Characterization of some novel GlcNAc inducible genes in *Candida albicans*

Thesis submitted to Jawaharlal Nehru University for the award of the degree of

DOCTOR OF PHILOSOPHY

PRIYA RANI

SPECIAL CENTRE FOR MOLECULAR MEDICINE

Jawaharlal Nehru University

New Delhi-110067, INDIA

2017

SPECIAL CENTRE FOR MOLECULAR MEDICINE
Jawaharlal Nehru University

CERTIFICATE

The research work embodied in this thesis entitled "**Characterization of some novel GlcNAc inducible genes in *Candida albicans***" has been carried out by Ms. Priya Rani under my guidance at the Special Centre for Molecular Medicine, Jawaharlal Nehru University, New Delhi, India. The work presented here is original and has not been submitted in part or full for any degree or diploma of any university/ Institution elsewhere.

Date: 21·07·17

Prof. Asis Datta

(Supervisor)

Prof. Suman K. Dhar

(Supervisor)

Priya Rani

(Candidate)

Prof. Vibha Tandon

(Chairperson)

Dedicated to my Parents, Teachers,

& Family…

Acknowledgements

Foremost, I am immensely grateful to my supervisor, Prof. Asis Datta for accepting me as his PhD student. I feel privileged to work under his guidance and to learn from his expertise in the field of research in Candida albicans. I acknowledge with gratitude his unwavering encouragement and for providing me with maximum facility to work. I sincerely wish to inculcate his positive temperament and passion for Science.

I am greatly indebted to Prof. Suman K. Dhar for showing faith in me and being my co-supervisor. I acknowledge his goodwill and advices.

Besides my supervisors, I would like to thank the rest of my doctoral committee: Prof. Debi P. Sarkar and Prof. K. Natarajan, for their insightful comments and encouragement, which incented me to widen my research from various perspectives.

I thank the present chairperson of SCMM Prof. Vibha Tandon and former chairpersons Prof. Chinmmay K. Mukhopadhyay and Prof. Suman K. Dhar for providing a scientifically rich environment. My sincere thanks to all the administrative and technical Staffs of SCMM for their relentless efforts to run the centre smooth. Special thanks to Mr. Nand Kumar, Niti, Naresh ji and Dr. Naminita Lahon, for providing the essential things at the right time.

With a special mention to Prof. S Gourinath, I sincerely thank him for providing me an opportunity to work in his lab, without his precious support it would not be possible to conduct the crystallization work. Gunjan has been very co-operative and helpful throughout our

collaboration work at his lab. I really enjoyed working in the positive and friendly environment of his lab.

I thank my fellow Labmates and Seniors, Dr. Hanumant, Dr. Swagata, Dr. Irfaan, Dr.Anil, Dr. Vinay, Dr.Abira, Karishma, Seikhu, DN Bhatt and Priyanka for their kind help and support at the time of need. I also thank Vipin and Dr. Sarjeet for their cooperation.

I feel lucky to have Sanju, Rajesh, Khadija, Shalini and Meetu as my batchmates. I enjoyed their company in free time. I would also thank Aakansha, Aaditi and Reema di for the good times we had.

My special thanks to Pallavi for her support and care. It is difficult to imagine my final years of Ph.D without you.

I extend my thanks to Mohit and all my friends from JNU who made my stay in the campus lively and memorable.

Last but the most important, I would like to thank my family: my parents and to my brother and sister for supporting me throughout my PhD and my life in general. They have always been my pillars of strength that kept me encouraged in my tough times.

Above all I thank the almighty God!

CONTENTS

Abbreviations

μ	micro
Å	Angstrom
APS	Ammonium per sulphate
β-ME	Beta- mercaptoethanol
bp	Base pair
BSA	Bovine serum albumin
CCP4	Collaborative Computational Project 4
COOT	Crystallographic Object-Oriented Toolkit
°C	degree Centigrade
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra amine
EtBr	Ethidium Bromide
GlcNAc	N-Acetyl glucosamine
HCL	Hydrochloric acid
IAA	Isoamylalcohol
IPTG	Isopropyl β-D-1- thiogalactopyranoside
Kb	Kilo Base
kDa	kilo Dalton
KCl	Potassium Chloride
M	Marker
$MgSO_4$	Magnesium Sulphate
MQ- Water	Milli-Q water

NaCl	Sodium Chloride
Na_2HPO_4	Sodium Hydrogen Phosphate
NaOAc	Sodium Acetate
NaOH	Sodium hydroxide
O.D.	Optical density
PAGE	Polyacrylamide Gel
PBST	Phosphate buffered Saline with Tween 20
PDB	Protein Data Bank
PEG	Polyethylene glycol
PMSF	Phenyl Methyl Suphonyl Flouride
rpm	Revolutions per minute
RMSD	Root Mean Square Deviation
SDS	Sodium Dodecyl Sulphate
Tris	Tris (hydroxymethyl) amino ethane
TAE	Tris-acetate-EDTA
TBE buffer	Tris/Borate/EDTA buffer
TBST	Tris Buffered Saline with Tween 20
TEMED	N,N,N',N', Tetramethylethylenediamine

REVIEW OF LITERATURE

*Candida albicans***: An insidious human fungal pathogen**

The epidemiology of invasive mycoses has changed over the past several decades. Fungal diseases are emerging as global threat to public health. With the rise in the population of patients undergoing organ or hematopoietic stem cell transplantation, chemotherapy, recipients of immunosuppressive therapy, major surgery, HIV and premature birth; fungal pathogens have become a major cause of nosocomial infection worldwide (Pfaller, 1996; Wisplinghoff et al., 2014). Medical conditions like diabetes, use of antibiotics, pregnancy and old age also increases the risk of fungal infection.

Around 600 species of fungi are human pathogens, *Candida albicans* being one of the most common. The species *C. albicans* belongs to family Saccharomycetaceae of Ascomycota Division in Fungi kingdom. *Candida albicans* usually resides as a harmless commensal organism in the micro-biome of most, if not all healthy individuals. It normally lives in the warm and moist surfaces of the skin, the mucosa of the oropharyngeal cavity, gastrointestinal tract and the genitourinary tract. However, this opportunistic fungus turns pathogenic in individuals with compromised immunity (Soll, 2002).

1.1 Various infections caused by *C. albicans* **-**

Localized infections:
This type of Infection is restricted to one location and is non-lethal. Superficial infection the skin by *Candida albicans* is termed 'Cutaneous candidiasis' and is most common cause of diaper rashes in infants. Cutaneous candidiasis is very common in warm moist creased areas of the skin like armpits and groin; it also affects nails, nail edges and corner of mouth (Naglik et al., 2011). Predisposing factors for cutaneous candidiasis includes diabetes, obesity, chemotherapy, steroid therapy and antibiotic therapy.
C. albicans colonizes the oral cavity of nearly 75% of the population. While healthy individuals remain asymptomatic, individuals with slightly compromised immunity can suffer from recalcitrant infection of oral candidiasis (OC) also called 'Thrush' (Cannon et al., 1995). In cases

of dysfunctional adaptive immune system like HIV the infection can also spread to oropharynx and or esophagus. Wearing dentures and old age also adds on to the risk of getting OC (Nasution, 2013).

In case of women, approximately 75% of all women suffer from vulvovaginal candidiasis (VVC) or commonly called yeast infection at least once in their lifetime, with high chances of reoccurrence (Achkar and Fries, 2010). High risk factors for VVC include diabetes, pregnancy, use of antibiotics, use of oral contraceptives and hormonal drugs.

Disseminated candidiasis

In contrast to superficial candidiasis, Systemic or disseminated candidiasis occurs when the fungi enters the blood stream affecting multiple organs resulting in very high crude mortality rate even with antifungal therapy (Campion et al., 2015; Lim et al., 2012). Patients with hospital stay especially in the intensive care unit, use of central venous catheters, gastrointestinal surgery (Yan et al., 2013), broad spectrum antibiotic therapy and neutropenia are highly susceptible to bloodstream infection by *C. albicans*.

Mortality

Recovery from localized mucocutaneous infection is possible by treatment with antifungals like fluconazole. On the contrary invasive candidemia is much more life-threatening. According to a study, the mortality rate for patients with systemic candidiasis was about 34%. Shockingly, when treatment is delayed the statistics raised to almost double; i.e., the mortality rate was around 78% when treatment was delayed for more than 48 hours (Pfaller and Diekema, 2007).

Treatment of *Candida albicans* and other fungal pathogens is extremely difficult since they are eukaryotes and share many of their biological processes with humans. Most antifungal drugs have toxic side effects and can be given to a dose that is usually fungistatic rather than fungicidal (Cannon et al., 2007). The chances of reoccurrence of infection are very high. Thorough understanding of the molecular basis of the mechanism of infection along with the biology of the pathogen is needed to find novel drug targets.

Attachment/ colonization

Biofilm formation

Invasion : Superficial infection

Deep seated infection
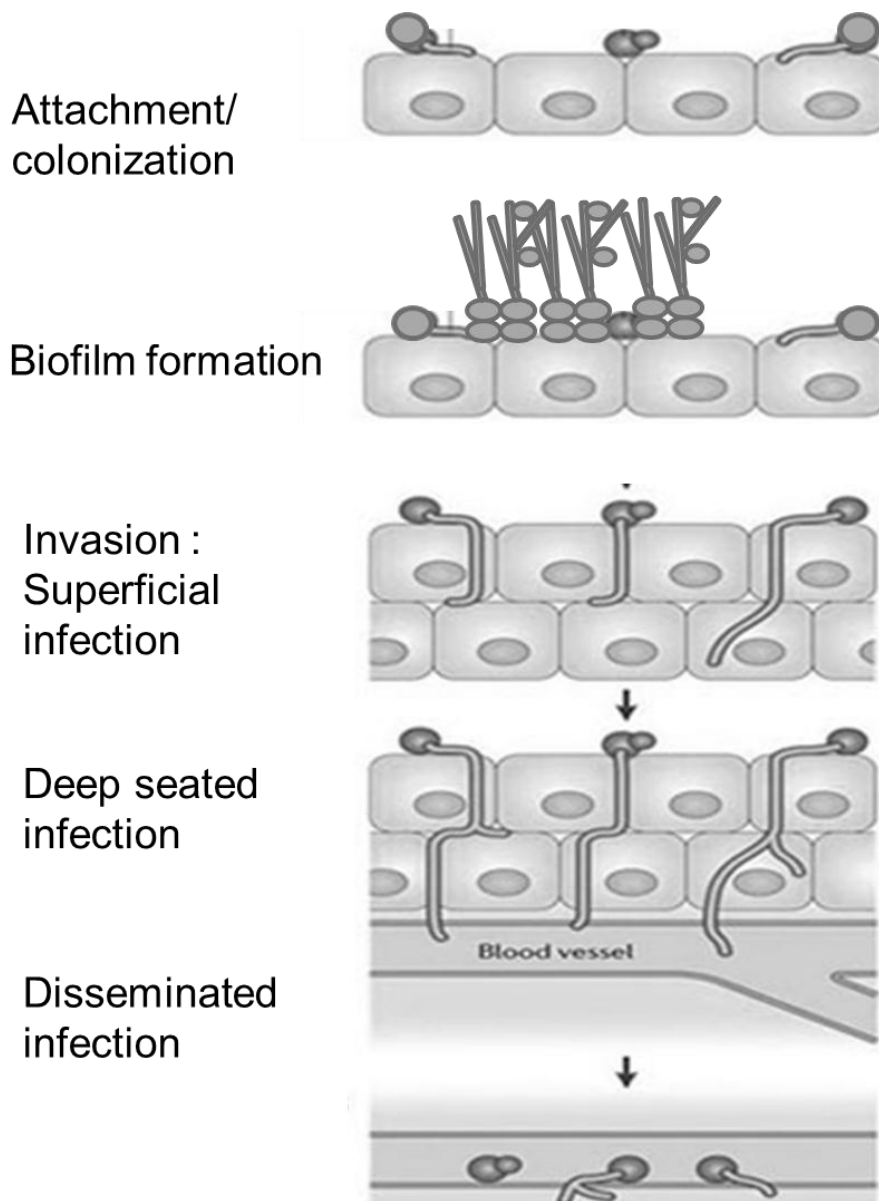
Disseminated infection

Blood vessel

Figure 1: Graphic representation of *Candida albicans* stages of infection. First two stages are commensal and later are pathogenic.

1.2 Genome of *C. albicans*

Candida albicans was amongst the first eukaryotic pathogens chosen for whole genome sequencing. The genome of *C. albicans* was sequenced by Stanford Genome Technology Centre in a duration of more than 10years culminating in 2004; Diploid and haploid assemblies are available at http://www.candidagenome.org (Berman and Sudbery, 2002)

C. albicans is a diploid organism with a haploid genome size of around 16MB seperated in eight pairs of chromosomes (Magee and Magee, 1987); encoding 6198 ORFs. Nearly 70% of these ORFs are still uncharacterized; 20% are unique and have no known counterparts in other available genome sequences. One distinct feature of *C. albicans* genome is the anomaly in the usage of codon CUG to encode serine instead of leucine (Santos and Tuite, 1995). This characteristic is also present in a group of related species including *Candida paropsilosis* and *Candida tropicalis* collectively known as CTG clade (Butler et al., 2009). The CUG clade of these pathogenic yeasts includes a total of 21 gene families encoding lipases, adhesions, oligopeptide transporters, cell wall mannoproteins, transcription factors, and ferric reductases (Butler et al., 2009).

The genes usually lack introns. Several ORFs encoding the conserved centromere proteins in the genome sequence were first reported in 2004 (Sanyal et al., 2004). Later in 2007, the centromere region of all eight chromosomes of *C. albicans* were characterized (Mishra et al., 2007) . Sequence analysis of all eight 3-4.5 kb *CEN* (centromere) regions revealed unique DNA sequence on each chromosome lacking any conserved DNA sequence motifs indicating epigenetic mode of inheritance. *C. albicans* CEN in general lacks DNA repeat elements and transposon except a few small inverted repeats and long terminal repeats occurring in the centromeric and pericentric regions in few chromosomes. Telomere sequences and telomerase homologues have also been reported (Singh et al., 2002).

Another interesting feature of genome of *C. albicans* clinical isolates is their variable karyotype. Karyotype changes in the clinical isolates are caused by expansion and contraction of the multiple repeat sequences (MRS), and reciprocal translocation events between these MRS repeats (Chu et al., 1993). Single or double copies of MRS are present in all chromosomes except chromosome 3 (Chindamporn et al., 1998). Events like nondisjunction are quite frequent

in *C. albicans*, helping the pathogen in adapting to stressful conditions. As an example, in strains grown on Sorbose as sole carbon source, the loss of chromosome 5 carrying a repressor for sorbose is commonly observed (Janbon et al., 1998). Likewise, strains resistant to fuconozol have increased chances of losing chromosome 4 or gaining chromosome 3 (Perepnikhatka et al., 1999).

Polymorphism in *Candida albicans*

Polymorphism a common trait amongst fungi is also evident in *Candida albicans*. Some fungi live as unicellular yeasts but most are multicellular filamentous organisms. Some fungi including *Saccharomyces cerevisiae* are dimorphic, living as both yeast and filamentous hyphae form. *Candida albicans* is polymorphic or pleomorphic; four classic cell types are Yeast, Pseudohyphae, Hyphae and Clamydospores. While Yeast and Hyphae type are well characterized, pseudohyphae and chlamydospores are somewhat less studied. Each of these cell types has different morphology, mode of division and virulence.

Yeast cell type has morphology similar to *Saccharomyces cerevisiae* i.e., round to oval. They reproduce by budding. After genome duplication the nucleus divides at the junction between the mother and budding daughter cells followed by cytokinesis giving rise to two unicellular yeast cells.

Hyphae cell types are thin and tube shaped. Cell division does not involve equal nucleus distribution between the mother-daughter cells. Even after cytokinesis hyphal cells remain tightly attached to one another in an end to end manner. After several rounds of cell division they form multicellular filamentous structures called mycelia.

Pseudohyphae cell types are ellipsoidal in shape and have features of both yeast and hyphae. Unlike Yeast and hyphae, *in vitro* conditions to induce stable, pure population of pseudohyphae are not known. Similar to hyphae, pseudohyphal cells also generate mycelia after multiple rounds of cell division and cells remain attached following cytokinesis. Similar to yeasts, nuclear division in pseudohyphae occurs at mother–daughter junctions and unlike hyphae, these junctions are demarcated by visible constrictions. In hyphal cells mother-daughter or septal junctions are not constricted and the cel walls are parallel.

Chlamydospores are large, spherical thick walled cells. *In vitro*, they are observed when the cells are under certain stress conditions like starvation and hypoxia. The cells at the distal ends of the mycelial filament also called suspensor cells produces chlamydospores. Within the suspensor parent cell the nuclear division takes place and the progeny nucleus migrates to the nascent chlamydospore, which stays attached to its mother cell.

Another recently identified cell type is GUT or gastrointestinally induced transition, which is formed when the *Candida albicans* white cells are exposed to mammalian gut (Pande et al., 2013).
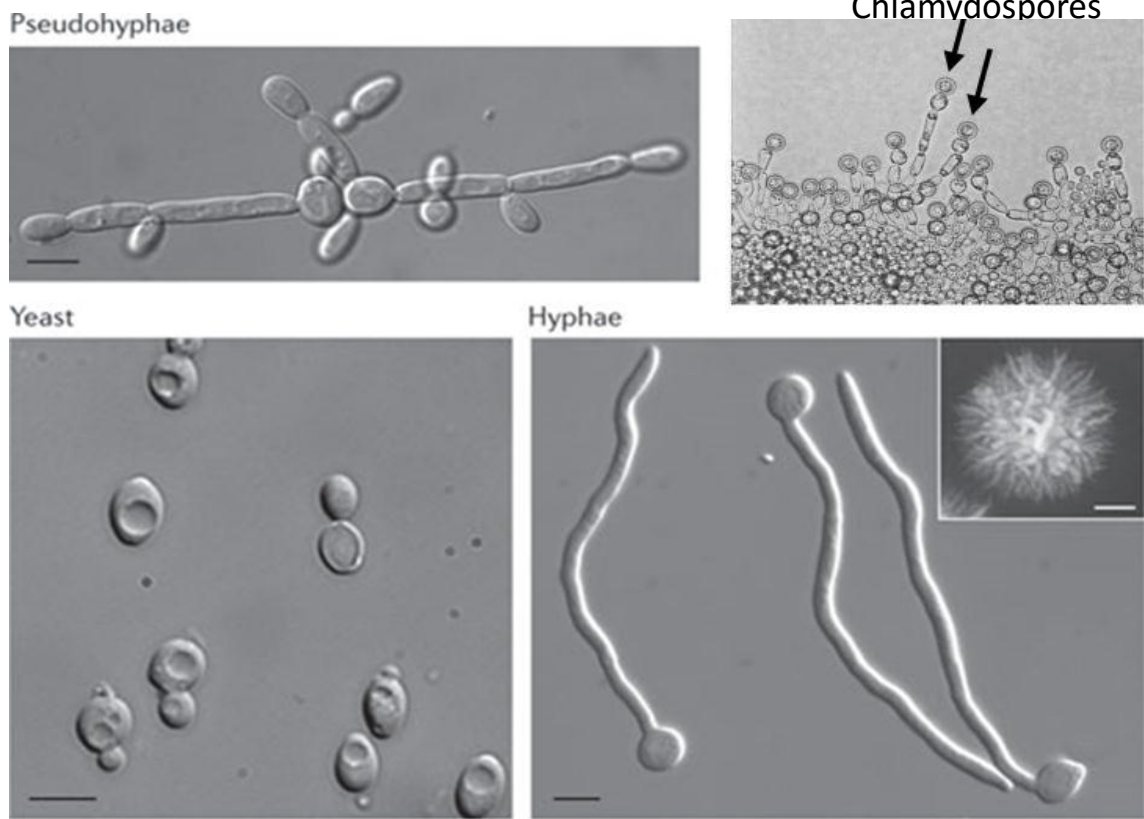
Figure 2: Morphology of Pseudohyphae, Chlamydospores, Yeast and Hyphae cell types. Inset in hyphae panel shows colony morphology of hyphae type on Spider medium (Sudbery, 2011)
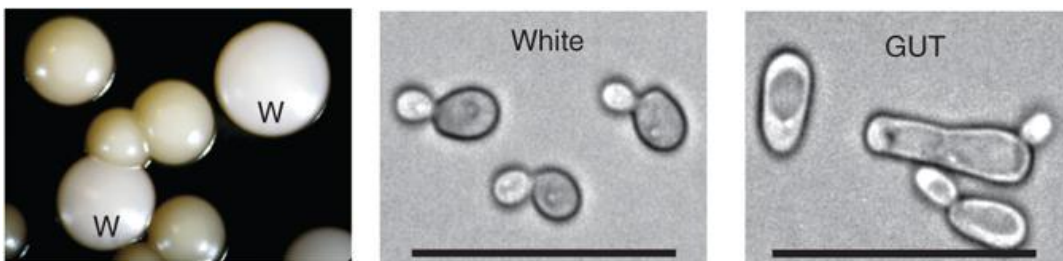

Figure 3: Colony morphology of GUT cells with white cells; light micrograph of white and GUT cell types (Pande et al., 2013)

Virulence and morphogenesis

The three cell types - yeasts, hyphae and pseudohyphae can not only propagate stably as the same cell type but can also interconvert, sensing cues from their microenvironment. Traditionally, hyphae and pseudohyphae forms were considered virulent while yeast form was considered commensal. Hyphae form can grow invasively on solid media and their tips exhibit characteristic of contact sensing or Thigmotropism. Hyphae type also express virulence factors like adhesins (Hyphal wall protein1 (Hwp1), agglutin like protein 3(Als3), Als10, factor activated 2(Fav 2) and Pga55), tissue lytic enzymes (secreted aspartyl protease 4 (Sap4), Sap5 and Sap6), oxidative stress response proteins (superoxide dismutase 5 (Sod5)), and cytolytic peptide toxin (extent of cell elongation protein 1 (Ece1)). The ability to interconvert between different cell types is required for virulence. Thus hyphal cells can actively penetrate epithelial barrier while the yeasts can only colonize tissue surfaces without causing any damage.

However in disseminated candidiasis, yeast, hyphae and pseudohyhae all forms seem to contribute to pathogenesis. All three cell types have been reported to be present infected tissues of both human patients as well as animals with disseminated candidiasis.

Moreover, mutants that are trapped as either yeast (lacking transcription factors Cph1 and Efg1 or cyclin Hgc1) or hyphae (lacking tup1 repressor) are both defective in bloodstream infection models, linking both form to pathogenesis. These findings suggest that the ability to interconvert between different cell types is required for pathogenesis.

Mating and parasexual life of *C. albicans*

Meiosis an essential event for sexual reproduction in organisms, is absent in C. albicans classifying it as Fungi imperfecti. For a long time until the identification of *mating type locus* (*MTL*) (Hull et al., 2000)*;* C. albicans was considered to be asexual. Later a *natural mating-*competent form was identified that mates naturally at a high frequency forming a tetraploid gamete revealing the alternative parasexual mode of reproduction in *C. albicans.* Till now, attempts to validate meiosis, and thus a complete sexual cycle, have not been successful. Even though, homologs of the whole collection of genes essential for meiosis in *Saccharomyces cerevisiae* are present in *C. albicans* genome.

Naturally occurring *C. albicans* strains are mostly diploid and heterozygous (**a**/α) at the mating type locus (MTL) and thus unable to mate or switch from white to opaque form (Anderson and Soll, 1987).

The white- opaque transition

Initially, the white-opaque transition was thought to be a cellular and colony morphology transition, limited to some specific strains. But the switch is epigenetic and each state is stable but capable of transitioning to the other state with a frequency rate much higher than that of standard mutation. The frequency can also be altered by external conditions like temperature. White-form cells are classic ovoid shaped, and form domed colonies, creamy in color. Opaque-form cells are elongated and have a pitted cell wall, not like the white form cells that have relatively smooth cell wall surface. Colonies formed by Opaque-form cells are flattened and more gray compared to the colonies generated by white-form cells. The opaque-form cells have reduced viability relative to white-form cells under many growth conditions. This feature allows in distinguishing colonies of the opaque-form cells as they can be more readily stained by vital dyes such as phloxine B.

At molecular level the ability to transition from white to opaque form is determined by the fact whether the cells are homozygous for the MTL locus. The MTL locus controls cell type; cells that are heterozygous for the MTL locus (MTLa/MTLα) are unable to switch. The inability of MTL heterozygous strains to switch is regulated by the repressor a1/α2. The a1 protein encoded
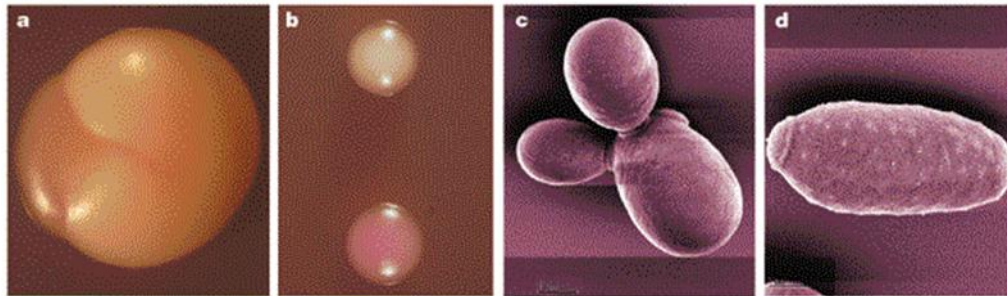
Figure 4: a. Colony morphology of a white colony with opaque sector; b. upper white, lower opaque colony on phloxin B; Scanning electron micrograph of c. white and d. opaque cells (Johnson, 2003)**.**

by the MTLa locus, and the α2 protein encoded by the MTLα locus assemble together to form the hetrodimeric repressor a1/α2; thus only heterozygous strains can form it. This repressor a1/α2 controls the expression of another transcription factor, encoded by the WOR1 gene. TheWor1p transcription factor acts as a primary controller of the opaque state. Ectopic expression of Wor1p in MTL homozygotes efficiently induces the opaque state, and even in MTL heterozygotes a pseudo-opaque state can be triggered by activating Wor1p expression. The epigenetic characteristics of the opaque state can be explained by the regulation of Wor1p. Chromatin immunoprecipitation experiments have found as many as five binding sites for Wor1p in its own promoter region suggesting that Wor1p is autoregulated. It means once Wor1p expression is established, a positive-feedback loop starts and it tends to remain on, keeping the cell in the opaque state.

Comparative analysis of the gene expression profiles of white and opaque form cells identifies many genes whose expression switches between cells in the white form and cells in the opaque form. Many of these genes are involved in controlling the distinct physiologies of the two different cell types. White-form expresses genes characteristic of a fermentative life style, on the other hand opaque form shows characteristics of an oxidative metabolism. Yet, many other genes expressed differently between the two cell types probably define the different growth patterns and cell surface structures displayed by the white- and opaque-form cells.

The transcription regulator, Efg1p a critical positive controller of the yeast-hyphae transition and theTup1p that is a negative regulator of pseudohyphal development are also involved in the white-opaque switching. Efg1p expression is high in white-form cells but it is not expressed in opaque-form cells; this expression is under control of a strong, white-phase specific promoter. Loss of EFG1 expression in white-form cells causes them to express some of the morphologies of opaque form cells, particularly the elongated cell shape, along with the expression of some opaque-specific genes. Ectopic overexpression of Efg1p shifts opaque-form cells to the white morphology which suggests that repression of EFG1 expression is apparently necessary to establish the opaque state. However, certain structural characteristics of the opaque state, particularly the surface pimples, were absent in the efg1-null mutants. Therefore, seemingly Efg1p may be regulating only a subset of the genes required in establishing the opaque state.

