78 | 0.30 |
| | R.TSINHK&VYR.V | 2.57 | 0.37 |
| | R.VGK+GTDEANGATEFDR.T | 4.88 | 0.62 |
| | R.FQTPAEK+HAFMGTLK+K+DLEN | 4.48 | 0.50 |
| | K.VAC#IGAWHPANVNWTVAR.A | 3.65 | 0.37 |
| **CA3534** | **CaSSB1 heat shock protein 70** | | |
| | R.QATK&DAGAIAGLNVLR.I | 3.84 | 0.47 |
| | R.QATK+DAGAIAGLNVLR.I | 4.63 | 0.56 |
| | R.AVTK&GMATR | 2.99 | 0.53 |
| | K.VTAVEK&STGR.S | 2.79 | 0.52 |
| | R.LIGDAAK+NQAALNPROK.N | 3.72 | 0.43 |
| | R.VNC#TENTLLGEFDLK.N | 3.62 | 0.41 |
| **CA1440** | **CaRPL17B RPL17B ribosomal protein L17.e** | | |
| | R.YAATPANPROAK&SASAR.G | 3.59 | 0.50 |
| | R.YAATPANPROAK+SASAR.G | 3.75 | 0.61 |
| | R.WPAK&SVNFVK&DLLR.N | 2.83 | 0.52 |
| | R.WPAK+SVNFVK+DLLR.N | 4.40 | 0.61 |
| | R.ETAQAINGWK+LEK+AQK+YLDQVLDHQR.A | 5.36 | 0.60 |
| | R.NAQANAEAK+GLDSSK.L | 4.92 | 0.39 |
| **CA3018** | **CaSAH1 S-adenosyl-L-homocysteine hydrolase (by homology)** | | |
| | R.EK&FPEFAK&TGPFDVGVHLLPK.V | 3.60 | 0.53 |
| | R.K+DIELSENEMPGLMYIR.K | 6.17 | 0.66 |
| | K.K+LNLILDDGGDLTSLVHEK.Y | 5.25 | 0.52 |
| | K.YGPSQPLK+GAR.I | 2.67 | 0.44 |
| | K.GETEEEYQWC#IEQQLFAFK.D | 5.49 | 0.63 |
| | K.FDNLYGC#R.E | 2.78 | 0.32 |
| **CA3483** | **CaCDC19 pyruvate kinase (by homology)** | | |

| | | | |
|---|---|---|---|
| | R.GVYPFIYDK+PSIENWQEDVENR.L | 4.58 | 0.61 |
| | R.SLNAGK+ISSHK.G | 2.97 | 0.25 |
| | K.KPTATTETC#AVAAVSAAYEQDAK.A | 5.69 | 0.62 |
| | R.AEVSDVGNAILDGADC#VMLSGETAK.G | 4.02 | 0.58 |
| | K.PVIC#ATQMLESMTYNPROR.P | 3.82 | 0.48 |
| | K.C#NLAAKPVIC#ATQMLESMTYNPROR.P | 3.43 | 0.38 |
| **CA6123** | **CaRPS15.3 40S ribosomal protein S15, 3-prime end (by homology)** | | |
| | R.AATEPNERPAVVK&THLR.N | 4.60 | 0.58 |
| | R.AATEPNERPAVVK+THLR.N | 3.68 | 0.49 |
| | R.AGNASSK&FMPLR | 3.29 | 0.25 |
| | R.AGNASSK+FMPLR | 3.35 | 0.5 |
| | R.GLDSK&PMGLIK&K&LR.A | 3.46 | 0.33 |
| | R.GLDSK+PMGLIK+K+LR.A | 3.95 | 0.44 |
| **CA2300** | **CaPMA1 plasma membrane H+-transporting ATPase 1** | | |
| | R.VGLTDDEVTK+R.R | 3.82 | 0.53 |
| | R.IVSEDC#LLQVDQSAITGESLAVDKR.S | 6.27 | 0.62 |
| | K.LSLHEPYTVEGVEPDDLMLTAC#LAASR.K | 6.27 | 0.65 |
| | K.LSAIESLAGVEILC#SDK.T | 4.40 | 0.52 |
| | R.SGDSC#YSSSTVK.T | 2.58 | 0.40 |
| | VIEFQPFDPVSKK&VTAIVESPEGER | 3.90 | 0.50 |
| **CA4412** | **CaPDA1 Pyruvate dehydrogenase alpha chain (by homology)** | | |
| | R.K+YVDEQVAAAEADAPPEAK.M | 5.72 | 0.72 |
| | R.NDPIAGLK+AVLLEK.E | 4.72 | 0.55 |
| | R.ISDDSWDFK+NK.T | 4.08 | 0.17 |
| | R.SSAMTEYYK+R.G | 2.89 | 0.41 |
| | K.LWNLPVIFAC#ENNK.Y | 4.91 | 0.65 |
| **CA4748** | **CaMLS1 malate synthase** | | |
| | R.DLLSPFVPNAK+ITEQGIR.A | 5.15 | 0.52 |
| | R.EVTLGC#DSC@WVAHPALVPVVLK.V | 3.60 | 0.43 |
| | R.EVTLGC#DSC#WVAHPALVPVVLK.V | 3.42 | 0.46 |
| | R.EVTLGC@DSC#WVAHPALVPVVLK.V | 3.12 | 0.23 |
| | R.EHSAGLNC#GR.W | 2.94 | 0.58 |
| **CA3909** | **CaCIT1.exon2 Citrate synthase, exon 2** | | |
| | K.LPAAIDSK&LDYGANLASLLGFGDNK.E | 6.15 | 0.51 |
| | R.GIK&GLVWEGSVLDPIEGIR.F | 5.07 | 0.55 |
| | R.GIK+GLVWEGSVLDPIEGIR.F | 6.21 | 0.65 |
| | R.SALPK+HVEELIDR.S | 3.31 | 0.54 |
| | K.AEEVK+QFK+K+EHGK.T | 3.13 | 0.06 |

