FROM YEAST TO HUMAN: UNRAVELING SPHINGOLIPID METABOLISM THROUGH MACROSCOPIC AND MICROSCOPIC ANALYSES

by

Fatma Betül Kavun Özbayraktar

B.S., Chemical Engineering, Boğaziçi University, 2003

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Chemical Engineering
Boğaziçi University
2011

To My Mother, Fatma

ACKNOWLEDGEMENTS

I would like to express my gratitude to my thesis advisor, Prof. Kutlu Ö. Ülgen, for her everlasting support, encouragement, help and kind attitude throughout my studies. I would like to thank the members of my thesis committee, Prof. Z. İlsen Önsan, Prof. Hale Saybaşılı, Prof. Dilek Kazan and Assoc. Prof. Z. Petek Çakar for accepting to be a member of my jury, critical reading of my thesis and their comments. I would like to acknowledge Dr. Judy Monthie-Doyum for proofreading of my review paper which was published in Biotechnology Journal in 2009.

This research was supported by Boğaziçi University Research Fund through projects 06HA504D. The financial support of Turkish Scientific Research Council (TÜBITAK) through project 104M362 is gratefully acknowledged. I also acknowledge the doctoral scholarship provided by TÜBITAK-BIDEB 2211. BÜVAK and TÜBITAK are acknowledged due to financial contributions regarding conferences.

I would like to thank Dr. Tunahan Çakır for his help regarding the codes in stoichiometric modelling presented in Chapter 4. I also thank Senem Tiveci regarding our collaboration in coding selective permissibility algorithm presented in Chapter 5 and endless scientific chats. I would like to thank to all academic and administrative staff of Boğaziçi University, Department of Chemical Engineering for their help and support throughout my education.

I would like to thank to my parents, especially to my mother, Fatma, for her never-ending support, patience and love. I believe she is the angel from heaven and I dedicate this dissertation to her. I would like to thank to my uncle, Mehmet, due to being a brother beyond being an uncle to me. I would like to deeply thank to Nahit for her kindness and support. I would like to thank to my husband, Hikmet, for his support, patience and love. Susam, you should know that your mother loves you, cares you and thinks about you.

My special thanks are due to all current and former members of Biosystems Engineering Research Group, especially Pınar, Tunahan, Yalçın, Saliha, Senem, Elif,

Ayça, Yasemen, Esra and Ceyda for providing a very-nice and very-friendly environment in our KB440, including scientific chats and gossips.

My 11 years in Boğaziçi University made me older and wiser and I will always remember these years with joy and full of friendship.

ABSTRACT

FROM YEAST TO HUMAN: UNRAVELING SPHINGOLIPID METABOLISM THROUGH MACROSCOPIC AND MICROSCOPIC ANALYSES

Sphingolipids are both structural and regulatory components of the cell, where they control processes decisive in cell's fate. The pharmacological manipulation of the sphingolipid metabolism in cancer therapeutics necessitates the detailed understanding of the pathway. Different methodologies based on dynamic, stoichiometric and protein interaction information are used to identify potential drug target enzymes among sphingolipid pathway components. All the enzymes in sphingolipid pathway were ranked according to their roles in controlling the metabolic network using metabolic control analysis. The physiologically connected reactions were identified by metabolic pathway analysis. The mathematical tools' efficiency for drug target identification performed in this study is validated by clinically available drugs. The first elaborate metabolic model of *Sacchoramyces cerevisiae* sphingolipid metabolism was reconstructed *in silico*. The model considers five different states of sphingolipid hydroxylation, rendering it unique among other models. We propose that *IPT1, GDA1, CSG* and *AUR1* gene deletions may be novel candidates of drug targets for cancer therapy according to the results of flux balance and variability analyses coupled with robustness analysis. Constructing the protein-protein interaction network of sphingolipids in *S. cerevisiae* enabled us to understand the topological properties of the network of 1591 nodes. 14 novel interactions are predicted using a newly developed integrated methodology employing sequence and structure based computational interaction prediction tools, orthology, expression profiles, co-localization and Gene Ontology terms. The function annotation of 11 uncharacterized proteins of the network is performed using another newly developed multi-dimensional hybrid method which combines the results from modules and neighbors. *LCB5* and *IPT1* are identified as potential drug targets for cancer, possessing the topological properties of an ideal drug target.

ÖZET

MAYADAN İNSANA: MAKROSKOPİK VE MİKROSKOPİK ANALİZLERLE SFİNGOLİPİD METABOLİZMASININ İNCELENMESİ

Sfingolipidler hem hücre yapısında hem de regülasyonda rol alırlar ve hücrenin akıbetini belirleyecek süreçleri kontrol ederler. Farmakolojik olarak sfingolipid metabolizmasının manipülasyonu kanser tedavisinde etkili bir metot olarak karşımıza çıkmaktadır ve bu da sfingolipid yolizinin detaylı olarak anlaşılmasını gerektirmektedir. Bu çalışmada dinamik, stokiyometrik temellere ve protein etkileşimlerine dayalı farklı metotlar kullanılarak, sfingolipid metabolizmasında rol alan olası ilaç hedefleri tespit edilmiştir. Metabolik kontrol analizi ile sfingolipid yolizinde rol alan enzimlerin metabolik ağ üzerindeki kontrol etkileri incelenmiştir. Fizyolojik olarak birbirine bağlı reaksiyonlar metabolik yolizi analizi kullanılarak tespit edilmiştir. Her iki analiz metodunun da ilaç hedefi tayinindeki verimliliği klinik olarak incelenen hedefler sayesinde sınanmıştır. Model organizma olarak seçilen *S. cerevisiae* mayasında sfingolipid metabolizmasının ayrıntılı metabolik modeli *in siliko* olarak literatürde ilk defa oluşturulmuştur. Burada kurulan modeli diğer çalışmalardan ayıran en önemli özellik beş farklı sfingolipid hidroksilleme aşamasının da dahil edilmesidir. Bu modelin akı denge, değişkenlik ve dayanıklılık analizleri kullanılarak incelenmesi sonucu *IPT1, GDA1, CSG* ve *AUR1* gen delesyonları kanser tedavisinde faydalı olabilecek olası ilaç hedefleri olarak belirlenmiştir. Mayada ve insanda sfingolipidler için protein-protein etkileşim ağları ilk defa bu çalışmada oluşturulmuştur. Protein dizilimine ve yapısına dayalı hesapsal etkileşim öngören araçlar, homoloji, gen ekspresyon profilleri, ortak yerleşim bilgisi ve gen ontolojisi terimleri bütünleştirilerek yeni bir metot geliştirilmiş ve mayada 14 yeni protein etkileşimi öngörülerek ağyapı etkileşim modelinde kullanılmıştır. Modül ve ağ komşusu bilgilerini, bir araya getiren çok yönlü hibrit bir metot daha geliştirilerek mayada 11 bilinmeyen proteine fonksiyon ataması gerçekleştirilmiştir. Topolojik özellikler açısından ideal ilaç hedefi özellikleri taşıyan *LCB5* ve *IPT1* genleri olası ilaç hedefi olarak önerilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3PS	3-phosphoserine
4-HPR	N-4-hydroxyphenyl retinamide, fenretinide
AAA1	Gene name
Aaa1p	Protein product of gene *AAA1*
ACBP	Acyl-Coenzyme A binding protein
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
B13	Ceramide analog
BID	BH3-interacting domain death agonist
BST	Biochemical systems theory
C16-CoA	C16 fatty acid
C26-CoA	C26 fatty acid, very long chain fatty acid
CCC	Concentration control coefficient
CDP	Cytidine diphosphate
CDP-E	Cytidine diphosphate ethanolamine
CEF	Control effective flux
Cer	Ceramide
CMP	Cytidine monophosphate
CO2	Carbon dioxide
CoA	Coenzyme A
CTP	Cytidine triphosphate
CVD	Cardiovascular disease
DAG	Diacylglycerol
DHCer	Dihydroceramide
DHS	Dihydrosphingosine
DHS-1-P	Dihydrosphingosine 1 phosphate
DP	Diphosphate
EFM	Elementary flux mode

EP	Ethanolamine phosphate
EPC	Edge percolated component
ER	Endoplasmic reticulum
FADD	Fas-associated death domain protein
FBA	Flux balance analysis
FCC	Flux control coefficient
FFAT	Two phenylalanines in an acidic tract
FVA	Flux variability analysis
G3P	Glycerol-3-phosphate
GD	Gaucher disease
GDP	Guanosine diphosphate
GM4	N-acetylneuraminyl-galactosylceramide
GMA	Generalized mass action
GMP	Guanosine monophosphate
GO	Gene ontolology
GPI	Glycosylphosphatidylinositol
H2O	Water
Hexco	Hexacosanoate
HPalmA	2-hydroxy-hexadecanal
I1P	Inositol-1-phosphate
IPC	Inositol phosphorylceramide
ITP	Inositol 1, 4, 5-triphosphate
KDHS	3-ketodihydrosphingosine
LCB	Long chain base
M(IP)2C	Mannose diinositol phosphorylceramide
Mal-CoA	Malonyl-coenzyme A
MCA	Metabolic control analysis
MCS	Membrane contact site
MIPC	Mannose inositol phosphorylceramide
ML	Metachromatic Leukodystrophy
MPA	Metabolic pathway analysis
NADP	Nicotinamide adenosine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate

NP	Niemann-Pick
O2	Oxygen
P	Phosphate
PA	Phosphatidic acid
PalCoA	Palmitoyl Coenzyme A
PalmA	Palmitaldehyde
PE	Phosphatidyl ethanolamine
PH	Pleckstrin homology
PHCer	Phytoceramide
PHS	Phytosphingosine
PHS-1-P	Phytosphingosine 1 phosphate
PI	Phosphatidyl inositol
PI45BP	Phosphoinositol 4, 5-biphosphate
PIN	Protein interaction network
pm	Plasma membrane
PSC833	Valspodar
PtdIns-4-P	Phosphatidylinositol-4-monophosphate
SA	Selection algorithm
SAP	Sphingolipid activator protein
SH3	Src homology-3
SPA	Selective permissibility algorithm
SPT	Serine palmitoyl transferase
START	Steroidogenic acute regulatory protein-related lipid transfer
TNF	Tumour necrosis factor
TNFR	Tumour necrosis factor receptor
TRADD	TNFR-associated death domain protein
VAP	Vesicle associated membrane protein-associated protein

1. INTRODUCTION

Sphingolipids constitute a biologically active lipid class that is significantly important from both structural and regulatory aspects. The manipulation of sphingolipid metabolism is currently studied as a novel strategy for cancer therapy. The basics of this therapeutic approach lie in the regulation property of sphingolipids on cellular processes important in a cell's fate, such as cell proliferation, apoptosis, cell cycle arrest, senescence and inflammation. Furhermore, the mutations in the enzymes catalyzing some specific reactions in the sphingolipid metabolism cause mortal lysosomal storage diseases, like Fabry, Gaucher, Niemann-Pick, Farber, Krabbe and Metachromatic Leukodystrophy. Therefore, the alteration of the sphingolipid metabolic pathway determines the choice between life and death. Understanding the sphingolipid metabolism and regulation is significant for the development of new therapeutic approaches for all sphingolipid-related diseases, as well as for cancer. An important feature of the sphingolipid metabolic pathway is the compartmentalization into endoplasmic reticulum, Golgi apparatus, lysosome and plasma membrane, and this compartmentalization makes the transport of sphingolipids critical for proper functioning. This research focuses on the structures, metabolic pathways, localization, transport mechanisms and diseases of sphingolipids in yeast (*Saccharomyces cerevisiae)* and human, and provides the latest comprehensive information on sphingolipid research. Computational systems biology approaches are utilized to construct models of the system under investigation and the newly constructed models are examined from macroscopic and microscopic aspects.

The ultimate goal of this study is to gain a detailed understanding of the sphingolipid metabolism and regulation, and use this understanding to identify critically important enzymes and proteins involved in sphingolipid pathways. The results from dynamic and stoichiometric approaches along with interactome studies are used to propose potential drug targets for cancer therapy. This research also gives the opportunity to compare the efficacy of computational systems biology tools in drug target identification process. The first elaborate stoichiometric model of yeast sphingolipid metabolism including hydroxylation reactions of sphingolipids and the first comprehensive sphingolipid protein-protein interaction network of both yeast and human are presented in this research. Two

brand new integrated methods for function annotation and protein-protein interaction prediction are also developed within the framework of this research.

The second chapter of the thesis starts with a detailed discussion of the motivation of studying sphingolipids. The information on sphingolipid structure and types is followed by the explanation of sphingolipid metabolism in yeast. The enzymes functioning in sphingolipid metabolism and their encoding genes are discussed in detail in this section. *De novo* sphingolipid metabolism starts in endoplasmic reticulum and continues in the Golgi apparatus and also in other compartments of the cell. The continuity of the metabolism requires the translocation of synthesized sphingolipids from one compartment to the other in a controlled and ordered way. The transport mechanisms (vesicular and non-vesicular) governing the sphingolipid metabolism of yeast are explained in this chapter. The human sphingolipid metabolism is also discussed. Diseases involving sphingolipid metabolism are divided into four sections as storage diseases, cancer, neurodegenerative diseases and cardiovascular pathology. Chapter 2 also includes up-to-date occurrence frequencies of sphingolipid diseases in Turkey.

In the third chapter drug target identification obtained by metabolic control and pathway analyses are discussed for sphingolipid metabolism. First, these two computational systems biology tools are explained in methods section and then the design of selection algorithms utilized in drug target identification process are mentioned and schematically shown in detail. The comparison of the results with the clinically known drug targets gave us the opportunity to compare the efficacy of the two computational systems biology tools in drug target identification. Finally an ultimate drug target set is achieved by assessing the results from both methods. The biological meanings of the candidate drug targets proposed in this research are also given at the end of the chapter.

The first section of the fourth chapter of the thesis focuses on the first elaborate stoichiometric network reconstruction of yeast sphingolipid metabolism. Five states of hydroxylation of ceramide and long chain bases make the newly constructed model unique among the others in literature. Flux balance analysis, flux variability analysis, minimal cut set calculations and robustness analysis along with the reactions of the stoichiometric model are given in the methods subsection. The validation of the newly constructed model

is discussed in the following subsections. The deletion mutants with experimentally verified phenotypic behaviors related to calcium sensitivity are simulated for the purpose of additional validation. The deletion mutant simulations are given in detail and novel drug targets are proposed under the title of *in silico* drug actions. The relationship between the viability issue of the cell and complex sphingolipid content of the cell are discussed in detail. The second section of this chapter summarizes the attempt of constructing the stoichiometric model of human sphingolipid metabolism. The lack of information and challenges encountered are discussed.

Chapter 5 gives the details of protein-protein interaction network of both yeast and human sphingolipids. Two scale-free, small-world and modular protein-protein interaction networks of 1591 proteins for yeast and 3097 proteins for human are constructed and their topological properties are investigated. Although the lack of interactome data in human limited the analysis of the human network, the yeast network is examined in detail and consequently used in drug target identification and function annotation. Two novel integrated methods for protein-protein interaction prediction and function annotation are developed within the framework of this research.

Finally, all the results of critical importance and the contributions in both methodology and modelling are summarized in “Conclusions and Recommendations” chapter (Chapter 6) along with the future directions.

2. MOLECULAR FACETS OF SPHINGOLIPIDS: MEDIATORS OF DISEASES

Sphingolipids are signaling molecules that are functional in vital cellular processes, like differentiation, migration, apoptosis, cell proliferation, cell cycle arrest, senescence and inflammation. Due to their regulatory property on cellular processes that determine a cell's fate, sphingolipids have the role of an arbiter in cancer biology and therapy (Dickson, 2008; Ogretmen, 2006). The knowledge of the modulation of the sphingolipid metabolism inducing apoptosis in cancer cells is expected to improve the effectiveness of cancer therapeutics. Other than cancer, sphingolipids also have roles in various human diseases, such as sphingolipidoses, diabetes, cardiovascular disease, microbial infections, Alzheimer's disease and other neurological disorders and immune dysfunctions. In order to understand the mechanisms underlying the above-mentioned diseases and consequently to determine the *in vivo* drug targets, the biology of sphingolipids and the related metabolic and signaling pathways need to be investigated in detail.

Since the study of sphingolipids is a young field of systems biology, simpler model organisms, such as *Saccharomyces cerevisiae*, have mostly been employed to unravel the relationship between sphingolipids and cellular processes leading to diseases (Dickson, 2008). The incentive for this model organism selection is that all the genes encoding the enzymes in sphingolipid metabolism are known in *S. cerevisiae*, and more importantly, most of the *S. cerevisiae* enzymes functioning in the sphingolipid pathway have homologues or orthologues in mammalian species. This homology is expected to help illuminate the sphingolipid metabolism in humans (Dickson, 2008; Hannun and Obeid, 2002; Modrak *et al.*, 2006). It is believed that the complete understanding of apoptosis and cancer and many other critical cellular processes coupled with major diseases can be achieved after the elucidation of the sphingolipid metabolism. Hence, in this chapter, the structures, metabolic pathways, localization, transport mechanisms and diseases of sphingolipids in *S. cerevisiae* and humans are discussed (Kavun-Ozbayraktar and Ulgen, 2009).

2.1. Sphingolipid Structure

The structural components of sphingolipids are classified into three parts: a sphingoid long-chain base, which forms the backbone of the sphingolipids, a fatty acid, and a polar head group, as seen in Figure 2.1. The type of sphingolipid and to which organism it belongs are determined from the identity of the long-chain base, fatty acid, and the polar hydrophilic head group. For example, 26 carbon long saturated fatty acid is an identifier of sphingolipids in *S. cerevisiae*. The fatty acid is amide linked to the carbon-2 of the sphingoid long-chain base, and the head group is linked to the carbon-1 position from the OH-group. A complex sphingolipid is named according to the attached head group. If a sugar is added to the long chain base, glycosphingolipids are formed in mammalian cells. When phosphorylcholine is added to the association of fatty acid and long-chain base, sphingomyelin is produced. Though there is usually one main type of fatty acid in *S. cerevisiae*, there are more than 20 types of fatty acids used in mammalian sphingolipid synthesis. The number of carbons, the degree of hydroxylation and the degree of saturation make the difference among them (Dickson, 2008; Futerman and Riezman, 2005).

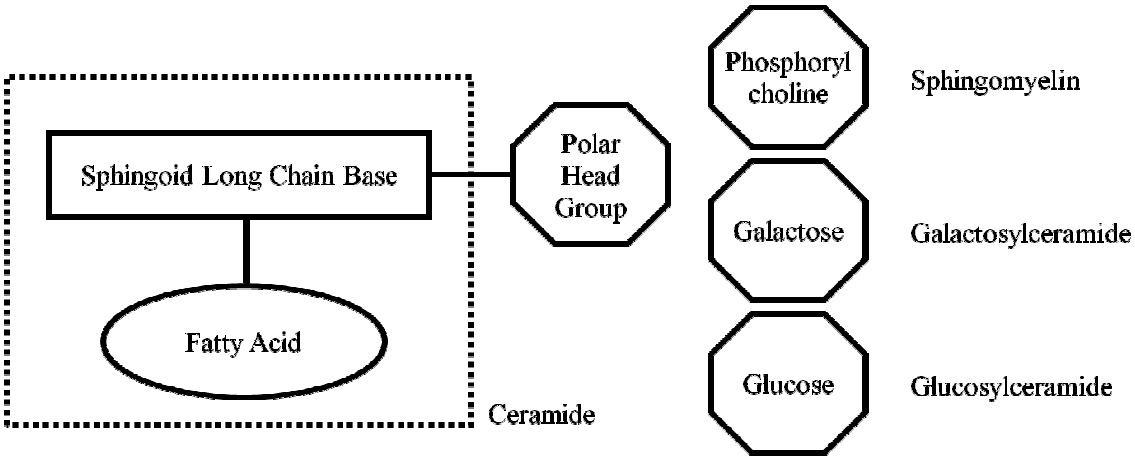


Figure 2.1. Sphingoid long chain base forms the back bone of all sphingolipids. Ceramide is formed by the addition of a fatty acid to the backbone. Addition of polar head groups leads to the formation of other complex sphingolipids

In *S. cerevisiae*, sphingoid long chain bases are linear alkanes with two hydroxyl groups on carbon-1 and carbon-3 positions and one amino group between them. Dihydrosphingosine and phytosphingosine are the two main types of sphingolipids. The

difference between these two long-chain bases is an extra hydroxyl group at the carbon-4 position of phytosphingosine. A 4, 5-double bond distinguishes the *S. cerevisiae* and mammalian long-chain bases. The same double bond is also the sole difference between dihydroceramide and ceramide in mammalian cells (Dickson and Lester, 1999).

One of the long-chain bases that is not found in *S. cerevisiae*, but is mammalian cells, is sphingosine. Most of the sphingosine comes from the degradation of complex sphingolipids/ceramide, whereas dihydrosphingosine is synthesized through the sphingolipid biosynthetic pathway reactions (Futerman and Riezman, 2005). The reason for the absence of sphingosine in *S. cerevisiae* cells may be the fact that there is no *in vivo* evidence of decomposition of complex sphingolipids, inositol phosphorylceramide (IPC), mannose inositol phosphorylceramide (MIPC) and mannose diinositol phosphorylceramide ($M(IP)_2C$), back to ceramide (Dickson and Lester, 1999). Although the double-bonded sphingosine is the essential long-chain base in mammalian sphingolipid synthesis, dihydrosphingosine and phytosphingosine also exist in small amounts in mammalian cells (Dickson and Lester, 2002).

2.2. Sphingolipid Metabolism of *S. cerevisiae*

S. cerevisiae sphingolipid metabolism is schematically depicted in Figure 2.2 (Alvarez-Vasquez *et al.*, 2004; Le Stunff *et al.*, 2002; MIPS; SGD; Sims *et al.*, 2004). In *S. cerevisiae* cells, sphingolipid synthesis is initiated by the condensation of serine with palmitoyl-CoA via serine palmitoyl transferase (SPT) enzyme (EC 2.3.1.50) in the endoplasmic reticulum. The main product of this conversion is 3-ketodihydrosphingosine (KDHS), and carbon dioxide (CO_2) and coenzyme A (CoA) are also released at the end of this reaction. The SPT enzyme active in *S. cerevisiae* has three sub-units: Lcb1p and Lcb2p are common in all organisms' serine palmitoyl transferases, whereas small hydrophobic sub-unit Tsc3p is specific to *S. cerevisiae*. The genes encoding these proteins in *S. cerevisiae* are YMR296C, YDR062W and YBR058C-A, respectively (Dickson, 2008). The active site of the SPT enzyme is located on the endoplasmic reticulum membrane facing the cytoplasm (Futerman and Riezman, 2005).

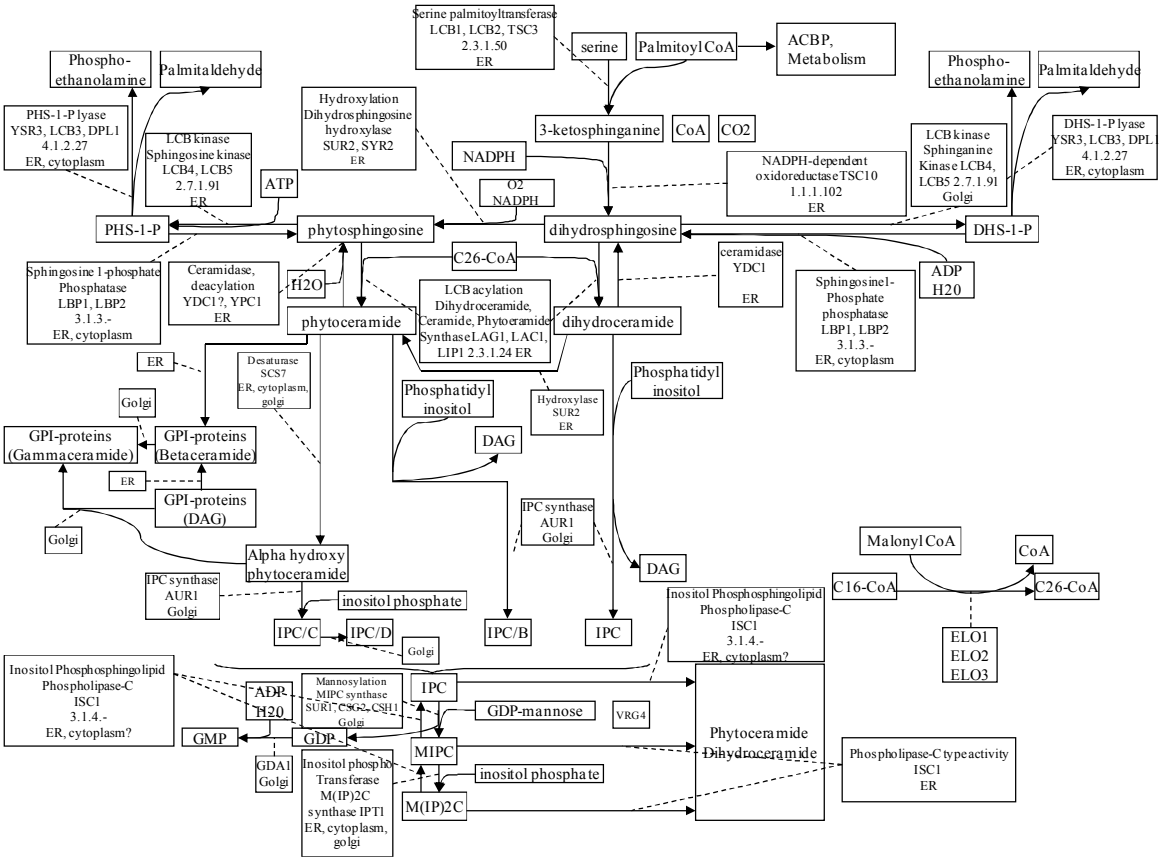


Figure 2.2. The arrows show the reactions and the dashed lines show the enzyme information. The enzyme names, gene names encoding the enzymes, Enzyme Commission numbers and probable locations are given in boxes connected to the dashed lines

The second reaction in the sphingolipid pathway is the reduction of KDHS to synthesize the first long-chain base, sphinganine (DHS). The enzyme responsible for this conversion is KDHS reductase encoded by Tsc10p present on the endoplasmic reticulum membrane. Similar to most of the sphingolipid metabolism enzymes, the active site of the reductase enzyme faces the cytoplasm (Dickson, 2008; Futerman and Riezman, 2005).

C_{26} fatty acid is added to the carbon-2 position of DHS by an amide bond to form dihydroceramide. Ceramide synthase (EC 2.3.1.24) is the enzyme responsible for the acylation of DHS on the endoplasmic reticulum membrane. There are two genes encoding ceramide synthase, which are *LAC1* and *LAG1*. For the functionality of ceramide synthase, Lip1p, a *S. cerevisiae* specific protein, is also necessary. The genes encoding these proteins in *S. cerevisiae* are YKL008C, YHL003C and YMR298W, respectively (Dickson, 2008; Futerman and Riezman, 2005).

Two types of long-chain bases (namely, dihydrosphingosine and phytosphingosine) exist in *S. cerevisiae* sphingolipid metabolism, where the one with only two hydroxyl groups (DHS) can be hydroxylated once more to obtain phytosphingosine, but the reverse reaction is not possible. This can be an indicator that the *S. cerevisiae* cell favors phytosphingosine (PHS) rather than dihydrosphingosine. 16, 18 or 20 carbon-DHS and 18 or 20 carbon-PHS exist in *S. cerevisiae.* The hydroxylation reaction attaches the extra hydroxyl group to the carbon-4 position of dihydrosphingosine. A similar metabolic conversion is valid for the transformation of dihydroceramide to phytoceramide. Both conversions take place in the membrane of endoplasmic reticulum. The hydroxylase enzyme is encoded by *SUR2* gene and has a sterol desaturase domain. It should be noted that there are two alternative paths to synthesize phytoceramide; either from the hydroxylation of dihydroceramide or directly from the acylation of PHS via the ceramide synthase enzyme (Dickson, 2008; Sims *et al.*, 2004).

The phosphorylation of dihydrosphingosine and phytosphingosine to their phosphates occurs by the LCB kinase enzyme. Although the kinase is encoded by *LCB4* and *LCB5* genes (the two isoforms), most of the sphingosine kinase activity is provided by Lcb4p. The kinase enzyme is found on the membranes of endoplasmic reticulum, the Golgi and endosome. In the double deletion mutant of *LCB4* and *LCB5* the long chain bases (DHS and PHS) were reported to accumulate. This finding indicates that the accumulation of dihydrosphingosine and phytosphingosine does not inhibit the growth in *S. cerevisiae* cells (Cowart and Obeid, 2007; Futerman and Riezman, 2005; Hait *et al.*, 2006; Sims *et al.*, 2004). The dihydrosphingosine-1-phosphate and phytosphingosine-1-phosphate can follow two alternative paths; either the dephosphorylation back to long-chain bases occurs via sphingosine-1-phosphate phosphatase enzyme in the endoplasmic reticulum, or the phosphates can breakdown into fatty aldehydes and ethanolamine phosphate. The latter path ends up in another metabolism, whereas the dephosphorylation process brings back the long chain base consumed by the phosphorylation step. Although the site of the phosphorylation step is unsettled, the location can be deduced from the site of dephosphorylation and the breakdown of long-chain base phosphates. The absence of transport of long-chain base phosphates from the Golgi apparatus to endoplasmic reticulum may indicate that the phosphorylation step also takes place in the endoplasmic reticulum. According to several literature sources, the phosphorylation and then dephosphorylation

are required for ceramide synthesis (Cowart and Obeid, 2007; Futerman and Riezman, 2005; Hait *et al.*, 2006; Sims *et al.*, 2004).

According to numerous authors, complex sphingolipids in *S. cerevisiae* are derived from phytoceramide, but not from dihydroceramide. In fact, with respect to this statement, phytoceramide in *S. cerevisiae* cell replaces the ceramide in mammalian cells, being the biologically active species (Perry and Ridgway, 2005; SGD). Dihydroceramide was also reported to be converted to complex sphingolipids in Vallee and Riezman (2005). At this stage of the metabolic process, ceramides are transported from the endoplasmic reticulum to the Golgi apparatus, where the addition of the polar head groups takes place to form the final complex sphingolipids that are abundantly found in plasma membrane. A phosphorylinositol head group is attached to the ceramide backbone to form the first complex sphingolipid of *S. cerevisiae*, IPC, by IPC synthase in conversion from both phytoceramide and dihydroceramide in the Golgi apparatus. The IPC synthase enzyme is embedded in the membrane with the N-terminal region of the enzyme within the Golgi lumen and the C-terminal region facing the cytosol, AUR1 is the gene encoding this enzyme in *S. cerevisiae* (Dickson, 2008; Perry and Ridgway, 2005; Sims *et al.*, 2004). The alternative path for the IPC synthesis consists of two metabolic steps. First, phytoceramide is desaturated by Scs7p to produce alpha hydroxyl phytoceramide, and then this intermediate metabolite is converted to IPC by Aur1p. Reggiori and Conzelmann (1998) made the discrimination between the IPCs synthesized from these two paths, and the complex sphingolipid formed in the absence of *SCS7* gene is called IPC/B. The main IPC synthesis branch was stated to be via alpha hydroxyl phytoceramide (Reggiori and Conzelmann, 1998; Sims *et al.*, 2004).

The second type of complex sphingolipid is MIPC that is synthesized in the Golgi apparatus from IPC. The conversion involves the addition of a mannose from guanosine diphosphate (GDP)-mannose to the previously added inositol in IPC. The mannosylation takes place via MIPC synthase that is the combined effect of the gene products of *SUR1, CSG2* and *CSH1*. According to Dickson (2008), there are two types of MIPC synthases: one is the complex of Sur1p and Csg2p, and the other is formed by the gene products of *CSG2* and *CSH1*. The end product of *S. cerevisiae* sphingolipid metabolism is $M(IP)_2C$, which is the most abundant complex sphingolipid. It is synthesized in the Golgi apparatus

by the addition of the second inositol phosphate to MIPC by the inositol phosphotransferase enzyme (Ipt1p). The decomposition of these complex sphingolipids back to dihydroceramide and phytoceramide is performed by inositol phosphosphingolipid phospholipase C, Isc1p, which has a phospholipase-C type activity (Dickson, 2008; Sims *et al.*, 2004).

2.3. Transport Mechanisms in *S. cerevisiae*

Dihydroceramide and phytoceramide synthesized in the endoplasmic reticulum should be transported to the Golgi apparatus where various head groups are attached to form complex sphingolipids. There are two transport mechanisms active in *S. cerevisiae* cells, which are vesicular ceramide transport and adenosine triphosphate (ATP)-independent non-vesicular ceramide transport. The main difference between the transport mechanisms of *S. cerevisiae* and mammalian cells is that there is no transporter protein between endoplasmic reticulum and the Golgi in *S. cerevisiae* cells. The membranes of these organelles are required to touch each other in order to transfer ceramide from the site of synthesis to the Golgi. Ceramide is first located on the outer leaflet of the Golgi membrane and then transferred to the inner leaflet that is the luminal side. This transfer on the Golgi membrane is proposed to be performed by a translocator protein without energy requirement. The IPC synthase, MIPC synthase and inositolphosphotransferase enzymes catalyzing the complex sphingolipid synthesis reactions are functional in the lumen of the Golgi apparatus. The non-vesicular and vesicular transport mechanisms differ in their destination point. The non-vesicular transport mechanism brings ceramide to *medial*-Golgi for IPC synthesis, whereas the vesicular transport takes ceramide to *cis*-Golgi for other regulatory purposes. The complex sphingolipids are transported to their site of action that is the plasma membrane via vesicular transport (Dickson, 2008; Funato *et al.*, 2002; Sims *et al.*, 2004).

Another transport mechanism that is thought to take place in *S. cerevisiae* operates from the plasma membrane to the endoplasmic reticulum membrane. *S. cerevisiae* can take the long-chain bases and their phosphates from outside and utilize them in cellular sphingolipid synthesis. The exact mechanism for this transport process is not described in *S. cerevisiae* but a kind of a translocator is supposed to be functional, though none has

been identified yet. The transport reaction for the removal of long-chain base phosphates from the *S. cerevisiae* cell is catalyzed by Yor1p. Although it is possible to decompose the phosphates synthesized, this transport is an alternative way of avoiding the accumulation of phosphates in the cell (Funato *et al.*, 2002).

2.4. Mammalian Sphingolipid Metabolism

In this study, a generic pathway of mammalian sphingolipid metabolism is reconstructed by collecting the reactions from several papers, as well as from KEGG database (Figure 2.3) (BRENDA; KEGG; Ogretmen and Hannun, 2004). The sphingolipid pathways of *S. cerevisiae* and mammalian cells resemble each other up to the synthesis of ceramide. Instead of converting dihydroceramide into phytoceramide, mammalian cells add a double bond to dihydroceramide and obtain biologically active ceramide. One other distinguishing feature of the mammalian sphingolipid metabolism from that of the *S. cerevisiae* is the complex sphingolipids. Mammalian cells do not add inositol phosphate and mannose to ceramide but add hydrophilic head groups, such as galactose, phosphatidylcholine, and glucose. Ceramide is one of the precursors of complex sphingolipids in mammalian cells. Additionally, some mammalian cells, like those in skin, intestines and kidney, are known to contain phytosphingosine-based sphingolipids.

Ceramide is converted to sphingomyelin by the addition of phosphatidylcholine via sphingomyelin synthase enzyme (EC 2.7.8.3), mainly in the luminal surface of the Golgi membrane, but the enzymatic activity of the synthase is also observed in the plasma membrane. *SMS1* and *SMS2* genes are responsible for the synthesis of the sphingomyelin synthase enzyme. Sms1p is located in the *trans* Golgi region, whereas the location of Sms2p is uncertain (Hanada *et al.*, 2007). Sms2p can function as an isozyme in the Golgi apparatus, or apart from the *de novo* synthesis, it may convert the ceramide produced in the plasma membrane to sphingomyelin.

Galactosylceramide is another complex sphingolipid synthesized by the galactosyltransferase enzyme (EC 2.4.1.47) that attaches a galactose to the sphingoid long-chain base backbone. Glucosylceramide synthase (EC 2.4.1.80) converts ceramide to glucosylceramide mainly in the Golgi apparatus (*cis* Golgi region), and glucosylceramide

can be further metabolized to lactosylceramide by the lactosylceramide synthase enzyme. Besides the Golgi apparatus, endoplamic reticulum and microsome also contain the glucosylceramide synthase enzyme. The first conversion to glucosylceramide takes place on the cytosolic face of the Golgi membrane, whereas the second conversion to lactosylceramide takes place on the luminal face of the Golgi apparatus (BRENDA; Hanada *et al.*, 2007; Kolter and Sandhoff, 2006; Perry and Ridgway, 2005). The breakdown of complex sphingolipids back to ceramide also takes place in mammalian cells.

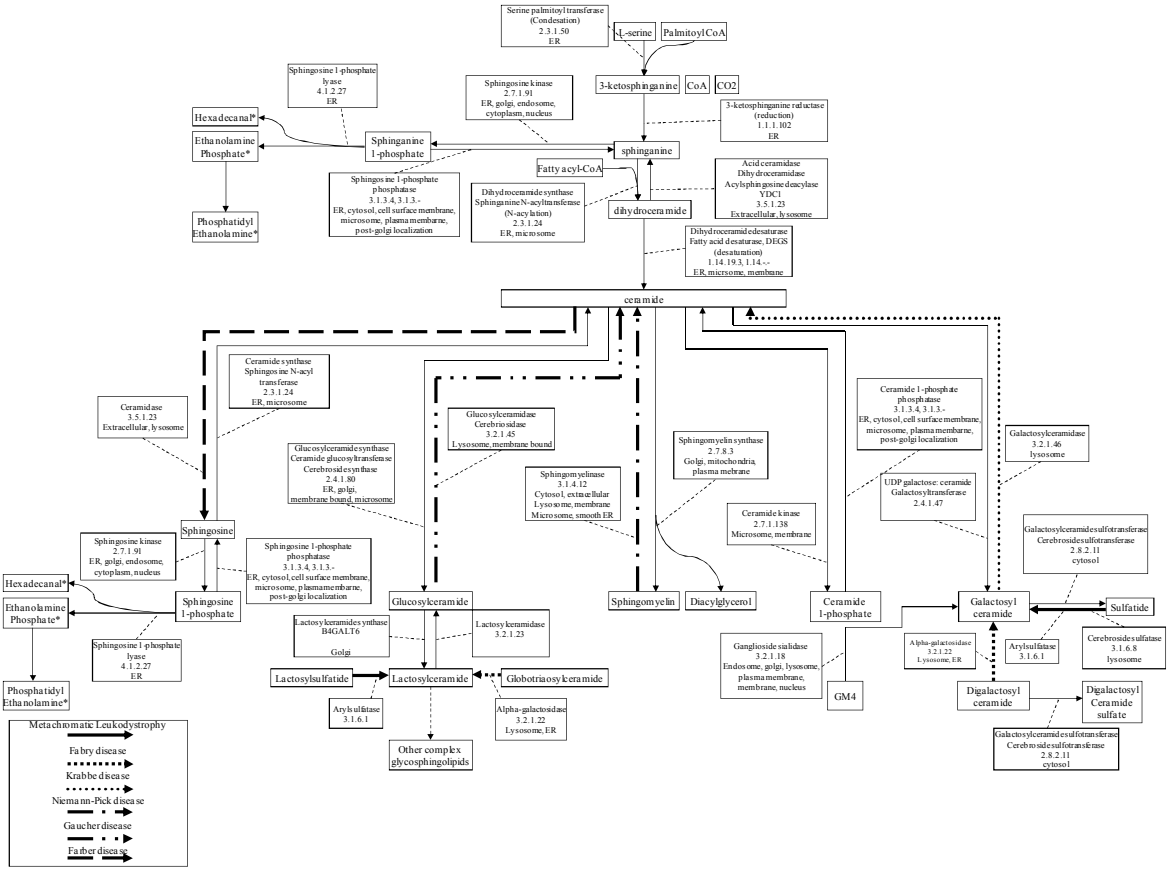


Figure 2.3. Human sphingolipid metabolism and sphingolipidoses. All the enzymes functioning in the pathway are indicated with their names, Enzyme Commission numbers and probable locations. The defective enzymes are highlighted with different arrows

Furthermore, ceramide can be broken down to sphingosine via ceramidase enzyme, which has numerous forms. *ASAH1, ASAH2, ASAH2B, ASAH2C, ASAH3* and *ASAH3L* are some of the genes encoding the enzyme in humans. The phosphorylation and

dephosphorylation of ceramide, sphingosine and dihydrosphingosine occur via kinase and phosphatase enzymes, respectively.

2.5. Transport Mechanisms in Mammalian Cells

The reactions of mammalian sphingolipid metabolism are scattered to various compartments of the cell. The intermediate metabolites should be transported from the site of the synthesis to the site where they will be metabolized, and the end products should be transferred to the location where they will function properly. So, it can be easily concluded that the transport mechanisms of sphingolipids are vital for the well-being of the cell. Another fact that indicates the importance and necessity of the sphingolipid transport is that sphingolipids are structural components of the membranes, and this lipid content is crucial for the well-being of the cell.

As in the case of *S. cerevisiae*, mammalian sphingolipid synthesis starts in the endoplasmic reticulum and continues in the same way until ceramide formation. Ceramide is a key metabolite having control effect on vital cellular processes, such as apoptosis. Since ceramide is a bio-effector in intracellular signaling pathways, its transport is very important. Besides the signaling functions, ceramide is necessary for complex sphingolipid synthesis in the Golgi apparatus. If ceramide travels within the cell from the endoplamic reticulum to the Golgi apparatus spontaneously by interbilayer movements, it may take days; so, the cell has two alternative transport mechanisms for ceramide, which are vesicular and non-vesicular transport, as it was the case in *S. cerevisiae* cells (Figure 2.4). Ceramide is first transported from the endoplasmic reticulum membrane to the cytosolic face of the Golgi apparatus membrane, and the second transport mechanism takes it to the luminal face of the Golgi apparatus, where it is metabolized to sphingomyelin. Since glucosylceramide formation takes place on the cytosolic face of the Golgi, another transport mechanism should play a role in the transfer of glucosylceramide to the luminal surface for its conversion to lactosylceramide (Perry and Ridgway, 2005).

Non-vesicular transport involves the attachment of ceramide to a carrier protein, whereas in the vesicular transport mechanism, ceramide is packed in a vesicle (COPII). Unlike the *S. cerevisiae* cell in mammals, the carrier transport protein that functions in

non-vesicular ceramide transport is identified to be CERT, and in some sources, the ceramide carried by CERT protein is reported to be used in sphingomyelin synthesis (Hanada *et al.*, 2007; Perry and Ridgway, 2005). CERT is a hydrophilic 68-kDa protein which has three functional domains: ceramide binding steroidogenic acute regulatory protein-related lipid transfer (START) domain, two phenylalanines in an acidic tract (FFAT) domain and pleckstrin homology (PH) domain. FFAT domain performs the attachment with VAP of endoplasmic reticulum (VAP is the part of the endoplasmic reticulum membrane). The START domain has ~210 amino acids on the carboxyl terminal of the protein, which can extract the firmly embedded ceramide from the endoplasmic reticulum membrane. The PH domain covers ~120 amino acids in the amino terminal of the protein, which directs the protein to the Golgi apparatus or the *trans* Golgi network. The transport by the CERT protein is based on the conformational changes of its domains; namely, domains are exposed and obscured in such an order that ceramide is carried from the endoplasmic reticulum to the Golgi apparatus.

When CERT is free from ceramide, its FFAT domain is exposed, and CERT is directed towards the endoplamic reticulum to get the ceramide to be transferred for sphingomyelin formation in the Golgi apparatus. Under this condition, ceramide is extracted from the endoplasmic reticulum and binds to the START domain of the transporter protein. This domain accomplishes the intermembrane transfer of ceramide from the endoplasmic reticulum to CERT. Perry and Ridgway (2005) reported that START domain can bind to d-erythro-ceramides with C_{14}, C_{16}, C_{18}, C_{20}, C_{16}-dihydroceramide, C_{16}-phytoceramide and short chain fluorescent ceramide analogs but not C_{22} and C_{24}, N-acyl chains. In fact, the polar residues at the ceramide binding domain, START, gives the information about which ceramide can bind to CERT. The attachment of ceramide to START domain damages the FFAT-VAP binding, so the loading of ceramide ends up in the concealing of FFAT domain and exposure of PH domain, which directs the ceramide-loaded CERT towards the Golgi apparatus through the cytosol. PH domain interacts with the Golgi membrane by phosphatidylinositol-4-monophosphate (PtdIns-4-P) and Arf1p. The cycle is completed when ceramide reaches the Golgi membrane and FFAT domain is exposed for the transportation of new ceramide molecules. Another alternative for the non-vesicular ceramide transport involves the use of the membrane contact sites (MCS) of endoplasmic reticulum and the Golgi apparatus. CERT binds both organelles at the same

time via MCS; the three domains of the transporter protein function simultaneously. FFAT domain binds to VAP of endoplasmic reticulum, ceramide binds to START domain, and finally, PH domain binds to the Golgi via PtdIns-4-P. Perry and Ridgway (2005) reported that an oxysterol-binding protein regulates ceramide transport by CERT and increases sphingomyelin synthesis. This statement also supported the hypothesis that ceramide transported by CERT is utilized in sphingomyelin synthesis. The half-time of the transport of ceramide by CERT is reduced from days to less than a minute (Perry and Ridgway, 2005; Hanada *et al.*, 2007).

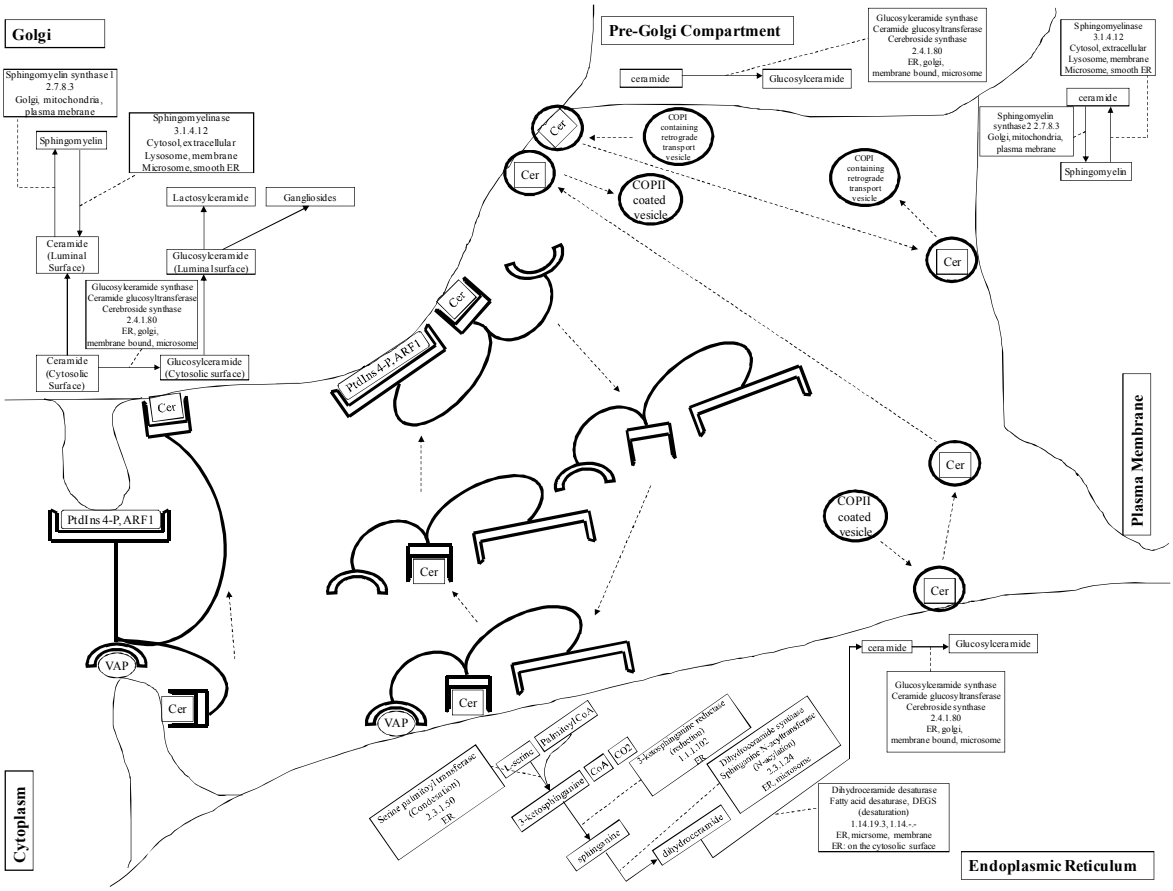


Figure 2.4. The first two mechanisms on the left show the non-vesicular transport of ceramide by CERT protein either on a membrane contact site or freely in the cytoplasm. The last two mechanisms on the right are the vesicular transport by COPII and COPI

Since ceramide transported by CERT is used in sphingomyelin synthesis, other transport mechanisms should exist for the synthesis of the remaining complex sphingolipids. The alternative transport mechanism involves COPII-mediated vesicular transport, though the exact molecular mechanism is not known yet. It is known that

sphingolipid synthesis starts in the endoplasmic reticulum, continues in the Golgi apparatus and finalizes in the plasma membrane, where they are enriched. So, as a conclusion, there must be a transport mechanism ending up in the plasma membrane. The only known sphingolipid transport to plasma membrane is by COPI vesicles directed from the Golgi apparatus. Furthermore, glycosphingolipids are specifically carried to the plasma membrane by vesicular exocytotic membrane flow (Kolter and Sandhoff, 2006; Perry and Ridgway, 2005; van Meer and Holthuis, 2000).

2.6. Diseases Involving Sphingolipid Metabolism

The genetic alterations in the enzymes of sphingolipid metabolism lead to the accumulation of various complex sphingolipids in the lysosome, causing lysosomal storage diseases that cover Fabry, Farber, Gaucher, Niemann-Pick, Krabbe, and Metachromatic Leukodystrophy. Sphingolipids and cancer relation is the most studied issue when diseases involving sphingolipids are considered, and apoptosis essentially plays the most important role. In cancer cells, bioactive sphingolipids, such as ceramide and sphingosine-1-phosphate, are utilized to induce apoptosis and cell proliferation, respectively. Apoptosis and sphingolipids cascade are also activated in Alzheimer's disease, due to the accumulation of amyloid beta protein. Angiogenesis is another cellular process regulated by sphingolipids, and this fact dictates an active role for sphingolipids in cardiovascular pathology (Kolter and Sandhoff, 2006; Ogretmen, 2006; Ogretmen and Hannun, 2004; Tardy *et al.*, 2004). An understanding of the sphingolipid component of these diseases is expected to lead to new approaches for their treatment.

2.6.1. Sphingolipid Storage Diseases (Sphingolipidoses)

Sphingolipidoses are rare but highly mortal, inherited, inborn diseases caused by modifications in sphingolipid metabolism, *i.e.*, by the pathological accumulation of sphingolipids. Hence, they are classified as storage diseases. It was suggested that lysosomal storage diseases are generally coupled with problematic cell growth and death (Tardy *et al.*, 2004). The defective part causing the sphingolipidoses occurs during the lysosomal degradation of sphingolipids, leading to the deposition of the non-degradable material in the lysosome. The mutation leading to a defect in the sphingolipid degradation

pathway can be either on an enzyme responsible for the hydrolysis of the complex sphingolipids or on a sphingolipid activator protein (SAP) functioning in the lysosomal degradation process (Marks and Pagano, 2002). The pathological accumulation of the sphingolipids that are not degraded due to mutant enzymes perturbs the cholesterol homeostasis and consequently changes the composition of the membrane leading to problematic sphingolipid transport. For example, the blockage of the Golgi targeting of lactosylceramide was observed in fibroblasts with sphingolipidosis (Pagano *et al.*, 2000; Pagano, 2003). It can thus be stated that there is a direct link between sphingolipid transport and sphingolipidoses.

The diseases are grouped and classified according to the sphingolipid that accumulates in the cell (Table 2.1 and Figure 2.3). The common indications of sphingolipidoses involve neurological symptoms, like dementia, major psychiatric illnesses, celebral palsy, developmental delay, rarefaction of the bones and several somatic symptoms, such as hepatomegaly and splenomegaly (Brady, 1967; Kolter and Sandhoff, 2006; Marks and Pagano, 2002; Pagano *et al.*, 2000; Pagano, 2003; Sandhoff, 1977; Tardy *et al.*, 2004).