The Tup1p is another transcriptional regulator that influences the opaque state (Braun et al., 2000) (Braun and Johnson, 2000). Loss of Tup1p, produces constitutive pseudohyphae like state in white-form cells and also drastically affects the morphology of opaque-form cells. Expression of some phase-specific genes is deregulated by loss of tup1 but the establishment of a mating-competent cell type is still permited.

White opaque regulator 1 (Wor1), acts opposite to Efg1; Efg1 and Wor1 oppose each other in a complex regulatory circuit along with many other transcription factors to control white opaque switching. Other transcription regulators includes a1-α2, Wor2, Wor3, Wor4, Ahr1 (adhesion and hyphal regulator 1), Czf1 (Brown et al., 1999) and many more. Wor1 expression is induced by Wor2, Wor3 and Wor4 to promote white-opaque switch; Wor1 inhibits Efg1 thereby inhibiting opaque to white transition. In contrast, Efg1 promotes opaque to white switch by inhibiting Wor1, thus inhibiting the white to opaque switch.

Table 1 | Features of *Candida albicans* cell types

	MTL locus Genotype	Cell shape	Unicellular versus multicellular	Special morphological features	In vitro inducing signals	Special functions	Host interactions
Yeast (white(a/α))	a/α	Round-to-oval	Unicellular	N/A	Default cell shape under most in vitro conditions	Biofilm formation (conventional)	Virulence (bloodstream model); commensalism (mouth, skin, vagina and gastrointestinal tract)
Hypha*	a/α	Tube	Multicellular	N/A	37°C, N-acetylglucosamine, serum, immersion in agar, hypoxia, hypercarbia and alkaline pH	Thigmotropism; biofilm formation (conventional)	Induced endocytosis; active penetration of host epithelial cells; virulence (mouth, vagina and bloodstream models)
Pseudohypha*	a/α	Elongated ellipsoid	Multicellular	Indented cell–cell junctions	Hypha-inducing cues‡	Biofilm formation (conventional)	Virulence (mouth, vagina and bloodstream)
Chlamydospore	a/α	Round	Multicellular§	Thick cell wall	Nutrient scarcity, hypoxia	Unknown	Unknown
White(a) and white(α)	a/Δ, a/a and α/Δ, α/α	Round-to-oval	Unicellular	N/A	37°C, glucose and alkaline pH	Biofilm formation (sexual)	Unknown
Opaque(a) and Opaque(α)	a/Δ, a/a and α/Δ, α/α	Ellipsoid	Unicellular	Surface pimples	N-acetylglucosamine, hypercarbia and acidic pH	Mating	High fitness in a neonatal mouse skin colonization model
Opaque(a/α)	a/α	Ellipsoid	Unicellular	Surface pimples	Nutrient scarcity, N-acetylglucosamine and hypercarbia	Unknown	High fitness in a neonatal mouse skin colonization model
Grey(a/α)	a/α	Ellipsoid	Unicellular	Smallest cell type	Nutrient abundance	Unknown	High fitness in an ex vivo tongue infection model
GUT	a/α	Ellipsoid	Unicellular	N/A	Unknown	Unknown	High fitness in a mouse gastrointestinal commensalism model

GUT, gastrointestinally induced transition; MTL, mating-type-like. *Please note that **a** and α-cells form hyphae and pseudohyphae under certain environmental conditions, but these cells types have not been well characterized. ‡Pseudohyphae arise as a subpopulation under most hypha-inducing conditions. §Chlamydospores are produced by the terminal cells of hyphae and pseudohyphae under nutrient-poor and oxygen-depleted conditions.

(Noble et al., 2016)

N-ACETYLGLUCOSAMINE: A MULTIFUNCTIONAL AMINO SUGAR

N-acetyl-D-glucosamine (GlcNAc or NAG) is a monosaccharide derivative of glucose; an amide between glucosamine and acetic acid. It has a molecular formula C8H15NO6 with a molar mass of 221.21 g/mol. GlcNAc plays various significant roles across several biological systems. It is an abundant hexose that as a monomer or as part of macromolecules plays fundamental role of maintaining cell structure in both prokaryotes and eukaryotes. It is a key component of bacterial cell wall peptidoglycan, fungal cell wall chitin, and extracellular matrix of animal cells.

Functions of GlcNAc in *Candida albicans*

As in other organisms, in *Candida albicans* also, GlcNAc is involved in a plethora of processes including structural function, energy metabolism and signaling.

1. As a Structural component

GlcNAc plays critical role in the formation of fungal cell wall. It is the building block for chitin, the innermost layer of the fungal cell wall. Chitin is a polymer of ββ-(1,4) linked GlcNAc. It is also involved in post translational modification of cell-surface proteins. It is a part of the N-linked polysaccharide chain which is added to glycosylated proteins and is a key component in synthesis of GPI-anchors that tethers proteins in plasma membrane. GPI anchored proteins are crucial in maintaining cell wall integrity and various other functions. Chitin, N-linked glycans, and GPI anchors are all synthesized using Uridine di-phosphate (UDP)-GlcNAc as a substrate. Thus UDP-GlcNAc, the end product of Leloir pathway (Milewski et al., 2006) is a crucial metabolite and its synthesis is tightly regulated.

2. **As a Carbon source**

Candida albicans can utilize GlcNAc as a sole carbon source. The mucosa, the primary sites of *Candida albicans* infection are rich in amino sugars like glucosamine and N-acetyl glucosamine. These amino sugars are believed to be produced through cell wall remodelling by endogenous bacteria in gastrointestinal tracts (Ghuysen, J.M; Hakenbeck, R .1994).

The GlcNAc catabolic pathway:

For the breakdown of GlcNAc in *C. albicans*, first it is transported inside the cells by GlcNAc permease *NGT1* (Alvarez and Konopka, 2007) and then the sequential action of GlcNAc kinase, GlcNAc-6-phosphate deacetylase (Yamada-Okabe et al., 2001) and glucosamine-6-phosphate deaminase (Natarajan and Datta, 1993), converts GlcNAc to fructose 6-phosphate which is then metabolized through the glycolytic Embden-Meyerhoff pathway.

The *NAG* regulon: A fascinating feature of this pathway in *C. albicans* is that all the genes involved in the catabolism of GlcNAc exist in a cluster (Kumar *et al.*, 2000). This cluster consists a set of six genes. Three of these genes are glucosamine-6-phosphate deaminase (*NAG1/CaNAG1*), GlcNAc-6-phosphate deacetylase (*DAC1/CaNAG2*) and GlcNAc kinase (*HXK1/CaNAG5*). The other three genes (*CaNAG3, CaNAG4* and *CaNAG6*) in the cluster were shown to be involved in drug sensitivity (Yamada-Okabe and Yamada-Okabe, 2002). *CaNAG3* and *CaNAG4* share significant sequence homology with *S. cerevisiae* Ypr156C and Ygr138C, moreover the two genes (*CaNAG3* and *CaNAG4)* also share more than 80% homology with each other. CaNAG3/TMP1 and CaNAG4/TMP2 function as multidrug efflux pumps (Sengupta and Datta, 2003). Surprisingly, homologs of *NAG1*, *DAC1* and *HXK1* are not found in *S. cerevisiae*, suggesting that either the pathway is absent or is present in an alternate form in *S. cerevisiae*. Another peculiar feature of the NAG cluster is that *NAG1* reading frame is opposite to that of *DAC1*, suggesting that the two genes are transcribed divergently from a bi-directional promoter (Kumar et al., 2000). *CaNAG3*, *CaNAG4*, *CaNAG5* and *CaNAG6* are transcribed from the distal to the proximal region, while *CaNAG1* and *CaNAG2* are transcribed from the proximal to the distal region of the cluster (Figure 5).

Mutants with disruption of the part of nag cluster containing *DAC1*, *NAG1* and *HXK1* were unable to grow on aminosugars, exhibited reduced hyphae formation under GlcNAc-inducible conditions and displayed avirulence in a murine model of systemic candidosis. Other than this, under other filament-inducing conditions excluding GlcNAc, the mutants showed hyperfilamentation and change in colony morphology. Conclusively, disruption of the pathway alters both virulence and morphogenesis of *C. albicans* (Singh et al., 2001).
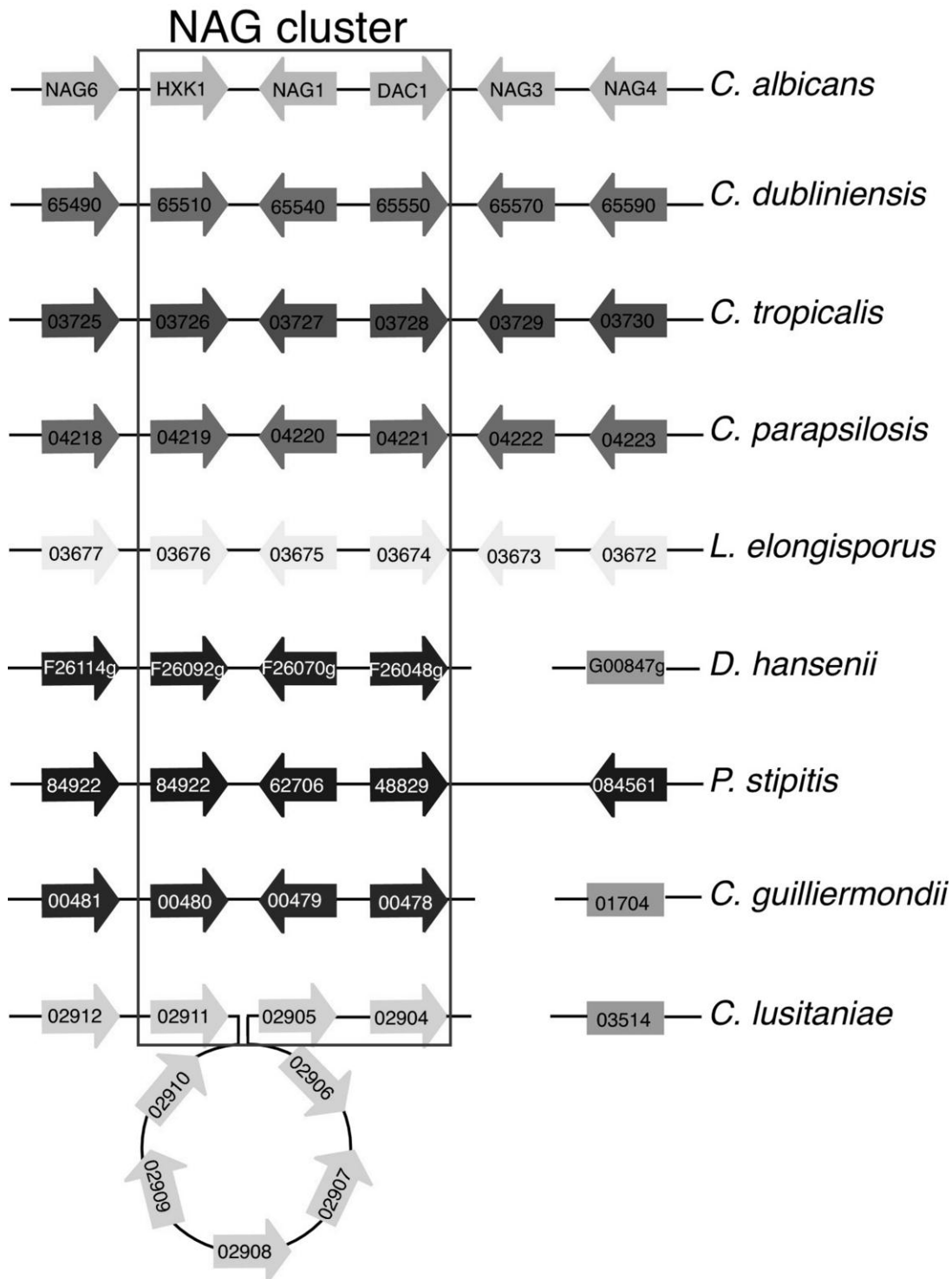
Figure 5: Gene order around the *N*-acetylglucosamine (NAG) cluster. Blocks represent chromosomes, homologs are organized in pillars Synteny of NAG enzymes is observed in all species except for *C. lusitaniae*, which has 5 intervening genes displayed as an insertion loop.(Fitzpatrick et al., 2010)
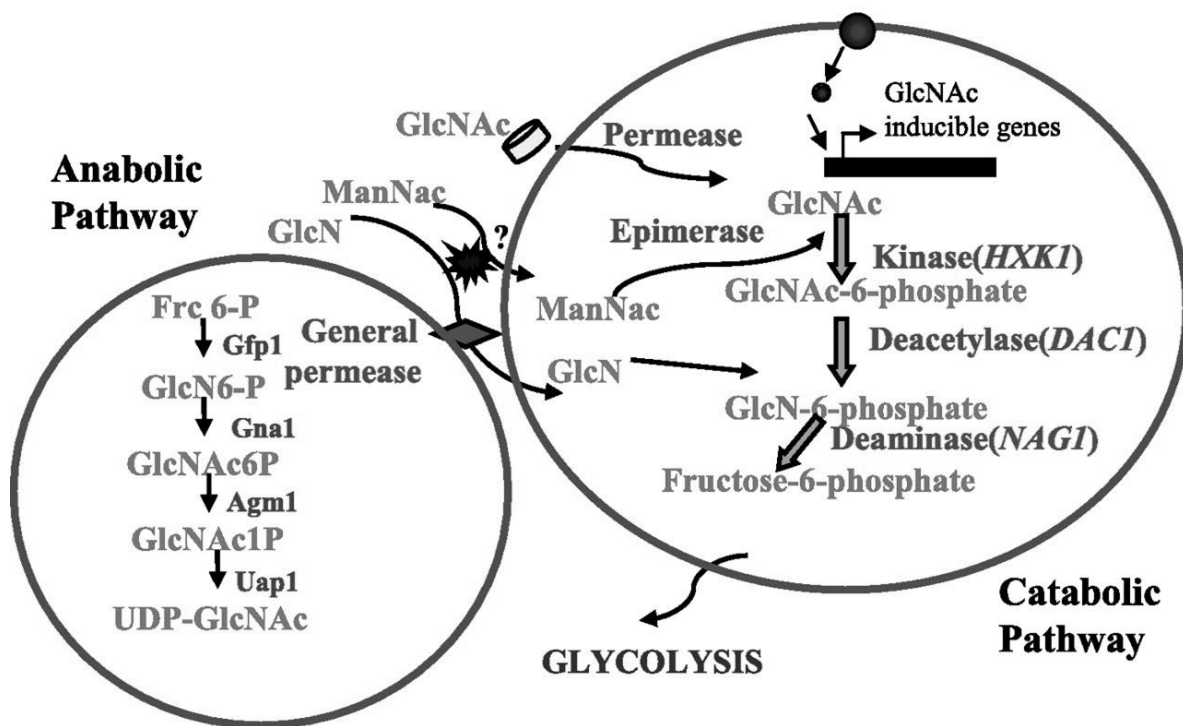
Figure 6: GlcNAc synthesis and degradation pathways in *C. albicans* (Biswas et al., 2007)

3. As a signaling molecule

GlcNAc in the extracellular medium is a potent inducer of yeast to hyphae transition in *C. albicans*, while other sugars are not (Simonetti et al., 1974). GlcNAc activates signaling pathways that stimulate yeast to hyphae switching, white to opaque switching via epigenetic changes, and gene expression of GlcNAc catabolic gene (Konopka, 2012; Naseem et al., 2012). Out of these, two sets of responses induced by GlcNAc are of special significance in understanding the mechanism of pathogenicity. One is the morphogenesis and second is expression of virulence genes.

Various Signaling pathways coordinating to regulate morphogenesis:

The transition from budding yeast to hyphal and pseudohyphal form is a key factor attributing to virulence of *C. albicans* (Sudbery 2011). Several signaling pathways contribute to activate this transition. The major pathways reported for morphogenesis in *C. albicans* are MAPK, Ras and cAMP-PKA signaling cascades. *T*he presence of cAMP is essential for this morphogenetic switch (the yeast-hyphal transition); however, cAMP levels are not the only metric for morphogenesis regulation. Recently a cAMP independent pathway for hyphal morphogenesis has also been identified in *C. albicans*. It is dependent on *Efg1* a transcription factor downstream to PKA; and not dependent on *Ras1*.

MAP Kinase Pathway

Three different MAPK (Mitogen-activated protein kinase) signalling pathways have been identified in *C. albicans*. One is *Hog1* MAPK cascade, mainly a stress response pathway mediating osmotic and oxidative stress. Second is *Mck1* MAPK cascade regulating cell wall integrity. Third is *Cek1* MAPK pathway, controlling cell wall formation during vegetative and filamentous growth. Both *Hog1* and *Mck1* MAPK pathways are activated under oxidative stress while *Cek1* MAPK pathway gets deactivated.

Cek1 mediated MAPK pathway controls mating and filamentation in *C. albicans*. The cascade follows through the kinases Cst20 (homolog of p21-activated kinase [PAK] kinase Ste20), CaSte7/Hst7 (homolog of MAP kinase kinase Ste7), and Cek1 (homolog of Fus3 and Kss1 MAP kinases). Downstream to them is a transcription factor, Acpr/Cph1, homolog of Ste12; regulates mating and pseudohyphal growth in *S. cerevisiae*. Disruption of any of the genes of the MAP

kinase cascade (Cst20, Hst7, or Cek1) or the transcription factor Cph1 results in filamentation defect on solid medium in response to many inducing conditions; whereas in case of liquid medium in response to serum , all these null mutants exhibit normal filamentation. Even though mutant of *cek1* MAP kinase was able form morphologically normal filaments in response to serum, it showed minor defect in growth on serum-containing medium. Additionally, the *cek1* mutant was also avirulent, possibly due to its growth defect; indicating involvement of Cek1 MAP kinase in more than one pathway or in crosstalk between other MAPK cascades. Along with the kinases (Cst20, Hst7 and Cek1) and transcription factor Cph1, the Cek1 MAPK pathway also consist of a phosphatase, Cpp1. Phenotypic analysis of homozygous mutant of CPP1 gene revealed that they were capable of derepressing hyphal production and resulted in hyperfilamentation. This hyperfilamentation can be suppressed by deletion of Cek1 MAPK. Cpp1 mutants also exhibited reduced virulence in both systemic and localized models of candidiasis. Another important component of the Cek1 MAPK pathway is a *GPA2*, a G-protein α subunit homologue. Null mutation of *GPA2* expressed invitro defects in filament growth in Spider and SLAD media as well as in embedded conditions although not in serum containing media. These defects cannot be rescued by addition of exogenous cyclic AMP. However over expression of *HST7* reversed the filamentation defect suggesting involvement of *GPA2* in Cek1 MAP kinase signaling pathway.

Another homolog of Fus3 of *S. cerevesiae* named Cek2; sharing 56% homology with Cek1 had also been identified in *C. albicans*. Similar to Cek1, Cek2 null mutants are also susceptible to cell wall interfering drugs. Role of Cek2 in cell wall biogenesis is still not clear and it may work in complement with cek1 but is not completely redundant to cek1.

cAMP-PKA Pathway

The well conserved cAMP-dependent protein kinase A (PKA) pathway is known to play important role in filamentation of *S. cerevisiae*, and other fungi including *C. albicans*. In *C. albicans*, the two important traits of hyphae formation and virulence are controlled by this signaling cascade. An increased level of cAMP accompanies the yeast to hyphal transition, and this transition is induced by inhibition of the cAMP phosphodiesterase. Moreover, cell-permeating PKA inhibitors, like myristoylated protein kinase inhibitor (myrPKI) amide and the

small-molecule PKA inhibitor H-89, can block this transition when induced by *N*-acetylglucosamine, but not when induced by serum.

The levels of cAMP are modulated by the phosphodiesterase Pde2, Pde2 is essential for morphogenesis and virulence (Bahn et al., 2003). Two essential components of the cAMP-dependent protein of kinase are a regulatory subunit, Bcy1, and a catalytic subunit Tpk1. Bcy1 required for localization Tpk1 to the nucleus and is apparently essential. In addition, Gpa2 the G protein homolog appears to function upstream of the CEK1 MAP kinase pathway, and is essential for hyphal development under several hyphal inducing conditions.

These signal transduction pathways work as a complex and interconnected network in sensing environmental cues to appropriately control morphogenesis and transcription. For instance Efg1, the APSES protein controls different programs such as hyphal development, white–opaque switching, chlamydospore formation and assembly of biofilms by receiving information through multiple pathways. In order to maintain the specificity of these responses, EfG1 activation must be phase specific. This might be achieved by regulating the levels of Efg1 either by post-translational modifications of Efg1 or by other transcription factors of the network functioning synergistically with or antagonistically to Efg1. Adenylyl cyclase and cAMP work as one of the regulatory circuits controlling Efg1 function.

Ras Pathway

C. albicans carries only one Ras homolog, Ras1, which is not essential for its survival. Studies on *ras1* mutant strains have shown severe defect in hyphal growth in response to serum and other hyphae inducing conditions. Moreover, while a dominant negative Ras1 mutation (Ras1A16) caused filamentation growth defect, a dominant active Ras1 mutation (Ras1V13) enhanced the hyphae formation. Since the in vitro morphological transition defects can be reversed by either supplementing the growth media with cAMP or by over expressing components of the filament inducing MAP Kinase cascade, it is evident that Ras1 functions upstream of both cyclic AMP as well as MAP Kinase pathway.

The Ras1 protein very likely activates two protein kinase A (PKA) isoforms Tpk1p and Tpk2p. Mutants of *tpk1* are defective in hyphal formation on solid media but not in liquid. On the other hand, in *tpk2* mutants, hyphal formation is partially affected on solid media but is blocked in
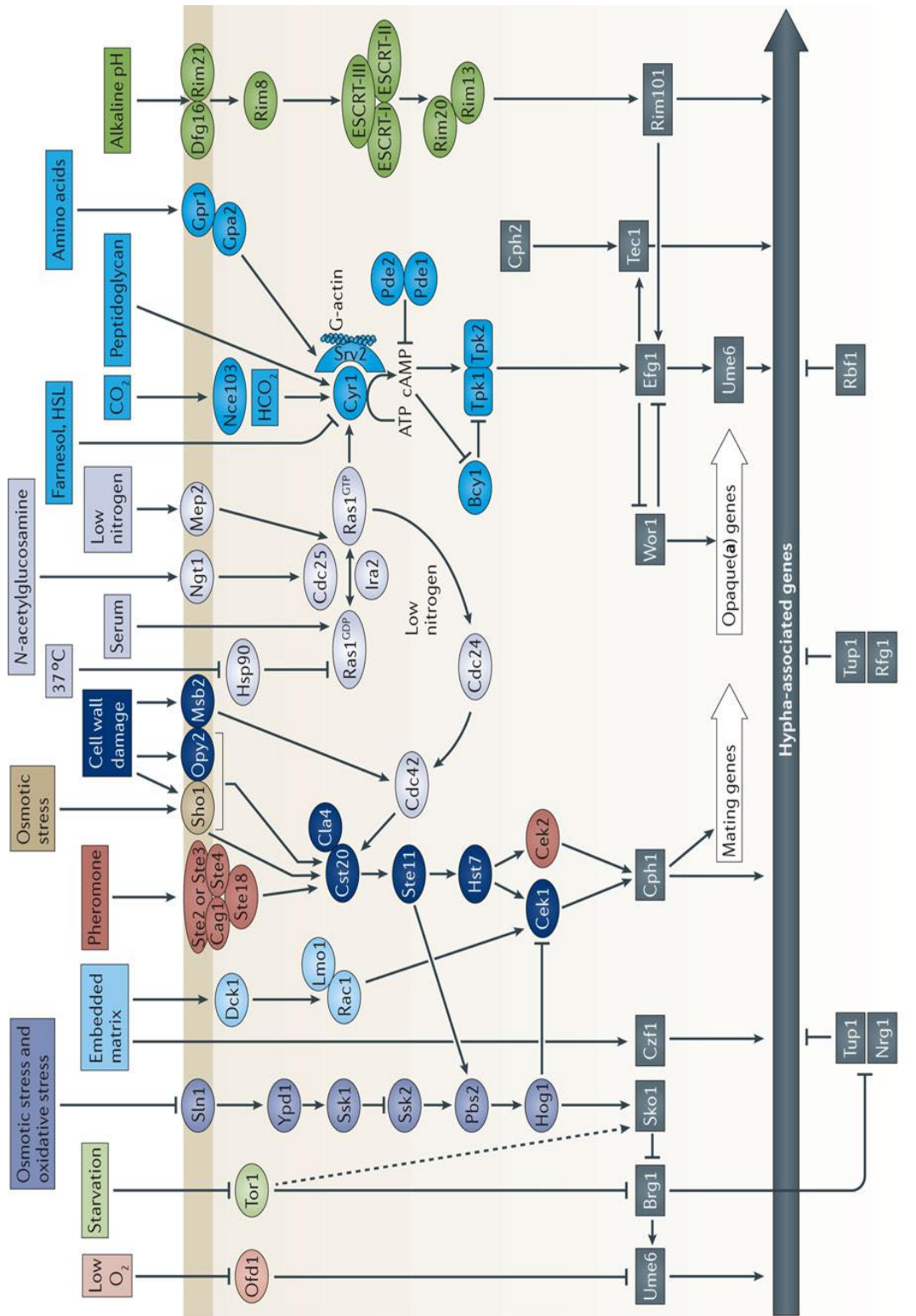
Figure 7: Signal transduction pathways that regulate morphogenesis (Noble et al., 2016).

liquid. *TPK2* null mutant containing a single allele of *TPK1* under a regulatable promoter is not viable at low expression levels.

GlcNAc regulates gene expression of N-acetylglucosamine catabolic pathway:

The capability to catabolize GlcNAc as a sole carbon is an important aspect of *Candida albicans* pathogenesis (Singh and Datta, 1979). The GlcNAc catabolic genes, encoding GlcNAc-6-phosphate deacetylase (*DAC1*), GlcN-6-phosphate deaminase (*NAG1*), and a hexokinase (*HXK1*), exist in a cluster called Nag regulon on chromosome 6 and are coordinately regulated at the transcriptional level(Kumar et al., 2000). The transcript levels for these genes when checked by northern blotting were induced within minutes of each other after GlcNAc induction. For instance, *NAG1* mRNA, which was absent in uninduced cells, was detected at 8 minutes of induction by GlcNAc, and reaches steady-state in the first 30 minutes of growth. Likewise, *DAC1* mRNA appeared at 16 minutes and *HXK1* mRNA at 4 minutes upon GlcNAc induction. They reach steady state at 30 minutes of growth in GlcNAc. These observations indicated that the coordinated expression of the catabolic pathway genes was achieved by activation of trans-acting factor(s) (Kumar *et al.*, 2000). The other 3 genes in the nag cluster were not induced by GlcNAc (Yamada-Okabe *et al*., 2001).

