RECONSTRUCTION OF INSULIN SIGNALING NETWORK IN *H. SAPIENS*

by

Pelin Ümit

B.S., Chemical Engineering, Boğaziçi University, 2007

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

Graduate Program in Chemical Engineering
Boğaziçi University
2009

dedicated to my beloved father

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and grateful thanks to my thesis advisor Prof. Kutlu Ö. Ülgen for her everlasting encouragement and support during my graduate studies. I also would like to thank my thesis jury; Prof. Hale Saybaşılı and Asst. Prof. Elif Özkırımlı Ölmez for their critical comments on my thesis.

I am sincerely grateful to my friends in Biosystems Engineering Research Group, especially Betül Kavun, Esra Börklü Yücel, Elif Dereli, Ceyda Kasavi, Yasemen Güngörmez Duygu Dikicioğlu, Ayça Cankorur, Dicle Hasdemir, Güray Kuzu and K. Yalçın Arga for their everlasting patience and support. I must also acknowledge Saliha Durmuş Tekir for her suggestions on the framework of the project and devoting her leisure time to seeking answers to my never–ending questions.

I would also like to thank my mother, Hatice Ümit, my father, Muharrem Ümit, my brother Mehmet Ümit and my aunt Zübeyde Pehlivan for their never ending love, support and encouragement through all my life. I am greatly indebted to these people for their support.

Last but certainly not the least I would like to thank with all my heart to my love, my fiancée, Erhan Abik for his everlasting love, continuous patience and support which makes everything bearable and beautiful.

I also acknowledge the financial supports provided by BU-BAP 09A501P and TÜBİTAK 104M362 and 108M649 projects.

ABSTRACT

RECONSTRUCTION OF INSULIN SIGNALING NETWORK IN *H. SAPIENS*

Insulin is an animal hormone that regulates the glucose levels in blood. It is medically used to treat diabetic patients. Diabetes is characterized by the failure of synthesizing, secreting or responding to insulin and it is the third most prevalent disease in the world. In order to understand the reasons underlying Type-2 diabetes, the insulin signaling network should be analyzed in detail as signaling through the insulin pathway is critical for the regulation of intracellular and blood glucose levels and for the avoidance of diabetes. Moreover, the topological structure of the insulin signaling pathway is still incomplete. Thus in the present study, a protein interaction network consisting of proteins that have statistically high probability of being biologically related to insulin signaling pathway was reconstructed using SPA algorithm and Gene Ontology (GO) annotations. The 27 insulin signaling annotated proteins in humans were used as the input and a candidate network of 416 proteins and 895 interactions was obtained. The graph theoretic analysis of this reconstructed insulin signaling network whose degree distribution approximates the power law model, yielded the common topological features of the biological networks, i.e. it has small world characteristics with scale free topology. An important response of the muscle cell to insulin stimulation is known to be the translocation of intracellular vesicles containing a glucose transporter (GLUT4) to the cell surface. Therefore, beside the reconstructed insulin signaling network, the metabolic insulin network was also investigated by dynamic modeling approaches. The controlling reactions of the metabolic model were determined as the phosphorylation of PI3-kinase (a lipid kinase) and AKT (protein kinase B) proteins. In agreement with the literature reports, the present study also indicates that the insulin receptor, PKC, PI3K and Akt have important roles in regulating insulin signaling. In addition to these molecules, GRB2 protein is determined to be a hub with high connectivity and must therefore be assessed with special care in drug targeting.

ÖZET

H. SAPİENS'DE İNSÜLİN SİNYALLEŞMESİ AĞ YAPISININ OLUŞTURULMASI

Insülin, kandaki glikoz seviyesini düzenleyen hayvansal bir hormondur. Tıbben diyabet hastalığının tedavisinde kullanılır. Diyabet, insülin sentezleme, salgılama veya insüline tepki vermede oluşan bir aksaklık olarak nitelendirilir ve dünyada en yaygın görülen üçüncü hastalıktır. Tip 2 diyabetin altında yatan nedenleri anlayabilmek için, insülin sinyalleşme ağının detaylı bir şekilde incelenmesi gereklidir, ve bu aşamada insülin yol izindeki sinyalleşme hücre içi ve kan glikoz seviyelerinin düzenlenmesi ve diyabetin önüne geçilebilmesi için çözümsel bir yoldur. Ancak, insülin sinyalleşmesinin topolojik yapısı halen tamamlanamamıştır. Bu çalışmada biyolojik açıdan insülin sinyalleşme yol iziyle alakalı olma olasılığı yüksek proteinleri içeren bir protein etkileşim ağı SPA algoritması ve Gen Ontolojisi (GO) terimlerini içeren istatistiksel yöntemler kullanılarak oluşturulmuştur. İnsülin sinyalleşmesiyle bağlantılı 27 insan proteini girdi olarak kullanılmış; 416 protein ve 895 etkileşimden oluşan aday bir ağ elde edilmiştir. Proteinler arası bağlantı dağılımı güç yasası modeline yaklaşan bu insülin sinyalleşme ağı grafik teorisi yöntemleri ile analiz sonucu, biyolojik ağların genel topolojik özelliklerine uygun bulunmuştur; oluşturulan ağyapının boyuttan bağımsız bir topoloji ile küçük dünya özelliklerine sahip olduğu görülmüştür. Kas hücresinin insülin uyarımına bilinen en önemli tepkisi, glikoz taşıyıcı (GLUT4) içeren hücre içi bir vesikülün hücre yüzeyine yer değiştirmesidir. Bundan dolayı, oluşturulan insülin sinyalleşme ağının yanı sıra dinamik modelleme yaklaşımları kullanılarak, metabolik insülin ağı da bu çalışma kapsamında incelenmiştir. Metabolik modeli kontrol eden reaksiyonlar, metabolik yolizi analizi yöntemleri kullanılarak PI3-kinaz (bir çeşit yağ kinazı) ve AKT (protein kinaz B) proteinlerinin fosforilasyonu olarak bulunmuştur. Aynı zamanda bu çalışma, literatürle uyum içinde olarak, insülin receptörlerinin, PKC, PI3K ve Akt enzimlerinin, insülin sinyalleşmesinin düzenlenmesinde önemli rolleri olduğunu ortaya koymuştur. Bu moleküllere ek olarak, yüksek bağlantıya sahip GRB2 proteininin bir merkez olduğu saptanmış, ve bu yüzden ilaç hedefi olarak dikkatle incelenmesi gerektiği anlaşılmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Mammalian cells integrate information from complex intracellular signal transduction pathways to make decisions in response to changes in the environment (Durmuş Tekir *et al.*, 2009). Systematic genome-wide and pathway-specific protein-protein interaction screens have generated a putative, organizing framework of the spatial interconnectivity of a large number of human proteins, including numerous therapeutically relevant disease-associated proteins. The intrinsic value for drug discovery is that these physical protein-protein interaction networks may contribute to a mechanistic understanding of the pathophysiology of disease and can aid in the identification and prioritization of tractable targets and generate hypotheses on how to best drug non-tractable, disease-associated targets (Ruffner *et al.*, 2007). Understanding cellular signaling therefore is central for gaining insight into the molecular mechanisms behind diseases as well as adaptation of living cells to changes in the environment. A very important signaling network in mammalian cells is insulin signaling pathway, because glucose homeostasis depends upon the balance between hepatic glucose production and glucose utilization by the major insulin-dependent tissues, such as liver, adipose, and muscle, and by insulin-independent tissues, such as brain and kidney. Failure of these target cells to respond to insulin, a state known as insulin-resistance, is a major attribute to the pathological manifestations associated with diabetes, an ever increasing epidemic of the 21st century (Cheatman and Kahn, 1995). Reconstruction of protein interaction networks that represent groups of proteins contributing to the same cellular function is a key step towards quantitative studies of signal transduction pathways. In recent years the high-throughput techniques such as transcriptomics, proteomics and metabolomics enabled the collection of genome scale experimental data and lead to the reconstruction of large-scale signaling networks (Papin and Palsson, 2004b; Oda *et al.* , 2005a; Oda *et al.* , 2005b; Oda and Kitano, 2006). Insulin signaling mechanism is very important for the regulation of glucose levels in blood and for the prevention of the diabetes. In order to understand the reasons underlying diabetes, the insulin signaling network was reconstructed in *Homo sapiens* and analyzed both structurally and dynamically in detail within the framework of the present study. The ultimate aim of this project was to identify any unknown components of the