| | | | |
|---|---|---|---|
| **CA6079** | **CaRPL18.exon2 Ribosomal Protein RPL18B (large subunit), exon 2 (by homology)** | | |
| | R.LLEFPK&ATIAALR.F | 4.20 | 0.43 |
| | R.APK&GQNTLIVR.G | 3.64 | 0.41 |
| | R.APK+GQNTLIVR.G | 3.12 | 0.42 |
| | R.ILK+NGGEAITLDQLALR.A | 5.59 | 0.59 |
| | R.LLEFPK+ATIAALR.F | 4.14 | 0.57 |
| **CA0127** | **CaHXK2.3f hexokinase II, 3-prime end (by homology)** | | |
| | R.DFDTTQSK&FALPAHMR.T | 4.46 | 0.56 |
| | R.DFDTTQSK+FALPAHMR.T | 3.85 | 0.40 |
| | R.WTK&GWSIDGIEGK.D | 4.04 | 0.45 |
| | R.EK+LGIETTEPER.K | 2.86 | 0.42 |
| | R.FSVC#GIAAIC#QK.R | 4.04 | 0.34 |
| **CA6141** | **CaRPL19A.3 Ribosomal protein L19.e, 3-prime end (by homology)** | | |
| | R.K+VWLDPNETTEIANANSR.S | 5.59 | 0.57 |
| | R.SLVEHIIAAK+AEALR.E | 5.35 | 0.61 |
| | R.LAASVIGVGK+R.K | 4.00 | 0.55 |
| | R.SLVEHIIAAK&AEALR.E | 5.19 | 0.54 |
| | R.K&VWLDPNETTEIANANSR.S | 4.56 | 0.57 |
| **CA5239** | **CaGND1 6-phosphogluconate dehydrogenase** | | |
| | K.ALK+GPQVTGESPITDK.K | 4.68 | 0.55 |
| | R.TGPSLMPGGNEK+AWPHIK.D | 3.72 | 0.25 |
| | K.SDGEPC@C#DWVGDAGAGHYVK.M | 4.46 | 0.34 |
| | K.SDGEPC#C#DWVGDAGAGHYVK.M | 5.37 | 0.44 |
| **CA1051** | **CaCHT2 chitinase 2 precursor** | | |
| | K.FADSAPNK&NIK.L | 3.72 | 0.38 |
| | K.FADSAPNK+NIK.L | 3.4 | 0.24 |
| | K.FAK&DTSK.N | 2.61 | 0.32 |
| | K.C#DSHFAGVSLWDASGAWLNTDEK.G | 5.08 | 0.58 |
| **CA1564** | **CaGAD1 Glutamate decarboxylase (by homology)** | | |
| | K.K&LDSK&LYTINEEFKPGLPVVAFR.F | 5.05 | 0.56 |
| | R.DK&YPELPQELLSTLLR.K | 3.77 | 0.56 |
| | R.DK+YPELPQELLSTLLR.K | 4.31 | 0.58 |
| | K.LLDEVEK+ER.G | 3.28 | 0.38 |
| **CA2011** | **CaRPS31 Ubiquitin fusion protein** | | |
| | R.TLSDYNIQK&ESTLHLVLR.L | 3.91 | 0.48 |
| | R.LIFAGK&QLEDGR.T | 3.60 | 0.36 |
| | R.LIFAGK+QLEDGR.T | 3.76 | 0.48 |