Ozkara and Topcu (2004) studied the sphingolipidoses in Turkey among 2057 children with various neurological symptoms (300 with sphingolipidoses). The occurrence frequency of the sphingolipidoses in Turkey is 1:21 666 compared to the world's figure of 1: 7000-8000. The incidence of Krabbe disease is 1: 100 000, Fabry disease is 1: 6 500 000, Gaucher disease is 1: 224 137 and Metachromatic Leukodystrophy disease is 1: 69 892 in Turkey. When these numbers are compared with the figures of the world frequencies (Fabry disease 1: 40 000-117 000, Gaucher disease 1:50 000-200 000, and Metachromatic Leukodystrophy disease 1: 40 000-100 000), it can be concluded that Fabry disease is the most rarely seen one in Turkey with respect to the rest of the world. The most common sphingolipidoses are Metachromatic Leukodystrophy and Krabbe, especially seen in northern and middle parts of Turkey (Kolter and Sandhoff, 2006; Ozkara and Topcu, 2004).

Table 2.1. Sphingolipid related diseases, the deficient enzymes causing the sphingolipidoses, the accumulated metabolite as a result of the enzyme deficiency and the blocked reactions

Disease	Deficient Enzyme	Accumulated Sphingolipid	Blocked Reaction
Fabry Disease	Alpha-galactosidase	Globotriaosylceramide Digalactosylceramide	Globotriaosylceramide-> Lactosylceramide Digalactosylceramide -> Galactosylceramide
Gaucher Disease	Glucosyl ceramidase	Glucosylceramide	Glucosylceramide -> Ceramide
Metachromatic Leukodystrophy	Arylsulfatase	Sulfatide Lactosylsulfatide	Sulfatide -> Galactosylceramide Lactosylsulfatide -> Lactosylceramide
Krabbe Disease	Galactosyl ceramidase	Galactosylsphingosine	Galactosylceramide -> Ceramide
Niemann-Pick Disease	Sphingomyelinase	Sphingomyelin	Sphingomyelin -> Ceramide
Farber Disease	Ceramidase	Ceramide	Ceramide -> Sphingosine

2.6.1.1. Fabry Disease. Fabry disease is an X-linked disease, mainly seen in males and also in heterozygous females. It occurs as a result of the deficiency in the alpha-galactosidase enzyme, which is responsible for the lysosomal hydrolysis of globotriaosylceramide to lactosylceramide and digalactosylceramide to galactosylceramide (Figure 2.3). As a result of this enzyme deficiency, globotriaosylceramide and digalactosylceramide accumulate in the lysosomes of endothelial, perithelial and smooth muscle cells of blood vessels, and this accumulation further affects the heart, kidneys, eyes, cornea and the autonomous nervous system cells. The enzyme deficiency is caused by approximately 180 different mutations. The clinical symptoms of the Fabry disease are severe pain in the extremities, renal failure, vascular cutaneous lesions, hypohidrosis and corneal and lenticular opacities. Death is observed in patients of Fabry disease in their 4^{th} and 5^{th} decades of life due to the renal, cardiac and/or cerebral complications. An animal model of the Fabry disease was developed to be used in the evaluation of treatment strategies. Some examples of treatments used in Fabry disease are adeno-associated virus

mediated therapy, enzyme replacement therapy, chemical chaperone approach and deoxygalactonojirimycin treatment (Eng *et al.*, 2007; Kolter and Sandhoff, 2006; Lukacs *et al.*, 2007; Sandhoff, 1977).

2.6.1.2. Gaucher Disease. Gaucher disease (GD) is the deficiency of glucosylceramidase enzyme degrading glucosylceramide to ceramide in the lysosome (Figure 2.3). As a result of this most common sphingolipidose, glucosylceramide accumulation in the lysosomes of the reticulo-endothelial system, including the liver, spleen, bone-marrow and lung, is observed. As glucosylceramide accumulation continues, it is distributed to other parts of the cell. This is due to the corrupted transport of glucosylceramide to the lysosomes. Besides the lysosome, glucosylceramide is transferred to other sub-cellular compartments. Hein *et al.* (2007) observed an increase in the concentrations of several lipids other than glucosylceramide, such as ceramide, di-hexosylceramide and trihexosylceramide and phosphatidylglycerol in diseased cells. There are three types of Gaucher disease caused by approximately 200 different mutations. The first type is seen most extensively and has nonneuropathic effects. Type II is an acute form of Gaucher disease, which is rarely seen and panethnic. It affects the nervous system in early onset with a life expectancy of less than two years. Type III has neurological symptoms in later onset, and the disease in Type III patients develops more slowly than that of the more severe second type. The patients may live to the 4^{th} decade of their life. The deficiency of the enzyme causes severe skin disorders, whereas the partial deficient enzyme has no visible phenotype related to skin. The clinical symptoms are anemia, thrombocytopenia, enlargement of spleen (25-fold), liver and lymph nodes and bone damage. An animal model of Type II GD was developed in mice using a deficient glucosylceramidase enzyme. The therapeutic approaches involve enzyme therapy for Type I and bone-marrow transplantation (Hein *et al.*, 2007; Kolter and Sandhoff, 2006; Sandhoff, 1977). Brady (1967) suggested several potential immunotherapy approaches, such as the treatment of the spleen and the extention of the life-span of erythrocytes of GD patients.

2.6.1.3. Metachromatic Leukodystrophy. The deficiency causing Metachromatic Leukodystrophy (ML) involves arylsulfatase enzyme that cleaves sulfated lipids. Sixty different types of mutations exist leading to this enzyme deficiency. The arylsulfatase enzyme converts sulfatide to galactosylceramide and lactosylsulfatide to lactosylceramide

(Figure 2.3). The excessive deposition of sulfatides (structural subunits of myelin in the nervous system) has been observed in brain and kidney tissues. Metachromatic leukodystrophy is classified into three types, which are late infantile, juvenile and adult forms. Clinical symptoms of the infantile form are loss of speech, blindness, quadriparesis, peripheral neuropathy and seizures. The adult form of the disease is rare, with clinical symptoms of gait clumsiness, incontinence and optic atrophy. Although the animal model of the ML disease is available, there exists no casual treatment other than the enzyme replacement (Brady, 1967; Kolter and Sandhoff, 2006; Sandhoff, 1977).

2.6.1.4 Niemann-Pick Disease. The recessively inherited Niemann-Pick disease (NP) Types A and B are caused by the deficiency of acid sphingomyelinase enzyme catabolizing the conversion of sphingomyelin to ceramide in lysosome (Figure 2.3). The trafficking of cholesterol is modified as a consequence of the disease, and the accumulation of sphingomyelin is observed. NP-Type A is the infantile form with neurological implications, and NP-Type B is the juvenile form without any effect on the brain, and both types are panethnic disorders. Some of the clinical symptoms of Niemann-Pick disease of Type A are microcytic anemia, decreased platelet count, osteoporosis, brownish-yellow color of the skin and psychomotor retardation. The enlargement of spleen and/or liver and progressive pulmonary infiltration are the symptoms of Niemann-Pick Type B disease. Although a mouse model for the evaluation of different therapeutic strategies is available, there is still no specific treatment of this disease (Brady, 1967; Kolter and Sandhoff, 2006; Sandhoff, 1977). Ory (2000) reported that Types A and B have sphingomyelinase enzyme deficiency, whereas Types C and D have normal enzyme levels. Type C is named as a cholesterol storage disease, due to problematic trafficking of intracellular cholesterol (Ory, 2000).

2.6.1.5. Farber Disease. Farber disease is a rare disorder caused by the deficiency of ceramidase and subsequent accumulation of ceramide in the lysosome. The ceramidase enzyme catalyzes the breakdown of ceramide into sphingosine and fatty acid (Figure 2.3). The disease affects the skin, joints, liver, spleen, lung and heart. It should be noted that the accumulation of ceramide in the lysosome does not lead to apoptosis of the cell, contradicting the general knowledge of ceramide-causing apoptosis. Bone-marrow

transplantation is utilized in the treatment of the Farber disease (Kolter and Sandhoff, 2006; Sandhoff, 1977).

2.6.1.6. Krabbe Disease. The membrane-associated galactosylceramidase enzyme hydrolyzes galactosylceramide to ceramide (Figure 2.3). The lack of this enzyme activity leads to Krabbe disease, which is not classified as a storage disorder, since the accumulation of galactosylceramide is not observed, due to the rapid death of the cells. However, the defective enzyme leads to the accumulation of galactosylsphingosine, which is toxic and destroys oligodendroglial cells in the central nervous system. Krabbe disease has two forms, infantile and late onset. The clinical symptoms consist of blindness, deafness, spastic paralysis, convulsions and dementia and psychomotor retardation. The strategies employed in the treatment of Krabbe disease are bone-marrow transplantation and peripheral enzyme replacement (Kolter and Sandhoff, 2006; Sandhoff, 1977; Tardy *et al.*, 2004).

2.6.2. Therapeutic Approaches for Sphingolipidoses

In order to eliminate the negative effects of sphingolipidoses, numerous treatment approaches are being studied, as mentioned above. The basic idea underlying the majority of these treatment strategies is the depletion of the accumulated lipid by increasing the degradation capacity of the lysosome or decreasing the substrate influx to the lysosome. The mostly used methods can be summarized as gene therapy, enzyme replacement therapy, bone-marrow transplantation, cell-mediated "cross correction" and substrate deprivation. As an example of substrate deprivation, the enzyme synthesizing glucosylceramide (glucosylceramide synthase) is inhibited so that there is less glucosylceramide to be degraded, and the accumulation of this sphingolipid is thus eliminated (Marks and Pagano, 2002). Although bone-marrow transplantation is applied in Gaucher and mild forms of Krabbe diseases, several problems arise, such as the absence of donors, high mortality and high morbidity. In the case of the enzyme replacement therapy, the defective enzyme is replaced with its purified soluble form. This method has been used in the treatment of Gaucher disease, but no recovery in the neurological symptoms has been observed (Marks and Pagano, 2002). Moreover, Pagano *et al.* (2000 and 2003) suggested the overexpression of Rab7 and Rab9 proteins as a possible therapeutic

approach. Rab proteins are known to be functional in membrane trafficking processes, including vesicle formation. They regulate the several steps of transport processes. Rab7 protein serves in the transport from early to late endosome and from there to lysosome. Rab9 protein functions in another step of lipid transport, from late endosome to the Golgi apparatus. The gene manipulation of rab proteins corrects the lipid transport mechanism so that the accumulated lipid content is reduced (Kolter and Sandhoff, 2006; Marks and Pagano, 2002; Pagano, 2003; Soldati *et al.*, 1995).

2.6.3. Sphingolipids and Apoptosis – Cancer

Sphingolipids are bioactive molecules which are also functional in the regulation of apoptosis and tumor development. Apoptosis is the programmed cell death that is necessary for the control of the cell population and proper development of multicellular organisms. A problem in the regulation of apoptosis turns out to be degenerative diseases (at increased rate of cell death) or malignant diseases (at reduced apoptosis rate with respect to the normal cell). There are a variety of different mechanisms controlling apoptosis. Various types of proteases, including lysosomal ones, are involved in the process of apoptosis, and they are termed “caspases”. Sphingolipids, like ceramide, sphingosine and sphingosine-1-phosphate, are bio-effectors that have signaling functions in the regulation of apoptosis. Experimental evidence for this statement exists, such that the addition of exogenous ceramide and sphingosine to the cell induces apoptosis; increasing sphingomyelinase activity also increases the ceramide content of the cell and induces apoptosis. Moreover, Tardy *et al.* (2004) reported the neuronal cell death in Niemann-Pick Type C mice, as a result of increased cellular sphingosine levels (Ariga *et al.*, 1998; Hirabayashi *et al.*, 2006; Levade *et al.*, 2002; Segui *et al.*, 2006; Tardy *et al.*, 2004).

The delicate balance between cell growth and death decides the cell’s fate, and sphingolipids’ role is extremely important in the preservation of this balance. Ceramide, sphingosine and some gangliosides are termed “tumor-suppressor” lipids that promote caspase activation and consequently induce apoptosis, whereas sphingosine-1-phosphate, ceramide-1-phosphate, lactosylceramide and galactosylceramide are “tumor-promoting” lipids that induce cell proliferation. Beside these effects, sphingosine-1-phosphate induces tumor angiogenesis and inflammation that accelerate the tumor development and cancer

promotion. A novel cancer therapy approach considers the use of sphingolipid metabolism to regulate the disrupted balance between cell death and growth. The dynamic balance between sphingomyelin and ceramide, as well as between ceramide and sphingosine-1-phosphate (sphingolipid rheostat), should therefore be investigated carefully (Maceyka *et al.*, 2005; Ogretmen, 2006; Ogretmen and Hannun, 2004; Okada *et al.*, 2009; Reynolds *et al.*, 2004; Segui *et al.*, 2006). It has been noted that the frequency of cancer is increasing in industrialized, as well as in developing countries. It is foreseen by the National Cancer Institute in the USA that three families out of four will be affected by cancer in the forthcoming years. Although there are not any reliable data for cancer statistics in Turkey, it has been the second-leading cause of death, according to the Governmental Statistics Institute.

Lysosomal storage diseases may also affect apoptosis from various aspects. The absence of apoptosis inducers and/or suppresors, due to a deficiency in lysosomal degradation, influences the regulation of apoptosis. The activity of lysosomal proteases may change in both directions, due to the problematic lysosomal degradation of sphingolipids. Moreover, sphingolipidoses change the lipid content of the membrane leading to problems in intracellular transport mechanisms, and this may cause irregular signaling in the regulation of apoptosis (Tardy *et al.*, 2004).

2.6.4. Sphingolipids and Neurodegenerative Diseases

Neurodegenerative diseases like Alzheimer's and Parkinson's diseases are due to the conditions posing problems in the initiation of apoptosis. However, programmed cell death is required for the normal development of the central and peripheral nervous system. In this study, we focus on Alzheimer's disease, which is the most common one among the neurodegenerative disorders affecting 20 million individuals worldwide. It is the sixth-leading cause of death. As many as 5.2 million people in the United States are living with Alzheimer disease (Alzheimer's Association). It is estimated that there are about 350 000 people suffering from Alzheimer's disease in Turkey. The frequency of dementia and Alzheimer disease in Istanbul, Turkey is similar to Western World figures (Gurvit *et al.*, 2008).

He *et al.* (2008) reports that sphingolipid metabolism, specifically the sphingolipid rheostat, does not function properly in cells affected by Alzheimer's disease. Excessive loss of neurons, senile plaques, deposition of hyperphosphorylated tau proteins and accumulation of amyloid beta peptide are clinical implications of Alzheimer's disease. The improper functioning and loss of neurons further leads to problems in short-term memory and emotional disturbances. The accumulated amyloid beta peptide activates sphingomyelinase, and the ceramide content of the brain cell increases due to this elevated activity. Lee *et al.* (2004) stated the amyloid beta protein-sphingomyelinase-ceramide apoptosis cascade as a novel and important signaling mechanism. Another alteration in the sphingolipid metabolism of patients with Alzheimer's disease is the decrease in the activity of ceramidase enzyme leading to ceramide accumulation. Satoi *et al.* (2005) suggested a different mechanism for the increased ceramide levels in Alzheimer's disease, such as that there is a slight increase in *de novo* sphingolipid synthesis, but a dramatic decrease in the glucosylceramide synthesis leading to elevated ceramide levels. The authors assumed that it is the inhibition of glucosylceramide synthase enzyme that causes the increase in the ceramide content of cells in Alzheimer's disease (Satoi *et al.*, 2005). Besides the changes in ceramide content of the brain, the sphingosine-1-phoshate level decreases, which in turn decreases the cellular proliferation rate. So, it can be summarized that the depletion of sphingosine-1-phosphate and the accumulation of ceramide and sphingosine may be the causes of abnormal apoptosis in Alzheimer's disease that ends up in neuronal cell death. Okada *et al.* (2008) hypothesized that sphingosine-1-phosphate also functions in the regulation of neurotransmitter secretion. This fact can be utilized in further studies of therapeutic approaches for neuronal disorders (Ariga *et al.*, 1998; Okada *et al.*, 2009; Xiao *et al.*, 2002; He *et al.*, 2010).

2.6.5. Sphingolipids and Cardiovascular Pathology

Cardiovascular disease (CVD) is the number one cause of death in the world. Approximately 17 million people die from CVD every year, *i.e.*, 8.1 million men and 8.6 million women die from CVD each year (American Heart Association). Turkey is one of the countries in which the frequency of cardiovascular diseases is considerably high. It is expected that the death rates will increase by 1.8 fold in women and by 2.3 fold in men by the year 2030.

Sphingolipids also have functions in numerous cardiovascular processes. One of them regulated by sphingolipids is angiogenesis that is the formation of new capillary networks. Angiogenesis is involved in wound healing, uterine development, cancer, rheumatoid arthritis and diabetic retinopathy. The sphingolipid that has regulatory function in angiogenesis is sphingosine-1-phosphate that also affects the development and function of the heart. Sphingosine-1-phosphate has its effect on heart rate via the control of calcium metabolism and ionic currents of the sinoatrial node. Sphingosine-1-phosphate reduces heart rate, blood pressure and ventricular contraction, and these consequences protect the cardiac cells under ischemic conditions. On the other side, the sphingomyelin-ceramide signaling pathway is also activated under ischemic conditions, and the activated pathway leads to ischemic acute renal failure and postischemic neuronal cell death (Levade *et al.*, 2001). Other sphingolipid related changes that occur under ischemic conditions are ceramide accumulation and subsequent free radical deposition. In addition to these, ceramide accumulation is also observed in atherosclerotic lesions depositing low density lipoproteins that favor the narrowing and hardening of the arteries. That accumulated ceramide may lead to apoptosis, and this may result in heart failure and arrhythmia. Sphingosine-1-phosphate also takes part in atherosclerosis (the thickening of plaque on the wall of arteries) by stabilizing the plaque. Ceramide, lactosylceramide and sphingosine-1-phosphate promote the adhesion and migration of monocytes (Levade *et al.*, 2001; Saba and Hla, 2004).

2.7. Concluding Remarks

Sphingolipids are a class of lipids that are found abundantly in plasma membrane functioning as both structural units of the membrane and as bio-effectors in the regulation of numerous cellular processes having importance in a cell's fate. The biosynthetic and catabolic pathways of sphingolipids are scattered to many parts of the cell, and the integrity and continuity of metabolism depends on the transport mechanisms of these sphingolipids. Although sphingolipids are involved in many diseases, the relationship between sphingolipids and cancer has been excessively studied in recent years, and therapeutic agents have been developed, as will be discussed below. We here summarize the potential role of sphingolipids in diseases related to cancer, lysosomal storage, neurodegenerative and cardiovascular pathology. For the development of further

therapeutic approaches and the identification of potential drug targets, the sphingolipids metabolism should be understood from all aspects, including metabolic, signaling and transport processes. The advances in our understanding of these cellular events are expected to facilitate future drug discovery.

The idea that sphingolipids could be potential drug targets for novel therapeutic agents has found some applications in medicine. New synthetic inhibitors of specific enzymes involved in sphingolipid metabolism have been developed for the therapy of cancer (Zeidan and Hannun, 2007). These target enzymes of the human sphingolipid pathway are serine palmitoyl transferase; the first step in sphingolipid synthesis, ceramide synthase, ceramidase, sphingosine kinase, glucosylceramide synthase and 1-O-Acylceramide synthase (Reynolds *et al.*, 2004). A number of drugs like valspodar (PSC 833), ceramide analog (B13), N-4- hydroxyphenyl retinamide (4-HPR), daunorubicin, safingol, butyldeoxynojirimycin and phenoxodiol targeting the above listed enzymes are in pre-clinical or clinical development (Fracasso *et al.*, 2001; Reynolds *et al.*, 2004; Samsel *et al.*, 2004; Tiwari *et al.* 2006; Zeidan and Hannun, 2007). Changing ceramide levels in favor of cancer therapy seems to be a promising strategy if the developed drugs targeting the newly identified sphingolipid enzymes are selective toward malignant cells but not healthy cells. With some ceramidase inhibitor trials, this selectivity seems to be possible, due to the hypersensitivity of malignant cells to ceramide perturbation (Reynolds *et al.*, 2004). Another evidence for the significance of accumulation of ceramide in cancer cells is that cancer cells escape from apoptosis by degrading, catabolizing or converting ceramide (Zeidan and Hannun, 2007). In order to obtain more efficient drugs in cancer therapy, drugs targeting two or more enzymes may be developed, or two or more drugs may be utilized at the same time. There is clinical evidence of simultaneously using drugs inhibiting ceramide catabolism and generating ceramide (Reynolds *et al.*, 2004; Zeidan and Hannun, 2007).

The most difficult step in drug discovery is the decision of the drug target enzymes. As the clinical aspects of sphingolipids are still not completely known, the identification of new drug targets by employing computational tools is a significant requirement.

3. DRUG TARGET IDENTIFICATION IN SPHINGOLIPID METABOLISM BY COMPUTATIONAL SYSTEMS BIOLOGY TOOLS: METABOLIC CONTROL ANALYSIS AND METABOLIC PATHWAY ANALYSIS

Sphingolipids comprise a class of complex lipids which are abundantly found in cell membrane. These membrane lipids do not only function as structural components of the cell membrane but they possess important roles in signal transduction (as second messengers) and in regulatory pathways such as cell cycle arrest, apoptosis, senescence and differentiation. The investigation of the processes that are regulated by bioactive sphingolipid signaling molecules demonstrates that sphingolipids are fundamental to cancer pathogenesis and therapeutics (Ogretmen, 2006). Ceramide, an essential building block of sphingolipids, has been suggested to be a reasonable key metabolite in cancer therapy inducing antiproliferative and apoptotic responses. Many anticancer drugs are reported to lead to increased endogenous ceramide levels. Ceramide's influence on cell growth is dynamically balanced by sphingosine-1-phosphate and this balance is termed as "sphingolipid rheostat". In cancer cells, this dynamic balance is missing (Maceyka *et al.*, 2005; Ogretmen, 2006; Ogretmen and Hannun, 2004; Radin, 2001a; Radin, 2001b; Radin, 2003; Reynolds *et al.*, 2004; Robert and Lester, 1999).

A novel cancer therapy approach resides on the adjustment of the sphingolipid metabolism to accumulate ceramide and to decrease sphingosine-1-phosphate based on the fact of ceramide being a "tumor-suppressor lipid" and sphingosine-1-phosphate being a "tumor-promoting lipid". The drugs using this fact are expected to resemble the antiproliferative and apoptotic responses in the cell by altering the sphingolipid levels. The advantages of such changes are not only in cancer therapy but also, enhancement of drug action and chemoprevention (Maceyka *et al.*, 2005; Ogretmen, 2006; Ogretmen and Hannun, 2004; Reynolds *et al.*, 2004).

There are several newly developed synthetic inhibitors of specific enzymes involved in sphingolipid metabolism (Zeidan and Hannun, 2007). The target enzymes of human

sphingolipid pathway are serine palmitoyl transferase, ceramide synthase, ceramidase, sphingosine kinase, glucosylceramide synthase, 1-O-acylceramide synthase. A number of drugs like N-4-hydroxyphenyl retinamide (4-HPR), valspodar (PSC 833), daunorubicin, ceramide analog (B13), safingol, butyldeoxynojirimycin and phenoxodiol targeting the above listed enzymes are in pre-clinical or clinical development (Fracasso *et al.*, 2001; Reynolds *et al.*, 2004; Samsel *et al.*, 2004; Tiwari *et al.*, 2006; Zeidan and Hannun, 2007). Changing ceramide levels in favor of cancer therapy seems to be a promising strategy if the developed drugs targeting the newly identified sphingolipid enzymes are selective toward malignant cells but not healthy cells. With some ceramidase inhibitor trials this selectivity seems to be possible due to the hypersensitivity of malignant cells to ceramide perturbation (Reynolds *et al.*, 2004). Another evidence for the significance of accumulation of ceramide in cancer cells is that cancer cells escape from apoptosis by degrading, catabolizing or converting ceramide (Zeidan and Hannun, 2007). Experimentalists observed elevated levels of acid ceramidase expression in human prostate cancer cell lines (Seelan *et al.*, 2000). The manipulation of these cellular processes gives the chance of developing new anti-cancer therapeutics and indeed, forms the basics of the novel cancer therapy method (Ogretmen, 2006).

The most difficult step in drug discovery is the decision of the drug target enzymes. Since clinical aspects of sphingolipids are still lacking information, the identification of drug targets by employing computational tools is a significant requirement. Metabolic control analysis is a powerful mathematical tool to solve this puzzle of drug target identification via considering the role of individual components within a metabolic network (Cascante *et al.*, 2002). As mentioned previously, sphingolipids regulate processes that are vital in terms of cell's fate. These processes are all very important for cancer initiation, progression and treatment (Ogretmen and Hannun, 2004). The identification of critical enzymes within sphingolipid metabolism facilitates the link between sphingolipids and cancer.

In the present study, the key points of sphingolipid metabolism in terms of ceramide and sphingosine-1-phosphate are investigated using two different mathematical frameworks, metabolic control analysis and metabolic pathway analysis. Metabolic control analysis is based on dynamic data whereas metabolic pathway analysis uses the

stoichiometry of the system. These two systems biology tools reveal the important metabolic steps of the sphingolipid pathway and using this information novel drug target enzymes can be recognized and potential drug target enzymes in human pathway can be proposed. Although the *S. cerevisiae* and human sphingolipid pathways are similar to each other, one should admit that *S. cerevisiae* sphingolipid pathway consists of only few complex sphingolipids when compared with that of mammalian cells/human. Thus, this study can be considered as an attempt to bridge the basic sphingolipid research to potential medical applications in cancer therapy (Kavun-Ozbayraktar and Ulgen, 2010).

3.1. Computational Methods

In order to gain a better understanding of sphingolipid metabolism, *S. cerevisiae* is taken as the model system. The availability of the detailed and complete enzymatic and genetic data of *S. cerevisiae* makes the organism eligible for computational works that can give reliable first guesses for human studies. The conservation of sphingolipid metabolic and regulatory pathways across species has been used in the identification of the corresponding mammalian orthologues (Cowart and Hannun, 2005; Dickson and Lester, 2002; Obeid *et al.*, 2002; Petranovic and Nielsen, 2008). In the present study, the protein sequence homologies of the enzymes that are responsible for the sphingolipid pathway activity were cross-checked against human (*Homo Sapiens*) proteome via NCBI (National Center for Biotechnology Information) protein blast utility. Almost all the enzymes involved in the sphingolipid reactions of the model system have a homolog in human proteome with a reasonable and exceptable E-value, which shows the similarity of the two protein sequences. Also the functional domains of the yeast and human sphingolipid homolog proteins were compared and almost complete overlap is observed. The *de novo* sphingolipid metabolism of yeast exactly matches to that of human with the exception of the double bonded ceramide in human which is missed in yeast cells. The complex sphingolipids of yeast and human differ but the potential drug targets proposed in the current study correspond to mainly *de novo* sphingolipid synthesis.

3.1.1. Sphingolipid Metabolic Pathway

This model system consists of three parts. The main part is the sphingolipid metabolism and the two auxiliary pathways are the fatty acid metabolism and phospholipid metabolism. They integrate sphingolipid metabolism to other cellular processes. The fatty acid metabolism reactions used in the model system are steps that correspond to the synthesis and elongation of fatty acids. The very long chain fatty acids (C_{26}-CoA) are employed as precursors in the synthesis of dihydroceramide (DHCer) and phytoceramide (PHCer). Phospholipid metabolism is also another important branch point at the initial step of sphingolipid, *i.e.* at palmitoyl-coenzyme A (Pal-CoA) level. Synthesis of phosphatidyl inositol (PI) is included in the model system due to the fact that; PI is a precursor metabolite for the synthesis of complex sphingolipid, inositol phosphoryl ceramide (IPC) from both dihydroceramide and phytoceramide.

The sphingolipid metabolic pathway (Table 3.1) for the dynamic model is taken from Alvarez-Vasquez *et al.* (2005) (Alvarez-Vasquez *et al.*, 2005) and explained below. The first reaction of the *de novo* sphingolipid synthesis is the condensation of serine and Pal-CoA via the enzyme serine palmitoyl transferase, r_{11}, and takes place in the endoplasmic reticulum. This reaction produces the precursor metabolite 3-ketodihydrosphingosine (KDHS) which is required for the synthesis of the first long chain base, dihydrosphingosine (DHS). KDHS is rapidly reduced to the long chain base, r_{13}, catalyzed by 3-ketodihydrosphingosine reductase in the endoplasmic reticulum. The condensation reaction can be named as the initializing step of *de novo* sphingolipid synthesis in *S. cerevisiae* (Cowart and Hannun, 2005; Obeid *et al.*, 2002; Robert and Lester, 1999).

After the formation of the sphingoid long chain bases in *S. cerevisiae*, there are two alternative paths that the long chain bases can follow: acylation or phosphorylation. Dihydrosphingosine (DHS) and phytosphingosine (PHS) are acylated to DHCer, r_{23}, and PHCer, r_{24}, respectively by ceramide synthase and the reactions take place in endoplasmic reticulum (Alvarez-Vasquez *et al.*, 2005; Obeid *et al.*, 2002; SGD). The enzyme responsible for the phosphorylation of the long chain bases to long chain base phosphates is sphingoid base kinase which catalyzes the synthesis of dihydrosphingosine-1-phosphate (DHS-P) and phytosphingosine-1-phosphate (PHS-P), r_{15} and r_{20}, respectively. The reverse

reactions, dephosphorylations, r_{14} and r_{19}, are also possible by the enzyme sphingoid-1-phosphate phosphatase in endoplasmic reticulum of *S. cerevisiae*. The phosphorylation and dephosphorylation steps are needed for the synthesis of sphingolipids. The phosphorylated long chain bases can be decomposed by the enzyme sphingosine-phosphate lyase to cytidine diphosphate ethanolamine (CDP-E), r_{17} and r_{18}. At this point sphingolipid pathway is again linked to the phospholipid pathway (Alvarez-Vasquez *et al.*, 2005; Obeid *et al.*, 2002).

Table 3.1. The list of reactions and enzymes used in MCA (The key reactions are indicated as bold.) It is taken from Alvarez-Vasquez *et al.* (2005). The underlined reaction numbers indicate the reactions that are further studied and developed in MPA

#	Reactions	Enzyme Name
1	Palm → Pal-CoA	Palmitoyl-Coenzyme A synthase
2	Pal-CoA → PA	Glycerol-3-phosphate acyltransferase
3	PA → DAG	Phosphatidate phosphatase
4	PA → CDP-D	Cytidine diphosphate-diacylglycerol synthase
5	CDP-D + Inositol → PI	Phosphatidyl-inositol synthase
6	G6P → Inositol	Inositol-1-phosphate synthase
7	PI + ATP →	Phosphoinositide kinase
8	Serine + CDP-D → PS	Phosphatidyl-serine synthase
9	3PS → Serine	Phosphoserine Phosphatase
10	Serine →	Serine hydroxymethyl transferase
11	Pal-CoA + Serine → KDHS	Serine Palmitoyl transferase
12	Pal-CoA →	Acyl-Coenzyme A binding protein
13	**KDHS → DHS**	**Ketodihydro-sphingosine reductase**
14	**DHS-P → DHS**	**Sphingoid base-1-phosphate phosphatase**
15	**DHS + ATP → DHS-P**	**Sphingoid Base Kinase**
16	PS →	Phosphatidyl-serine decarboxylase
17	**DHS-P → Pal-CoA + CDP-E**	**Sphingosine phosphate lyase**
18	**PHS-P → Pal-CoA + CDP-E**	**Sphingosine phosphate lyase**
19	**PHS-P → PHS**	**Sphingoid base-1-phosphate phosphatase**
20	**ATP +PHS → PHS-P**	**Sphingoid Base Kinase**
21	**DHS → PHS**	**Hydroxylase**
22	**DHCer → DHS**	**Dihydro-ceramidase**
23	**C_{26}-CoA + DHS → DHCer**	**Ceramide Synthase**
24	**C_{26}-CoA + PHS → PHCer**	**Ceramide Synthase**
25	**PHCer → PHS**	**Phyto-ceramide alkaline ceramidase**
26	**DHCer → PHCer**	**Hydroxylase**
27	**PHCer →**	**Glycosyphosphatidylinositol Remodelase**
28	**IPC → DHCer**	**Inositol phospho-sphingolipid phospholipase C**

Table 3.1. The list of reactions and enzymes used in MCA (The key reactions are indicated as bold.) It is taken from Alvarez-Vasquez *et al.* (2005). The underlined reaction numbers indicate the reactions that are further studied and developed in MPA (cont.)

#	Reactions	Enzyme Name
29	**DHCer + PI → DAG + IPC**	**Inositol phosphory-ceramide synthase**
30	**PHCer + PI → DAG + IPC**	**Inositol phosphory-ceramide synthase**
31	**IPC → PHCer**	**Inositol phospho-sphingolipid phospholipase C**
32	CDP-E + DAG →	Diacylglycerol-ethanolamine phosphotransferase
33	DAG →	Diacylglycerol-choline phosphotransferase
34	AC + CoA + ATP → Ac-CoA	Acetyl-Coenzyme A Synthetase
35	Ac-CoA + ATP → Mal-CoA	Acetyl-Coenzyme A Carboxylase
36	Mal-CoA + Ac-CoA → Pal-CoA	Fatty Acid Synthethase
37	Mal-CoA + Pal-CoA → C_{26}-CoA	Very Long Chain Fatty Acid Synthase
38	**IPC → IPCpm**	**Transport to plasma membrane**
39	**IPC → MIPC**	**Mannosyl Inositol Phosphoceramide Synthase**
40	**MIPC → PHCer**	**Inositol phospho-sphingolipid phospholipase C**
41	**MIPC → MIPCpm**	**Transport to plasma membrane**
42	**MIPC → DHCer**	**Inositol phospho-sphingolipid phospholipase C**
43	**$M(IP)_2C$ → DHCer**	**Inositol phospho-sphingolipid phospholipase C**
44	**$M(IP)_2C$ → PHCer**	**Inositol phospho-sphingolipid phospholipase C**
45	**$M(IP)_2C$ → $M(IP)_2Cpm$**	**Transport to plasma membrane**
46	**IPCpm → IPC**	**Transport from plasma membrane**
47	**MIPCpm → MIPC**	**Transport from plasma membrane**
48	**$M(IP)_2Cpm$ → $M(IP)_2C$**	**Transport from plasma membrane**
49	**MIPC + PI → $M(IP)_2C$ + DAG**	**Mannosyldiinositol Phosphorylceramide Synthase**

DHCer and PHCer can be converted back to long chain bases of DHS and PHS, r_{22} and r_{25}, by dihydroceramide and phytoceramide alkaline ceramidases in the endoplasmic reticulum (Alvarez-Vasquez *et al.*, 2005; Obeid *et al.*, 2002). Hydroxylation of DHS and DHCer to PHS, r_{21}, and PHCer, r_{26}, respectively, are catalyzed by sphingosine hydroxylase encoded by *SUR2* in endoplasmic reticulum (SGD). This enzyme is responsible for the adjustment of balance between the two *S. cerevisiae* sphingoid long chain bases (Alvarez-Vasquez *et al.*, 2005; Obeid *et al.*, 2002). This step does not take place in mammalian cells, where dihydroceramide is desaturated to form ceramide but not phytoceramide (Ogretmen, 2006).

In *S. cerevisiae* three types of complex sphingolipids containing *myo*-inositol are synthesized: inositol phosphorylceramide (IPC), r_{29} and r_{30}, mannosylinositol

phosphorylceramide (MIPC), r_{39}, and mannosyldiinositol phosphorylceramide ($M(IP)_2C$), r_{49}. As the only alteration of ceramide in *S. cerevisiae* is by the addition of inositol phosphate (IP), these three complex sphingolipids are the end-products of *S. cerevisiae* sphingolipid metabolism. This modification by IP occurs only in *S. cerevisiae* but not in mammalian cells; therefore IPC, MIPC, $M(IP)_2C$ may not be found in mammalian cells. In *S. cerevisiae*, complex sphingolipid synthesis reactions are reported to occur in the Golgi apparatus (Obeid *et al.*, 2002; Robert and Lester, 1999).

There are 49 irreversible reactions in *S. cerevisiae* sphingolipid metabolism. The enzymes corresponding to each reaction are listed in Table 3.1. 31 metabolites with their abbreviations are given in Table 3.2. The rate equations are taken from Alvarez -Vasquez *et al.* (2005) who used Biochemical Systems Theory (BST) and Generalized Mass Action (GMA) representation. The rate expressions for each reaction represent the effect of the substrate concentration, the enzyme activity and the modifier concentration (Alvarez-Vasquez *et al.*, 2005). An example for the rate expression of a hypothetical reaction is given by Eq. (3.1).

$$A + B \xrightarrow[M]{X} C$$
$$v = rate\,cons\tan t \times c_A^{p1} \times c_B^{p2} \times c_M^{p3} \times X^{p4} \qquad (3.1)$$

where v : rate

c_A : Concentration of metabolite A.

c_B : Concentration of metabolite B.

c_M : Concentration of modifier M.

X : Enzyme activity.

p : Kinetic order.

Table 3.2. The list of metabolites and abbreviations (The key metabolites are indicated as bold.)

Metabolite Abbreviations	Metabolite Names
3PS	3-phosphate serine
AC	Acetate
Acetyl-CoA	Acetyl-Coenzyme A
ATP	Adenosine tri-phosphate
C_{26}-CoA	hexasanoyl-CoA (n-C26:0 CoA)
CDPDAG	Cytidine-diphosphate diacylglycerol
CDP-E	Cytidine-diphosphate ethanolamine
CoA	Coenzyme A
DAG	Diacylglycerol
DHS	**Dihydro-sphingosine**
DHS-P	**Dihydro-sphingosine-1-phosphate**
DHCer	**Dihydro-ceramide**
G6P	Glucose 6 phosphate
INO	Inositol
IPC	**Inositol phosphoryl-ceramide**
IPCpm	**Inositol Phosphorylceramide- plasma membrane**
KDHS	3-keto-dihydrosphingosine
$M(IP)_2C$	**Mannosyldiinositol Phosphorylceramide**
$M(IP)_2Cpm$	**Mannosyldiinositol Phosphorylceramide- plasma membrane**
Mal-CoA	Malonyl-Coenzyme A
MIPC	**Mannosylinositol Phosphorylceramide**
MIPCpm	**Mannosylinositol Phosphorylceramide- plasma membrane**
PA	Phosphatidic acid
Pal-CoA	Palmitoyl-Coenzyme A
Palm	Palmitate
PHS	**Phyto-sphingosine**
PHS-P	**Phyto-sphingosine-1-phosphate**
PHCer	**Phyto-ceramide**
PI	Phosphatidyl inositol
PS	Phosphatidyl serine
SER	Serine

3.1.2. Metabolic Control Analysis (MCA)

Metabolic control analysis (Heinrich and Rapoport, 1974a; Heinrich and Rapoport, 1974b; Kacser and Burns, 1973) is classified as a post-genomic tool to understand the principles governing a metabolic network's control which is distributed among various enzymatic steps. In this study, metabolic control analysis was employed to quantify the biochemical control of the enzymes constituting the sphingolipid metabolism on the

sphingolipid fluxes and concentrations, and hence to determine the candidate enzymes in the metabolic network appropriate as drug targets for antiproliferative and apoptotic responses.

Control coefficients quantitatively describe the relative change in the systematic parameter such as the reaction flux or metabolite concentration with respect to the enzyme activity. When the amount of change in the enzyme activity is same as the change in the systematic property then the control coefficient is unity. A high control coefficient indicates that the change in the enzyme activity is completely reflected to systematic parameter under investigation. The same rationale applies for a low control coefficient such that the change in the enzyme activity has negligible effect on the systematic parameter (Cascante *et al.*, 2002).

The flux control coefficient (C_i^J) of Eq. (3.2a) illustrates the strength of response in the reaction flux (J), to a change in the enzyme concentration (E_i). The concentration control coefficient (C_i^S) of Eq. (3.2b) illustrates the strength of response in the metabolite concentration (S), to a change in the enzyme concentration (E_i) (Fell, 1992). The concentration control coefficients can take both positive and negative values; a negative control coefficient indicates a negative control effect on the concentration of the metabolite of interest.

$$C_i^J = \frac{\partial J}{\partial E_i} \cdot \frac{E_i}{J} = \frac{\partial \ln |J|}{\partial \ln E_i} \tag{3.2a}$$

$$C_i^S = \frac{\partial S}{\partial E_i} \cdot \frac{E_i}{S} = \frac{\partial \ln |S|}{\partial \ln E_i} \tag{3.2b}$$

As there are various important metabolites/fluxes that have to be considered, three different algorithms were developed to apply a selection criterion for the control coefficients and to identify the vital steps of the pathway (Figure 3.1). The set of key metabolites consists of DHS, PHS, DHCer, PHCer, DHS-P, PHS-P and all complex

sphingolipids at the site of production and plasma membrane (pm) (IPC, IPCpm, MIPC, MIPCpm, $M(IP)_2C$, $M(IP)_2Cpm$) (indicated as bold in Table 3.2).

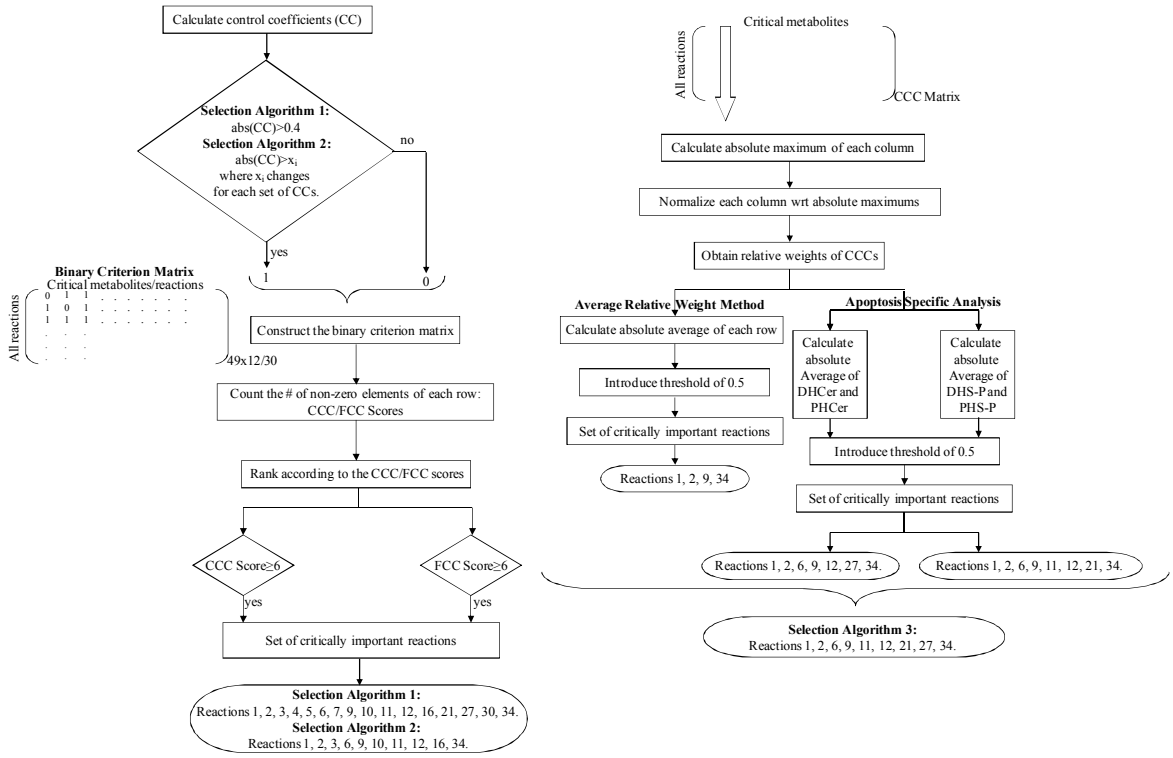


Figure 3.1. The schematic representation of three selection algorithms utilized in the analysis and filtration of MCA results. First two algorithms are shown in the same flowchart representation whereas third algorithm is depicted in a separate flowchart

3.1.2.1. Selection Criterion 1. In the first algorithm a global threshold value of 0.4 is chosen for the concentration control coefficients and a matrix that consists of +1, 0 and -1 is constructed. The columns of this matrix contain the set of selected metabolites and the rows contain the enzymatic steps. As there are 49 enzymatic steps and 12 metabolites of interest, the above described binary criterion matrix had the dimensions of 49×12. Absolute $|+1/-1|$ values were used to represent the concentration control coefficients of the metabolites and reactions which are absolutely greater than the global threshold value of 0.4. Thus, this first algorithm does not discriminate between the negative and positive control coefficients. Then the number of non-zero elements in each row was counted and recorded. The enzymes for each reaction step were ranked according to this score representing the reactions that are important for most of the metabolites constituting the set. The enzymes (reactions) that affect at least 50% and more of the 12 critical metabolites

with concentration control coefficients higher than 0.4 were selected as the significant enzymes and included in the set of potential drug targets.

The same procedure was applied to the analysis of flux control coefficients. A set of critical reactions were chosen among the ones that are producing and consuming the 12 important metabolites in the previous step of the algorithm. These are the reactions numbered 13-15, 17-31, 38-49 and indicated as bold in Table 3.1. The binary criterion matrix had the dimensions of 49×30 where the columns corresponded to the selected reactions instead of the metabolites. In the final step of determining the potential drug target enzymes, the enzymes (reactions) that manipulate at least 50% and more of the 30 important reactions with flux control coefficients higher than 0.4 were selected (Figure 3.1).

3.1.2.2. Selection Criterion 2. In the second algorithm instead of using a global threshold value for both concentration and flux control coefficients, local threshold values are used for each control coefficient examination. The decision of the boundaries, *i.e.* the local threshold values, is made individually according to the gaps between the coefficients, *i.e.* the highest differences between the coefficients were taken as break points. After the determination of the threshold values for each control coefficient, the algorithm follows as the previous one (Figure 3.1).

3.1.2.3. Selection Criterion 3. The basic principle underlying in both algorithms was the utilization of the threshold values and the enzymes over a specified threshold are treated equally. In order to use the information embedded in control coefficients more efficiently, a third algorithm is designed where a matrix of concentration control coefficients of the selected critical metabolites is constructed and the maximum values of the absolute control coefficients in each column is found separately. Then concentration control coefficients are normalized with respect to the absolute maximum value of each column so that the matrix at this stage consists of the relative weights of the concentration control coefficients. This stage is a branch point of the algorithm where further analysis involves two approaches, *i.e.* (i) overall interpretation by average relative weights and (ii) apoptosis specific analysis by relative weights.

First an overall interpretation of the weighted concentration control coefficients are performed which involves averaging the weights for each enzymatic step without considering the direction of action, *i.e.* signs of the control coefficients. The important reactions from average weights of both concentration control coefficients and flux control coefficients are selected by using a threshold value of 0.5.

Second, in apoptosis specific analysis, the constructed critical metabolite set is partitioned into two parts according to the knowledge about the action of metabolites on apoptosis in mammalian cells. As mentioned formerly, ceramide induces apoptosis whereas sphingosine-1-phosphate promotes cell growth and proliferation. Hence, the concentration control coefficients of DHCer and PHCer are examined separately, and the phosphorylated forms of long chain bases are analyzed accordingly (Figure 3.1).

Metabolic control analysis was performed using the software GEPASI version 3.30 (Mendes, 1993). The ordinary differential equations employed in the metabolic control analysis were also solved under MATLAB R2007a using the stiff differential equation solver with the variable order method (Appendix D).

3.1.3. Metabolic Pathway Analysis (MPA)

Metabolic pathway analysis is a powerful stoichiometric modelling approach for systems biology and used here in the identification of physiologically significant sets of connected reactions in sphingolipid metabolic network. The results of metabolic pathway analysis are examined and used in potential drug target identification process for cancer therapy.

In order to study the inherent network properties, the metabolic reaction network of Alvarez-Vasquez *et al.* (2005) is reconstructed; the reactions are balanced stoichiometrically using Expasy, KEGG, BRENDA databases and the genome scale models from the literature (Forster *et al.*, 2003; Duarte *et al.*, 2004; Kuepfer *et al.*, 2005); and the lumped reactions are decomposed as shown in Figure 3.2.

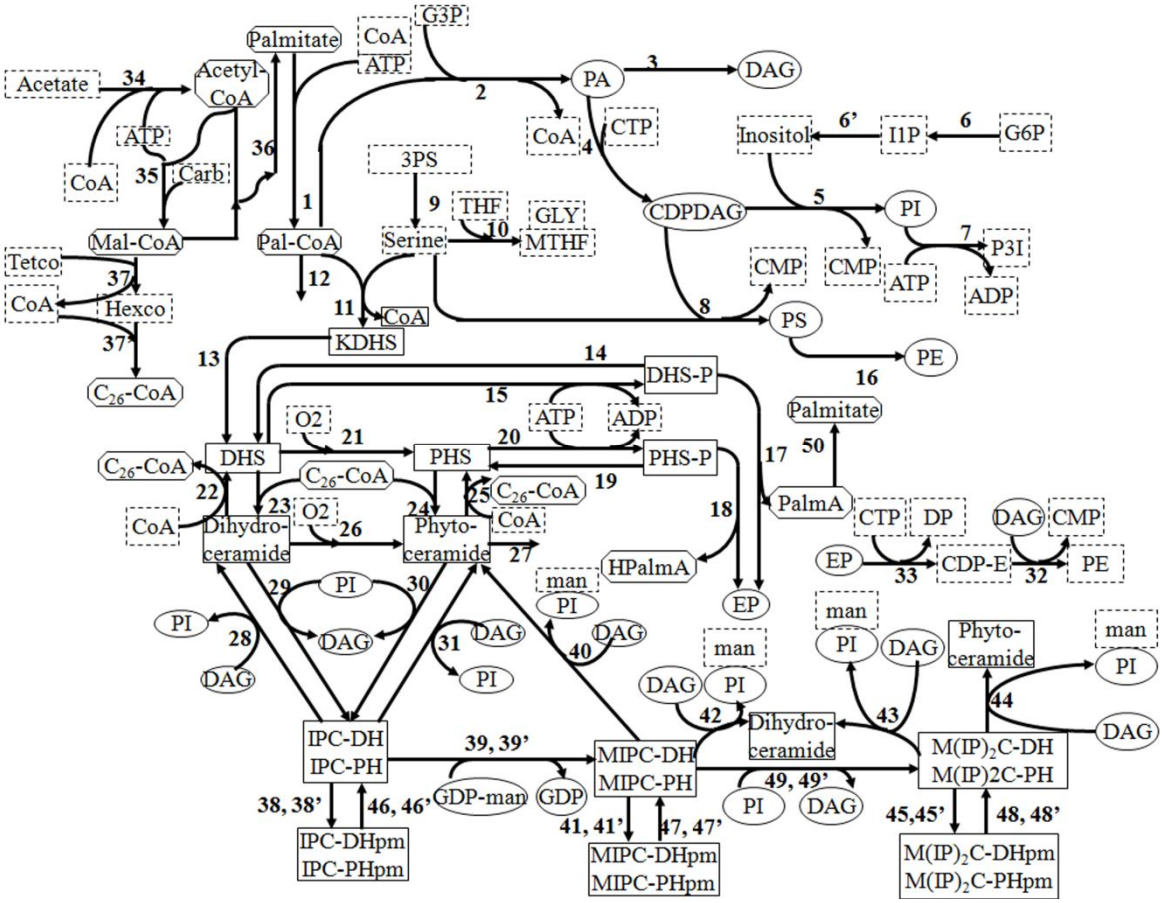


Figure 3.2. The metabolites involved in fatty acid metabolism are shown in octagon boxes; the phospholipids are in oval boxes, and the sphingolipids are shown in rectangular boxes. The dashed boxes show the metabolites that do not fit the above classifications

3.1.3.1. Improvements in Reaction System. The lumped reaction for the synthesis of C_{26}-CoA from malonyl-coenzyme A (Mal-CoA) is represented by two distinct reactions: First Mal-CoA is converted to hexacosanoate (Hexco), r_{37}, and then converted to C_{26}-CoA, r_{37}', and Pal-CoA is no more a substrate for this reaction. The reaction 36 is modified so that in the newly reconstructed network, palmitate is synthesized by r_{36}', instead of Pal-CoA. The reaction 6 is also decomposed to two separate reactions, first glucose-6-phosphate (G6P) is converted to inositol-1-phosphate (I1P), r_6, and then it is converted to inositol, r_6'.

The products of the decomposition reactions of the phosphorylated long chain bases show also differences with the ones in Alvarez-Vasquez *et al.*'s (2005) pathway. DHS-P and PHS-P are decomposed to ethanolamine phosphate (EP) and palmitaldehyde (PalmA), r_{17} and EP and 2-hydroxy-hexadecanal (HPalmA), r_{18}, respectively in the newly

reconstructed network. The synthesized EP from these decomposition reactions is converted to cytidine diphosphate-ethanolamine (CDP-E) on Figure 3.2 (r_{33}). The products of 17th and 18th reactions catalyzed by lyases were Pal-CoA and CDP-E for both long chain base phosphates in the older version of the metabolic network. But such distinct reactions are not available in the literature and also when the elements' balance is controlled for the products, it is impossible to balance the elements. A new reaction is added to the metabolic network such that the conversion of PalmA (the decomposition reactions' product) to palmitate, r_{50}, and then to Pal-CoA becomes possible, r_1.