insulin network through protein-protein interaction data, and to investigate potential drug targets for the improvement of insulin resistance or defective insulin signaling.

In the last decade, for the analysis of large-scale networks, several methods have been developed to understand the structure of the systems of molecular interactions (Jeong *et al.*, 2000; Barabasi and Oltvai, 2004). Within the framework of this thesis, the graph theoretic analysis was used to detect the components that are well or poorly connected in the interaction network and the results can guide experimental studies on protein interactions (Durmuş Tekir *et al.*, 2009). An important response of the muscle cell to insulin stimulation is known to be the translocation of intracellular vesicles containing a glucose transporter (GLUT4) to the cell surface. Therefore, beside the reconstructed insulin signaling network, the metabolic insulin network was also investigated by dynamic modeling approaches. For analyzing the metabolic network (such as insulin metabolism) of species with completely sequenced genomes (e.g. *H. sapiens*), it is necessary to have special tools to handle the problems caused by the high complexity of those networks (Bork *et al.*, 1999). The program of CellNetAnalyzer / FluxAnalyzer V. 9.2 (Klamt *et al.*, 2003) was here used to perform metabolic pathway analysis (MPA), that is, to determine the elementary flux modes (EFM) and the control effective fluxes (CEF). In addition to these, metabolic control analysis (MCA) was performed to provide a quantitative description of the substrate flux in response to the changes in the system parameters. Moreover, the linear paths, beginning at ligands and ending at phenotypes, were found by NetSearch algorithm and investigated in detail.

Following this brief introduction, the second chapter gives information about cell signaling and introduces general signal transduction pathways and signaling molecules. Cell signaling studies are given in two parts as the reconstruction of signaling networks and their analysis both dynamically and structurally. The chapter also includes information about metabolic insulin network. The third chapter gives the computational methods used in this research. The structural analysis techniques such as graph theory and pathway analysis are explained in detail. The fourth chapter is focused on the results of reconstruction and analyses of the insulin signaling network. Finally, in the fifth chapter of "Conclusions and Recommendations", the summary of the main results and the main contributions to the research area are given with the recommendations for future work.

2. BACKGROUND ASPECTS

2.1 Cell Signaling

The study of how cells communicate impinges on all aspects of biology, from development to disease. Like all living things, all cells must continually sense their surrounding environment and make decisions on the basis of that information. Single cell organisms must be able to tell which nutrients are nearby and regulate their metabolic processes accordingly. Cells in multicellular organisms such as human must sense the presence of neighbouring cells and hormones when making decisions such as to whether to proliferate, move or die. These processes all require the transfer of information (Figure 2.1) from detection systems referred to as receptors through intermediate molecules within the cell, to cause changes in the expression of genes and the activity of enzymes — the specialized protein machinery that carries out all of a cell's functions. 'Signal transduction', 'cell signaling' or simply 'signaling' is the study of the mechanisms by which this transfer of biological information comes about. Signaling can be studied at the level of the individual cell or the whole organism. For individual cells, signaling is crucial to decisions about division, specialization, death and metabolic control. In more specialized cells it is central to immunity and the transmission of nerve impulses, to name but two examples. At the level of whole multicellular organisms, signaling controls growth and development, as well as aspects of metabolism and behaviour. Not surprisingly, then, signaling malfunctions underlie many human diseases (i.e. cancer – much of the basic work on the regulation of cell proliferation has been carried out to obtain a better understanding of cancer - AIDS, obesity, diabetes, asthma and infectious diseases). It can be hoped that in depth characterization of signaling pathways will eventually lead to an ability to intervene in diseases in which those pathways are defective (Downward, 2001).

In general, a trans-membrane protein complex works as a receptor in an intracellular signaling pathway. Extracellular signaling molecule (ligand) binds to the receptor and activates the signal transmission. A cascade of intracellular signaling proteins transmits the signal to the target proteins and finally initiates a response (Figure 2.1). The response may be alteration of metabolism if the target protein is a metabolic enzyme. On the other hand,

if the target protein functions as a transcriptional regulator, the response may be changes in gene expression. Also, the signal may be transmitted to a cytoskeletal protein, which alters the shape or location of the cell.