| | | | |
|---|---|---|---|
| | K.IQDK+EGIPPDQQR.L | 3.58 | 0.49 |
| **CA2454** | **CaRPL23B.3 ribosomal protein L23.e, 3-prime end (by homology)** | | |
| | R.LPAAAAGDMVMATVK&K&GKPELR.K | 4.37 | 0.12 |
| | R.K&K&VMPAIVIR.Q | 2.91 | 0.33 |
| | R.NLYVLAVK+GTGAR.L | 3.83 | 0.54 |
| | R.LPAAAAGDMVMATVK+K+GK+PELR.K | 3.62 | 0.36 |
| **CA2811** | **CaRPS10 Ribosomal protein 10** | | |
| | R.K&DWFDIK&APTTFENR.N | 3.85 | 0.61 |
| | R.NVGK&TLINR.S | 3.32 | 0.37 |
| | R.VFEVC#LADLQGSEDHSYR.K | 6.30 | 0.42 |
| | R.EVSNC#TLAQLTSK.L | 3.47 | 0.50 |
| **CA3081** | **CaCEF3 translation elongation factor 3** | | |
| | K.IFK+IEGTPR.R | 3.42 | 0.40 |
| | K.C#PSAQSYYELGASDLEFR.F | 5.65 | 0.63 |
| | K.FIPQLISC#IAK.P | 3.30 | 0.31 |
| | K.STLINVLTGELLPTTGEVYVHENC#R.I | 4.80 | 0.54 |
| **CA3304** | **CaRPL82 60S ribosomal protein L7a.e.B (by homology)** | | |
| | K.NFGIGQSIQPK&R.N | 3.31 | 0.34 |
| | R.LK&VPPSIAQFSQTLDK.N | 3.78 | 0.56 |
| | R.LTK+EAAAIAEGK.T | 3.19 | 0.38 |
| | K.LVLIANDVDPIELVVFLPALC#K.K | 5.20 | 0.62 |
| **CA3800** | **CaRPL7A.3 60S Ribosomal Protein L7-A, 3-prime end** | | |
| | R.K&FQHFIQGGDTGNR.E | 4.91 | 0.54 |
| | R.K+FQHFIQGGDTGNR.E | 4.75 | 0.49 |
| | R.AAAYQK+EYTEAER.S | 3.90 | 0.49 |
| | R.EQFINALVK+QMN | 3.48 | 0.37 |
| **CA4220** | **IPF8762 unknown function** | | |
| | R.K&IAQGDNGKGQQVSR.Y | 3.90 | 0.50 |
| | R.K+IAQGDNGK+GQQVSR.Y | 5.13 | 0.67 |
| | R.K+YATEVPPNFNPROR.K | 4.44 | 0.62 |
| | K.VHADHEK+IVNIPISK.A | 3.21 | 0.33 |
| **CA4362** | **CaATP2 F1F0-ATPase complex, F1 beta subunit (by homology)** | | |
| | K.TVFIQELINNIAK&AHGGFSVFTGVGER.T | 4.65 | 0.66 |
| | R.GISELGIYPAVDPLDSK+SR.L | 4.91 | 0.57 |
| | R.EMK+ETGVINLEGDSK.V | 3.84 | 0.57 |
| | R.IINVVGEPIDDRGPIEC#K.E | 3.08 | 0.39 |

| | | | |
|---|---|---|---|
| **CA4399** | **CaPRA1 pH-regulated antigen** | | |
| | K.SSILFLC@DDLDDK&C@K.N | 5.22 | 0.13 |
| | K.SSILFLC@DDLDDK+C@K.N | 5.21 | 0.16 |
| | K.SK&TNIFWAGDLLHR.F | 4.81 | 0.47 |
| | R.FGSK+SPFFR.K | 3.18 | 0.54 |
| **CA5206** | **CaGPH1 Glycogen phosphorylase (by homology)** | | |
| | K.LEDPNYDYLTNLGK+LK.K | 4.50 | 0.13 |
| | R.QANPROK+LAALIAEK.L | 3.29 | 0.54 |
| | R.VSIIEESPK+SVR.M | 3.02 | 0.38 |
| | R.LAAC#FVDSLSSK.N | 3.71 | 0.48 |
| **CA1662** | **CaRPL28.3f Ribosomal protein, 3-prime end (by homology)** | | |
| | R.TNLDK&YHPGYFGK&VGMR.Y | 5.15 | 0.57 |
| | R.TNLDK+YHPGYFGK+VGMR.Y | 3.43 | 0.36 |
| | R.LPEVPVIVK&AR.F | 4.05 | 0.55 |
| **CA0345** | **CaACH1 acetyl-coenzyme-A hydrolase (by homology)** | | |
| | R.TQFFDK&HLSMFPQDLTYGFYTK.D | 5.08 | 0.43 |
| | R.GIANHLIEFLEHEVK+QGR.L | 5.24 | 0.55 |
| | R.TGEQC#IDLFK.H | 3.05 | 0.45 |
| **CA2561** | **CaCAR2 ornithine aminotransferase (by homology)** | | |
| | R.LAPPLVISEEDLLK+GVDSIR.K | 4.96 | 0.55 |
| | K.EVMSTLEPGSHGSTYGGNPROLAC#R.V | 5.01 | 0.48 |
| | K.HNVLLIC#DEIQTGIAR.T | 4.81 | 0.40 |
| **CA0622** | **IPF14171 unknown function** | | |
| | R.K&QAAGALGPELSR.K | 4.83 | 0.58 |
| | R.K&QAAGALGPELSR.K | 4.55 | 0.53 |
| | K.QK+ITDQIAYLK.Y | 2.91 | 0.30 |
| **CA1021** | **CaRPS4 ribosomal protein S4, 3-prime end** | | |
| | R.LAAPSHWMLDK&LSGTYAPR.P | 4.13 | 0.48 |
| | R.LAAPSHWMLDK+LSGTYAPR.P | 3.34 | 0.37 |
| | R.ISAEEAAYK+LGK.V | 4.36 | 0.33 |
| **CA0915** | **CaKAR2 dnaK-type molecular chaperone (by homology)** | | |
| | R.QATK&DAGTIAGLNVLR.I | 4.47 | 0.59 |
| | R.QATK+DAGTIAGLNVLR.I | 4.93 | 0.57 |
| | R.IPK+VQELLEGFFDGK.K | 4.67 | 0.51 |
| **CA2862** | **CaHHF21 histone H4** | | |