However, how levels of GlcNAc are sensed and then subsequently transduced is still being explored. One study has identified Ngs1, GCN5-related *N*-acetyltransferase as sensor that binds GlcNAc once it enters the cell via Ngt1. Ngs1 has a N- terminal *β-N*-acetylglucosaminidase homology domain via which it binds to GlcNAc and C-terminal GCN5-related *N*-acetyltransferase domain. Binding of GlcNAc to the N terminal pocket activates the C terminal GNAT domain, which is required for GlcNAc-induced promoter histone acetylation and transcription. Constitutive targeting of Ngs1 to the promoters of GlcNAc-inducible genes is through the transcription factor Rep1(Su et al., 2016).

Another study has found that Ngs1 is important for growth not only on GlcNAc but also on maltose. A new transcription factor Ron1 has been identified which acts specifically in response to GlcNAc. Ron1 contains an Ndt80 like DNA binding domain. Mutation of *NGS1* or *RON1* in *Candida albicans* renders inability to induce GlcNAc catabolic genes in response to GlcNAc induction. In addition to that, *ngs1Δ* mutant were highly defective in filamention in response to

GlcNAc, while *ron1Δ* mutants were partially defective. These results suggest Ngs1 and Ron1 to be critical regulators of GlcNAc catabolism as well as filamentation (Naseem et al., 2017).

GlcNAc catabolism regulates extracellular pH

A study has shown that GlcNAc catabolism is not necessary to induce hyphae formation signaling in *Candida albicans* (Naseem et al., 2011); the same group has later reported that, GlcNAc catabolism raises the extracellular pH which then indirectly stimulates filamentation (Naseem et al., 2015). While cell growth on glucose medium leads to acidification of the surrounding medium; growth on GlcNAc on the contrary raises the surrounding medium's pH since the cells exported excess nitrogen generated from GlcNAc catabolism as ammonia(Vylkova et al., 2011). This phenonmenon is supported by the studies on mutant strains lacking GlcNAc metabolic genes (*hxk1Δ nag1Δ dac1Δ triple mutant*). It was seen that the mutants were capable of inducing hyphae formation without inducing hyphal specific genes at low pH. Although, the mutants induced hyphal-specific genes when buffered to a higher pH (>5), since this mimicked the effects of GlcNAc catabolism(Naseem and Konopka, 2015). Thus it can be concluded that GlcNAc impacts Candida albicans morphologenesis hence virulence via multiple ways; either by directly regulating cell signaling as well as indirectly by changing ambient pH.

GlcNAc can also induces cell death in *Candida albicans*

In the expanding list of the cellular processes regulated in response to GlcNAc, one interesting finding is that it can act as a signal to induce cell death in *Candida albicans*. Nutrient sensing from the surround microenvironment is one of the important adaptation strategies of pathogens. One study reports that *Candida albicans* cells when grown on water only survived for weeks by entering G_0 phase; on the other hand when grown on water with GlcNAc, it lead to loss quick loss of viability. This phenomenon was termed GICD (GlcNAc induced cell death). GlcNAc triggered upregulation of ribosomal biogenesis genes, altered mitochondrial metabolism and ROS (reactive oxygen species) accumulation, leading to rapid cell death by both apoptosis and necrosis. GCID involves multiple pathways, including cAMP dependent and cAMP independent pathways. This result adds to the understanding that, *Candida albicans* uses GlcNAc as a signal of nutrient availability which then regulates multiple cellular processes in a coordinated manner

to maximize efficient nutrient use. This adaptability of *Candida albicans* enhances its survival and pathogenicity inside the host(Du et al., 2015).

CHAPTER 1

Chapter 1

Cloning, expression, purification and crystal structure analysis of Gig2 (GlcNAc inducible gene 2) protein of *Candida albicans*

Introduction

N-acetyl glucosamine (GlcNAc), an amino-sugar present abundantly at mucosal linings is utilized by many pathogens as an alternate carbon source and a signaling molecule. GlcNAc metabolism and signaling plays an important role in regulating virulence of *Candida albicans.* In *Candida albicans,* GlcNAc is picked up from the extracellular medium by a transporter Ngt1. Inside the cell, the molecule activates various signaling events including expression of its own catabolic enzymes. Catabolism of GlcNAc starts with its phosphorylation by HxK1, followed by deacetylation and deamination via Dac1 and Nag1, forming Glucose-6-phosphate which is fed into the glycolysis cycle. While these 3 genes directly involved in GlcNAc catabolism are well established, secondary pathways bifurcating from the metabolic intermediates formed are yet to be studied. GIG2 or N-Acetylglucosamine (GlcNAc) Inducible Gene 2 was identified as one of the novel genes upregulated in response to GlcNAc in *Candida albicans* based on microarray analysis. Gig2 was found to be upregulated in Dac1 and Nag1 mutants, but not in Hxk1 mutants in response to GlcNAc; indicating its involvement in some pathway bifurcating from GlcNAc-6-phosphate when not being utilized in glycosis. GlcNAc-6-phosphate is a central molecule for several anabolic and catabolic pathways.

Preliminary studies on Gig2 mutants have found them to be sensitive to H_2O_2 induced oxidative stress. Mutants were also attenuated in virulence in murine model of systemic candidiasis. Based on the results of protein-metabolite complex purification and metabolite identification by using UPLC coupled ESI-MS and solution state NMR it was proposed that gig2 may be involved in N-acetyl neuraminic acid (sialic acid) metabolism in *Candida albicans* but the exact catalytic reaction carried out by gig2 is not known.

GIG2 gene of *Candida albicans* encoded by orf19.4783 translates into a 474 amino acids long protein. Gig2 is a member of DUF1479 family of proteins. DUFs or Domains of unknown function represents uncharacterized protein families present in pfam database. Based on BLAST

results Gig2 is conserved amongst members of Enterobacteriaceae, pathogenic fungi, and plants (Fig 1.1) but not present in humans. Closest PDB homolog is from *Yersenia pestis* [PDB ID: 4RGK] sharing 29% identity but uncharacterized function. No member of the DUF1479 family has been characterized yet.

Domain of Unknown Function (DUF)

The advancement in techniques like High throughput genome sequencing has made possible to sequence the complete genome of *Candida albicans* available at CGD (Candida genome database). However, a large proportion of the genome encodes putative proteins whose structure and function are unknown. For functional annotation, tools like multiple sequence alignment and Hidden Markov models (HMM) helps in clustering of proteins into domains and families based on conserved residues. The pfam database has a large collection of protein families that do not contain any protein of known function and are placed under the category called DUF (Domain of Unknown Function) families. Chris Ponting, was the first to introduce the DUF prefixed naming scheme.

One approach to understand the function of DUF proteins is by studying their structure. Several structural genomics programmes have attempted this approach to understand the function of DUFs via structure determination. Structures of more than 250 DUF families have already been solved. These structural analysis have revealed that about two third of these DUFs had structural similarity with already known structures therefore expected to be evolutionarily diverged from existing protein families, on the other hand about one third had novel protein folds (Jaroszewski et al., 2009).

DUF 1479

Gig2 belongs to DUF 1479 family, containing a 420 residues long DUF 1479 domain. Other members of the family include several hypothetical enterobacterial proteins, none with a clue of function. Protein sequence similarity search identifies PDB hits with less than 30% homology. Therefore, this study attempts to crystallize Gig2 protein in order to solve its structure by employing X-ray crystallography.
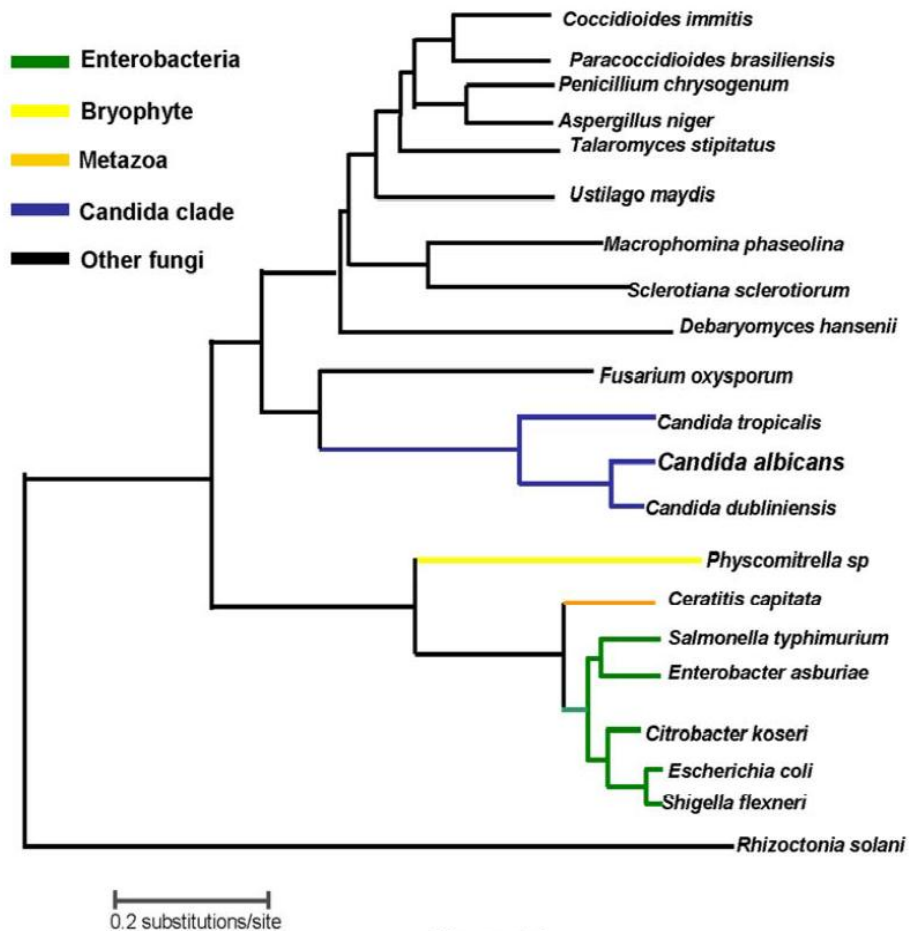
Figure1.1: Evolutionary conservation of Gig2 (Ghosh et al., 2014)

Materials and Methods

1.1. Cloning of *GIG2* gene

Gene *GIG2* of *Candida albicans* was PCR amplified using gene specific oligonucleotides with respective restriction sites for cloning. The amplified *GIG2* DNA fragment was purified by agarose gel extraction. The purified *GIG*2 DNA fragment was digested with *BamH*I and *Sac*I restriction enzymes. The vector, pET28a was also digested with *BamH*I and *Sac*I restriction enzymes followed by ligation with digested *GIG2* DNA fragment for constructing pET28a-gig2. Competent *E. coli* DH5α cells were transformed with the ligation mixture.

1.2. Screening of putative clones

The transformants obtained were screened for presence of insert by mini-preparations of colony's plasmid DNA and their restriction digestion analysis. Positive clones were selected on the basis of a fall-out of gig2 at 1449bp after digestion with *BamH*I and *Sac*I respectively (Fig1.2B).

1.3 Sequencing of the plasmid of positive clone:

The sequence of insert was confirmed by automated sequencing at DNA Sequencing Facility at NIPGR Delhi. The observed sequence was checked by using BLAST at ncbi.nlm.nih.gov. The sequence of insert was found to be same as in genomic DNA and no mutations were observed in the sequence of *GIG*2 gene.

1.4. Recombinant gig2 protein was expressed in *E. coli* BL21

1.4.1. *E. coli* BL21 (codon plus) cells were transformed with pET28a-gig2. Cells were grown till 0.6 OD at 600nm and induced by 0.5mM and 1mM IPTG at 16°C, 22°C, 24°C, 28°C and 30°C for 3 to 16 hours. Samples were collected before and after induction. Cell lysates were prepared from these samples by adding lysis buffer (1X) and boiling them for 20 min. Cell lysates were then subjected to 10 % SDS-PAGE. Gels were stained, destained and analyzed. Recombinant gig2 protein expression was observed at 55kDa in *E. coli* BL21 cells after 5 hours of induction. Optimum expression was obtained with 0.5 mM IPTG at 22°C for 8 hours (Figure 1.3)

1.5. Purification of recombinant gig2

1.5.1. Purification of gig2 protein under native conditions from cytosol:

The recombinant gig2 was purified from the cell pellet under native condition; employing Ni-NTA affinity chromatography. Cell pellet stored at -80°C after thawing was resuspended in lysis buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 0.5% NP40, 10mM imidazole, 0.5mg/ml lysozyme) and subjected to repeated cycles of freeze-thaw for cell lysis and then was sonicated at 25% amplitude 30s pulse 30s rest (6-7cycles). The lysate was the centrifuged at 20,000rpm for 30mins at 4°C. Filtered lysate was incubated with Ni-NTA resin for 45min at 4°C. The protein-bound beads were washed three times with wash buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 30 mM imidazole). Protein was eluted in elution buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 300 mM imidazole) at 4°C in batch method. (Figure 3)

1.5.2. Sequencing of the 55KDa band by MALDI/TOF:

The 55KDa purified protein band was excised from the SDS PAGE gel and processed by in gel trypsinization method for extraction of peptides which was analyzed by Mass spectrometry facility at AIRF, JNU Delhi. The result identified the protein as gig2 of *Candida albicans*.

1.5.3. Further purification of Ni-NTA purified protein by GPC:

After the recombinant gig2 protein was purified by Ni-NTA resin in bulk, it was concentrated using Amicon ultracentrifugal filters upto 20mg/ml and was loaded on a gel filteration chromatography Superdex 200 10/300 GL column (GE Healthcare), which was previously equilibrated with buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 5% glycerol). Fractions of the peak from 80- 90 ml were collected, checked on 10% SDS-PAGE gel and concentrated upto 8mg/ml. The purified concentrated protein is then used for crystallization.

1.6. Crystallization trials

Crystallization trials for purified Gig2 protein were performed by Mosquito robotics through Hanging Drop Vapor Diffusion method. The ratio of the precipitant to the protein was chosen 1:1 and the drop size for each was kept at 350nL. Commercial crystallization screens from Hampton Research and Molecular Dimensions were used for these trials. Proteins concentrated up to 8-10mg were used for putting the drops. The 96 well plate crystallization plates were kept at 16°C and At 4°C

inside the incubator and observed at regular interval under microscopes. Once after getting optimum hits for Gig2 it was repeated in 24 well plates manually for further optimization and crystal growth.

1.7. Data collection

Gig2 crystals from several drops were mounted on 0.2mm Nylon loops and were soaked in cryoprotectant which was mother liquor itself. These soaked crystals were quickly flash frozen in liquid nitrogen. Data was collected on BM14 beamline at the ESRF (Grenoble, France) at a wavelength of 0.97Å. The distance between the detector and the crystal was 267.9mm. A total of 380 images were collected with 0.5 Å oscillation using CCD detector (Marresearch) and the data was indexed and scaled using HKL 2000(Otwinowski and Minor, 1997). SCALEPACK2MTZ from the CCP4 suite(Winn et al., 2011) was run to convert the scaled reflections to MTZ files and these reflections were further used for molecular-replacement trials.

Results

1.3.1 Cloning of the coding region of *GIG2*

Sequence of the *GIG2* gene encoded by orf19.4783 in *Candida albicans* (Strain SC5314) was retrieved from *Candida* genome database at http://www.candidagenome.org/. Forward (5'ACTGAT<u>GGATCC</u>ATGTCTCCTTCCAAATTAT3') and reverse (5'TATCAT<u>GAGCTC</u>CTTAGCAGCATGGTGA3') primers were designed manually using Oligo software for the full length *GIG2* gene flanked with restriction sites for *BamH*I and *Sac*I restriction enzymes respectively. The PCR amplified *GIG2* gene of 1449 bps (Fig1.2A) was ligated into the pET28a (Novagen) expression. The clone was confirmed by plasmid isolation and restriction digestion with *BamH*I and *Sac*I (Fig1.2B). The cloned gene was sequenced at DNA Sequencing Facility at NIPGR Delhi which showed no mutation. Highly purified cloned construct of pET28a-gig2 was purified and stored at -20 $^{\circ}$C for further use.

1.3.2 Expression and Purification of Gig2

Transformation of recombinant plasmid in BL21 codon plus *E.coli* cells yielded a good number of colonies. Optimum expression was obtained with 0.5 mM IPTG at 22 $^{\circ}$C for 8 hours (Figure 1.2). The protein purification strategy involved two steps of chromatography techniques: affinity and gel filtration. From Ni-NTA affinity chromatographic column, the 6X His-tagged protein was eluted at 200mM imidazole concentration indicating high affinity. The SDS-PAGE analysis showed the Ni-NTA column purified protein to be nearly 70% pure (Fig 1.3) and also the molecular mass weight to be in agreement with the expected value of ~ 55 kDa. To further confirm the identity of this 55kDa purified protein band, it was excised from the SDS PAGE gel and processed for extraction of peptides which was analyzed by Mass spectrometry facility at AIRF, JNU Delhi. The result confirmed the protein as Gig2 of *Candida albicans*.

Fractions collected after NiNTA purification were pooled and loaded on to the pre-equilibrated Superdex 200 10/300 GL column (GE Healthcare) as mentioned in Materials and Methods section. The Gig2 protein elutes as a single peak corresponding to a molecular weight of ~ 55 kDa (Fig. 1.4); as per the standard (Fig. 1.4 inset) which indicates Gig2 protein exists as a monomer in GPC buffer. Further, SDS-PAGE analysis showed GPC purified protein to be more than 95% pure.
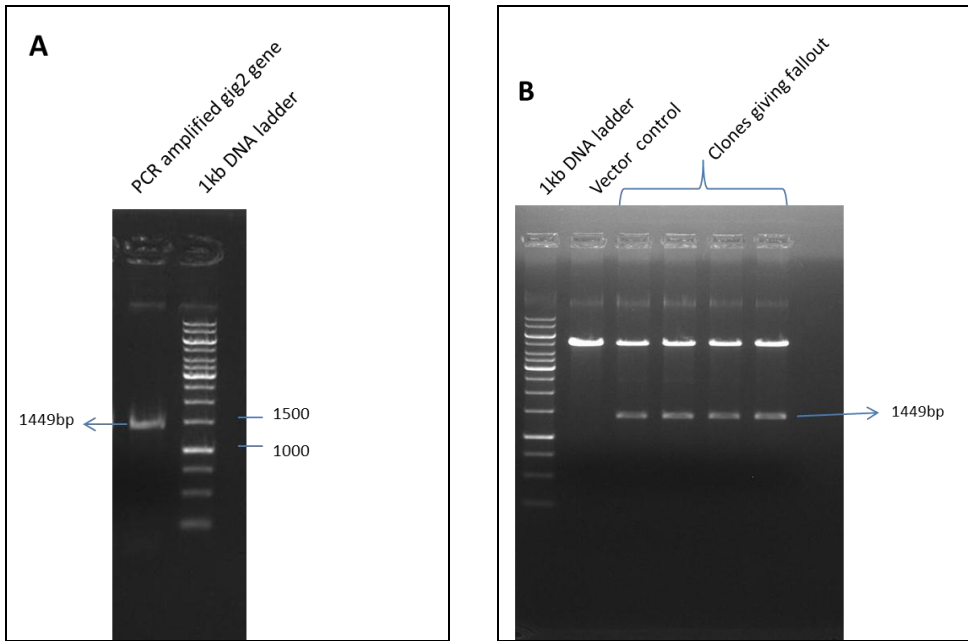
Figure 1.2: [A] shows the PCR amplification of *gig*2 from *C. albicans* (strain SC5314) genomic DNA as template. [B] shows the restriction digestion of the putative clone with 1449bp fall out of gig*2*.
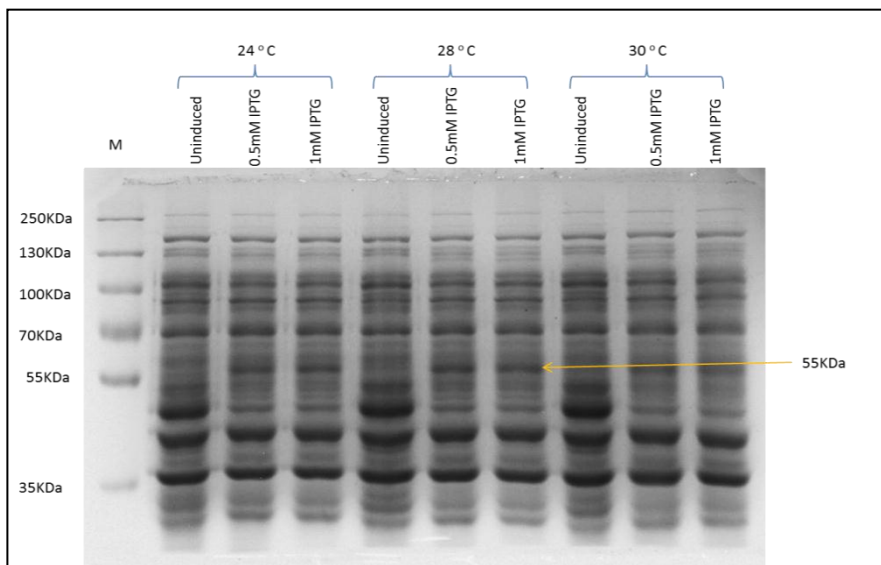

Figure 1.3: It shows the recombinant Gig2 protein expression in BL21 (codon plus) cells transformed with pET28a-gig2 plasmid. The arrow shows the induction band of Gig2 at 55kDa.
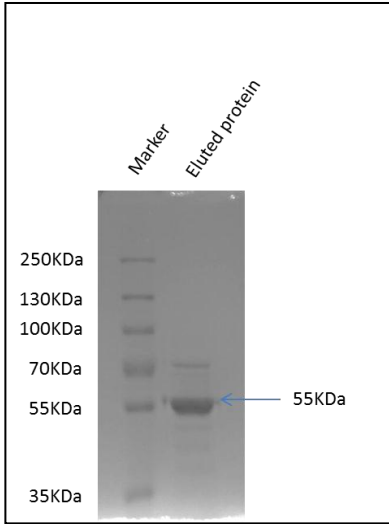
Figure 1.4: The purified Gig2 protein eluted from Ni-NTA resin is shown following SDS-PAGE and coomassie staining.
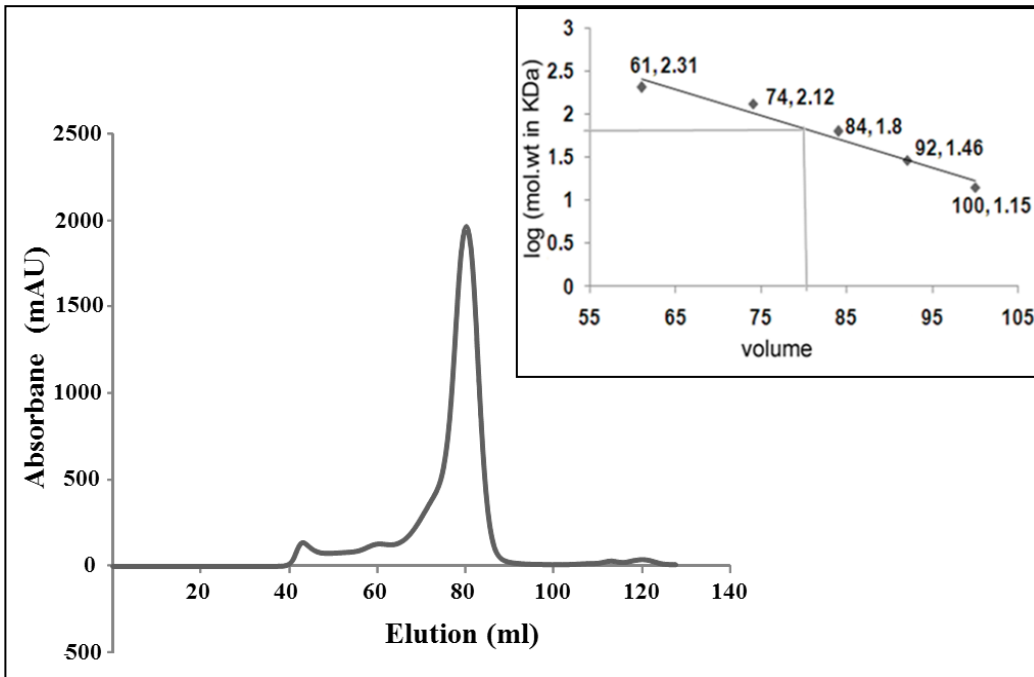

Figure 1.5: Size-exclusion chromatography elution profile of Gig2 purification. Fractions (80–90 ml) containing purified Gig2 were pooled, concentrated and used for crystallization. Inset: Standard curve for Superdex 200 10/300 GL column showing elution volume of different Molecular weight protein markers.

1.3.3. Crystallization of Gig2

Purified protein was concentrated to 8-10 mg/ml and was used for crystallization experiments. Different commercially available crystallization screens (Hampton Research and Molecular Dimensions) were set up using Mosquito crystallization robot (TTP Labtech) using greiner 96 well crystallization plates (350 nl drop with protein solution and reservoir solution in 1:1 ratio) and the plates were kept at 16^0C.

Initial attempts for crystallization yielded small fine needle shaped crystals that appeared after 3-4 days in **Morpheus screen** (Molecular Dimensions) in conditions; A4 [12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD ,0.03 M of each divalent cation $MgCl_2$ and $CaCl_2$, 0.1 M MES/imidazole pH 6.5] (Fig. 2A), A8 [12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD, 0.03 M of each divalent cation, 0.1 M MOPS/HEPES-Na pH 7.5] and A12 [12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD, 0.03 M of each divalent cation,0.1 M bicine/Trizma base pH 8.5].Further crystallisation optimisation experiments were done manually using hanging drop method in 24 well plates (Corning). A4 and A8 crystals were very small rods while A12 crystals were small fine needles. These crystals were used as seeds for macro-seeding in both A4 and A8 conditions by varying pH from 6.5 to 7.5 by a unit of 0.2 and by varying the percentage-ratio of PEG 3350 /PEG 1000 from 12.5% / 12.5% to 14% /11% ; 15% / 10% and 16% / 9%. Best diffractable crystals were obtained in the condition; 11% w/v PEG 1000, 14% w/v PEG 3350, 12.5% v/v MPD , 0.01 M of each $MgCl_2$ and $CaCl_2$, 0.1 M MES/imidazole pH 6.9. The crystals obtained were slightly bigger than the initial crystals obtained without seeding.