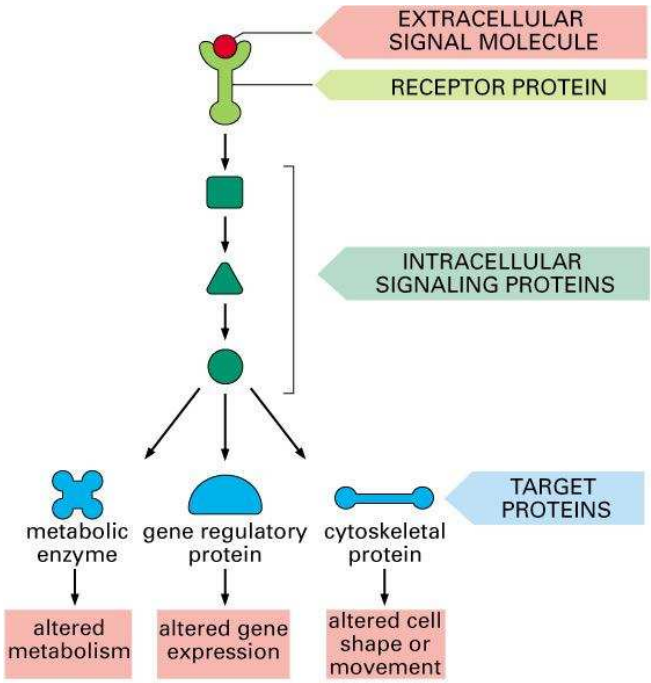


Figure 2.1. Simplified view of an intracellular signaling pathway (Alberts *et al.*, 2002)

The general organization is very similar for all intracellular signaling pathways. In Figure 2.2, the complexity of intracellular signaling pathways is viewed in general terms. As seen in the figure, the complexity comes from the different levels of signal transmission, large number of elements present in the mechanism and the differentiation of the individual proteins with the same function.

2.1.1 Ligand

An extracellular signaling molecule is called 'ligand'. Yeast cells communicate with one another for mating by secreting several kinds of small peptides. However, in higher animals, cells communicate by means of hundreds of signaling molecules, including proteins, small peptides, amino acids, nucleotides, steroids, retinoids, fatty acid derivatives, and even dissolved gases such as nitric oxide and carbon monoxide. Most of these signaling molecules are secreted from the signaling cell by exocytosis or released by

diffusion through the plasma membrane. On the other hand, some signaling molecules remain tightly bound to the cell surface and influence only cells that contact the signaling cell (Alberts *et al.*, 2002).

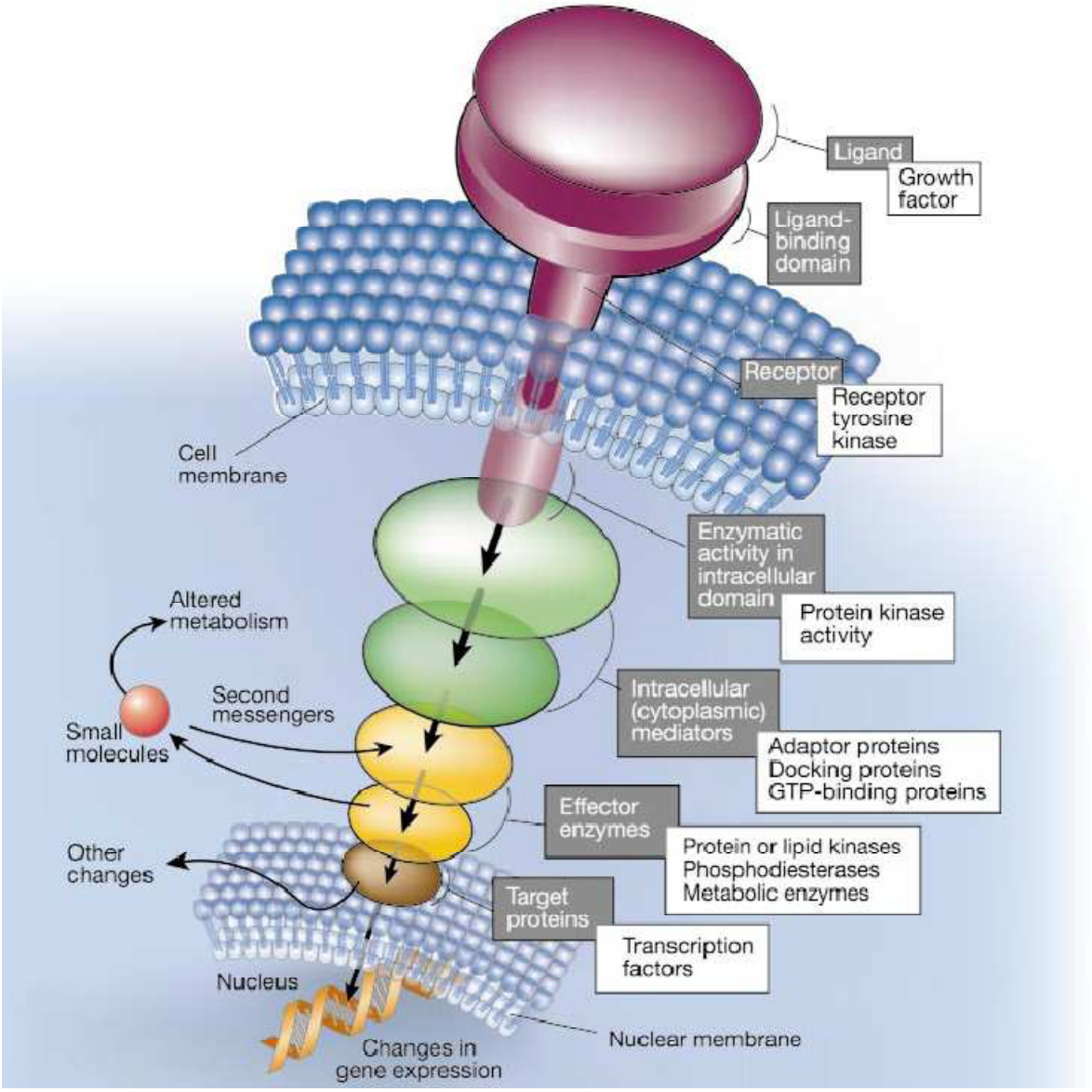


Figure 2.2. A more detailed intracellular signaling pathway (Downward, 2001)

2.1.2 Receptor

Regardless of the nature of the signal, the target cell responds by means of a specific protein called a receptor. It specifically binds the signaling molecule and then initiates a response in the target cell. Receptors are most commonly found at the cell surface, where they bind to extracellular molecules that cannot penetrate the plasma membrane – the lipid boundary between the cell and the outside world.

All water-soluble signaling molecules, as well as some lipid-soluble ones, bind to specific receptor proteins on the surface of the target cells they influence. These cell surface receptor proteins act as signal transducers: they bind the ligand with high affinity and convert this extracellular event into one or more intracellular signals that alter the behavior of the target cell. As a result of engagement of these receptor proteins by their binding partners (‘ligands’), a signal is transferred across the plasma membrane. In most cases, the receptor itself spans the membrane. Ligand binding causes a change in the shape of the protein; this change is transmitted from the extracellular part of the receptor to part inside the cell (Alberts *et al.*, 2002; Downward, 2001).