The complex sphingolipids synthesized from DHCer and PHCer are marked as follows: IPC produced from PHCer is named as IPC-PH and represented as IPC-DH if it is synthesized from DHCer. The difference of reactions 38, 39, 41, 45, 46, 47, 48 and 49 from their primed versions are their source, *i.e.* whether they are produced from DHCer or PHCer. This distinction is made to follow the two types of long chain bases, DHS and PHS (Table 3.3).

Table 3.3. The balanced reaction list with the enzymes catalyzing the reactions used in metabolic pathway analysis

#	Reactions	Enzyme names
1	ATP + Palm + CoA -> AMP + DP + Pal-CoA	Palmitoyl-CoA synthase
2	Pal-CoA + G3P -> CoA + PA	Glycerol-3-Phosphate Acyltransferase
3	H2O + PA -> DAG + P	Phosphatide phosphatase
4	PA + H + CTP -> CDPDAG + DP	CDP-DAG syhthase
5	CDPDAG + INO -> PI + CMP + H	Phosphatidylinositol synthase
6	G6P -> I1P	Inositol-1-Phosphate Synthase
6'	I1P + H2O -> INO + P	Inositol producing and missing rxn
7	ATP + PI -> ADP + H + P3I	Phosphatidylinositol kinase
8	CDPDAG + SER -> CMP + H + PS	Phosphatidylserine Synthase
9	3PS + H2O -> P + SER	Phosphoserine Phosphatase
10	SER + THF -> GLY + H2O + MTHF	Serine Hydroxymethyl Transferase
11	H + Pal-CoA + SER -> KDHS + CO2 + CoA	Serine C-palmitoyltransferase
12	Pal-CoA -> Pal-CoAxt	Acyl-CoA-Binding Protein
13	KDHS + H + NADPH -> NADP + DHS	3-Dehydrosphinganine reductase
14	H2O + DHS-P -> P + DHS	Sphingoid base-phosphate phosphatase (sphinganine 1-phosphatase)
15	ATP + DHS -> ADP + H + DHS-P	Sphingolipid long chain base kinase (sphinganine kinase)
16	H + PS -> CO2 + PE	Phosphatidylserine Decarboxylase

Table 3.3. The balanced reaction list with the enzymes catalyzing the reactions used in metabolic pathway analysis (cont.)

#	Reactions	Enzyme names
17	DHS-P -> EP + PalmA	Sphinganine phosphate lyase
18	PHS-P -> HPalmA + EP	Phytosphingosine phosphate lyase
19	H2O + PHS-P -> P + PHS	Sphingoid base-phosphate phosphatase (phytosphingosine 1-phosphate phosphatase)
20	ATP + PHS -> ADP + H + PHS-P	Sphingolipid long chain base kinase (phytosphingosine kinase)
21	H + NADPH + O2 + DHS -> H2O + NADP + PHS	Phytosphingosine synthesis
22	DHCer + CoA + H -> C26-CoA + DHS	Dihydroceramide Alkaline Ceramidase
23	C26-CoA + DHS -> DHCer + CoA + H	Ceramide-1 synthase (26C)
24	C26-CoA + PHS -> PHCer + CoA + H	Ceramide-2 synthase (26C)
25	PHCer + CoA + H -> C26-CoA + PHS	Phytoceramide Alkaline Ceramidase
26	DHCer + H + NADPH + O2 -> PHCer + H2O + NADP	Ceramide-1 hydroxylase (26C)
27	PHCer -> PHCerxt	GPI remodelase
28	DAG + IPC-DH -> DHCer + PI	Inositol Phosphosphingolipid Phospholipase C
29	DHCer + PI -> DAG + IPC-DH	Inositol phosphorylceramide synthase (26C)
30	PHCer + PI -> DAG + IPC-PH	Inositol phosphoryceramide synthase (26C)
31	DAG + IPC-PH -> PHCer + PI	Inositol Phosphosphingolipid Phospholipase C
32	DAG + CDP-E -> CMP + H + PE	Diacylglycerol-Ethanolamine Phosphotransferase
33	EP + CTP + H -> CDP-E + DP	Phosphoethanolamine Cytidylyltransferase
34	ATP + AC + CoA -> AMP + DP + Ac-CoA	Acetyl-Coenzyme A Synthetase
35	Ac-CoA + ATP + Carb -> ADP + H + Mal-CoA + P	Acetyl-Coenzyme A Carboxylase
36	Ac-CoA + 7 Mal-CoA + 14 NADPH + 7 H -> Palm + 7 CO2 + 14 NADP + 8 CoA + 6 H2O	Fatty Acid Synthetase
37	3 H + Mal-CoA + 2 NADPH + Tetco -> CO2 + CoA + H2O + Hexco + 2 NADP	Very long chain fatty acid synthase
37'	ATP + CoA + Hexco -> AMP + C26-CoA + DP	Very long chain fatty acid synthase
38	IPC-DH <-> IPC-DHpm	Transport of IPCDH to plasma membrane
38'	IPC-PH <-> IPC-PHpm	Transport of IPCPH to plasma membrane
39	GDPman + IPC-DH -> GDP + H + MIPC-DH	Mannose-inositol phosphorylceramide synthase (26C)
39'	GDPman + IPC-PH -> GDP + H + MIPC-PH	Mannose-inositol phosphorylceramide synthase (26C)
40	H2O + DAG + MIPC-PH -> PHCer + PI + man	Inositol Phosphosphingolipid Phospholipase C
41	MIPC-DH <-> MIPC-DHpm	Transport of MIPCDH to plasma membrane
41'	MIPC-PH <-> MIPC-PHpm	Transport of MIPCPH to plasma membrane
42	H2O + DAG + MIPC-DH -> DHCer + PI + man	Inositol Phosphosphingolipid Phospholipase C
43	H2O + 2 DAG + M(IP)2C-DH -> DHCer + 2 PI + man	Inositol Phosphosphingolipid Phospholipase C
44	H2O + 2 DAG + M(IP)2C-PH -> PHCer + 2 PI + man	Inositol Phosphosphingolipid Phospholipase C
45	M(IP)2C-DH <-> M(IP)2C-DHpm	Transport of $M(IP)_2CDH$ to plasma membrane
45'	M(IP)2C-PH <-> M(IP)2C-PHpm	Transport of $M(IP)_2CPH$ to plasma membrane
49	MIPC-DH + PI -> DAG + M(IP)2C-DH	Mannosyldiinositol Phosphorylceramide synthase (26C)
49'	MIPC-PH + PI -> DAG + M(IP)2C-PH	Mannosyldiinositol Phosphorylceramide synthase (26C)
50	PalmA + NAD + H2O -> Palm + NADH	

3.1.3.2. Elementary Flux Modes. Metabolic pathway analysis algorithm involves the calculation of elementary flux modes of the sphingolipid system. Elementary flux modes (EFM) of a metabolic system specify all the possible sets of enzymes that can operate without disturbing the valid steady state of the system (Cakir *et al.*, 2004; Klamt *et al.*, 2003; Klamt *et al.*, 2005; Schuster *et al.*, 1999; Schwarz *et al.*, 2005). These identified functional pathways cannot be decomposed into smaller meaningful parts and they can be used as the linear basis for the expression of the whole network states. The basic principles that should be satisfied during computation of elementary flux modes are (i) a pseudo steady state condition (ii) a feasibility condition and (iii) a non-decomposability condition.

First, elementary flux modes were determined using CellNetAnalyzer Version 6.3. (Klamt *et al.*, 2007), and then a set of key metabolites was constructed including DHCer, PHCer, DHS-P and PHS-P due to the significance of ceramide and sphingosine-1-phosphate in apoptosis. The elementary flux modes in that all these key metabolites participate are selected. The reactions that are common in all the selected elementary flux modes are identified and the enzymes catalyzing these common reactions are noted as candidate potential drug targets. The selection of the EFMs that all the elements of the key set participate in and the common reactions occurring in these EFMs is performed by an algorithm coded in MATLAB R2007a (Appendix E).

3.1.3.3. Control Effective Flux Analysis. The control effective flux (CEF) of a reaction represents the weighted sum of all the fluxes of that reaction observed in all the possible routes (elementary flux modes). So it can be considered as the quantifier of importance of a reaction at steady states. In literature, the control effective flux values were used in the identification of enzymes that can serve as potential drug targets (Cakir *et al.*, 2004; Cakir *et al.*, 2007).

The CEF calculation necessitates the definition of substrate uptake reaction and objective reactions. The substrate uptake reaction is chosen to be the 11^{th} reaction catalyzed by Serine palmitoyl transferase, since it is the pioneering reaction of the *de novo* sphingolipid synthesis. In fact the substrates to the sphingolipid metabolism could have been selected as serine and Pal-CoA, but as we could not discriminate between the two in terms of importance, the next step that is the initial step of *de novo* synthesis is selected as

the substrate uptake reaction. The objective reaction issue is resolved by taking into account the key metabolite set employed previously in elementary flux mode analysis. The reactions synthesizing DHCer and PHCer are selected (r_{23} and r_{24}) as objective reactions. They serve the aim of accumulation of ceramide in cancer therapy via the manipulation of the sphingolipid pathway. Another facet of the same cancer therapy approach is the depletion of DHS-P and PHS-P, this fact is represented in the objective reaction set by including reactions 17 and 18. Therefore, the first objective function set consists of reactions 17, 18, 23 and 24. The second set additionally contains reactions 14 and 19, converting DHS-P and PHS-P back to the long chain bases. However, the two objective reaction sets make no difference in terms of the consequences so one of them will be discussed in the following sections.

3.2. Results

Ceramide and sphingosine-1-phosphate, being the sphingolipid rheostat, are known to play critical roles in cell proliferation and cell's death. This fact about the two opposing signaling molecules leads us to investigate the effects of all ceramide related metabolites using metabolic control analysis and metabolic pathway analysis.

3.2.1. Drug Target Identification by Metabolic Control Analysis

The discovery of the key enzymes functioning in the sphingolipid metabolism is essential for both understanding the metabolism and identifying drug targets for cancer therapy. The control coefficients belonging to the metabolic system given in Table 3.1 are used to identify the critically important enzymes that can be used as potential drug targets in future applications. A high concentration control coefficient (CCC) indicates that a change in the enzyme concentration is strongly reflected in the desired metabolite concentration. For analyzing and processing the CCCs, the set of metabolites is grouped into two parts as they lead to different cellular processes. The 'tumor suppressor' lipids (DHS, PHS, DHCer, PHCer, complex sphingolipids) are examined in a different category than the "tumor-promoting' lipids of DHS-P and PHS-P. The enzymes that catalyze the reactions containing this set of metabolites and the related fluxes and also have high control coefficients are selected as a good starting point in drug target identification. The

ultimate drug designed to target the enzyme with high or optimum control coefficient is expected to have an efficient outcome that can be achieved by a small concentration introduced to the organism.

3.2.1.1. Selection Algorithms (SA). Using the first selection algorithm, the following enzymes are identified as candidates for being potential drug targets (Figure 3.3): palmitoyl-coenzyme A synthase (r_1), glycerol-3-phosphate acyltransferase (r_2), inositol-1-phosphate synthase (r_6), phosphoserine-phosphatase (r_9), acyl-coenzyme A binding protein (r_{12}), serine hydroxymethyl transferase (r_{10}), serine palmitoyltransferase (r_{11}), acetyl-coenzyme A synthetase (r_{34}), phosphatidate phosphatase (r_3), phosphatidylserine decarboxylase (r_{16}), phosphoinositol kinase (r_7), 4-hydroxylase (r_{21}), CDP-diacylglycerol synthase (r_4), phosphatidylinositol synthase (r_5), in the order of importance (The enzyme names are taken from Alvarez-Vasquez *et al.*, 2005).

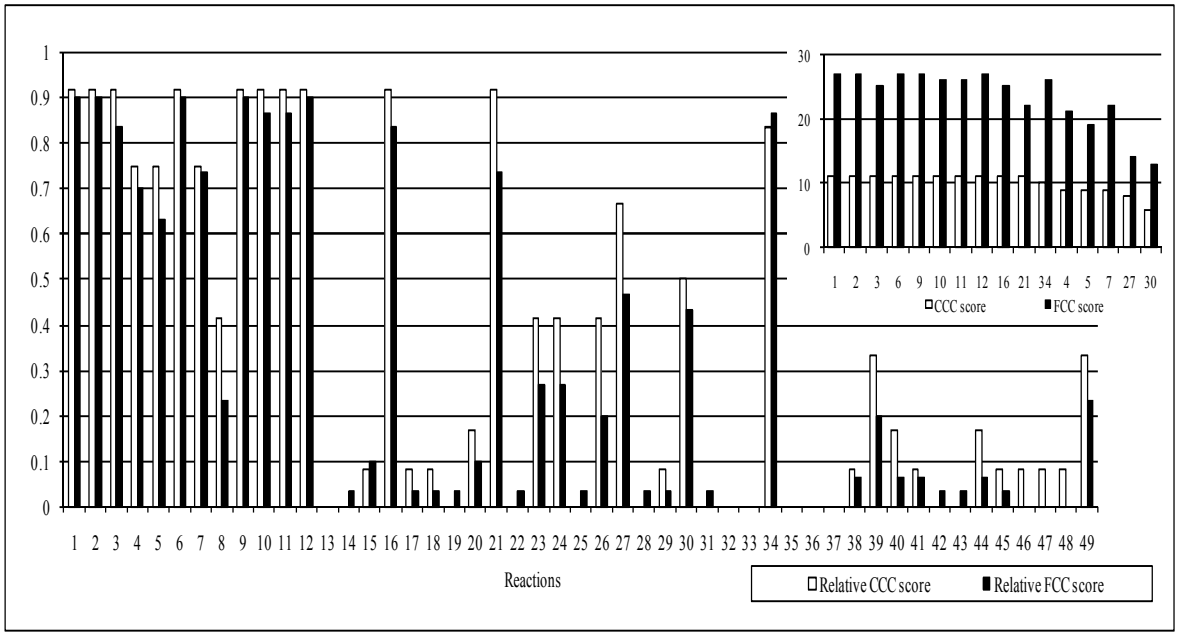


Figure 3.3. The relative scores of the control coefficients obtained from the first selection algorithm utilizing a global threshold value. On the insert, the CCC and FCC scores of the selected 16 critically important reactions are ranked on a bar graph representation

By examining Figure 3.3, it can be seen that the critically significant candidate enzymes for being potential drug targets coincide for both concentration and flux control coefficients. Although the analysis of relative concentration control coefficient (CCC)

scores indicates that the enzymes glycosylphosphatidylinositol (GPI) Remodelase and IPC synthase catalyzing the 27th and 30th reactions can also be considered as significant, the relative flux control coefficient (FCC) scores of the reactions 27 and 30 are less than 0.5, namely 0.467 and 0.433, respectively. An FCC of 0.433 means that a change in the concentration of IPC synthase (rxn 30) affects 13 out of 30 reaction fluxes.

In the second algorithm the use of local thresholds caused the limit of enzyme selection to increase from 0.4 to 0.8, meaning that the enzymes manipulating more than 80 % of 12 critical metabolites and 30 reactions are identified as candidates of drug targets. The concentration and flux control coefficients again point out to the same candidate enzymes (Figure 3.4). The candidate set includes the enzymes of the reactions 1, 2, 3, 6, 9, 10, 11, 12, 16 and 34. The second algorithm gives a more refined set of enzymes as expected.

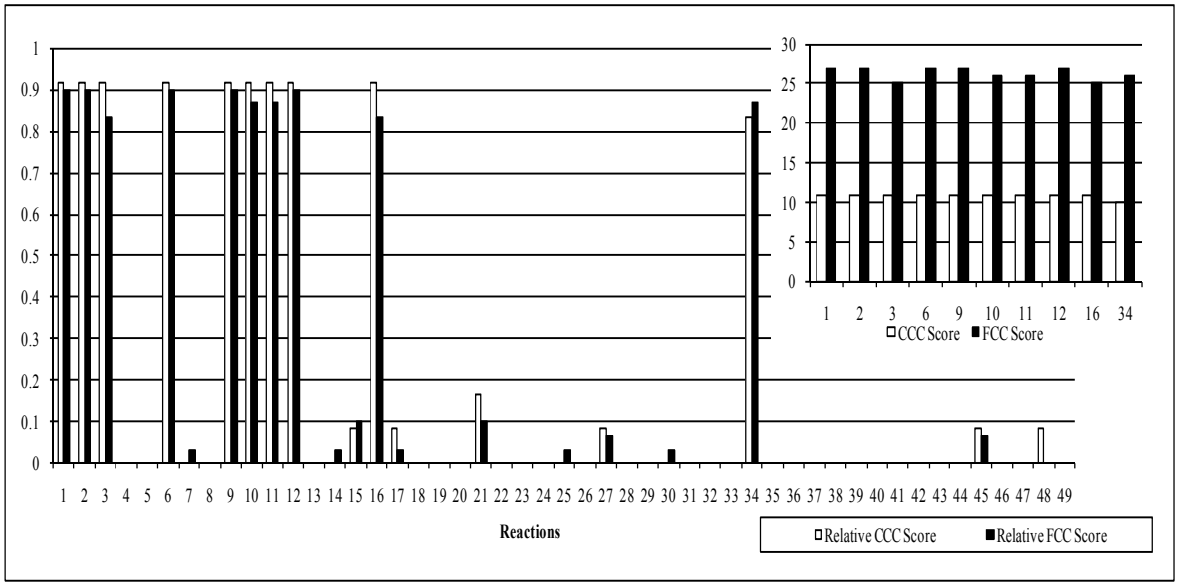


Figure 3.4. The relative scores of control coefficients obtained from the second selection algorithm utilizing a local threshold value. On insert, the CCC and FCC scores of the 10 critically important reactions are ranked and shown on a bar graph representation

The important reactions (r_1, r_2, r_9 and r_{34}) obtained from the third algorithm with 'average relative' weights of control coefficients compose a small subset of the previously identified reaction sets (Figure 3.5). These four reactions are related to the production and

consumption of the fatty and amino acids required for the initiation of sphingolipid synthesis in *S. cerevisiae*.

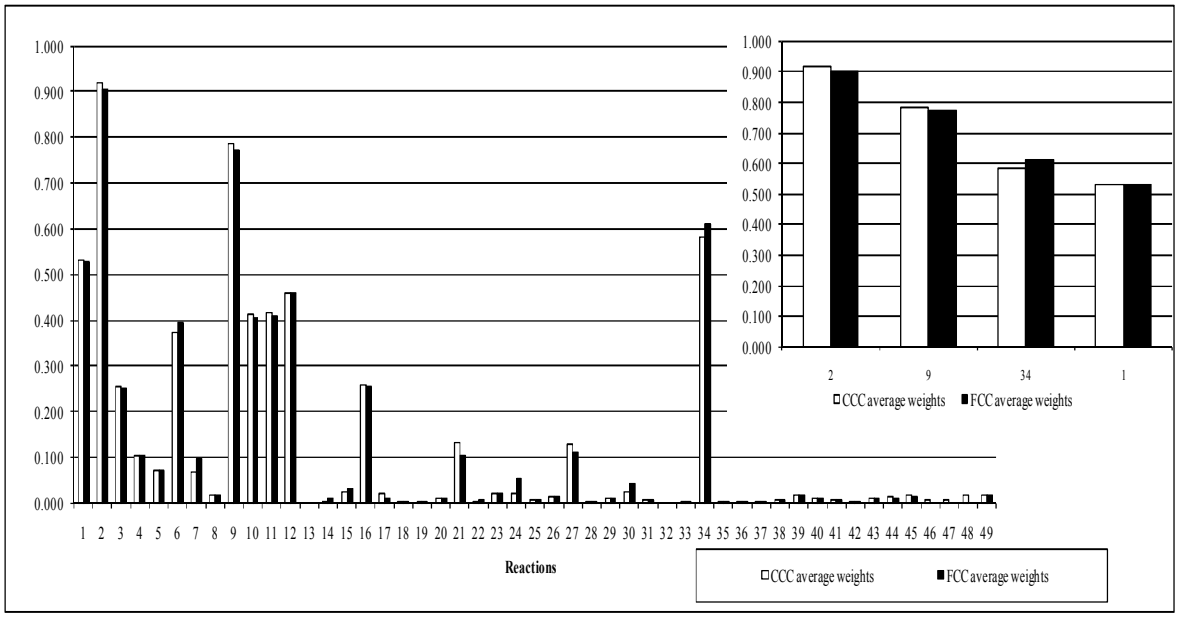


Figure 3.5. The graphical representation of the 'average' relative weights of both CCCs and FCCs (third algorithm). The important reactions selected (r_1, r_2, r_9 and r_{34}) compose a small subset of the previously identified reaction sets and shown in the insert

3.2.1.2. Apoptosis Specific Analysis. In apoptosis specific analysis (3rd algorithm), the constructed critical metabolite set is partitioned into two parts according to the knowledge about the action of metabolites on apoptosis in mammalian cells. The CCCs of the phosphorylated forms of the long chain bases are examined separately from that of the DHCer and PHCer utilizing the matrix of 'relative' weights of the control coefficients. Consequently, the enzymes catalyzing the reactions 1, 2, 6, 9, 11, 12, 21, 27 and 34 are selected as candidates for being drug targets (Table 3.4 and Figure 3.6).

When the signs of the control coefficients are also taken into account instead of the absolute values, the activities of the enzymes belonging to the reactions 6, 9, 11 and 12 have to be decreased to obtain a decrease in the concentrations of DHS-P and PHS-P. The increase in the activities of the enzymes catalyzing the reactions 1, 2, 21 and 34 are found to cause negative influences on the concentrations of phosphorylated sphingoid bases of *S. cerevisiae*.

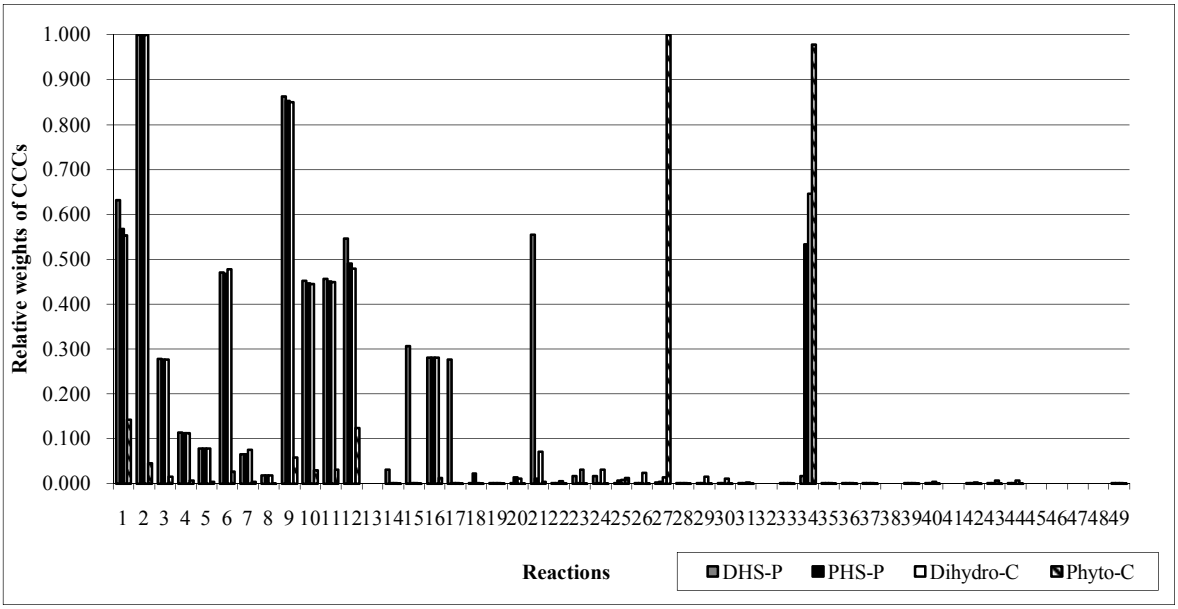


Figure 3.6. The relative weights of CCCs of phosphorylated forms of long chain bases, dihydroceramide and phytoceramide for the selected enzymes. The enzymes catalyzing the reactions 1, 2, 6, 9, 11, 12, 21, 27 and 34 are selected in apoptosis specific analysis

When the signs of flux control coefficients and their effects on ceramides are considered it can be seen that the enzymatic activity of reaction 34 should be increased in order to have an increase in the concentrations of both ceramides (DHCer and PhCer) in *S. cerevisiae*, whereas the enzyme activity of reaction 27 should be decreased to obtain a similar effect on ceramide concentration.

Table 3.4. Relative weights of concentration control coefficients of potential drug target enzymes identified by apoptosis specific analysis

DHCer		**PHCer**		**DHS-P**		**PHS-P**	
R2	1	**R27**	-1	**R2**	-1	**R2**	-1
R9	-0.8503	**R34**	0.9782	**R9**	0.8627	**R9**	0.8524
R34	0.646			**R1**	-0.6312	**R1**	-0.567
R1	0.5538			**R21**	-0.555	**R34**	-0.5331
R12	-0.4796			**R12**	0.5465	**R12**	0.491
R6	-0.4777			**R6**	0.4713	**R6**	0.4683
				R11	0.456		

Moreover, the enzymes catalyzing the reactions 6, 9, and 12 have negative effects and those catalyzing the reactions 1 and 2 have positive effects on DHCer but they have almost no influence on the concentration of PHCer.

3.2.1.3. Perturbations on Candidate Enzymes. Metabolic control analysis indicates that in order to be utilized in a cancer therapy approach (increasing ceramide concentration and decreasing phosphorylated long chain base concentrations), the activities of the enzymes catalyzing reactions 1, 2, 21 and 34 should be increased while the enzymatic activities belonging to reactions 6, 9, 11, 12, and 27 should be reduced.

The effects of these enzymes were studied at a 15 % perturbation in their activities. The change in enzyme activities is kept low as it is required that the ultimate goal (*i.e.* ceramide accumulation) should be attained by a minimum disruption of the rest of the sphingolipid metabolism. The simulation results were examined using the relative changes in the DHCer, PHCer, DHS-P and PHS-P concentrations with respect to the wild type concentrations at the 24^{th} hour of perturbation (Table 3.5).

More than 50 % increases with respect to the wild type are reached in DHCer concentrations (except for the enzymes catalyzing reactions 21, 27 and 34) (Table 3.5). The most effective potential drug target enzyme for DHCer is glycerol-3-phosphate acyltransferase (r_2) with a 92.23 % increase. The most effective enzyme that can be utilized as a candidate drug target for PHCer accumulation is the enzyme catalyzing reaction 27 with a 37.95 % increase in the metabolite concentration. When the changes attained with the same enzymatic manipulations (15 % increase/decrease) are compared for the two ceramides, it can be seen that DHCer is almost all the times more vulnerable to perturbations whereas PHCer is more robust. This result may indicate the biological significance of PHCer in *S. cerevisiae* cells. The critically important components of metabolic networks should be robust to alterations (Barkai and Leibler, 1997).

Palmitoyl-coenzyme A synthase (r_1), glycerol-3-phosphate acyltransferase (r_2), phosphoserine-phosphatase (r_9) and acyl-coenzyme A binding protein (r_{12}) contribute to the decline in the concentration of DHS-P. When PHS-P is considered a 15 % increase in the activity of glycerol-3-phosphate acyltransferase (r_2) is found to produce an 88 %

decrease in PHS-P concentrations after 24 hours from the perturbation (Table 3.5). As it can be seen from Table 3.5, the manipulation of the selected enzyme activities gives the desired end results, *i.e.* increase in DHCer and PHCer concentrations and decrease in DHS-P and PHS-P concentrations as required for the application of the novel cancer therapy approach.

Table 3.5. The perturbation studies on candidate enzymes. The effect of 15 % perturbations in candidate enzymes on key metabolite concentrations at the 24th hour of the simulations. The most effective changes in each key metabolite are highlighted

Concentrations @24 hours	**r1 +**	**r2 +**	**r21 +**	**r34 +**	**r6 -**	**r9 -**	**r11 -**	**r12 -**	**r27 -**	**WT**
DHCer	0.0670	0.0767	0.0263	0.0580	0.0672	0.0730	0.0609	0.0667	0.0441	0.0399
% change DHCer	67.92	**92.23**	-34.09	45.36	68.42	82.96	52.63	67.17	10.53	
PHCer	0.0374	0.0333	0.0529	0.0544	0.0432	0.0351	0.0430	0.0375	0.0727	0.0527
% change PHCer	-29.03	-36.81	0.38	3.23	-18.03	-33.40	-18.41	-28.84	**37.95**	
DHS-P	0.0006	0.0006	0.0008	0.0010	0.0008	0.0006	0.0008	0.0006	0.0010	0.0010
% change DHS-P	**-40.00**	**-40.00**	-20.00	0.00	-20.00	**-40.00**	-20.00	**-40.00**	0.00	
PHS-P	0.0010	0.0006	0.0052	0.0031	0.0016	0.0007	0.0016	0.0010	0.0051	0.0050
% change PHS-P	-80.00	**-88.00**	4.00	-38.00	-68.00	-86.00	-68.00	-80.00	2.00	
IPC	0.1471	0.1622	0.0893	0.1583	0.0973	0.1570	0.1438	0.1469	0.1643	0.1151
% change IPC	27.80	40.92	-22.42	37.53	-15.46	36.40	24.93	27.63	42.75	
MIPC	0.1877	0.2029	0.1255	0.1989	0.1768	0.1976	0.1843	0.1874	0.2050	0.1541
% change MIPC	21.80	31.67	-18.56	29.07	14.73	28.23	19.60	21.61	33.03	
$M(IP)_2C$	0.0124	0.0128	0.0103	0.0127	0.0091	0.0127	0.0123	0.0124	0.0129	0.0113
% change $M(IP)_2C$	9.73	13.27	-8.85	12.39	-19.47	12.39	8.85	9.73	14.16	

3.2.2. Drug Target Identification by Metabolic Pathway Analysis

Metabolic pathway analysis was performed for the improved system of reactions which are shown on Figure 3.2. The newly developed metabolic pathway was constructed in such a way that the very first substrates and the end products are consistent with the system employed in the metabolic control analysis. Since the metabolic pathway analysis

is a stoichiometric tool the reactions in the improved model are stoichiometrically balanced including the cofactors and hydrogen, and the lumped reactions are decomposed. The current sphingolipid pathway used in metabolic pathway analysis has 54 reactions and 41 transport reactions and 67 metabolites.

3.2.2.1. Elementary Flux Modes (EFMs). Elementary flux modes are here used to identify the functionally and structurally important enzymatic steps in the *S. cerevisiae* sphingolipid metabolism and hence degree of control of reactions. The present system of sphingolipid reactions resulted in 2446 elementary flux modes which were studied in terms of two different approaches, *i.e.* common reactions among selected EFMs and control effective fluxes.

The key metabolites, DHCer, PHCer, DHS-P and PHS-P, are found to participate in 164 out of 2446 calculated EFMs. The common reactions available in all the 164 EFMs were reactions 1, 2, 9, 11, 13, 15, 20, 23, 26, 32 and 33. Furthermore the transport reactions of adenosine triphosphate (ATP), 3-phosphoserine (3PS), glycerol-3-phosphate (G3P), cytidine triphosphate (CTP), reduced nicotinamide adenosine dinucleotide phosphate (NADPH), oxygen (O_2), adenosine monophosphate (AMP), diphosphate (DP), phosphate (P), cytidine monophosphate (CMP), carbondioxide (CO_2), nicotinamide adenosine dinucleotide phosphate (NADP) and phosphatidyl ethanolamine (PE) were also identified as common reactions occurring in all the selected EFMs of the key metabolites set.

In an additional EFM analysis, a second set of critical metabolites were constructed where the end products of the sphingolipid pathway that is the complex sphingolipids, IPC, MIPC and $M(IP)_2C$ synthesized from both DHCer and PHCer were also taken into consideration beside the elements of the first set. However, there were no EFMs in which all the constituent metabolites of the second set participated.

3.2.2.2. Control Effective Fluxes (CEFs). The information for the identification of potential drug targets was extracted from the determined elementary flux modes via control effective flux (CEF) analysis. The calculated CEF values for the internal fluxes excluding the transport reactions are given on Table 3.6 using the first objective function set. The

first 10 reactions were identified as candidates of potential drug target enzymes. The candidate enzymes catalyze the reactions 34, 35, 1, 9, 13, 2, 4, 5, 30 and 39' in the order of their importance.

Table 3.6. The calculated CEF values for the internal fluxes excluding the transport reactions. The normalized CEF values with respect to the maximum CEF of internal fluxes are given in the second column. The first 10 reactions are highlighted

Reactions	CEF	Normalized CEF	Reactions	CEF	Normalized CEF
34	**11.9544**	**1**	26	1.1521	0.1
35	**10.6808**	**0.89**	6	1.1358	0.1
1	**7.5022**	**0.63**	6'	1.1358	0.1
9	**4.0306**	**0.34**	40	1.133	0.09
13	**4**	**0.33**	42	1.133	0.09
2	**3.5022**	**0.29**	43	1.1008	0.09
4	**3.3047**	**0.28**	44	1.1008	0.09
5	**3.2741**	**0.27**	17	0.9297	0.08
30	**3.2647**	**0.27**	50	0.9297	0.08
39'	**2.966**	**0.25**	24	0.7618	0.06
29	2.6435	0.22	25	0.5586	0.05
39	2.5272	0.21	27	0.3245	0.03
20	2.4725	0.21	41'	0.3084	0.03
32	2.2348	0.19	38'	0.2987	0.02
33	2.2348	0.19	45'	0.2114	0.02
15	2.0972	0.18	3	0.1975	0.02
37	1.7652	0.15	41	0.1207	0.01
37'	1.7652	0.15	38	0.1163	0.01
23	1.5619	0.13	45	0.0862	0.01
49'	1.5246	0.13	8	0.0306	0
21	1.5083	0.13	16	0.0306	0
18	1.3051	0.11	10	0	0
36	1.2737	0.11	12	0	0
49	1.2735	0.11	22	0	0
7	1.2368	0.1	28	0	0
14	1.1675	0.1	31	0	0
19	1.1675	0.1			

As a concluding remark that can be dictated from the metabolic pathway analysis results, the reactions 1, 2, 4, 5, 9, 11, 13, 15, 20, 23, 26, 30, 32, 33, 34, 35 and 39' may be stated to be potential drug targets (Table 3.7). When the candidate enzymes proposed for being drug targets achieved by metabolic pathway analysis are put side by side with the known potential drug target enzymes in pre-clinical or clinical development stages of drug discovery, it can be seen that metabolic pathway analysis can identify the known targets

such as the enzymes of reactions 11, 15, 20 and 23. The mathematical tools' efficiency for drug target identification performed in this study is thus validated by the clinically available drugs.

Table 3.7. Potential drug targets that are determined using three selection algorithms of MCA results, two different computations using MPA results (EFM, CEF). The enzymes that are known drug targets in clinical development stages are also indicated as bold

Reactions	1	2	3	4	5	6	7	9	10	**11**	12	13	**15**	16	**20**	21	**23**	26	27	30	32	33	34	35	39'
MCA S.C. 1	+	+	+	+	+	+	+	+	+	+	+			+		+			+				+		
MCA S.C. 2	+	+	+			+		+	+	+	+			+									+		
MCA S.C. 3	+	+				+		+		+	+					+			+				+		
MPA EFM	+	+						+		+		+	+		+		+	+			+	+			
MPA CEF	+	+		+	+			+				+								+			+	+	+

3.3. Discussion

We combined the results of the two mathematical tools (MCA and MPA) to obtain a more complete view and reached an ultimate set of potential drug target enzymes (Table 3.7). These common enzymes that are indicated by both MCA and MPA results are of the reactions 1, 2, 9, 11 and 34 and they are identified by at least four out of five of the methods discussed here. Moreover, if the target enzymes that appear in at least one of the MCA selection algorithms and one of the MPA computations are also included in this final set, the enzymes catalyzing the reactions 4, 5 and 11 emerge to be the additional candidates. Besides these, the known drug target enzymes from the literature that are identified by the MPA results could also be included in the concluding set (r_{11}, r_{15}, r_{20}, r_{23}). The target enzymes that are determined by all the three selection algorithms of MCA but not MPA, *i.e.* enzymes belonging to reactions 6 and 12, can be added to the candidate list due to their reliability by passing all selection criterions of MCA. Another reason for including reaction 6 in the ultimate set of potential drug targets is that the manipulation of the enzyme activity of reaction 6 does not disturb the complex sphingolipid concentrations

much with the percent change values of -15.46 %, 14.73 % and -19.47 % for the IPC, MIPC and $M(IP)_2C$, respectively (Table 3.5). An ideal drug is expected not to affect the metabolism other than its targets. Although not identified by MPA results, reactions 21 and 27 from MCA results are also taken into account due to reaction 27's essential effect on PHCer accumulation (37.95 % change, Table 3.5) and reaction 21's almost null effect on complex sphingolipids. The manipulation of the enzyme activity of reaction 21 has its desired effects on ceramide accumulation and long chain base phosphate depletion but it does not change the complex sphingolipid content of the cell with respect to wild type, *i.e.* -22.42 %, -18.56 % and -8.85 % alterations in IPC, MIPC and $M(IP)_2C$'s concentrations, respectively (Table 3.5).

The drug targets proposed in this study are found to be more conserved than the other enzymes in the sphingolipid network (Appendix A: Table A.1).

3.3.1. Biological Meaning of the Potential Drug Targets

The first reaction catalyzed by palmitoyl-coenzyme A synthase, supplies the system with the fatty acid, Pal-CoA, which is needed for the very first step of *de novo* sphingolipid synthesis, *i.e.* condensation of Pal-CoA and serine in the endoplasmic reticulum to yield 3-ketodihydrosphingosine. The significance of the glycerol-3-phosphate acyltransferase enzyme catalyzing the second reaction comes from the fact that the second reaction is a branch point for the consumption of Pal-CoA. Pal-CoA can enter the *de novo* sphingolipid synthesis via reaction 11 or it can shift to the fatty acid metabolism via reaction 2. This enzyme is on the path of supplying substrates to the sphingolipid metabolism. The cell decides to enhance the production of either sphingolipids or fatty acids at this branch point. Cytidine diphosphate diacylglycerol (CDP-DAG) synthase (r_4) and phosphatidylinositol synthase (r_5) have roles in the phospholipids pathway and they serve as the source of PI required for the synthesis of complex sphingolipids. Their importance relies on the fact that they function on the path of complex sphingolipid synthesis. 4-Hydroxylase (r_{21}) is also proposed to be a potential drug target according to the results of metabolic control analysis; this reaction yields the conversion of DHS to PHS. In some literature resources (Obeid *et al.*, 2002; Ogretmen and Hannun, 2004; Olivera and Spiegel, 2001; Radin, 2003), it is reported that DHCer is not biologically active in mammalian cells, whereas the same

remark cannot be stated for *S. cerevisiae* at least with the experimental evidence available currently (Jiang *et al.*, 2004; Kraveka *et al.*, 2007; Tserng and Griffin, 2004) reporting the biological function of DHCer in apoptosis in human cancer cells. The significance of the conversion of DHS to PHS in *S. cerevisiae* may emphasize a more active role to PHS when compared with DHS just as in mammalian cells. Acetyl-coenzyme A synthetase is the enzyme catalyzing the 34th reaction, functioning in the fatty acid synthesis. The precursor fatty acid that is necessary for the production of the initiator fatty acid Pal-CoA of the sphingolipid pathway is produced by reaction 34. The significance and controlling effect of the enzyme is due to this reason.

Inositol-1-phosphate synthase functions as the catalyst of reaction 6 that supplies inositol to the sphingolipid metabolism. In the model system under investigation, reaction 6 is the only source for inositol and it was classified as a critically important enzymatic step. Inositol is essential for the synthesis of complex sphingolipids because it is the precursor metabolite of phosphatidyl inositol. Phosphoserine-phosphatase functions as the supplier of serine (r_9) is necessitated for the beginning of *de novo* sphingolipid production. Serine palmitoyl transferase enzyme catalyzing reaction 11 is one of the drug targets described in the literature (Reynolds *et al.*, 2004; Zeidan and Hannun, 2007), drugs targeting this enzyme are now at pre-clinical or clinical development stages. This step is also identified in this study as one of the significant control points that can be used in cancer therapy. The importance of this enzymatic step comes from the fact that it is the initial reaction of the *de novo* sphingolipid synthesis. 12th and 27th reactions are two exchange reactions which deplete the essential fatty acid needed for the initiation of *de novo* sphingolipid synthesis: Pal-CoA and PHCer, respectively. Since these reactions are the direct consumers of the two essential metabolites of sphingolipid metabolism, the two exchange reactions are specified and proposed to be candidates for being potential drug targets. Acyl-Coenzyme A binding protein is responsible for the exchange of Pal-CoA in reaction 12 and it is the GPI remodelase enzyme that executes reaction 27. Reactions 15 and 20 are catalyzed by sphingolipid long chain base kinase and responsible for the formation of DHS-P and PHS-P, respectively. These reactions are the ones that directly produce the phosphorylated forms of the long chain bases.

3.4. Concluding Remarks

When the two systems biology tools employed in the current study are compared in terms of their potency in potential drug target identification, it can be seen that metabolic pathway analysis results overlap more with the known drug targets in the literature. Metabolic control analysis results, on the other hand, are more informative since one can gather information about the direction of action, *i.e.* whether one should inhibit or activate the target enzyme. Thus, MCA results go one step further on the way to drug discovery and drug design.

In conclusion, the enzymes catalyzing the reactions 1, 2, 4, 5, 6, 9, 11, 12, 15, 20, 21, 23, 27 and 34 are proposed as putative drug targets by both MCA and MPA. The activities of palmitoyl-coenzyme A synthase (r_1), glycerol-3-phosphate acyltransferase (r_2), CDP-DAG synthase (r_4), phosphatidylinositol synthase (r_5), 4-hydroxylase (r_{21}), ceramide sythase (r_{23}) and acetyl-coenzyme A synthethase (r_{34}) should be increased to obtain the ceramide accumulation and long chain base phosphate depletion and the activities of Inositol-1-phosphate synthase (r_6), phosphoserine-phosphatase (r_9), serine palmitoyl transferase (r_{11}), acyl-coenyzme A binding protein (r_{12}), sphingolipid long chain base kinase (r_{15} and r_{20}) and GPI remodelase (r_{27}) should be decreased in order to reach the same cancer therapy essentials. Integration of the results of the two theoretical frameworks as we have done in this study can be utilized in future in other drug target identification studies.

In order to obtain more efficient drugs in cancer therapy; drugs targeting two or more enzymes may be developed or two or more drugs may be utilized at the same time course. There is clinical evidence of using drugs (Reynolds *et al.*, 2004; Zeidan and Hannun, 2007) both inhibiting ceramide catabolism and generating ceramide simultaneously.

4. STOICHIOMETRIC NETWORK RECONSTRUCTION AND ANALYSIS OF SPHINGOLIPID METABOLISM

The network of metabolic reactions of sphingolipid metabolism for both yeast and human are reconstructed and analyzed. The newly constructed stoichiometric networks are used to understand the governing mechanisms of sphingolipids in both organisms (Kavun-Ozbayraktar and Ulgen, 2011).

4.1. Stoichiometric Network Reconstruction and Analysis of Yeast Sphingolipid Metabolism Incorporating Different States of Hydroxylation

Sphingolipids are complex lipids found mostly in the plasma membrane and function as signaling molecules. Their structure is formed around a long chain sphingoid base backbone with an amide-linked fatty acid attached to it. Although there are numerous types of long-chain bases, sphingosine, dihydrosphingosine (DHS) and phytosphingosine (PHS) are the most widely encountered sphingoid bases in nature (Cuvillier and Levade, 2003); the latter two are also the major types of long chain bases of sphingolipid metabolism in *S. cerevisiae* (Dickson and Lester, 1999; Dickson and Lester, 2002).

The pathway starting with the rate limiting step of condensation of serine and palmitoyl-coenzyme A in the endoplasmic reticulum (serine palmitoyl transferase) and continuing up to dihydroceramide is called the *de novo* sphingolipid synthesis (Alvarez-Vasquez *et al.*, 2004; Alvarez-Vasquez *et al.*, 2005; Hannun *et al.*, 2001; Obeid *et al.*, 2002). The second subsection of the sphingolipid metabolism in *S. cerevisiae* is the phosphorylation of long chain bases (sphingosine kinase). This pathway is finalized by the breakdown of the phosphorylated long chain bases to fatty acid and phospholipid derivatives (sphingosine phosphate lyase). Another alternative for the phosphorylated long chain bases is the dephosphorylation back to DHS and PHS. The third and last part of the current sphingolipid pathway is the synthesis of complex sphingolipids taking place in the Golgi apparatus. There are three types of complex sphingolipids in *S. cerevisiae*, namely inositol phosphoryl ceramide (IPC), mannose-inositol phosphorylceramide (MIPC) and

mannose-diinositol phosphorylceramide ($M(IP)_2C$) (Dickson and Lester, 1999; Dickson and Lester, 2002; Dickson *et al.*, 2006; Kavun Ozbayraktar and Ulgen, 2009).

Sphingolipid-based therapeutics against cancer is based on the fact that sphingolipids regulate biological processes like cell death and survival playing important roles in cell's fate. There is a delicate balance between different types of sphingolipids since ceramide, sphingosine and sphingosine-1-phosphate are acting in opposing biological processes and therefore termed as sphingolipid rheostat (Spiegel and Milstein, 2002). Ceramide and sphingosine are known as tumor suppressor lipids inducing apoptosis whereas sphingosine-1-phosphate is known as tumor promoting lipid owing to its regulatory function in cell proliferation (Taha *et al.*, 2006; Zeidan and Hannun, 2007). The essential fact of cancer therapeutics with sphingolipid metabolism manipulation is that ceramide accumulation induces apoptosis only in cancerous cell lines, such as leukemia and breast carcinoma cells (Ogretmen and Hannun, 2002).

The pharmacological manipulation of the sphingolipid metabolism in cancer therapeutics necessitates the detailed understanding of the pathway and its underlying principles. Constraint-based modelling, which is one of the gains of genomic revolution, has been used as an identifier of critical reactions of the metabolic system of interest during drug discovery and development studies (Trawick and Schilling, 2006). In the present study, a metabolic model for the sphingolipid metabolism in *S. cerevisiae* coupled with the reactions of phospholipid and fatty acid metabolisms was developed taking five different states of sphingolipid hydroxylation into account, rendering it unique among other models. It is aimed to clarify the significance of hydroxylation on sphingolipids and hence to interpret the preferences of the cell between different metabolic pathway branches under different stress conditions. By the application of flux balance analysis (FBA), *in silico* gene deletion simulations were performed to elaborate the capabilities of the present metabolic network against several perturbations. Employing the computational systems biology tools (metabolic control and metabolic pathway analyses) we previously identified several candidate enzymes, to be used as drug targets in novel cancer therapy approaches (Kavun Ozbayraktar and Ulgen, 2010). We found that the enzymes encoded by the genes of *INO1*, *SER2, GUP1, ACB1, CDS1, LCB1*/*LCB2*, *LCB4*/*LCB5*, and *GAT1*/*GAT2* have important roles in achieving an increase in ceramide content of the cell and hence inducing

apoptosis. The metabolic consequences of the drug actions on these candidate enzymes were also investigated using the newly constructed *in silico* metabolic model which was validated against experimentally verified deletion phenotypes. The *in silico* model is further used in proposal of potential drug target enzymes, whose inhibition may induce apoptosis in cancerous cells. The possible flux ranges of each metabolic reaction were calculated by flux variability analysis and the flexibility of the network was then evaluated. The structural robustness of the selected reactions is studied using the concept of minimal cut sets.

4.2. Materials and Methods

Yeast (*S. cerevisiae*) is a good initial point for the investigation of mammalian sphingolipid metabolism as many yeast sphingolipid genes have homologs or orthologs in mammalian cells (Obeid *et al.*, 2002). In the present stoichiometric modelling of the *Saccharomyces cerevisiae* sphingolipid metabolism, 60 internal metabolic reactions and 77 intracellular metabolites are taken into account. The main mathematical tools employed are flux balance analysis, flux variability analysis and metabolic pathway analysis.

4.2.1. The Sphingolipid Metabolic Pathway

The metabolic pathway used in this study (Figure 4.1) is reconstructed by literature mining (Alvarez-Vasquez *et al.*, 2004; Alvarez-Vasquez *et al.*, 2005; BRENDA; Expasy; Guilllas *et al.*, 2001; Haak *et al.*, 1997; KEGG; Kohlwein *et al.*, 2001; Reggoiri and Conzelmann, 1998; Stunff *et al.*, 2002) and consists of fatty acid synthesis, phospholipid synthesis and sphingolipid synthesis reactions. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes are given in Appendix B, Table B.1 and the metabolite abbreviations list is given Table B.2. The genes encoding some of the important enzymes are given in italic capital letters next to the corresponding reactions in Figure 4.1.The metabolic steps other than the sphingolipid synthesis are the supporting reactions that are required for the beginning of *de novo* sphingolipid production and complex sphingolipid synthesis (Sims *et al.*, 2004). Phosphatidyl inositol (PI) that is necessary for the production of complex sphingolipids in *S. cerevisiae* is synthesized in the phospholipid metabolism. Fatty acid metabolism supplies two vital fatty acids, palmitoyl

coenzyme A (PalCoA) and very long chain fatty acid (in this study C_{26}CoA), required for the initiation of *de novo* sphingolipid synthesis and the synthesis of two main *S. cerevisiae* ceramides (dihydroceramide and phytoceramide), respectively. Complex sphingolipids synthesized from dihydroceramide, phytoceramide, and their hydroxylated states, are either sent to the plasma membrane (reactions 61-75, a pool designed for accumulation) or directed to further syntheses reactions (reactions 35-53) (Perry and Ridgway, 2005; Vallee and Riezman, 2005).

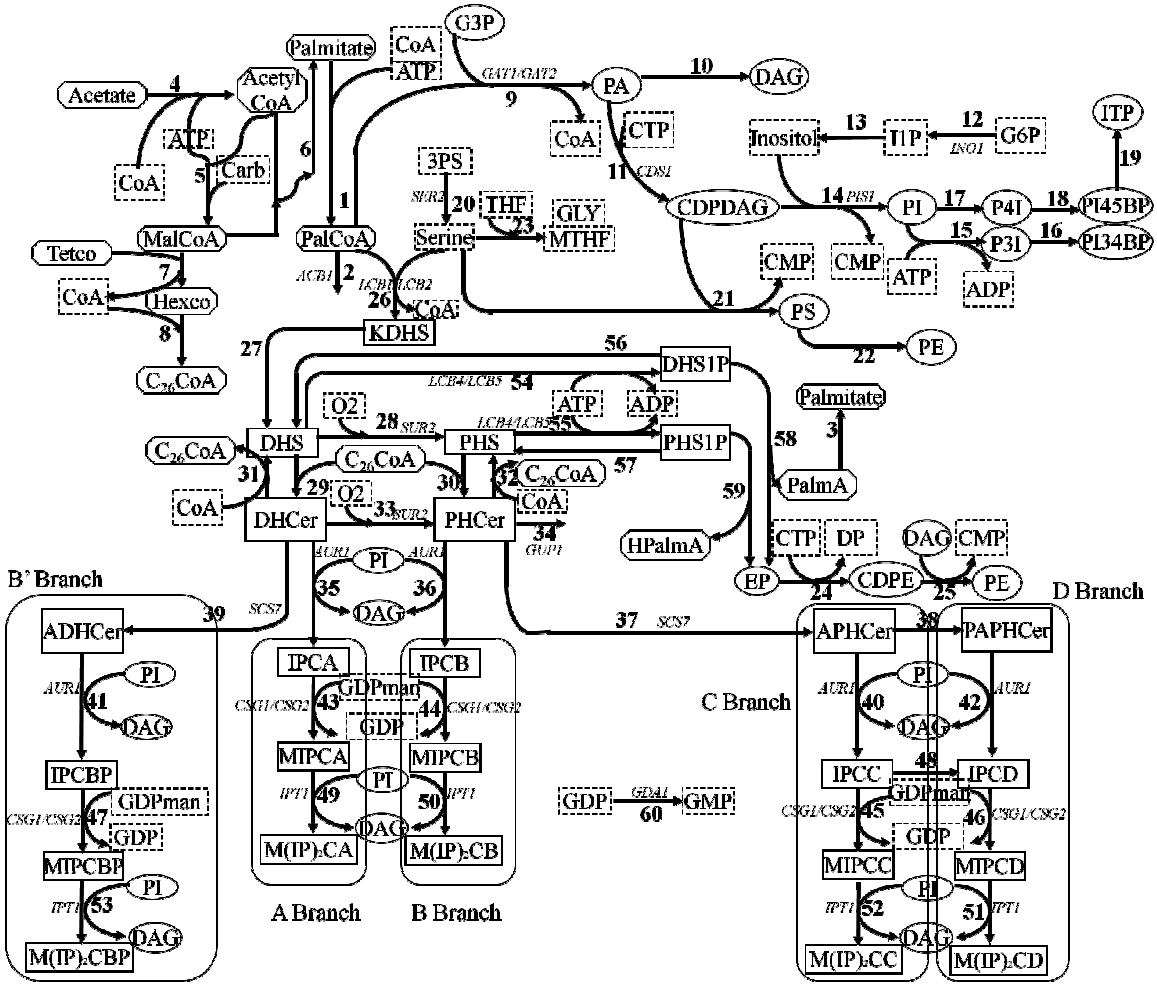


Figure 4.1. The rectangular boxes belong to the sphingolipid class. The oval boxes represent the phospholipids and the octagons correspond to the fatty acids. The dashed lined boxes are the cofactors and metabolites that do not fit to the above three categories

The state of hydroxylation is a very important factor in differentiation of complex sphingolipids. The identity, structure and location of sphingolipids are all affected by the state of hydroxylation (Dunn *et al.*, 1998; Haak *et al.*, 1997). In order to see the metabolic

effects of hydroxylation products on the sphingolipid content of the cell, five states of hydroxylation are included in the model system: 1) Hydroxylation of dihydrosphingosine to phytosphingosine (reaction 28) 2) Hydroxylation of dihydroceramide to phytoceramide (reaction 33) 3) Hydroxylation of dihydroceramide leading to B' branch of complex sphingolipid synthesis (reaction 39) 4) Hydroxylation of phytoceramide leading to C branch of complex sphingolipid synthesis (reaction 37) and 5) Further hydroxylation of alpha-hydroxyphytoceramide leading to D branch of complex sphingolipid synthesis (reaction 38).