Later crystals were also observed after 10-15days in **Pact Priemere screen** (Molecular Dimensions) in condition, 0.2 M Magnesium chloride hexahydrate, 0.1 M Tris 8.0, and 20 % w/v PEG 6000. These crystals were also like fine needle but slightly larger than the previous conditions found in Morpheus screen. Crystallisation condition was the refined manually using hanging drop method in 24 well plates (Corning). The crystals repeated in 24 well plates also grew up as small fine needles. These crystal were used as seeds for macro-seeding. Better long diffractable rod shaped crystal grew in condition 0.1 M Magnesium chloride hexahydrate, 0.1 M Tris 8.0, and with PEG 6000 concentration (w/v) ranging from 25% to 30%.
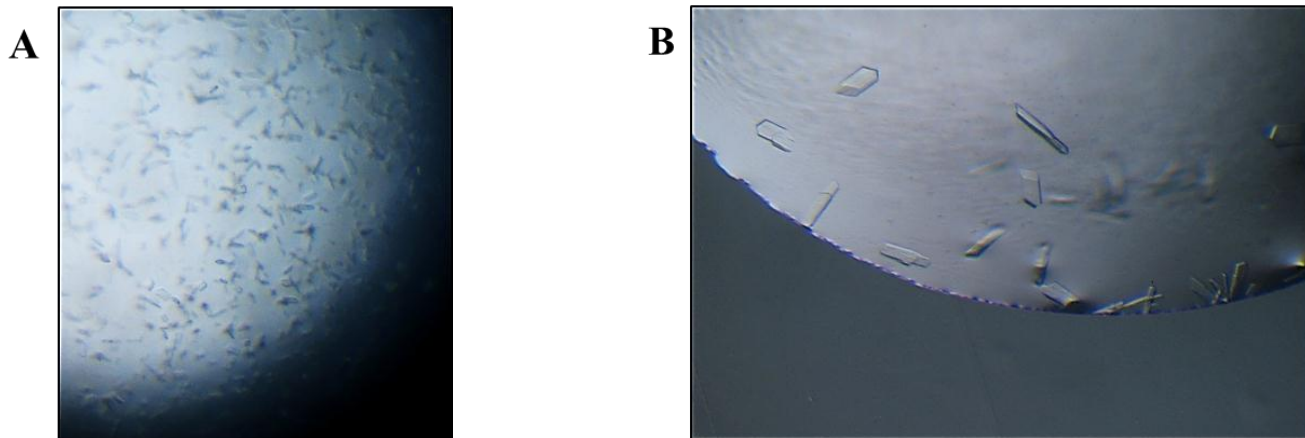
Figure 1.6: Crystal images of Gig2 crystals: (A) Initial hits of Gig2 obtained in crystallization condition A8 (12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD, 0.03 M of each Mgcl$_2$ and CaCl$_2$, 0.1 M MOPS/HEPES-Na pH 6.5) of Morpheus Screen from Molecular Dimensions. (B) Examples of Gig2 crystal obtained after macro-seeding with initial small crystals in crystallization condition; 11% w/v PEG 1000, 14% w/v PEG 3350, 12.5% v/v MPD , 0.01 M of each divalent cation MgCl$_2$ and CaCl$_2$, 0.1 M MES/imidazole pH 6.9. One of these crystals was used for data collection at BM14 beamline, ESRF, Grenoble, France.
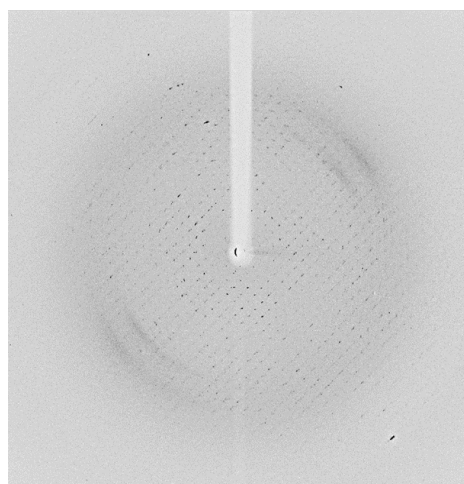

Figure 1.7 Gig2 diffraction image obtained using a Marresearch CCD detector on BM14 beamline at ESRF.

1.3.4 Structure of Gig2

Data collection and structure determination

The initial screening yielded crystals under many conditions in Morpheus crystallization screen (Molecular Dimensions) but the diffractable crystals were obtained in 11% w/v PEG 1000, 14% w/v PEG 3350, 12.5% v/v MPD, 0.01 M of each divalent cation, 0.1 M MES/imidazole pH 6.9. The crystals diffracted at 2.4 Å resolution at ESRF (Table 1) in $P2_1$ space group with unit cell parameters a=59.59, b= 54.43, c=73.29 and β= 102.70Å. The calculated Mathews coefficient was 1.92 Å^3/Da with a solvent content of 35.92%. This corresponded to the presence of a single monomer of Gig2 protein in an asymmetric unit. No structure with more than 30% identity was available in PDB. During structure solution the data was cut off at 2.7 Å resolution which has an ⟨$I/\sigma(I)$⟩ equal to 2. The phase problem was solved by molecular replacement using PHASER from the CCP4 suite, using putative Oxidoreductase from *Salmonella typhimurium LT2* (PDB ID 2CSG), as a search model which had a sequences identity of only 27%. The template PDB was trimmed by chainsaw and used for molecular replacement. PHASER yielded a top solution with LLG score of 87, RFZ= 9.2 and TFZ= 6.4. First refinement cycle using REFMAC5 from CCP4 suite resulted in an R/R$_{free}$ of 0.47/0.50. The structure was refined to an R/R$_{free}$ of 0.37/0.47 by undergoing 7 cycles of refinement, each followed by manual building in COOT. Complete structure could not be solved as density could not be built for residues in the loop region. The structure needed further refinement.

In an attempt to get better diffractable crystals another crystallization screen Pact Priemere (Molecular Dimensions) where crystals were observed after 10-15days was manually optimized. By optimization and macroseeding better long diffractable rod shaped crystals grew in condition 0.1 M Magnesium chloride hexahydrate, 0.1 M Tris 8.0, and with PEG 6000 concentration (w/v) ranging from 25% to 30%. These crystals diffracted in $P2_1$ space group with unit cell parameters a=60.65, b= 54.54, c=74.85 and β= 105.61° (Table1). The PDB file from low resolution data of 2.7 Å, was used as a template for structure solution .The template PDB was used for molecular replacement by Phaser MR in CCP4 suite. PHASER yielded a single solution with LLG score of 1837, RFZ= 34.9 and TFZ= 26.0. First refinement cycle using REFMAC5 (Murshudov et al., 2011) from CCP4 suite resulted in an R/R$_{free}$ of 0.39/0.42. . The resulting model was then submitted to ARP/wARP classic for automated model building (Langer et al., 2008), which successfully built about 90% of the missing loops and side chains of the model with good-quality

electron density. The remaining parts of the polypeptide were built manually using COOT (Emsley et al., 2010). The structure was refined to an R/R_{free} of 0.18/0.22 by undergoing 3 cycles of refinement, each followed by manual building in COOT. The structure is currently under further refinement. Latest refinement statistics are given in Table 2.

Overall structure of Gig2 from *Candida albicans*:

Crystal structure of GlcNAc inducible gene 2, Gig2 from *Candida albicans* consists of a single domain with 26 alpha helices and 13 beta sheets (Fig 1.10). Only one monomer is present in an asymmetric unit in $P2_1$ space group. The core of the structure has eight antiparallel beta sheets arranged in a β-jelly roll barrel. A close observation depicts that the jelly-roll forms a cupin domain corresponding to the active site and containing a Fe^{2+} in the crystal structure. Cupin superfamily has two characteristic motifs:

G-$(X)_5$-HXH-$(X)_{3,4}$-E-$(X)_6$-G as Motif 1, and G-$(X)_5$-PXG-$(X)_2$-H-$(X)_3$ –N as Motif 2. Each motif contains active site residues and are organised into two β-strands. Gig2 also harbours the characteristic motifs of cupin fold. Motif 1 was located between residues 184-216 as G-$(X)_7$-HXD-$(X)_4$-E-$(X)_{13}$-G and harbours the catalytic H^{192} and D^{194}. Motif 2 in Gig2 is G-$(X)_4$-P-H-$(X)_6$ –N located between residues 344-362, is slightly different from the characteristic motif 2 in sequence comparision but structurally it still has two β-sheet present in motif2, part of cupin fold.

The two motifs are separated by 126 amino acids i.e. greater than 100 amino acids as has been reported previously in case of cupin family eukaryotic transcription factors and dioxygenases.

Metal coordination and Active site analysis:

The experimental electron density map in the crystal structure of Gig2 shows presence of a metal ion; which was confirmed by X-ray fluorescence scanning (XFS) measurements. XFS gave emission equivalent to characteristic Kα for Fe, clearly revealing the presence of the Fe ion (Fig 1.8). The iron is present in ferrous state coordinated by the side chains of His192, Glu194, and His355 located in the β-barrel. The ferrous state is confirmed by the octahedral arrangement of three water molecules and active site residues His192, Glu194, and His355 surrounding Fe atom. The active site is partially masked by two loops $L_{231-256}$, L_a1 and $L_{302-328}$, L_a2. They may be involved in controlling the entry and exit of ligand until the reaction takes place. In the crystal

structure Loop L_a1 is incomplete with the density of 10 residues missing from the structure (Fig 1.13).

Active site similarity:

A Dali search for structural comparison resulted in similar proteins with rmsd ranging from 2.2 – 10.1 which states that Gig2 has less structural similarity with the available structures in PDB. Gig2 active site shared maximal similarity with 2-OG dependant iron containing mononuclear dioxygenases which are quite diverse in types of ligands involved and biochemical reactions undertaken by them. In 2-oxo glutarate dependant mononuclear dioxygenases, two of the water molecules in the octahedral coordination are occupied by the cofactor, α-keto glutarate and the third is occupied by molecular oxygen. The active site of gig2 can also accommodate α-keto glutarate as seen in the docking studies but further validation needs to be done.
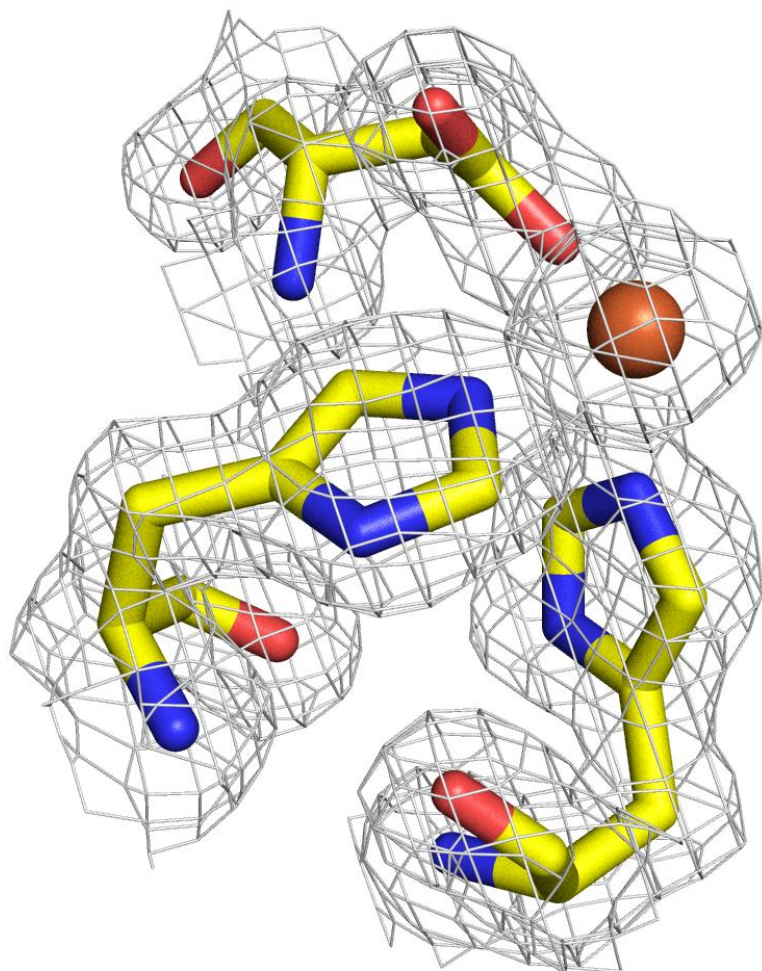
Fig 1.8: A portion of Gig2 molecule with metal ion shown with 2Fo-Fc electron density map carved at 1.6σ cutoff.

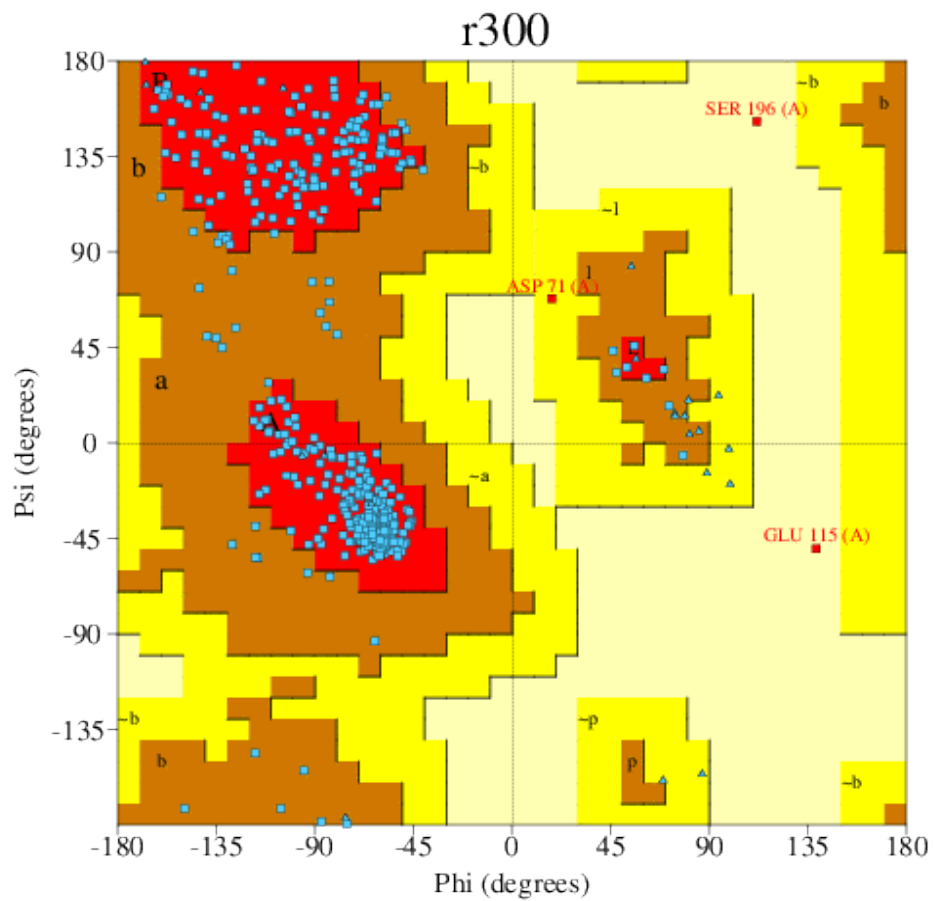Fig 1.9: Ramachandran plot of the main chain torsion angles (φ,ψ) for the final refined model of Gig2. The plot was calculated with the program PROCHECK (Laskowski et al., 1993). Glycine residues are identified as triangles, non-glycine residues by squares.

Table 1: Data collection and refinement statistics

Values for the outer shell are given in parentheses.

	Morpheus condition	Pact premier condition
Diffraction source	Beamline BM14, ESRF	Beamline BM14, ESRF
Wavelength (Å)	0.97	0.97
Temperature (K)	100	100
Detector	CCD, Marresearch	CCD, Marresearch
Crystal-detector distance (mm)	267.91	168.6
Rotation range per image (°)	0.5	0.5
Total rotation range (°)	190	190
Exposure time per image (s)	7.8	6.1
Space group	$P2_1$	$P2_1$
a, b, c (Å)	59.59, 54.43, 73.29	60.65, 54.54, 74.85
α, β, γ (°)	90, 102.70, 90	90, 105.61, 90
Mosaicity (°)	0.3	0.6
Resolution range (Å)	50.00–2.46 (2.50–2.46)	36.04 - 1.695 (1.756 - 1.695)
Total No. of reflections	68853	203858
No. of unique reflections	16747	52149 (4894)
Completeness (%)	99.5 (92.7)	0.99
Redundancy	3.8 (3.0)	0.45
$\langle I/\sigma(I)\rangle$	7.4 (1.3)	17.4 (1.4)
R_{merge}	0.2 (0.8)	0.1 (1.0)
R_{sym}	0.2 (0.8)	0.1 (0.1)
$R_{\text{r.i.m.}}$	0.262 (0.950)	0.06 (0.6)
Overall B factor from Wilson plot (Å^2)	25.92	18.55

mean $I/\sigma(I)$ is < 2 at 2.46Å.

Table 2: Structure solution and refinement statistics

	Morpheus condition	Pact premier condition
Resolution range (Å)	71.502–2.600 (2.667–2.600)	36.04 - 1.695 (1.756 - 1.695)
Completeness (%)	99.9	99.0
No. of reflections, working set	13557 (952)	52093 (4889)
No. of reflections, test set	732 (64)	2580 (257)
Final R_{cryst}	0.376 (0.351)	0.19 (0.27)
Final R_{free}	0.485 (0.446)	0.22 (0.29)
No. of non-H atoms	2970	4005
Protein	395	3752
Ion	0	1
R.m.s. deviations :		
Bonds (Å)	0.023	0.019
Angles (°)	2.06	2.01
Average B factors (Å^2)	20.6	24.29
Ramachandran plot:		
Most favoured (%)	67	97
Allowed (%)	20	1.5
Outliers (%)	10	1.1
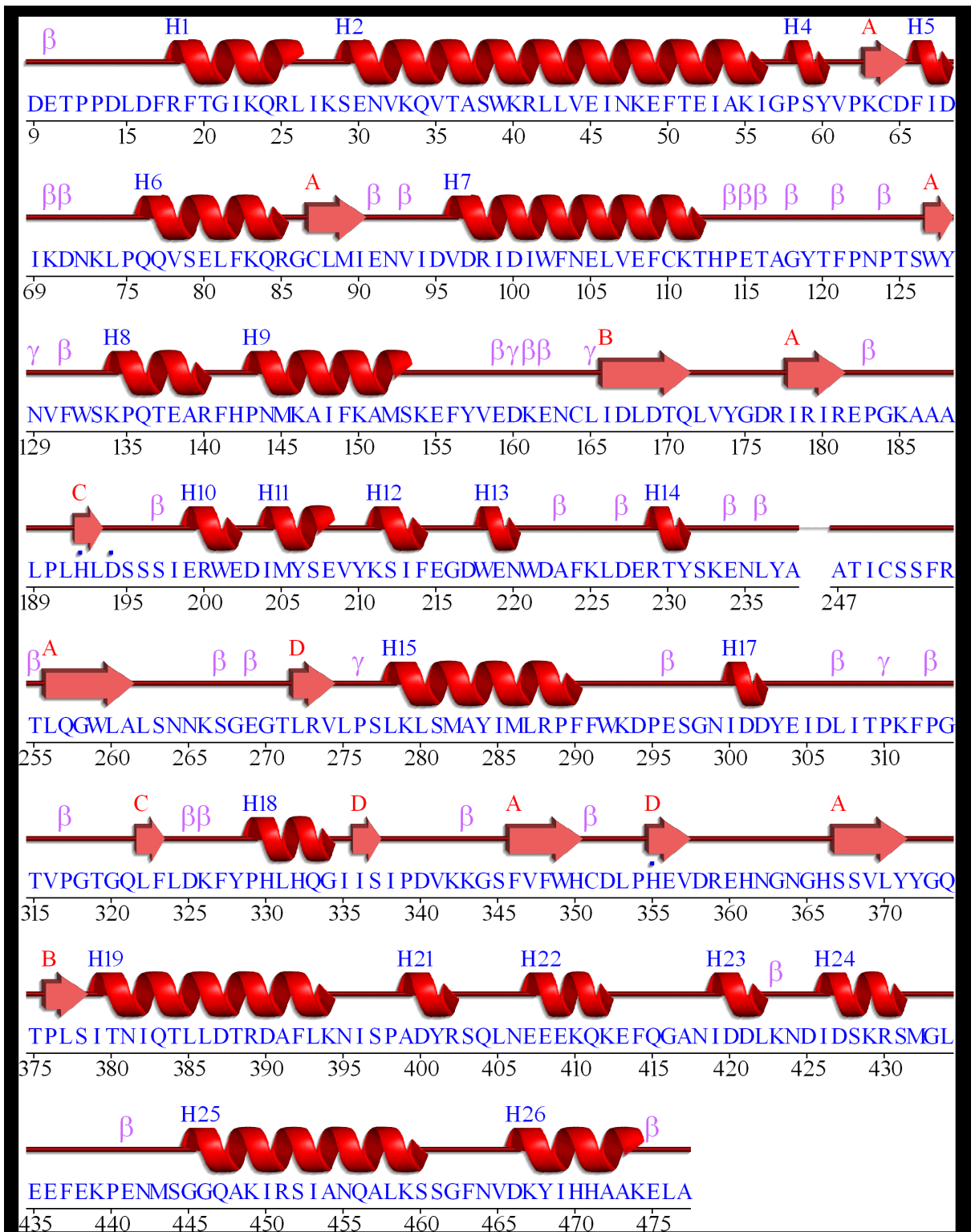
Fig1.10: Linear view of the complete Gig2 protein structure with alpha helices, beta and gamma turn and corresponding amino acid residues (Figure generated using PDBsum).
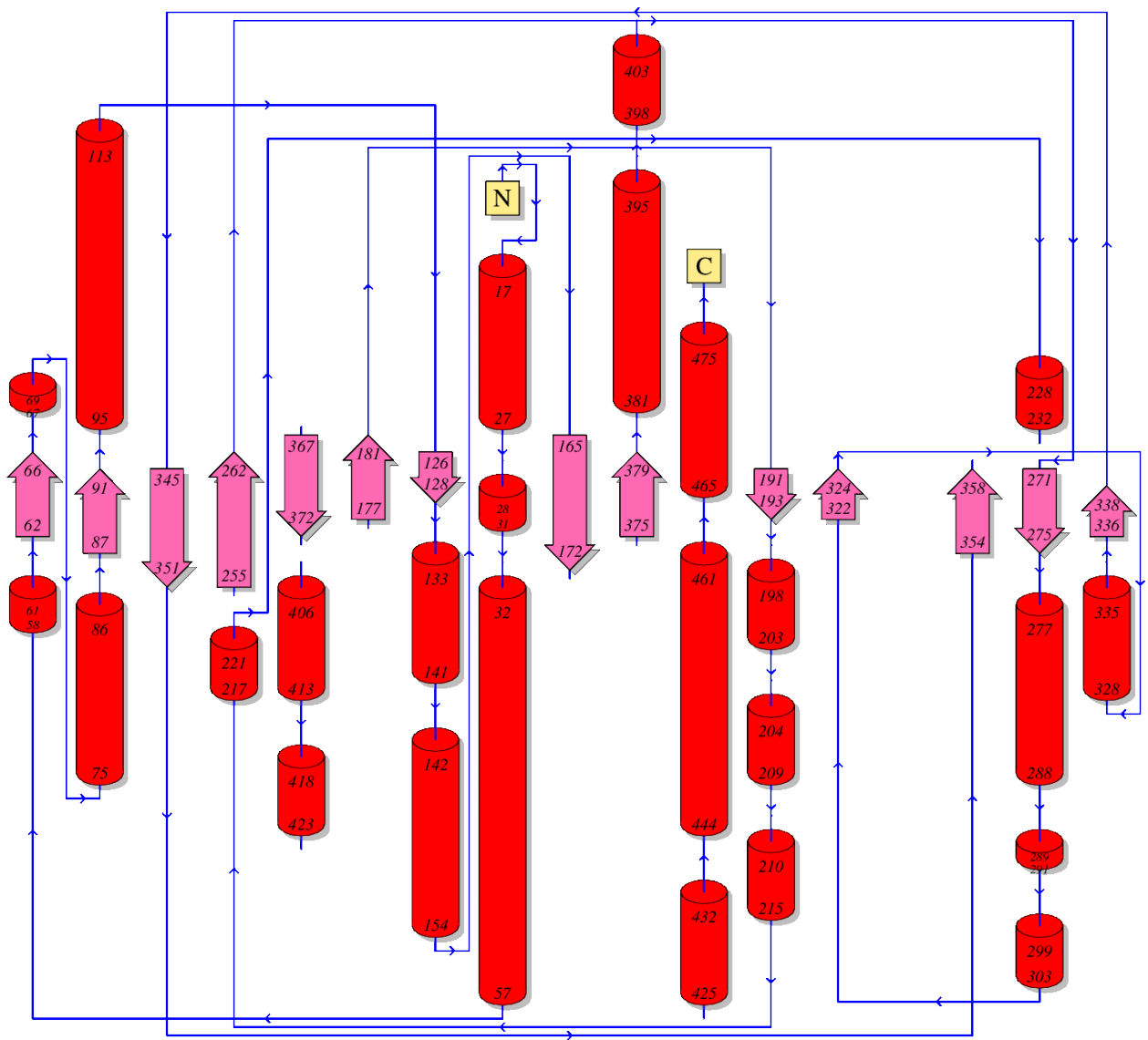
Figure 1.11: Topology diagram of complete Gig2 protein depicting the amino acid residues contributing to the secondary structure (alpha helix and beta turns) (Figure generated using PDBsum).
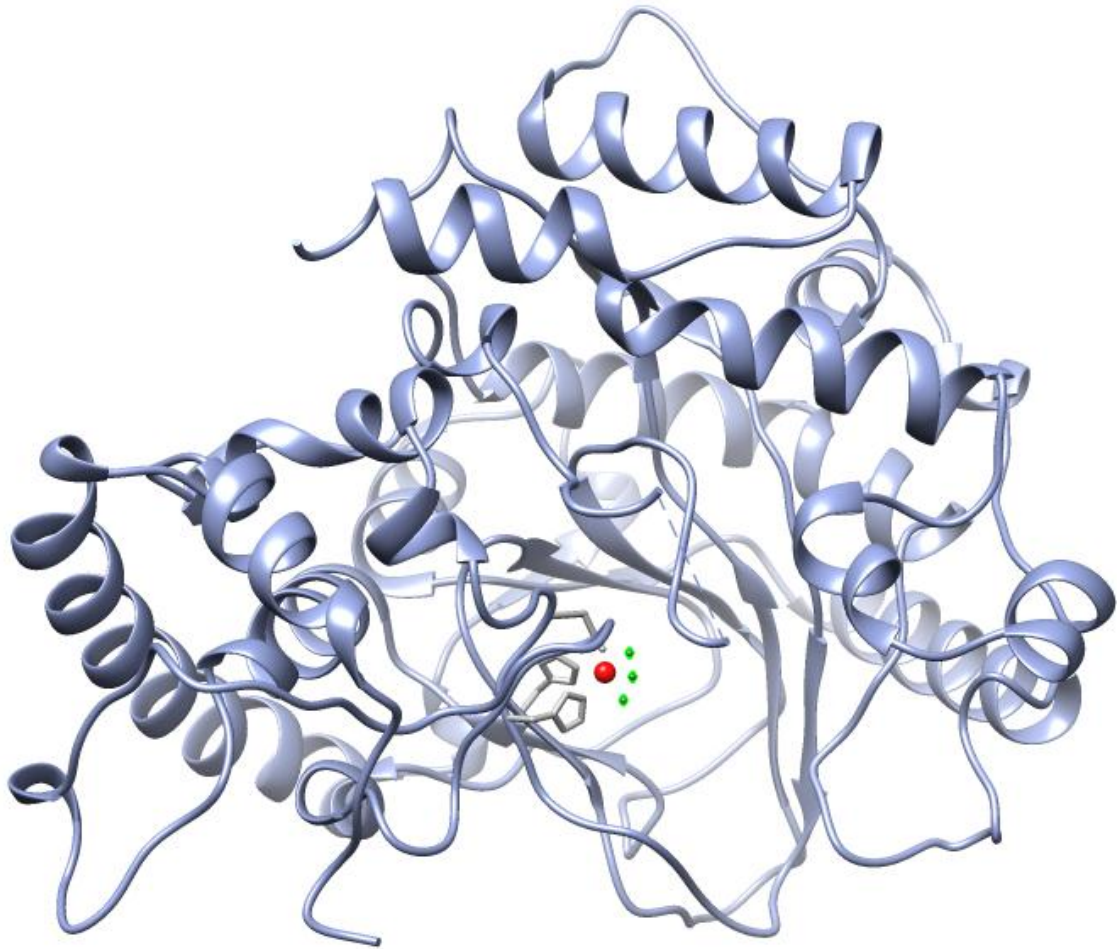
Fig 1.12: Quaternary structure of Gig2 protein. Ribbon drawing of Gig2 with Fe^{2+} ion shown as red sphere. The figure was prepared with UCSF chimera.
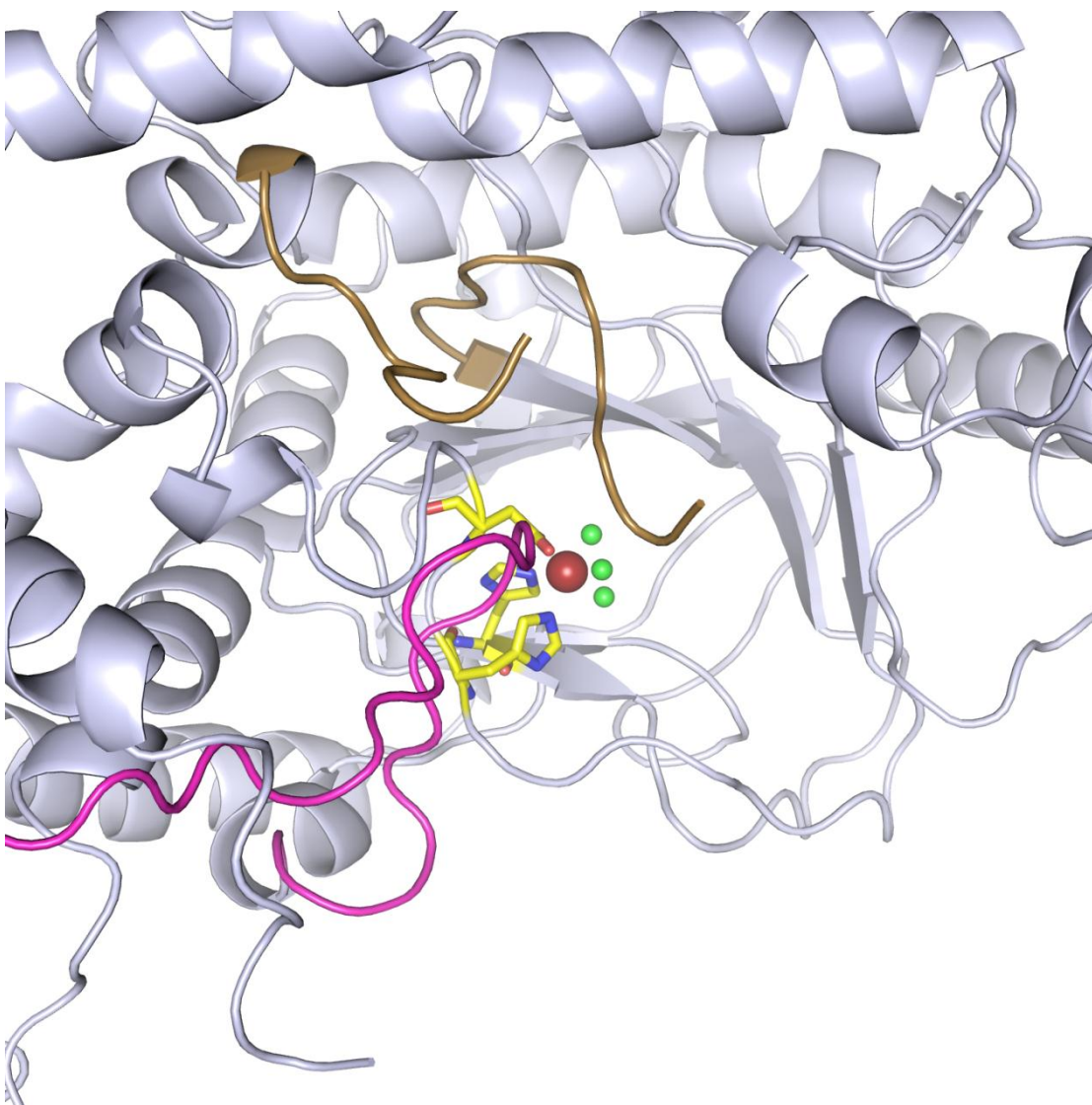
Figure 1.13: 3D structure the two loops $L_{231-256}$, L_a1 (coloured golden) and $L_{302-328}$, L_a2 (coloured pink) partially masking the active sit of Gig2. Loop L_a1 is incomplete with the density of 10 residues missing from the crystal structure (Figure generated using UCSF chimera software).
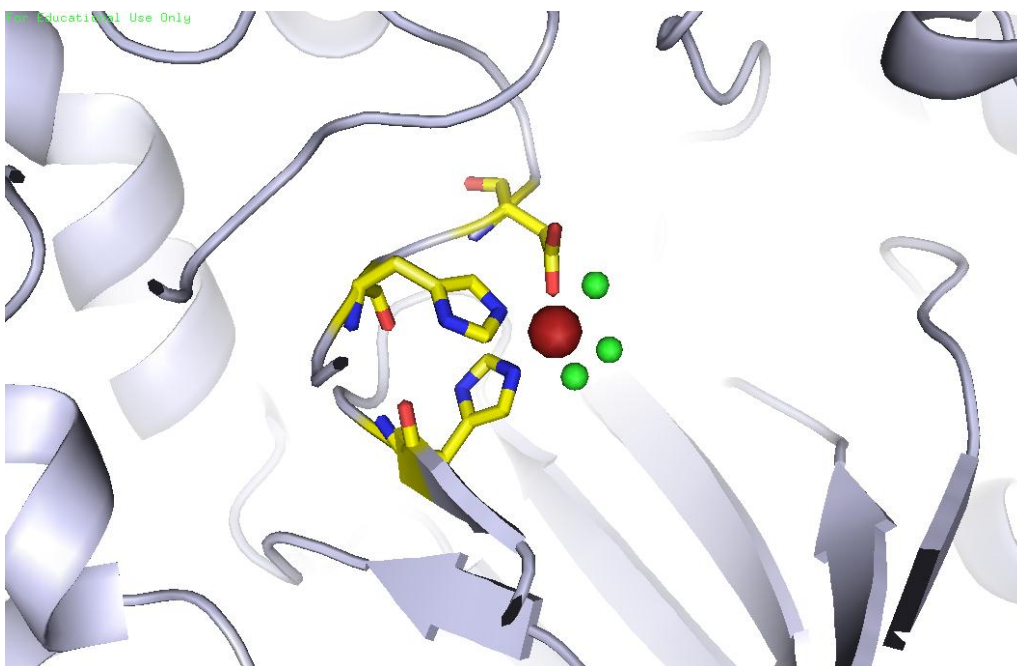
Fig 1.14: Active Site of Gig2 protein showing metal ion Fe^{2+} (coloured red) held by the 2 Histidines and 1 Aspartic acid inside the cupin fold domain; three water molecules are represented in green (Figure generated using UCSF chimera).
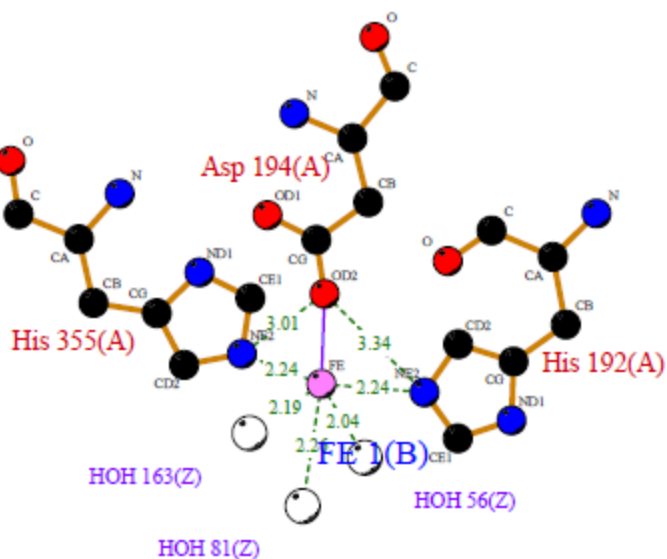
Fig 1.15: 2D representation of residual interaction of Gig2 with Fe^{2+} ;green dashed lines represent hydrogen bonds(Figure generated using Ligplot).