There are three known classes of cell-surface-receptor proteins: ion-channel linked, G-protein linked, and enzyme linked. In some cases, the receptors called intracellular receptors can be inside the target cell. Ion-channel linked receptors are involved in rapid synaptic signaling between electrically excitable cells in which ion permeability of the plasma membrane changes. G-protein linked receptors act indirectly to regulate the activity of target protein, called a G-protein. Enzyme linked receptors either functions directly as enzymes or they are associated with enzymes. When receptors are inside the target cell, the ligand has to enter the cell to activate them. These signaling molecules therefore must be sufficiently small and hydrophobic to diffuse across the plasma membrane (Alberts *et al.*, 2002).

2.1.3 Features of Signaling

The main features of signaling are specificity, amplification, desensitization/ adaptation and integration, which are explained shortly in the following subsections.

2.1.3.1 Specificity: A signaling molecule fits specifically to the binding site on its complementary receptor (Figure 2.3); other signaling molecules do not fit. Each cell is programmed to respond to specific combinations of signaling molecules. Different cells can respond differently to the same chemical signal since a single signaling molecule often has different effects on different target cells (Nelson and Cox, 2004).

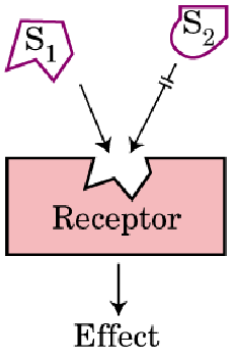


Figure 2.3. Specificity of receptor (Nelson and Cox, 2004)

2.1.3.2 Amplification: When enzymes activate enzymes, the number of affected molecules increases geometrically in an enzyme cascade (Figure 2.4) (Nelson and Cox, 2004).

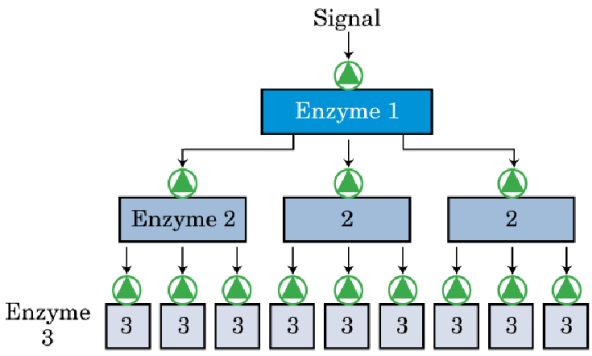


Figure 2.4. Signal amplification (Nelson and Cox, 2004)

2.1.3.3 Desensitization/Adaptation: Receptor activation triggers a feedback circuit that shuts off the receptor or removes it from the cell surface (Figure 2.5) (Nelson and Cox, 2004).

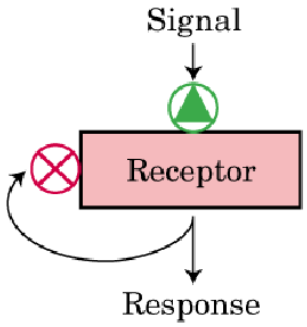


Figure 2.5. Feedback of signaling (Nelson and Cox, 2004)

2.1.3.4 Integration: When two signals have opposite effects on a metabolic characteristic such as the concentration of a second messenger, or the membrane potential, the regulatory outcome results from the integrated input from both receptors (Figure 2.6) (Nelson and Cox, 2004).

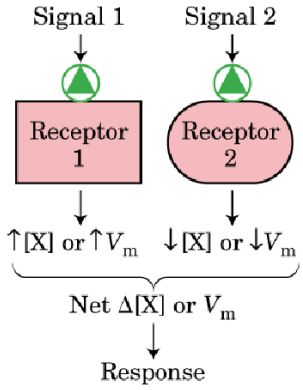


Figure 2.6. Signal integration (Nelson and Cox, 2004)

2.1.4 Types of Cell Signaling

In biological systems, signal transmission occurs mostly through two mechanisms: (i) protein-protein interactions and enzymatic reactions such as protein methylation, phosphorylation and dephosphorylation (post-translational modifications) or (ii) protein degradation or production of intracellular messengers (Bhalla and Iyengar, 1999).

Papin and Palsson classified the signal transduction events into four basic categories according to input-output relationships (Figure 2.7). The classical case of a transduced signal relates a single input to a single output. Some outputs require the concatenation of

multiple inputs. Other signaling interactions occur in which the transduction of a single input generates multiple outputs, a type of signal pleiotropy. Complex signaling events arise as multiple inputs trigger interacting signaling cascades that result in multiple outputs (Papin and Palsson, 2004a).

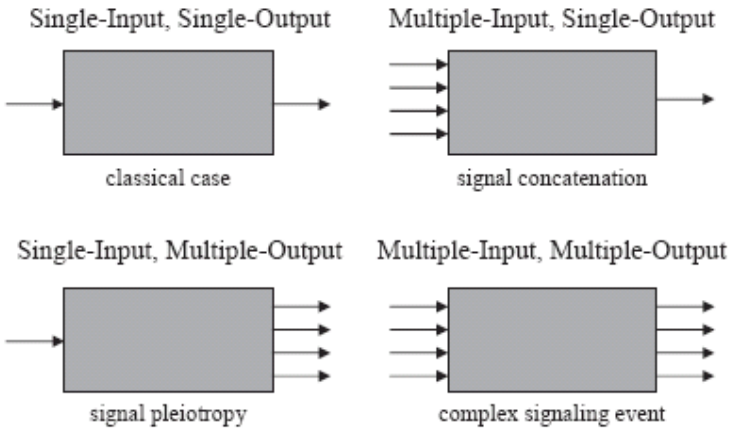


Figure 2.7. Classification of signal transduction input–output relationships (Papin and Palsson, 2004a)

2.1.5 Signaling Networks

The cell signaling mechanisms have been studied extensively over the past 20 years and the advent of high-throughput technologies is enabling the reconstruction of large scale signaling networks. After careful reconstruction of signaling networks, their properties must be described within an integrative framework that accounts for the complexity of the cellular signaling network and that is amenable to quantitative modeling (Papin *et al.*, 2005). The network reconstruction with high-throughput technologies and mathematical modeling to analyze network properties are two main parts of this section.

Recent genomic technologies have resulted in increasingly more detailed descriptions of signaling mechanisms, which have generated reconstructions of ever-larger signaling networks (Papin and Palsson, 2004a; Oda *et al.*, 2005a; Oda *et al.*, 2005b; Oda and Kitano, 2006). Such reconstructions will enable a systemic understanding of signaling network function, which is crucial for studying diseases as diverse as asthma and cancer

(Finkel and Gutkind, 2003). The characterization of signaling properties that arise from whole-cell function requires integrated, mathematical descriptions of the relationships between different cellular components (Levchenko, 2003; Weng *et al.*, 1999).