In order to perform flux balance analysis and flux variability analysis, the reactions must be balanced stoichiometrically. Three genome scale models (Duarte *et al.* 2004; Forster *et al.*, 2003; Kuepfer *et al.*, 2005) and the databases of BRENDA, Expasy, KEGG, were utilized in order to get the complete stoichiometry of the reactions. The genes and the enzymes catalyzing the reactions, Enzyme Commision (EC) numbers and many other information were gathered from the databases.

4.2.2. Flux Balance Analysis (FBA)

The metabolic capabilities of the current sphingolipid system is interpreted using flux balance analysis (Bonarius *et al.*, 1997; Fell and Small, 1986; Kaufmann *et al.*, 2003; Stephanopoulos *et al.*, 1998) based on the physicochemical constraints. A distinct metabolic flux distribution as a representative of the metabolic phenotype is achieved by the use of a biologically significant objective function among the set of possible distributions (Varma and Palsson, 1994). Using the steady state assumption, the mass balance equations written around each metabolite reduce to Eq. (4.1) which builds the basis of flux balance analysis (Varma and Palsson, 1994).

$$S \cdot v = 0 \tag{4.1}$$

S is the stoichiometric matrix of m (number of metabolites) rows and n (number of reactions) columns. The unknown fluxes of the metabolic network are stored in the vector v of dimension n. The vector of unknown metabolic reaction rates include all the internal fluxes and transport fluxes. For such biological systems the number of metabolic reactions

usually surpasses the number of metabolites, *i.e.* underdetermined system. Therefore, one has to restrict multiple solutions by introducing other constraints to the mass balance equations, such as inequality constraints: $0 \le v_i \le \infty$ for irreversible reactions, $-\infty \le v_i \le \infty$ for reversible reactions. The constraints have also been applied to the transport reactions of the uptake metabolites, which have been taken from experimental measurements and the dynamic model of Alvarez-Vasquez *et al.* (2004 and 2005), (Table 4.1).

In order to achieve a distinct flux distribution representing the related phenotype dictated by the genotype, an objective function is defined by Eq. (4.2) (Varma and Palsson, 1994; Edwards and Palsson, 2000a and 2000b; Ramakrishna *et al.*, 2001).

$$\begin{aligned} &\text{objective function : minimize/maximize } Z = \sum c_i v_i \\ &\text{such that} \\ &S \cdot v = 0 \\ &(l.b.)_i \le v_i \le (u.b.)_i \end{aligned} \tag{4.2}$$

where c_i denotes the weights of the fluxes forming the objective function.

Table 4.1. The constraints used in the flux balance and flux variability analyses (Alvarez-Vasquez *et al.*, 2004; Alvarez-Vasquez *et al.*, 2005)

Reaction #	**Reaction name**	**Constraints (μM/min)**
1	Palmitoyl-CoA synthase	59
79	Uptake of ATP	60-120
80	Uptake of 3-phosphoserine	0-40

In both individual and combined objective function trials, the linear programming of MOSEK optimization software (MOSEK version 5.0.0.57) working under MATLAB R2007a was used to solve the optimization problem (Appendix F).

For the selection of appropriate objective function to be used in FBA, the simulation results need to be validated by comparison with the experimental phenotype information.

As experimental flux data for sphingolipid related reactions have not been available, a dynamic model was reconstructed using the equations given in Alvarez-Vasquez *et al.* (2004 and 2005). The ordinary differential equations of the dynamic model were solved by the stiff differential equation solver of MATLAB R2007a with the variable order method. Moreover, the lumped reactions in the dynamic model system of Alvarez-Vasquez *et al.* (2004 and 2005) were decomposed to deal with the elementary reactions. The steady state flux values of both models (dynamic and stoichiometric) were normalized with respect to the flux of reaction 1, that is the synthesis of palmitoyl-coenzyme A from palmitate, and these normalized flux values were plotted with respect to each other for comparison. The slope and the goodness of fit (R-square) of the line fitted to these data points were examined and the closeness to 1 was used as the criterion for the determination of the most appropriate objective function to be used in further simulation studies.

In order to eliminate the multiplicity of resultant flux distributions (FBA), a second objective function (minimization of the Euclidean norm of fluxes) following the first one was used. Flux minimization reproduces the cell's attitude towards the efficient use of metabolites with minimal effort (Bonarius *et al.*, 1996; Cakir *et al.*, 2007; Holzhutter, 2004). The solution of the second optimization problem was attained by the quadratic programming of MOSEK optimization software (MOSEKversion 5.0.0.57). Among the multiple solutions of the first optimization, the minimum flux distribution is selected by means of second optimization algorithm depicted above (Appendix F).

4.2.3. Flux Variability Analysis

In order to identify the ranges of each reaction satisfying the same quantitative objective function value, flux variability (Mahadevan and Schilling, 2003) of each metabolic step was calculated at the chosen first objective function. Two sets of optimization problems (Eq. 4.3) are subsequently solved to achieve the upper and lower bound values of each reaction fluxes. Each metabolic step is maximized/minimized one at a time with fixed values of the objective function obtained from flux balance analysis and its corresponding flux value establishes the upper/lower limit of the corresponding reaction's range.

$$\begin{aligned} &\text{objective function : maximize/minimize } v_i \\ &\text{such that} \\ &S \cdot v = 0 \\ &f^T v = Z \\ &(l.b.)_i \le v_i \le (u.b.)_i \text{ for i = 1 to n} \end{aligned} \tag{4.3}$$

where f is the objective function vector, n is the number of reactions and Z is the objective function value from the previous flux balance analysis study. The range of each reaction is employed to determine the robustness of the system (Appendix F).

4.2.4. Minimal Cut Sets and Robustness Analysis

The concept of minimal cut sets is used for calculating the structural robustness of the network functions as well as for checking the objective function used in flux balance analysis and the criteria used in proposing new drug targets. Minimal cut sets introduce all the possible knock-out strategies for the failure mode of a selected network function (*e.g.* ATP consumption). The elementary flux modes of the network were calculated using CellNetAnalyzer version 9.4. (Klamt *et al.*, 2007), and then minimal cut sets and network robustness were determined. The fragility coefficients and robustness scores are calculated in Microsoft Excel. Fragility coefficient is the average reciprocal of the average size of all minimal cut sets in which each reaction participates. Robustness score is the complement of fragility score to 1 (Klamt, 2006).

4.3. Results and Discussion

In this study, it is aimed (i) to discover the metabolic effects of gene deletions in sphingolipid metabolism via a small-scale stoichiometric model of *S. cerevisiae* and (ii) to determine the preferences of the cell between different branches of ceramide hydroxylation under several stress conditions. In order to validate the present reconstructed stoichiometric model, the phenotypes of the deletions of numerous genes were simulated and the results were compared with the experimental findings reported in literature as explained below.

4.3.1. Flux Distribution of Wild Type Strain

The flux distribution of the wild type strain was obtained after the two sequential optimizations (linear programming with the maximization of $M(IP)_2C$-from C and D branches- production and quadratic programming with the minimization of all fluxes as the objective functions). The first step of stoichiometric modelling is the determination of the objective function since there is no a priori information about it. In order to gain knowledge about the effect of each metabolite on the system objective function, the production/consumption rates of all the constituent metabolites are both maximized and minimized one by one as the objective functions of the sphingolipid metabolism using a simplified version of the metabolic pathway in order to compare our results with the dynamic model of Alvarez-Vasquez *et al.* (2004 and 2005). The simplified version of the metabolic *in silico* model only discriminates between dihydroceramide and phytoceramide, disregarding further hydroxylation reactions due to limitations of the dynamic model. This simplified version of the *in silico* model was only used for the determination of the objective function, thereafter the extended version was utilized in further simulation studies. A subset of the objective functions producing possible and mathematically plausible flux distributions without any thermodynamically infeasible futile cycles in the solution space was determined; the subset consists of the individual minimization of production of dihydrosphingosine-1-phosphate, phytoshingosine-1-phosphate, 3-ketodihydrosphingosine, dihydrosphingosine, phytosphingosine, hexasanoyl-coenzyme A, malonyl-coenzyme A and uptake of glycerol-3-phosphate, serine; and maximization of IPC synthesized from dihydroceramide only, MIPC and $M(IP)_2C$ synthesized from both dihydroceramide and phytoceramide, and finally plasma membrane forms of MIPC and $M(IP)_2C$ synthesized again from both sources of ceramides. All possible pairwise combinations from this set of objective functions were also tested: *e.g.* minimum G3P uptake and maximum production of $M(IP)_2C$ synthesized from both dihydroceramide and phytoceramide. The steady state flux distributions (solution of the stoichiometric model via flux balance analysis with two successive optimizations) are compared with the dynamic model results using the slope and R^2 values of fitted lines in Microsoft Excel (see the computational methods for details). The objective function, resulting in the highest similarity with the dynamic model is the maximum production rates of $M(IP)_2C$ (from both

dihydroceramide and phytoceramide). $M(IP)_2C$ is the most complex and abundant sphingolipid of *S. cerevisiae.*

$M(IP)_2C$ synthesis produces two molecules of DAG with a growth stimulatory signal, rendering its synthesis an appreciated objective by the cell (Schneiter, 1999). The experimental findings from the literature (Dunn *et al.*, 1998) revealed that IPC-C and IPC-D are the only available IPC species in wild type *S. cerevisiae* strains. This fact signifies the maximum synthesis of $M(IP)_2C$ from alpha-hydroxyphytoceramide (C branch of hydroxylation) and further hydroxylation product of alpha-hydroxyphytoceramide (D branch of hydroxylation). Hence, it is a biologically relevant objective function for the *in silico* sphingolipid metabolic model of yeast.

Palmitoyl-Coenzyme A synthesis from palmitate (reaction 1: pioneering reaction of *de novo* sphingolipid biosynthesis) was specified as 59 μM/min according to literature flux data (Alvarez-Vasquez *et al.*, 2004; Alvarez-Vasquez *et al.*, 2005). The ATP consumption of the cell can be noted from Fig. 2 as 118 μM/min. The wild type strain representing the normal (healthy) case synthesizes 19.7 μM/min of complex sphingolipids and transports them to the plasma membrane. The plasma membrane sphingolipid content consists of 52.6 % $M(IP)_2C$-C, from the first hydroxylation of phytoceramide and 47.4 % $M(IP)_2C$-D, from the second hydroxylation branch of phytoceramide; it can be concluded that the wild type strain predominantly utilizes alpha-hydroxyphytoceramide to synthesize complex sphingolipids. This finding is also verified by the experimental data reported in the literature (Haak *et al.*, 1997; Dunn *et al.*, 1998). The growth inhibitory phytosphingosine and phytoceramide do not accumulate (reaction 34) and the growth stimulatory DAG is synthesized at a rate of 39.3 μM/min in the *in silico* model.

The flux distribution in Fig. 2 reveals that, the *in silico* wild type cell prefers to use phytosphingosine to further synthesize complex sphingolipids. This simulation result of *in silico* wild type strain is in agreement with several literature sources. Saba *et al.* (1996) and Mandala (2001) pointed to the phytosphingosine preference of *S. cerevisiae* cells in ceramide synthesis. The inhibition of growth by phytoceramide (but not by dihydroceramide) in yeast may be considered as another supportive information for the preference of yeast cells towards phyto-species. Phytoceramide-induced apoptosis

encourages the use of yeast phytoceramide as the counterpart of mammalian ceramide in cancer therapy (Hwang *et al.*, 2001; Mandala, 2001).

When all the deletion mutants are examined from the perspective of long chain base preference of the cell, 51 out of 70 deletion mutants studied prefer to use hydroxylation products of phytoceramide instead of dihydroceramide in the synthesis of complex sphingolipids. For example, the in silico double mutant *sur2Δcsg2Δ*, when supplied with exogenous dihydrosphingosine does not uptake it, but synthesizes only a small amount by *de novo* process and converts it to IPC-A and IPC-BP with the total plasma membrane sphingolipid content of 3.0636 μM/min. However, the total plasma membrane sphingolipid content of the *in silico sur2Δcsg2Δ* mutant with free exogenous phytosphingosine uptake is 20.3230 μM/min. So, these simulation results agree with the statement of Haak *et al.* (1997): The alterations due to *sur2Δ* deletion are corrected by the addition of phytosphingosine but not dihydrosphingosine. Eventually, we can conclude that the present *in silico* model correctly predicts the phyto- preference of yeast cells.

4.3.2. Validation of the *in silico* Model by Deletion Mutant Studies

As the first step in the validation of the *in silico* model of yeast sphingolipid metabolism, we selected the deletion mutants with the experimentally available phenotypes, *i.e.* single mutations of *sur2Δ*, *scs7Δ*, *gda1Δ* and double mutations of *scs7Δsur2Δ* and *gda1Δscs7Δ*. Moreover, the single mutation of *csg2Δ*, the double mutations of *sur2Δcsg2Δ*, *scs7Δcsg2Δ,* and finally a triple mutation of *sur2Δscs7Δcsg2Δ* which were also employed for the validation purpose will be discussed under the subsection of calcium sensitivity.

Sur2p is responsible for the first state of hydroxylation, *i.e.,* conversion of dihydrosphingosine to phytosphingosine and conversion of dihydroceramide to phytoceramide. Thus, *sur2Δ* mutant lacks the ability of synthesizing phytosphingosine and phytoceramide. The plasma membrane sphingolipids of the *in silico* mutant consist of IPC-A, $M(IP)_2C$-A and IPC-BP. The *in silico* mutant cell compensates the lack of phytosphingosine derivatives by using dihydrosphingosine and dihydroceramide. Half of

the dihydrosphingosine synthesized *de novo* (reaction 27) is transferred to fatty acid and phospholipid metabolisms via reaction 58 catalyzed by long chain base-phosphate lyase. The second hydroxylation state that is synthesis of alpha-hydroxydihydroceramide from dihydroceramide and synthesis of alpha-hydroxyphytoceramide from phytoceramide is catalyzed by Scs7p. *In silico scs7Δ* mutant generates and transports IPC-A, IPC-B, $M(IP)_2C$-A and $M(IP)_2C$-B to the plasma membrane, since alpha-hydroxylation of both types of ceramides is missing. Abnormal sphingolipid accumulation is reported for *SCS7* gene deletion in *Saccharomyces* Genome Database (SGD), which is in agreement with our *in silico* modelling results. *In silico gda1Δ* mutation lacks the ability to convert IPC to MIPC since it cannot metabolize the product (GDP) of mannosylation step. Instead of mannosylation, the mutant synthesizes and delivers all types of IPC, mainly IPC-A and IPC-B, to the plasma membrane. Approximately, half of the (55 %) *de novo* produced long chain bases are phosphorylated and directed to other lipid metabolisms in the cell. The plasma membrane sphingolipid contents of the *in silico* single mutants are in agreement with the experimentally reported data (Dunn *et al.*, 1998; Haak *et al.*, 1997; Schneiter, 1999).

The double mutations of the genes mentioned above (*scs7Δsur2Δ*, *gda1Δscs7Δ*) were also examined *in silico*. The *in silico scs7Δsur2Δ* double mutant accumulates IPC-A and its derivative $M(IP)_2C$-A in the plasma membrane since the *in silico* mutant cell cannot synthesize phytosphingosine and phytoceramide due to the lack of *SUR2* gene and it cannot hydroxylate the synthesized dihydroceramide to obtain BP-series of complex sphingolipids due to the lack of *SCS7* gene. The phenotypic characateristics of *sur2Δ* in the sphingolipid metabolism is also observed in the *in silico* double deletion mutant; almost 50 % of the synthesized dihydrosphingosine is unused in complex sphingolipid synthesis and phosphorylated to be directed to other lipid pathways. Finally, *in silico gda1Δscs7Δ* double deletion mutant accumulates IPC-A and IPC-B in the plasma membrane. The simulation results of these double mutations are consistent with the experimental findings of Haak *et al.* (1997) and Dunn *et al.* (1998) validating our stoichiometric model.

4.3.3. Calcium Sensitivity in Sphingolipid Metabolism

The deletion mutants with experimentally verified phenotypic behaviors related to calcium sensitivity were found from literature and simulated for the purpose of additional validation. *In silico csg2Δ* mutation (lack of mannose-inositol phosphorylceramide (MIPC) synthase activity) causes the accumulation of mainly IPC-C and also IPC-D in the plasma membrane. There are no MIPC and $M(IP)_2C$ syntheses in the *in silico* mutant and that is also verified with the mutant phenotypes reported in SGD. The accumulated IPC-C is the reason that makes the mutant cells calcium-sensivitive (Beeler *et al.*, 1998; Dunn *et al.*, 1998; Haak *et al.*, 1997). The *in silico* mutant cell utilizes phytosphingosine and phytoceramide instead of dihydro-series to build up complex sphingolipids just as the wild type strain. 57 % of the long chain bases synthesized *de novo* in the mutant cell are directed to other lipid metabolisms via reactions 58 and 59 catalyzed by the long chain base-phosphate lyase enzyme.

Using the fact of IPC-C accumulation leading to calcium sensitivity as the driving force, the deletion mutants that accumulate IPC-C in the plasma membrane of the *in silico* cell are examined. The *in silico* mutants of *ipt1Δ* and *gda1Δ* accumulate IPC-C in the plasma membrane. These mutant strains may also be calcium sensitive due to IPC-C accumulation.

Suppression of calcium sensitivity: *CSG1* and *CSG2* genes encode MIPC synthase, so the mutant cells cannot synthesize the most abundant mature sphingolipid of yeast, $M(IP)_2C$ due the absence of the precursor metabolite MIPC. Beeler *et al.* (1998) stated that calcium sensitivity of the *CSG1* and *CSG2* deletion mutants can be suppressed by any mutation that reduces IPC-C accumulation. All the possible double and triple deletion mutations of *csg* mutants are simulated *in silico* and the resulting flux distributions are examined for reduction in IPC-C accumulation in the plasma membrane.

The accumulation of IPC-C causing calcium sensitivity in *csg2Δ* mutants is eliminated by *sur2Δ* in the *sur2Δcsg2Δ* double mutants whose simulation results show that mainly IPC-A and also IPC-BP are synthesized and delivered to the plasma membrane. As a characteristic of the *sur2Δ,* all the dihydrosphingosine synthesized cannot be utilized in

the complex sphingolipid synthesis, and almost half of it is transferred to other lipid synthesizing pathways, *i.e.,* fatty acid and phospholipid metabolisms. When supplied with exogenous phytosphingosine, the double deletion mutant restores the complex sphingolipids' flux distribution of *csg2Δ* mutant, eliminating the effect of *SUR2* gene deletion and its phenotypic characteristics in terms of sphingolipid metabolism. The second *in silico* double deletion mutant studied is *scs7Δcsg2Δ*, where the *in silico* mutant cell accumulates IPC-A and IPC-B in the plasma membrane, (similar to *in silico gda1Δscs7Δ* double mutant), but not IPC-C and as a result no calcium sensitivity are observed. Since mannosylation step is not executed, approximately 55 % of the unused long chain bases (both dihydrosphingosine and phytosphingosine) are phosphorylated and oriented to fatty acid and phospholipid pathways.

CSG2 gene deletion coupled with any of the following gene deletions of *LAC1/LAG1/LIP1* (ceramide synthase) and *AUR1* (IPC synthase) in addition to *SUR2* (hydroxylase) and *SCS7* (ceramide hydroxylase) mentioned above; or coupled with the blockage of any of the following reactions, *i.e.,* reaction 1 (synthesis of palmitoyl-Coenzyme A from palmitate), reaction 4 (synthesis of acetyl-Coenzyme A from acetate), reaction 5 (synthesis of malonyl-Coenzyme A from acetyl-Coenzyme A), reaction 7 (synthesis of hexacosanoate from malonyl-Coenzyme A), reaction 8 (synthesis of C_{26}-Coenzyme A from hexacosanoate), reaction 9 (synthesis of phosphatidic acid from palmitoyl-Coenzyme A), reaction 11 (synthesis of CDP-DAG from phosphatidic acid), reaction 14 (synthesis of phosphatidyl inositol from inositol), reaction 20 (synthesis of serine from 3-phosphoserine), reaction 26 (condensation of serine and palmitoyl-Coenzyme A), reaction 27 (*de novo* synthesis of dihydrosphingosine), reaction 37 (alpha-hydroxylation of phytoceramide) and reaction 40 (synthesis of IPC-C from alpha-hydroxyphytoceramide) have no IPC-C accumulation in the plasma membrane (See Appendix B, Table B.1 for gene names). A further investigation of the mutants showed that *in silico sur2Δcsg2Δ, scs7Δcsg2Δ* double mutants and *CSG2* gene deletion mutant lacking either hydroxylation of phytoceramide (reaction 37) or synthesis of IPC-C from alpha-phytoceramide (reaction 40) hinder IPC-C accumulation without complete shutdown of complex sphingolipid synthesis. These blocked reactions/genes consequently help to eliminate the calcium sensitivity of *CSG2* gene deletion mutants without affecting the viability of the mutants (IPCs are essential sphingolipids). Eventually, we can propose

these four mutants as suppressors of calcium sensitivity of *csg2Δ* mutants. Haak *et al.* (1997) also reported *sur2Δcsg2Δ, scs7Δcsg2Δ* double mutations as eliminators of calcium sensitivity, which validates our modelling results. *In silico CSG2* gene deletion mutant lacking hydroxylation of phytoceramide (reaction 37) synthesizes IPC-A, IPC-B and IPC-BP and transports them to the plasma membrane; and *CSG2* gene deletion mutant that cannot synthesize IPC-C from alpha-phytoceramide (reaction 40) synthesizes only IPC-D. *In silico csg2Δ* mutants that cannot synthesize IPC-D either from IPC-C via reaction 48 or from alpha-hydroxyphytoceramide via reactions 38 and 42 accumulate 50 % more IPC-C than *csg2Δ* mutant itself, so they are more susceptible to calcium sensitivity.

The triple mutant of *sur2Δscs7Δcsg2Δ* accumulates IPC-A in the plasma membrane of the *in silico* cell. 50 % of the dihydrosphingosine synthesized *de novo* is not used in the complex sphingolipid synthesis since the mannosylation step is missing due to the lack of the MIPC synthase enzyme. The unused long chain base is phosphorylated and directed to phospholipid and fatty acid metabolisms. When the *in silico* triple mutant is supplied with exogenous phytosphingosine, the mutant can also synthesize some IPC-B, but it cannot utilize all the supplied phytosphingosine in the IPC synthesis, instead it selects to use 58 % of the exogenously supplied phytosphingosine in the other lipid pathways via reaction 59. Unlike the other *in silico* deletion studies performed here, 13 % of the exogenous phytosphingosine is utilized in the synthesis of phytoceramide and it is directed to glycosylphosphatidylinositol (GPI)-remodelling. The results of the single, double and triple mutations of *CSG2* genes simulated here are consistent with the experimental findings of Haak *et al.* (1997) and Dunn *et al.* (1998).

4.3.4. *In silico* Drug Actions

The genes encoding the enzymes proposed as the potential drug targets in Chapter 3 (Kavun Ozbayraktar and Ulgen, 2010) were taken as the basis in selecting the deletion mutants to be studied *in silico* in this section. In Kavun Ozbayraktar and Ulgen (2010) inositol-1-phosphate synthase (*INO1,* reaction 12), acyl-coenzyme A binding protein (*ACB1,* reaction 2), sphingosine kinase (*LCB4/LCB5* reaction 54 and 55), phosphoserine phosphatase (*SER2,* reaction 20), CDP-DAG synthase (*CDS1,* reaction 11), glycerol-3-phosphate acyltransferase (*GAT1* and *GAT2,* reaction 9), and GPI remodelase (*GUP1,*

reaction 34), were identified as the enzymes that should be manipulated in order to cause the accumulation of ceramides and the depletion of long chain base phosphates, that is, to induce apoptosis. Serine palmitoyl transferase (over-expression of *LCB1/LCB2* encoded enzyme) and sphingosine kinase (inhibition of *LCB4/LCB5* encoded enzyme) are commercial drug targets and have been in clinical trial stages of drug development. In order to understand the metabolic consequences of enzyme inhibitions, the deletion mutants of the genes encoding these drug target enzymes were simulated *in silico*. The flux distributions were compared with that of the wild type strain and some of the significant results were discussed below.

In silico INO1 deletion mutant: The deletion of *INO1* gene blocks the conversion of glucose-6-phosphate to inositol-1-phosphate (reaction 12) which is the necessary metabolite for the production of inositol. Inositol can also be transported from the extracellular space into the cell, so the metabolic synthesis of inositol from glucose-6-phosphate is not the only inositol source for the use of cellular processes. When external inositol uptake is allowed, the *in silico* wild type strain prefers to transport almost all the inositol required for the cell from the extracellular space, consequently, *in silico ino1Δ* mutant gives almost the same flux distribution as the wild type strain. This information is in agreement with the experimental finding of *INO1* gene deletion resulting in a viable strain (SGD). When the wild type strain is not supplied with exogenous inositol, the *in silico* cell synthesizes the necessary inositol from inositol-1-phosphate (reaction 13) at an amount similar to that transported from the extracellular space (when inositol is available to the cell). It can be concluded that the source of inositol is equally accessible to the cell.

In order to investigate the effects of complete absence of inositol, the external transport of inositol into the cell is also hindered in the *in silico INO1* deletion mutant. Inositol is a versatile biological compound that has important regulatory effects in both its own synthesis and phospholipid pathway (Shaldubina *et al.*, 2002). In the case of *in silico INO1* deletion mutant without external inositol transport, *S. cerevisiae* cells cannot produce the basic phospholipid (phosphatidyl inositol) required for the synthesis of *S. cerevisiae*-specific complex sphingolipids (IPC, MIPC, $M(IP)_2C$) and the fluxes of the phospholipid pathway (reactions 9, 11, 14, 15, 16, 17, 18, 19, and 20) are decreased with respect to those of wild type (Figure 4.2 and Table 4.2).

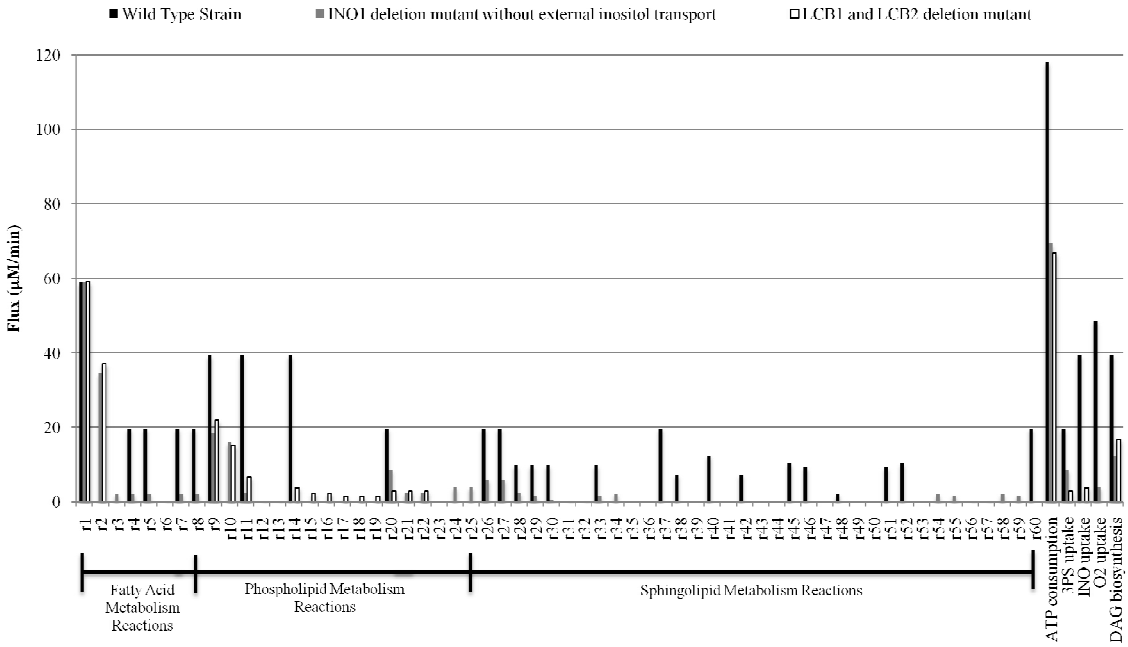


Figure 4.2. Flux (μM/min) distributions of wild type strain, *INO1* deletion mutant without external inositol transport and *LCB1*/*LCB2* double deletion mutant

Moreover, the *de novo* sphingolipid synthesis pathway does not reach its end products, *i.e.* complex sphingolipids, but ceases at the level of dihydroceramide and phytoceramide. The *de novo* synthesis of dihydrosphingosine (reaction 27) decreased to 30.7 % of its wild type value in the *in silico* mutant without exogenous inositol. In contrast to the wild type strain, the *in silico* cell does not utilize all the dihydrosphingosine and dihydroceramide to synthesize phytosphingosine and phytoceramide, but 40.7 % of the *de novo* synthesized dihydrosphingosine is hydroxylated to phytosphingosine (reaction 28). 35.8 % of the *de novo* synthesis is used in the generation of phytoceramide from both hydroxylation of dihydrosphingosine and then conversion to phytoceramide via reactions 28 and 30, and direct hydroxylation of dihydroceramide to phytoceramide via reaction 33. The synthesis of phytoceramide without any utilization in the formation of complex sphingolipids due to the absence of inositol, leads to its own accumulation (reaction 34) and further direction to GPI remodelling. GPI remodelling is activated when there is no inositol available to the cell. It can be concluded that *INO1* deletion along with the elimination of external inositol transport leads to a shift from complex sphingolipid synthesis to glycosylphosphatidylinositol (GPI)-anchored proteins synthesis.

Table 4.2. Metabolic fluxes (μM/min) of hydroxylation products of ceramide and important branch points of both fatty acid and phospholipid metabolisms belonging to the *in silico INO1* deletion mutant without external inositol transport and wild type strain

Reaction Name	**Wild Type (μM/min)**	***INO1* deletion mutant without external inositol transport (μM/min)**
IPC-A synthesis (reaction 35)	0	0
IPC-B synthesis (reaction 36)	0	0
IPC-C synthesis (reaction 40)	12.4273	0
IPC-D synthesis (reaction 42)	7.2492	0
IPC-B' synthesis (reaction 41)	0	0
C_{26}-CoA synthesis (reaction 8)	19.6765	2.1613
PA synthesis, fore-branch of phospholipid reactions (reaction 9)	39.353	18.4935
PI synthesis (reaction 14)	39.353	0
de novo sphingolipid synthesis (reaction 26)	19.6765	6.0399

The phosphorylation of long chain bases (both dihydrosphingosine and phytosphingosine) takes place in the *in silico* mutant without exogenous inositol (reactions 54 and 55) and their breakdown products are further utilized in fatty acid (palmitaldehyde) and phosholipid (ethanolamine phosphate) metabolisms (reactions 58 and 59). Eventually, it can be concluded that the decrease in the amount of inositol, causes an increase in the rate of phosphorylation of dihydrosphingosine and phytosphingosine with respect to the wild type strain. Finally, one can notice the decrease in the synthesis of fatty acids due to the decreased ceramide synthesis when inositol is completely missing (approximately 89 % flux decrease in reaction 4, 5, 7 and 8).

When inositol is not available, serine is not consumed in *de novo* sphingolipid metabolism as much as the wild type strain consumes but instead it is utilized in the phospholipid metabolism starting with reaction 21 and ends up in phosphatidyl ethanolamine, which is not available in the wild type strain. As a consequence of decreased sphingolipid synthesis, there is 41 % decrease in the ATP consumption of the *in silico* mutant cell when compared with that of the wild type strain. ATP is the cell's valid energy currency in most of the cases; a considerable decrease in ATP consumption arose the question about the abnormalities that may take place in the *in silico* cell. Without exogenous inositol the DAG consumption of the *in silico* mutant cell decreased 69 % with respect to wild type flux and that may be another evidence for the problematic growth, when the growth stimulatory effect of DAG and lack of essential sphingolipid IPC are taken into consideration (Schneiter, 1999). Dominguez *et al.* (1978) reported abnormal functioning of yeast cells under inositol deprivation, and it can be concluded that the experimental findings reported in literature overlaps with the present *in silico* modelling results.

In silico LCB4 and *LCB5* double deletion mutant: In order to examine the effect of inactive sphingosine kinase in response to drug action, *in silico LCB4* and *LCB5* double deletion mutant is simulated by blocking the reactions of phosphorylation of long chain bases, both dihydrosphingosine and phytosphingosine (reactions 54 and 55, respectively). Since there is no phosphorylation and subsequent dephosphorylation of the long-chain bases, the degradation of long-chain base phosphates are also eliminated consequently (reaction 58 and 59). According to the metabolic information gathered through the the current *in silico* model of yeast sphingolipid metabolism, *in silico* wild type strain does not phosphorylate long chain bases, as a result of this, long chain base-phosphate phosphatase and lyase enzymes' activities do not appear in the flux distribution. So, the double deletion mutant of *LCB4* and *LCB5* genes gives almost the same flux distribution as the wild type strain in agreement with experimentally investigated viability of the mutant (SGD). The second objective function (minimization of all fluxes) utilized in flux balance analysis may have caused the inactivity of the phosphorylation of long chain bases branch in the *in silico* wild type strain. The *in silico* cell prefers the direct usage of the *de novo* synthesized long chain bases in the biosynthesis of complex sphingolipids (the first objective function of flux balance analysis) without subsequent phosphorylation and dephosphorylation

processes before the integration of exogenous long chain bases into the synthesis of complex sphingolipids as depicted in Mao *et al.* (1997). The requirement of the cell for the phosphorylation of long chain bases followed by dephosphorylation may be due to the regulatory effects which are not taken into consideration in our current model system. Another reason for the inactivity of phosphorylation and dephosphorylation in the *in silico* wild type strain would be the necessity of phosphorylation of long chain bases and subsequent dephosphorylation for the integration of exogenous long chain bases do not apply for the *de novo* synthesized long chain bases. The long chain base kinase, long chain base-phosphate phosphatase and lyase enzymes become active under stress conditions (*i.e.,* in several *in silico* deletion mutants).

In silico GAT1 and *GAT2* double deletion mutant: *In silico GAT1* and *GAT2* (encoding glycerol-3-phosphate acyltransferase) double deletion mutant does not synthesize any of the complex sphingolipids due to the lack of phosphatidyl inositol (PI). The mutant has a defective phospholipid metabolic pathway, and consequently, complete of shutdown of complex sphingolipid and DAG biosyntheses coupled with 35 % decrease in ATP consumption were observed with respect to the wild type strain. Phytoceramide accumulation leading to the induction of apoptosis is the most significant finding of the *in silico* deletion simulation which is partially in agreement with the decreased competitive fitness of *GAT2* gene reported in literature (SGD).

In silico CDS1 deletion mutant: *CDS1* deletion mutant is another *in silico* strain with a defective phospholipid pathway. Due to the lack of PI, complex sphingolipid biosynthesis does not take place; *de novo* sphingolipid pathway ceases at the level of phytoceramide which is then accumulated via reaction 34. Phytoceramide accumulation along with 40 % decrease in ATP consumption and 66 % decrease in DAG biosynthesis of the *in silico* cell with respect to the wild type strain supports the inviability of the mutant (SGD).

In silico LCB1 and *LCB2* double deletion mutant: Reynolds *et al.* (2004) and Modrak *et al.* (2006) reported that serine palmitoyl transferase enzyme is a potential drug target in cancer therapy, working with the principle of over-expression. In order to see the metabolic consequences of its inhibition, it is simulated by the blockage of the flux passing

through condensation of serine and palmitoyl-Coenzyme A (reaction 26). But since this enzyme is the starting point of the *de novo* sphingolipid synthesis and does not have any alternatives, the sphingolipid pathway of the model system is totally hindered. Since palmitoyl-Coenzyme A cannot be integrated into the sphingolipid pathway, the flux is directed to the other Acyl-Coenzyme A consuming processes represented by reaction 2 and phospholipid pathway. As it can be seen from Fig. 2, the fluxes of phospholipid pathway reactions (reactions 10, 15, 16, 17, 18, 19, 21 and 22) increased due to the shutdown of sphingolipid pathway. These reactions belong to inositol phosphate biosynthesis that were inactive in the *in silico* wild type strain, which invested inositol in phosphatidyl inositol synthesis to be used in complex sphingolipid generation. The fluxes of fatty acid production (reaction 4, 5, 6, 7 and 8) ceased since there is no consumption via ceramide synthesis. The same rationale is valid for the unused serine; less serine is synthesized via reaction 20 and it is purged to the synthesis of phosphatidyl serine and phosphatidyl ethanolamine (reactions 21 and 22). The metabolic flux through reactions 20-22 is 2.8355 μM/min for *lcb1*Δ*lcb2*Δ mutant showing a significant decrease in serine production (85.6 %) when compared to that of the wild type. 43.5 % decrease in ATP consumption, and 57.4 % decrease in DAG biosynthesis in the *in silico* mutant cell coupled with the absence of IPC are in agreement with the reported inviability of the *lcb1Δlcb2Δ in vitro* double mutant (SGD).

It should be noted here that the drug target enzyme (encoded by *INO1* gene) identified by the dynamic modelling approach integrated with regulatory effects in Chapter 3 (Kavun Ozbayraktar and Ulgen, 2010) is also confirmed by the present stoichiometric modelling approach.

4.3.5. Metabolic Effects of Gene Deletions of Sphingolipid Metabolism

In order to see the metabolic effects of all gene deletions and to use the results in proposing novel potential drug targets, we blocked all the reactions in the *in silico* model one at a time, and obtained the corresponding flux distribution after two sequential optimizations. These calculations simulate the case where a specific inhibitor for each reaction is introduced into the system. It is known that one gene may encode an enzyme

that can catalyze numerous reactions, so; in order to simulate such gene deletions, one has to block all the reactions that are catalyzed by that specific enzyme.

Sphingolipid metabolism related viability: The first complex sphingolipid of *S. cerevisiae*, IPC is essential for the cell (Schneiter, 1999). If IPC is missing in the cell, the rest of the complex sphingolipids cannot be synthesized; so, the complete absence of complex sphingolipids is considered as a criterion for problematic growth/inviability (SGD). If any one of the complex sphingolipids is available, we conclude that the deletion mutant is viable. Since the *in silico* stoichiometric model deals with steady state composition of the sphingolipids, it cannot discriminate the time-dependent distribution of sphingolipids. Therefore, it is assumed that if a mature sphingolipid exists, IPC also exists.

The sphingolipid content of the plasma membrane were examined for all the deletion mutants simulated in the current study. On Table 4.3, the deletion phenotypes of sphingolipid genes taken from *Saccharomyces* Genome Database (SGD) are tabulated and the complex sphingolipids in the *in silico* cell found by the model are given in the last column. *In silico* *ino1Δ* mutant without exogenous inositol, *lcb1Δlcb2Δ, tsc10Δ, gat1Δgat2Δ, cds1Δ, ser2Δ, aur1Δ and lac1Δlag1Δlip1Δ* mutants do not have complex sphingolipids transported to the plasma membrane. The viability/growth information of core sphingolipid metabolism genes is in agreement with SGD mutant phenotypes (Table 4.3). Besides, complete absence of mature sphingolipids were observed in *in silico* mutants when any of the reactions 1 (synthesis of palmitoyl-Coenzyme A from palmitate), 4 (synthesis of acetyl-Coenzyme A from acetate), 5 (synthesis of malonyl-Coenzyme A from acetyl-Coenzyme A), 7 (synthesis of hexacosanoate from malonyl-Coenzyme A), 8 (synthesis of C_{26}-Coenzyme A from hexacosanoate), 9 (synthesis of phosphatidic acid from palmitoyl-Coenzyme A), 11 (synthesis of CDP-DAG from phosphatidic acid), 14 (synthesis of phosphatidyl inositol from inositol), 20 (synthesis of serine from 3-phosphoserine), 26 (condensation of serine and palmitoyl-Coenzyme A) and 27 (*de novo* synthesis of dihydrosphingosine) were blocked (See Appendix B, Table B.1 for gene names and Table B.4 for total complex sphingolipid content of the deletion mutants). This finding also indicates growth problems or inviability of these mutants. Viability is a very important criterion in proposing novel potential drug targets. For some of the above-

mentioned mutants, the lack of complex sphingolipids is due to defective *de novo* sphingolipids synthesis, for example, *ser2Δ,* and *lcb1Δlcb2Δ* mutants.

Table 4.3. The deletion phenotypes of sphingolipid genes taken from *Saccharomyces* Genome Database (SGD) along with the *in silico* complex sphingolipids content

Genes	**Deletion Phenotype**	**Mutant Phenotype**	**Reactions**	**Total *in silico*Sphingolipid Content of the Plasma Membrane**
LAC1	Viable		29 30	no complex sphingolipids
LAG1	Viable			
LIP1	Inviable			
AUR1	Inviable	inositol phosphoceramide accumulation: decreased ceramides accumulation: increased	35 36 40 41 42	no complex sphingolipids
LCB1	Inviable		26	no complex sphingolipids
LCB2	Inviable			
TSC3	Inviable			
TSC10	Inviable		27	no complex sphingolipids
SUR2	Viable	polyphosphate accumulation: abnormal	28 33	IPC-A $M(IP)_2C$-A IPC-BP
YDC1	Viable	apoptosis: increased ceramides accumulation: decreased	31	$M(IP)_2C$-C $M(IP)_2C$-D
YPC1	Viable		32	$M(IP)_2C$-C $M(IP)_2C$-D
SCS7	Viable	sphingolipid accumulation: abnormal	37 39	IPC-A IPC-B $M(IP)_2C$-A $M(IP)_2C$-B $M(IP)_2C$-C

Table 4.3. The deletion phenotypes of sphingolipid genes taken from *Saccharomyces* Genome Database (SGD) along with the *in silico* complex sphingolipids content (cont.)

Genes	Deletion Phenotype	Mutant Phenotype	Reactions	Total *in silico* Sphingolipid Content of the Plasma Membrane
CCC2	Viable		38 48	$M(IP)_2C$-C
CSG1	Viable			
CSG2	Viable	mannosylinositol phosphorylceramide (MIPC) accumulation: decreased mannosyldiinositol phosphorylceramide (M(IP)2C) accumulation: decreased inositol phosphorylceramide (IPC) accumulation: increased	43 44 45 46 47	IPC-C IPC-D
IPT1	Viable		49 50 51 52 53	IPC-A IPC-B IPC-C IPC-D IPC-BP MIPC-A MIPC-B
LCB4	Viable	sphingosine 1-phosphate accumulation: absent	54 55	$M(IP)_2C$-C $M(IP)_2C$-D
LCB5	Viable	glycogen accumulation: decreased		
LCB3	Viable	resistance to sphinganine: increased	56 57	$M(IP)_2C$-C $M(IP)_2C$-D
LBP2	Viable			
DPL1	Viable	sphingosine 1-phosphate accumulation: increased phosphorylated sphingoid bases accumulation: increased resistance to phytosphingosine: decreased resistance to D-erythro-sphingosine: decreased	58 59	$M(IP)_2C$-C $M(IP)_2C$-D

10 (underlined mutations) out of 18 *in silico* mutants, that do not synthesize any complex sphingolipids, integrate the unused long chain bases into ethanolamine phosphate, *i.e.,* to glycerophospholipid synthesis via phosphorylation by long chain base kinase (reactions 54 and 55) and then degradation by long chain base-phosphate lyase (reactions 58 and 59). This is a one-way route that interrelates the sphingolipid turnover to glycerophospholipid synthesis (Schneiter, 1999). Moreover, when there is a significant decrease in total complex sphingolipid synthesis (more than 85 % with respect to wild type), the unused long chain bases are also directed to glycerophospholipid metabolism via

same route. That situation occurred in *in silico* deletion mutants of *gda1Δ, sur2Δ, scs7Δ* and *ipt1Δ* and when the reaction 37 (hydroxylation of phytoceramide) is blocked.

One common property of these above-mentioned deletion mutants with either complete shutdown or highly reduced synthesis of complex sphingolipids is the significantly low ATP consumption (34-49 % decrease in ATP consumption, see Appendix B, Table B.4).

The *in silico* mutants that consume significantly less ATP when compared to wild type strain, also direct palmitoyl-coenzyme A which is unused by the sphingolipid metabolism, to other acyl-coenzyme A consuming processes. Palmitoyl-coenzyme A may follow three metabolic routes in the current stoichiometric model: 1) Direct integration to *de novo* sphingolipid synthesis by the condensation of serine and palmitoyl-coenzyme A (reaction 26), 2) Indirect integration to the complex sphingolipid pathway via conversion to phosphatidyl inositol in the phospholipid metabolism (reactions 9, 10, 11 and 14), and 3) Transfer to other acyl-coenzyme A consuming process (reaction 2). The same set of deletion mutants listed above undergoes 53.5 - 100 % reduction in diacylglycerol (DAG) biosynthesis with respect to wild type strain. Since DAG is a growth stimulating signaling molecule with anti-apoptotic properties (Gomez-Munoz, 1998; Schneiter, 1999), the decrease in its generation with respect to wild type strain is a critical change that should be taken into consideration. *In vitro* experiments of *ser2Δ, gda1Δ* and *sur2Δ* mutants showed decreased competitive fitness of the mutant cells (SGD) confirming our stoichiometric model results. The significant reduction in ATP consumption, DAG biosynthesis, complex sphingolipid synthesis (or its complete absence) and increased transfer of palmitoyl-Coenzyme A to reaction 2 are all coupled with each other and lead to either inviability or growth problems (See Appendix B, Table B.4).

DAG versus ceramide; new potential drug targets: Diacyglycerol (DAG), phosphatidic acid (PA), ceramides, long chain bases and their phosphorylated forms are all second messengers synthesized as effectors of intracellular responses to extracellular stimuli. The delicate balance among the lipid second messengers controls the cell's fate by switching between cell proliferation and death. DAG and PA have regulatory roles in the stimulation of cell proliferation via activating protein kinase C and mitogen activated

protein kinase, respectively. Ceramide and sphingosine induce cell death, whereas sphingosine-1-phosphate counteracts to eliminate their apoptotic effects. The anti-mitogenic signals of ceramide are predominated by the mitogenic signals of DAG meanwhile DAG does not nullify consequences of other signaling events. The protection of healthy cells from death in cancer therapy may be provided by the struggle between ceramide and DAG (Gomez-Munoz, 1998; Mathias *et al.*, 1998; Ohanian and Ohanian, 2001a). The understanding of these mechanisms may provide new targets to be used in therapeutic intervention.

The *in silico* model has four different routes of DAG biosynthesis: (i) Hydrolysis of phosphoinositol 4, 5-biphosphate (PI45BP) to DAG and inositol 1, 4, 5-triphosphate (ITP) via reaction 19, (ii) Conversion of palmitoyl-Coenzyme A to DAG over phosphatidic acid (PA) through reactions 9 and 10, (iii) IPC and DAG synthesis from ceramides through reactions 35, 36, 40, 41 and 42, and (iv) $M(IP)_2C$ and DAG synthesis from MIPC through reactions 49-53. The first route yields two messengers: DAG and ITP. ITP increases cytosolic calcium levels and regulates cellular functions through interaction with calmodulin kinases. The second path of DAG biosynthesis does not involve an increase in calcium level and its consequences last longer than the first path. It can be concluded that different DAG biosynthesis paths end up in the regulation of different signaling mechanisms with different consequences emphasizing the significance of the source of DAG (Gomez-Munoz, 1998; Ohanian and Ohanian, 2001a). The *in silico* wild type strain prefers to synthesize DAG from IPC and $M(IP)_2C$ (routes iii and iv). However, the deletion mutants with reduced DAG biosynthesis, synthesize DAG mostly from reaction 10 (route ii), which is a more sustained route without any accompanied calcium release. These *in silico* deletion mutants also have reduced *de novo* synthesis of long chain bases, coupled with reduced PA/DAG ratio varying between 0.2-0.6 when compared with *in silico* wild type strain (See Appendix B, Table B.4). This fact is in agreement with the experimental findings of Gomez-Munoz (1998) who reported that increased sphingosine content leads to increased PA/DAG ratio due to significant reduction in DAG and corresponding increase in PA.

Yeast phytoceramide, but not dihydroceramide, inhibits proliferation and induces apoptosis. The arise of phytoceramide induced cell death makes the sphingolipid an

economical alternative to be used in cancer therapy (Hwang *et al.*, 2001). Another information that may help to propose novel potential drug targets for cancer therapy is the accumulation of phytosphingosine which also leads to growth inhibition in yeast (Schneiter, 1999). When the *in silico* cell accumulates phytoceramide and/or phytosphingosine, it cannot integrate them in the biosynthesis of complex sphingolipids but it directs them to GPI-remodelling, *i.e.,* reaction 34. The mutants, in which reaction 9 (*GAT1/GAT2,* glycerol-3-phosphate acyltransferase), reaction 11 (*CDS1,* CDP-DAG synthase) and reaction 14 (*PIS1,* phosphatidyl inositol synthase) are blocked, and *aur1Δ* mutant and *ino1Δ* mutant without exogenous inositol have more flux directed through reaction 34 to accumulate phytoceramide and/or phytosphingosine. Ceramide accumulation of *aur1Δ* mutant is verified with the mutant phenotype information reported in *Saccharomyces* Genome Database (SGD). Besides, these mutants have reduced consumption of ATP and synthesis of DAG coupled with the absence of complex sphingolipids. All the metabolic properties that guide the *in silico* cell to growth inhibition/apoptosis, which is a required property in cancer therapy, are available in five deletion mutants mentioned above. They have DAG/phytoceramide ratios varying between 0-6, which can be considered low with respect to that of wild type strain (See Appendix B, Table B.4). In proliferating cells of human T cell line, on the other hand, an increase was observed in the ratio of DAG/ceramide (Ohanian and Ohanian, 2001b).

4.3.6. Flexibility of the System and Flux Variability Analysis (FVA)

To get a deep insight as well as to observe the effect of variability on flux distribution and to determine the possible range of each flux with the flexibility of the system, the flux variability analysis was performed for the potential drug target enzymes proposed in this study (using the objective function of maximum $M(IP)_2C$-C and $M(IP)_2C$-D synthesis with the constraints used in flux balance analysis). The fluxes are classified according to their activity in the metabolism as ‘never been used’, ‘having constant flux values’ or ‘having variable flux values’ (Vo *et al.*, 2004).

98.4 % of the reactions in wild type strain have variable fluxes and 0.8 % of the reactions are ‘never been used’ (Table 4.4). The variable fluxes of the wild type strain reduce the restrictions of the solution space and make the wild type strain most flexible

among all simulations. *IPT1, GDA1*, *CSG* and *AUR1* gene deletions have the highest percentage of variable fluxes among the deletion mutants listed in Table 4.4, 86.4 %, 72.0 %, 71.2 % and 56.8 %, respectively (See Figure 4.3 for flux variation of deletion mutants). In other gene deletion mutants almost half of the pathway becomes idle. *INO1* gene deletion without any inositol transport into the cell, *GAT1/2, CDS1* and *PIS1* gene deletions render the *in silico* cell rigid due to high percentages of un-used reactions, 52.0 %, 66.4 %, 54.4 % and 52.0 %, respectively. Serine palmitoyl transferase (encoded by *LCB1* and *LCB2* genes) is the pioneering enzyme of the sphingolipid metabolism without any bypassing branches; therefore, its blockage increases the number of reactions never been used (56 %), *i.e.*, almost half of the reactions are unused. The low number of variable fluxes (43.2 %) in this double mutant points out that this gene deletion makes the system more rigid and idle, and hence prevents the efficient functioning of the system. Over-expression (but not inhibition) of serine palmitoyl transferase enzyme, on the other hand, is targeted in cancer therapy approaches.