| | | | |
|---|---|---|---|
| | R.VVLK&QFLENVIR.D | 3.93 | 0.29 |
| | R.DNIQGITK+PAIR.R | 4.02 | 0.49 |
| | R.DAVTYTEHAK+R.K | 2.99 | 0.54 |
| **CA3208** | **CaSRB1GDP-mannose pyrophosphorylase** | | |
| | R.FVEK&PVEFVGNR.I | 3.90 | 0.54 |
| | K.KDDSPFFVLNSDVIC#DYPFK.E | 5.40 | 0.52 |
| | K.DFLSGTC#LYLTSLSK.K | 3.65 | 0.46 |
| **CA3278** | **CaRPS3 Ribosomal protein S3.e (by homology)** | | |
| | R.AK&SQK&YADGFMIHSGQPTR.D | 4.69 | 0.49 |
| | K.AK&LLSGLPIR.R | 3.66 | 0.32 |
| | R.FK+LSPEGIAIYAER.V | 4.32 | 0.59 |
| **CA3559** | **CaPGI1 Glucose-6-phosphate isomerase** | | |
| | K.VK&AEGATGGLVPHK.E | 4.46 | 0.58 |
| | K.VK+AEGATGGLVPHK.E | 4.34 | 0.56 |
| | R.K&MPVDGKDTAQEVDDVLK.H | 3.45 | 0.33 |
| **CA3883** | **CaGCV2 Glycine decarboxylase P subunit** | | |
| | R.AENIGVEIVEIPLATK+QGVEELSR.I | 3.98 | 0.53 |
| | K.HC#AHEFILDLR.E | 3.49 | 0.37 |
| | K.TFALSHGGGGPGQAPVC#VK.E | 3.16 | 0.40 |
| **CA3923** | **CaADH2 alcohol dehydrogenase I (by homology)** | | |
| | K.LMEEGK+ILGR.Y | 3.89 | 0.47 |
| | K.YSGVC#HTDLHAWK.G | 3.78 | 0.41 |
| | R.IPAGTDLANVAPILC#AGVTVYK.A | 5.57 | 0.62 |
| **CA4446** | **CaICL1 Isocitrate lyase** | | |
| | R.GTLK&INHPSSQQADK.L | 3.14 | 0.34 |
| | K.TVSFTFGALDPIHVAQMAK.Y | 3.21 | 0.45 |
| | K.LFHEAVIDEIK+NGNYSNK.D | 4.70 | 0.37 |
| **CA0290** | **CaPOT14 acetyl-CoA acetyltransferase (by homology)** | | |
| | R.TPIGSFQGTLSSLTYSDLGAHAVK&AALNK.V | 4.68 | 0.51 |
| | R.TPIGSFQGTLSSLTYSDLGAHAVK+AALNK.V | 4.64 | 0.56 |
| | K.AGNALK&QGK.F | 2.85 | 0.34 |
| **CA4474** | **CaSSC1 Mitochondrial heat shock protein 70-related** | | |
| | R.TTPSIVAFTK&DGER.L | 2.74 | 0.51 |
| | R.TTPSIVAFTK+DGER.L | 4.04 | 0.6 |
| | R.QAVVNPROSDTLFATK+R.L | 3.70 | 0.60 |