2.1.5.1 Signaling Network Reconstruction: Network reconstruction includes a chemically accurate representation of all of the biochemical events that are occurring within a defined signaling network, and incorporates the interconnectivity and functional relationships that are inferred from experimental data. It provides the framework for the application of mathematical methods that can quantitatively describe the properties of signaling networks. Contextual specificity is a crucial consideration in answering five questions for signaling network reconstruction: What proteins and other network components participate? What are the ligand-receptor interactions? What are the receptor-intracellular component interactions? What are the intracellular component interactions? What are the intracellular component-DNA interactions? Genome annotation, biochemical experimentation, cell-physiology characterizations, expression arrays, and other such data sources each provide different types of datum that answer these questions and contribute to the reconstruction of a given cellular signaling network (Papin *et al.*, 2005).

Signaling network reconstruction can be approached in three different ways (Figure 2.8). The first approach consists of reconstructions of highly connected nodes in networks. Such reconstructions involve comprehensively listing the compounds and reactions that are associated with a given protein, ion or metabolite. The second approach to network reconstruction involves forming linear pathways that connect signaling inputs to signaling outputs. For example, such a pathway might be delineation of all of the steps from the binding of a growth factor to its receptor through to the subsequent activation of a transcription factor that induces the expression of target genes. The third approach consists of identifying signaling modules. Such modules historically consist of groups of compounds and proteins that function together under certain conditions. In addition, these modules have led to detailed kinetic analyses that traced the concentrations of various effector proteins and helped to understand processes such as feedback mechanisms (Papin *et al.*, 2005).

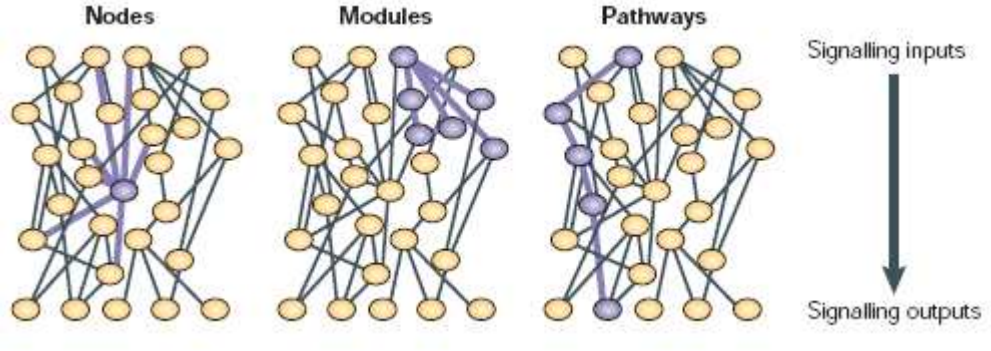


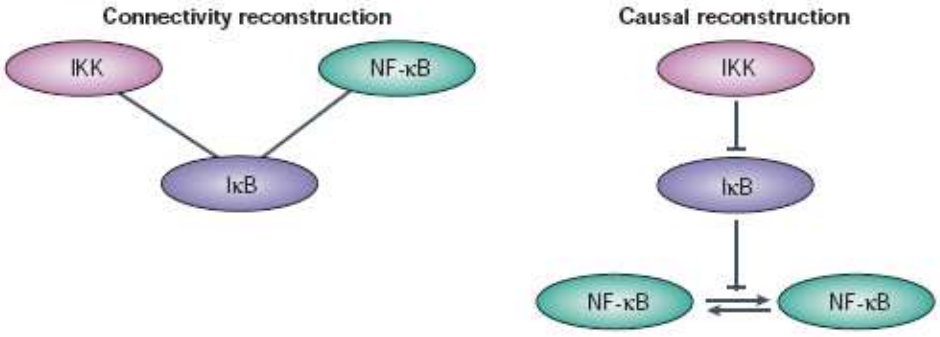
Figure 2.8. Approaches in reconstruction of signaling networks (Papin *et al.*, 2005)

Reactions amongst components in signaling networks are chemical transformations. There are three levels of resolution in reconstructions (Figure 2.9). A connectivity reconstruction lists the associations between network components (for example, nuclear factor (NF)-κB is functionally connected to IκB kinase (IKK) through the inhibitor of NF-κB (IκB)). A more detailed causal reconstruction describes cause-and-effect relationships and is often analyzed with differential equations (for example, IKK interacts with the IκB–NF-κB complex such that NF-κB is activated). As signaling reactions are chemical transformations, they can also be represented by a more mechanistic description – for example, a stoichiometric matrix. This representation accounts for all chemical events that occur in a given network. For example, one IKK complex binds to and phosphorylates one IκB-NF-κB complex with two ATP molecules, which leads to the degradation of IκB and the nuclear localization of NF- κB (R_1-R_4). This relationship can be written out as a series of stoichiometric equations and its accompanying matrix (Papin *et al.*, 2005).

2.1.5.2 Mathematical Analysis of Signaling Networks: Large-scale signaling networks are complex. Their complexity necessitates the use of methods from systems sciences, which are quite mathematical. Structural and dynamic analyses, which can provide different results that can be integrated to characterize the properties of reconstructed signaling networks, measure the time-invariant/topological and the time-variant properties of a network, respectively (Papin *et al.*, 2005).

Structural Network Analysis: Large-scale networks can undergo structural analysis in their entirety, as this does not require an extensive knowledge of the parameters that have been determined from detailed experimentation. Structural analyses of connectivity

reconstructions can generate hypotheses regarding the structure of the global network as well as the function of individual proteins. Recently published examples illustrate the analyses that have led to hypotheses concerning global, modular and individual protein function (Jeong *et al.*, 2001; Rives and Galitski, 2003; Bu *et al.*, 2003). Initial structural analyses of causal reconstructions of signaling networks have also highlighted the value of these analyses in describing network properties (Schuster *et al.*, 2000). So far, stoichiometric analyses of signaling networks are limited, owing to a lack of corresponding reconstructions. This analysis of a stoichiometric reconstruction has led to descriptions of protein synthesis requirements and energy demands of signaling networks, as well as mathematical definitions of network properties such as crosstalk and pathway redundancy (Papin *et al.*, 2005).