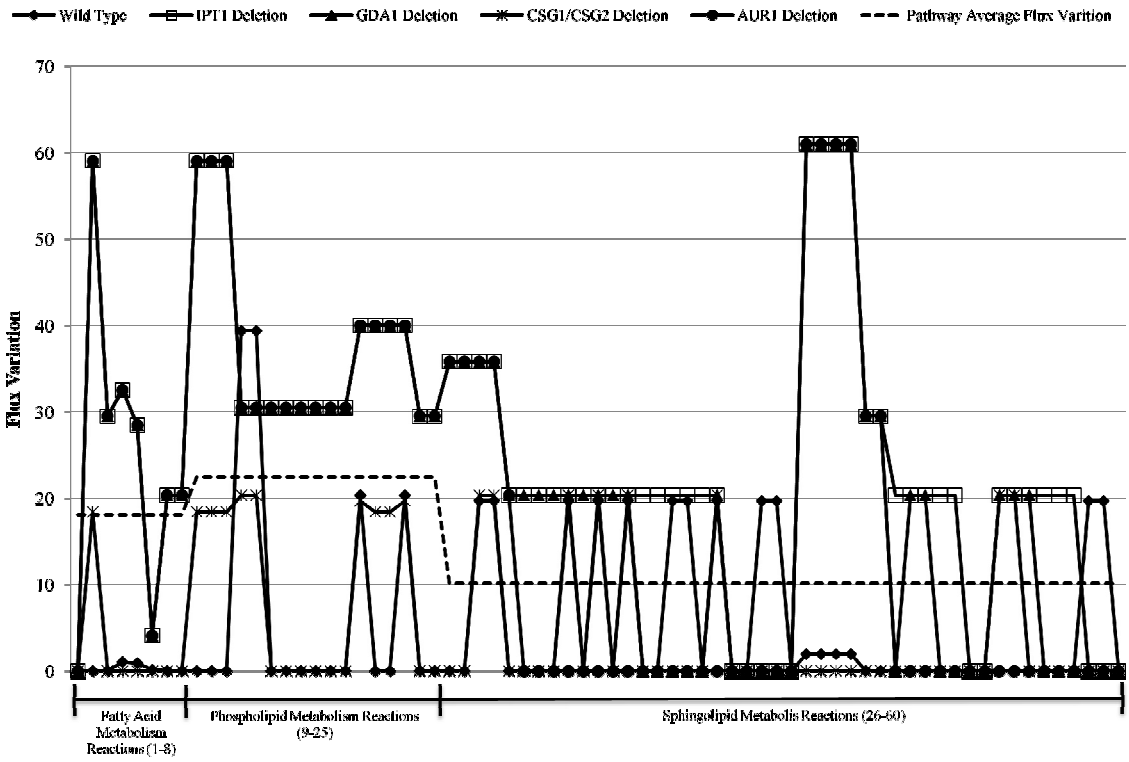


Figure 4.3. The ranges of flux variation (μM/min) for five simulations. The reaction order is according to the list given in Table B.1. The average flux ranges for all 10 simulations listed in Table 4.4 are also calculated and shown with a dashed line (---)

Table 4.4. Flux variability analysis results of nine potential drug target enzymes proposed according to flux balance analyses results

Mutations/ Blocked Reactions	% never used	% constant non-zero flux value	% variable flux values
Wild Type	0.8	0.8	98.4
IPT1 Deletion (Reactions 49-53)	12.8	0.8	86.4
GDA1 Deletion (Reaction 60)	27.2	0.8	72.0
SUR2/CSG2 Deletion with exogenous phytosphingosine (Reactions 28, 33 and 43-47)	33.3	2.4	64.3
CSG1/CSG2 Deletion (Reactions 43-47)	27.2	1.6	71.2
AUR1 Deletion (Reactions 35, 36 and 40-42)	42.4	0.8	56.8
INO1 Deletion without inositol transport (Reaction 12)	52.0	0.8	47.2
GAT1/GAT2 Deletion (Reaction 9)	66.4	0.8	32.8
CDS1 Deletion (Reaction 11)	54.4	0.8	44.8
PIS1 Deletion (Reaction 14)	52.0	0.8	47.2

Next, the average value of flux variation according to all the deletion mutations and wild type strain are calculated to examine the most flexible/variable part of the metabolic network in terms of subpathways. The phospholipid metabolism has the highest range of flux variation whereas the sphingolipid metabolism has the smallest range, *i.e.,* it is the most rigid part (Figure 4.3). The rigidity of the sphingolipid metabolism may be due to the studied gene deletions mostly affecting the sphingolipid biosynthesis reactions.

The internal reactions that have variable flux values are further investigated in order to identify the common ones among the deletion mutants and wild type strain. These common reactions include almost all the fatty acid synthesis reactions (reactions 2 and 4-8)

and the reactions functioning in the serine metabolism (reactions 20 and 23). The most flexible part of the sphingolipid metabolism for all *in silico* deletion mutants and wild type strain consists of *de novo* synthesis of dihydrosphingosine (reactions 26 and 27), both types of ceramide synthesis (reactions 29 and 30), and their breakdown (reactions 31 and 32) and finally phosphorylation and subsequent dephosphorylation of both long chain bases (reactions 54-57) (See Appendix B, Table B.3 for details).

Human ceramide synthase (*LASS1-6*) and ceramidase (*ACER1-3*, *ASAH1-2* and *ASAH2B-C*) enzymes are reported as sphingolipid based drug targets for cancer and their drugs are at the stages of clinical trials (Reynolds *et al.*, 2004; Zeidan and Hannun, 2007). In the sphingolipid metabolism of *in silico S. cerevisiae*, the syntheses of dihydroceramide and phytoceramide (reactions 29 and 30, respectively) are catalyzed by a type of ceramide synthase and the reverse reactions (reactions 31 and 32) are catalyzed by ceramidases. It can be concluded that the drug target enzymes under clinical trials catalyze the most flexible reactions in the metabolic network.

The inhibition action of a drug on a target enzyme should not leave the whole system idle; but the flexibility of the system must rather be protected. A high flexibility makes the system under investigation more stable and resistant to stress conditions increasing the range of solution space. In that case, the feasible solutions of the metabolic network span more of the solution space. In the light of the above information, one can conclude that *IPT1, GDA1, CSG* and *AUR1* gene deletions might be better candidates for drug targeting in terms flux variability results.

4.3.7. Minimal Cut Sets and Robustness Scores

The yeast sphingolipid metabolism has 5464 elementary flux modes. These modes are further used to calculate minimal cut sets and finally robustness scores. The structural robustness of five branches of complex sphingolipids with different states of hydroxylation were calculated and compared with each other (Table 4.5). The most robust (close to 1) branch is the biosynthesis of IPC-C and IPC-D with robustness scores of 0.69 and 0.68, respectively. This finding supports our objective function selection in flux balance analysis

and the wild type yeast cell's preference towards C and D branches of complex sphingolipid synthesis.

Table 4.5. Results of minimal cut set calculations and list of structural robustness scores

Target Reaction	# of Elementary Flux Modes (EFM) (over 5464)	# of Minimal Cut Sets (MCS)	# of reactions involved in EFMs selected	Network Fragility Coefficient	Network Robustness Score
Reaction 35: IPC-A biosynthesis	808	227	146	0.3551	0.6449
Reaction 36: IPC-B biosynthesis	908	313	141	0.3612	0.6388
Reaction 40: IPC-C biosynthesis	2261	7999	165	0.3046	0.6954
Reaction 41: IPC-B' biosynthesis	582	291	142	0.3676	0.6324
Reaction 42: IPC-D biosynthesis	1210	5131	141	0.3190	0.6810
Reaction 142: PHCer accumulation PHCerxt ->	281	5435	165	0.2605	0.7395
Reaction 159: DAG biosynthesis DAGxt ->	1702	15311	165	0.1154	0.8846
Reaction 163: ATP consumption -> ATPxt	5459	28768	169	0.0529	0.9471

The reduced DAG biosynthesis and ATP consumption are the criteria used in potential drug target proposal. These two common properties of potential drug targets lead to the robustness scores of 0.88 and 0.95 (both values are very close to 1), respectively

(Table 4.5). It can be concluded that our criteria for selecting drug targets have high structural robustness. Finally, the procedure of the methodology utilized in the present work is depicted schematically on Figure 4.4.

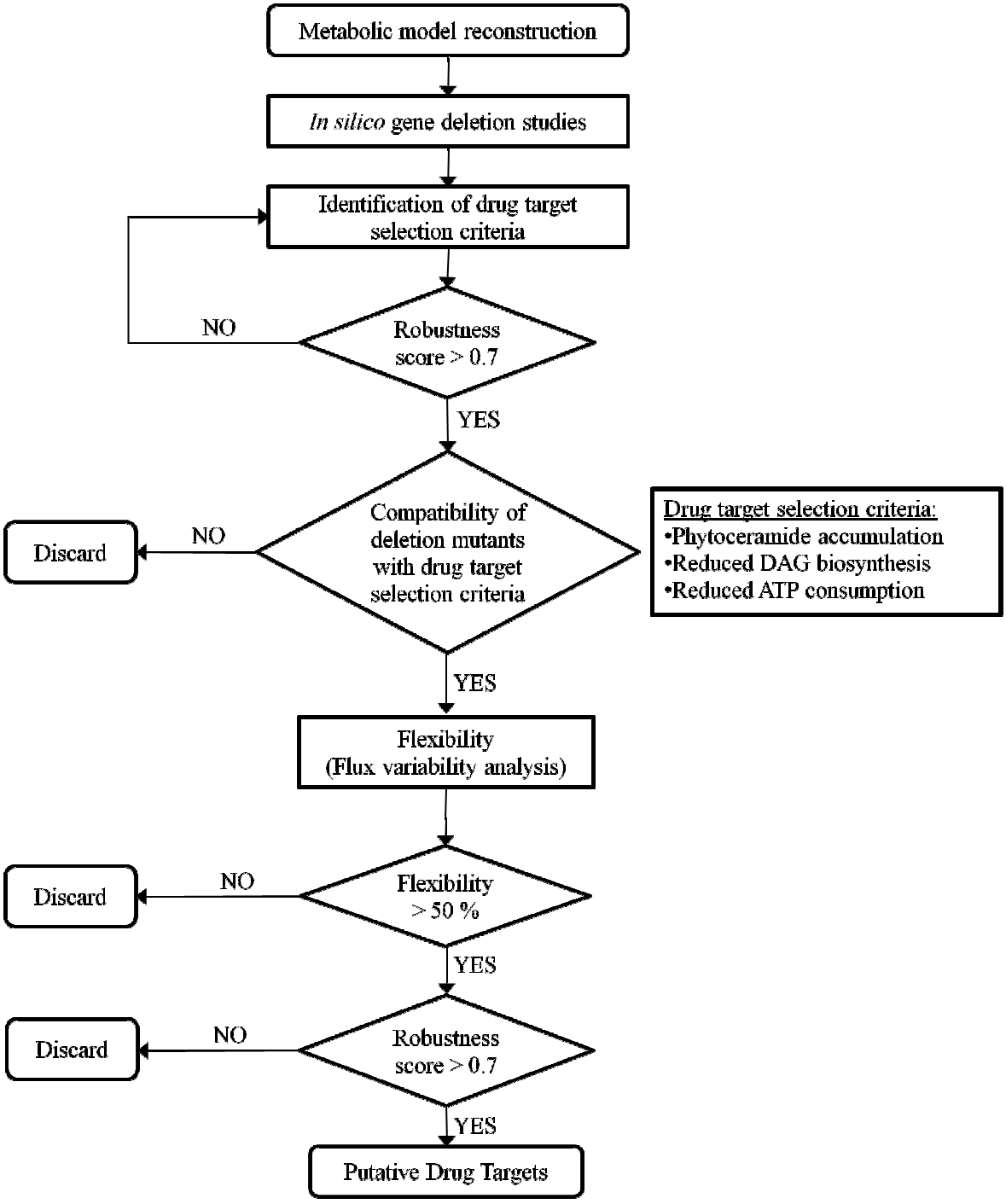


Figure 4.4. Summary of potential drug target selection algorithm

4.4. Concluding Remarks

The regulational aspects of sphingolipids make the sphingolipid metabolism an appropriate target for cancer therapeutics. The delicate metabolic balance between tumor-suppressor and tumor-promoter lipids plays a very important role in cell's fate. A simple organism, *S. cerevisiae*, is selected as a model organism for the *in silico* analysis and the first comprehensive stoichiometric model of yeast sphingolipid metabolism is reconstructed. Although there are distinct variations between yeast and human complex sphingolipids, the current study constitutes the first attempt for unraveling the sphingolipid metabolism in detail. Five states of hydroxylation are taken into consideration in the reconstructed metabolic model, elucidating the significance of hydroxylation on sphingolipids. Hydroxylation products of phytoceramide are found to be the most favored species in the biosynthesis of complex sphingolipids which is in agreement with the experimental findings depicted in several literature sources (Mandala, 2001; Saba *et al.*, 1996). The phytoceramide derived complex sphingolipid syntheses also constitute the most robust part of the metabolic model. Phytoceramide (counterpart of mammalian ceramide) accumulation, reduced DAG biosynthesis and ATP consumption and reduced complex sphingolipid synthesis (since IPC is essential for the viability of cell) are the metabolic criteria utilized to specify the conditions that are either necessary to induce apoptosis or valid during apoptosis. The validated *in silico* model along with these criteria is used to propose potential drug targets for inducing cell death. When the results of the *in silico* drug action simulations are integrated with the flexibility of the system, *IPT1, GDA1, CSG* and *AUR1* gene deletions might be novel and appropriate candidates for drug targeting in cancer therapy. Hence, our findings may guide experimentalist in experiment design with beneficial starting points.

4.5. Stoichiometric Network Reconstruction and Analysis of Human Sphingolipid Metabolism

The pathway of human sphingolipid metabolism is reconstructed by collecting the reactions from several papers, as well as from KEGG database (Figure 2.3) (KEGG; BRENDA; Ogretmen and Hannun, 2004). In human sphingolipid biosynthesis pathway a double bond is added to dihydroceramide to obtain biologically active ceramide. Ceramide

is one of the precursors of complex sphingolipids in human. Human cells do not add inositol phosphate and mannose to ceramide but add hydrophilic head groups, such as galactose, phosphatidylcholine, and glucose during the synthesis of complex sphingolipids.

Ceramide is converted to sphingomyelin by the addition of phosphatidylcholine via sphingomyelin synthase enzyme (EC 2.7.8.3), mainly in the luminal surface of the Golgi membrane, but the enzymatic activity of the synthase is also observed in the plasma membrane. *SMS1* and *SMS2* genes are responsible for the encoding of the sphingomyelin synthase enzyme. Sms1p is located in the *trans* Golgi region, whereas the location of Sms2p is uncertain (Hanada *et al.*, 2007). Sms2p can function as an isozyme in the Golgi apparatus, or apart from the *de novo* synthesis, it may convert the ceramide produced in the plasma membrane to sphingomyelin.

Galactosylceramide is another complex sphingolipid synthesized by the galactosyltransferase enzyme (EC 2.4.1.47) that attaches a galactose to the sphingoid long-chain base backbone. Glucosylceramide synthase (EC 2.4.1.80) converts ceramide to glucosylceramide mainly in the Golgi apparatus (*cis* Golgi region), and glucosylceramide can be further metabolized to lactosylceramide by the lactosylceramide synthase enzyme. Besides the Golgi apparatus, endoplamic reticulum and microsome also contain the glucosylceramide synthase enzyme. The first conversion to glucosylceramide takes place on the cytosolic face of the Golgi membrane, whereas the second conversion to lactosylceramide takes place on the luminal face of the Golgi apparatus (BRENDA; Hanada *et al.*, 2007; Perry and Ridgway, 2005; Kolter and Sandhoff, 2006). The breakdown of complex sphingolipids back to ceramide also occurs in human.

Furthermore, ceramide can be broken down to sphingosine via ceramidase enzyme, which has numerous forms. *ASAH1, ASAH2, ASAH2B, ASAH2C, ASAH3* and *ASAH3L* are some of the genes encoding this enzyme in human. The phosphorylation and dephosphorylation of ceramide, sphingosine and dihydrosphingosine occur via kinase and phosphatase enzymes, respectively.

Glycosphingolipids, as the name implies, consists of a sugar added to the ceramide existing as a substrate in this concept. Glycosphingolipids exist in a membrane of the cell,

where the hydrophobic ceramide is attached to the extracytoplasmic leaflet of the membrane, exposing its hydrophilic sugar content to the extracytoplasmic space. There are alternative pathways producing glycosphingolipids: either from *de novo* synthesis or from salvage/recycling pathways in endosome. In a slowly dividing cell, salvage pathway is more active whereas in a rapidly dividing cell, de novo pathway becomes more active (Maccioni, 1999). Lacto-series and neo-lacto-series glycosphingolipid biosynthesis roots from lactosylceramide. Most of the differentiation reactions involve the activity of either a galactosyltransferase or a fucosyltransferase. Ganglio-series glycosphingolipids also diverges from lactosylceramide. Four main pathway branches starting with the synthesis of precursors (GA_2, GM_3, GD_3 and GT_3) constitute ganglioside biosynthesis. Glucosylation step constitutes the rate-limiting step in ganglioside biosynthesis, which takes place on the cytosolic surface of the Golgi compartment. Galactose added gangliosides are mostly found in nerve tissue and they function in central nervous system (Bektas and Spiegel, 2004). Finally, globo-series glycosphingolipid biosynthesis starts with the conversion of lactosylceramide to globotriaosylceramide and then to globoside. Certain types of enzymes that can accept different substrates in different cell types and development stages are responsible for glycosphingolipid diversity (Kolter *et al.*, 2002).

Although we do not have as much information on the function of glycosphingolipids as we have on ceramide, it is known that they also have critical functions in pathological processes. Gangliosides GM_3 and GD_3 are melanoma-associated antigens that have functional roles in metastasis. Glucosylceramide functions in multidrug resistance in cancer cells. Reduced GD_3 synthase activity prevents apoptosis. This finding on GD_3 points out a similar relation of ceramide and apoptosis for GD_3-apoptosis (Bektas and Spiegel, 2004). Ganglioside GM_1 has anti-apoptotic effects just like sphingosine-1-phosphate. Cavallini (1999) reports that GM_1 enhances sphingosine-1-phosphate synthesis by activating sphingosine kinase and it is also known that exogenous GM_1 have anti-apoptotic effects. Due to these effects GM_1 is used as therapeutics in Alzheimer disease (Cavallini, 1999; Bektas and Spiegel, 2004). Although signalling effects of sphingolipids and glycosphingolipids are supported by many experimental findings from literature, the precise mechanisms are not understood yet. Complex interconversions of sphingolipids and glycosphingolipids make it difficult to understand the main reason of sphingolipids' apoptotic and anti-apoptotic effects. Is it only ceramide that induces apoptosis and are the

changes in rest of the sphingolipids a response to ceramide accumulation or do they also induce/prevent apoptosis due to some distinct mechanisms? Hence, the examination and analysis of sphingolipid metabolism and metabolic consequences of cellular disturbances become extremely important to enlighten our understanding of human sphingolipids.

The best objective function in the modelling of in *silico* healthy cell would be the maximization of glucosylceramide, GM_1 and sphingosine-1-phosphate. On the other hand the best objective function in the modelling of *in silico* cancerous cell would be the maximization of ceramide and GD_3 and minimization of sphingosine-1-phosphate and GM_1. Although we have some intuitions about the objective function selection, which is the first step of metabolic flux analysis, one needs to validate the chosen objective function by experimental fluxome data for the continuation of the modelling study. In literature no flux data are available for human sphingolipid biosynthesis pathway. Deletion mutant phenotype information for genes involved in human sphingolipid metabolism and/or a dynamic model would also be used in validating an *in silico* model, but such information is unfortunately missing for human sphingolipid metabolism.

5. RECONSTRUCTION AND ANALYSES OF SPHINGOLIPID PROTEIN-PROTEIN INTERACTION NETWORKS

Protein-protein interaction network of sphingolipid metabolism, regulation and transport are reconstructed and analysed in terms of topological properties for yeast and human.

5.1. Reconstruction and Analysis of Yeast Sphingolipid Protein Interaction Network and Application of Hybrid Function Annotation Methodology

Sphingolipids are signaling molecules that control vital cellular processes like cell proliferation, differentiation, cell cycle arrest, and apoptosis. They are also structural components of the membrane. The difficulty introduced by their existence in membrane is that all the information related to the mechanisms and structures governing sphingolipids and their vital role in cell's fate are not yet clearly understood as a whole (Kavun Ozbayraktar and Ulgen, 2009; Riezman, 2006). After the development of numerous high-throughput techniques in the post-genomic era and the inevitable flood of protein-protein interaction data, the use of interaction networks in the understanding of cell's functional organization and analysis by graph theoretical applications emerged widely (Barabasi and Oltvai, 2004; Grindrod and Kibble, 2004; Jeong *et al.*, 2000). Thus, by analyzing protein interaction networks it may be possible to unravel the principles of sphingolipid metabolism and consequently many related diseases such as Fabry, Gaucher, Niemann-Pick, Farber, Krabbe, Metachromatic Leukodystrophy and cancer (Kavun Ozbayraktar and Ulgen, 2009).

The biosynthesis journey of sphingolipids in yeast cell starts in the endoplasmic reticulum and ends in the Golgi apparatus and plasma membrane. The enzymes that take part in this biosynthesis route of sphingolipids are mostly membrane proteins with transmembrane domains. This property loads uncertainty to the topology of the proteins along with the mechanisms governing their protein-protein interactions. Although there have been much information related to sphingolipid metabolism at many specific points, little is known about the big picture integrating their biosynthesis, regulation, transport and function. In this study, the protein-protein interaction network of yeast (*S. cerevisiae*)

sphingolipid metabolism was constructed, and its topology was elaborately analyzed. The sphingolipid metabolism proteins, some of which were reported as drug targets for cancer therapy in literature as well as in Chapter 3 and 4 (Kavun Ozbayraktar and Ulgen, 2010), are examined in terms of connectivity, modularity and their orientation in the whole network. We also developed a hybrid algorithm integrating sequence and structure based interaction prediction tools, orthology, expression profiles, co-localization information and GO terms to link the isolated sphingolipid proteins to the constructed network.

Each protein-protein interaction serves a function to implement the global behavior of the cell. Only a small number of proteins' functions have been determined by experimental techniques, and the rest relies on the predictions from computational methods (Louie *et al.*, 2009). Although annotation transfer via sequence similarity is a traditional tool (Hawkins *et al.*, 2006; Hawkins *et al.*, 2008; Louie *et al.*, 2009; Verspoor *et al.*, 2006), homologous sequences are not always identified at acceptable significance threshold values. The drawback of sequence similarity may be eliminated by the use of domains, *i.e.,* protein family databases such as PFAM (Finn *et al.*, 2010) and InterPro (Hunter *et al.*, 2009). Annotation transfer by domains becomes inadequate for membrane proteins, whose structures are not clearly known (Riezman, 2006). Although half of the genetically linked proteins function in the same cellular pathway, genetic interactions may be useful only as partial evidence in function assignment to an uncharacterized protein (Lee *et al.*, 2007). Genes encoding proteins with similar function tend to be co-expressed and co-localized (Chen and Xu, 2005; Xu *et al.*, 2008). Physical interactions and modules are also used in function prediction of uncharacterized proteins (Sharan *et al.*, 2007; Xu *et al.*, 2008). Novel and more reliable approach in function annotation would be integration of each dimension into a hybrid method, which takes into account all the evidence available in literature. In the present study, the uncharacterized proteins that are revealed to function in sphingolipid related processes are identified and a putative function is annotated based on a multi-dimensional approach incorporating data gathered from functional modules, expression profiles, sequence similarity, genetic interactions, physical interactions and domains.

5.2. Computational Methods

The computational systems biology tools utilized in the reconstruction of protein-protein interaction networks and their analysis are described in detail in the following sections. The here developed integrated method of protein-protein interaction prediction and the hybrid method for function annotation are also discussed.

5.2.1. Reconstruction of Protein Interaction Network and Selective Permissibility Algorithm

Selective permissibility algorithm, SPA, (Arga *et al.*, 2007) is utilized to reconstruct the yeast sphingolipid protein interaction network. The method originates from a set of proteins, called core proteins. The proteins that function in sphingolipid metabolism, sphingolipid related processes and have evidence that they are related to sphingolipids (by an unknown mechanism) are all taken into consideration following a detailed literature survey (AmiGO; KEGG; MIPS; Schneiter, 1999; SGD). The gene ontology (GO) terms of process, function and component are obtained from AmiGO (31/05/2010) and used to build an annotation table. The GO terms of the core proteins in the annotation table are used as a measure to add new proteins and reconstruct the network. SPA is an iterative procedure that converges when there are no new proteins that could be added to the protein interaction network according to the interactome data and annotation table. The loop starts with the identification of the first neighbors of the core proteins and the selection of them according to having at least one GO term of each three categories of process, function and component from the annotation table. Only the physical interactions are used in neighbor finding process, these interactions are taken from BioGRID: Database of Protein and Genetic Interactions-v3.0.64 (Stark *et al.*, 2006). The same procedure is applied to all the consecutive neighbors until no new proteins could be appended to the network. Finally the repeats and self-loops are removed, and the proteins isolated from the rest of the network are identified. The neighbor finding process is performed using Microsoft EXCEL, and the elimination process is performed using a code written in MATLAB R2007a (Appendix G).

The computational methods utilized in the reconstruction and analysis of human sphingolipid protein-protein interaction network are described in this section in detail. The

physical protein-protein interactions are taken from BioGRID: Database of Protein and Genetic Interactions-v3.0.64 (Stark *et al.*, 2006). The gene ontology (GO) terms of process, function and component are obtained from AmiGO (13/12/2010).

5.2.2. Methodology for the Prediction of Missing Protein Interactions

The proteins that are not linked to the bulk interaction network due to the lack of interactome data are called isolated proteins and if these isolated proteins are among the core proteins, then the necessary but unknown connections should be revealed and added to the network. Four different sources of information coupled with two computational protein-protein interaction prediction methods (one sequence based and the other structure based) are used to predict the missing interactions required to rewire the isolated core proteins to the network. Since it is known that genes, which encode interacting proteins, are mostly co-expressed (Sharan *et al.*, 2007), we identified the genes that have similar expression profiles with the isolated proteins. The co-expressed proteins are identified from *Saccharomyces* Genome Database (SGD), where the top 20 similarly expressed genes of the isolated query gene are listed with a Pearson correlation coefficient greater than 0.8.

Orthologous proteins are mostly involved in similar processes to serve the same function (Yu *et al.*, 2004). In the current study, the percent amino acid identity of proteins is defined as the ratio of the number of matching amino acid residues over total number of amino acid residues in the query protein. Based on orthology and function relationship, we used the concept of interologs (Walhout *et al.*, 2000) to find out missing interactions of isolated proteins. The interologs are identified based on the conservation of the interaction partners across species. Five experimentally tractable model organisms (*Caenorhaditis elegans, Drosophila melanogaster, Helicobacter pylori, Homo sapiens, Mus musculus* and *Saccharomyces cerevisiae*) are used in interaction transfer via orthology (Yu *et al.*, 2004). BLASTP (Altschul *et al.*, 1990) analysis is performed among non-redundant protein sequences in five organisms, the best-matching candidate is determined as the ortholog of the query protein from each species with E-value less than 10^{-70}, percent identity greater than 80 %, and along with these two criteria, the selected ortholog proteins are reviewed proteins in UniProt. The interaction partners of the orthologous proteins are determined in the corresponding organisms using BioGRID data. These interaction partners are then

filtered according to the GO terms in the annotation table of yeast and the orthologs of the suitable interaction partners are found in yeast using the same threshold values for E-value and percent identity depicted above. The interaction partners that are transferred via orthology are also controlled against the annotation table. It should be noted that the interaction transfer via orthology should be rephrased as interaction transfer via homology for the case of *S. cerevisiae*.

Besides the interaction data taken from BioGRID, the IntAct molecular interaction database (Aranda *et al.*, 2009) and Yeast Resource Center (YRC) Informatics Platform, Public Data Repository are used to find potential interaction partners to isolated proteins.

Protein-protein interaction prediction engine (PIPE v2) is used to predict novel protein-protein interactions based on the amino acid sequence using a moving 20 amino acid window. The sensitivity of the method for yeast is reported as 57 %, specificity is 89 % and accuracy is 75 % (Pitre *et al.* 2006). Struct2Net (Singh *et al.*, 2006) is another computational protein-protein interaction prediction algorithm but distinctively it is based on protein structure. The method uses homology to find out the resemblance of any two proteins with known protein complexes and then calculates the interaction energy using logistic regression (Singh *et al.*, 2006). The authors performed a whole yeast genome interaction prediction study, from which the predictions for isolated proteins are included as supportive data in our algorithm.

In order to rewire the isolated proteins to the newly reconstructed network the proteins obtained from all data sources are pooled and screened using the annotation table constructed with GO terms of the core proteins. For each isolated protein a binary matrix having putative interaction partners on the rows and 9 selection criteria on the columns is formed and then scored as 1 or 0. The score “1” indicates that putative interaction partner satisfies the related criterion. The 9 selection criteria are (i) prediction by PIPE, (ii) prediction by Struct2Net, (iii) existence in IntAct and YRC public data repository, (iv) orthology, (v) being similarly expressed, (vi) being co-localized to the same cellular compartment with the isolated protein of interest, (vii) having lipid related GO terms, (viii) having sphingolipid related GO terms, and (ix) being a member of the newly constructed network. The data obtained from IntAct, YRC public data repository, PIPE and Struct2Net

designate the existence of a physical interaction with changing reliability. The corresponding ones and zeros of these four sources in each row of the matrix are summed to give the "interaction score" which shows the confidence of each putative interaction partner. It is necessary but not enough to have interaction score greater than zero; it should be supported by other information sources. We sum up the rest of the selection criteria (*v-ix*) to constitute the "supportive score". Finally, "overall score" is defined as the product of interaction score, supportive score, co-localization information (1 or 0), and state of belonging to the newly constructed network (1 or 0). The overall score is designed such that a putative interaction partner is taken into consideration in further calculations if and only if it is co-localizated with the isolated protein of interest, it already exists in the newly constructed network, and it has a positive interaction and supportive score simultaneously. It is more likely that the isolated core proteins of sphingolipid metabolism may interact with the members of the constructed network. The putative interaction partners that have overall score greater than or equal to 3 are taken as interactors to rewire the isolated proteins and to be used in further topological analysis. The putative interaction partners with overall score equal to 2 are examined one by one in terms of GO terms, and if a protein which is directly related to sphingolipid metabolism or if the same function as the core protein under investigation, is encountered, then the interaction between them is also accepted to be used in the network. A "pseudo overall score" is also defined that discards the co-localization information from the overall score. The proteins with pseudo overall score greater than or equal to 3 are examined one by one in terms of GO component terms in order to eliminate false negatives, *i.e.* not to disregard any protein pair due to seemingly different but actually same GO component terms (*e.g.,* membrane and integral to membrane).

After the selection of probable interaction partners of isolated proteins, we tried to support the novel interactions via domain-domain interactions. The interaction of two proteins is explained by the binding of at least two domains at atomic level. DOMINE (Raghavachari *et al.*, 2008) is a database that consists of known (from PDB structures) and predicted (from 8 different computational methods) domain-domain interactions with changing confidence levels. Experimentally determined domain-domain interactions from PDB entries have the highest confidence (Raghavachari *et al.*, 2008). Manually curated Pfam A domains of isolated proteins and probable interaction partners are found from

InterPro v28.0 (Hunter *et al.*, 2009) and DOMINE database is used to control whether there is a known or predicted domain-domain interaction for these domain pairs.

5.2.3. Graph Theoretical Approach

The network theory is applied to characterize the newly constructed network of yeast sphingolipid metabolism in terms of organization and structure. Graph theoretical applications help illuminate the metabolic, regulatory and transport processes governing sphingolipid metabolism and detailed understanding of the role of individual proteins on these processes is aimed. Graph theoretical analyses (network diameter, shortest path lengths, characteristic path length, degree distribution, clustering coefficients, and heterogeneity) are achieved by Cytoscape network visualization software v2.6.3 (Shannon *et al.*, 2003) using NetworkAnalyzer v2.6.1. plug-in (Assenov *et al.*, 2008). CytoHubba plug-in is used in the determination of hubs and bottleneck proteins (Lin *et al.*, 2008). BINGO v2.3. plug-in of Cytoscape network visualization software (Maere *et al.*, 2005) is used to identify the GO terms enriched in the network using the GO terms downloaded from AmiGO (31/05/2010). Statistical calculations are performed in MATLAB R2007a by binocdf command.

5.2.4. A Hybrid Function Annotation Methodology

The proteins which have both unknown process and function GO terms are classified as uncharacterized in the current study. These proteins constitute the unknown protein set, whose functions are attempted to be predicted. A gold-standard set (referred to as gold-set) of proteins is formed for the evaluation of the performance of the function prediction methodology developed here. GO evidence codes are utilized as the criteria of being a member of the gold-set, *i.e.* the proteins with at least one GO term determined from direct assay experiments for each process, function and component categories are selected among our protein interaction network (PIN). This criterion guaranteed a high-quality test set for the evaluation of our function prediction method which is first applied to the gold-set; the GO annotations of gold-set proteins are discarded from the analysis and they are treated as uncharacterized proteins during testing. True positives and false positives are calculated to further evaluate the precision and prediction performance of our method. When our

algorithm passed the statistical test on the set of gold-standard proteins, the same method is applied for function prediction of unknown proteins. In this study, the function prediction is performed by using the GO biological process terms that are not inferred from electronic annotation. The terms coming from electronic annotation are error-prone and it is not safe to base a prediction upon another prediction.

A module-assisted approach is employed in the first basic step of the current function prediction algorithm. The members of a functional module come together to perform a specific function in the cell, so one can assign a function to a protein involved in a functional module based on the enriched GO terms of that module. The molecular complex detection algorithm (MCODE) identifies highly interconnected regions, namely modules (Bader and Hogue, 2003). MCODE v1.2 plug-in is used in Cytoscape network visualization software v2.6.3 (Shannon *et al.*, 2003) to determine the modules that contain at least one protein from the set of interest (gold or unknown). The unknown density of the modules is calculated as the ratio of the number of proteins of interest (gold-proteins during testing, unknown proteins during function prediction) over total number of proteins in the module. The modules that have unknown density greater than 15 % are eliminated, due to the decreased confidence level of obtaining the true enriched GO terms which will be employed in function prediction of unknown proteins. BINGO plug-in of Cytoscape network visualization software (Maere *et al.*, 2005) is used to find out the most frequent and statistically significant (p-value less than 0.05) GO process terms encountered in each module with unknown density less than or equal to 15 %. The proteins that share the same frequent and significant GO process terms are also identified.

Every source of information has its own advantages and disadvantages in function prediction, so a more comprehensive method is the one that can integrate many to obtain a multi-dimensional approach in function prediction (Sharan *et al.*, 2007; Joshi *et al.*, 2004). Three different information sources (genetic interactions, expression profiles and sequence similarity) are utilized as supporting data to GO term enrichment results of modules to be used in function prediction. Annotation transfer from co-expression data is based on the idea that proteins with similar functions also have similar expression profiles (Forslund and Sonnhammer, 2008; Joshi *et al.*, 2004; Massjouni *et al.*, 2006; Zhu *et al.*, 2007). The top 20 similarly expressed genes of the query gene are obtained from *Saccharomyces* Genome

Database with a Pearson correlation coefficient greater than 0.8. Annotation transfer from sequence similarity is also another widely used method (Forslund and Sonnhammer, 2008; Jung *et al.*, 2010). BLASTP is performed for the protein of interest against the whole module with an expect threshold value of 10^{-5}. Genetic interactions which also indicate functionally linked gene products (Mani *et al.*, 2008), are taken from BioGRID v3.0.64. We have integrated these three methods, which are all widely used in annotation transfer separately, in module-assisted function prediction approach to obtain an indication of reliability of each module before function prediction. The proteins that have genetic interactions, similar expression profiles and sequence similarity with the proteins of interest (gold or unknown) are found and the participation percentages of them in each frequent and significant GO process term are calculated for each module separately. The most frequent and significant GO process term containing proteins with participation percentages greater than or equal to 20 % (from any of the three supportive information sources) are those whose function annotation assignments are performed at high confidence. The other frequent and significant GO process terms with greater than 20 % support are also assigned to the proteins of interest but with low confidence.

To further support function assignment, neighborhood counting method (Sharan *et al.*, 2007) is integrated to the results of the module-assisted approach. Being a direct method of function prediction, neighborhood counting is based on the fact that two physically interacting proteins share similar functions and it makes use of first interaction partners of the protein of interest in the whole yeast proteome. GO term enrichment for the first neighbors of proteins of interest is performed by using BINGO plug-in. The most frequent and significant GO process terms are compared with those of the related modules; the common terms are identified. The common terms which have average frequency greater than or equal to 30 % are assigned to the proteins of interest at high confidence.

Finally, the repeated keywords encountered in enriched GO process terms of both neighbors and modules are counted and their frequencies are calculated as the number of occurrence among the statistically significant and frequent terms. Among the different threshold values tested in gold-set, 14 and more repeats in 100 significant and frequent terms are assigned as low confidence function annotations due to the sensitivity. The

algorithm of the newly developed hybrid function annotation method is schematically summarized in Figure 5.1.

For the proteins of unknown-set, domain analysis is performed using Pfam2GO, Interpro2GO (Hunter *et al.*, 2009) and Multipfam2GO (Forslund and Sonnhammer, 2008) mappings. Pfam2GO and Interpro2GO map domains to GO annotations and Multipfam2GO maps sets of domains to GO annotations. Manually curated Pfam A domains of unknown proteins are found from InterPro v28.0 (Hunter *et al.*, 2009) and their GO annotations are used to support our low and high confidence function predictions.

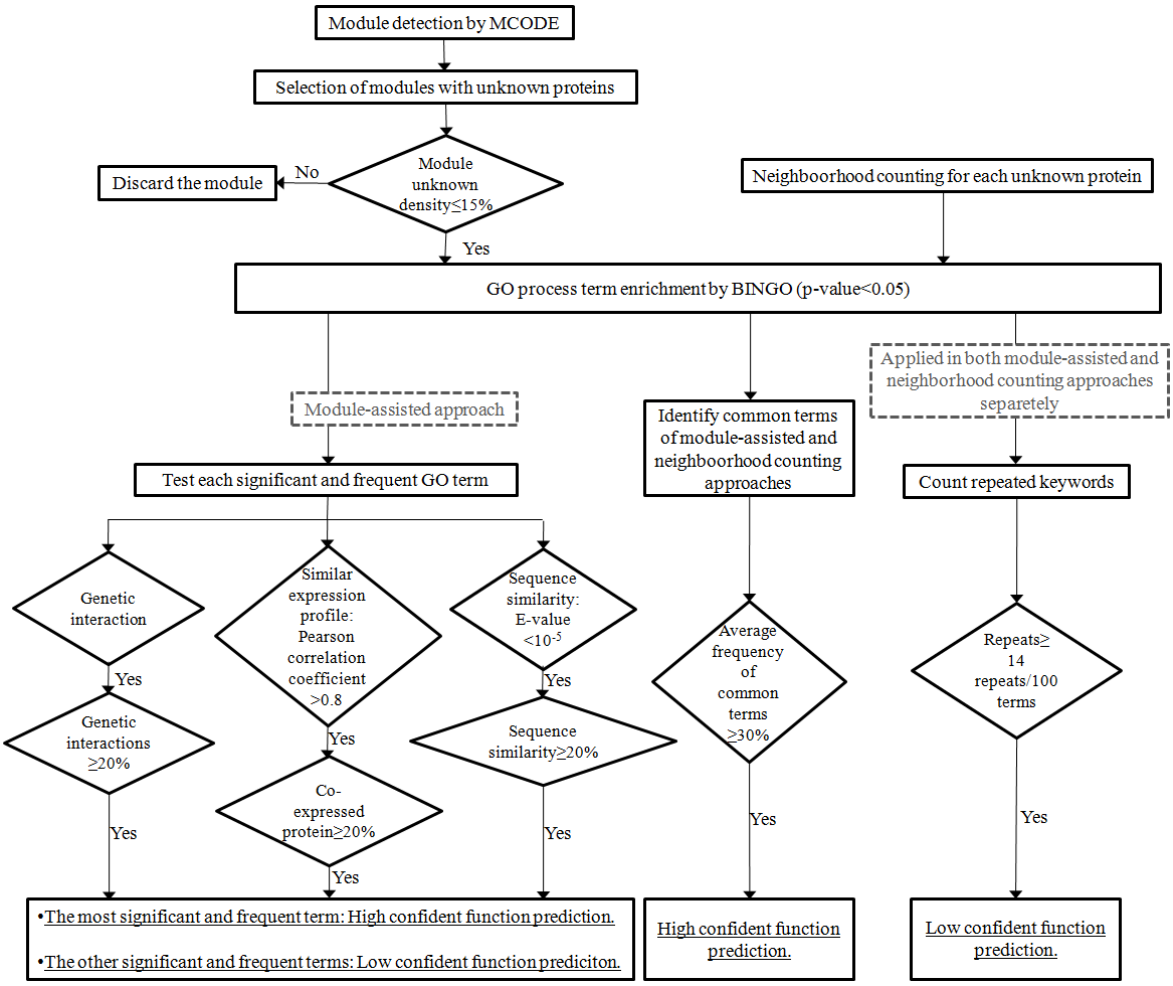


Figure 5.1. Four different routes are used in function annotation of uncharacterized proteins. The routes originate from either a module-assisted or neighborhood counting approach or the combination (grey dashed boxes show the classification)

5.3. Results and Discussion

The newly reconstructed yeast sphingolipid protein-protein interaction network is examined and topologically analyzed. The results are further investigated to use in function annotation to uncharacterized sphingolipid proteins.

5.3.1. Reconstruction of the Yeast Sphingolipid Protein Interaction Network

32 core proteins, which are used as the seeds of yeast sphingolipid PIN, are selected by manual examination of GO terms, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway database, SGD and MIPS (Munich Information Center for Protein Sequences). The proteins that take part in our previous dynamic and stoichiometric models of sphingolipid metabolism (Kavun Ozbayraktar and Ulgen, 2010) are also included in core protein set. 91 % of the core proteins function directly in sphingolipid metabolic process and its sibling processes (see Figure 5.2 for detailed GO process term distribution), the rest (9 %) also function in sphingolipid related processes however the precise mechanisms of sphingolipid relationship are not clearly understood yet. 72 % of the core proteins are in endoplasmic reticulum, where the biosynthesis of sphingolipids begins. The annotation table, which is constructed by using the GO terms of core proteins, consists of mainly sphingolipid related terms, beside the unknown terms. Unknown terms give the opportunity to identify novel proteins that function in sphingolipid processes.

Six runs are required to obtain convergence in SPA, in the seventh run no new proteins are added to the network (Appendix C, Figure C.1). Totally, 1591 unique proteins are attained in yeast sphingolipid PIN with 5341 edges linking them. The resulting network is undirected and self-loops are eliminated as the homomer relationships are not considered in this study (Chiang *et al.*, 2007). Five isolated core proteins are observed; the genes that encode them are: YDR294C, YER019W, YGR143W, YKR053C, and YGR212W. YDR294C encodes dihydrosphingosine phosphate lyase, which is responsible for the degradation of dihydrosphingosine phosphates into ethanolamine phosphate and palmitaldehyde (Kim *et al.*, 2000). YER019W encodes inositol phosphosphingolipid phospholipase C, which is required for the breakdown of complex sphingolipids into ceramide (Sawai *et al.*, 2000). YGR143W encodes a membrane protein that is functional in

beta-glucan synthesis and sphingolipid biosynthesis (Roemer *et al.*, 1993). YKR053C encodes dihydrosphingosine 1-phosphate phosphatase that is responsible for the dephosphorylation of phosphorylated dihydrosphingosine back to dihydrosphingosine (Mao *et al.*, 1997). Finally, YGR212W encodes N-acetyltransferase, which shows resistance to sphingolipid biosynthesis inhibitor myriocin (Momoi *et al.*, 2004). The lack of information in the interactome is the reason why these five proteins are left isolated. Although they are known to be functioning in sphingolipid biosynthesis, the precise mechanisms in terms of protein interactions are not known yet.

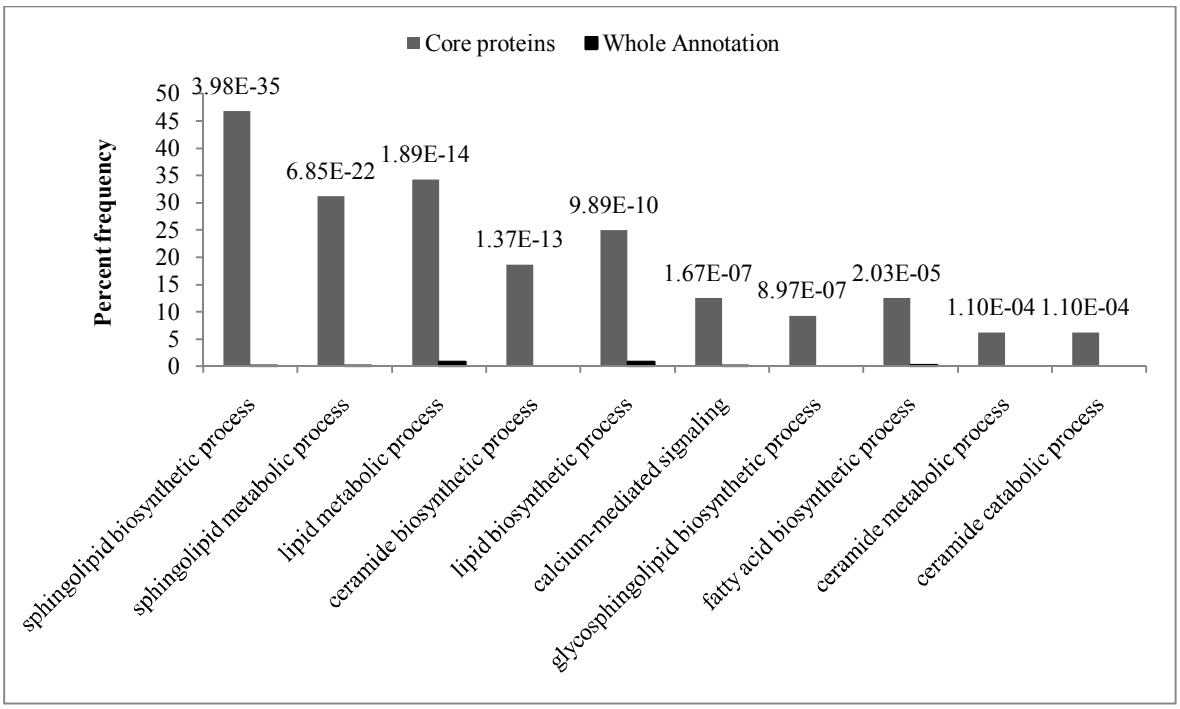


Figure 5.2. GO process term distribution for core proteins. GO process terms enriched among core proteins with p-values calculated with respect to whole annotation

5.3.2. Protein Interaction Prediction to Connect the Isolated Core Proteins to the Bulk Network

An integrated approach of protein interaction prediction is used to rewire five isolated core proteins to the rest of the network. Table 5.1 gives the binary matrix constructed and utilized in interaction prediction algorithm. Four putative interaction partners are predicted and appended to the network for dihydrosphingosine phosphate lyase (YDR294C) with interaction score equal to one and supportive score equal to four.

YDR062W, YMR296C, and YJL167W are interaction partners predicted via Struct2Net and YKL008C is suggested as a putative interaction partner with evidence from IntAct. All are co-localized with YDR294C, have lipid related GO process terms, and in fact three of them have directly sphingolipid GO process terms. The putative interaction partners for dihydrosphingosine phosphate lyase with sphingolipid GO process terms are two components of serine palmitoyltransferase (YDR062W, YMR296C) and one component of ceramide synthase (YKL008C). Reiss *et al.* (2004) reported that sphingosine phosphate lyase enzyme activity induces apoptosis by increasing ceramide generation (Reiss *et al.*, 2004). The lyase is a modulator of ceramide metabolism, pointing out its relation with ceramide synthase and serine palmitoyl transferase, which is the first step of *de novo* ceramide synthesis.

Two putative interaction partners are predicted for inositol phosphosphingolipid phospholipase C (YER019W). The interaction score of two is obtained with evidence from IntAct and PIPE for YCR068W, which is co-expressed and co-localized with YER019W, has lipid related GO terms and already exists in the network. The other predicted partner is YHR135C, with interaction evidence from PIPE and it is already available in the overall network, co-expressed and co-localized with YER019W. Although inositol phosphosphingolipid phospholipase C is reported to function in mitochondria and regulate sphingolipid metabolism in the organelle, a bulk of the complex sphingolipids in yeast are found in plasma membrane (Kitagakia *et al.*, 2007). The predicted protein-protein interactions of inositol phosphosphingolipid phospholipase C with YCR068W gene product (functioning in lipid catabolic process) and YHR135C gene product (functioning in endocytosis) may indicate the degradation of complex sphingolipids in plasma membrane and the existence of an unknown transport mechanism of degradation product ceramide, which controls macroautophagy (Kavun Ozbayraktar and Ulgen, 2009; Scarlatti *et al.*, 2004).

Five putative interaction partners are suggested for YGR143W, the details of the prediction, interaction scores, and supportive scores are listed in Table 5.1. Although the involvement of the protein in complex sphingolipid biosynthesis is known, the precise mechanism has not been elaborated yet (Thevissena *et al.*, 2005).

Table 5.1. Binary matrix developed in interaction prediction algorithm. Putative interaction partners selected for each isolated core protein are shown with their interaction and supportive scores. The sources of information used to build the supportive scores are shown in grey

Isolated core proteins	Putative Interaction Partners	PIPE	Struct2Net	IntAct and/or YRC	Orthology Search	Interaction Score	Similarly expressed proteins	Co-localization	Lipid related GO terms	Sphingolipid related GO terms	Member of network	Supportive score	Overall score	Pseudo overall score	DDI Domine
YDR294C	YDR062W	0	1	0	0	1	0	1	1	1	1	4	4	4	0
	YKL008C	0	0	1	0	1	0	1	1	1	1	4	4	4	0
	YMR296C	0	1	0	0	1	0	1	1	1	1	4	4	4	0
	YJL167W	0	1	0	0	1	1	1	1	0	1	4	4	4	0
YER019W	YCR068W	1	0	1	0	2	1	1	1	0	1	4	8	8	0
	YHR135C	1	0	0	0	1	1	1	0	0	1	3	3	3	0
YGR143W	YFR042W	1	0	0	1	2	0	1	0	0	1	2	4	4	0
	YLR096W	1	0	0	1	2	0	1	0	0	1	2	4	4	0
	YKL126W	1	0	0	0	1	0	1	1	0	1	3	3	3	0
	YDR122W	1	0	0	0	1	1	1	0	0	1	3	3	3	0
	YOL129W	0	0	1	0	1	1	1	0	0	1	3	3	3	0
YKR053C	YCR007C	1	0	0	0	1	1	1	0	0	1	3	3	3	1
	YGR295C	1	0	0	0	1	1	1	0	0	1	3	3	3	1
YGR212W	YPL204W	1	0	0	0	1	0	1	0	0	1	2	2	2	0

Two putative interaction partners (YCR007C, YGR295C) are predicted for dihydrosphingosine 1-phosphate phosphatase (YKR053C) at low confidence. They are the only interactors confirmed by the domain-domain interactions from DOMINE database (Raghavachari *et al.*, 2008). The proteins involving in sphingolipid related processes are mostly membrane proteins whose structures are not clearly known, therefore the domains and domain-domain interactions are not all identified yet (Riezman, 2006). The absence of a domain-domain interaction for membrane proteins generally indicates the lack of information rather than its real absence.

YGR212W do not have putative interaction partners with overall score greater than three. In order to assign at least one interaction partner with low confidence, we examined all the candidates with overall score of two in terms of GO terms. The selected putative interaction partner for YGR212W is YPL204W, which takes part in vesicle-mediated transport from endoplasmic reticulum to Golgi apparatus (Table 5.1). This process is significantly important in the transport and biosynthesis of sphingolipids (Futerman, 2006). YGR212W gene product functions in response to drug (myriocin), so that transport mechanisms gain importance. 36 % of the novel putative interaction partners predicted in this study already have interactions with at least 10 and more proteins in the reconstructed network. Finally, with the addition of predicted interactions for isolated core proteins, the network of yeast sphingolipid PIN has 1591 nodes and 5355 edges without self-loops and isolated proteins.