| | | | |
|---|---|---|---|
| CA4696 | CaHTA1 Histone H2A (by homology) | | |
| | R.SAK&AGLTFPVGR.V | 2.73 | 0.43 |
| | R.SAK+AGLTFPVGR.V | 3.56 | 0.5 |
| | R.K&GNYAQR.I | 2.52 | 0.22 |
| CA5714 | IPF2431 similar to Saccharomyces cerevisiae Tsa1p thiol | | |
| | R.GIFLIDPK&GVLR.Q | 3.95 | 0.47 |
| | R.GIFLIDPK+GVLR.Q | 3.86 | 0.43 |
| | K.YGEVC#PANWHPGDETIKPSPEASK.E | 5.54 | 0.57 |
| CA4671 | CaGPM1 phosphoglycerate mutase (by homology) | | |
| | R.AGELLK+EAGINVDVLHTSK.L | 4.81 | 0.43 |
| | K.QYLSAQASAVEADPK+SK.V | 5.25 | 0.06 |
| CA2047 | CaRPL10 Ribosomal protein L10 (by homology) | | |
| | R.INK&MLSC@AGADR.L | 3.24 | 0.50 |
| | R.K+GNLEANLQQFPNYQYSA | 3.68 | 0.56 |
| CAALS4.5f agglutinin-like protein, 5-prime end | | | |
| | F.K&FITDQTSIDLVADGR.T | 0.05 | 4.00 |
| | E.K&STVASSDR.I | 0.26 | 1.00 |
| CA0567 | CaRPS23 Ribosomal protein S23 (by homology) | | |
| | R.WADQAYK+AR.L | 3.46 | 0.47 |
| | K.VTAFVPNDGC#LNFVDENDEVLLAGFGR.R | 4.51 | 0.58 |
| CA0828 | IPF17186 unknown function | | |
| | K.K+ITGFTDIGEDILGVTDIMK.K | 6.21 | 0.49 |
| | K.GGVVSAVC#HGPAIFENLNDPK.T | 6.51 | 0.57 |
| CA0870 | CaIPP1 inorganic pyrophosphatase (by homology) | | |
| | K.IPDGKPENQFAFSGEC#K.N | 4.53 | 0.51 |
| | K.AKGDNDPLDVC#EIGEK.V | 4.04 | 0.32 |
| CA0972 | CaCYP1 cyclophilin (peptidylprolyl isomerase), mitochondrial (by homology) | | |
| | K.LYDDVVPK&TAENFR.A | 3.80 | 0.55 |
| | K.K&IESFGSGSGATSK.K | 3.66 | 0.60 |
| CA1239 | CaHSP60 Heat Shock Protein 60 (HSP60) | | |
| | R.RPLLIIAEDVDGEALAAC#ILNK.L | 4.15 | 0.40 |

| | | | |
|---|---|---|---|
| | K.NVAAGC#NPROMDLR.R | 3.36 | 0.55 |
| **CA1298** | **CaRPL32 ribosomal protein L32** | | |
| | K.DLDVLLLHTK&SYAAEIASSVSSR.K | 4.38 | 0.35 |
| | R.K+QK+GIDSCVR.R | 2.98 | 0.33 |
| **CA2031** | **CaRPL10A L10A ribosomal protein** | | |
| | R.LLGPTLSK+AGK.F | 2.75 | 0.22 |
| | R.PNMTIC#IFGDAFDVDR.A | 4.18 | 0.53 |
| **CA2065** | **CaQCR2 Ubiquinol--cytochrome-c reductase 40KD chain II** | | |
| | R.VPSSGASSALIGIPVK+PADFGK.Y | 5.48 | 0.37 |
| | K.IAAK+ESATDLTK.L | 3.72 | 0.45 |
| **CA2250** | **CaMIS11 mitochondrial C1-tetrahydrofolate synthase precursor** | | |
| | K.ASFITPVPGGVGPMTVAC#LLDNVVIGAK.K | 4.22 | 0.55 |
| | K.AGIPITC#EDIGC#AGALTALLK.D | 3.47 | 0.35 |
| **CA2416** | **CaIFD4 Putative aryl-alcohol dehydrogenase (by homology)** | | |
| | K.GVIPIAGVSK+FEQAEELVGIFK.V | 3.86 | 0.51 |
| | R.DADK+IIVDR.V | 2.73 | 0.05 |
| **CA2582** | **CaTAL1 transaldolase (by homology)** | | |
| | R.QFGK+DAVTLLTELESR.F | 4.25 | 0.61 |
| | K.EK+AEIALDR.L | 2.59 | 0.27 |
| **CA2634** | **CaZWF1 glucose-6-phosphate dehydrogenase (by homology)** | | |
| | R.YSK+DFWIPEAYEALIR.D | 4.40 | 0.53 |
| | R.LFYLALPPSVFTTVC#EK.V | 3.83 | 0.48 |
| **CA2708** | **CaRPS6A ribosomal protein S6 (by homology)** | | |
| | R.K&FFGLTK&EDDVR.D | 5.02 | 0.46 |
| | R.GC#IVAQDLSVLALSIVK.Q | 5.58 | 0.58 |
| **CA2818** | **CaRPL13 Ribosomal protein** | | |
| | R.QNK+SQETFDANVAR.L | 4.03 | 0.55 |
| | R.NEK+K+YK+GIR.E | 3.08 | 0.45 |
| **CA2857** | **CaHSP70 Heat shock protein of HSP70 family** | | |
| | K.SK&LDASEIEEVTK.A | 4.57 | 0.52 |