Stoichiometric reconstruction

R_1: IKK + IκB–NF-κB + 2 ATP → IKK–IκB–NK-κB–ATP
R_2: IKK–IκB–NK-κB–ATP → IKK + IκBpp–NF-κB + 2 ADP
R_3: IκBpp–NF-κB → IκBpp (degraded) + NF-κB (cytoplasm)
R_4: NF-κB (cytoplasm) → NF-κB (nucleus)

Compounds	Reactions			
	R_1	R_2	R_3	R_4
IKK	−1	+1	0	0
IKK complex	+1	−1	0	0
IκBpp–NF-κB	0	+1	−1	0
IκBpp (degraded)	0	0	+1	0
NF-κB (cytoplasm)	0	0	+1	−1
IκB–NF-κB	−1	0	0	0
ATP	−2	0	0	0
ADP	0	+2	0	0
NF-κB (nucleus)	0	0	0	+1

Figure 2.9. Three levels of resolution in reconstructions (Papin *et al.*, 2005)

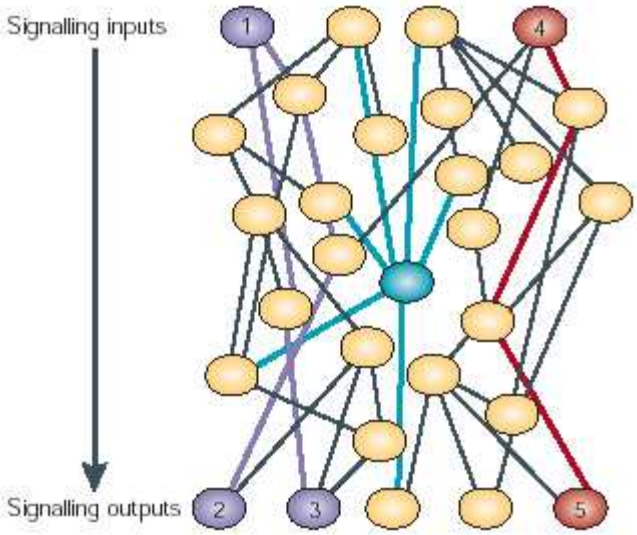


Figure 2.10. Structural analyses of signaling networks (Papin *et al.*, 2005)

Structural analyses can identify components that are well or poorly connected (and therefore of potential interest for drug targeting). For example, the blue component is the most highly connected in the schematic of a signaling network (Figure 2.10), and so drugs that inhibit the activity of this hypothetical component could have the broadest effect on the functions of the network. Structural analyses can also characterize which signaling inputs generate which signaling outputs. For example, in the schematic of signaling network (Figure 2.10) the signaling inputs 1 and 4 can generate signaling outputs 2, 3 and 5.

Dynamic Network Analysis: A dynamic analysis of a reconstructed signaling network can be carried out once the associated kinetic parameters are known. The timescales that are associated with signaling processes can be estimated, and crudely divided into two groups: signaling activities and signaling responses. Signaling activities typically occur rapidly. For example, most protein conformational changes, kinase/phosphatase reactions (Goodman *et al.*, 1998; Vuong *et al.*, 1991), and the physical movement of signaling compounds by diffusion or cytoskeleton-dependent mechanisms (Teruel *et al.*, 2000; Theurkauf *et al.*, 1994) occur over a timeframe that ranges from fractions of a second to seconds. However, signaling responses can occur over a wider range of timescales. Signaling responses that are coupled with metabolic processes or intermediate phenotypes can occur over a timeframe of fractions of a second (Stryer, 1995; Neves *et al.*, 2002) as

can elements of chemotactic and mechanotransduction behaviour (Stryer, 1995). However, other signaling responses occur over a timescale that is an order of magnitude slower. For example, transcriptional events (Zubay, 1973; McAdams and Arkin, 1998), cellular growth (Alberts *et al.*, 2002) and receptor internalization (Bomsztyk *et al.*, 1989; Chang *et al.*, 1996; Jullien *et al.*, 2002) require several minutes, or longer, in response to a signal. This timescale separation is a crucial consideration for dynamic network analyses and can lead to simplifications that enable more thorough analyses, which would otherwise be difficult (Papin *et al.*, 2005).

As numerical values for kinetic parameters are typically difficult to obtain (Bailey *et al.*, 2001), dynamic analyses are usually only carried out for causal and stoichiometric reconstructions (Figure 2.9) of smaller cellular signaling network reconstructions. These studies have analyzed complex network properties. The coupling of experimental data with these mathematical analyses can enable the identification of previously unknown signaling mechanisms. An elegant study that shows the benefit of integrative experimental and mathematical analyses deciphered the importance of particular IκB isoforms in feedback loops that involved the NF-κB signaling module. Predictions were experimentally verified in knockout mouse models (Hoffman *et al.*, 2002). The WNT signaling module, which is important for development as well as oncogenesis, was recently represented using an extensive set of kinetic reactions (Lee *et al.*, 2003). Predictions were made for dynamic profiles of concentrations of β-catenin and other signaling mediators, and these matched experimental results. These dynamic analyses of signaling modules show the complex properties that can be studied once the reconstruction of only a limited number of reactions have been completed and experimental data are integrated with the model predictions (Papin *et al.*, 2005).

The existing biological knowledge for a given system is composed of several types of datum (Figure 2.11). Each datum type provides unique information that can be incorporated into a chemically accurate reconstruction (for example, a stoichiometric matrix). The first type of datum includes the identification of components and endpoints of a network (for example, genome sequencing or genome-wide location analysis). The second type of datum characterizes the interactions between network components (for example, yeast two-hybrid and immunoprecipitation data identify protein–protein

interactions and protein complexes). The third datum type describes the network behaviour of the integrated components (for example, perturbation analysis and cDNA arrays delineate how entire networks function under various conditions). Each of these results provides unique types of datum that can be used to generate a cellular network reconstruction (for example, genome sequencing enables the annotation of the genes that are present in a given organism). With a network reconstruction, dynamic and structural analysis techniques can be used to describe emergent properties of the network, and generate new hypotheses. These characterizations then expand and revise the foundation of biological knowledge for the given system. This process can be iterated to offer increasingly more accurate descriptions of a given biochemical network (Papin *et al.*, 2005).

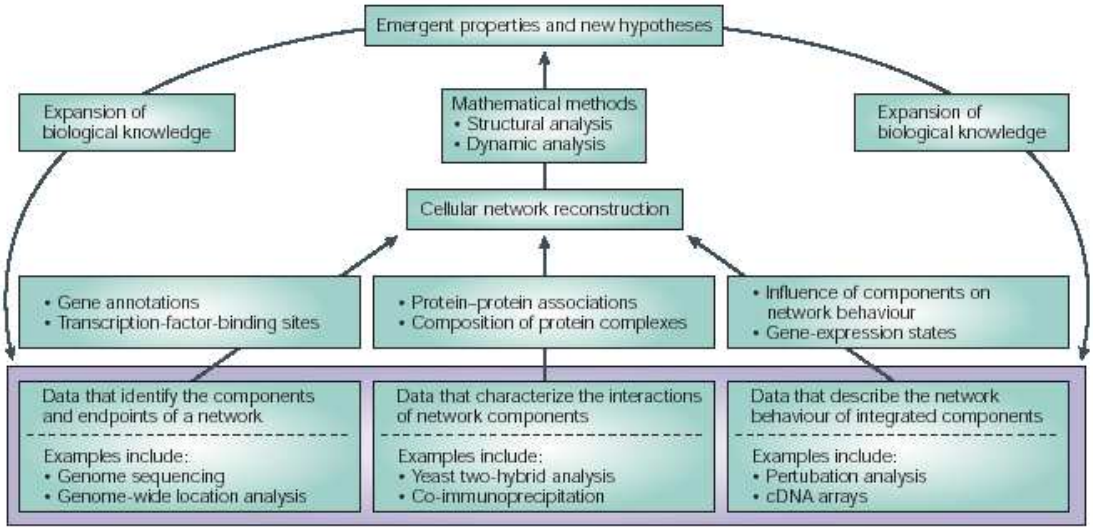


Figure 2.11. Integrative and iterative process of cellular signaling network reconstruction (Papin *et al.*, 2005)