5.3.3. Graph Theoretical Analysis of the Protein Interaction Network

For efficient discussion of topological analysis of yeast sphingolipid PIN, the network of whole yeast proteome (5655 nodes and 51932 edges) is also constructed and analyzed in terms of graph theoretical parameters, by this means it is aimed to understand how well our network represents the whole proteome in terms of topological properties. Table 5.2 gives the topological parameters of both networks under investigation in a comparative manner with examples from literature.

Table 5.2. The topological properties of newly constructed protein interaction network are compared with both whole yeast protein interaction network and reference values from different literature sources

Topological parameters	Sphingolipid PIN	Whole yeast PIN	References
Degrees	1591	5655	-
Edges	5355	51932	-
Degree exponent	1.68	1.41	1.8 (*S. cerevisiae* PIN Li *et al.*, 2006) 1.6 (*C. elegans* PIN Li *et al.*, 2006) 2.0 (*D. melanogaster* PIN Li *et al.*, 2006) 2.34 (*S. cerevisiae* PIN from MIPS Yook *et al.*, 2004) 2.50 (*S. cerevisiae* PIN from DIP Yook *et al.*, 2004)
Network heterogeneity	1.34	2.61	-
Characteristic path length	3.97	2.77	4.2 (*S. cerevisiae* PIN Li *et al.*, 2006) 4.8 (*C. elegans* PIN Wu *et al.*, 2005; Li *et al.*, 2006) 4.5 (*D. melanogaster* PIN Li *et al.*, 2006) 4.4 (D. melanogaster PIN Wu *et al.*, 2005) 5.0 (*S. cerevisiae* PIN Wu *et al.*, 2005) 4.1 (*H. pylori* PIN Wu *et al.*, 2005) 6.8 (*H. sapiens*PIN Wu *et al.*, 2005) 5.5 (*E. coli*PIN Wu *et al.*, 2005) 3.6 (*M. Musculus* PIN Wu *et al.*, 2005)
Network diameter	10	6	13 (*S. cerevisiae* PIN Wu *et al.*, 2005) 14 (*C. elegans* PIN Wu *et al.*, 2005) 11 (*D. melanogaster* PIN Wu *et al.*, 2005) 9 (*H. pylori* PIN Wu *et al.*, 2005) 21 (*H. sapiens*PIN Wu *et al.*, 2005) 16 (*E. coli*PIN Wu *et al.*, 2005) 9 (*M. Musculus* PIN Wu *et al.*, 2005)
Clustering coefficient	0.132	0.236	0.33 (*S. cerevisiae* PIN Wu *et al.*, 2005) 0.06 (*C. elegans* PIN Wu *et al.*, 2005) 0.02 (*D. melanogaster* PIN Wu *et al.*, 2005) 0.08 (*H. pylori* PIN Wu *et al.*, 2005) 0.21 (*H. sapiens*PIN Wu *et al.*, 2005) 0.62 (*E. coli*PIN Wu *et al.*, 2005) 0.15 (*M. Musculus* PIN Wu *et al.*, 2005)

Yeast sphingolipid PIN contains 28 % of the nodes and 10 % of the edges of the whole interactome data. A power law degree distribution is obtained for yeast sphingolipid PIN with a degree exponent of 1.68 and a goodness of fit of 0.9. The degree exponent less than three indicates a scale-free network, which in turn points out the existence of hub proteins (Barabasi and Oltvai, 2004). As degree exponent gets much smaller than three, then the importance of hubs in the network becomes more significant. Most of the biological networks are scale-free networks, with a few highly connected nodes and most

of the nodes with just a few degrees; shortly the network is not uniform (Barabasi and Oltvai, 2004). The existence of hub proteins (essentially the existence of non-hub proteins) in biological networks makes the network robust for disturbances (Yu *et al.*, 2004). Non-uniformity of the network's degree distribution is also quantified with network heterogeneity (Dong and Horvath, 2007). Yeast sphingolipid PIN is heterogeneous (1.34) but it is not as heterogeneous as the whole interactome (2.61). This is due to the fact that sphingolipid network's hub proteins do not have as much interactions as the whole interactome's hubs have.

Characteristic path length, which evaluates the accessibility/interconnectivity of the network, introduces the concept of small-world (Li *et al.*, 2006; Watts and Strogatz, 1998). The criterion of being a small-world network is defined as having characteristic path length on the order of natural logarithm of the total number of nodes in the network (Cohen and Havlin, 2003). Characteristic path length of yeast sphingolipid PIN is calculated to be 4, which is less than the natural logarithm of total number of the nodes (1591). It can be concluded that the newly constructed network of yeast sphingolipids is a small-world network (easily interconnected, four edges is sufficient to link any two nodes on average). Network diameter for yeast sphingolipid PIN is calculated as 10, whereas a smaller diameter of 6 is obtained for the whole yeast proteome. The average clustering coefficient of yeast sphingolipid PIN is calculated as 0.132, which shows the tendency of nodes of the network towards clustering (Barabasi and Oltvai, 2004). When this value is compared with the ones obtained from literature and with the whole yeast interactome network's average clustering coefficient, we can conclude that our network is modular as most of the complex and real networks (Ravasz *et al.*, 2002).

The proteins that are functioning in the accessibility of the network, *i.e.,* involving in many of the shortest path lengths are called bottlenecks. Bottleneck proteins, like hubs proteins, are significantly important connection points, which should be examined in detail. Betweenness centrality is used to identify bottleneck proteins (Yu *et al.*, 2007). In order to determine whether degree or betweenness centrality is decisive in essentiality of our network proteins, we identified both hubs (from degree) and bottlenecks (from betweenness centrality) and compared the ratio of essentiality. Top ten highest degree proteins are selected as hubs and top ten proteins with highest betweenness centrality are

determined as bottlenecks using CytoHubba (Lin *et al.*, 2008). Essentiality information of each gene is obtained from *Saccharomyces* Genome Deletion Project, SGD, and GENECENSUS. Cumulative binomial distribution is used to calculate p-values (*i.e.,* significance) for selecting essential proteins among each control group (hub-nonbottleneck, bottleneck-nonhub, and bottleneck-hub) against testing group of nonhub-nonbottleneck proteins. The number of genes in each group, percentage of essential genes and corresponding p-values are given in Table 5.3 and Figure 5.3 (part a).

Table 5.3. The significance of essentiality among bottleneck and hub proteins is calculated with respect to testing group of nonhub-nonbottleneck proteins. Hubs are selected only according to degree at this stage

	Total # of genes	**Essential # of genes**	**Percent of essential genes**	**P-value**
Nonhub-nonbottleneck (Testing group)	1565	185	11.8	-
Hub-nonbottleneck	3	1	33.3	3.39×10^{-86}
Bottleneck-nonhub	3	2	66.7	0 ($<10^{-324}$)
Bottleneck-hub	7	2	28.6	1.44×10^{-57}

As it can be seen from Figure 5.3 (part a), hub-nonbottleneck proteins are significantly essential with respect to nonhub-nonbottleneck proteins, but the level of significance is much more considerable for bottleneck-nonhub proteins. It can be concluded that bottleneck proteins are more likely to be essential proteins in yeast sphingolipid PIN. Degree is known to be a better indicator of essentiality in interaction networks, but exceptionally if bottleneck-nonhub proteins are involved in signaling processes, then betweenness centrality is reported to be a better indicator of essentiality (Yu *et al.*, 2007). The bottleneck-nonhub proteins (YPL204W, YNL272C, and YFR024C-A) function in signal transduction and regulatory processes (AmiGO), verifying the importance of bottlenecks in essentiality for the network under investigation. YPL204W encodes a protein kinase, which regulates vesicular trafficking, DNA repair, and chromosome segregation. YNL272C gene product is essential for post-Golgi vesicle

transport. Although YFR024C-A gene product has a Src homology-3 (SH3) domain that performs an assembly and/or regulatory role, the exact function of the protein is not known yet (SGD).

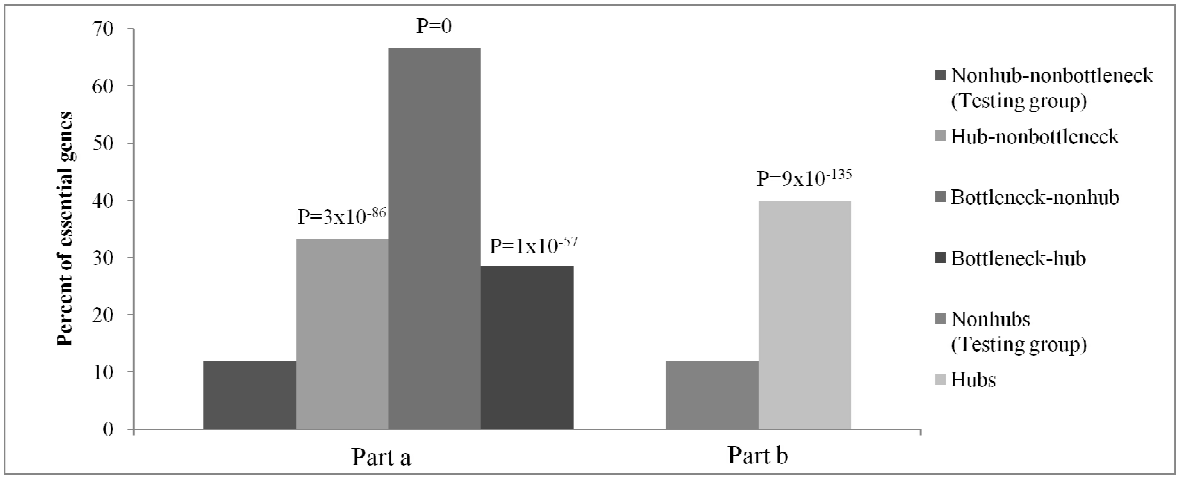


Figure 5.3. Part a shows the significance of essentiality among bottleneck and hub (degree) proteins calculated with respect to nonhub-nonbottleneck proteins. Part b shows only the significance of hubs (composite hub selection method) calculated with respect to nonhubs

Based on the above results, a composite hub selection method is utilized. Cytohubba plug-in offers four different calculation methods for the determination of hubs. These are degree, bottleneck, edge percolated component (Chin and Samanta, 2003), and maximum neighborhood component (Lin *et al.*, 2008). Edge percolated component (EPC), is a stochastic method that takes into account not only the direct links of a node but all the links of the network. Thus, essential nodes that have low degree are also identified via EPC (Chin and Samanta, 2003). Maximum neighborhood component gives score according to adjacent subnetworks of a node (Lin *et al.*, 2008). In the present study, the top 20 proteins that are identified as significantly important according to these four methods are compared with each other and the proteins that are determined from at least three of the methods are accepted as hubs of the network. 10 hubs for yeast sphingolipid PIN and 14 hubs for the whole yeast interactome network are identified according to this composite method and four proteins marked with an asterisk in Table 5.4 are hubs in both networks. 40 % of the hub proteins of yeast sphingolipid PIN and 12 % of the nonhubs are essential. Hence it can be concluded that hubs are significantly found in essential proteins (p-value=9.34×10^{-135}, see Figure 5.3 (part b)). GO enrichment results of the selected hubs of yeast sphingolipid

PIN indicate that half of the hubs have kinase activity and function in protein amino acid phosphorylation process with a significance level of $5x10^{-05}$. These GO terms are also enriched among hubs of the whole yeast interactome. Although hubs of yeast sphingolipid PIN show the same characteristics with those of whole yeast interactome, they also have the signature of lipid related terms like ketoreductase activity, lipid biosynthetic process, and fatty acid biosynthetic process (Figure 5.4). Nine out of ten hub proteins are conserved in human proteome (Appendix C, Table C.1). Either the whole protein is conserved with an E-value of less than 10^{-47} or the functional domain is conserved.

Table 5.4. Hub proteins are selected according to the composite hub selection method. Hub proteins of both whole yeast PIN and newly constructed sphingolipid PIN are given in a comparative manner. The asterisk (*) shows the proteins that are hubs in both networks

Whole yeast interactome hubs	**Yeast sphingolipid PIN hubs**
YBL032W	YBR106W
YBR160W*	YBR159W
YDL229W	YBR160W*
YDR381W	YGL137W
YDR510W	YGR060W
YFR004W	YHR135C*
YGL122C	YJL164C*
YHR082C	YJR059W
YHR135C*	YNL154C*
YJL164C*	YNL272C
YLL039C	
YNL154C*	
YNL197C	
YPL240C	

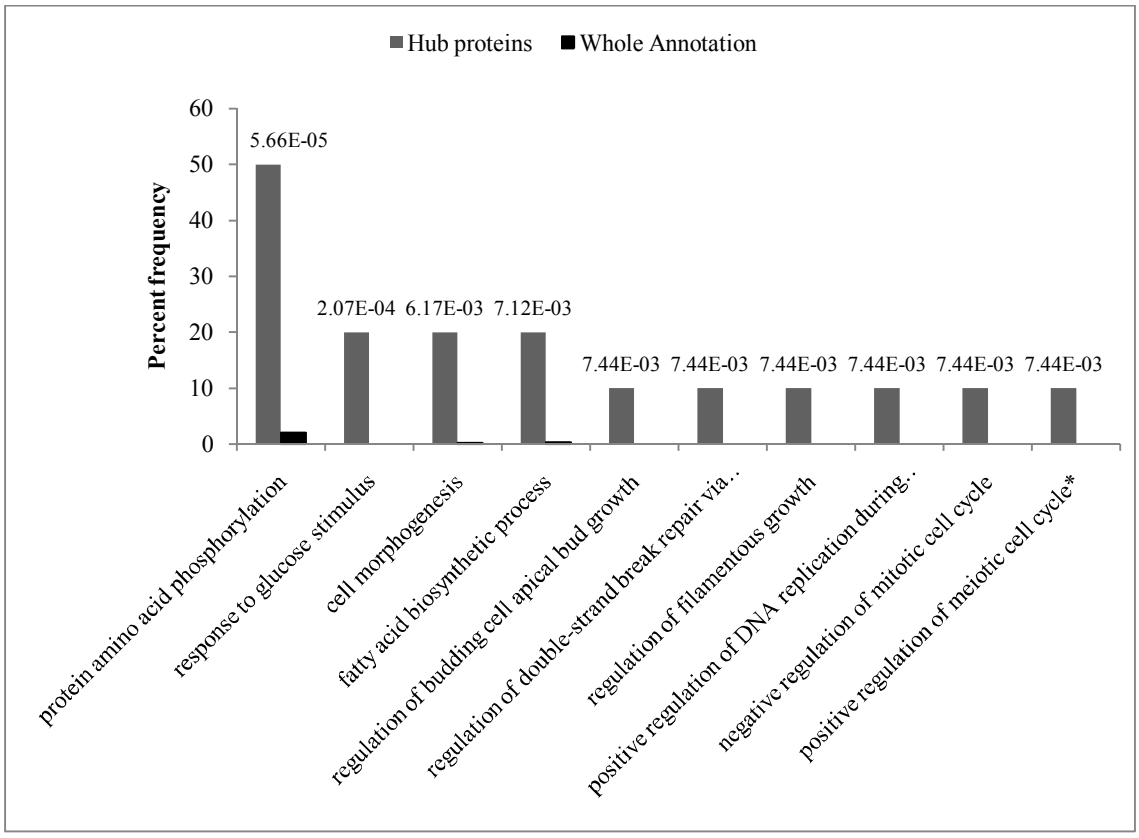


Figure 5.4. The GO process terms enriched among hub proteins of yeast sphingolipid PIN with p-values calculated with respect to whole annotation

5.3.4. Function Annotation to Uncharacterized Proteins

The topological analysis of the yeast sphingolipid network showed that it possesses the characteristic properties of biological networks. Network properties such as modularity can be safely used as the basis in function annotation of uncharacterized proteins. In the present study, we developed a hybrid multi-dimensional function annotation method using module assisted and neighbor counting approaches. There are 282 uncharacterized proteins with unknown process and function terms in our network. The gold-set contains 244 proteins with GO process, function and component terms inferred from direct assay experiments. Only 49 gold-proteins are involved in modules (found by MCODE) and 13 of them can be examined and annotated according to the unknown density of the modules. Our hybrid function annotation method enables us to correctly predict at least one GO term for 13 gold-proteins. Table C.2 gives detailed results for 13 gold-proteins. For example, we predicted four high confidence (ergosterol biosynthetic process, alcohol metabolic process, cellular lipid metabolic process, and cellular metabolic compound salvage) and two low

confidence (biosynthetic process and metabolic process) GO terms for YGL001C. In reality, YGL001C functions in ergosterol biosynthetic process, alcohol metabolic process, and cellular lipid metabolic process in aggrement with our findings. Our method has 85 % prediction rate for terms inferred from direct assay experiments, but if the source of evidence of the predicted GO term is not taken into consideration, the method has a 100 % prediction rate. Overall we have 34 correctly predicted GO terms (true positives) and only five incorrect GO terms (false positives), which makes 87 % precision for the hybrid function annotation method developed in the current study.

Subsequent to the validation of the hybrid function annotation method using gold-proteins and achievement of satisfactory results; we applied the same methodology for function annotation of unknown proteins. 32 of 282 unknown proteins are involved in twelve MCODE clusters, and only five clusters including 11 unknown proteins are suitable for function annotation purposes (unknown density less than 15 %). Only expression data give supportive information for function annotation of unknown proteins, no support could be found from genetic interactions and sequence similarity. The detailed results of function annotation to 11 unknown proteins are given in Table 5.5.

Table 5.5. 11 unknown proteins are involved in 5 different clusters. Unknown density of each cluster is tabulated. GO process term assignments at high and low confidence are given. The superscript ^{M} stands for the results of module-assisted, the superscript ^{N} stands for the results of neighborhood counting, the superscript ^{C} stands for the combination of both approaches

ORF name	Cluster	Unknown density	High confidence^{C}			High confidence^{M}		Low confidence^{M,N}		Low confidence^{M}	
			GO process term	Frequency in neighboorhood	Frequency in module	GO process term	% overall support	Keyword	% repeats	GO process term	% support
YHR140W	3	7.7	-	-	-	Ergosterol biosynthetic process	22	Metabolic process	50	Alcohol metabolic process	25
								Lipid	33.3	Cellular lipid metabolic process	40
								Transport	31.2		
								Biosynthetic process	18.7		
YPL264C	4	6.3	-	-	-	Glycoprotein metabolic process	20	Vesicle/vesicular	36.4	-	-
								Transport	31.8		
								Golgi	31.8		
								Vesicle-mediated	18.2		
								Glyco	18.2		
YDR307W	4	6.3	-	-	-	Glycoprotein metabolic process	20	Vesicle/vesicular	36.4	-	-
								Golgi	31.8		
								Transport	27.3		
								Vesicle-mediated	18.2		
								Glyco	18.2		
								Fatty acid	17.4		
								Biosynthetic process	17.4		

Table 5.5. 11 unknown proteins are involved in 5 different clusters. Unknown density of each cluster is tabulated. GO process term assignments at high and low confidence are given. The superscript ^{M} stands for the results of module-assisted, the superscript ^{N} stands for the results of neighborhood counting, the superscript ^{C} stands for the combination of both approaches (cont.)

ORF name	Cluster	Unknown density	High confidence^{C}			High confidence^{M}		Low confidence^{M,N}		Low confidence^{M}	
			GO process term	Frequency in neighboorhood	Frequency in module	GO process term	% overall support	Keyword	% repeats	GO process term	% support
YMR086W	13	4.1	-	-	-	Vesicle-mediated transport	25	Translation	42.9	Cellular carbohydrate metabolic process	25
								Regulation	42.9	Protein amino acid phosphorylation	22
								Negative regulation	28.6		
								Metabolic process	15.4		
YJR015W	13	4.1	-	-	-	Vesicle-mediated transport	25	Transport	27.3	Nucleotide metabolic process	30
								Metabolic process	15.4		
								Regulation	14.1		
YLR427W	13	4.1	-	-	-	Vesicle-mediated transport	25	Regulation	21.7	-	-
								Metabolic process	15.4		
YMR031C	13	4.1	-	-	-	Vesicle-mediated transport	25	Translation	57.1	Cellular carbohydrate metabolic process	38
								Regulation	28.6		
								Metabolic process	15.4		
YJL151C	6	9.1	-	-	-	-	-	Metabolic process	23.5	-	-
								Regulation	21.1		
								Transport	19.2		
								Biosynthetic process	17.3		

Table 5.5. 11 unknown proteins are involved in 5 different clusters. Unknown density of each cluster is tabulated. GO process term assignments at high and low confidence are given. The superscript ^{M} stands for the results of module-assisted, the superscript ^{N} stands for the results of neighborhood counting, the superscript ^{C} stands for the combination of both approaches (cont.)

ORF name	Cluster	Unknown density	High confidence^{C}			High confidence^{M}		Low confidence^{M,N}		Low confidence^{M}	
			GO process term	Frequency in neighboorhood	Frequency in module	GO process term	% overall support	Keyword	% repeats	GO process term	% support
YKL065C	11	8.3	-	-	-	-	-	Transport	52.4	-	-
								Regulation	45.4		
								Membrane	19		
								Metabolic process	15.1		
YHL048W	11	8.3	-	-	-	-	-	Regulation	45.4	-	-
								Transport	25		
								Metabolic process	15.1		
YDR032C	6	9.1	-	-	-	-	-	Transport	33.3	-	-
								Metabolic process	23.5		
								Regulation	20.8		

YHR140W is predicted to function in ergosterol biosynthetic process with high confidence. YPL264C and YDR307W appear in the same functional cluster, and they are predicted to function in glycoprotein metabolic process with high confidence. YDR307W gene product has dolichyl-phosphate-mannose-protein mannosyltransferase domain (PF02366), which is mapped to mannosyltransferase activity and protein amino acid O-linked glycosylation terms in Pfam2GO. This GO term mapping of the functional domain in YDR307W supports our finding of glycoprotein metabolic process. When the low confidence keyword predictions are gathered together to achieve a plausible function annotation, we attain vesicle-mediated transport for both uncharacterized proteins. Princeton University, Laboratory of Bioinformatics and Functional Genomics also annotated the function of vesicle-mediated transport for YDR307W gene product in agreement with our finding. YMR086W, YJR015W, YLR427W, and YMR031C take part in the same functional module. Vesicle-mediated transport is the predicted GO term for these four proteins at high confidence. The predicted low-confidence keywords are regulation, transport, and metabolic process. Protein amino acid phosphorylation term is also annotated to YMR086W with low confidence; the same term is also given on the web page of Laboratory of Bioinformatics and Functional Genomics, Princeton University. Only low confidence terms (transport, metabolic process, and regulation) can be assigned to the following unknown proteins; YJL151C, YDR032C, YKL065C, and YHL048W. Pfam A domain (PF05529) of YKL065C is annotated to intracellular protein transport supporting our low confidence finding.

5.3.5. Determination of Potential Drug Targets and their Topological Analysis

A novel strategy used in cancer therapy is the manipulation of sphingolipid metabolism to induce apoptosis (Dickson, 2008; Ogretmen, 2006). In our previous studies, potential drug target enzymes (critically important) were identified using both dynamical (Chapter 3), (Kavun Ozbayraktar and Ulgen, 2010) and stoichiometric models (Chapter 4). There are also commercially available drug targets involved in sphingolipid metabolism for cancer therapy. We first attempted to identify the characteristic topological properties that make any protein a drug target using the newly constructed yeast sphingolipid PIN. Specifically closeness centrality, betweenness centrality, degree, clustering coefficient and average shortest path length of the potential drug targets appearing in sphingolipid PIN are

examined (Table 5.6). The drug target proteins are not among hubs and bottlenecks. We compare the topological parameters of the drug targets with those of the hub proteins. The average shortest path lengths of these drug targets (ranging from 2.74 to 3.29) are higher than those of the hubs. Having small shortest path length means that it is easy to reach the nodes in the neighborhood, which also means that any attack/perturbation on that node spreads over others easily. Drug targets should affect the required mechanisms easily but all the neighbor nodes should not be affected, this introduces the concept of side effects (Berger and Iyengar, 2009). Closeness centrality is used to quantify how fast information travels between nodes (Newman, 2003) and betweenness centrality is used to measure the amount of control that one node has over the network (Yoon *et al.*, 2006). The betweenness centrality of the drug targets (ranging from 0.0063 to 0.0912) are smaller than those of the hub proteins, indicating that drug targets do not have control effect over the network as strong as the hubs have. High betweenness centrality diminishes the possibility of a node being an ideal drug target (Hwang *et al.*, 2008). However, the closeness centrality is almost as high as the hub proteins' (ranging from 0.30 to 0.36), indicating that the drug targets spread information to reachable nodes rather fast. The clustering coefficients of drug targets are higher than that of the hubs on average. Drug targets are more likely to be involved in functional modules than hubs, whereas hubs generally stand in between numerous functional processes having control effect over all (Berger and Iyengar, 2009). The most frequent and significant GO process term in the modules that involve drug targets is transport process. Our drug targets are not hub proteins that have strong control effect over the network but they stand on a topological position such that in case of their perturbation, the flow of information is prevented. In sphingolipid metabolism compartmentalization is very important (Kitagaki *et al.*, 2007) and any disruption in their transport may affect the whole metabolism.

Table 5.6. Topological properties of drug targets. Average shortest path length, betweenness centrality, closeness centrality, clustering coefficient, and degree are tabulated for commercial and potential drug target proteins proposed in Chapter 3 and 4 (Kavun Ozbayraktar and Ulgen, 2009)

Commercial drug targets	**Gene name**	**Average shortest path length**	**Betweenness centrality**	**Closeness centrality**	**Clustering coefficient**	**Degree**
Serine palmitoyl transferase	YMR296C	4.00	0.000328	0.250118	0.33	4
	YDR062W	3.55	0.002300	0.281565	0.20	12
	YBR058C-A	4.49	0	0.222939	1.00	2
Sphingoid base kinase	YOR171C	4.16	0.000207	0.240617	0	2
	YLR260W	3.92	0.000000	0.255217	0.00	1
Ceramide synthase	YKL008C	3.16	0.013388	0.315978	0.23	24
	YHL003C	3.60	0.002830	0.277438	0.11	12
	YMR298W	3.98	0.000005	0.251383	0.67	3
Drug tagets from stoichiometric model	**Gene name**	**Average shortest path length**	**Betweenness centrality**	**Closeness centrality**	**Clustering coefficient**	**Degree**
MIPC synthase	YBR036C	3.46	0.004273	0.288933	0.23	28
$M(IP)_2C$ synthase	YDR072C	4.03	0	0.248282	0	1
IPC synthase	YKL004W	3.51	0.001848	0.285048	0.18	14
MIPC synthase	YPL057C	4.02	0.000028	0.248865	0.50	4

Table 5.6. Topological properties of drug targets. Average shortest path length, betweenness centrality, closeness centrality, clustering coefficient, and degree are tabulated for commercial and potential drug target proteins proposed in Chapter 3 and 4 (Kavun Ozbayraktar and Ulgen, 2009), (cont.)

Drug tagets from dynamic model	Gene name	Average shortest path length	Betweenness centrality	Closeness centrality	Clustering coefficient	Degree
Palmitoyl-Coenzyme A synthase	YBR041W	3.55	0.003262	0.281865	0.19	20
	YOR317W	3.60	0.000565	0.277729	0	5
	YMR246W	3.83	0.002279	0.261298	0	7
Glycerol-3-phosphate acyltransferase	YKR067W	3.84	0.000028	0.260187	0.33	3
	YBL011W	3.18	0.003016	0.314789	0.03	15
Cytidine diphosphate-diacylglycerol synthase	YBR029C	-	-	-	-	-
Phosphatidyl-inositol synthase	YPR113W	-	-	-	-	-
Inositol-1-phosphate synthase	YJL153C	3.42	0.005114	0.292602	0.05	12
Phosphoserine Phosphatase	YGR208W	-	-	-	-	-
Serine palmitoyl transferase	YMR296C	4.00	0.000328	0.250118	0.33	4
	YDR062W	3.55	0.002300	0.281565	0.20	12
	YBR058C-A	4.49	0.000000	0.222939	1.00	2
Acyl-Coenzyme A binding protein	YGR037C	3.37	0.003321	0.296974	0.04	11
Sphingoid base kinase	YOR171C	4.16	0.000207	0.240617	0	2
	YLR260W	3.92	0	0.255217	0	1
4-Hydroxylase	YDR297W	3.64	0.001427	0.274422	0.42	19

Table 5.6. Topological properties of drug targets. Average shortest path length, betweenness centrality, closeness centrality, clustering coefficient, and degree are tabulated for commercial and potential drug target proteins proposed in Chapter 3 and 4 (Kavun Ozbayraktar and Ulgen, 2009), (cont.)

Drug tagets from dynamic model	**Gene name**	**Average shortest path length**	**Betweenness centrality**	**Closeness centrality**	**Clustering coefficient**	**Degree**
Ceramide synthase	YKL008C	3.16	0.013388	0.315978	0.23	24
	YHL003C	3.60	0.002830	0.277438	0.11	12
	YMR298W	3.98	0.000005	0.251383	0.67	3
Glycosylphosphatidylinositol remodelase	YGL084C	-	-	-	-	-
Acetyl-Coenzyme A Synthetase	YAL054C	3.84	0.000856	0.260357	0	3
	YLR153C	3.35	0.002915	0.298928	0.12	12

The knowledge achieved from topological analysis of commercial and predetermined putative drug targets leads us to search for novel plausible drug targets in sphingolipid metabolism. The top 100 proteins with low betweenness centrality as well as with high closeness centrality are selected among the proteins with low degree in the newly reconstructed network. These proteins are screened for the ones that do not have any unknown GO term and that function directly in sphingolipid metabolism. This algorithm pointed out two proteins, sphingoid base kinase (YLR260W) and $M(IP)_2C$ synthase (YDR072C). Sphingoid base kinase is responsible for the phosphorylation of sphingoid long-chain bases, namely dihydrosphingosine and phytosphingosine in yeast. The enzyme is already a commercially available drug target. Although YLR260W gene product performs a minor fraction of total kinase activity (Dickson *et al.*, 2006), it is the best candidate in terms of topological properties. The phosphorylated sphingoid base in human is known for its tumor promoting properties due to its regulatory function in cell proliferation (Taha *et al.*, 2006; Zeidan and Hannun, 2007). $M(IP)_2C$ synthase is responsible for the biosynthesis of $M(IP)_2C$, which is the most abundant sphingolipid found in yeast plasma membrane (Dickson *et al.*, 1997) and the end product of *de novo* sphingolipid biosynthesis. The enzyme has also been identified as a novel potential drug target in our previous stoichiometric modeling study in terms of its metabolic consequences (Chapter 4).

5.4. Concluding Remarks

Experimental studies of sphingolipid metabolism have generally focused on a small part of the whole sphingolipid system. Computational tools are additionally required to obtain an insight about the complete mechanism governing sphingolipid related processes. We here aimed to determine all the proteins and their interactions responsible for metabolic, regulatory, and transport processes of sphingolipids in yeast. Selective permissibility algorithm is utilized to obtain a final comprehensive network of 1591 proteins. Five isolated proteins are rewired to the network by developing an algorithm which integrates the computational interaction prediction tools, orthology, expression profiles, co-localization information, and gene ontology data. The ultimate sphingolipid protein interaction network shares the same topological properties as all biological networks. The newly constructed network forms an adequate platform to assign function to

uncharacterized proteins. The here developed multi-dimensional function annotation method proves to be efficient for the test set (gold-proteins) and gives seven high confidence function annotations for 11 uncharacterized proteins of the present network. The topological analysis of commercial drug targets reported in literature outlines the characteristics of an ideal drug target. Sphingolipid protein interaction network is further used to identify novel potential drug targets for cancer. A drawback of the current study is that it barely uses secondary structures of proteins due to the lack of complete structural information of membrane proteins. Although gene ontology terms inferred from electronic annotation are not utilized in the current study, error propagation due to deficiencies of databases is inevitable. Nevertheless the present methodology along with the constructed protein interaction network can be employed to computationally predict missing protein interactions, annotate function, and propose drug targets for other sphingolipid related diseases.

5.5. Human Sphingolipid Protein-Protein Interaction Network Reconstruction and Analysis

The ultimate goal of this study is to examine and understand the human sphingolipid metabolism. The information gained so far can be further used in proposing cancer therapeutics based on manipulation of sphingolipid metabolism. This section is devoted to the discussion of human sphingolipid protein interaction network reconstruction and analysis.

5.5.1. Network Reconstruction

In order to obtain a plausible protein interaction network representing the sphingolipid metabolism, signalling and transport, one has to select the pioneering core proteins carefully. We have first started with 19 proteins functioning in the *de novo* sphingolipid metabolism listed in Table 5.7. These 19 proteins include the enzymes that are responsible for the *de novo* biosynthesis of ceramide starting from condensation of serine and palmitoyl-coenzyme A. The enzymes responsible for phosphorylation and dephosphorylation of ceramide and long chain bases are also included. Since there are still a large number of proteins with unknown function, process and component Gene Ontology

(GO) terms in the human proteome, the unknown GO terms are not appended to the annotation table utilized in this study. The reason of this exclusion is to avoid basing network reconstruction process upon excessive uncertain data. Eight runs of selective permissibility algorithm resulted in a network of 1708 proteins and 6446 edges. However, only three out of 19 core proteins are wired to the network reconstructed. Serine palmitoyl transferase 1, sphingosine kinase 1 and lipid phosphate phosphohydrolase 2 are the proteins that are connected to the newly constructed network. Remaining 16 core proteins are found to be isolated due to the lack of interaction data. 16/19 isolated proteins give a very low coverage rate of core proteins, and this ratio points out a failure in network reconstruction process.

We have reconstructed a second network using the same core proteins but we have appended the unknown GO terms to the annotation table. The reason of this addition is to examine the effect of excluding\including unknown GO terms on the number of isolated core proteins. A larger network of protein interactions are obtained with 1845 nodes and 6744 edges at the end of eight runs of selective permissibility algorithm, but still 15 core proteins remained unwired to the network. It can be concluded that the problem of a very high ratio of isolated core proteins could not be solved by the inclusion of unknown GO terms to the annotation table.

In order to unravel the reason and eliminate the isolation of core proteins, a more extensive set of core proteins are selected. KEGG pathway database is used to identify all the sphingolipid related proteins. 84 core proteins are identified functioning in sphingolipid metabolism, glycosphingolipid biosynthesis of lacto, neolacto, globo and ganglio series. The most significant GO process term enriched among core proteins is sphingolipid metabolic process with a p-value of 1.62×10^{-8}. The core proteins are most significantly and frequently found in the Golgi cisterna membrane (p-value = 5.28×10^{-6}, see Table 5.8 for details). Eight runs of selective permissibility algorithm resulted in a network of 3097 nodes and 11064 edges. Sequential addition of proteins at each run is summarized in Figure 5.5. 13 out of 84 core proteins are connected to the bulk network. There is still a very high ratio of isolated core proteins. Although most of the proteins functioning in sphingolipid metabolism are taken into consideration, the isolated proteins still remained unbound. It can be concluded that the incompleteness of the network reconstruction

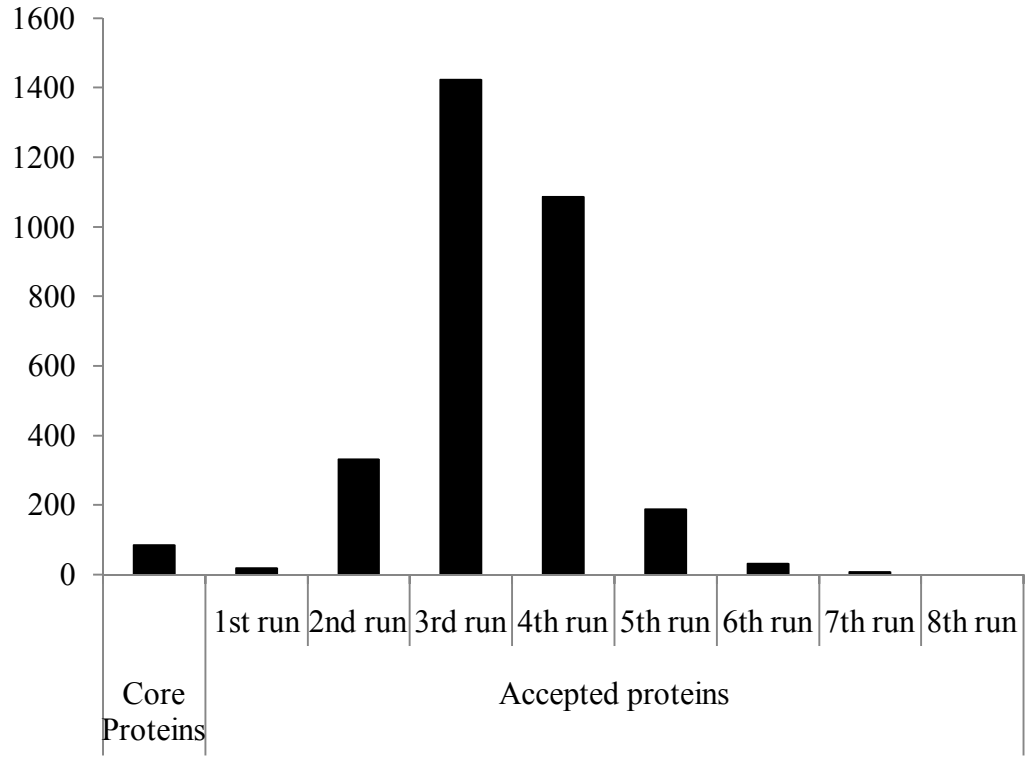
process is due to the lack of interaction data for sphingolipid metabolism related proteins. It is not reliable to computationally predict interactions for 85 % of the core proteins; it may lead to error propagation to base predictions upon other predictions.

Table 5.7. 19 core proteins belonging to *de novo* sphingolipid metabolism

Official Symbol	**Uniprot ID**	**Protein Name**
SPTLC1	O15269	Serine palmitoyltransferase 1
SPTLC3	Q9NUV7	Serine palmitoyltransferase 3
SPTLC2	O15270	Serine palmitoyltransferase 2
KDSR	Q06136	3-ketodihydrosphingosine reductase
ASAH1	Q13510	Acid ceramidase
ASAH2	Q9NR71	Neutral ceramidase
ACER1	Q8TDN7	Alkaline ceramidase 1
ACER2	Q5QJU3	Alkaline ceramidase 2
ACER3	Q9NUN7	Alkaline ceramidase 3
DEGS2	Q6QHC5	Sphingolipid delta(4)-desaturase
DEGS1	O15121	Sphingolipid delta(4)-desaturase
SPHK2	Q9NRA0	Sphingosine kinase 2
SPHK1	Q9NYA1	Sphingosine kinase 1
SGPP1	Q9BX95	Sphingosine-1-phosphate phosphatase 1
SGPP2	Q8IWX5	Sphingosine-1-phosphate phosphatase 2
PPAP2A	O14494	Lipid phosphate phosphohydrolase 1
PPAP2C	O43688	Lipid phosphate phosphohydrolase 2
PPAP2B	O14495	Lipid phosphate phosphohydrolase 3
CERK	Q8TCT0	Ceramide kinase

The 13 core proteins that are wired to the network are the most studied sphingolipid related proteins in literature. Serine palmitoyl transferease 1 is responsible for the pioneering reaction of *de novo* sphingolipid biosynthesis. Sphingolipid delta (4) desaturase converts dihydroceramide into active signalling molecule, ceramide. Lipid phosphate phosphohydrolases dephosphorylate their substrates to yield ceramide, sphingosine and dihydrosphingosine. Sphingosine kinase is responsible for the phosphorylation of long

chain bases to yield sphingosine-1-phosphate with anti-apoptotic effects. The delicate balance between tumour-suppressor lipids (sphingosine and ceramide) and tumour-promoting lipid (sphingosine-1-phosphate) is named sphingolipid rheostat and constitutes the most studied part of sphingolipid biology. The interaction data of proteins that are responsible for the balance of cell death and proliferation are known whereas the other parts of sphingolipid metabolism still remain uncertain. Ceramide glucosyltransferase has known links to be connected to the newly constructed network. Two disease (galactosialidosis, gangliosidosis, sialodosis) gene products, beta-galactosidase and sialidase, are among the 13 core proteins. Beta-galactosidase is responsible for the conversion of lactosylceramide to glucosylceramide in the sphingolipid biosynthesis pathway and it also functions in the ganglioside biosynthesis during the synthesis of GM_2 to GM_1. The absence of this enzyme leads to galactosialidosis and GM_1 gangliosidosis, which are both autosomal recessive lysosomal storage disorders. Sialidase is responsible for the conversion of ganglioside GM_4 to galactosylceramide. Sialidase enzyme deficiency causes autosomal recessive lysosomal storage disorders: sialidosis and galactosialidosis (KEGG). Beta-galactosidase and sialidase are among the non-hub proteins with only one and two degrees, respectively. The reason of low connectivity of the disease gene products may be due to the lack of interaction data, and not due to actual topological properties of the proteins. Low connectivity (missing of interaction data) reduces the chance of identifying the functional modules in which these proteins take place. Five proteins from glycosphingolipid biosynthesis pathways have known interactions to be connected to the network.



Core Proteins	Accepted proteins							
	1st run	2nd run	3rd run	4th run	5th run	6th run	7th run	8th run
84	18	332	1422	1086	188	31	8	0
Total number of proteins at the end of each run								
0th run	**1st run**	**2nd run**	**3rd run**	**4th run**	**5th run**	**6th run**	**7th run**	**8th run**
84	102	434	1856	2942	3130	3161	3169	3169

Figure 5.5. Number of proteins added to the newly constructed network at each run of selective permissibility algorithm

Table 5.8. GO enrichment analysis for 84 core proteins of sphingolipid metabolism. The top three most significant and frequent GO process, component and function terms are listed in a comparative manner with respect to the whole annotation

GO-ID	**Enriched GO process term**	**P-value**	**Cluster frequency**	**Total frequency**
6665	Sphingolipid metabolic process	1.62E-08	5/13 38.4%	44/14846 0.2%
5975	Carbohydrate metabolic process	1.20E-04	5/13 38.4%	295/14846 1.9%
46521	Sphingoid catabolic process	8.88E-03	1/13 7.6%	1/14846 0.0%
GO-ID	**Enriched GO component term**	**P-value**	**Cluster frequency**	**Total frequency**
32580	Golgi cisterna membrane	5.28E-06	4/13 30.7%	69/16926 0.4%
5764	Lysosome	3.20E-03	3/13 23.0%	158/16926 0.9%
5794	Golgi apparatus	3.57E-03	5/13 38.4%	889/16926 5.2%
GO-ID	**Enriched GO function term**	**P-value**	**Cluster frequency**	**Total frequency**
3831	Beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase activity	5.07E-05	2/13 15.3%	3/16253 0.0%
3945	N-acetyllactosamine synthase activity	5.07E-05	2/13 15.3%	4/16253 0.0%
4308	Exo-alpha-sialidase activity	5.07E-05	2/13 15.3%	4/16253 0.0%

5.5.2. Network Analysis

The topological parameters of both yeast and human sphingolipid networks in a comparative manner with examples from literature are given in Table 5.9. A power law degree distribution is obtained for human sphingolipid protein interaction network with a degree exponent of 1.68 (Figure 5.6). The degree exponent less than three indicates a scale-free network, which in turn points out the existence of hub proteins (Barabasi and Oltvai, 2004). As degree exponent gets much smaller than three, then the importance of hubs in the network becomes more significant. Most of the biological networks are scale-free networks, with a few highly connected nodes and most of the nodes with just a few degrees; shortly the network is not uniform (Barabasi and Oltvai, 2004). The existence of hub proteins in biological networks makes the network robust for disturbances (Yu *et al.*, 2004). Non-uniformity of a network's degree distribution is also quantified with network heterogeneity (Dong and Horvath, 2007). The human sphingolipid network is heterogeneous (1.66) and it is more heterogeneous than the yeast sphingolipid network (1.34). Characteristic path length, which evaluates the accessibility/interconnectivity of the network, introduces the concept of small-world (Watts and Strogatz, 1998; Li *et al.*, 2006). The criterion of being a small-world network is defined as having characteristic path length on the order of natural logarithm of the total number of nodes in the network (Cohen and Havlin, 2003). The characteristic path length of human sphingolipid network is calculated to be 4.10, which is less than the natural logarithm of total number of the nodes (3097). It can be concluded that the newly constructed network of human sphingolipids is a small-world network (easily interconnected, four edges is sufficient to link any two nodes on average). The network diameter for human sphingolipid network is calculated as 13, which is consistent with the range of data obtained from the literature surveys. The average clustering coefficient of human sphingolipid network is calculated as 0.13, which shows the tendency of the nodes of the network towards clustering (Barabasi and Oltvai, 2004). When this value is compared with the ones obtained from the literature (Table 5.9), we can conclude that our network is modular as most of the complex and real networks (Ravasz *et al.*, 2002). Although we have an incomplete network of protein interactions (covering only 15 % of the core proteins), the topological properties of the network is similar to properties of all biological networks.

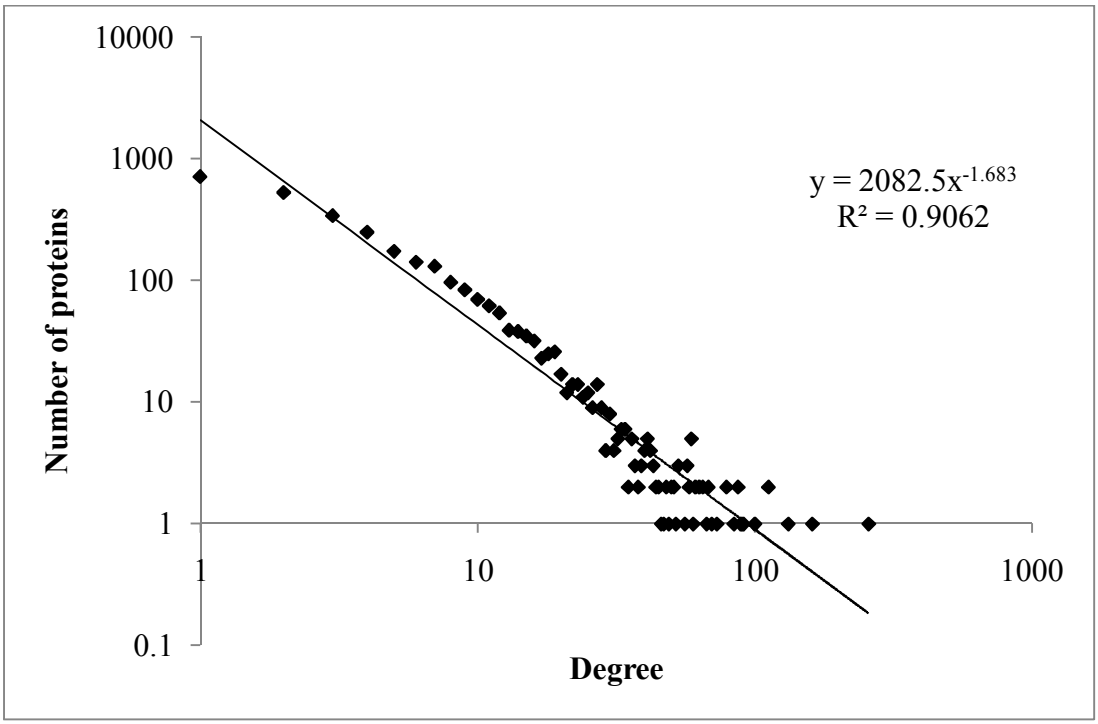


Figure 5.6. A power law degree distribution is obtained for the network with 3097 nodes. Degree exponent of 1.68 is shown

Signal transduction and apoptosis are the most significantly and frequently enriched GO process terms in the newly constructed network of 3097 proteins. Although degree is known to be a better indicator of essentiality in interaction networks, bottleneck proteins are more likely to be essential in networks that have proteins involving in mostly signalling processes (Yu *et al.*, 2007). As 591 out of 3097 proteins function in signal transduction processes, it can be concluded that bottleneck proteins should be examined instead of proteins with highest degree, namely hubs (Table 5.10). When first ten bottleneck proteins are examined, it is seen that they also have high degree. It can be concluded that bottleneck proteins are among highest degree proteins in human sphingolipid protein interaction network. Almost all the bottleneck proteins are also hub proteins in the newly constructed network.

Table 5.9. Topological properties of human sphingolipid protein interaction network

Topological parameters	Yeast Sphingolipid PIN	Human Sphingolipid PIN	References
Degrees	1591	3097	-
Edges	5355	11064	-
Degree exponent	1.68	1.68	1.8 (*S. cerevisiae* PIN Li *et al.*, 2006) 1.6 (*C. elegans* PIN Li *et al.*, 2006) 2.0 (*D. melanogaster* PIN Li *et al.*, 2006) 2.34 (*S. cerevisiae* PIN from MIPS Yook *et al.*, 2004) 2.50 (*S. cerevisiae* PIN from DIP Yook *et al.*, 2004)
Network heterogeneity	1.34	1.66	-
Characteristic path length	3.97	4.10	4.2 (*S. cerevisiae* PIN Li *et al.*, 2006) 4.8 (*C. elegans* PIN Wu *et al.*, 2005; Li *et al.*, 2006) 4.5 (*D. melanogaster* PIN Li *et al.*, 2006) 4.4 (D. melanogaster PIN Wu *et al.*, 2005) 5.0 (*S. cerevisiae* PIN Wu *et al.*, 2005) 4.1 (*H. pylori* PIN Wu *et al.*, 2005) 6.8 (*H. sapiens*PIN Wu *et al.*, 2005) 5.5 (*E. coli*PIN Wu *et al.*, 2005) 3.6 (*M. Musculus* PIN Wu *et al.*, 2005) 3.7 (*H. sapiens*PIN Bader and Madduri, 2007)
Network diameter	10	13	13 (*S. cerevisiae* PIN Wu *et al.*, 2005) 14 (*C. elegans* PIN Wu *et al.*, 2005) 11 (*D. melanogaster* PIN Wu *et al.*, 2005) 9 (*H. pylori* PIN Wu *et al.*, 2005) 21 (*H. sapiens*PIN Wu *et al.*, 2005) 16 (*E. coli*PIN Wu *et al.*, 2005) 9 (*M. Musculus* PIN Wu *et al.*, 2005) 14 (*H. sapiens*PIN Bader and Madduri, 2007)
Clustering coefficient	0.132	0.130	0.33 (*S. cerevisiae* PIN Wu *et al.*, 2005) 0.06 (*C. elegans* PIN Wu *et al.*, 2005) 0.02 (*D. melanogaster* PIN Wu *et al.*, 2005) 0.08 (*H. pylori* PIN Wu *et al.*, 2005) 0.21 (*H. sapiens*PIN Wu *et al.*, 2005) 0.62 (*E. coli*PIN Wu *et al.*, 2005) 0.15 (*M. Musculus* PIN Wu *et al.*, 2005)

Polyubiquitin have distinct functions in numerous processes like DNA repair, endoplasmic reticulum-associated degradation, cell-cycle regulation, lysosomal degradation, kinase modification and endocytosis (UniProt). Cellular tumour antigen p53 acts as a tumour suppressor in many tumour types inducing apoptosis. Epidermal growth factor receptor, growth factor receptor-bound protein 2 and RAF proto-oncogene serine/threonine-protein kinase function in the epidermal growth factor receptor (EGFR) signalling pathway, which is a part of cancer signalling pathway. NF-kappa-B essential

modulator functions in the tumour necrosis factor (TNF) alpha signalling pathway, which is also another signalling pathway important in cancer (Human Protein Reference Database). It can be concluded that almost all the bottleneck proteins have significantly important functions in vital processes.

Table 5.10. First ten bottleneck proteins of human sphingolipid protein interaction network are listed with their degree

#	UniProt ID	Protein description	Degree
1	P0CG48	Polyubiquitin-C	257
2	P04637	Cellular tumour antigen p53	161
3	P00533	Epidermal growth factor receptor	112
4	P37840	Alpha-synuclein (Non-A beta component of AD amyloid)	112
5	P35222	Catenin beta-1	87
6	P10275	Androgen receptor (Dihydrotestosterone receptor)	87
7	P36897	TGF-beta receptor type-1	79
8	P62993	Growth factor receptor-bound protein 2	132
9	P04049	RAF proto-oncogene serine/threonine-protein kinase	60
10	Q9Y6K9	NF-kappa-B essential modulator	70

5.5.3. Sphingolipid signalling in the regulation of the tumour necrosis factor (TNF) alpha induced cell death

TNF-alpha controls many immune responses including induction of apoptosis. A model representing the molecular mechanism of sphingolipid generation and TNF-alpha receptor 1 activation is available in literature (Malagarie-Cazenave *et al.*, 2002). 189 proteins are reported in Human Protein Reference Database as members of TNF-alpha signalling pathway. 124 out of 189 proteins are found in the newly constructed network of human sphingolipids. Human Protein Reference Database also reports 873 proteins that are transcriptionally regulated by TNF-alpha signalling pathway. 313 of these proteins are also involved in human sphingolipid protein interaction network. In this section of the study, it is aimed to check whether the working model of regulation of sphingolipid generation following TNF-alpha receptor 1 activation (Malagarie-Cazenave *et al.*, 2002) is captured in the newly constructed network.