| | | | |
|---|---|---|---|
| | R.QATK+DAGTIAGLNVMR.I | 4.28 | 0.53 |
| **CA2998** | **CaLPD1 dihydrolipoamide dehydrogenase (by homology)** | | |
| | R.AK+TNMDTDGFVK.F | 4.06 | 0.55 |
| | R.IVTSTGILSLK+EVPER.L | 3.70 | 0.28 |
| **CA3123** | **CaRPL6.3 ribosomal protein, 3-prime end** | | |
| | K.TPLLK&QYLAASFSLK.N | 4.17 | 0.61 |
| | K.NLEDNTLLVSGPFK+VNGVPLR.R | 4.74 | 0.44 |
| **CA3147** | **IPF9803 unknown function** | | |
| | R.GQEDPDASAAASNAK+R.T | 5.05 | 0.68 |
| | K.LQEYK+DVEAK.H | 3.33 | 0.45 |
| **CA3327** | **CaNCP1 NADPH-cytochrome P450 reductase** | | |
| | K.GK&AWVDVPFEFIIENVQHLTPR.Y | 4.91 | 0.54 |
| | R.NC#VHVEFDISDSNLK.Y | 4.07 | 0.56 |
| **CA3530** | **IPF9926 alkaline phosphatase (by homology)** | | |
| | R.K&DSEYPNLVEQTQVALNALTEHTK.D | 6.60 | 0.56 |
| | R.K+DSEYPNLVEQTQVALNALTEHTK.D | 6.56 | 0.67 |
| **CA3546** | **CaACO1 aconitate hydratase (by homology)** | | |
| | R.NHYTGK+YDGVPQTAAAYR.D | 3.49 | 0.55 |
| | K.EVYDFLSTAC#AK.Y | 3.13 | 0.47 |
| **CA3689** | **CaRPS22A ribosomal protein S15a.e.c10 (by homology)** | | |
| | R.FNVK&INDIER.W | 4.26 | 0.37 |
| | R.FNVK+INDIER.W | 3.48 | 0.38 |
| **CA3867** | **CaPHR2 pH-regulated protein 2** | | |
| | K.K&SNTDASAFVK.A | 3.39 | 0.40 |
| | K.NLGIPIFFSEYGC#NEVR.P | 4.03 | 0.67 |
| **CA3924** | **CaTKL1 transketolase 1** | | |
| | R.K+LSENVINALHGK.I | 3.78 | 0.36 |
| | K.FNPROK+DPNWINR.D | 3.30 | 0.36 |
| **CA4001** | **CaRPL12 ribosomal protein** | | |
| | K.NVK+HSGNIPLDEIFEIAR.K | 4.26 | 0.48 |
| | K.EILGTAQSVGC#R.V | 3.24 | 0.45 |

| | | | |
|---|---|---|---|
| **CA4148** | **CaIDH2 Isocitrate dehydrogenase (NAD+) subunit 2, mitochondrial (by homology)** | | |
| | R.TGDLK+GTATTTR.F | 3.48 | 0.52 |
| | K.GLANPROTALLLSSC#MMLR.H | 4.24 | 0.58 |
| **CA4697** | **CaHTB1 Histone H2B (by homology)** | | |
| | R.LILPGELAK&HAVSEGTR.A | 4.97 | 0.63 |
| | R.LILPGELAK+HAVSEGTR.A | 4.63 | 0.67 |
| **CA4753** | **CaIDH1.3 isocitrate dehydrogenase (NAD+) subunit1, mitochondrial, 3-prime end (by homology)** | | |
| | R.HLGLNDHADK+ISK.A | 3.96 | 0.10 |
| | K.ANIMK+LGDGLFR.Q | 3.75 | 0.48 |
| **CA4862** | **CaEFB1 translation elongation factor eEF1beta** | | |
| | K.AFQK&EFPQFTR.W | 3.80 | 0.39 |
| | K.AFQK+EFPQFTR.W | 3.87 | 0.40 |
| **CA5164** | **CaMDH1 Mitochondrial malate dehydrogenase precursor (by homology)** | | |
| | R.FAGAVLDGLAGEKDVIEC#TFVDSPLFK.D | 5.83 | 0.53 |
| | K.DVIEC#TFVDSPLFKDEGVDFFSTK.V | 3.59 | 0.52 |
| **CA5255** | **CaACT1 actin (by homology)** | | |
| | R.VAPEEHPVLLTEAPMNPROK&SNR.E | 4.76 | 0.44 |
| | R.DLTNHLSK+ILSER.G | 4.39 | 0.54 |
| **CA5343** | **CaRPL4B Ribosomal protein L4B (by homology)** | | |
| | R.NVPGVETASVK&HLGLLQLAPGAHLGR.F | 3.74 | 0.25 |
| | R.SGQAAFGNMC#R.G | 3.26 | 0.48 |
| **CA1982** | **CaADK1 adenylate kinase, cytosolic (by homology)** | | |
| | R.TIPQAEK+LDSMLESR.K | 5.00 | 0.52 |
| | R.DIVEK&LASLPFQK&VAHR.I | 4.24 | 0.51 |
| **CA6068** | **CaRPS19A.3 ribosomal protein S19.e, 3-prime end (by homology)** | | |
| | R.GFRPHK&HVDASGSINR.K | 4.21 | 0.46 |
| | R.PHK+HVDASGSINR.K | 3.64 | 0.46 |
| **CA0855** | **CaAMYG1 glucoamylase** | | |
| | R.K&NPROFGLLVALDAEGTASGK.L | 5.58 | 0.64 |
| | R.K+NPFGLLVALDAEGTASGK.L | 4.5 | 0.5 |
| **CA6105** | **CaFAS2.5f fatty-acyl-CoA synthase, alpha chain,** | | |