2.2 Insulin Signaling in *H. Sapiens*

Diabetes is characterized by the failure to synthesize, secrete or respond to insulin and is the third most prevalent disease in the Western world, affecting at least 3% of the population of Europe and North America and about 100 million worldwide. People with diabetes have increased blood glucose (sugar) levels for one or more of the following three reasons: insulin is not being produced, insulin production is insufficient, and/or the body is

resistant to the effects of insulin. There are two types of diabetes. Lack of severe reduction in insulin secretion due to autoimmune destruction of β cells is responsible for Type 1 diabetes. In at least 75 % of the diabetics, the level of circulating insulin is normal, or even above normal, and the disease is caused by the major target tissues for insulin (muscle, fat, liver) becoming resistant to this hormone which is responsible for Type 2 diabetes. In light of these facts, the pathogenesis and metabolic mechanisms of Type 2 diabetes is utterly important, it is therefore important to elucidate the mechanism by which insulin signals to the cell interior (Cohen *et al.*, 1997).

2.2.1 Type-2 Diabetes

At the whole-body level, glucose is used as a source of energy, to provide fuel for skeletal muscle and other tissues. In the postprandial state after a meal, insulin is secreted from the pancreatic β-cells into the circulation to promote glucose uptake into insulin-sensitive tissues and to decrease the hepatic glucose production. When glucose levels are low, the pancreatic α-cells releases glucagon into the circulation that promotes glucose production from the liver. This system has to be tightly regulated to achieve glucose homeostasis.

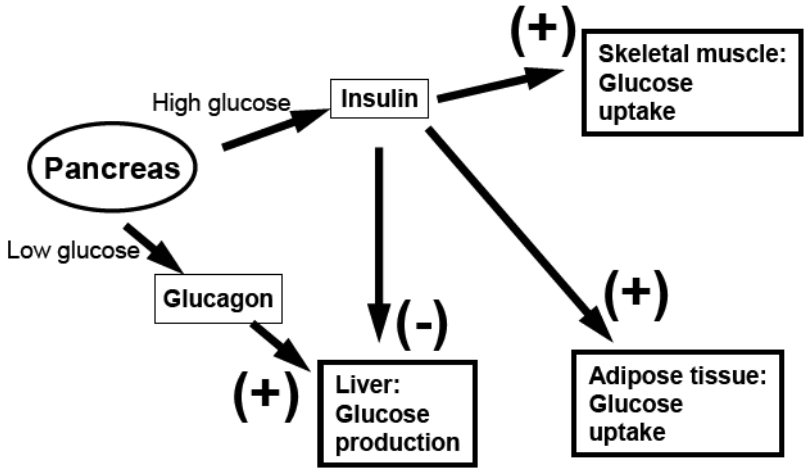


Figure 2.12. Maintenance of glucose homeostasis (Karlsson, 2005)

High glucose levels in the circulation trigger the pancreas to release insulin. Insulin release leads to an increased glucose uptake in skeletal muscle and adipose tissue, and a decreased glucose production from the liver. On the other hand, low glucose levels lead to glucagon release to increase glucose production from the liver as shown in Figure 2.12. Type 2 diabetes is a metabolic disorder where the efficacy of insulin is insufficient to regulate glucose uptake and hepatic glucose production (Karlsson, 2005). Thus, in order to understand the reasons underlying Type-2 diabetes, insulin signaling network should be analyzed in detail.

2.2.2 Glucose Transport

Glucose is a hydrophilic molecule and thus, cannot diffuse through the lipid bilayer of cell membrane. Therefore, glucose entry into the cell needs to be facilitated by membrane transporters. There are mainly two different families of glucose transporters in humans, the sodium dependent glucose transporter family (SGLT) and the facilitative glucose transporters (GLUT) (Joost *et al.*, 2002; Scheepers *et al.*, 2004). While the sodium dependent glucose co-transporters are mainly involved in glucose absorption in the intestine and the kidneys, GLUT family members are facilitative glucose transporters. At present there are 14 identified genes coding for individual proteins of the GLUT family (GLUT1-14). The GLUT's have different tissue expression and different affinity for glucose (Scheepers *et al.*, 2004). In skeletal muscle GLUT4 is the predominantly expressed isoform (Birnbaum, 1989; Fukumoto *et al.*, 1989) and it is localized intercellularly and translocated to the surface membrane in response to insulin (Karlsson, 2005; Hirshman *et al.*, 1990; Kristiansen *et al.*, 1996; Ryder *et al.*, 2000).

Impaired insulin-stimulated glucose transport in skeletal muscle is the rate determining step in the reduced glycogen synthesis observed in insulin resistant Type 2 diabetic patients (Cline *et al.*, 1999). Firstly, insulin binds to the insulin receptor (1), thereby initiating a signaling cascade through the insulin signaling pathway (2). The intracellular membrane bound GLUT4-containing vesicles are then translocated to the cell surface membrane (3), where they ultimately integrate GLUT4 in the membrane (4) which then facilitates glucose transport into the muscle cell (5).