Discussion

Gig2 was identified as a novel uncharacterised gene upregulated by GlcNAc, a potent modulator of morphogenesis and virulence in human pathogenic fungi *Candida albicans*. Compromised virulence of Gig2 mutants projected it a as potential drug target against *Candida albicans* (Ghosh et al., 2014). But being a member of DUF family, its function is non-reported in any organism; therefore the present study was undertaken with objective to determine the crystal structure of Gig2 to decipher its function. For this Gig2 protein was purified after heterologous over-expression in *E.coli* BL21 codon plus cells and crystallization trials with different commercially available screens were done. Intial hits were obtained in 3 different conditions of Morpheus screen (Molecular Dimensions) after 3-4 days forming fine needle shaped crystals. Manual optimization of these conditions was done and macro-seeding was employed to yield diffractable crystals (Table1). Structure solution of the 2.7 Å resolution data was done by molecular replacement, using a putative Oxidoreductase from *Salmonella typhimurium LT2* (PDB ID 2CSG) as template sharing only 27% sequence identity. However, the complete structure could not be built as density for residues in the loop region was missing. To get better diffractable crystals other crystallization conditions were screened. This time, pact priemere screen was manually optimized followed by macroseeding and better long rod shaped could be obtained. These crystals diffracted at 1.7 Å resolution (Table1) and the data was solved by molecular replacement using the previous 2.7 Å resolution file as template (Table 2). Around 90% of the structure is solved with good electron density.

Structural analysis of Gig2 depicts that it exists as a monomer with a single domain made of 26 alpha helices and 13 beta sheets. The core region contains eight antiparallel beta sheets arranged in a β-jelly roll fold forming a cupin domain; a conserved barrel domain found in functionally diverse range of enzymes, transcription factors and storage proteins (Dunwell et al., 2004). Of the two characteristic motifs of cupin superfamily, Motif1: $G-(X)_5-HXH-(X)_{3,4}-E-(X)_6-G$, and Motif 2: $G-(X)_5-PXG-(X)_2-H-(X)_3 -N$; Gig2 had the characteristic motif 1 located between residues 184-216 as $G-(X)_7-HXD-(X)_4-E-(X)_{13}-G$ and motif 2 as $G-(X)_4-P-H-(X)_6 -N$ located between residues 344-362, which is slightly different in sequence from the characteristic motif 2 but structurally still has two β-sheet present in motif2 forming the cupin fold. Each of the two

motifs contain active site residues, acting as a ligands for binding active site metal ion. Gig2 has H^{192} and D^{194} from Motif1 and H^{355} from Motif 2 holding Fe^{2+} in the active site.

The 2-His-1-carboxylate facial triad is a common feature of the catalytic sites of mononuclear non-heme iron(II) enzymes (Hegg and Que, 1997). Active site similarity search for Gig2 with Dali server found less structural similarity with already reported PDB structures. The maximum similarity found was with 2-oxoglutarate dependant mononuclear dioxygenases. The 2-oxoglutarate dependant mononuclear dioxygenases comprises of an extensive family of enzymes catalysing a variety of oxidative transformation. These enzymes activate molecular oxygen (O_2) for incorporation into a diverse range of primary substrates using 2-oxoglutarate (2OG) as co-substrate to form desired product(s), succinate and CO_2 (Martinez and Hausinger, 2015). The variety of oxidative transformations carried by Fe(II)/ 2OG dependent oxygenases includes hydroxylation, halogenation, ring closure, desaturation, epimerization, ring expansion, and epoxidation reactions (Martinez and Hausinger, 2015).

Based on the structural information of Gig2 it can be concluded that it is a Mononuclear non-heme dioxygenase, possibly using 2OG as Co-substrate. Predicting and identifying the primary substrate of Gig2 is yet another challenge considering its structural uniqueness and extensive possible list of substrates.

CHAPTER 2

Chapter 2

Characterization of some other novel GlcNAc inducible genes in _Candida albicans_ -

A Hap41 a putative CCAAT binding transcription factor

Introduction

Candida albicans is a commensal pathogen where utilization of carbon and iron proves to be two important phenomena for establishment and persistence in the host. It diverged from the respiro-fermentative yeast _Saccharomyces cerevisiae_ approximately 300 million years ago and has acquired a repertoire of specialized genes during the evolutionary process. Moreover, a widely reprogrammed transcriptional circuitry also aids in the utilization of different metabolic programmes by the fungus. The CCAAT binding factor (CBF; the Hap2p/Hap3p/Hap4p/Hap5p complex) is one such transcriptional regulator responsible for activation of many of the genes involved in respiratory metabolism (Johnson et al., 2005), as well as different genes needed for metabolic functions, such as ammonia assimilation and iron homeostasis(Chen et al., 2011; Hsu et al., 2013). The CBF is basically a hetero-oligomeric transcriptional activator that is highly conserved structurally amongst multitude of organisms ranging from yeast to humans. Despite this structural conservation, the genes regulated by this transcription factor vary widely with the organism being considered (Johnson et al., 2005). The generality coupled with evolutionary divergence of such transcription factors help us to understand the niche-specific adaptation of pathogens like _C. albicans_ from an evolutionary view-point.

The _Saccharomyces cerevisiae_ CCAAT-binding factor is a hetromultimer composed of four subunits, termed Hap2p, Hap3p, Hap4p, and Hap5p which are involved in the transcriptional activation of numerous genes that encode proteins involved in respiration as well as other genes (Forsburg and Guarente, 1989a)(Forsburg and Guarente, 1989b)(DeRisi, 1997). The Hap complex binds to promoter elements containing the penta-nucleotide sequence 5-CCAAT-3. The CCAAT-binding factor is unique among DNA-binding regulator proteins as it requires three heterologous subunits for DNA binding activity. Subunits Hap2p, Hap3p and Hap5p are necessary and sufficient for binding to the target promoters (McNabb et al., 1995). With the use of functional complementation of HAP2, HAP3 and HAP5 mutants of S. cerevisiae, their respective homologues from other eukaryotes were cloned (Hooft van Huijsduijnen et al., 1990;

McNabb et al., 1997). While understanding the molecular interaction between the three subunits and DNA has been the subject of extensive studies in yeast as well as in mammalian homologs; the same is not the case with Hap4p. Hap4p is not necessary for in vitro DNA binding; it interacts directly with the Hap2/3/5p trimer. A transcriptional activation domain is present at its C-terminus (Forsburg and Guarente, 1989c). Studies on Hap5p from *S. cerevisiae* have shown it contains a 32- amino-acid fungus-specific conserved domain that is required for Hap4p to interact with the Hap2/3/5p heterotrimer and is hence termed Hap4p recruitment domain (McNabb et al., 1995).

In *candida albicans*, the CCAAT binding factor comprises of single homologs of Hap2p and Hap5p and two different homologs of Hap3p, named Hap31 and Hap32 [(Baek et al., 2008; McNabb et al., 1995; Singh et al., 2011). In addition, there are three putative homologs of Hap4, namely Hap41, Hap42, and Hap43 [(Singh et al., 2011)(Johnson et al., 2005). Studies have shown that deletion of either *HAP2* or *HAP5* leads to complete loss of DNA-binding activity. (Johnson et al., 2005),(Baek et al., 2008)(McNabb et al., 1995).It is plausible that Hap31 and Hap32 may interact individually with Hap2 and Hap5 to form DNA-binding complexes with differing regulatory functions via interaction with the three Hap4-like proteins (Singh et al., 2011)

In *C. albicans* the CCAT binding complex was initially reported to regulate components of the mitochondrial electron transport chain in response to different sugars. Accumulating evidences now highlight the involvement of Hap 43 as both a positive and negative regulator for iron homeostasis (Singh et al., 2011) and is also known to pleiotropically regulate many virulence traits *in vitro* (Hsu et al., 2013).While Hap42 and Hap43 are involved in iron utilization programme very little is known about the function of HAP41 (Chakravarti et al., 2017). Earlier studies show that *hap41* mutant have significantly reduced levels of GI tract colonization and unlike other hap genes *HAP41* does not respond to iron deprivation conditions (Pérez et al., 2013).

Materials and Methods

2.1. Strains and Plasmids

All the strains and plasmids were used have been listed in table2.1.

2.2. Media and Solutions

The media and solutions used have been listed in Appendix A.1.

2.3. Growth and Maintenance of strains

Candida albicans strain SC5314 was routinely cultured in YPD or SD media at 30°C at 200rpm shaking. For GlcNAc induction 10 ml preculture was grown for 16Hrs and 1% preculture was inoculated into SD medium for the bulk culture and was allowed to grow up-to 0.8 O.D., cells were then washed twice with 0.3% KH_2PO_4 and induced with GlcNAc in SN medium for indicated time points. *E.coli* strains were cultured in Luria-Bertani broth or agar plates containing 50μg/ml Ampicillin. For auxotrophic strains uridine was added 80μg ml^{-1} and Arginine and Leucine was added 50 μg ml^{-1}.

2.4. Storage of *C.albicans* and *E.coli* strains

Bacterial and *C.albicans* strains were stored as glycerol stocks at - 80°C. To the overnight grown cultures, sterile glycerol was added to the final concentration of 15%.

2.5 Cloning of Hap41 gene (*CaHAP41*)

Standard procedures were used for cloning and manipulation of DNA fragments (Sambrook et al., 1989) .

2.5.1 PCR amplification of *CaHAP41* gene

The gene was amplified from genomic DNA of *C. albicans* strain SC-5314. The primers were designed to amplify the entire ORF along with the respective promoter. For cloning in T-vector, PCR was done with Amplitaq (Applied Biosystems). For large scale PCR, reactions were set in multiple vials. The PCR products were pooled and were run on a 0.8% preparative agarose gel.

Table 2.1 Strains and plasmids used in this study

Strains	Relevant genotype and description	Source or reference
Candida albicans		
SC5314	Wild-type	W.A. Fonzi
SN152	*ura3/*::imm434::*URA3/ura3*::imm434 *iro1::IRO1/iro1*::imm434 *his*1::hisG/*his1*::hisG *leu2/leu2 arg4/arg4*	S.M. Noble SM, A.D. Johnson
iHAP41-GFP	*CaHAP41-GFP1-ADH1T-URA3/CaHAP41*	In this study
hap41Δ/ hap41Δ	*ura3/*::imm434::*URA3/ura3*::imm434 *iro1::IRO1/iro1*::imm434 *hap41*::HIS1/*his1*::hisG *HAP41/LEU2 arg4/arg4*	In this study
hap5Δ/ hap5Δ	*ura3Δ::imm434/ura3Δ::imm434 his1Δ::hisG/his1Δ::hisG arg4Δ::hisG/arg4Δ:hisG hap5Δ::URA3/hap5Δ::HIS1*	Fungal genetic stock centre
E.coli strains		
DH5α	F– Φ80*lacZ*ΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Invitrogen Life Technology
Plasmids		
pGEM-T Easy	T-vector for cloning Taq amplified PCR products	Promega
pGFP-URA3	Plasmid containing GFP with URA3 marker	Cheryl A. Gale
CIP-ADH1.mCherry.Hap5	Hap5 cloned in CIP-ADH1.mCherry plasmid	In this study

Table 2.2: Primer sequences used in this study

Hap41 T-F	ATGTATACTAACATTATTCTAGC
Hap41 T-R	GCTTACATTCATCAAATATTGATCAC
HAP41-GP-F1	AGCTTATTCTGGATCTTTCTCGTACTACTTCATCTTATAGTGATCAATAT TTGATGAATGTAGGTTCTAAAGGTGAAGAATTATT
HAP41-UR-R1	AATCACATCTCATCACATCACATCACATCAAAAATAAAATCCGTTAATTT TTTACTGTTCCTTCTAGAAGGACCACCTTTGATTG
HAP-DEL-F	GTTGTTGTTGTTTTTTGGAATTGAAACAAATAATCTAACAAATCTTCTTC CTTATTCCAACAAGCTCGGATCCACTAGTAACG
HAP-DEL-R	CACATCTCATCACATCACATCACATCAAAAATAAAATCCGTTAATTTTTT ACTGTTCCTGCCAGTGTGATGGATATCTGC
CHECK-MUT1F	TATACTAACATTATTCTAGC
CHECK-MUT2F	GTTCTCTCAATTCATATCTGTTCC
CHECK-HIS1R	TCTTCGGTTCTTTCAAGGATC
CHECK-LEU2R	TCTACCAATACCGAATCTAATG
CHECK-ARG4R	TGGTTCAGGTAGATATTC
HAP5-mChF	TGGTCATAATGATGAAGCTACGTATGAAAATTTTAACGGTTACCAAAAC AATTATGGTGGTGTTTCAAAAGGTGAAGAAGA
Hap5-mChR	ATTATTACAAAATCAAACACTATTTTTAAAATGAACGAAAAAAAAAAAAA AATCCTGAATTCCGGAATATTTATGAGAAAC
NAG-r1	AGGAGTGTGCCAACTACGAAAAG
NAG-r2	CCCAACCCACCTAAGAACAAATC
NGT-r1	CGGCTCGTTGTTGTCATTCA
NGT-r2	CCTTGAGCTGCCCATAAAAGA
ICL-r1	ATGCCTTACACTCCTATTG
ICL-r2	TTAAGCCTTGGCTTTGG
GAL-r1	GAATCAACAAAAATCCCCTTAGCA
GAL-r2	CTTCACATCATTGGCTTTACATACTTC

The bands were cut using a sterile blade and eluted by using QIAGEN Min Elute Gel Extraction Kit. 1μl of the eluted bands were run on a fresh gel and quantitated visually.

2.5.2. Vector Preparation:

For routine use, the *CaHAP41* sequence was cloned in pGEMT-EASY vector (Promega).

2.5.3 Ligation

Ligation reactions were set up using 50ng of T-vector and the purified PCR products, at vector: insert ratio of 1:3 nanomoles in 10μl reaction volume containing 1X ligation buffer and 1μl T4 DNA ligase. The reactions were incubated at 4°C for 16 hours.

2.5.4 Transformation and screening of recombinants

5μl of the ligation mixes were used to transform DH5α *E.coli* cells 100μl of the transformation mixes were plated on LB amp plates and incubated at 37°C. The recombinant clones were selected by isolating plasmid DNA by alkaline lysis and running along with a 3kb plasmid as control. The clones showing retarded migration were selected for further confirmation by restriction digestions. Digestions were set up in a 20μl reaction volume with 1xNEB3 buffer and 2 units of NotI enzyme. The digestions were incubated at 55°C for 2 hours. The digestions were stopped by heating the reactions at 65°C for 20mins. They were loaded on a 0.8% gel.

2.5.5 DNA Sequence Analysis

The plasmid pGEMT-Hap41 was sequenced by automated sequencer according to the manufacturer's instructions. The sequence homology searches were done using BLAST (NCBI) algorithm. Sequence alignment was carried out with CLUSTALW program. The restriction sites were determined by WEBCUTTER2. The promoters were analyzed with TRANSFAC program. The protein domain was analyzed by NCBI Conserved Domain Search and DAS Transmembrane Prediction Server.

2.6. Subcellular localization of *C.albicans HAP41* protein.

2.6.1. Generation of *HAP41*-GFP PCR module

GFP tagging: The plasmid pGFP-URA3 was kind gift from Cheryl A. Gale. This plasmid harbours a codon optimized GFP sequence along with *URA3* marker and a *C. albicans ADH1* terminator located downstream of GFP sequence. PCR was performed using this cassette as template and appropriate gene specific primers namely HAP41-GP-F1 and HAP41-UR-R1. The PCR product was used for transforming a *C. albicans ura* strain, *CAI4*. Transformants were selected by plating the transformation mix on the appropriate selective medium. To identify transformants in which the cassette had correctly integrated into the target gene sequence, genomic DNA was prepared and used as template in PCR reactions, using one primer that annealed within the transformation module (Tag1) and a second primer that annealed to the target gene locus outside the altered region (Tag2). The same strain (HAP41-GFP) was reconfirmed by Western analysis using anti-GFP antibody.

2.6.2 Western Blotting

The cells containing the GFP-epitopes were lysed. The lysates were prepared for gel electrophoresis. SDS gel electrophoresis was done with the lysate. The protein was transferred to a PVDF membrane in transfer buffer containing 20% methanol, 24 mM Tris base, 194 mM glycine. The membrane was transferred to a tray and incubated for 1 hour at +15 to +25°C (or overnight at 4°C) with a 1:10 dilution of Western Blocking Reagent (Cat. No. 11 921 673 001) in phosphate buffered saline with 0.1% Tween 20, pH 7.5 (PBST).The membrane was washed three times with PBST. The Anti-GFP antibody concentrate was diluted to 1 µg/ml in a 1:20 dilution of Western Blocking Reagent. The membrane was incubated with this diluted Anti-GFP antibody for 1–2 hours at +15 to +25°C with gentle rotation. The membrane was washed three times with 1x Wash Buffer .Goat Anti-Mouse IgG (H+L) HRP Conjugate was diluted 1:4000 into a 1:20 dilution of Western Blocking Reagent. The membrane was incubated with this diluted antibody solution for 1 hour at RT with gentle rotation. The membrane was again washed three times with 1x Wash Buffer. Detection solutions A and B (Amersham ECL plus Western Blotting Detection Reagent) were mixed in a ratio of 40:1. The excess wash buffer was drained from the membranes and the membrane was placed on a clean Saran wrap with the protein side up. The

mixed detection reagent was pipetted onto the membrane and incubated for 5 mins at room temperature. The membrane was exposed to X-ray film and developed within 1-5 mins interval.

2.6.3 Study of subcellular localization of Hap411 under different Carbon sources.

Overnight cultures of SC5314, HAP41-gfp were inoculated into fresh YPD medium to an optical density of 0.1(measured at 600 nm) and 1% inoculum from overnight grown culture was added to fresh media and grown till the cultures reached at OD600 of 0.8.The cells were then pelleted and further washed with sterile MQ and resuspended in the YNB with 2% GLcNAc/Glucose/Lactose medium and kept at 37C with shaking. Microscopic analysis for the subcellular localization was done after 30minutes of induction. Photographs were taken in confocal microscope.

2.7 Gene disruption of *HAP41* in *C. albicans*

The deletion for *HAP41* (orf19.740) was constructed in *C. albicans* strain SN152 using the method described in literature (Noble and Johnson, 2005). Precisely, PCR primers (HAP41-DEL-F and HAP41-DEL-R) containing ~80 bases of sequence homologous to the sequence flanking the open reading frame of Ca*HAP41* were used to amplify either *LEU2* or *HIS1* or the *ARG4* selectable marker gene (from plasmids pSN40 and pSN52 respectively). Integration of the deletion cassettes at the appropriate sites was verified by performing PCR using combinations of primers that flanked the integration as well as the primers that annealed within the integrated cassettes (CHECK-MUT1, CHECK-LEU2 and CHECK-HIS1). Complemented strains were constructed by transforming a plasmid carrying one wild-type copy of Ca*HAP41* and *ARG4* selectable marker gene into the *C. albicans* genome (primer, CHECK-ARG4 used for confirmation). For revertant preparation, Hap41 cloned in Cip30 was used.