| | | | |
|---|---|---|---|
| | **5-prime end** | | |
| | R.LGQQLIDNC#K.Q | 2.78 | 0.44 |
| | R.TYDSYWNWAR.Q | 3.53 | 0.66 |
| CA0156 | **CaSCW1 glucanase (by homology)** | | |
| | K.SSC#GASAILFTAFNDLWK.A | 3.55 | 0.58 |
| CA0408 | **CaSAC6.5f actin filament bundling protein, fimbrin** | | |
| | R.LPFDTETFQIFDEC#R.D | 4.59 | 0.45 |
| CA0495 | **UCF1 unknown function** | | |
| | K.SEALPLDLDNIK&PMDHLQPVPK&TR.S | 4.66 | 0.45 |
| CA4159 | **CaALD5 aldehyde dehydrogenase (NAD+) (by homology)** | | |
| | R.VPLVC#GQIIPWNFPLLMASWK.L | 4.67 | 0.58 |
| CA0517 | **CaHEM13 by homology S. cerev.: coproporphyrinogen III oxidase** | | |
| | K.WC#DEYFFIK.H | 3.39 | 0.47 |
| CA0685 | **CaADO1 adenosine kinase** | | |
| | K.LVC#LGNPROLLDLQAK.V | 4.51 | 0.52 |
| CA0801 | **CaCHC1 clathrin heavy chain (by homology)** | | |
| | R.LAQIC#GLNLIIDAEELPELVK.T | 4.61 | 0.38 |
| CA1111 | **CaRIB3 3,4-dihydroxy-2-butanone 4-phosphate synthase (by homology)** | | |
| | R.GHTEAAVQLSTLAGLQPAGVIC#ELVR.D | 5.82 | 0.62 |
| CA1245 | **CaIMH3.exon2 IMP dehydrogenase, exon 2** | | |
| | K.GSITK&FVPYLYNGLQHSLQDIGIK.S | 4.3 | 0.34 |
| CA1484 | **CaARO4 3-dehydro-deoxyphosphoheptonate aldolase** | | |
| | R.C#IVIVGPC#SIHDPPQALEYGK.R | 3.54 | 0.36 |
| CA1541 | **CaBGL21 endo-beta-1,3-glucanase (by homology)** | | |
| | K.EALQNYLPK&ISVSTIK.I | 4.14 | 0.55 |
| CA1716 | **IPF17237 unknown function** | | |
| | R.K&ELNHLEEEGQNLLSQFVAYVK.S | 4.80 | 0.57 |

| | | | |
|---|---|---|---|
| CA1755 | CaPDI1 protein disulfide-isomerase precursor (by homology) | | |
| | R.GK&LNIVGLDASLFGR.H | 3.14 | 0.48 |
| CA1845 | IPF11123 similar to Saccharomyces cerevisiae Sdh2p suc | | |
| | K.C#GPMVLDALLK.I | 3.24 | 0.34 |
| CA2270 | CaVMA4 H+-transporting ATPase E chain | | |
| | R.ILSTK+DEVLHEIFDEAEAELK.K | 3.31 | 0.43 |
| CA2391 | CaADH5 probable alcohol dehydrogenase (by homology) | | |
| | K.ASGLNLK&K&DLPVNKPGAGQLLLK.V | 4.44 | 0.33 |
| CA2457 | CaMCR1 NADH-cytochrome-b5 reductase (by homology) | | |
| | K.SIALIGGGTGITPLYQLLHQITSNPROK&DNTK.V | 3.48 | 0.24 |
| CA2522 | CaGCV1 glycine cleavage T protein (by homology) | | |
| | R.IGLTSK+GPSPR.D | 2.89 | 0.41 |
| CA2558 | CaRBT5 repressed by TUP1 protein 5 | | |
| | R.IYDQLPECAK+ECVK.Q | 4.08 | 0.32 |
| CA2645 | CaYNK1 Nucleoside diphosphate kinase (by homology) | | |
| | R.NVC#HGSDSVESANK.E | 4.29 | 0.52 |
| CA2661 | CaAAT1 aspartate aminotransferase (by homology) | | |
| | K.IPSFALC#QSFAK.N | 3.90 | 0.50 |
| CA2675 | CaGSP1 GTP-binding protein (by homology) | | |
| | K.SNYNFEK&PFLWLAR.K | 4.13 | 0.50 |
| CA2861 | CaHHT21 Histone H3 | | |
| | R.FQK+STELLIR.K | 3.22 | 0.46 |
| CA3199 | CaFBP1 Fructose-1,6-bisphosphatase | | |
| | R.VLYEC#FPMALLMEQAGGSAVTIK.G | 3.81 | 0.42 |
| CA3284 | CaADE6 5 -phosphoribosylformyl glycinamidine synthetas | | |