12 functional modules are calculated involving the proteins functioning in TNF-alpha signalling network. These modules are classified according to their scores calculated by BINGO. The score is the product of the density of the module and the number of members of the module. The module with the highest score (Figure 5.7) involves four proteins from TNF-alpha signalling pathway and six proteins that are transcriptionally regulated by TNF-alpha signalling pathway. This module has 13 highly connected proteins with a density of 0.5 (Table 5.11). 77 % of the members of the module are related to TNF-alpha signalling. Despite its incompleteness, the resultant newly constructed network of protein interactions represent TNF-alpha signalling pathway with a high coverage rate.

Table 5.11. 13 Protein members of the TNF-alpha signalling module

UniProt ID	**Protein description**
P19438	Tumor necrosis factor receptor 1
Q13158	FAS-associated death domain protein
Q14790	Caspase-8
Q92851	Caspase-10
P55957	BH3-interacting domain death agonist
P45983	Mitogen-activated protein kinase 8 (MAP kinase 8)
P61586	Transforming protein RhoA
P52565	Rho GDP-dissociation inhibitor 1
P15311	Ezrin
P26038	Moesin
P25445	Tumor necrosis factor receptor superfamily member 6
P48023	Tumor necrosis factor ligand superfamily member 6
O14763	Tumor necrosis factor receptor superfamily member 10B

Since we are dealing with an incomplete network due the lack of interaction data, the module with the highest score is selected and examined in further sections of this study. Activation of pro-apoptotic gene products (p-value=1.85×10^{-14}, frequency=54 %) and induction of apoptosis by extracellular signals (p-value=3.05×10^{-13}, frequency=62 %) are most significantly and frequently enriched GO process terms in this functional module. In the plasma membrane, TNF-alpha binds to TNF-alpha receptor 1 (TNFR1, P19438), triggering the trimerisation of TNFR1 which then induces the recruitment of two adapters

to the death domain of TNFR1: TNFR-associated death domain protein (TRADD, Q15628) and Fas-associated death domain protein (FADD, Q13158). The recruitment of two adapter proteins ends in the autocatalytic cleavage of pro-caspase-8 that activates caspase-8 (Q14790). Caspase-8 cleaves BH3-interacting domain death agonist (BID, P55957). The cleavage translocates BID to mitochondria and activates the release of cytochrome c which is an apoptogenic factor resulting in the activation of executioner caspases and finally induction of apoptosis. The trimerisation TNFR1 and binding of two adapter proteins (TRADD and FADD) stimulate ceramide generation via degradation of sphingomyelin into ceramide (Malagarie-Cazenave *et al.*, 2002). The module mimics the binding of FADD to TNFR1, activation of caspase-8 and cleavage of BID (Figure 5.6). Ezrin (P15311) and Moesin (P26038) function in the module, and they both involve in connections of major cytoskeletal structures to the plasma membrane (UniProt). Probably, they have critical roles through the binding process of TNF-alpha to TNFR1. In the functional module under the investigation, two proteins (P61586 and P52565) regulate a signal transduction pathway linking plasma membrane receptors to the assembly of focal adhesions and action stress fibers. Xu *et al.* (2006) reported that TNF-alpha stimulation provided an increase in protein activity of P61586. This information explains the existence of these proteins in the functional module. It is also known that TNF-alpha signals via unknown second messengers lead to a weak stimulation of MAP kinase (Westwick *et al.*, 1995), which is also available in the module. It can be concluded that the functional module depicted in Figure 5.7 represents the working model of regulation of sphingolipid generation and TNF-alpha induced apoptosis.

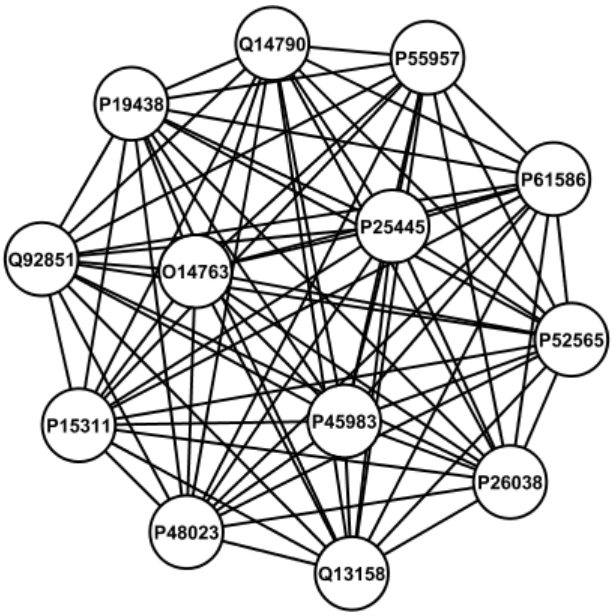


Figure 5.7. TNF-alpha signalling module

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

In the present thesis work, different systems biology tools are utilized to unravel sphingolipid metabolism in both yeast and human. Experimental studies of sphingolipid metabolism have generally focused on a small part of the whole sphingolipid system. Computational tools are additionally required to obtain an insight about the complete mechanism governing sphingolipid related processes. Diverse perspectives based on dynamic and stoichiometric modelling approaches as well as protein-protein interaction networks are attained within the framework of the research presented here. The ultimate goal of this work is to achieve a set of candidate drug target enzymes that can be further utilized in drug design studies for cancer therapy. The following conclusions are drawn integrating all the results gained.

6.1.1. Drug Target Identification in Sphingolipid Metabolism by Computational Systems Biology Tools: Metabolic Control Analysis and Metabolic Pathway Analysis

- In this study, metabolic control analysis and metabolic pathway analysis are performed in a comparative manner to identify potential drug target enzymes in sphingolipid pathway. Our aim was to rank the enzymes in sphingolipid pathway according to their roles in controlling the metabolic network by metabolic control analysis and to identify the physiologically connected reactions, *i.e.* biologically significant and functional modules of network by metabolic pathway analysis.
- According to the results of metabolic control and pathway analyses; the activities of palmitoyl-coenzyme A synthase, glycerol-3-phosphate acyltransferase, CDP-DAG synthase, phosphatidylinositol synthase, 4-hydroxylase, ceramide sythase and acetyl-coenzyme A synthethase should be increased to obtain ceramide accumulation and long chain base phosphate depletion and the activities of inositol-1-phosphate synthase, phosphoserine-phosphatase, serine palmitoyl transferase, acyl-coenyzme A binding protein, sphingolipid long chain base kinase and GPI

remodelase should be decreased in order to reach the same cancer therapy essentials.,

- When metabolic control and pathway analyses are compared in terms of their potency in potential drug target identification, it can be seen that the results of metabolic pathway analysis overlap more with the known drug targets in literature. The results of metabolic control analysis, on the other hand, are more informative since one can gather information about the direction of action, *i.e.* whether one should inhibit or activate the target enzyme. Thus, metabolic control analysis results go one step further on the way to drug discovery and drug design.

6.1.2. Stoichiometric Network Reconstruction and Analyses of Sphingolipid Metabolism

- The first comprehensive stoichiometric model of yeast sphingolipid metabolism is reconstructed with 60 internal metabolic reactions and 77 intracellular metabolites. The *in silico* model considers five states of hydroxylation, elucidating the significance of hydroxylation on sphingolipids.
- The hydroxylation products of phytoceramide in yeast are found to be the most favored species in the biosynthesis of complex sphingolipids as opposed to dihydroceramide derived complex sphingolipids.
- The phytoceramide derived complex sphingolipid syntheses constitute the most robust part of the metabolic yeast model.
- The significant reduction in ATP consumption, DAG biosynthesis, and complex sphingolipid synthesis (or its complete absence) are all coupled with each other and lead to either inviability or growth problems in yeast.
- Using the fact of IPC-C accumulation leading to calcium sensitivity as the driving force, the deletion mutants that accumulate IPC-C in the plasma membrane of the *in silico* cell are examined. The *in silico* yeast mutants o*f ipt1Δ* and *gda1Δ* accumulate IPC-C in the plasma membrane. These mutant strains may be calcium sensitive due to IPC-C accumulation.
- *In silico sur2Δcsg2Δ, scs7Δcsg2Δ* double mutants and *CSG2* gene deletion mutant lacking either hydroxylation of phytoceramide or synthesis of IPC-C from alpha-phytoceramide hinder IPC-C accumulation without complete shutdown of complex

sphingolipid synthesis. These four mutants are proposed as suppressors of calcium sensitivity of *csg2Δ* mutants.

- *IPT1, GDA1, CSG* and *AUR1* gene deletions may be novel candidates of drug targets for cancer therapy according to the results of flux balance and variability analyses coupled with robustness analysis.
- To get a deep insight as well as to observe the effect of variability on flux distribution and to determine the possible range of each flux with the flexibility of the system, the flux variability analysis was performed for the potential drug target enzymes proposed in this study. 98.4 % of the reactions in wild type strain have variable fluxes and 0.8 % of the reactions are 'never been used'. *IPT1, GDA1*, *CSG* and *AUR1* gene deletions have the highest percentage of variable fluxes among the deletion mutants studied, variabilities of 86.4 %, 72.0 %, 71.2 % and 56.8 % , respectively. The inhibition action of a drug on a target enzyme should not leave the whole system idle; but the flexibility of the system must rather be protected. A high flexibility makes the system under investigation more stable and resistant to stress conditions increasing the range of solution space.

6.1.3. Reconstruction and Analyses of Sphingolipid Protein-Protein Interaction Networks

- 32 core sphingolipid proteins are utilized in the reconstruction of a final yeast network of 1591 nodes and 5355 edges. The constructed network has a scale-free, small-world and modular structure, as it is the case for biological networks.
- 14 novel interactions in yeast network are predicted using a newly developed integrated methodology employing sequence and structure based computational interaction prediction tools, orthology, expression profiles, co-localization information and Gene Ontology terms.
- The function annotation of 11 uncharacterized proteins of the yeast network is successfully performed using the here developed multi-dimensional hybrid method which combines the results from modules and neighbors, and examines them by information gathered from genetic interactions, expression profiles, and sequence similarity.

- Sphingoid base kinase (YLR260W) and mannose diinositol-phosphorylceramide synthase (YDR072C) are identified as novel potential drug targets for cancer. Since they possess the topological properties of an ideal drug target with low control over neighboring proteins, they may reduce side effects of the drug, but spread information fast to reachable nodes.
- 84 core sphingolipid proteins are utilized in the reconstruction of a final human network of 3097 nodes and 11064 edges. The constructed network has a scale-free, small-world and modular structure, as it is in the case of biological networks. It can be concluded that the incompleteness of the network reconstruction process is due to the lack of interaction data for sphingolipid metabolism related proteins. It is not reliable to computationally predict interactions for 85 % of the core proteins; it may lead to error propagation to base predictions upon other predictions.
- The functional module representing TNF-alpha induced apoptosis including TNF-alpha receptor 1, Fas-associated death domain protein, caspase-8 and BH3-interacting domain death agonist is successfully depicted in the human sphingolipid protein-protein interaction network despite the lack of interactome data.

6.2. Recommendations

Within the framework of this thesis, drugs targeting each enzyme were assessed one at a time (Chapter 3), and the conclusions are drawn accordingly. Another option would be to develop drugs targeting two or more enzymes at the same time course. Inhibition and/or activation of two or more enzymes at the same time may introduce different control mechanisms and different side effects that should be considered in detail. In order to obtain more efficient drugs in cancer therapy; this would be the ideal method. There is also clinical evidence of using drugs both inhibiting ceramide catabolism and generating ceramide simultaneously in literature (Reynolds *et al.*, 2004; Zeidan and Hannun, 2007).

During the reconstruction of human sphingolipid metabolic model (Chapter 4), the stoichiometric model could not be completed due to lack of fluxome data which are necessary for the validation of the model. It would have been the ideal case to perform an experimental fluxome analysis concerning human sphingolipid reactions.

The structure based interaction prediction methods could not be used efficiently (Chapter 5) and the proposed drug target proteins could not be examined in terms of their structure. A drawback of the current study is that it barely uses secondary structures of proteins due to the lack of complete structural information of membrane proteins. It is the weak interactions as well as environment dependent interactions that form the native structures of membrane proteins. Any classical method for determining protein structure (X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy) becomes insufficient and challenging for most of the membrane proteins. Distinct solution preparation techniques should be used in finding structures of membrane proteins in their native states (Cross *et al.*, 2010). It would be a promising future study to find the three dimensional native structures of potential drug target enzymes proposed in this research.

During the reconstruction of human sphingolipid protein-protein interaction network, the lack of core protein interaction data has limited the completeness of the network. A revolutionary new approach of aptamer-based multiplexed proteomic technology may enlighten the missing pieces and may enable one to reconstruct a complete network (Gold *et al.*, 2010; personal communication with Dr. Larry Gold). The new technology of proteomics would help to construct validated cancer models and will surely replace microarray technologies in the near future.

Another promising approach would be to construct a network integrating metabolites and proteins at the same time. Such an integrated network would help discover the precise mechanisms of sphingolipid metabolites inducing apoptosis via protein interactions.

APPENDIX A: SUPPLEMENTARY TABLE FOR CHAPTER 3

In this section, the supplementary table mentioned in Chapter 3 (Drug Target Identification in Sphingolipid Metabolism by Computational Systems Biology Tools: Metabolic Control Analysis and Metabolic Pathway Analysis) are provided. Table A.1 cannot be given here explicitly since its dimensions are not appropriate for printing. Instead, only few representative columns are given. The complete form of Table A.1 is given in CD, it is saved as AppendixA_TA1.xlsx.

Table A.1. The homology and conservation between the enzymes of yeast and human sphingolipid pathways

Reaction #	Yeast Enzyme Name	ORF	Yeast Domains- InterPro Database	Human Homolog	E value	Positives	S-Prot ID	Human Domains- InterPro Database
11	Serine palmitoyl transferase	YMR296C	IPR001917. Aminotrans_II_pyridoxalP_BS. IPR004839. Aminotransferase_I/II. IPR015421. PyrdxlP-dep_Trfase_major_sub1.	serine palmitoyltransferase subunit 1 isoform a	2.00E-69	235/467 (50%)	O15269.1	IPR001917. Aminotrans_II_pyridoxalP_BS. IPR004839. Aminotransferase_I/II. IPR015421. PyrdxlP-dep_Trfase_major_sub1.
		YDR062W	IPR001917. Aminotrans_II_pyridoxalP_BS. IPR004839. Aminotransferase_I/II. IPR015421. PyrdxlP-dep_Trfase_major_sub1.	serine palmitoyltransferase, long chain base subunit 2	2.00E-124	310/449 (69%)	O15270	IPR001917. Aminotrans_II_pyridoxalP_BS. IPR004839. Aminotransferase_I/II. IPR015421. PyrdxlP-dep_Trfase_major_sub1.
		YBR058C-A		-	-			

Table A.1. The homology and conservation between the enzymes of yeast and human sphingolipid pathways (cont.)

Reaction #	Yeast Enzyme Name	ORF	Yeast Domains- InterPro Database	Human Homolog	E value	Positives	S-Prot ID	Human Domains- InterPro Database
13	Ketodihydro sphingosine reductase	YBR265W	IPR002198. DH_sc/Rdtase_SDR. IPR002347. Glc/ribitol_DH. IPR016040. NAD(P)-bd_dom.	3-ketodihydrosphingosine reductase precursor [Homo sapiens]	5.00E-13	130/310 (41%)	Q06136.1	IPR002198. DH_sc/Rdtase_SDR. IPR002347. Glc/ribitol_DH. IPR016040. NAD(P)-bd_dom.
14	Sphingoid base-1-phosphate phosphatase	YJL134W	IPR000326. P_Acid_Pase_2/haloperoxidase.	Sphingosine-1-phosphate phosphatase 1 [Homo sapiens]	3.00E-15	142/338 (42%)	Q9BX95.2	IPR000326. P_Acid_Pase_2/haloperoxidase.
15, 20	Sphingoid base kinase	YOR171C	IPR001206. Diacylglycerol_kinase_cat.	Sphingosine kinase [Homo sapiens]	2.00E-25	178/389 (45%)	Q9NYA1	IPR001206. Diacylglycerol_kinase_cat.
		YLR260W	IPR001206. Diacylglycerol_kinase_cat.	Sphingosine kinase 1 isoform 2 [Homo sapiens]	2.00E-22	103/194 (53%)	Q9NYA1	IPR001206. Diacylglycerol_kinase_cat.
17	Sphingosine phosphate lyase	YDR294C	IPR002129. PyrdxlP-dep_de-COase. IPR015421. PyrdxlP-dep_Trfase_major_sub1.	Sphingosine-1-phosphate lyase 1 [Homo sapiens]	8.00E-101	303/501 (60%)	O95470.3	IPR002129. PyrdxlP-dep_de-COase. IPR015421. PyrdxlP-dep_Trfase_major_sub1.
18	Sphingosine phosphate lyase	YDR294C	IPR002129. PyrdxlP-dep_de-COase. IPR015421. PyrdxlP-dep_Trfase_major_sub1.	Sphingosine-1-phosphate lyase 1 [Homo sapiens]	8.00E-101	303/501 (60%)	O95470.3	IPR002129. PyrdxlP-dep_de-COase. IPR015421. PyrdxlP-dep_Trfase_major_sub1.

Table A.1. The homology and conservation between the enzymes of yeast and human sphingolipid pathways (cont.)

Reaction #	Yeast Enzyme Name	ORF	Yeast Domains- InterPro Database	Human Homolog	E value	Positives	S-Prot ID	Human Domains- InterPro Database
19	sphingoid base-1-phosphate phosphatase	YJL134W	IPR000326. P_Acid_Pase_2/haloperoxidase.	Sphingosine-1-phosphate phosphatase 1 [Homo sapiens]	3.00E-15	142/338 (42%)	Q9BX95.2	IPR000326. P_Acid_Pase_2/haloperoxidase.
21	4-Hydroxylase	YDR297W	IPR006694. Fatty_acid_hydroxylase.	Sterol-C4-methyl oxidase-like isoform 2 [Homo sapiens]	1.00E-19	89/161 (55%)	Q15800	IPR006694. Fatty_acid_hydroxylase.
				Cholesterol 25-hydroxylase [Homo sapiens]	8.00E-09	63/140 (45%)	O95992.1	IPR006694. Fatty_acid_hydroxylase.
				Sterol-C5-desaturase [Homo sapiens]	2.00E-06	71/148 (47%)	O75845.2	IPR006694. Fatty_acid_hydroxylase.
22	Dihydro-ceramidase	YPL087W	IPR008901. APHC.	Phytoceramidase, alkaline [Homo sapiens]	3.00E-40	145/260 (55%)	Q9NUN7	IPR008901. APHC.

Table A.1. The homology and conservation between the enzymes of yeast and human sphingolipid pathways (cont.)

Reaction #	Yeast Enzyme Name	ORF	Yeast Domains- InterPro Database	Human Homolog	E value	Positives	S-Prot ID	Human Domains- InterPro Database
23	Ceramide synthase	YKL008C	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom. IPR013599. TRAM1.	LAG1 homolog, ceramide synthase 1 isoform 2 [Homo sapiens]	5.00E-11	105/258 (40%)	P27544	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom.
		YHL003C	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom. IPR013599. TRAM1.	LAG1 homolog, ceramide synthase 1 isoform 2 [Homo sapiens]	4.00E-08	60/123 (48%)	P27544	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom.
		YMR298W		-	-			
24	Ceramide synthase	YKL008C	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom. IPR013599. TRAM1.	LAG1 homolog, ceramide synthase 1 isoform 2 [Homo sapiens]	5.00E-11	105/258 (40%)	P27544	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom.
		YHL003C	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom. IPR013599. TRAM1.	LAG1 homolog, ceramide synthase 1 isoform 2 [Homo sapiens]	4.00E-08	60/123 (48%)	P27544	IPR016439. Longevity_assurance_LAG1_LAC1. IPR006634. TLC-dom.
25	Phytoceramide Alkaline Ceramidase	YBR183W	IPR008901. APHC.	Phytoceramidase, alkaline [Homo sapiens]	3.00E-38	150/271 (55%)	Q9NUN7	IPR008901. APHC.

Table A.1. The homology and conservation between the enzymes of yeast and human sphingolipid pathways (cont.)

Reaction #	Yeast Enzyme Name	ORF	Yeast Domains- InterPro Database	Human Homolog	E value	Positives	S-Prot ID	Human Domains- InterPro Database
26	Hydroxylase	YDR297W	IPR006694. Fatty_acid_hydroxylase.	Sterol-C4-methyl oxidase-like isoform 2 [Homo sapiens]	1.00E-19	89/161 (55%)	Q15800	IPR006694. Fatty_acid_hydroxylase.
				Cholesterol 25-hydroxylase [Homo sapiens]	8.00E-09	63/140 (45%)	O95992.1	IPR006694. Fatty_acid_hydroxylase.
				Sterol-C5-desaturase [Homo sapiens]	2.00E-06	71/148 (47%)	O75845.2	IPR006694. Fatty_acid_hydroxylase.
27	Glycosylphosphatidylinositol remodelase	YGL084C	IPR004299. MBOAT_fam.	Hedgehog acyltransferase [Homo sapiens]	8.00E-13	92/182 (50%)	Q5VTY9	IPR004299. MBOAT_fam.
28, 31, 40, 42, 43, 44	Inositol Phospho-sphingolipid Phospholipase C	YER019W	IPR005135. Endo/exonuclease/phosphatase.	Sphingomyelin phosphodiesterase 2, neutral [Homo sapiens]	5.00E-29	140/319 (43%)	O60906.2	IPR005135. Endo/exonuclease/phosphatase.

Table A.1. The homology and conservation between the enzymes of yeast and human sphingolipid pathways (cont.)

Reaction #	Yeast Enzyme Name	ORF	Yeast Domains- InterPro Database	Human Homolog	E value	Positives	S-Prot ID	Human Domains- InterPro Database
29, 30	Inositol phosphoryl-ceramide synthase	YKL004W	IPR000326. P_Acid_Pase_2/haloperoxidase.	Phosphatidic acid phosphatase type 2 domain containing 3, isoform	0.26	36/73 (49%)	Q8NBV4	IPR000326. P_Acid_Pase_2/haloperoxidase.
39, 39'	Mannosyl inositol phospho ceramide synthase	YPL057C	IPR007577. Glyco_trans_s_bd.	Alpha-1,4-N-acetylglucosaminyltransferase [Homo sapiens]	0.05	42/87 (48%)	Q9UNA3.1	IPR007652. A1-4-GlycosylTfrase_cons-reg. IPR007577. Glyco_trans_s_bd.
				Alpha-1,4-N-acetylglucosaminyltransferase [Homo sapiens]	0.066	42/87 (48%)	Q9UNA3	IPR007652. A1-4-GlycosylTfrase_cons-reg. IPR007577. Glyco_trans_s_bd.
49, 49'	Mannosyl diinositol phosphoryl ceramide synthase	YDR072C	IPR000326. P_Acid_Pase_2/haloperoxidase.	-	-			

APPENDIX B: SUPPLEMENTARY TABLES FOR CHAPTER 4

In this section, four supplementary tables mentioned in Chapter 4 (Stoichiometric Network Reconstruction and Analysis of Sphingolipid Metabolism) are provided.

Table B.1. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes

Reaction #	Enzyme Name	Reaction	Associated Genes
1	Palmitoyl-CoA synthase	ATP + Palm + CoA -> AMP + DP + PalCoA	FAA1
			FAA2
			FAA3
			FAA4
			FAT1
2	Acyl-CoA-Binding Protein	PalCoA -> PalCoAxt	ACB1
3	Long-chain-aldehyde dehydrogenase	PalmA + NAD + H2O -> Palm + NADH	HFD1-putative
4	Acetyl-Coenzyme A Synthetase	ATP + AC + CoA -> AMP + DP + AcCoA	ACS1
			ACS2
5	Acetyl-Coenzyme A Carboxylase	AcCoA + ATP + Carb -> ADP + H + MalCoA + P	HFA1
			ACC1
7	Fatty Acid Synthetase	AcCoA + 7 MalCoA + 14 NADPH + 7 H -> Palm + 7 CO2 + 14 NADP + 8 CoA + 6 H2O	ACP1
			CEM1
			MCT1
			OAR1
			ACB1
			FAS1
			FAS2
8	Very long chain fatty acid synthase	ATP + CoA + Hexco -> AMP + C26CoA + DP	FAT1
			TSC13 ?
9	Glycerol-3-Phosphate Acyltransferase	PalCoA + G3P -> CoA + PA	GAT1
			GAT2
			YDR018C
			SLC1 ?(2.3.1.51)

Table B.1. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes (cont.)

Reaction #	Enzyme Name	Reaction	Associated Genes
10	Phosphatidate phosphatase	H2O + PA -> DAG + P	DPP1
			LPP1
			PAH1
11	CDP-DAG synthase	PA + H + CTP -> CDPDAG + DP	CDS1
12	Inositol-1-Phosphate Synthase	G6P -> I1P	INO1
13	Inositol monophosphatase	I1P + H2O -> INO + P	INM1
14	Phosphatidylinositol synthase	CDPDAG + INO -> PI + CMP + H	PIS1
			YNR025C
15	Phosphatidylinositol kinase	ATP + PI -> ADP + H + P3I	VPS34
			ATG14
			VPS38
			VPS30
16	Phosphatidylinositol-3-phosphate 4-kinase	P3I + ATP -> PI34BP + ADP + H	VPS34 ?
			ATG14 ?
			VPS38 ?
			VPS30 ?
17	Phosphatidylinositol-4-kinase	PI + ATP -> P4I + ADP + H	STT4
			PIK1
			LSB6 ?? Type II
18	Phosphatidylinositol-4-phosphate 5-kinase	P4I + ATP -> PI45BP + ADP + H	MSS4
19	Phospholipase C	PI45BP + H2O -> ITP + DAG + H	PLC1
20	Phosphoserine Phosphatase	3PS + H2O -> P + SER	SER2
			YNL010W-similar to SER2
21	Phosphatidylserine Synthase	CDPDAG + SER -> CMP + H + PS	CHO1
22	Phosphatidylserine Decarboxylase	H + PS -> CO2 + PE	PSD1
			PSD2

Table B.1. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes (cont.)

Reaction #	Enzyme Name	Reaction	Associated Genes
23	Serine Hydroxymethyl Transferase	SER + THF -> GLY + H2O + MTHF	SHM1
			SHM2
24	Phosphoethanolamine Cytidylyltransferase	EP + CTP + H -> CDPE + DP	MUQ1
25	Diacylglycerol-Ethanolamine Phosphotransferase	DAG + CDPE -> CMP + H + PE	EPT1
26	serine C-palmitoyltransferase	H + PalCoA + SER -> KDHS + CO2 + CoA	LCB1
			LCB2
			TSC3
27	3-Dehydrosphinganine reductase	KDHS + H + NADPH -> NADP + DHS	TSC10
28	Phytosphingosine synthesis, 4-hydroxylase	H + NADPH + O2 + DHS -> H2O + NADP + PHS	SUR2
29	Ceramide-1 synthase	C26CoA + DHS -> DHCer + CoA + H	LAG1
			LAC1
			LIP1 ?
30	Ceramide-2 synthase	C26CoA + PHS -> PHCer + CoA + H	LAG1
			LAC1
			LIP1 ?
31	Dihydroceramide Alkaline Ceramidase	DHCer + CoA + H -> C26CoA + DHS	YDC1
32	Phytoceramide Alkaline Ceramidase	PHCer + CoA + H -> C26CoA + PHS	YPC1
33	Ceramide-1 hydroxylase	DHCer + H + NADPH + O2 -> PHCer + H2O + NADP	SUR2
34	GPI remodelase	PHCer -> PHCerxt	GUP1
35	Inositol phosphorylceramide synthase	DHCer + PI -> DAG + IPCA	AUR1
36	Inositol phosphorylceramide synthase	PHCer + PI -> DAG + IPCB	AUR1
37	Desaturase	PHCer + H + NADPH + O2 -> APHCer + H2O + NADP	SCS7
38	Cu(+2)-transporting P-type ATPase	APHCer + H + NADPH + O2 -> PAPHCer + H2O + NADP	CCC2
39	Hydroxylase	DHCer + H + NADPH + O2 -> ADHCer + H2O + NADP	SCS7
40	Inositol phosphorylceramide synthase	APHCer + PI -> IPCC + DAG	AUR1
41	Inositol phosphorylceramide synthase	ADHCer + PI -> IPCBP + DAG	AUR1
42	Inositol phosphorylceramide synthase	PAPHCer + PI -> IPCD + DAG	AUR1

Table B.1. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes (cont.)

Reaction #	Enzyme Name	Reaction	Associated Genes
43	Mannose-inositol phosphorylceramide synthase	GDPman + IPCA -> GDP + H + MIPCA	CSG1
			CSG2
44	Mannose-inositol phosphorylceramide synthase	GDPman + IPCB -> GDP + H + MIPCB	CSG1
			CSG2
45	Mannose-inositol phosphorylceramide synthase	IPCC + GDPman -> GDP + H + MIPCC	CSG1
			CSG2
46	Mannose-inositol phosphorylceramide synthase	IPCD + GDPman -> GDP + H + MIPCD	CSG1
			CSG2
47	Mannose-inositol phosphorylceramide synthase	IPCBP + GDPman -> GDP + H + MIPCBP	CSG1
			CSG2
48	Cu(+2)-transporting P-type ATPase	IPCC + H + NADPH + O2 -> IPCD + H2O + NADP	CCC2
49	Mannose-(inositol-P)2-ceramide synthase	MIPCA + PI -> DAG + M(IP)2CA	IPT1
50	Mannose-(inositol-P)2-ceramide synthase	MIPCB + PI -> DAG + M(IP)2CB	IPT1
51	Mannose-(inositol-P)2-ceramide synthase	MIPCD + PI -> DAG + M(IP)2CD	IPT1
52	Mannose-(inositol-P)2-ceramide synthase	MIPCC + PI -> DAG + M(IP)2CC	IPT1
53	Mannose-(inositol-P)2-ceramide synthase	MIPCBP + PI -> DAG + M(IP)2CBP	IPT1
54	Sphingolipid long chain base kinase (sphinganine)	ATP + DHS -> ADP + H + DHS1P	LCB4
			LCB5
55	Sphingolipid long chain base kinase (phytosphingosine)	ATP + PHS -> ADP + H + PHS1P	LCB4
			LCB5
56	Sphingoid base-phosphate phosphatase (sphinganine 1-phosphatase)	H2O + DHS1P -> P + DHS	LCB3
			LBP2
57	Sphingoid base-phosphate phosphatase (phytosphingosine 1-phosphate)	H2O + PHS1P -> P + PHS	LBP2 ?
58	Sphinganine phosphate lyase	DHS1P -> EP + PalmA	DPL1
59	Phytosphingosine phosphate lyase	PHS1P -> HPalmA + EP	DPL1 ?

Table B.1. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes (cont.)

Reaction #	Enzyme Name	Reaction	Associated Genes
60	Guanosine diphosphatase	GDP + H2O -> GMP + H + P	GDA1
61	Transport to plasma membrane	IPCA <-> IPCApm	
62	Transport to plasma membrane	IPCB <-> IPCBpm	
63	Transport to plasma membrane	MIPCA <-> MIPCApm	
64	Transport to plasma membrane	MIPCB <-> MIPCBpm	
65	Transport to plasma membrane	M(IP)2CA <-> M(IP)2CApm	
66	Transport to plasma membrane	M(IP)2CB <-> M(IP)2CBpm	
67	Transport to plasma membrane	IPCC <-> IPCCpm	
68	Transport to plasma membrane	IPCD <-> IPCDpm	
69	Transport to plasma membrane	IPCBP <-> IPCBPpm	
70	Transport to plasma membrane	MIPCC <-> MIPCCpm	
71	Transport to plasma membrane	MIPCD <-> MIPCDpm	
72	Transport to plasma membrane	MIPCBP <-> MIPCBPpm	
73	Transport to plasma membrane	M(IP)2CC <-> M(IP)2CCpm	
74	Transport to plasma membrane	M(IP)2CD <-> M(IP)2CDpm	
75	Transport to plasma membrane	M(IP)2CBP <-> M(IP)2CBPpm	
76	Transport to extracellular space	H <-> Hxt	
77	Transport to extracellular space	H2O <-> H2Oxt	
78	Transport to extracellular space	ACxt -> AC	
79	Transport to extracellular space	ATPxt -> ATP	
80	Transport to extracellular space	3PSxt -> 3PS	
81	Transport to extracellular space	G6Pxt -> G6P	
82	Transport to extracellular space	G3Pxt -> G3P	
83	Transport to extracellular space	CTPxt -> CTP	
84	Transport to extracellular space	INOxt -> INO	
85	Transport to extracellular space	THFxt -> THF	
86	Transport to extracellular space	NADPHxt -> NADPH	
87	Transport to extracellular space	O2xt -> O2	

Table B.1. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes (cont.)

Reaction #	Enzyme Name	Reaction	Associated Genes
88	Transport to extracellular space	Carbxt -> Carb	
89	Transport to extracellular space	Tetcoxt -> Tetco	
90	Transport to extracellular space	GDPmanxt -> GDPman	
91	Transport to extracellular space	AMP -> AMPxt	
92	Transport to extracellular space	DP -> DPxt	
93	Transport to extracellular space	P -> Pxt	
94	Transport to extracellular space	CMP -> CMPxt	
95	Transport to extracellular space	ADP -> ADPxt	
96	Transport to extracellular space	GLY -> GLYxt	
97	Transport to extracellular space	MTHF -> MTHFxt	
98	Transport to extracellular space	CO2 -> CO2xt	
99	Transport to extracellular space	NADP -> NADPxt	
100	Transport to extracellular space	PE -> PExt	
101	Transport to extracellular space	HPalmA -> HPalmAxt	
102	Transport to extracellular space	GMP -> GMPxt	
103	Transport to extracellular space	man -> manxt	
104	Transport to extracellular space	Palmxt -> Palm	
105	Transport to extracellular space	CoAxt -> CoA	
106	Transport to extracellular space	IPCApm -> IPCApmxt	
107	Transport to extracellular space	IPCBpm -> IPCBpmxt	
108	Transport to extracellular space	MIPCApm -> MIPCApmxt	
109	Transport to extracellular space	MIPCBpm -> MIPCBpmxt	
110	Transport to extracellular space	M(IP)2CApm -> M(IP)2CApmxt	
111	Transport to extracellular space	M(IP)2CBpm -> M(IP)2CBpmxt	
112	Transport to extracellular space	IPCCpm -> IPCCpmxt	
113	Transport to extracellular space	IPCDpm -> IPCDpmxt	
114	Transport to extracellular space	IPCBPpm -> IPCBPpmxt	

Table B.1. The metabolic reactions, the enzymes catalyzing each step and the genes encoding these enzymes (cont.)

Reaction #	Enzyme Name	Reaction	Associated Genes
115	Transport to extracellular space	MIPCCpm -> MIPCCpmxt	
116	Transport to extracellular space	MIPCDpm -> MIPCDpmxt	
117	Transport to extracellular space	MIPCBPpm -> MIPCBPpmxt	
118	Transport to extracellular space	M(IP)2CCpm -> M(IP)2CCpmxt	
119	Transport to extracellular space	M(IP)2CDpm -> M(IP)2CDpmxt	
120	Transport to extracellular space	M(IP)2CBPpm -> M(IP)2CBPpmxt	
121	Transport to extracellular space	DAG -> DAGxt	
122	Transport to extracellular space	NADH -> NADHxt	
123	Transport to extracellular space	NADxt -> NAD	
124	Transport to extracellular space	PI34BP -> PI34BPxt	
125	Transport to extracellular space	ITP -> ITPxt	

Table B.2. The metabolite abbreviations list

Metabolite Names	**Abbreviations**
O-phospho-L-serine; 3-phosphoserine	3PS
Acetate	AC
Acetyl-Coenzyme A	AcCoA
Alpha-hydroxy-dihydroceramide	ADHCer
Adenosine diphosphate	ADP
Adenosine monophosphate	AMP
Alpha-hydroxy-phytoceramide	APHCer
ATP	ATP
Hexasanoyl-Coenzyme A (n-C26:0 CoA)	C26CoA
Bicarbonate	Carb
Cytidine diphosphate-diacylglycerol	CDPDAG
Cytidine diphosphate-ethanolamine	CDPE
Cytidine monophosphate	CMP
Carbon dioxide	CO2
Coenzyme A	CoA
Cytidine triphosphate	CTP
1,2 diacylglycerol	DAG
Dihydroceramide	DHCer
Sphinganine, dihydrosphingosine	DHS
Sphinganine-1-phosphate, dihydrosphingosine-1-phosphate	DHS1P
Diphosphate	DP
Ethanolamine phosphate	EP
Glycerol-3-phosphate	G3P
Glucose-6-phosphate	G6P
Guanosine diphosphate	GDP
Guanosine diphosphate-D-mannose	GDPman
Glycine	GLY
Guanosine monophosphate	GMP
Hydrogen	H
Water	H2O
Hexacosanoate	Hexco
2-hydroxy-hexadecanal	HPalmA
Inositol-1-phosphate	I1P
Inositol	INO
Inositol phosphorylceramide-A (from dihyroceramide)	IPCA
Inositol phosphorylceramide-B (from phytoceramide)	IPCB
Inositol phosphorylceramide-B' (from alpha-hydroxy-dihydroceramide)	IPCBP
Inositol phosphorylceramide-C (from alpha-hydroxy-phytoceramide)	IPCC
Inositol phosphorylceramide-D (from alpha-dihydroxy-phytoceramide)	IPCD
1D-myo-Inositol 1,4,5-trisphosphate	ITP
3-dehydrosphinganine, 3-ketodihydrosphingosine	KDHS
Mannose-diinositol phosphorylceramide-A (from dihyroceramide)	M(IP)2CA
Mannose-diinositol phosphorylceramide-B (from phytoceramide)	M(IP)2CB
Mannose-diinositol phosphorylceramide-B' (from alpha-hydroxy-dihydroceramide)	M(IP)2CBP
Mannose-diinositol phosphorylceramide-C (from alpha-hydroxy-phytoceramide)	M(IP)2CC

Table B.2. The metabolite abbreviations list (cont.)

Metabolite Names	**Abbreviations**
Mannose-diinositol phosphorylceramide-D (from alpha-dihydroxy-phytoceramide)	M(IP)2CD
Malonyl-Coenzyme A	MalCoA
Mannose-inositol phosphorylceramide-A (from dihyroceramide)	MIPCA
Mannose-inositol phosphorylceramide-B (from phytoceramide)	MIPCB
Mannose-inositol phosphorylceramide-B' (from alpha-hydroxy-dihydroceramide)	MIPCBP
Mannose-inositol phosphorylceramide-C (from alpha-hydroxy-phytoceramide)	MIPCC
Mannose-inositol phosphorylceramide-D (from alpha-dihydroxy-phytoceramide)	MIPCD
5,10, methylenetetrahydrofolate	MTHF
Nicotinamide adenine dinucleotide	NAD
Nicotinamide adenine dinucleotide- reduced	NADH
Nicotinamide adenine dinucleotide phosphate	NADP
Nicotinamide adenine dinucleotide phosphate - reduced	NADPH
Oxygen	O2
Phosphate	P
Phosphatidyl-1D-myo-3-inositol	P3I
Phosphatidyl-1D-myo-4-inositol	P4I
Phosphatidic acid, phosphatidate	PA
Palmitoyl-Coenzyme A	PalCoA
Palmitate	Palm
Hexadecanal, palmitaldeyhe	PalmA
Alpha-dihydroxy-phytoceramide	PAPHCer
Phosphatidyl ethanolamine	PE
Phytoceramide	PHCer
Phytoshingosine	PHS
Phytoshingosine-1-phosphate	PHS1P
Phosphatidyl inositol	PI
Phosphatidyl-1D-myo-inositol 3,4-bisphosphate	PI34BP
Phosphatidyl-D-myo-inositol 4,5-bisphosphate	PI45BP
Phosphatidyl serine	PS
Serine	SER
Tetracosanoate	Tetco
Tetrahydrofolate	THF
Plasma membrane	pm
Extracellular space	xt

Table B.3. List of internal reactions that have always variable fluxes in all simulations according to flux variability analyses

Common Variable Reaction #s	**Common Variable Reactions in all *in silico*deletion mutants and wild type strain**	**Pathway**
2	PalCoA -> PalCoAxt	**Fatty Acid Reactions**
4	ATP + AC + CoA -> AMP + DP + AcCoA	
5	AcCoA + ATP + Carb -> ADP + H + MalCoA + P	
6	AcCoA + 7 MalCoA + 14 NADPH + 7 H -> Palm + 7 CO2 + 14 NADP + 8 CoA + 6 H2O	
7	3 H + MalCoA + 2 NADPH + Tetco -> CO2 + CoA + H2O + Hexco + 2 NADP	
8	ATP + CoA + Hexco -> AMP + C26CoA + DP	
20	3PS + H2O -> P + SER	**Phospholipid Rxns**
23	SER + THF -> GLY + H2O + MTHF	

Table B.3. List of internal reactions that have always variable fluxes in all simulations according to flux variability analyses

Common Variable Reaction #s	**Common Variable Reactions in all *in silico*deletion mutants and wild type strain**	**Pathway**
26	H + PalCoA + SER -> KDHS + CO2 + CoA	**Sphingolipid Reactions**
27	KDHS + H + NADPH -> NADP + DHS	
29	C26CoA + DHS -> DHCer + CoA + H	
30	C26CoA + PHS -> PHCer + CoA + H	
31	DHCer + CoA + H -> C26CoA + DHS	
32	PHCer + CoA + H -> C26CoA + PHS	
54	ATP + DHS -> ADP + H + DHS1P	
55	ATP + PHS -> ADP + H + PHS1P	
56	H2O + DHS1P -> P + DHS	
57	H2O + PHS1P -> P + PHS	

Table B.4. Percent changes of total complex sphingolipid content, ATP consumption, DAG biosynthesis and ratio of transfer rate of palmitoyl-coenzyme A to non-sphingolipid processes for all deletion mutants with respect to wild type strain and list of PA/DAG, and DAG/Ceramide ratios. INF stands for infinity

Mutants/ Blocked Reactions	Total Plasma Membrane Sphingolipid Content (% change wrt WT)	ATP Consumption (Reaction 79) (% change wrt WT)	DAG Biosynthesis (Reaction 121) (% change wrt WT)	Palmitoyl-CoA to non-sphingolipid processes (Reaction 2) (Ratio wrt WT)	PA/DAG	DAG/Ceramide
1	-100.0	-49.2	-100.0	0	INF	INF
2	0.0	0.0	0.0	0	1.0	INF
3	0.0	0.0	0.0	1	1.0	INF
4	-100.0	-38.7	-72.3	341535	0.5	INF
5	-100.0	-38.7	-72.3	341535	0.5	INF
6	0.0	0.0	0.0	1	1.0	INF
7	-100.0	-38.7	-72.3	341535	0.5	INF
8	-100.0	-38.7	-72.3	341535	0.5	INF
9	-100.0	-35.0	-100.0	531266	INF	0
10	0.0	0.0	0.0	1	1.0	INF
11	-100.0	-40.4	-66.4	349554	0.0	5.6
12	0.0	0.0	0.0	1	1.0	INF
13	0.0	0.0	0.0	1	1.0	INF
14	-100.0	-41.2	-69.2	344962	0.2	5.6
15	0.0	0.0	0.0	4	1.0	INF
16	0.0	0.0	0.0	4	1.0	INF
17	0.0	0.0	0.0	1	1.0	INF
18	0.0	0.0	0.0	1	1.0	INF

Table B.4. Percent changes of total complex sphingolipid content, ATP consumption, DAG biosynthesis and ratio of transfer rate of palmitoyl-coenzyme A to non-sphingolipid processes for all deletion mutants with respect to wild type strain and list of PA/DAG, and DAG/Ceramide ratios. INF stands for infinity (cont.)

Mutants/ Blocked Reactions	Total Plasma Membrane Sphingolipid Content (% change wrt WT)	ATP Consumption (Reaction 79) (% change wrt WT)	DAG Biosynthesis (Reaction 121) (% change wrt WT)	Palmitoyl-CoA to non-sphingolipid processes (Reaction 2) (Ratio wrt WT)	PA/DAG	DAG/Ceramide
19	0.0	0.0	0.0	1	1.0	INF
20	-100.0	-41.4	-53.4	376568	0.3	INF
21	0.0	0.0	0.0	1	1.0	INF
22	0.0	0.0	0.0	1	1.0	INF
23	0.0	0.0	0.0	1	1.0	INF
24	0.0	0.0	0.0	1	1.0	INF
25	0.0	0.0	0.0	1	1.0	INF
26	-100.0	-43.4	-57.4	371190	0.4	INF
27	-100.0	-43.4	-57.4	371190	0.4	INF
28	0.0	0.0	0.0	1	1.0	INF
29	0.0	0.0	0.0	1	1.0	INF
30	0.0	0.0	0.0	4	1.0	INF
31	0.0	0.0	0.0	1	1.0	INF
32	0.0	0.0	0.0	1	1.0	INF
33	0.0	0.0	0.0	1	1.0	INF
34	0.0	0.0	0.0	1	1.0	INF
35	0.0	0.0	0.0	1	1.0	INF
36	0.0	0.0	0.0	1	1.0	INF

Table B.4. Percent changes of total complex sphingolipid content, ATP consumption, DAG biosynthesis and ratio of transfer rate of palmitoyl-coenzyme A to non-sphingolipid processes for all deletion mutants with respect to wild type strain and list of PA/DAG, and DAG/Ceramide ratios. INF stands for infinity (cont.)

Mutants/ Blocked Reactions	Total Plasma Membrane Sphingolipid Content (% change wrt WT)	ATP Consumption (Reaction 79) (% change wrt WT)	DAG Biosynthesis (Reaction 121) (% change wrt WT)	Palmitoyl-CoA to non-sphingolipid processes (Reaction 2) (Ratio wrt WT)	PA/DAG	DAG/Ceramide
37	-85.5	-36.9	-63.1	330163	0.4	INF
38	0.0	0.0	0.0	1	1.0	INF
39	0.0	0.0	0.0	1	1.0	INF
40	0.0	0.0	0.0	1	1.0	INF
41	0.0	0.0	0.0	1	1.0	INF
42	0.0	0.0	0.0	1	1.0	INF
43	0.0	0.0	0.0	4	1.0	INF
44	0.0	0.0	0.0	1	1.0	INF
45	0.0	0.0	0.0	1	1.0	INF
46	0.0	0.0	0.0	1	1.0	INF
47	0.0	0.0	0.0	1	1.0	INF
48	0.0	0.0	0.0	1	1.0	INF
49	0.0	0.0	0.0	1	1.0	INF
50	0.0	0.0	0.0	1	1.0	INF
51	0.0	0.0	0.0	1	1.0	INF
52	0.0	0.0	0.0	4	1.0	INF
53	0.0	0.0	0.0	1	1.0	INF
54	0.0	0.0	0.0	1	1.0	INF

Table B.4. Percent changes of total complex sphingolipid content, ATP consumption, DAG biosynthesis and ratio of transfer rate of palmitoyl-coenzyme A to non-sphingolipid processes for all deletion mutants with respect to wild type strain and list of PA/DAG, and DAG/Ceramide ratios. INF stands for infinity (cont.)

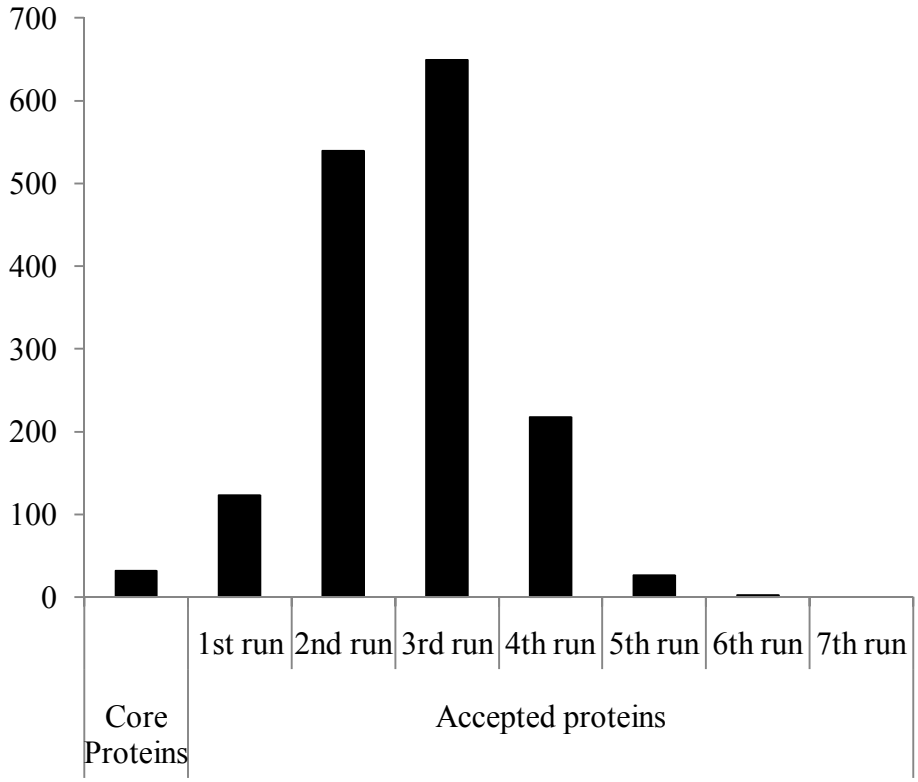
Mutants/ Blocked Reactions	Total Plasma Membrane Sphingolipid Content (% change wrt WT)	ATP Consumption (Reaction 79) (% change wrt WT)	DAG Biosynthesis (Reaction 121) (% change wrt WT)	Palmitoyl-CoA to non-sphingolipid processes (Reaction 2) (Ratio wrt WT)	PA/DAG	DAG/Ceramide
55	0.0	0.0	0.0	1	1.0	INF
56	0.0	0.0	0.0	1	1.0	INF
57	0.0	0.0	0.0	1	1.0	INF
58	0.0	0.0	0.0	1	1.0	INF
59	0.0	0.0	0.0	1	1.0	INF
60	-84.8	-35.5	-64.5	330632	0.4	INF
SUR2-28, 33	-84.9	-37.2	-61.3	333442	0.4	INF
LAC1/LAG1/LIP1-29, 30	-100.0	-38.8	-72.1	341820	0.5	INF
AUR1-35, 36, 40, 41, 42	-100.0	-34.1	-69.7	335497	0.4	5.4
SCS7-37, 39	-85.7	-36.8	-63.4	330417	0.4	INF
CCC2-38, 48	0.0	0.0	0.0	1	1.0	INF
CSG1/CSG2-43, 44, 45, 46, 47	3.3	1.6	-48.4	183836	1.0	INF
IPT1-49, 50, 51, 52, 53	-84.8	-35.5	-64.5	330632	0.4	INF
LCB4/LCB5-54, 55	0.0	0.0	0.0	1	1.0	INF
LCB3/LBP2-56, 57	0.0	0.0	0.0	1	1.0	INF
DPL1-58, 59	0.0	0.0	0.0	1	1.0	INF
SUR2/CSG2	-84.4	-36.0	-62.5	333875	0.4	INF
SCS7/CSG2	-85.2	-35.6	-64.7	330899	0.4	INF

Table B.4. Percent changes of total complex sphingolipid content, ATP consumption, DAG biosynthesis and ratio of transfer rate of palmitoyl-coenzyme A to non-sphingolipid processes for all deletion mutants with respect to wild type strain and list of PA/DAG, and DAG/Ceramide ratios. INF stands for infinity (cont.)

Mutants/ Blocked Reactions	**Total Plasma Membrane Sphingolipid Content (% change wrt WT)**	**ATP Consumption (Reaction 79) (% change wrt WT)**	**DAG Biosynthesis (Reaction 121) (% change wrt WT)**	**Palmitoyl-CoA to non-sphingolipid processes (Reaction 2) (Ratio wrt WT)**	**PA/DAG**	**DAG/Ceramide**
SCS7/SUR2	-85.3	-37.0	-61.8	334048	0.4	INF
WT-exogenous DHS and PHS	3.3	1.6	3.3	183836	1.0	INF
GDA1	-84.8	-35.5	-64.5	330632	0.4	INF
GDA1/SCS7	-85.2	-35.6	-64.7	330899	0.4	INF
INO1	0.0	0.0	0.0	1	1.0	INF
INO1 deletion no inositol transport	-100.0	-41.2	-69.2	344962	0.2	5.6
WT no inositol transport	0.0	0.0	0.0	1	1.0	INF
LCB1/LCB2	-100.0	-43.4	-57.4	371190	0.4	INF
LCB4/LCB5	0.0	0.0	0.0	1	1	INF
GUP1	0.0	0.0	0.0	1	1	INF
GAT1/GAT2	-100.0	-35.0	-100.0	531266	INF	0
CDS1	-100.0	-40.4	-66.4	349554	0	5.6
SER2	-100.0	-41.4	-53.4	376568	0.3	INF
ACB1	0.0	0.0	0.0	0	1.0	INF

APPENDIX C: SUPPLEMENTARY TABLES FOR CHAPTER 5

In this section, one supplementary figure and two supplementary tables mentioned in Chapter 5 (Reconstruction and Analyses of Sphingolipid Protein-Protein Interaction Networks) are provided. Table C.1 cannot be given here explicitly since its dimensions are not appropriate for printing. Instead, only few representative columns are given. The complete form of Table C.1 is given in CD, it is saved as AppendixC_TC1.xlsx.