2.7.1 Morphological studies:

Induction of filaments in liquid media

For studying the germ tube formation in various liquid media, the wild type and the mutant strains were patched on YPD plates and grown at 30°C overnight. The freshly patched cells were inoculated in 10 ml YPD, and grown at 30°C with 200rpm shaking for 6 to 7 hours till they

reached the logarithmic growth phase. The cells were then harvested by centrifuging at 5,000 rpm for 5 minutes, and were washed twice with sterile MQ-water. They were resuspended in 5ml MQ-water and starved for 10 hours at 30°C, with shaking at 120 rpm (Delbrück and Ernst, 1993).

The starved cells were resuspended at the concentration of 0.5 OD_{600}/1ml in various induction media such as 2.5 mM GlcNAc in salt base (A1), 2.5mM GlcNAc and 2.5mM glucose in salt base, Spider pH 7.2 (A1) and Serum (20%) (A1). They were induced at 37°C for 4 hours in case of GlcNAc induction and for 2 hours for Spiders and Serum induction. Cells were fixed with 4% Formaldehyde, and germ tube formation was examined under a light microscope.

Study of Morphogenesis on Solid Plates

Cells were grown in 10ml YPD medium at 30°C for 2 days to obtain a saturated culture. They were counted using a haemocytometer. The formula used to calculate the number of cells per ml was: average number of cells per square x dilution factor x 10^4. An aliquot of the diluted cells was plated on YPD plates to count the CFU. Cells were then plated on Spider plates and SLAD plates at a concentration of 80-100 cells per plate. Spider plates were incubated at 37°C for 7 days. SLAD plates were incubated at 37°C for 10 days.

Growth on different Carbon source medium supplemented with or without Antimycin

In order to determine the effect of disruption of *HAP41* on the growth of the mutant strain on GlcNAc containing media, spot dilution assay was performed. Cells were serially diluted and spotted on SD (2% glucose), SN (2% N-acetyl glucosamine) and Lactate (2%) ±4 µg/ml antimycin A plates. Plates were incubated at 30°C and 37°C for 4 days.

Microscopy

The photographs of germ tubes were taken at ×300 magnifications in a Nikon 80i microscope. Leitz Aristophot microscope. The photographs of the plates were taken in a Leica MZ6 stereomicroscope.

2.7.2 Quantitative PCR

For all reverse-transcription-PCR (RTPCR) experiments, total RNA isolated using TriPure isolation reagent was treated with RNase-free DNase I (Invitrogen) to remove any residual DNA. About 500 ng of this RNA was used for single-stranded cDNA synthesis by using a High-Capacity cDNA reverse transcription kit (Applied Biosystems), and the cDNA was used for quantitative RT-PCR (qRT-PCR) with SYBR green PCR master mix on an ABI Prism 7000 real-time PCR apparatus (Applied Biosystems). The comparative threshold cycle (*CT*) method (2___*CT*) was used to determine relative gene expression (Schmittgen and Livak, 2008). Control reactions without reverse transcriptase were carried out for each cDNA preparation and ascertained that no amplification was obtained, as judged by high *CT* values and gel analysis.

2.3 Results and Discussion

2.3.1. Cloning of *HAP41* gene

Candida albicans HAP41 gene including the upstream promoter region of the ORF 19.740 was amplified from the genomic DNA (Fig2.1.A). Primers used are mentioned in table 2.1. The amplified product was cloned in pGEMT easy vector (Promega). Transformant colonies were screened by miniprep alkaline lysis method (Fig 2.1.B) and digestion with NotI. Two of the positive clones were then used for plasmid purification by plasmid purification kit (Promega). Fig 2.1.C shows fallout of 1.5kb and vector backbone of 3kb on NotI digestion. The plasmid pGEMT-Hap41 was sequenced and the sequence matched with *HAP41*sequence obtained from the Candida genome database.

2.3.2. Epitope tagging of *HAP41* gene

For chromosomal tagging of *HAP41*gene PCR based cloning in plasmid pGFP-URA3 was done. Figure 2.2[A] shows graphic representation of the pGFP-URA3 cassette. PCR was performed using this cassette as template and appropriate gene specific primers namely HAP41-GP-F1 and HAP41-UR-R1 (Table 2.1). It yielded a 3.1 kb PCR product (Figure 2.2 [B].) The PCR product was used for transforming a *C. albicans ura* minus strain, *CAI4*. Transformants were selected by plating the transformation mix on the SD-URA selective medium. Colonies obtained were screened for correct integration by PCR, using one primer that annealed within the transformation module (Tag1) and a second primer that annealed to the target gene locus outside the altered region (Tag2). The same strain (HAP41-GFP) was reconfirmed by Western analysis using anti-GFP antibody

2.3.3 *HAP41* is inducible in presence of GlcNAc

Earlier performed microarray analysis under Glycerol versus GlcNAc and Glucose versus GlcNAc conditions, identified several genes that show differential regulation in response to GlcNAc. In both the conditions 19.740 came up as a prominently induced gene, the fold of induction being ~ 10 folds in case of Glycerol vs. GlcNAc and ~ 8 folds in case of Glucose versus GlcNAc. To further confirm the specificity of its expression, we have carried out western blot analysis for strain possessing this gene C-terminally tagged with GFP tag.

The western blot analysis revealed that its expression is specific only for GlcNAc and gets upregulated within 15-30 mins of induction after which there is a decrease in the protein level. In the control strain i.e. *CAI-4* no band could be observed (Figure 2.3B). Expression of Hap41-GFP is inducible in absence of glucose while Hap5-mchery is constitutively expressed (Fig2.3C).

2.3.4 Sub-cellular localization of Hap41:

For protein localization data, we used Hap41-GFP strain that has chromosomally tagged *HAP41* at the C-terminal region with GFP tag using a PCR-fusion based strategy. C-terminal Hap41-GFP-fusion protein localized to cytoplasm in the presence of 2% GlcNAc or lactate, since under this concentration, effects of GlcNAc catabolism are prominent. No detectable signals were observed in glucose. This reconfirmed that this gene is induced by GlcNAc and translated into a functional protein.

2.3.5 Preparation of *hap41* mutant

The deletion for *HAP41* (ORF19.740) was constructed in *C. albicans* strain SN152following a previously described method (Noble and Johnson, 2005). Using the PCR primers (HAP-DEL-F and HAP-DEL-R) either either *LEU2* or *HIS1* selectable marker gene were amplified from plasmids pSN40 and pSN52 as respective templates. Integration of the deletion cassettes at the appropriate sites was verified by PCR. Separate combinations of primers (CHECK-MUT1-F, CHECK-LEU2-R, CHECK-HIS1-R and CHECK-HAP41-del-R) that flanked the integrated cassette and the primers that annealed within the integrated cassettes were used. Revertant strain was constructed by transforming the *hap41*mutant strain with a plasmid carrying one wild type copy of *CaHAP41* along with an *ARG4* selectable marker gene (primers, CHECK-MUT1-F and CHECK-ARG4 were used for confirmation).

2.3.5.1 The *hap41* mutant exhibited little defect in filamentation under inducing conditions.

The effect of of *HAP41* disruption on the colony morphology of *C. albicans* was investigated on different growth and filamentation media like YPD (37°C), Spider (37°C) and GlcNAc(30°C) plates. *hap41* homozygous mutants did not show significant defects in hyphal formation under all of the above mentioned conditions. Thus we concluded that *HAP41* was not required for filamentation process in *C. albicans* (Fig 2.6A).

2.3.5.2 *hap41* mutants showed reduced growth on GlcNAc and lactate plates.

The *hap41* mutant strains were also tested for its ability to grow on alternate carbon source, GlcNAc and non-fermentable carbon source, lactate by spotting serial dilutions of cells onto a GlcNAc, lactate or dextrose (as the carbon source) containing solid agar medium with or without antimycin A. The homozygous mutant showed compromised growth on both GlcNAc and lactate containing media showing that this gene is indispensable for growth on GlcNAc or lactate. This effect was more prominent in presence of antimycinA an inhibitor of oxidative phosphorylation.

2.3.5.3 q-PCR

The relative expression of three previously known GlcNAc inducible genes: *NGT1*, *NAG1*, *GAL10* along with a Lactate inducible gene *ICL1* was checked in *HAP41* and *HAP5* mutants grown on glycerol, GlcNac and Lactate. It was observed that there was downregulation in the expression of all four genes in *Hap41* as well as *HAP5* mutants in GlcNAc condition. Expression of Lactate inducible *ICL1* was also downregulated in both mutants.
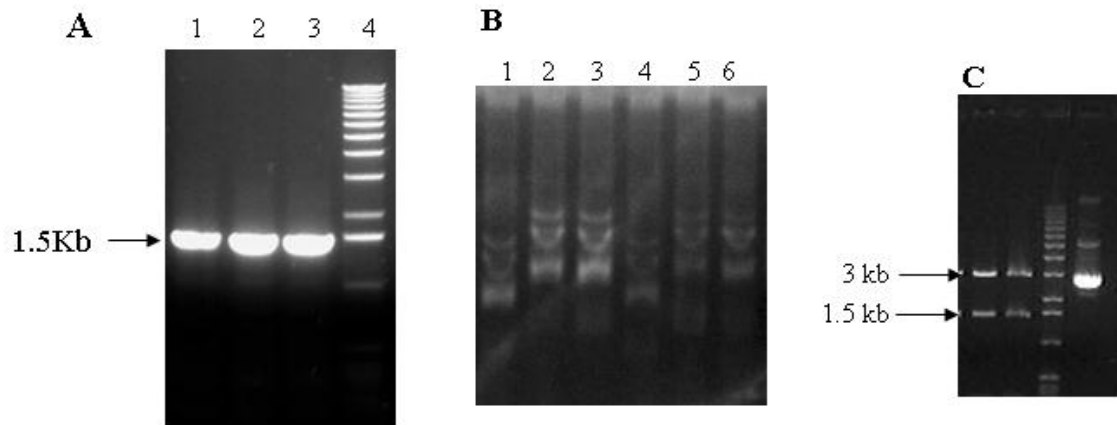
Figure 2.1: [A] Amplification of *HAP41* from *C.albicans.* [B] Plasmid isolation showing shifts (lanes 2, 3, 5 and 6). [C] pGEMT-HAP41 digestion with NotI to release inserts
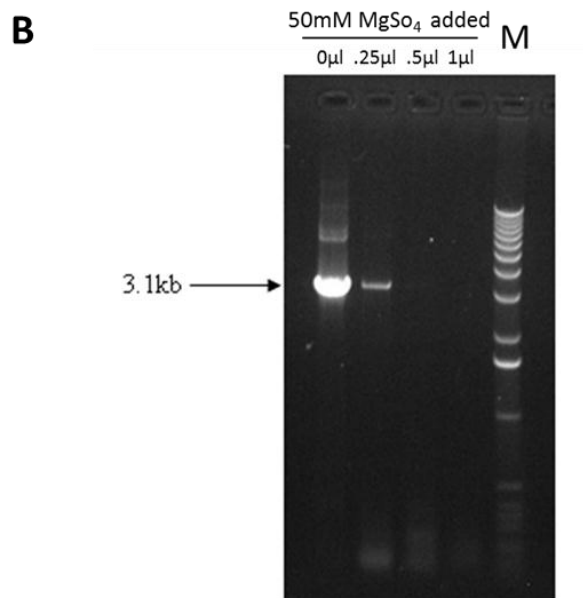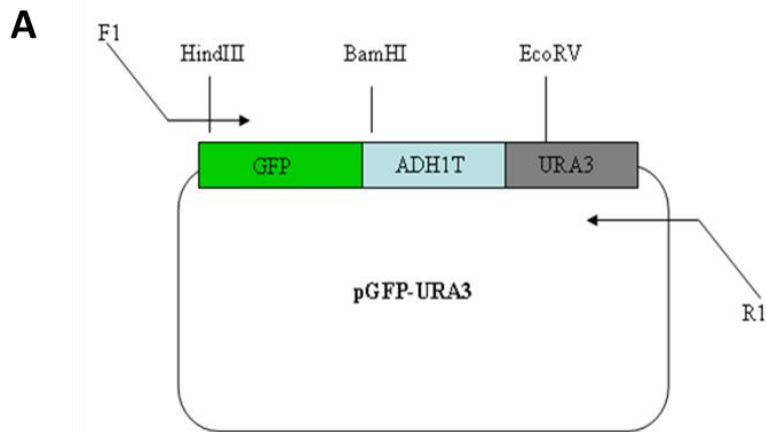
Figure2.2: Schematic representation of GFP tagging module and its PCR amplification. [A] Plasmid carrying the cassette used to generate tags at the 3'- ends of the *C. albicans* gene. Green box: codon-optimized GFP; blue box: *C. albicans* ADH1 termination sequence (ADH1T); Grey box: *C. albicans* auxotrophic markers URA3. Oligonucleotide primers (F1, R1,) used to generate the tag cassettes by PCR are described in text [B] 3.1 kb PCR product using primers F1 and R1. Amplification optimization by addition of different concentration of $MgSO_4$ to the PCR reaction mix
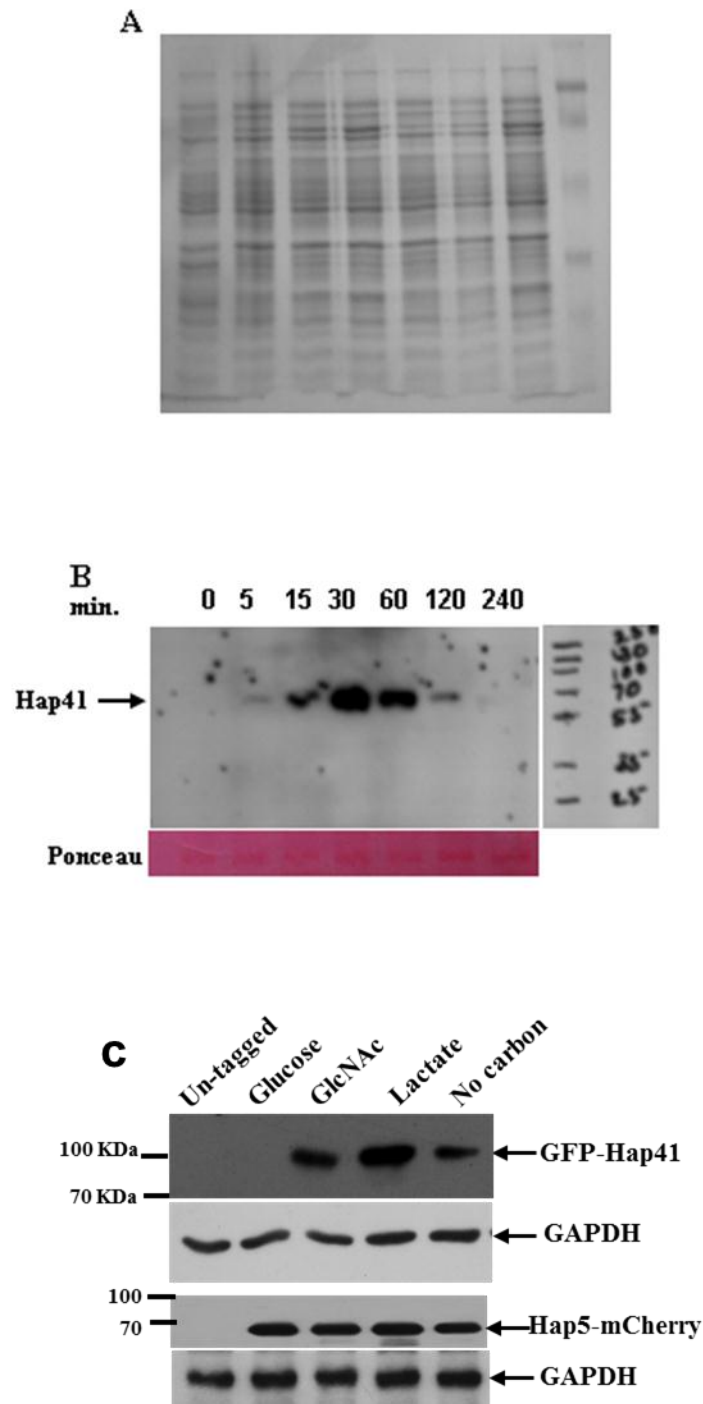
Figure 2.3: Western Blotting analysis showing the expression of HAP41 gene. [A] Coomassie stained gel. [B] Western blot showing the time kinetics for Hap41 protein expression. [C] Western blot showing expression of GFP-Hap41 and Hap5-mCherry grown in different carbon sources.
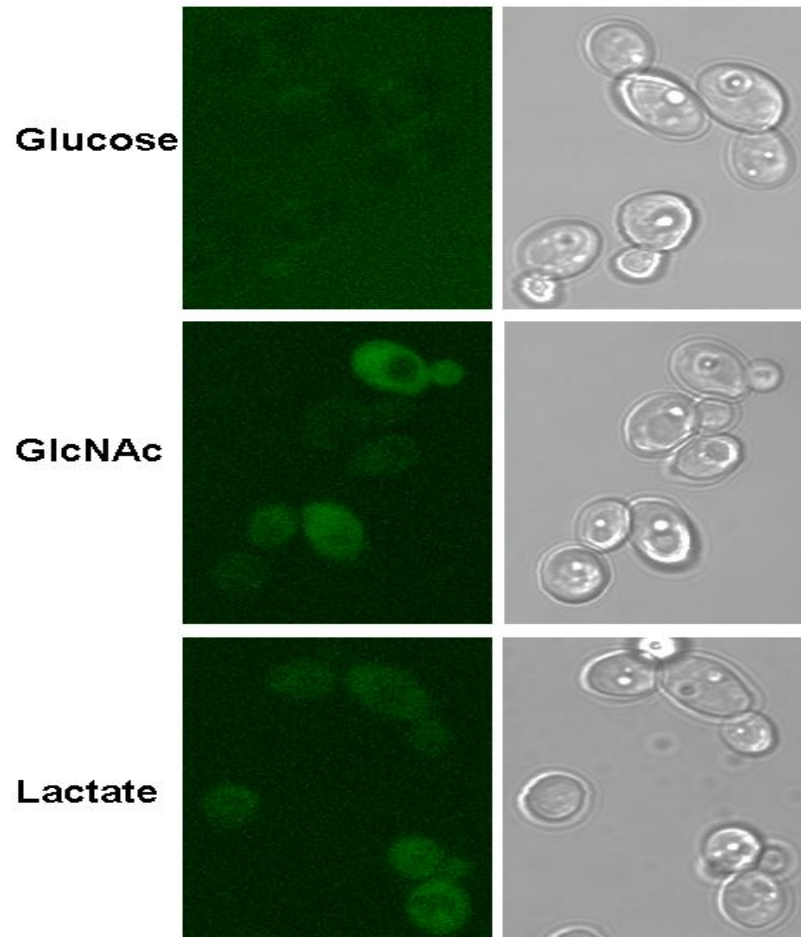
Figure 2.4: Hap41-GFP localization on different carbon sources.Hap41 localizes to cytoplasm when induced by GlcNAc and Lactate but not by Glucose.
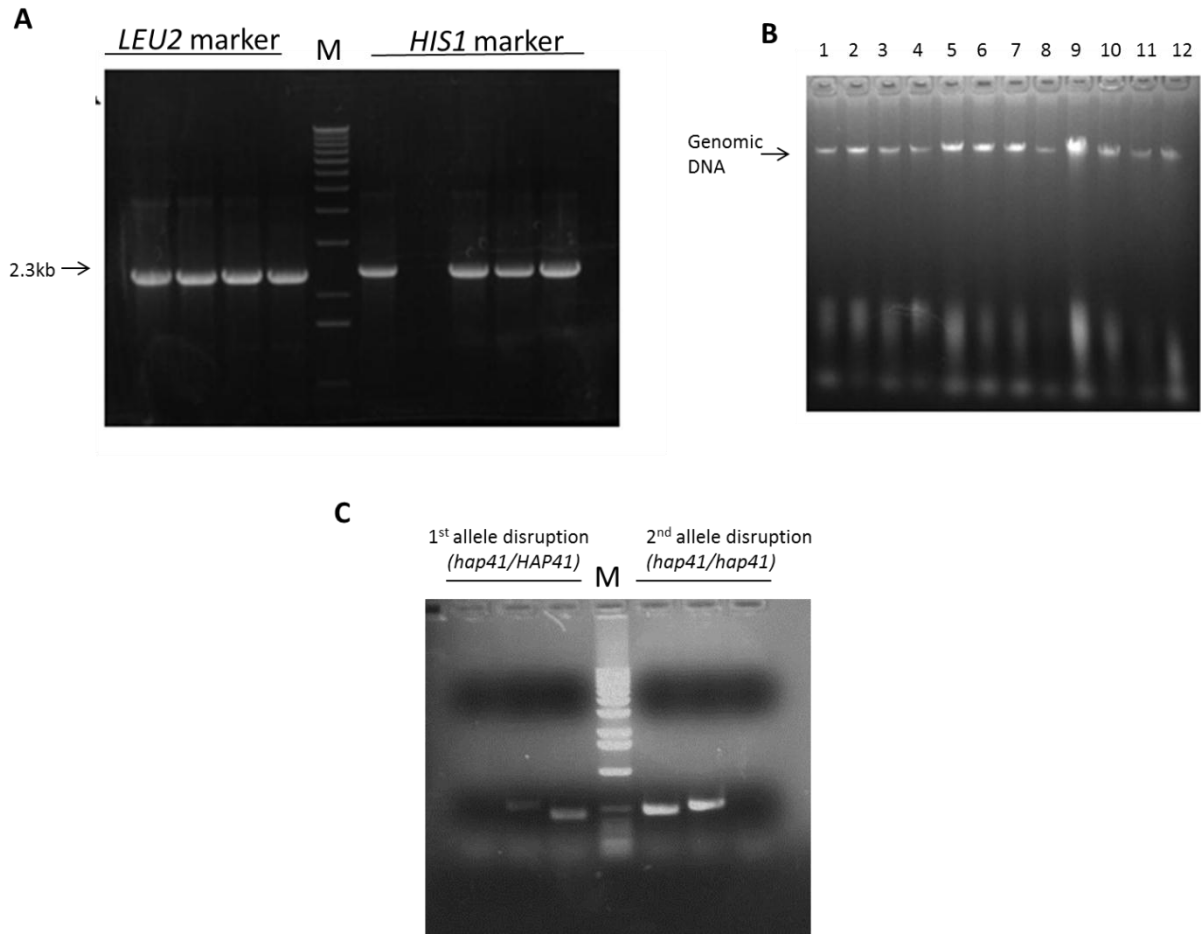
Figure 2.5: *hap41* mutant preparation: [A] PCR amplification of Selection markers (*LEU2*/*HIS1*) with Hap41 del F and R primers using plasmids SN40 and SN52 as template DNA. [B] Genomic DNA isolation of transformants, [C] PCR confirmation of marker integration and gene disruption. First and second allele transformant checked using CHECK MUT1 F with CHECK *HIS1*R, CHECK *LEU2R* and CHECK HAP4-DEL-R primers.
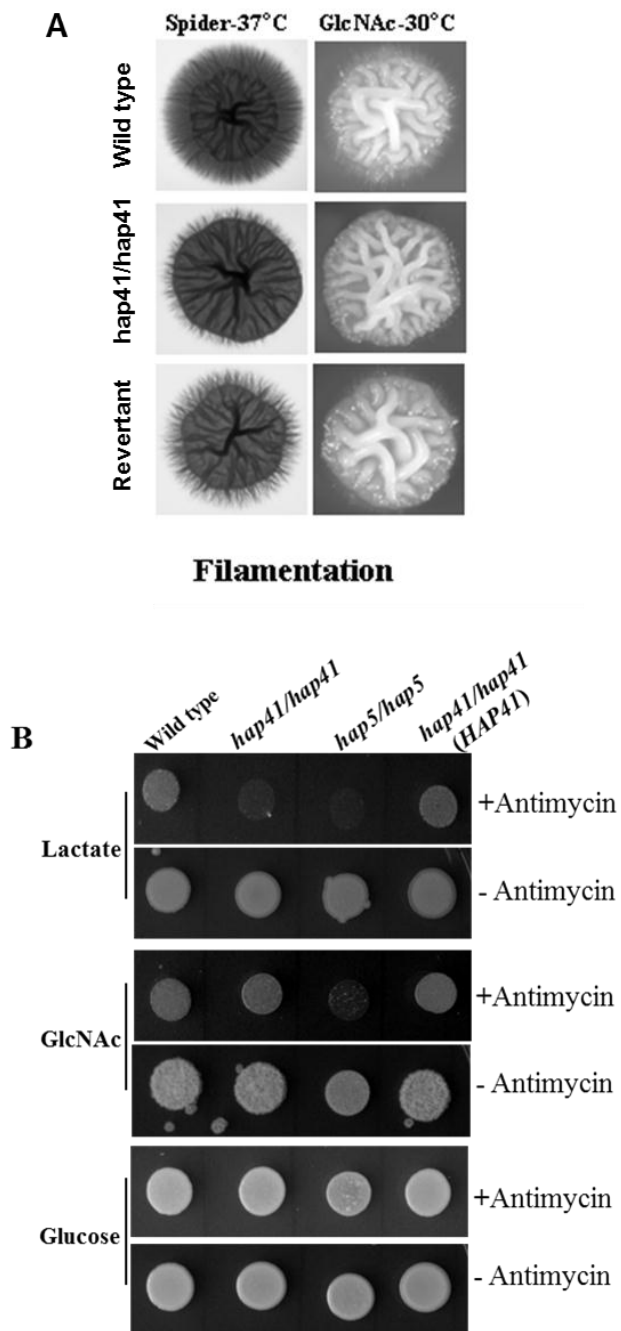
Figure2.6: [A] Growth of *hap41*mutant along with wild type and *HAP41* revertant on filamentation inducing media Spider and GlcNAc; [B] growth pattern of Hap41 and Hap5 mutants on different sugar sources supplemented with or without 4µg/mL antimycin A.
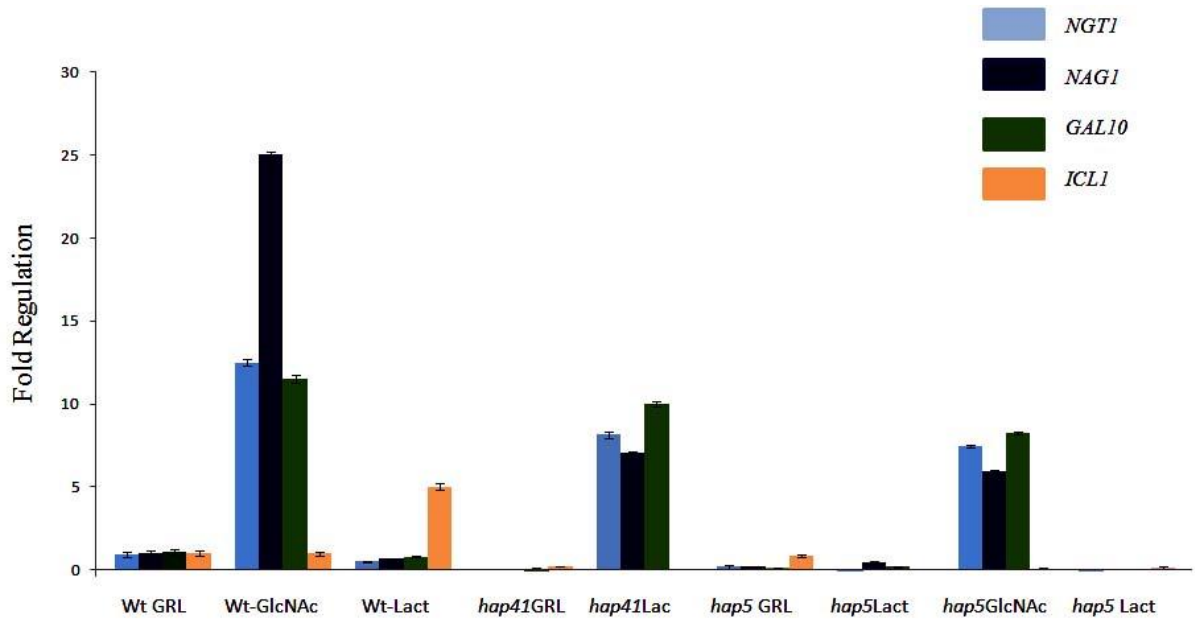
Figure 2.7: Graphical representation of the variable expression for the four selected genes [NGT1, NAG1, GAL10 and Isocitrate lyase (ICL)] in three different carbon sources Glycerol (GRL), GlcNAc and Lactate (Lact). All the four genes were upregulated in presence of N-acetyl glucosamine in wild type but downregulation was observed in case of hap41 and hap5 mutants

Discussion

Hap 41 is a putative Hap4 like transcription factor in *Candida albicans*. Hap 4 like proteins lack direct DNA binding ability and interact with core CBP complex via Hap5; acting as regulatory co-transcription factors. In General, eukaryotic transcriptional activator proteins are bipartite in nature, with distinct DNA binding and transactivation domains. Of the three Hap 4 like orthologs in *Candida albicans,* Hap42 and Hap43 have been reported to have roles in iron metabolism and oxidative stress response (Chakravarti et al., 2017). Hap41 does seem to be playing part in iron metabolism suggesting it might be having some novel function.