| | | | |
|---|---|---|---|
| | K.MC#VAESLLNIFAADIPSLDHVK.L | 4.50 | 0.45 |
| CA3333 | CaCDC48 microsomal ATPase (by homology) | | |
| | R.AAAPTVVFLDELDSIAK+AR.G | 4.76 | 0.63 |
| CA3469 | CaMNT1 Mannosyltransferase involved in n-linked and o-linked glycosylation | | |
| | K.LYC#DIDYDIFK.W | 3.16 | 0.41 |
| CA3479 | CaATP4 F1F0-ATPase complex, F0 subunit B (by homology) | | |
| | R.IETVSELK&NVVATTEDLFALSK.E | 5.15 | 0.45 |
| CA3746 | CaARA1 D-arabinose dehydrogenase (by homology) | | |
| | R.YIC#DPWGYGIGFR.W | 3.38 | 0.49 |
| CA3978 | CaCPR3 cyclophilin (peptidylprolyl isomerase), mitochondrial (by homology) | | |
| | K.KIESC#GTSSGTPSASVVIEESGEAEK | 6.66 | 0.60 |
| CA4078 | CaVMA2 H+-transporting ATPase (by homology) | | |
| | K.IPIFSASGLPHNEIAAQIC#R.Q | 6.55 | 0.59 |
| CA4123 | CaAMS1 alpha-mannosidase (by homology) | | |
| | R.K+LAEQLLGDK.I | 2.56 | 0.22 |
| CA4127 | IPF6629 unknown function | | |
| | K.DHASLTAC#ANPROIPLDLK.S | 4.91 | 0.54 |
| CA4349 | CaFUM12.5f Fumarate hydratase, 5-prime end (by homology) | | |
| | K.VNEELGALDPK&LSAAIQQAATEVAEGK.L | 5.28 | 0.50 |
| CA4378 | IPF11299 unknown function | | |
| | R.K+VLHHEEPTVR.D | 3.40 | 0.31 |
| CA4416 | IPF11888 unknown function | | |
| | R.VK&FPIIADPER.K | 3.29 | 0.44 |
| CA4513 | CaADE17 5-aminoimidazole-4-carboxamide ribotide transferase | | |
| | K.VDFVVC#NLYPFK.E | 3.26 | 0.40 |
| CA4570 | IPF9550 similar to Saccharomyces cerevisiae Osm1p | | |

| | | | |
|---|---|---|---|
| | osmotic growth protein | | |
| | R.LGGSSLLEC#VVFGR.F | 3.70 | 0.50 |
| CA4572 | CaKES1 involved in ergosterol biosynthesis (by homology) | | |
| | K.K&PLNPROFLGEVFVGK.W | 3.51 | 0.48 |
| CA4959 | CaHSP90 heat shock protein | | |
| | R.ELISNASDALDK+IR.Y | 3.63 | 0.54 |
| CA5050 | CaBMH2 similar to Saccharomyces cerevisae Bmh2p suppre | | |
| | K.IC#EDILSVLSDHLITSAQTGESK.V | 5.74 | 0.60 |
| CA5587 | CaSOD21.5f Superoxide dismutase by homology, 5-prime end | | |
| | K.ISLPK&IDWALDALEPYISK.E | 4.16 | 0.53 |
| CA5638 | IPF470 putative glutamine-tRNA ligase (by homology) | | |
| | K.MYPIYDFC#VPVVDSIEGVTHALR.T | 4.71 | 0.61 |
| CA5816 | CaACC1 acetyl-coenzyme-A carboxylase (by homology) | | |
| | R.LPPK&LDDGLTALVER.T | 3.94 | 0.50 |
| CA5914 | CaHHT3 histone H3 | | |
| | R.K+SAPVSGGVK+K+PHR.Y | 5.06 | 0.21 |
| CA5947 | CaKRS1 Lysyl-tRNA synthetase (by homology) | | |
| | R.IAPELFLK+ELVVGGMDR.V | 2.95 | 0.44 |
| CA6000 | CaFDH11.3 glutathione-dependent formaldehyde dehydrogen | | |
| | R.ATQGK+GVMPDGTPR.F | 4.43 | 0.58 |
| CA6010 | CaTRX1 thioredoxin (by homology) | | |
| | K.ENNLVIVDFFATWC@GPC#K.M | 3.44 | 0.62 |
| CA2082 | CaNTF2 nuclear transport factor (by homology) | | |
| | R.DIVEK&LASLPFQK&VAHR.I | 4.24 | 0.51 |
| CA4783 | IPF3358 ubiquinol-cytochrome-c reductase (by homology) | | |
| | K.QYLSAQASAVEADPK+SK.V | 5.25 | 0.06 |

Note: K& indicates that lysine residue was labeled by sulfo-NHS-LC-biotin. K+ indicates that lysine residue was labeled by sulfo-NHS-SS-biotin. C# indicates that cysteine residue was labeled by iodoacetyl-$PEO_2$-biotin. C@ shows that cysteine residue was modified by IA before trypsin digestion. M* shows that methionine was oxidized. Green highlighted proteins were those reported for the first time.

# VITA

Jiang Qian was born in Zhejiang, China on Aug 25, 1977. He obtained his B.S. in Chemical Engineering from Henan University of Technology in 2000. He then worked at an administrative committee for the Hangzhou Bay Chemical Developing Zone for a year. Later he went back to the Henan University of Technology in fall 2001 and studied for his M.S. degree in protein engineering till graduation in 2004. Upon completion his M.S he was admitted to the Ph.D program in chemistry at University of New Orleans. He completed his Ph.D study under the supervision of Dr. Yang Cai in 2009.