Core Proteins	Accepted proteins						
	1st run	2nd run	3rd run	4th run	5th run	6th run	7th run
32	123	539	649	218	27	3	0
Total number of proteins at the end of each run							
0th run	**1st run**	**2nd run**	**3rd run**	**4th run**	**5th run**	**6th run**	**7th run**
32	155	694	1343	1561	1588	1591	1591

Figure C.1. Progress of network reconstruction using selective permissibility algorithm (SPA). The number of proteins that are added at the end of each run and the total number of proteins can be observed cumulatively in the data table

Table C.1. Conservation of hub proteins. Human homologs with sequence identity and coverage are tabulated. The domains of hubs and homologs (yeast and human) are given with the corresponding alignment regions. Domain sequence similarity results are also shown

#	Yeast hub protein	Description	Human homolog	E-value	Yeast protein domain	Domain name	Human domain
1	YHR135C	Casein kinase I isoform	Casein kinase I, gamma 3, isoform CRA_c	3.74E-97	PF00069	Protein kinase domain	PF00069 PF12605
2	YBR159W	Microsomal beta-keta reductase	Steroid dehydrogenase homolog hydroxysteroid (17-beta) dehydrogenase 12, isoform CRA_b	7.57E-39	PF00106	Short chain dehydrogenase	PF00106
3	YJL164C	cAMP-dependent protein kinase catalytic subunit	cAMP-dependent protein kinase catalytic subunit alpha isoform I	4.45E-96	PF00069	Protein kinase domain	PF00069
4	YJR059W	Putative serine/threonine protein kinase involved in regulation of ion transport across plasma membrane; enhances spermine uptake	Testis-specific serine/threonine-protein kinase 6	2.08E-18	PF00069	Protein kinase domain	PF00069
5	YNL154C	Casein kinase I isoform	Casein kinase 1, gamma 3, isoform CRA_c	9.85E-99	PF00069	Protein kinase domain	PF00069
6	YBR160W	Catalytic subunit of the main cell cycle cyclin-dependent kinase	Cyclin-dependent kinase 3, isoform CRA_a	1.31E-108	PF00069	Protein kinase domain	PF00069
7	YBR106W	Probable membrane protein, involved in phosphate transport	Ig heavy chain V region precursor	2.22E+00	PF10032	Phosphate transporter	-
8	YGL137W	Essential beta'-coat protein of the COPI coatomer, involved in ER-to-Golgi and Golgi-to-ER transport	Coatomer subunit beta	0.00E+00	PF04053	Coatomer WD associated region	PF04053
					PF00400	WD domain, G-beta repeat	PF00400
9	YNL272C	Guanyl-nucleotide exchange factor for the small G-protein Sec4p	rab-3A-interacting protein isoform alpha 1	5.85E-06	PF06428	GDP/GTP exchange factor Sec2p	PF06428
10	YGR060W	C-4 methyl sterol oxidase	C-4 methylsterol oxidase isoform 1	4.14E-48	PF04116	Fatty acid hydroxylase superfamily	PF04116

Table C.2. 13 gold proteins are involved in 6 different clusters. GO process term assignments with high and low confidence are given. The incorrect GO term assignments are shown with an asterisk (*). The superscript ^{M} stands for the results of module-assisted approach, the superscript ^{N} stands for the results of neighborhood counting approach, the superscript ^{C} stands for the combination of both approaches

ORF name	Cluster	Unknown density	High confidence^{C}			High confidence^{M}		Low confidence$^{M, N}$	
			GO process term	Frequency in neighboorhood	Frequency in module	GO process term	% overall support	Keyword	% repeats
YGL001C	3	7.7	Ergosterol biosynthetic process	60	66.6	Ergosterol biosynthetic process	100	Biosythetic process	25
			Alcohol metabolic process	50	58.3			Metabolic process	50
			Cellular lipid metabolic process	30	33.3				
			Cellular metabolic compound salvage*	35	50				
YPL058C	6	13.6	-	-	-	-	-	Transport	59.1
YLR144C	6	13.6	-	-	-	-	-	Actin	13.9
								Metabolic process	21.4
YBR177C	6	13.6	-	-	-	-	-	Metabolic process	21.4
								Biosythetic process	20
YPR173C	8	10	Late endosome to vacuole transport	33.3	44.4	-	-	Vacuole/vacuolar	27.3
			Establishment of protein localization	38	44.4			Vesicle/vesicular	27.3
			Ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway*	19	88.8			Transport	18.2
YDR264C	9	9.1	-	-	-	-	-	Vesicle/vesicular*	35.3
								Golgi	29.4
								Transport	35.3
								Fusion	23.5
YBR080C	14	13	-	-	-	Protein amino acid phosphorylation*	60	Transport	20
								Fusion	17.1
								Vesicle/vesicular	20
								Golgi	14.3

Table C.2. 13 gold proteins are involved in 6 different clusters. GO process term assignments with high and low confidence are given. The incorrect GO term assignments are shown with an asterisk (*). The superscript ^{M} stands for the results of module-assisted approach, the superscript ^{N} stands for the results of neighborhood counting approach, the superscript ^{C} stands for the combination of both approaches (cont.)

ORF name	Cluster	Unknown density	High confidence^{C}			High confidence^{M}		Low confidence$^{M, N}$	
			GO process term	Frequency in neighboorhood	Frequency in module	GO process term	% overall support	Keyword	% repeats
YDR477W	14	13	-	-	-	Protein amino acid phosphorylation	60	Regulation	25
								Metabolic process	17.9
YGR262C	14	13	-	-	-	Protein amino acid phosphorylation	60	-	-
YJL164C	14	13	-	-	-	Protein amino acid phosphorylation	60	-	-
YKL210W	14	13	-	-	-	Protein amino acid phosphorylation*	60	Ubiquitination	33.3
								Metabolic process	33.3
YLR096W	14	13	-	-	-	Protein amino acid phosphorylation	60	Regulation	50
YGR040W	20	12.5	-	-	-	Protein amino acid phosphorylation	100	Regulation	33.3
								Growth	18.2

APPENDIX D: MATLAB CODES UTILIZED IN CHAPTER 3 FOR THE SOLUTION OF DIFFERENTIAL EQUATIONS

The MATLAB codes used in Chapter 3 during dynamic modelling are given in the attached CD as .m files under the directory “AppendixD”. The code solves the stiff differential equations using the variable order method. The rate equations of the reactions are also given in the code.

APPENDIX E: MATLAB CODES UTILIZED IN CHAPTER 3 FOR THE ANALYSIS OF ELEMENTARY FLUX MODES

The MATLAB code used in Chapter 3 during metabolic pathway analysis is given in the attached CD as .m file under the directory “AppendixE”. The selection of the elementary flux modes that all the elements of the key set participate in and the common reactions occurring in these elementary flux modes is performed by the code.

APPENDIX F: MATLAB CODES UTILIZED IN CHAPTER 4 FOR STOICHIOMETRIC MODELLING

The MATLAB codes used in Chapter 4 during stoichiometric modelling are given in the attached CD as .m files under the directory “AppendixF”. The codes responsible for the reconstruction of stoichiometric matrix from the reaction list, flux balance analyses by linear and quadratic programming and flux variability analyses are given. The gene deletion mutant simulations and *in silico* drug action studies are all given in the codes.

APPENDIX G: MATLAB CODES UTILIZED IN CHAPTER 5 FOR THE RECONSTRUCTION OF PROTEIN-PROTEIN INTERACTION NETWORK

The MATLAB codes used in Chapter 5 during the reconstruction of protein-protein interaction network of sphingolipids by the selective permissibility algorithm are given in the attached CD as .m files under the directory "AppendixG". The codes that are used for the reconstruction of the annotation table, conversion of the Gene Ontology number IDs to the Gene Ontology terms, finding the unknown Gene Ontology terms and lipid and/or sphingolipid related Gene Ontology terms are given. The code that eliminates the candidate proteins that do not fit to the network with respect to the annotation table is also given.

REFERENCES

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, 1990, "Basic Local Alignment Search Tool", *J. Mol. Biol.*, Vol. 215, pp. 403-410.

Alvarez-Vasquez, F., K. J. Sims, L. A. Cowart, Y. Okamoto, E. O. Voit and Y. A. Hannun, 2005, "Simulation and validation of modelled sphingolipid metabolism in *Saccharomyces cerevisiae*", *Nature*, Vol. 433, pp. 425-430.

Alvarez-Vasquez, F., K. J. Sims, Y. A. Hannun and E. O. Voit, 2004, "Integration of kinetic information on yeast sphingolipid metabolism in dynamical pathway models", *Journal of Theoretical Biology*, Vol. 226, pp. 265-291.

Alzheimer's Association, www.alz.org, latest time of access: November 2008.

American Heart Association, www.americanheart.org, latest time of access: November 2008.

AmiGO: the Gene Ontology, http://amigo.geneontology.org/cgi-bin/amigo/go.cgi, latest time of access: December 2010.

Aranda, B., P. Achuthan, Y. Alam-Faruque, I. Armean, A. Bridge *et al.*, 2009, "The IntAct molecular interaction database in 2010", *Nucleic. Acids Res.*, Vol. 38 (suppl 1), pp. D525-D531.

Arga, K. Y., Z. I. Onsan, B. Kirdar, K. O. Ulgen and J. Nielsen, 2007, "Understanding Signaling in Yeast: Insights From Network Analysis", *Biotechnol. Bioeng.*, Vol. 97, pp. 1246-1258.

Ariga, T., W. D. Jarvis and R. K. Yu, 1998, "Role of sphingolipid-mediated cell death in neurodegenerative diseases", *JLR*, Vol. 39, pp. 1-16.

Assenov, Y., F. Ramirez, S.-E. Schelhorn, T. Lengauer and M. Albrecht, 2008, "Computing topological parameters of biological networks", *Bioinformatics*, Vol. 24, pp. 282-284.

Bader, G. D. and C. W. Hogue, 2003, "An automated method for finding molecular complexes in large protein interaction networks", *BMC Bioinformatics* Vol. 4, pp. 2-29.

Barabasi, A. L. and Z. N. Oltvai, 2004, "Network Biology: Understanding the cell's functional organization", *Nat. Rev. Genet.*, Vol. 5, pp. 101-113.

Barkai, N. and S. Leibler, 1997, "Robustness in simple biochemical networks", *Nature*, Vol. 387, pp. 913-917.

Beeler, T., D. Bacikova, K. Gable, L. Hopkins, C. Johnson, H. Slife and T. Dunn, 1998, "The *Saccharomyces cerevisiae* TSC10/YBR265w gene encoding 3-ketosphinganine reductase is identified in a screen for temperature-sensitive suppressors of the Ca^{2+}-sensitive *csg2Δ* mutant", *J. Biol. Chem.*, Vol. 273, pp. 30688-30694.

Bektas, M. and S. Spiegel, 2004, "Glycosphingolipids and cell death", *Glycoconjugate Journal*, Vol. 20, pp. 39-47.

Berger, S. I. and R. Iyengar, 2009, "Network analyses in systems pharmacology", *Bioinformatics*, Vol. 25, pp. 2466-2472.

BLAST: Basic Local Alignment Search Tool, http://blast.ncbi.nlm.nih.gov, latest time of access: October 2010.

Bonarius, H. P. J., V. Hatzimanikatis, K. P. H. Meesters, C. D. de Gooijer, G. Schmid and J. Tramper, 1996, "Metabolic flux analysis of hybridoma cells in different culture media using mass balances", *Biotechnol. Bioeng.*, Vol. 50, pp. 299-318.

Bonarius, H. P. J., G. Schmid and J. Tramper, 1997, "Flux analysis of underdetermined metabolic networks: the quest for the missing constraints", *Trends Biotechnol.*, Vol. 15, pp. 308-314.

Brady, R., 1967, "Enzymatic abnormalities in diseases of sphingolipid metabolism", *Clinical Chemistry*, Vol. 14, pp. 565-577.

BRENDA: Enzyme Database, http://www.brenda-enzymes.info/, latest time of access: December 2010.

Çakir, T., S. Alsan, H. Saybasili, A. Akin and K. O. Ulgen, 2007, "Reconstruction and flux analysis of coupling between metabolic pathways of astrocytes and neurons: Application to cerebral hypoxia", *Theor. Biol. Med. Model.*, Vol. 4, pp. 48-66.

Çakır, T., B. Kırdar, Z. I. Önsan, K. Ö. Ülgen and J. Nielsen, 2007, "Effect of carbon source perturbations on transcriptional regulation of metabolic fluxes in *Saccharomyces cerevisiae*", *BMC Systems Biology*, Vol. 1, pp. 18-28.

Çakır, T., C. S. Tacer and K. O. Ülgen, 2004, "Metabolic pathway analysis of enzyme-deficient human red blood cells", *BioSystems*, Vol. 78, pp. 49-67.

Cascante, M., L. G. Boros, B. Comin-Anduix, P. d. Atauri, J. J. Centelles and P. W.-N. Lee, 2002, "Metabolic control analysis in drug discovery and disease", *Nat. Biotechnol.*, Vol. 20, pp. 243-249.

Cavallini, L., R. Venerando, G. Miotto and A. Alexandre, 1999, "Ganglioside GM1 protection from apoptosis of rat heart fibroblasts", *Arch. Biochem. Biophys.*, Vol. 370, pp. 156-162.

Chen, Y. and D. Xu, 2005, "Genome-scale protein function prediction in yeast Saccharomyces cerevisiae through integrating multiple sources of high-throughput data", *Pac. Symp. Biocomput.*, Vol. 10, pp. 471-482.

Chiang, T., D. Scholtens, D. Sarkar, R. Gentleman and W. Huber, 2007, "Coverage and error models of protein-protein interaction data by directed graph analysis", *Genome Biol.*, Vol. 8, p. R186.

Chin, C.-S. and M. P. Samanta, 2003, "Global snapshot of a protein interaction network—a percolation based approach", *Bioinformatics*, Vol. 19, pp. 2413-2419.

Cohen, R. and S. Havlin, 2003, "Scale-Free Networks are Ultrasmall", *Phys. Rev. Lett.*, Vol. 90, p. 58701.

Cowart, L. A. and Y. A. Hannun, 2005, "Using genomic and lipidomic strategies to investigate sphingolipid function in the yeast heat-stress response", *Biochemical Society Transactions*, Vol. 33, pp. 1166-1169.

Cowart, L. A. and L. M. Obeid, 2007, "Yeast sphingolipids: Recent developments in understanding biosynthesis, regulation, and function", *Biochimica et Biophysica Acta*, Vol. 1771, pp. 421-431.

Cuvillier, O. and T. Levade, 2003, "Enzymes of sphingosine metabolism as potential pharmacological targets for therapeutic intervention in cancer", *Pharmacol. Res.*, Vol. 47, pp. 439-445.

Dickson, R. C., Sumanasekera, C., Lester, R.L., 2006, "Functions and metabolism of sphingolipids in *Saccharomyces cerevisiae*", *Prog. Lipid Res.*, Vol. 45, pp. 447-465.

Dickson, R. C., 2008, "New insights into sphingolipid metabolism and function in budding yeast", *J. Lipid Res.*, Vol. 49, pp. 909-921.

Dickson, R. C. and R. L. Lester, 1999, "Metabolism and selected functions of sphingolipids in the yeast *Saccharomyces cerevisiae*", *Biochimica et Biophysica Acta*, Vol. 1438, pp. 305-321.

Dickson, R. C. and R. L. Lester, 2002, "Sphingolipid functions in *Saccharomyces cerevisiae*", *Biochimica et Biophysica Acta*, Vol. 1583, pp. 13-25.

Dickson, R. C., E. E. Nagiec, G. B. Wells, M. M. Nagiec and R. L. Lester, 1997, "Synthesis of Mannose-(inositol-P)2-ceramide, the Major Sphingolipid in *Saccharomyces cerevisiae*, Requires the *IPT1* (YDR072c) Gene", *J. Biol. Chem.*, Vol. 272, pp. 29620-29625.

Dickson, R. C., C. Sumanasekera and R. L. Lester, 2006, "Functions and metabolism of sphingolipids in *Saccharomyces cerevisiae*", *Prog. Lipid. Res.* , Vol. 45, pp. 447-465.

Dominguez, A., M. V. Elorza, E. Santos, J. R. Villanueva and R. Sentandreu, 1978, "Inositol deficiency in yeast: Metabolic, enzymatic and autoradiographic studies", *Antonie Van Leeuwenhoek*, Vol. 44, pp. 341-352.

Dong, J. and S. Horvath, 2007, "Understanding network concepts in modules", *BMC Syst. Biol.*, Vol. 1, p. 24.

Duarte, N. C., Herrgard, M.J., Palsson, B.O., 2004, "Reconstruction and Validation of *Saccharomyces cerevisiae* iND750, a Fully Compartmentalized Genome-Scale Metabolic Model", *Genome Res.*, Vol. 14, pp. 1298-1309.

Duarte, N. C., M. J. Herrgård and B. Ø. Palsson, 2004, "Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model", *Genome Res.*, Vol. 14, pp. 1298-1309.

Dunn, T. M., D. Haak, E. Monaghan and T. J. Beeler, 1998, "Synthesis of monohydroxylated ınositolphosphorylceramide (IPC-C) in *Saccharomyces cerevisiae* requires Scs7p, a protein with both a cytochrome b5-like domain and a hydroxylase/desaturase domain", *Yeast*, Vol. 14, pp. 311-321.

Edwards, J. S. and B. O. Palsson, 2000a, "The *Escherichia coli* MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities", *Proc. Natl. Acad. Sci. U. S. A.*, Vol. 97, pp. 5528-5533.

Edwards, J. S. and B. O. Palsson, 2000b, "Metabolic flux balance analysis and the in silico analysis of *Escherichia coli* K-12 gene deletions", *BMC Bioinformatics*, Vol. 1, p. 1.

Eng, C. M., J. Fletcher, W. R. Wilcox, S. Waldek, C. R. Scott *et al*., 2007, "Fabry disease: Baseline medical characteristics of a cohort of 1765 males and females in the Fabry Registry", *J. Inherit. Metab. Dis.*, Vol. 30, pp. 184-192.

Expasy Proteomics Server, http://www.expasy.ch/, latest time of access: January 2009.

Fell, D. A., 1992, "Metabolic control analysis - a survey of its theoretical and experimental development", *Biochem. J.*, Vol. 286, pp. 313-330.

Fell, D. A. and J. R. Small, 1986, "Fat synthesis in adipose tissue. an examination of stoichiometric constraints", *Biochem. J.*, Vol. 238, pp. 781-786.

Finn, R. D., J. Mistry, J. Tate, P. Coggill, A. Heger *et al.*, 2010, "The Pfam protein families database", *Nucl. Acids Res.*, Vol. 38 (suppl 1), pp. D211-D222.

Forslund, K. and E. L. L. Sonnhammer, 2008, "Predicting protein function from domain content", *Bioinformatics*, Vol. 24, pp. 1681-1687.

Förster, J., I. Famili, P. Fu, B. O. Palsson and J. Nielsen, 2003, "Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network", *Genome Res.*, Vol. 13, pp. 244-253.

Förster, J., I. Famili, P. Fu, B. Ø. Palsson and J. Nielsen, 2003, "Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network", *Genome Res.*, Vol. 13, pp. 244-253.

Fracasso, P. M., M. F. Brady, D. H. Moore, J. L. Walker, P. G. Rose, L. Letvak, T. M. Grogan and W. P. McGuire, 2001, "Phase II study of paclitaxel and valspodar (PSC 833) in refractory ovarian carcinoma: A gynecologic oncology group study", *Journal of Clinical Oncology*, Vol. 19, pp. 2975-2982.

Funato, K., B. Vallee and H. Riezman, 2002, "Biosynthesis and trafficking of sphingolipids in the yeast *Saccharomyces cerevisiae*", *Biochemistry*, Vol. 41, pp. 15105-15114.

Futerman, A. H., 2006, "Intracellular trafficking of sphingolipids: Relationship to biosynthesis", *Biochim. Biophys. Acta*, Vol. 1758, pp. 1885-1892.

Futerman, A. H. and H. Riezman, 2005, "The ins and outs of sphingolipid synthesis", *TRENDS in Cell Biology*, Vol. 15, pp. 312-318.

GENECENSUS, http://bioinfo.mbb.yale.edu/genome, latest time of access: December 2010.

Gepasi: Biochemical Kinetics Simulator, http://www.gepasi.org/, latest time of access: August 2007.

Gomez-Munoz, A., 1998, "Modulation of cell signalling by ceramides", *Biochim. Biophys. Acta*, Vol. 1391, pp. 92-109.

Grindrod, P. and M. Kibble, 2004, "Review of uses of network and graph theory concepts within proteomics", *Expert Rev Proteomics*, Vol. 1, pp. 89-98.

Guilllas, I., Kirchman, P.A., Chuard, R., Pfefferli, M., Jiang, J.C., Jazwinski, S.M., Conzelmann, A., 2001, "C26-CoA-dependent ceramide synthesis of *Saccharomyces cerevisiae* is operated by Lag1p and Lac1p ", *EMBO J.*, Vol. 20, pp. 2655-2665.

Gurvit, H., M. Emre, S. Tinaz, B. Bilgic, H. Hanagasi, H. Sahin, E. Gurol, J. T. Kvaloy and H. Harmanci, 2008, "The prevalence of dementia in an urban Turkish population", *AJADD*, Vol. 23, pp. 67-76.

Haak, D., K. Gable, T. Beeler and T. Dunn, 1997, "Hydroxylation of *Saccharomyces cerevisiae* ceramides requires Sur2p and Scs7p", *J. Biol. Chem.*, Vol. 272, pp. 29704-29710.

Hait, N. C., C. A. Oskeritzian, S. W. Paugh, S. Milstien and S. Spiegel, 2006, "Sphingosine kinases, sphingosine 1-phosphate, apoptosis and diseases", *Biochimica et Biophysica Acta*, Vol. 1758, pp. 2016-2026.

Hanada, K., K. Kumagai, N. Tomishige and M. Kawano, 2007, "CERT and intracellular trafficking of ceramide", *Biochimica et Biophysica Acta*, Vol. 1771, pp. 644-653.

Hannun, Y. A., C. Luberto and K. M. Argraves, 2001, "Enzymes of sphingolipid metabolism: From modular to integrative signaling", *Biochemistry*, Vol. 40, pp. 4893-4903.

Hannun, Y. A. and L. M. Obeid, 2002, "The Ceramide-centric universe of lipid-mediated cell regulation: Stress encounters of the lipid kind", *JBC*, Vol. 277, pp. 25847-25850.

Hawkins, T., M. Chitale, S. Luban and D. Kihara, 2008, "PFP: Automated prediction of gene ontology functional annotations with confidence scores using protein sequence data", *Proteins*, Vol. 74, pp. 566-582.

Hawkins, T., S. Luban and D. Kihara, 2006, "Enhanced automated function prediction using distantly related sequences and contextual association by PFP", *Protein Sci.*, Vol. 15, pp. 1550-1556.

He, X., Y. Huang, B. Li, C.-X. Gong and E. H. Schuchman, 2010, "Deregulation of sphingolipid metabolism in Alzheimer's disease", *Neurobiology of Aging*, Vol. 31, pp. 398-408.

Hein, L. K., P. J. Meikle, J. J. Hopwood and M. Fuller, 2007, "Secondary sphingolipid accumulation in a macrophage model of Gaucher disease", *Molecular Genetics and Metabolism*, Vol. 92, pp. 336-345.

Heinrich, R. and T. A. Rapoport, 1974, "A linear steady-state treatment of enzymatic chains. Critique of the crossover theorem and a general procedure to identify interaction sites with an effector", *Eur. J. Biochem.*, Vol. 42, pp. 97-105.

Heinrich, R. and T.A.Rapoport, 1974, "A linear steady-state treatment of enzymatic chains", *Eur. J. Biochem.*, Vol. 42, pp. 89-95.

Hirabayashi, Y., Y. Igarashi and A. H. M. Jr., 2006, *Sphinolipid Biology*, Springer, Verlag-Tokyo.

Holzhutter, H. G., 2004, "The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks", *Eur. J. Biochem.*, Vol. 271, pp. 2905-2922.

Hunter, S., R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman *et al.*, 2009, "InterPro: the integrative protein signature database", *Nucleic Acids Res.*, Vol. 37 (Database Issue), pp. D211–D215.

Hwang, O., G. Kim, Y. J. Jang, S. W. Kim, G. Choi, H. J. Choi, S. Y. Jeon, D. G. Lee and J. D. Lee, 2001, "Synthetic phytoceramides induce apoptosis with higher potency than ceramides", *Mol. Pharmacol.*, Vol. 59, pp. 1249-1255.

Hwang, W.-C., A. Zhang and M. Ramanathan, 2008, "Identification of Information Flow-Modulating Drug Targets: A Novel Bridging Paradigm for Drug Discovery", *Clin. Pharmacol. Ther.*, Vol. 84, pp. 563-572.

Jeong, H., B. Tombor, R. Albert, Z. N. Oltvai and A. L. Barabasi, 2000, "the large-scale organization of metabolic networks", *Nature*, Vol. 407, pp. 651-654.

Jiang, Q., J. Wong, H. Fyrst, J. D. Saba and B. N. Ames, 2004, "γ-Tocopherol or combinations of vitamin E forms induce cell death in human prostate cancer cells by interrupting sphingolipid synthesis", *PNAS*, Vol. 101, pp. 17825-17830.

Joshi, T., Y. Chen, J. M. Becker, N. Alexandrov and D. Xu, 2004, "Genome-Scale Gene Function Prediction Using Multiple Sources of High-Throughput Data in Yeast *Saccharomyces cerevisiae*", *OMICS*, Vol. 8, pp. 322-333.

Ju, T.-C., S.-D. Chen, C.-C. Liu and D.-I. Yang, 2005, "Protective effects of S-nitrosoglutathione against amyloid β-peptide neurotoxicity", *Free Radical Biology & Medicine*, Vol. 38, pp. 938-949.

Jung, J., G. Yi, S. A. Sukno and M. R. Thon, 2010, "PoGO: Prediction of Gene Ontology terms for fungal proteins", *BMC Bioinformatics*, Vol. 11, p. 215.

Kacser, H. and J. A. Burns, 1973, "Control of enzyme flux", *Symp. Soc. Exp. Biol.*, Vol. 27, pp. 65-104.

Kauffman, K. J., Prakash, P., Edwards, J.S., 2003, "Advances in flux balance analysis", *Curr. Opin. Biotechnol.*, Vol. 14, pp. 491-496.

Kavun, F. B. and K. Ö. Ülgen, 2005, "Metabolic control analysis and metabolic pathway analysis of sphingolipid metabolism for anticancer drug target identification", In: ECB12, *12th European Congress on Biotechnology,* Copenhagen, Danmark 21-24 August 2005.

Kavun, F. B. and K. Ö. Ülgen, 2007, "Drug target identification in sphingolipid metabolism by computational systems biology tools: MCA and MPA", In: *6th International Meeting of the Sphingolipid Club*, Bilbao, Spain, 22-25 November 2007.

Kavun-Ozbayraktar, F. B. and K. O. Ulgen, 2009, "Molecular facets of sphingolipids: Mediators of diseases", *Biotechnol. J.*, Vol. 4, pp. 1028-1041.

Kavun-Ozbayraktar, F. B. and K. O. Ulgen, 2010, "Drug target identification in sphingolipid metabolism by computational systems biology tools: Metabolic control analysis and metabolic pathway analysis", *J. Biomed. Inform.*, Vol. 43, pp. 537-549.

Kavun-Ozbayraktar, F. B. and K. O. Ulgen, 2011, "Stoichiometric network reconstruction and analysis of yeast sphingolipid metabolism incorporating different states of hydroxylation", *Biosystems*, in press.

KEGG: Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/, latest time of access: December 2010.

Kim, S., H. Fyrst and J. Saba, 2000, "Accumulation of Phosphorylated Sphingoid Long Chain Bases Results in Cell Growth Inhibition in Saccharomyces cerevisiae", *Genetics*, Vol. 156, pp. 1519-1529.

Kitagaki, H., L. A. Cowart, N. Matmati, S. V. d. Avalos, S. A. Novgorodov, Y. H. Zeidan, J. Bielawski, L. M. Obeid and Y. A. Hannun, 2007, "*Isc1* regulates sphingolipid metabolism in yeast mitochondria", *Biochim. Biophys. Acta*, Vol. 1768, pp. 2849-2861.

Klamt, S., 2006, "Generalized concept of minimal cut sets in biochemical networks", *BioSystems*, Vol. 83, pp. 233-247.

Klamt, S., J. Gagneur and A. v. Kamp, 2005, "Algorithmic approaches for computing elementary modes in large biochemical reaction networks", *IEE Proc.-Syst. Biol.*, Vol. 152, pp. 249-255.

Klamt, S., J. Saez-Rodriguez and E. D. Gilles, 2007, "Structural and functional analysis of cellular networks with CellNetAnalyzer", *BMC Systems Biology*, Vol. 1, pp. 2-15.

Klamt, S., J. Stelling, M. Ginkel and E. D. Gilles, 2003, "FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps", *Bioinformatics*, Vol. 19, pp. 261-269.

Kohlwein, S. D., S. Eder, C. S. Oh, C. E. Martin, K. Gable, D. Bacikova and T. Dunn, 2001, "Tsc13p is required for fatty acid elongation and localizes to a novel structure at the nuclear-vacuolar ınterface in *Saccharomyces cerevisiae*", *Mol. Cell. Biol.* , Vol. 21, pp. 109-125.

Kolter, T., R. L. Proia and K. Sandhoff, 2002, "Combinatorial Ganglioside Biosynthesis", *The Journal of Biological Chemistry*, Vol. 277, pp. 25859-25862.

Kolter, T. and K. Sandhoff, 2006, "Sphingolipid metabolism diseases", *Biochimica et Biophysica Acta*, Vol. 1758, pp. 2057-2079.

Kraveka, J. M., L. Li, Z. M. Szulc, J. Bielawski, B. Ogretmen, Y. A. Hannun, L. M. Obeid and A. Bielawska, 2007, "Involvement of dihydroceramide desaturase in cell cycle progression in human neuroblastoma cells", *The Journal of Biological Chemistry*, Vol. 282, pp. 16718-16728.

Kuepfer, L., U. Sauer and L. M. Blank, 2005, "Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*", *Genome Res.*, Vol. 15, pp. 1421-1430.

Lee, I., R. Narayanaswamy and E. M. Marcotte, 2007, "Bioinformatic Prediction of Yeast Gene Function", *Meth. Microbiol.*, Vol. 36, pp. 597-628.

Lee, J.-T., J. Xu, J.-M. Lee, G. Ku, X. Han, D.-I. Yang, S. Chen and C. Y. Hsu, 2004, "Amyloid-β peptide induces oligodendrocyte death by activating the neutral sphingomyelinase-ceramide pathway", *JCB*, Vol. 164, pp. 123-131.

Levade, T., N. Augé, R. J. Veldman, O. Cuvillier, A. Nègre-Salvayre and R. Salvayre, 2001, "Sphingolipid mediators in cardiovascular cell biology and pathology", *Circ. Res.*, Vol. 89, pp. 957-968.

Levade, T., S. Malagarie-Cazenave, V. Gouazé, B. Ségui, C. Tardy, S. Betito, N. Andrieu-Abadie and O. Cuvillier, 2002, "Ceramide in apoptosis: A revisited role", *Neurochemical Research*, Vol. 27, pp. 601-607.

Li, D., J. Li, S. Ouyang, J. Wang, S. Wu, P. Wan, Y. Zhu, X. Xu and F. He, 2006, "Protein interaction networks of *Saccharomyces cerevisiae, Caenorhabditis elegans* and *Drosophila melanogaster* : Large-scale organization and robustness", *Proteomics*, Vol. 6, pp. 456-461.

Lin, C.-Y., C.-H. Chin, H.-H. Wu, S.-H. Chen, C.-W. Ho and M.-T. Ko, 2008, "Hubba: hub objects analyzer-a framework of interactome hubs identification for network biology", *Nucleic Acids Res.*, Vol. 36 (suppl 2), pp. W438-W443.

Louie, B., R. Higdon and E. Kolker, 2009, "A Statistical Model of Protein Sequence Similarity and Function Similarity Reveals Overly-Specific Function Predictions", *PLoS One*, Vol. 4, p. e7546.

Lukacs, Z., R. Hartung, M. Beck, A. Keil and E. Mengel, 2007, "Direct comparison of enzyme measurements from dried blood and leukocytes from male and female Fabry disease patients", *J. Inherit. Metab. Dis.*, Vol. 30, p. 614.

Maccioni, H. J. F., J. L. Daniotti and J. A. Martina, 1999, "Organization of ganglioside synthesis in the Golgi apparatus", *Biochimica et Biophysica Acta*, Vol. 1437, pp. 101-118.

Maceyka, M., S. Milstien and S. Spiegel, 2005, "Sphingosine kinases, sphingosine-1-phosphate and sphingolipidomics", *Prostaglandins & other Lipid Mediators*, Vol. 77, pp. 15-22.

Maere, S., K. Heymans and M. Kuiper, 2005, "BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks", *Bioinformatics*, Vol. 21, pp. 3448-3449.

Mahadevan, R. and C. H. Schilling, 2003, "The effects of alternate optimal solutions in constraint-based genome-scale metabolic models", *Metab. Eng.*, Vol. 5, pp. 264-276.

Mandala, S. M., 2001, "Sphingosine-1-phosphate phosphatases", *Prostaglandins Other Lipid Mediat.*, Vol. 64, pp. 143-156.

Mani, R., R. P. St.Onge, J. L. Hartman, G. Giaever and F. P. Roth, 2008, "Defining genetic interaction", *PNAS*, Vol. 105, pp. 3461-3466.

Mao, C., M. Wadleigh, G. M. Jenkins, Y. A. Hannun and L. M. Obeid, 1997, "Identification and Characterization of *Saccharomyces cerevisiae* Dihydrosphingosine-1-phosphate Phosphatase", *J. Biol. Chem.*, Vol. 272, pp. 28690-28694.

Marks, D. L. and R. E. Pagano, 2002, "Endocytosis and sorting of glycosphingolipids in sphingolipid storage disease", *TRENDS in Cell Biology*, Vol. 12, pp. 605-613.

Massjouni, N., C. G. Rivera and T. M. Murali, 2006, "VIRGO: computational prediction of gene functions", *Nucleic Acids Res*, Vol. 34 (suppl 2), pp. W340-W344.

Mathias, S., L. A. Penna and R. N. Kolesnick, 1998, "Signal transduction of stress via ceramide", *Biochem. J.*, Vol. 335, pp. 465-480.

Meer, G. v. and J. C. M. Holthuis, 2000, "Sphingolipid transport in eukaryotic cells", *Biochimica et Biophysica Acta*, Vol. 1486, pp. 145-170.

MIPS: Institute of Bioinformatics and Systems Biology, http://www.helmholtz-muenchen.de/en/ibis, latest time of access: December 2010.

Modrak, D. E., D. V. Gold and D. M. Goldenberg, 2006, "Sphingolipid targets in cancer therapy", *Mol. Cancer Ther.*, Vol. 5, pp. 200-208.

Momoi, M., D. Tanoue, Y. Sun, H. Takematsu, Y. Suzuki, M. Suzuki, A. Suzuki, T. Fujita and Y. Kozutsumi, 2004, "*SLI1* (YGR212W) is a major gene conferring resistance to the sphingolipid biosynthesis inhibitor ISP-1, and encodes an ISP-1 N-acetyltransferase in yeast", *Biochem J*, Vol. 381, pp. 321-328.

Newman, M. E. J., 2003, "A measure of betweenness centrality based on random walks", *Soc. Networks* Vol. 27, pp. 39-54.

Obeid, L. M., Y. Okamoto and C. Mao, 2002, "Yeast sphingolipids: Metabolism and biology", *Biochimica et Biophysica Acta*, Vol. 1585, pp. 163-171.

Ogretmen, B., 2006, "Sphingolipids in cancer: Regulation of pathogenesis and therapy", *FEBS Letters*, Vol. 580, pp. 5467-5476.

Ogretmen, B. and Y. A.Hannun, 2004, "Biologically active sphingolipids in cancer pathogenesis and treatment", *Nature Reviews Cancer*, Vol. 4, pp. 604-616.

Ogretmen, B. and Y. A. Hannun, 2002, "Updates on functions of ceramide in chemotherapy-induced cell death and in multidrug resistance", *Drug Resist. Update*, Vol. 4, pp. 368-377.

Ohanian, J. and V. Ohanian, 2001a, "Sphingolipids in mammalian cell signalling", *Cell. Mol. Life Sci.*, Vol. 58, pp. 2053-2068.

Ohanian, J. and V. Ohanian, 2001b, "Lipid second messenger regulation: The role of diacylglycerol kinases and their relevance to hypertension", *J. Hum. Hypertens.*, Vol. 15, pp. 93-98.

Okada, T., T. Kajimoto, S. Jahangeer and S.-I. Nakamura, 2009, "Sphingosine kinase/sphingosine 1-phosphate signalling in central nervous system", *Cellular Signalling*, Vol. 21, pp. 7-13.

Olivera, A. and S. Spiegel, 2001, "Sphingosine kinase: A mediator of vital cellular functions", *Prostaglandins & other Lipid Mediators*, Vol. 64, pp. 123-134.

Ory, D. S., 2000, "Niemann-Pick type C: A disorder of cellular cholesterol trafficking", *Biochimica et Biophysica Acta*, Vol. 1529, pp. 331-339.

Ozkara, H. A. and M. Topcu, 2004, "Sphingolipidoses in Turkey", *Brain & Development*, Vol. 26, pp. 363-366.

Pagano, R. E., 2003, "Endocytic trafficking of glycosphingolipids in sphingolipid storage diseases", *Phil. Trans. R. Soc. Lond. B*, Vol. 358, pp. 885-891.

Pagano, R. E., V. Puri, M. Dominguez and D. L. Marks, 2000, "Membrane traffic in sphingolipid storage diseases", *Traffic*, Vol. 1, pp. 807-815.

Perry, R. J. and N. D. Ridgway, 2005, "Molecular mechanisms and regulation of ceramide transport", *Biochimica et Biophysica Acta*, Vol. 1734, pp. 220-234.

Petranovic, D. and J. Nielsen, 2008, "Can yeast systems biology contribute to the understanding of human disease?" *Trends in Biotechnology*, Vol. 26, pp. 584-590.

Pitre, S., F. Dehne, A. Chan, J. Cheetham, A. Duong *et al.*, "PIPE: a protein-protein interaction prediction engine based on the re-occurring short polypeptide sequences between known interacting protein pairs", *BMC Bioinformatics*, Vol. 7, pp. 365.

Princeton University, Laboratory of Bioinformatics and Functional Genomics, http://function.princeton.edu/genesite/genesite, latest time of access: December 2010.

Radin, N. S., 2001, "Apoptotic death by ceramide: Will the real killer please stand up?" *Medical Hypotheses*, Vol. 57, pp. 96-100.

Radin, N. S., 2001, "Killing cancer cells by poly-drug elevation of ceramide levels, A hypothesis whose time has come?" *Eur. J. Biochem.*, Vol. 268, pp. 193-204.

Radin, N. S., 2003, "Killing tumours by ceramide-induced apoptosis : a critique of available drugs", *Biochem. J.*, Vol. 371, pp. 243-256.

Raghavachari, B., A. Tasneem, T. M. Przytycka and R. Jothi, 2008, "DOMINE: a database of protein domain interactions ", *Nucl. Acids Res.*, Vol. 36 (suppl 1), pp. D656-D661.

Ramakrishna, R., J. S. Edwards, A. Mcculloch and B. O. Palsson, 2001, "Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints", *Am. J. Physiol. Regulatory Integrative Comp. Physiol.*, Vol. 280, pp. 695-704.

Ravasz, E., A. L. Somera, D. A. Mongru, Z. N. Oltvai and A.-L. Barabasi, 2002, "Hierarchical Organization of Modularity in Metabolic Networks", *Science*, Vol. 297, pp. 1551-1555.

Reggiori, F. and A. Conzelmann, 1998, "Biosynthesis of inositol phosphoceramides and remodeling of glycosylphosphatidylinositol anchors in *Saccharomyces cerevisiae* are mediated by different enzymes", *J. Biol. Chem.*, Vol. 273, pp. 30550-30559.

Reiss, U., B. Oskouian, J. Zhou, V. Gupta, P. Sooriyakumaran, S. Kelly, E. Wang, A. H. Merrill and J. D. Saba, 2004, "Sphingosine-phosphate Lyase Enhances Stress-induced Ceramide Generation and Apoptosis", *J. Biol. Chem.*, Vol. 279, pp. 1281-1290.

Reynolds, C. P., B. J. Maurera and R. N. Kolesnick, 2004, "Ceramide synthesis and metabolism as a target for cancer therapy", *Cancer Letters*, Vol. 206, pp. 169-180.

Riezman, H., 2006, "Organization and functions of sphingolipid biosynthesis in yeast", *Biochem. Soc. Trans.*, Vol. 34, pp. 367-369.

Roemer, T., S. Delaney and H. Bussey, 1993, "*SKN1* and *KRE6* Define a Pair of Functional Homologs Encoding Putative Membrane Proteins Involved in Beta-Glucan Synthesis", *Mol. Cell Biol.*, Vol. 13, pp. 4039-4048.

Saba, J. D. and T. Hla, 2004, "Point-counterpoint of sphingosine 1-phosphate metabolism", *Circ. Res.*, Vol. 94, pp. 724-734.

Saba, J. D., L. M. Obeid and Y. A. Hannun, 1996, "Ceramide: An intracellular mediator of apopotosis and growth supression", *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, Vol. 351, pp. 233-241.

Saccharomyces Genome Deletion Project, http://www-sequence.stanford.edu/group/yeast_deletion_project/deletions3, latest time of access: October 2010.

Samsel, L., G. Zaidel, H. M. Drumgoole, D. Jelovac, C. Drachenberg, J. G. Rhee, A. M. Brodie, A. Bielawska and M. J. Smyth, 2004, "The ceramide analog, B13, induces apoptosis in prostate cancer cell lines and inhibits tumor growth in prostate cancer xenografts", *Prostate*, Vol. 58, pp. 382-393.

Sandhoff, K., 1977, "The Biochemistry of sphingolipid storage diseases", *Anqew. Chem. Int. Ed. Engl.*, Vol. 16, pp. 273-285.

Satoi, H., H. Tomimoto, R. Ohtani, T. Kitano, T. Kondo *et al.*, 2005, "Astroglial expression of ceramide in alzheimer's disease brains: A role during neuronal apoptosis", *Neuroscience*, Vol. 130, pp. 657-666.

Sawai, H., Y. Okamoto, C. Luberto, C. Maoi, A. Bielawska, N. Domae and Y. A. Hannun, 2000, "Identification of *ISC1* (YER019w) as Inositol Phosphosphingolipid Phospholipase C in *Saccharomyces cerevisiae*", *J. Biol. Chem.*, Vol. 275, pp. 39793-39798.

Scarlatti, F., C. Bauvy, A. Ventruti, G. Sala, F. Cluzeaud, A. Vandewalle, R. Ghidoni and P. Codogno, 2004, "Ceramide-mediated Macroautophagy Involves Inhibition of Protein Kinase B and Up-regulation of Beclin 1", *J. Biol. Chem.*, Vol. 279, pp. 18384-18391.

Schneiter, R., 1999, "Brave little yeast, please guide us to Thebes: sphingolipid function in *S. cerevisiae*", *BioEssays*, Vol. 21, pp. 1004-1010.

Schuster, S., T. Dandekar and D. A. Fell, 1999, "Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering", *TIBTECH*, Vol. 17, pp. 53-60.

Schwarz, R., P. Musch, A. v. Kamp, B. Engels, H. Schirmer, S. Schuster and T. Dandekar, 2005, "YANA – a software tool for analyzing flux modes, gene-expression and enzyme activities", *BMC Bioinformatics*, Vol. 6, p. 135.

Seelan, R. S., C. Qian, A. Yokomizo, D. G. Bostwick, D. I. Smith and W. Liu, 2000, "Human acid ceramidase is overexpressed but not mutated in prostate cancer", *Genes, Chromosomes & Cancer*, Vol. 29, pp. 137-146.

Ségui, B., N. Andrieu-Abadie, J.-P. Jaffrézou, H. Benoist and T. Levade, 2006, "Sphingolipids as modulators of cancer cell death:Potential therapeutic targets", *Biochimica et Biophysica Acta*, Vol. 1758, pp. 2104-2120.

SGD: Saccharomyces Genome Database, http://www.yeastgenome.org/, latest time of access: December 2010.

Shaldubina, A., S. Ju, D. L. Vaden, D. Ding, R. H. Belmaker and M. L. Greenberg, 2002, "*Epi*-inositol regulates expression of the yeast *INO1* gene encoding inositol-1-P synthase", *Mol. Psychiatry*, Vol. 7, pp. 174-180.

Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, 2003, "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks", *Genome Res.*, Vol. 13, pp. 2498-2504.

Sharan, R., I. Ulitsky and R. Shamir, 2007, "Network-based prediction of protein function", *Mol. Syst. Biol.*, Vol. 3, p. 88.

Sims, K. J., S. D. Spassieva, E. O. Voit and L. M. Obeid, 2004, "Yeast sphingolipid metabolism: clues and connections", *Biochem. Cell Biol.*, Vol. 82, pp. 45-61.

Singh, R., J. Xu and B. Berger, 2006, "Struct2Net: Integrating Structure into Protein-Protein Interaction Prediction", *Pac. Symp. Biocomput.*, Vol. 11, pp. 403-414.

Soldati, T., C. Rancan, H. Geissler and S. R. Pfeffer, 1995, "Rab7 and Rab9 are recruited onto late endosomes by biochemically distinguishable processes", *JBC*, Vol. 270, pp. 25541-25548.

Somsen, O. J. G., M. A. Hoeben, E. Esgalhado, J. L. Snoep, D. Visser, R. T. J. M. Van Der Heijden, J. J. Heijnen and H. V. Westerhoff, 2000, "Glucose and the ATP paradox in yeast", *Biochem. J.*, Vol. 352, pp. 593-599.

Spiegel, S. and S. Milstien, 2002, "Sphingosine 1-Phosphate, a key cell signaling molecule", *J. Biol. Chem.*, Vol. 277, pp. 25851-25854.

Stark, C., B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz and M. Tyers, 2006, "BioGRID: a general repository for interaction datasets", *Nucleic Acids Res.*, Vol. 34 (Database Issue), pp. D535-D539.

Stephanopoulos, G., A. A. Aristidou and J. Nielsen, 1998, *Metabolic engineering: Principles and methodologies*, Academic Press, San Diego.

Stunff, H., I. Galve-Roperh, C. Peterson, S. Milstien and S. Spiegel, 2002, "Sphingosine-1-phosphate phosphohydrolase in regulation of sphingolipid metabolism and apoptosis", *J. Cell Biol.*, Vol. 158, pp. 1039-1049.

Stunff, H. L., C. Peterson, H. Liu, S. Milstien and S. Spiegel, 2002, "Sphingosine-1-phosphate and lipid phosphohydrolases", *Biochimica et Biophysica Acta*, Vol. 1582, pp. 8-17.

Taha, T. A., T. D. Mullen and L. M. Obeid, 2006, "A house divided: ceramide, sphingosine, and sphingosine-1-phosphate in programmed cell death", *Biochim. Biophys. Acta*, Vol. 1758, pp. 2027-2036.

Tardy, C., N. A. Abadie, R. Salvayr and T. Levade, 2004, "Lysosomal storage diseases: Is impaired apoptosis a pathogenic mechanism?" *Neurochemical Research*, Vol. 29, pp. 871-880.

Thevissena, K., J. Idkowiak-Baldys, Y.-J. Im, J. Takemoto, I. E. J. A. François *et al.*, 2005, "*SKN1*, a novel plant defensin-sensitivity gene in Saccharomyces cerevisiae, is implicated in sphingolipid biosynthesis", *FEBS Lett.*, Vol. 579, pp. 1973-1977.

Tiwari, M., A. Kumar, R. A. Sinha, A. Shrivastava, A. K. Balapure, R. Sharma, V. K. Bajpai, K. Mitra, S. Babu and M. M. Godbole, 2006, "Mechanism of 4-HPR-induced apoptosis in glioma cells: Evidences suggesting role of mitochondrial-mediated pathway and endoplasmic reticulum stress", *Carcinogenesis*, Vol. 27, pp. 2047-2058.

Trawick, J. D. and C. H. Schilling, 2006, "Use of constraint-based modeling for the prediction and validation of antimicrobial targets", *Biochem. Pharmacol.*, Vol. 71, pp. 1026-1035.

Tserng, K.-Y. and R. L. Griffin, 2004, "Ceramide metabolite, not intact ceramide molecule, may be responsible for cellular toxicity", *Biochem. J.*, Vol. 380, pp. 715-722.

UniProt: Universal protein resource, http://www.uniprot.org, latest time of access: December 2010.

Vallee, B. and H. Riezman, 2005, "Lip1p: a novel subunit of acyl-CoA ceramide synthase", *The EMBO Journal*, Vol. 24, pp. 730-741.

Varma, A. and B. O. Palsson, 1994, "Metabolic flux balancing: Basic concepts, scientific and pratical use", *Biotechnology (N.Y.)*, Vol. 12, pp. 994-998.

Verspoor, K., J. Cohn, S. Mniszewski and C. Joslyn, 2006, "A categorization approach to automated ontological function annotation", *Protein Sci.*, Vol. 15, pp. 1544-1549.

Vo, T. D., H. J. Greenberg and B. O. Palsson, 2004, "Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data", *J. Biol. Chem.*, Vol. 279, pp. 39532-39540.

Walhout, A. J., R. Sordella, X. Lu, J. L. Hartley, G. F. Temple, M. A. Brasch, N. Thierry-Mieg and M. Vidal, 2000, "Protein interaction mapping in C. elegans using proteins involved in vulval development", *Science*, Vol. 287, pp. 116-122.

Watts, D. J. and S. H. Strogatz, 1998, "Collective dynamics of 'small-world' networks", *Nature*, Vol. 393, pp. 440-442.

Wu, X.-R., Y. Zhu and Y. Li, 2005, "Analyzing Protein Interaction Networks via Random Graph Model", *Int. J. Inf. Technol.*, Vol. 11, pp. 125-132.

Xiao, A. Y., X. Q. Wang, A. Yang and S. P. Yu, 2002, "S light impairment of Na^+,K^+ -ATPase synergistically aggravates ceramide- and β-amyloid-induced apoptosis in cortical neurons", *Brain Research*, Vol. 955, pp. 253-259.

Xu, T., L. Du and Y. Zhou, 2008, "Evaluation of GO-based functional similarity measures using *S. cerevisiae* protein interaction and expression profile data", *BMC Bioinformatics*, Vol. 9, p. 472.

Yook, S.-H., Z. N. Oltvai and A.-L. Barabási, 2004, "Functional and topological characterization of protein interaction networks", *Proteomics*, Vol. 4, pp. 928-942.

Yoon, J., A. Blumer and K. Lee, 2006, "An algorithm for modularity analysis of directed and weighted biological networks based on edge-betweenness centrality", *Bioinformatics*, Vol. 22, pp. 3106-3108.

YRC:Yeast Resource Center, Informatics Platform, Public Data Repository, http://www.yeastrc.org/pdr, latest time of access: September 2010.

Yu, H., P. M. Kim, E. Sprecher, V. Trifonov and M. Gerstein, 2007, "The Importance of Bottlenecks in Protein Networks: Correlation with Gene Essentiality and Expression Dynamics", *PLoS Comput. Biol.*, Vol. 3, pp. 713-720.

Yu, R. K., E. Bieberich, T. Xia and G. Zeng, 2004, "Regulation of ganglioside biosynthesis in the nervous system", *J. Lipid Res.*, Vol. 45, pp. 783-793.

Zeidan, Y. H. and Y. A. Hannun, 2007, "Translational aspects of sphingolipid metabolism", *Trends Mol. Med.*, Vol. 13, pp. 327-336.

Zhu, X., M. Gerstein and M. Snyder, 2007, "Getting connected: analysis and principles of biological networks", *Genes Dev.* Vol. 21, pp. 1010-1024.