This work seeks to investigate its role in GlcNAc metabolism and signaling since it was found to be upregulated by GlcNAc in the previous microarray data from our lab. To study the expression profile of Hap41 it was chromosomally tagged at the C-terminal region with GFP tag following a PCR-based strategy (Gola et al., 2003) (Fig2.2). The western blot analysis of Hap41 expression after GlcNAc induction at different time points revealed that its expression is specific for GlcNAc and gets upregulated within 15-30 mins of induction after which there is a decrease in the protein level (Figure 2.3). Early induction within 15-30mins also indicates that GlcNAc catabolism may not be required for its induction. Expression of Hap5 is constitutive while Hap41 is expressed in the absence of glucose and induced in response to GlcNAc or Lactate (Fig2.3C). This differential expression of Hap41 suggests its regulatory role in the Hap complex during the utilization of GlcNAc or non-fermentative carbon sources like Lactate. Cellular localization studies of Hap41 under different sugars show its localization in the cytoplasm in the presence of GlcNAc or Lactate but not in the presence of glucose after 30 minutes of induction. These results reapprove induction of Hap41 expression by GlcNAc and downregulation in presence of glucose. Since Hap41 is a transcription factor specifically up-regulated in response to Lactate or GlcNAc, its nuclear localization pattern or dynamics need to be studied. Time-lapse live imaging microscopy studies have to be carried out to study the Hap41 localization dynamics in order to understand the molecular mechanism of Hap complex mediated gene regulation in *Candida albicans*.

hap41 mutants exhibited deceased filamentation on Spider plates although there was not much difference in filamentation on GlcNAc plates when compared with wild type strain (Figure 2.6A). *hap5* mutant showed completely arrested growth on both GlcNAc and Lactate medium,

whereas *hap41* mutant was unable to grow on Lactate medium while it showed little difference on GlcNAc plates (Figure 2.6B). The defect was more prominent in the presence of antimycinA, an inhibitor of cytochrome c reductase which inhibits oxidative phosphorylation. These results indicate that *Hap41* is indeed required for growth on GlcNAc and Lactate.

In quantitative PCR, the three previously known GlcNAc inducible genes: Ngt1, Nag1 and Gal10, expression is suppressed in *hap41* and *hap5* mutants. Lactate inducible Isocitrate lyase (*ICL1*) expression is also downregulated in both mutants. Isocitrate lyase is a glyoxylate cycle enzyme which helps in metabolic adaptation in the absence of glucose (Ishola et al., 2016). Our studies indicate that Hap complex with Hap41 in regulatory role, as a positive regulatory factor involved in the expression of metabolic genes for utilizing alternate carbon sources in the absence of glucose. Our present study and various recent studies (Chakravarti et al., 2017; Singh et al., 2011; Thiébaut et al., 2017) indicate that in Candida albicans, Hap complex has acquired additional features to adapt under various adverse stress conditions in the host.

SUMMARY

SUMMARY

Candida albicans is an important human fungal pathogen, keeping in mind its high frequency of isolation from human host and the associated morbidity and mortality rates. It resides as a harmless commensal in the micro-biome of most healthy individuals and can turn pathogenic in cases with compromised immunity leading to localized infection or even disseminated candidiasis(Pfaller, 1996; Wisplinghoff et al., 2014). Recovery from localized mucocutaneous infection is still possible by treatment with antifungals however invasive candidiasis is much likely to be life threatening. With around 34% mortality rate, Systemic candidiasis is the leading cause of nosocomial infections.

Candida albicans is a eukaryotic pathogen sharing much similarity in overall biological processes with humans making its treatment particularly difficult along with toxic side effects. Thorough understanding of the biology of this pathogen and its mechanism of infection at molecular level is required to find novel drug targets.

Polymorphism or pleomorphism is a characteristic feature in the biology of *Candida albicans* closely associated with its pathogenicity and virulence. Morphogenesis in *Candida albicans* is triggered by many environmental cues, one of them being GlcNAc; an amino sugar abundantly present at the mucosa, the primary site of infection.

GlcNAc serves multiple vital functions in *Candida albicans,* like being a structural component, providing an alternative carbon source as well as acting as a signaling molecule regulating several different biological processes. GlcNAc catabolism and signaling are key to its virulence and pathogenicity.

Although the major catabolic pathway of GlcNAc utilization has been well established in *Candida albicans*, various possible secondary pathways bifurcating from the catabolic intermediates are still elusive. Previous work from our lab on the transcriptome analysis by microarray of *Candida albicans* cells under GlcNAc inducible conditions have identified several novel genes with previously unknown function to be differentially regulated. Gig2 and Hap41, the focus of the present study are two such genes which were found to be induced by GlcNAc.

Major findings of this study are as follows:

- Crystal structure determination of Gig2 was undertaken with the purpose to identify its function in *Candida albicans*. Gig2 crystallizes as a monomer in P2$_1$ space group. The structure is solved to 1.7 Å resolution. It is made of a single domain with 26 alpha helices and 13 beta sheets. The core of the protein contains a cupin fold carrying Fe^{2+} in the catalytic site a characteristic feature of mononuclear non heme dioxygenases.

- The crystal structure of Gig2 has very less structural similarity with the known PDB structures and amongst them also most similar structures are uncharacterized proteins. Thus the primary substrate of Gig2 enzyme could not be identified.

- Hap41was chromosomally tagged at the C terminus with a GFP tag(Gola et al., 2003) to study its expression and localization.

- Hap41 protein expression was induced by GlcNAc and reached its peak within 15-30 minutes as seen in the western blot analysis.

- Hap41p localizes to the cytoplasm in GlcNAc and Lactate medium as seen by fluorescent microscopy of Hap41-GFP.

- Hap41 deletion mutant was prepared using PCR based strategy (Noble and Johnson, 2005).

- *hap41* mutants showed very less difference in the filament formation on GlcNAc however on spider medium filamentation was reduced. Growth on GlcNAc and lactate was compromised and was even more compromised when antimycinA, an inhibitor of cytochrome c was used. This indicates involvement of Hap41 in regulation of non-glycolytic alternative metabolic pathways.

- qPCR of three GlcNac inducible genes *NAG1*, *NGT1, GAL10* and Lactate inducible *ICL1* indicated relative downregulation in *hap41* and *hap5* mutants. This is a direct link of positive regulation of GlcNac metabolic genes by Hap complex.

In conclusion, this study identifies Gig2 as a mononuclear non-heme dioxygenase based on the crystal structure analysis. Although the substrate identification for Gig2 could not be done, it remains as a subject of future studies. Once the substrate is identified, site directed mutational studies of the key residues will be fruitful in designing inhibitors against this enzyme. Alongside, Hap41 a putative CAAT binding co-transcription factor is identified as functional protein in *Candida albicans* whose expression is induced by GlcNAc and Lactate in the absence of

glucose. Future work on the molecular mechanism of dynamics of interaction of Hap41 with the rest of the core CBP complex will enhance our understanding of GlcNAc and Lactate regulated gene expression.

Appendix1

APPENDIX 1: MEDIA AND SOLUTIONS

All % shown are on a W/V basis unless mentioned otherwise. All solutions and media were made in Milli-Q and Milli-RO water respectively. All solutions and media were sterilized by autoclaving at 15 lb/ sq inch for 15 minutes or filter sterilized by passing through a $0.22\mu M$ Millipore filter. Media used were from Invitrogen or Difco. Chemicals used were of analytical grade (Qualigens or Merck) or molecular biology grade mostly from Sigma, Invitrogen or USB. Restriction enzymes and DNA modifying enzymes were obtained from NEB, MBI-Fermentas, Roche, Promega, Amersham or Perkin Elmer. Buffers provided with the enzymes were used.

A1.1. Bacterial Media:

1. Luria Bertani (LB) Medium (Invitrogen): 1%Tryptone, 1% NaCl, 0.5% Yeast Extract

2. LB agar: LB with 1.5% Agar (Invitrogen).

Antibiotics were added to the media, after autoclaving, and cooling to <50°C, at indicated concentrations. Ampicillin was added to the concentration of 50μg/ml, to liquid medium and 75μg/ml, to Agar medium. Streptomycin was added at the concentration of 10μg/ml to liquid media and 15μg/ml to Agar medium.

3. Terrific Broth: 1.2% Bacto-tryptone (Difco), 2.4% Yeast extract (Invitrogen), 0.4% (v/v) glycerol. Supplemented with one tenth volume 10X Phosphate Buffer (2.31g K_2HPO_4, 12.54 g KH_2PO_4 in 100ml H_2O).

4. 2 X L Broth:	2% Bactotryptone, 1% Yeast Extract, 0.1% NaCl, 0.2% glucose.

A1.2. Media for growth, maintenance and morphogenetic studies of *Candida albicans*:

1. YPD:	1% Yeast extract, 2% Peptone, 2% Dextrose
2. YPD Agar:	YPD with 2% agar
3. SD:	0.67% Yeast nitrogen base without amino-acids (YNB w/o aa), 2% dextrose, supplemented with Uridine at the concentration of 25 μg/ml, when required.
4. SD Agar:	SD with 2.0% agar.
5. SN:	0.67% Yeast nitrogen base without amino-acids (YNB w/o aa); 2% GlcNAc.
6. SN Agar:	SN with 2% agar.
7. 5-FOA Plate:	0.67% YNB w/o aa, 2% Dextrose, 1mg/ml 5-Fluoroorotic Acid (filter sterilized and added when media is <50°C), 25μg/ml Uridine and 2% Agar.
8. GPK:	0.5% Glucose, 0.5% Peptone and 0.3% K_2HPO_4

9. NPK: 0.5% GlcNAc, 0.5% Peptone and 0.3% K_2HPO_4

10. Spider (pH 7.2):   1% Nutrient Broth, 1% Mannitol, 0.2% K_2HPO_4,
                                     1.35% Bactoagar for plates. Adopted from Liu et
                                     al., 1994.

11. SLAD: 0.17% YNB w/o aa and $(NH_4)_2SO_4$, 2% Dextrose,
                                     2% Bactoagar (washed several times with
                                     autoclaved water) and 50μM filter sterilized
                                     $(NH_4)_2SO_4$, added after autoclaving. Adopted from
                                     Gimenco et al. 1992.

12. 2.5mM GlcNAc in salt base: 0.45% NaCl and 0.335% YNB w/o aa. Adopted
                                     from Delbruck and Ernst, 1993.

13. 20% Serum: 20% Bovine Calf Serum in YPD.

14. SCAA   0.67% YNB (w/o aminoacids); 2% Casamino acids
                                     (Difco). Adopted from Leuker et al., 1997.

A1.3. Reagents and Buffers used for bacterial transformation, plasmid and genomic DNA isolation, and molecular cloning:

1. Trituration Buffer: 100mM $CaCl_2$, 70mM $MgCl_2$, 40mM NaOAc (pH
                                     5.5).

2. X-Gal: 20 mg/ml X-gal, in DMF. Stored at -20°C in dark.

3. IPTG: 200 mg/ml IPTG, in water, filter sterilized and
                                   stored at -20°C.

4. Solution I: 25mM Tris-HCl, pH 8.0; 10mM EDTA; 50mM
                                   Dextrose.

5. Solution II: 1% SDS, 0.2N NaOH.

6. Solution III: Potassium acetate, 3M with respect to Potassium
                                   and 5M with respect to acetate, as per Sambrook et
                                   al., 1989.

7. TE (pH 8.0):   10 mM Tris-HCl (pH 8.0), 1mM EDTA (pH 8.0).

8. TE (pH 7.6):   10mM Tris-HCl (pH 7.6), 1mM EDTA (pH 8.0)

9. 3M NaOAc (pH 5.2):   40.81gm of sodium acetate trihydrate was
                                   dissolved in 80ml water. pH was set to 5.2
                                   with glacial acetic acid. The volume was
                                   adjusted to 100ml. Sterilized by autoclaving.

10. 5M NH₄OAc.

38.5gm of ammonium acetate was dissolved in water and the volume was adjusted to 100ml, autoclaved.

11. Phenol, Qualigens:

Redistilled at 180°C and stored frozen at minus 20°C in small parts.

12. TE Saturated phenol:

Prepared according to Ausubel et al (1994)

13. Phenol: Chloroform: IAA:

25 parts of TE saturated phenol, 24 parts of chloroform and 1 part of Isoamylalcohol.

14. RNase A:

10 mg/ml. Stock was prepared as described in Sambrook et al., (1989).

15. CTAB:

Hexadecyltrimethyl ammonium bromide (CTAB) (Sigma H-6269) 10% in 0.7M NaCl.

16. Proteinase K:

20mg/ml in water.

17. NaCl 5M:

29.22gm of NaCl was dissolved in 80ml water and the volume was adjusted to 100ml. Sterilized by autoclaving.

18. Acrylamide gel elution buffer:

0.5M ammonium acetate, 10mM EDTA (pH 8.0), 0.1% SDS (optional).

19. Lysozyme: 10mg/ml in 10mM TrisCl pH 8.0.

20. 2X PEG/NaCl: 13% (w/v) PEG 8000 and 1.6M NaCl.

A1.4. Reagents and Buffers used for Agarose Gel Electrophoresis and Polyacrylamide Gel Electrophoresis:

1. 50X TAE: 242g Tris base, 57.1ml Glacial acetic acid and 100ml 0.5M EDTA, pH 8.0 for 1 litre.

2. Ethidium Bromide: 10mg/ml in water.

3. EndoR (6X):   30% Ficoll 400, 60mM EDTA pH 8.0, 0.6% SDS and 0.06% bromophenol blue.

6. 30% Acrylamide: 29:1 Acrylamide and Bis-acrylamide was prepared and deionised as recommended in Sambrook et al., 1989.

7. 25% APS: Prepared freshly by dissolved at 250mg/ml in water.

8. 10X TBE: 54g Tris base, 27.5g boric acid and 20ml 0.5M EDTA was added for 1litre soln.

A1.5. Reagents and Buffers used for *C.albicans* transformation and Genomic DNA isolation:

1. LATE buffer: Lithium acetate 100mM, Tris-HCl pH 7.5 10mM, EDTA 1mM.

2. PLATE buffer: 40% PEG 3350 in LATE buffer.

4. Breaking Buffer: 10mM Tris-HCl, pH 8.0, 1% SDS, 1mM EDTA, 2% Triton-X-100 and 100mM NaCl.

Appendix2

APPENDIX 2: COMMONLY USED TECHNIQUES

A2.1. PCR Amplification and Optimization

PCR was carried out in an MJ Research PTC100 thermocycler, with a heated bonnet.

A2.1.1. Amplification

Typically the amplification reactions were done for ~30 cycles, with specific alterations arrived at, empirically to optimize yield.   The program was as follows,

94°C denaturation for 2 minutes only in the first cycle

Optimum annealing temperature $T_{a \text{ OPT}}$ for 30 seconds, for all cycles

 94°C denaturation for 1 minute for all cycles

 72°C extension for 1 minute for all cycles

 Cycling for ~30 times

 72°C for 10 minutes to fill the incomplete extension products,

 4°C for a minimum of 10 minutes

A2.1.2. Optimizations

1. Optimization of Annealing Temperature:

The annealing temperature was calculated using the following parameters according to Rychlik et al., 1990.

$$T_{a \text{ Primer}} = 0.3 T_{m \text{ Primer}} + 0.7 \, T_{m \text{ Product}}$$

$$T_{m \text{ Product}} = 0.41(\%G+C) + 16.6 \, (\log_{10} [K^+]) + (675/l) + 81.5 \text{ °C,}$$

Where, l defines the length of the product in bp, and $[K^+]$ represents the potassium ion concentration in 1x PCR buffer which is normally provided as 50 mM.

And $T_{m\ Primer} = 2\times (A+T) + 4 \times (G+C)$, where A, T, G and C are the number of the respective nucleotides in the oligonucleotide primer sequence.

The $T_{a\ OPT}$ was usually taken as the lower of the two $T_{a\ Primers}$ calculated as above.

The T_a thus arrived at, was quite close to the ideal temperature, just a difference of 0.7°C.

2. Optimization of $MgCl_2$:

The Mg^{++} ion concentration of the amplification reaction was optimized empirically, varying the concentration from 1mM to 4mM in increments of 0.5 mM to obtain the best   amplification.

3. Optimization of yield**:**

Using the above parameters, the template amount, and number of cycles were varied and an ideal combination was arrived at to maximize the yield.

A2.2. Gel Electrophoresis of DNA fragments

1. Agarose Gel Electrophoresis:

Agarose gel electrophoresis of DNA was routinely carried out in 1X TAE, in the presence of 0.5µg/ml ethidium bromide (A1.5.2) both in gel and in running buffer. For most purposes 0.8% gel was used. Depending on the expected size of the band, the gel strength was varied from 0.7%, for visualizing uncut genomic DNA, to 2.0% for checking total RNA. Before performing any southern or northern, the integrity of the nucleic acid was also verified by such methods. For the exact measurement of very small fragments, ranging from 75 to 500bp of DNA, nondenaturing Polyacrylamide Gel Electrophoresis (PAGE) was carried out as described later (A2.2.3). All procedures were followed from Sambrook et al., 1989, or Ausubel et al., 1994.

2. Gel Elution of inserts from Agarose Gel:

Large scale digestion of the plasmid containing the DNA of interest was set with appropriate enzymes to release the insert. Preparatory gel was casted using low melting agar. The digested DNA was loaded on the gel and run at low voltage till the band of interest was well separated from the vector backbone. The band was cut out with a new scalpel. Equal volume of TE

saturated phenol was added to it and kept in minus 80°C for 1 hour. It was then centrifuged at 13,000 rpm at room temperature, and the aqueous layer was transferred into a new tube. The DNA was extracted once with chloroform: isoamyl alcohol and precipitated by adding $1/10^{th}$ volume of 3M NaOAc (A1.4.9) and 2.5 vols of chilled ethanol. The precipitate was washed, dried and dissolved in 20μl TE. Alternatively the DNA was prepared from the gel slice using QIAEX II Gel Extraction Kit (QIAGEN). PCR products were purified from the gel similarly.

A2.3. Cloning of gene of interest in vectors

Vectors most commonly used in these studies are pUC19, pBS KS+, pGEM-T Easy and pGP704.

Enzymatic manipulations of DNA, like restriction digestion, Klenow end filling, dephosphorylation by CIAP, ligation, etc. were carried out essentially as described in Sambrook et al., 1989 and following the manufacturers' instructions, wherever applicable. All the preparative digestions for the preparation of inserts and vectors were generally set up in 20 μl volume with 10-15 μg of plasmid DNA. Digests were resolved on 0.7-1.0% agarose gel and appropriate DNA fragments were cut out from the gel. As reference, marker DNA ladders were loaded along with the samples. DNA was eluted from the gel piece either using QIAEX II Gel Extraction Kit (QIAGEN) or by phenol freeze method. DNA amounts were empirically estimated. Modifications like end polishing, dephosphorylation, etc. were carried out and recovered DNA was further estimated. Ligation reactions were set up at vector: insert ratio of 1:4 (in general) in 10 μl reaction volume containing 1X ligase buffer and 3.3 units of T4 DNA ligase (NEB). For blunt end ligation, additionally, PEG (MW 8000) was added to a concentration 15% and incubated at 20°C for 16 hours. Sticky end ligation was carried out at 16°C. Cloning in pGEM-T Easy was carried out as per manufacturers' instructions.

A2.4. Transformation of *E. coli*

A2.4.1. Hanahan Method for preparation of competent cells

1. Preparation of *E. coli* DH5α Competent Cells

DH5α was tested for recA as described in Maniatis et al. (1982). Chemically competent cells were made according to the method described by Hanahan (1991) with some modifications.

DH5α cells were streaked on fresh LB-agar plate from glycerol stock. A single colony was inoculated from the freshly streaked plate, in a preculture of 10ml of 2XL medium supplemented with 0.2% glucose (A1.1.4). It was grown at 30°C at 200 rpm shaking for 16 hours.

The overnight grown preculture was diluted 1:100 times in fresh 2XL supplemented with 0.2% glucose and incubated at 30°C with shaking at 200rpm till A_{600} reached 0.45. The bacterial culture was then chilled on ice-water bath for 2 hours. The cell culture was transferred to chilled SM34 tubes and pelleted at 5000 g for 5 min at 4°C in a chilled rotor. The supernatant was discarded and the pellet was resuspended in half the culture volume of freshly made trituration buffer (A1.4.1). The cell mix was incubated on ice-water bath for 45 minutes. Cells were pelleted by centrifugation at 5000 g for 10 minutes at 4°C. Competent cells were then resuspended in trituration buffer in a $1/20^{th}$ volume of the starting culture volume. 80% glycerol was added to it to a final concentration of 15% (v/v).

Cells were aliquoted on ice and immediately frozen at –80°C in microfuge tubes and stored at the same temperature until further use.

2. Transformation of competent cells

For each transformation, a 100 μl aliquot of the competent cells was used. Cells were thawed on ice. 5-10 μl of ligation mix or 1-20ng of plasmid DNA was mixed with 1X TE pH 8.0 (A1.4.7) to make a final volume of 50 μl. 100 μl cells was added to it and incubated for 45 minutes on ice with gentle tapping at regular intervals. The cells were subjected to heat shock at 37°C in a water bath for 5 minutes and then chilled on ice for 5 minutes. 3.85 ml of 2XL broth was added and was allowed to outgrow for 90 minutes at 37°C at 200 rpm shaking. As a positive control of transformation, 1ng of intact pBluescript II vector was used and a negative control was kept without addition of DNA. Transformation efficiency was determined by counting the number of colonies / μg of DNA on selection plate. In case of transformation of ligation mix, a vector only control ligation was also transformed along with vector-insert ligation, to check the efficiency of dephosphorylation reaction of the vector.

3. Plating of Transformation Mix

40 μl of 20 mg/ml X-gal (A1.4.2) and 4 μl of 200 mg/ml IPTG (A1.4.3) were spread on 25 ml LB-agar plate containing 75μg/ml ampicillin. An aliquot of 200 μl out of the 4 ml-transformation mix was spread on each plate. The plates were dried in laminar flow air and then incubated in 37°C incubator, agar side up until the colonies appeared (16 hours). The plates were then stored at 4°C until blue and white coloured colonies became distinct.

In the case where blue/ white selection was not present, transformation mixes were plated on LB ampicillin plates.

A2.5. Screening and Analysis of Recombinants

A2.5.1. Plasmid isolation

The recombinant colonies (white), from vector and insert plate were picked up and patched on a fresh LB-agar plate containing ampicillin and simultaneously inoculated into 1.0 ml TB medium containing 50μg/ml ampicillin. The cells in liquid culture were grown at 37°C in an incubator shaker at 200 rpm for 18-20 hours and the plate was incubated in 37°C incubator until the patched colonies appeared.

The cultures were processed to isolate plasmid DNA by the miniprep protocol and analyzed by restriction digestion.

1. Small Scale Plasmid DNA Isolation

Small scale plasmid DNA isolation was carried out by modified alkaline lysis method of Sambrook . (1989). Cells were pelleted at 12,000 g for 20 seconds at room temperature (RT). The supernatants containing the media were removed by aspiration, leaving the bacterial pellet as dry as possible. The pellets were resuspended completely by vortexing in 50 μl of Solution I, TEG (A1.4.4). 100 μl of freshly prepared Solution II (A1.4.5) was added to them and the tubes were inverted few times to mix well. They were incubated at RT for 5 minutes till a visible cell lysis occurred and the liquids became transparent.

75 µl Solution III (A1.4.6) was added to the cell lysates and mixed well. They were incubated on ice for 10 minutes to allow the precipitation of cell debris. They were then centrifuged at 12,000 g for 7 minutes at RT. The pellets of bacterial debris were removed using sterile toothpicks.

The plasmids were precipitated by the addition of equal volumes of isopropanol to the supernatants. They were mixed by inversion and incubated at RT for 15-30 minutes. The plasmid pellets were recovered by centrifugation at 12,000 g for 5 minutes at RT. The supernatants were removed by gentle aspiration; pellets were washed once with 70% ethanol and dried under vacuum. The pellets were dissolved in 20 µl TE, pH 8.0.

2 µl of each sample were separated on a 0.8 percent agarose gel along with the marker plasmid which was the uncut vector used for ligation. Recombinants were selected based on their slower migration in comparison to the vector. The putative positive clones were verified by restriction digestion with suitable enzymes.

2. Medium Scale Plasmid DNA Isolation

A confirmed positive clone was inoculated for Midi-preparation of plasmid DNA in 50ml TB. Alkaline lysis method was done as described by Sambrook et al. (1989) with modifications as described below.

A single colony containing the plasmid from LB-agar plate supplemented with Ampicillin (75 µg/ml) was inoculated in 50 ml TB medium supplemented with the same antibiotic. The cells were grown overnight (16-18 hours) at 200 rpm in a 37°C incubator shaker. The culture was harvested by centrifugation at 8000 g at 4°C for 5 minutes. The cells were resuspended in freshly made 5 ml of Solution I (A1.4.4) by vortexing. After thorough suspension, lysozyme (A1.4.19) was added to a final concentration of 1mg/ml and incubated at RT for 10 minutes.

10 ml of freshly prepared solution II (A1.4.5) was added to the suspension and was mixed gently by inversion and incubated at RT for 10 minutes. 7.5 ml of 3M NaOAc, pH 5.2 (Soln. III, A1.4.6) was added to it and mixed thoroughly by inverting the tube gently. The tube was stored on ice for 20 minutes. The lysate was centrifuged at 12,000 g for 15 minutes at 4°C and the supernatant was carefully transferred to a fresh tube avoiding any precipitate.

The plasmid DNA was precipitated by addition of 0.6 volume of isopropanol at RT for 30 minutes. It was centrifuged at RT at 12,000 g. The pellet was washed with 70% ethanol and dried under vacuum. It was then dissolved in 2 ml TE and collected in a SM-24 tube. To it DNase free RNase A was added to a final concentration of 20µg/ml and incubated at 37°C for 30 minutes. The plasmid DNA was precipitated by PEG precipitation. NaCl and PEG was added at a final concentration of 0.8M and 6.5% respectively (A1.4.20); mixed by gentle inversion and incubated in ice-water bath for one hour.

The solution was centrifuged at 10,000 rpm for 10 mins. The supernatant was removed and the pellet was washed with 70% alcohol, dried and dissolved in 500µl of TE. This was then transferred in a fresh micro centrifuge tube. The plasmid DNA was extracted twice with phenol: chloroform: isoamyl alcohol, once with chloroform: isoamyl alcohol and precipitated by adding 1/10th volume of sodium acetate and 2.5 volume of absolute alcohol.

Plasmid DNA was pelleted at 12,000 g at RT for 15 minutes; washed twice with 70% alcohol and dried under vacuum. It was dissolved in 200 µl TE, and estimated spectrophotometrically. The routine yield was 400-800 µg.

A2.6. Isolation of Genomic DNA from bacterial strains

A modification of the method of Murray and Thompson (1980) was used for DNA extraction. *Vibrio cholerae* strains were inoculated in 10ml LB and grown at 37°C for 16 hours. Cells were pelleted by spinning the culture at 10,000 rpm for 5 mins. The cells were resuspended in 567µl of TE buffer. 30µl of 10% SDS and 3µl of freshly prepared proteinase K solution (20mg/ml) (A1.4.16) were added to it. It was then incubated at 37°C for 1 hour. 100µl of 5M NaCl solution (A1.4.17) was added to it, followed by the addition of 80µl of CTAB/NaCl (A1.4.15) solution heated to 65°C. The solution was incubated at 65°C for 10mins. The DNA was extracted once with phenol: chloroform: isoamyl alcohol (25:24:1) and once with chloroform: isoamyl alcohol. The aqueous solution was transferred to a fresh tube and equal volume of isopropanol was added to it. The DNA pellet was washed with 70% alcohol, dried and resuspended in 30µl TE.

A2.7. Transformation of *C.albicans*

Transformation was done using the lithium acetate method as described by Braun and Johnson (1997). A single well isolated colony of ura minus strain of *C.albicans*, was inoculated in 10 ml of YPD supplemented with Uridine. It was grown at 30°C at 200rpm shaking for 16-18 hrs. 1% inoculum was added to 50ml culture and grown till the culture reached mid log phase. The cells were then pelleted by centrifuging at 7,000 rpm at RT for 10 minutes. The pellet was resuspended in 5ml LATE buffer (A1.6.1). The cell mix was centrifuged at 7,000 rpm for 10 minutes. The supernatant was discarded and the cells were resuspended in 500µl LATE buffer.

100µl of competent cells, 5µl of 10mg/ml sheared, single stranded, salmon sperm DNA and 2-10 µg of digested DNA were mixed together. The ssDNA was denatured by boiling for 5 minutes and chilling on ice for 5 minutes prior to addition to the transformation mix. The transformation mix was incubated at 30°C for 30 minutes. Then 700µl of PLATE buffer (A1.6.2) was added to it and vortexed for few seconds. It was then incubated at 30°C overnight. Next day, heat shock was given at 42°C for one hour. Cells were pelleted and washed with 1 ml of TE pH 8.0. The cells were finally resuspended in 200 µl of TE. Around 100µl of the transformation mix was plated on appropriate plates.

A2.8. Genomic DNA Isolation from *C. albicans*

The *C.albicans* strains were inoculated in 10 ml YPD and grown at 30°C with shaking at 200 rpm, till saturation. The cells were collected by centrifugation at 5,000 rpm for 5 minutes, resuspended in 0.5 ml distilled water and transferred to micro centrifuge tubes. The cells were washed by giving a brief spin at 5,000 rpm for 2 minutes and the supernatants were decanted. The pellets were resuspended in the residual liquid by vortexing.

0.2 ml of freshly prepared breaking buffer (A1.6.4) was added to each tube. To it 0.2 ml phenol: chloroform: isoamylalcohol and 0.3 g acid-washed glass beads were added. The cells were vortexed at full speed for 4 minutes and spun at 10,000 rpm for 5 minutes in a microcentrifuge. The aqueous layer was transferred to a new tube and 1ml of chilled ethanol was added to each of them.

The tubes were spun at 10,000 rpm for 2 minutes. The supernatants were discarded. The pellets were washed with 70% ethanol, air-dried, and resuspended in 0.4 ml 1X TE (pH 8.0).

30 µg RNaseA (A1.4.14) was added and incubated at 37°C for 15 minutes. 10µl 4M ammonium acetate and 1ml absolute alcohol was added to each and mixed gently by inverting the tubes. The tubes were centrifuged at 10,000 rpm for 2 minutes. Pellets were washed with 70% ethanol, air-dried, and resuspended in 50 µl 1X TE (pH 8.0).

2 µl of each sample was loaded and checked on a gel, DNA quantity was estimated visually.

A2.9. RNA isolation from *Candida albicans*

Total RNA was isolated from 10 ml of culture either by the hot phenol method (Ausubel et al., 1994) or by phenol and guanidine thiocyanate method using the monophasic TriPure Isolation Reagent (BOEHRINGER MANNHEIM).

1. Hot Phenol Method

The cell pellets from various conditions were collected in microfuge tubes. The pellets were resuspended in TES buffer (A1.8.1) to which 250 µl phenol saturated in DEPC treated water was added and vortexed. The mixture was heated at 65°C for 1 hour with intermittent vortexing. The hot mixture was then chilled on ice for 5 minutes. 200 µl TES was added to the mixture and vortexed vigorously again. The mixture was then centrifuged at 12000 g for 10 minutes at RT. The proteins and unlysed cell debris formed a thick interphase, above the organic phase. The aqueous phase was carefully removed to another tube, without disturbing the interphase, even if it meant leaving some of the aqueous phase behind. The aqueous phase was extracted with chloroform once, and the RNA was precipitated by adding 1/6th volume 2M NaOAc pH 4.2 (A1.8.3) and twice the volume ethanol. The RNA was recovered by centrifugation at 12000 g for 10 minutes at RT in a microfuge. The pellet was washed in 70% ethanol in DEPC water, dried and dissolved in DEPC treated water.

2. TriPure Method of RNA Isolation

TriPure isolation reagent was added to the polypropylene tubes containing cell pellet at room temperature. 1 ml reagent was used for each ~10^7 cells. Equal volumes of 0.45 mm chilled glass beads were added and homogenized for 3 minutes on cyclomixer. This was incubated at room temperature for 5 minutes to ensure complete dissociation of the nucleoprotein complexes. For phase separation, chloroform was added, 0.2 ml for each 1 ml of TriPure reagent used. The tubes were capped, vortexed vigorously for 15 seconds and incubated at room temperature for 10 minutes. The tubes were centrifuged at 12,000 g for 15 minutes at 4°C, to separate the solution into three phases. The colourless upper aqueous phase was collected in a new tube and the RNA was precipitated by adding isopropanol (0.5 ml for each 1 ml TriPure reagent used) and incubated at room temperature for 10 minutes. RNA was pelleted down by centrifuging the tubes at 12,000 g for 10 minutes at 4°C; supernatant was discarded and the pellet was given 75% ethanol wash. It was then air-dried and dissolved in DEPC-treated RNase free water incubating at 55°C for 10 minutes.

A2.10. Microscopy

For study of morphogenesis in liquid media, the cells were fixed in 4% formaldehyde and observed at x300 magnification in a Leiz Aristophot microscope. The photographs of the plates were taken in a Leica MZ6 stereomicroscope.

A.2.11. SDS-Polyacrylamide gel electrophoresis

SDS-PAGE was carried out under reducing conditions. The separating gels (12-14% arylamide as per need) was prepared using acrylamide (acrylamide:bis-acrylamide=29:1) in 1.5% Tris-Cl pH 8.8, 0.1% (w/v) SDS, 0.04% (w/v) APS and TEMED. After polymerization of separating gel, stacking gel was poured. The stacking gel contained 4% acrylamide in 0.5% Tris-Cl pH 6.8, 0.1% (w/v) SDS, 0.04% (w/v) APS and TEMED. Prewarmed samples and 4X SDS-PAGE loading dye [125 mM Tris-Cl pH 6.8, 4% (w/v) SDS, 10% (w/v) 2-mercaptoethanol, 20% (v/v) glycerol and 0.2% (w/v) bromophenol blue] were mixed to 1X dye concentration and reboiled for 3 min. After electrophoresis, proteins were fixed in the gel by incubating in fixing solution (45% methanol, 10% acetic acid) and detected by Coomassie Brilliant Blue (0.25% CBB R-250

in fixing solution) staining for 30 minutes to 1 hr. The gels were destained in the fixing solution and scanned using Gel Doc (Laemmli UK, 1970).

A.2.12. Protein Estimation

The amount of protein in a sample was estimated by the Bradford Dye-Binding Protein Assay using BSA as the standard (Bradford, 1976). The micro assay was done in Spectramax-M2 and the concentration of protein is determined. This method involves the measurement of the binding of Coomassie Brilliant Blue G-250 dye to proteins and measurement of the absorbance of the dye-protein complex at 595 nm. Its advantages as a protein concentration measuring procedure over other chemical methods as it is fast and only requires one reagent, its color intensity is relatively stable over a period of an hour so more flexibility in reading the absorbance is provided and it has fewer substances that interfere with it 5 µl of protein sample was mixed 250 µl of the working solution in a microtitre plate and incubated at 37°C for 25 min. The absorbance was taken at 595 nm. BSA of known concentration was used as standard.

Estimation of the protein concentration by measuring absorbance at A280

Concentration of the protein was determined by measuring the absorbance at 280 nm.

Protein concentration (mg/ml) = Absorbance(280)

The O.D. of a protein was calculated by equation:

O.D.= (No. of tryptophan residue * 5500) + (No. of Tyrosine residue *1490) + (No. of Cystine *125)

Absorbance at 280 nm of protein is mainly due to the tryptophan and tyrosine amino acid residues present, and partly depends upon cystine (double bond). Above formula works very well for the protein containing tryptophan residue, but sensitivity goes down for protein without it. Concentration of the proteins which do not contain tryptophan can also be accurately measured by refractometer. Refractometer actually measures the degree to which light ray bends, when it crosses the interface between two medium of different densities. Protein present in the solution increases its density, which causes pronounced bending of light ray.

REFERENCES

REFERENCES

Achkar, J.M., and Fries, B.C. (2010). Candida infections of the genitourinary tract. Clin. Microbiol. Rev. *23*, 253–273.

Alvarez, F.J., and Konopka, J.B. (2007). Identification of an N-acetylglucosamine transporter that mediates hyphal induction in Candida albicans. Mol. Biol. Cell *18*, 965–975.

Anderson, J.M., and Soll, D.R. (1987). Unique phenotype of opaque cells in the white-opaque transition of candida albicans. J. Bacteriol. *169*, 5579–5588.

Baek, Y.-U., Li, M., and Davis, D.A. (2008). Candida albicans ferric reductases are differentially regulated in response to distinct forms of iron limitation by the Rim101 and CBF transcription factors. Eukaryot. Cell *7*, 1168–1179.

Bahn, Y.S., Staab, J., and Sundstrom, P. (2003). Increased high-affinity phosphodiesterase PDE2 gene expression in germ tubes counteracts CAP1-dependent synthesis of cyclic AMP, limits hypha production and promotes virulence of Candida albicans. Mol. Microbiol. *50*, 391–409.

Berman, J., and Sudbery, P.E. (2002). Candida Albicans: a molecular revolution built on lessons from budding yeast. Nat. Rev. Genet. *3*, 918–930.

Biswas, S., Van Dijck, P., and Datta, A. (2007). Environmental sensing and signal transduction pathways regulating morphopathogenic determinants of Candida albicans. Microbiol. Mol. Biol. Rev. *71*, 348–376.

Braun, B.R., and Johnson, A.D. (2000). TUP1, CPH1 and EFG1 make independent contributions to filamentation in Candida albicans. Genetics *155*, 57–67.

Braun, B.R., Head, W.S., Wang, M.X., and Johnson, A.D. (2000). Identification and characterization of TUP1-regulated genes in Candida albicans. Genetics *156*, 31–44.

Brown, D.H., Giusani, A.D., Chen, X., and Kumamoto, C.A. (1999). Filamentous growth of Candida albicans in response to physical environmental cues and its regulation by the unique CZF1 gene. Mol. Microbiol. *34*, 651–662.

Butler, G., Rasmussen, M.D., Lin, M.F., Santos, M.A.S., Sakthikumar, S., Munro, C.A., Rheinbay, E., Grabherr, M., Forche, A., Reedy, J.L., et al. (2009). Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature *459*, 657–662.

Campion, E.W., Kullberg, B.J., and Arendrup, M.C. (2015). Invasive Candidiasis. N. Engl. J. Med. *373*, 1445–1456.

Cannon, R.D., Holmes, A.R., Mason, A.B., and Monk, B.C. (1995). Oral Candida: Clearance, Colonization, or Candidiasis? J. Dent. Res. *74*, 1152–1161.

Cannon, R.D., Lamping, E., Holmes, A.R., Niimi, K., Tanabe, K., Niimi, M., and Monk, B.C. (2007). Candida albicans drug resistance - Another way to cope with stress. Microbiology *153*, 3211–3217.

Chakravarti, A., Camp, K., McNabb, D.S., and Pinto, I. (2017). The iron-dependent regulation of the candida albicans oxidative stress response by the CCAAT-binding factor. PLoS One *12*.

Chen, C., Pande, K., French, S.D., Tuch, B.B., and Noble, S.M. (2011). An iron homeostasis regulatory circuit with reciprocal roles in candida albicans commensalism and pathogenesis. Cell Host Microbe *10*, 118–135.

Chindamporn, A., Nakagawa, Y., Mizuguchi, I., Chibana, H., Doi, M., and Tanaka, K. (1998). Repetitive sequences (RPSs) in the chromosomes of Candida albicans are sandwiched between two novel stretches, HOK and RB2, common to each chromosome. Microbiology *144*, 849–857.

Chu, W.S., Magee, B.B., and Magee, P.T. (1993). Construction of an SfiI macrorestriction map of the Candida albicans genome. J. Bacteriol. *175*, 6637–6651.

Delbrück, S., and Ernst, J.F. (1993). Morphogenesis-independent regulation of actin transcript levels in the pathogenic yeast Candida albicans. Mol. Microbiol. *10*, 859–866.

DeRisi, J.L. (1997). Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. Science (80-.). *278*, 680–686.

Du, H., Guan, G., Li, X., Gulati, M., Tao, L., Cao, C., Johnson, A.D., Nobile, C.J., and Huang, G. (2015). N-Acetylglucosamine-induced cell death in candida albicans and its implications for adaptive mechanisms of nutrient sensing in yeasts. MBio *6*.

Dunwell, J.M., Purvis, A., and Khuri, S. (2004). Cupins: the most functionally diverse protein superfamily? Phytochemistry *65*, 7–17.

Fitzpatrick, D. a, O'Gaora, P., Byrne, K.P., and Butler, G. (2010). Analysis of gene evolution and metabolic pathways using the Candida Gene Order Browser. BMC

Genomics *11*, 290.

Forsburg, S.L., and Guarente, L. (1989a). Communication between mitochondria and the nucleus in regulation of cytochrome genes in the yeast Saccharomyces cerevisiae. Annu. Rev. Cell Biol. *5*, 153–180.

Forsburg, S.L., and Guarente, L. (1989b). Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. Genes Dev. *3*, 1166–1178.

Forsburg, S.L., and Guarente, L. (1989c). Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. Genes Dev. *3*, 1166–1178.

Ghosh, S., Rao, K.H., Bhavesh, N.S., Das, G., Dwivedi, V.P., and Datta, A. (2014). N-acetylglucosamine (GlcNAc)-inducible gene GIG2 Is a novel component of GlcNAc metabolism in Candida albicans. Eukaryot. Cell *13*, 66–76.

Gola, S., Martin, R., Walther, A., Dünkler, A., and Wendland, J. (2003). New modules for PCR-based gene targeting in Candida albicans: Rapid and efficient gene targeting using 100 bp of flanking homology region. Yeast *20*, 1339–1347.

Hegg, E.L., and Que, L. (1997). The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. Eur. J. Biochem. *250*, 625–629.

Hooft van Huijsduijnen, R., Li, X.Y., Black, D., Matthes, H., Benoist, C., and Mathis, D. (1990). Co-evolution from yeast to mouse: cDNA cloning of the two NF-Y (CP-1/CBF) subunits. EMBO J. *9*, 3119–3127.

Hsu, P.C., Chao, C.C., Yang, C.Y., Ye, Y.L., Liu, F.C., Chuang, Y.J., and Lan, C.Y. (2013). Diverse hap43-independent functions of the Candida albicans CCAAT-binding complex. Eukaryot. Cell *12*, 804–815.

Hull, C.M., Raisner, R.M., and Johnson, A.D. (2000). Evidence for mating of the "asexual" yeast Candida albicans in a mammalian host. Science *289*, 307–310.

Ishola, O.A., Ting, S.Y., Tabana, Y.M., Ahmed, M.A., Yunus, M.A., Mohamed, R., Lung Than, L.T., and Sandai, D. (2016). The Role of Isocitrate Lyase (ICL1) in the Metabolic Adaptation of Candida albicans Biofilms. Jundishapur J. Microbiol. *9*, e38031.

Janbon, G., Sherman, F., and Rustchenko, E. (1998). Monosomy of a specific

chromosome determines L-sorbose utilization: a novel regulatory mechanism in Candida albicans. Proc. Natl. Acad. Sci. U. S. A. *95*, 5150–5155.

Jaroszewski, L., Li, Z., Krishna, S.S., Bakolitsa, C., Wooley, J., Deacon, A.M., Wilson, I.A., and Godzik, A. (2009). Exploration of uncharted regions of the protein universe. PLoS Biol. *7*.

Johnson, A. (2003). The biology of mating in Candida albicans. Nat. Rev. Microbiol. *1*, 106–116.

Johnson, D.C., Cano, K.E., Kroger, E.C., and McNabb, D.S. (2005). Novel regulatory function for the CCAAT-binding factor in Candida albicans. Eukaryot. Cell *4*, 1662–1676.

Konopka, J.B. (2012). N-acetylglucosamine (GlcNAc) functions in cell signaling. Scientifica (Cairo). *2012*, 631–632.

Kumar, M.J., Jamaluddin, M.S., Natarajan, K., Kaur, D., and Datta, A. (2000). The inducible N-acetylglucosamine catabolic pathway gene cluster in Candida albicans: discrete N-acetylglucosamine-inducible factors interact at the promoter of NAG1. Proc. Natl. Acad. Sci. U. S. A. *97*, 14218–14223.

Laskowski, R. a., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Lim, C.S.Y., Rosli, R., Seow, H.F., and Chong, P.P. (2012). Candida and invasive candidiasis: Back to basics. Eur. J. Clin. Microbiol. Infect. Dis. *31*, 21–31.

Magee, B.B., and Magee, P.T. (1987). Electrophoretic karyotypes and chromosome numbers in Candida species. J. Gen. Microbiol. *133*, 425–430.

Martinez, S., and Hausinger, R.P. (2015). Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-dependent Oxygenases. J. Biol. Chem. *290*, 20702–20711.

McNabb, D.S., Xing, Y., and Guarente, L. (1995). Cloning of yeast HAP5: a novel subunit of a heterotrimeric complex required for CCAAT binding. Genes Dev. *9*, 47–58.

McNabb, D.S., Tseng, K.A., and Guarente, L. (1997). The Saccharomyces cerevisiae Hap5p homolog from fission yeast reveals two conserved domains that are essential for assembly of heterotetrameric CCAAT-binding factor. Mol. Cell. Biol. *17*, 7008–7018.

Milewski, S., Gabriel, I., and Olchowy, J. (2006). Enzymes of UDP-GlcNAc biosynthesis in yeast. Yeast *23*, 1–14.

Mishra, P.K., Baum, M., and Carbon, J. (2007). Centromere size and position in Candida albicans are evolutionarily conserved independent of DNA sequence heterogeneity. Mol. Genet. Genomics *278*, 455–465.

Naglik, J.R., Moyes, D.L., Wächtler, B., and Hube, B. (2011). Candida albicans interactions with epithelial cells and mucosal immunity. Microbes Infect. *13*, 963–976.

Naseem, S., and Konopka, J.B. (2015). N-acetylglucosamine Regulates Virulence Properties in Microbial Pathogens. PLoS Pathog. *11*, e1004947.

Naseem, S., Gunasekera, A., Araya, E., and Konopka, J.B. (2011). N-acetylglucosamine (GIcNAc) induction of hyphal morphogenesis and transcriptional responses in Candida albicans are not dependent on its metabolism. J. Biol. Chem. *286*, 28671–28680.

Naseem, S., Parrino, S.M., Buenten, D.M., and Konopka, J.B. (2012). Novel roles for GlcNAc in cell signaling. Commun. Integr. Biol. *5*, 156–159.

Naseem, S., Araya, E., and Konopka, J.B. (2015). Hyphal growth in Candida albicans does not require induction of hyphal-specific gene expression. Mol. Biol. Cell *26*, 1174–1187.

Naseem, S., Min, K., Spitzer, D., Gardin, J., and Konopka, J.B. (2017). Regulation of Hyphal Growth and N-Acetylglucosamine Catabolism by Two Transcription Factors in Candida albicans. Genetics *206*.

Nasution, A. (2013). Virulence Factor and Pathogenicity of Candida albicans in Oral Candidiasis. World Dent. *4*, 267–271.

Natarajan, K., and Datta, A. (1993). Molecular cloning and analysis of the NAG1 cDNA coding for glucosamine-6-phosphate deaminase from Candida albicans. J. Biol. Chem. *268*, 9206–9214.

Noble, S.M., and Johnson, A.D. (2005). Strains and strategies for large-scale gene deletion studies of the diploid human fungal pathogen Candida albicans. Eukaryot. Cell *4*, 298–309.

Noble, S.M., Gianetti, B.A., and Witchley, J.N. (2016). Candida albicans cell-type switching and functional plasticity in the mammalian host. Nat. Rev. Microbiol. *15*, 96–108.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pande, K., Chen, C., and Noble, S.M. (2013). Passage through the mammalian gut triggers a phenotypic switch that promotes Candida albicans commensalism. Nat Genet *45*, 1088–1091.

Perepnikhatka, V., Fischer, F.J., Niimi, M., Baker, R.A., Cannon, R.D., Wang, Y.K., Sherman, F., and Rustchenko, E. (1999). Specific chromosome alterations in fluconazole-resistant mutants of Candida albicans. J. Bacteriol. *181*, 4041–4049.

Pérez, J.C., Kumamoto, C.A., and Johnson, A.D. (2013). Candida albicans Commensalism and Pathogenicity Are Intertwined Traits Directed by a Tightly Knit Transcriptional Regulatory Circuit. PLoS Biol. *11*.

Pfaller, M.A. (1996). Nosocomial candidiasis: emerging species, reservoirs, and modes of transmission. Clin. Infect. Dis. *22 Suppl 2*, S89-94.

Pfaller, M.A., and Diekema, D.J. (2007). Epidemiology of invasive candidiasis: a persistent public health problem. Clin. Microbiol. Rev. *20*, 133–163.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor laboratory press.

Santos, M.A.S., and Tuite, M.F. (1995). The CUG codon is decoded in vivo as serine and not leucine in Candida albicans. Nucleic Acids Res. *23*, 1481–1486.

Sanyal, K., Baum, M., and Carbon, J. (2004). Centromeric DNA sequences in the pathogenic yeast Candida albicans are all different and unique. Proc. Natl. Acad. Sci. *101*, 11374–11379.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative CT method. Nat. Protoc. *3*, 1101–1108.

Sengupta, M., and Datta, A. (2003). Two membrane proteins located in the Nag regulon of Candida albicans confer multidrug resistance. Biochem. Biophys. Res. Commun. *301*, 1099–1108.

Simonetti, N., Strippoli, V., and Cassone, A. (1974). Yeast-mycelial conversion induced by N-acetyl-D-glucosamine in Candida albicans.

Singh, B., and Datta, A. (1979). Induction of N-acetylglucosamine-catabolic pathway in spheroplasts of Candida albicans. Biochem. J. *178*, 427–431.

Singh, P., Ghosh, S., and Datta, A. (2001). Attenuation of virulence and changes in

morphology in Candida albicans by disruption of the N-acetylglucosamine catabolic pathway. Infect. Immun. *69*, 7898–7903.

Singh, R.P., Prasad, H.K., Sinha, I., Agarwal, N., and Natarajan, K. (2011). Cap2-HAP complex is a critical transcriptional regulator that has dual but contrasting roles in regulation of iron homeostasis in Candida albicans. J. Biol. Chem. *286*, 25154–25170.

Singh, S.M., Steinberg-Neifach, O., Mian, I.S., and Lue, N.F. (2002). Analysis of telomerase in Candida albicans: Potential role in telomere end protection. Eukaryot. Cell *1*, 967–977.

Soll, D.R. (2002). Candida and Candidiasis. ASM Press. Washingt. 123–142.

Su, C., Lu, Y., and Liu, H. (2016). N-acetylglucosamine sensing by a GCN5-related N-acetyltransferase induces transcription via chromatin histone acetylation in fungi. Nat. Commun. *7*, 12916.

Sudbery, P.E. (2011). Growth of Candida albicans hyphae. Nat. Rev. Microbiol. *9*, 737–748.

Thiébaut, A., Delaveau, T., Benchouaia, M., Boeri, J., Garcia, M., Lelandais, G., and Devaux, F. (2017). The CCAAT-Binding Complex Controls Respiratory Gene Expression and Iron Homeostasis in Candida Glabrata. Sci. Rep. *7*, 3531.

Vylkova, S., Carman, A.J., Danhof, H.A., Collette, J.R., Zhou, H., and Lorenz, M.C. (2011). The fungal pathogen candida albicans autoinduces hyphal morphogenesis by raising extracellular pH. MBio *2*.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. *67*, 235–242.

Wisplinghoff, H., Ebbers, J., Geurtz, L., Stefanik, D., Major, Y., Edmond, M.B., Wenzel, R.P., and Seifert, H. (2014). Nosocomial bloodstream infections due to Candida spp. in the USA: Species distribution, clinical features and antifungal susceptibilities. Int. J. Antimicrob. Agents *43*, 78–81.

Yamada-Okabe, T., and Yamada-Okabe, H. (2002). Characterization of the CaNAG3, CaNAG4, and CaNAG6 genes of the pathogenic fungus Candida albicans: possible involvement of these genes in the susceptibilities of cytotoxic agents. FEMS Microbiol. Lett. *212*, 15–21.

Yamada-Okabe, T., Sakamori, Y., Mio, T., and Yamada-Okabe, H. (2001). Identification and characterization of the genes for N-acetylglucosamine kinase and N-acetylglucosamine-phosphate deacetylase in the pathogenic fungus Candida albicans. Eur. J. Biochem. *268*, 2498–2505.

Yan, L., Yang, C., and Tang, J. (2013). Disruption of the intestinal mucosal barrier in Candida albicans infections. Microbiol. Res. *168*, 389–395.