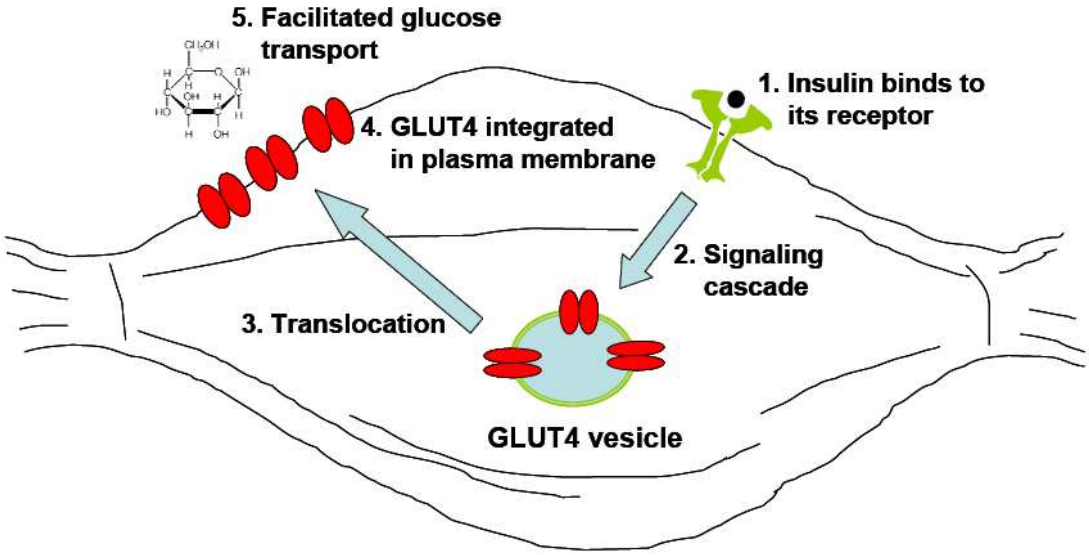


Figure 2.13. Principle of insulin-induced glucose uptake in skeletal muscle (Karlsson, 2005)

2.2.3 Pathways of Insulin Signaling

Insulin is an essential peptide hormone for maintaining glucose homeostasis and regulating carbohydrate, lipid and protein metabolism. Insulin signaling controls several processes by a complex, highly integrated network (Figure 2.14). In the presence of insulin, the insulin receptor (IR) phosphorylates insulin receptor substrate proteins (IRS proteins) that are linked to the activation of two main signaling pathways: the phosphatidylinositol 3-kinase (PI3K)-Akt/protein kinase B (PKB) pathway, which is responsible for most of the metabolic actions of the insulin, and the Ras-mitogen activated protein kinase (MAPK) pathway, which regulates expression of some genes and cooperates with the PI3K pathway to control cell growth and differentiation (Taniguchi *et al.*, 2006). The phenotypes of some nodes which have roles in insulin signaling network are explained in Table 2.1.

Table 2.1. Phenotypes of mouse models with deletion of components of the insulin signaling pathway (Li and Zhang, 2007)

Gene	Phenotype
Insulin receptor	Severe diabetes, postnatal death at 3-7 days
IGF1 receptor	Growth retardation, normal glucose homeostasis
IRS-1	Insulin and IGF1 resistance, β cell hyperplasia, metabolic syndrome
IRS-2	Insulin resistance, reduced β cell mass, type 2 diabetes
IRS-3	No apparent phenotype
IRS-4	No apparent phenotype
P85α (hetero)	Improved insulin sensitivity
Akt2	Insulin resistance and glucose intolerance
PTP1B	Improved insulin sensitivity, resistance to high-fat diet induced obesity
SHIP2	Improved insulin sensitivity
GLUT4	Cardiac hypertrophy, normal glucose homeostasis, severe insulin resistance, glucose intolerance, hyperinsulinemia, insulin resistance

Although a simple scheme of divergent pathways looks sufficient to explain insulin signaling, the critical nodes become apparent when one looks at the number of gene and protein isoforms that are involved in the activation of Akt and the generation of metabolic effects. For instance, the IR has two splice isoforms, which are usually co-expressed in cells that also express the highly related insulin-growth factor-1 receptor (IGF1R) that can be activated by insulin. Both the IR and IGF1R can phosphorylate at least six known substrate proteins that are capable of interacting with eight known forms of the PI3K regulatory subunit which associate with three forms of the PI3K catalytic subunit that can generate phosphaditylinositol-3,4,5-triphosphate (PIP_3), which leads to the activation of the three known isoforms of Akt. When the combinatorial possibilities in these first steps are accounted, there are over 1,000 possible combinations, and this number increases further if the compartmentalization, kinetics and downstream components of the signaling network are taken into consideration. Among the hundreds of molecules that have been shown to be involved in the insulin-signaling pathway, the IR/IRS, the PI3K and Akt proteins have

been identified as the three best-defined critical nodes. Each protein is essential for insulin action, consists of several isoforms with unique functions. The first critical node in the insulin signaling network is the IR and the associated IRS proteins which have a role in insulin resistant, hyperinsulinaemic states, including obesity and Type-2 diabetes. The second one, PI3K activates critical regulators of insulin signaling and the third one, Akt mediates most of the PI3K-mediated metabolic actions of insulin through the phosphorylation of several substrates, including other kinases, signaling proteins and transcription factors which are explained below (Taniguchi *et al.*, 2006).

2.2.3.1 PI3K/Akt Pathway: PI3 kinase plays a pivotal role in the metabolic and mitogenic actions of insulin. Activated PI3K specifically phosphorylates PI substrates to produce PI(3)P, PI(3,4)P2 and PI(3,4,5)P3. Acting as second messengers, these phospholipids recruit the PI3K-dependent serine/threonine kinases (PDK1) and Akt from cytoplasm to the plasma membrane by binding to the "pleckstrin homology domain" (PH domain) of kinases. Activated Akt phosphorylates and regulates the activity of many downstream proteins involved in multiple aspects of cellular physiology. Among others, Akt phosphorylates and regulates components of the glucose transporter 4 (GLUT4) complex, protein kinase C (PKC) isoforms, all of which are critical in insulin-mediated metabolic effects (Li and Zhang, 2007).

2.2.3.2 Ras/MAPK Pathway: The Ras/MAPK pathway is activated by insulin following the binding of GRB2 and the guanyl nucleotide-exchange factor SOS to cognate phosphotyrosines on IRS proteins SHC and GAB1. This binding triggers the activation of the small GTPase Ras and the subsequent activation of Raf, which triggers a kinase cascade that results in the phosphorylation and activation of the dual-specificity kinases MEK1 (MAPK and ERK kinase 1) and MEK2, which in turn phosphorylate MAPK/ERK1 and ERK2 on threonine and tyrosine residues. This pathway is necessary for the growth and differentiation aspects of insulin actions (Taniguchi *et al.*, 2006).

2.2.3.3 GLUT4 Translocation: An important response of the muscle cell to insulin stimulation is the translocation of intracellular vesicles containing a glucose transporter (GLUT4) to the cell surface. Insulin promotes glucose uptake by muscle and adipose tissue

via stimulation of GLUT4 from intracellular sites to the plasma membrane. Attenuated GLUT4 translocation and glucose uptake by muscle and fat cells following insulin stimulation represent a prime defect in insulin resistance (Li and Zhang, 2007). The PI3 kinase /Akt pathway has been demonstrated to be upstream of GLUT4 translocation.

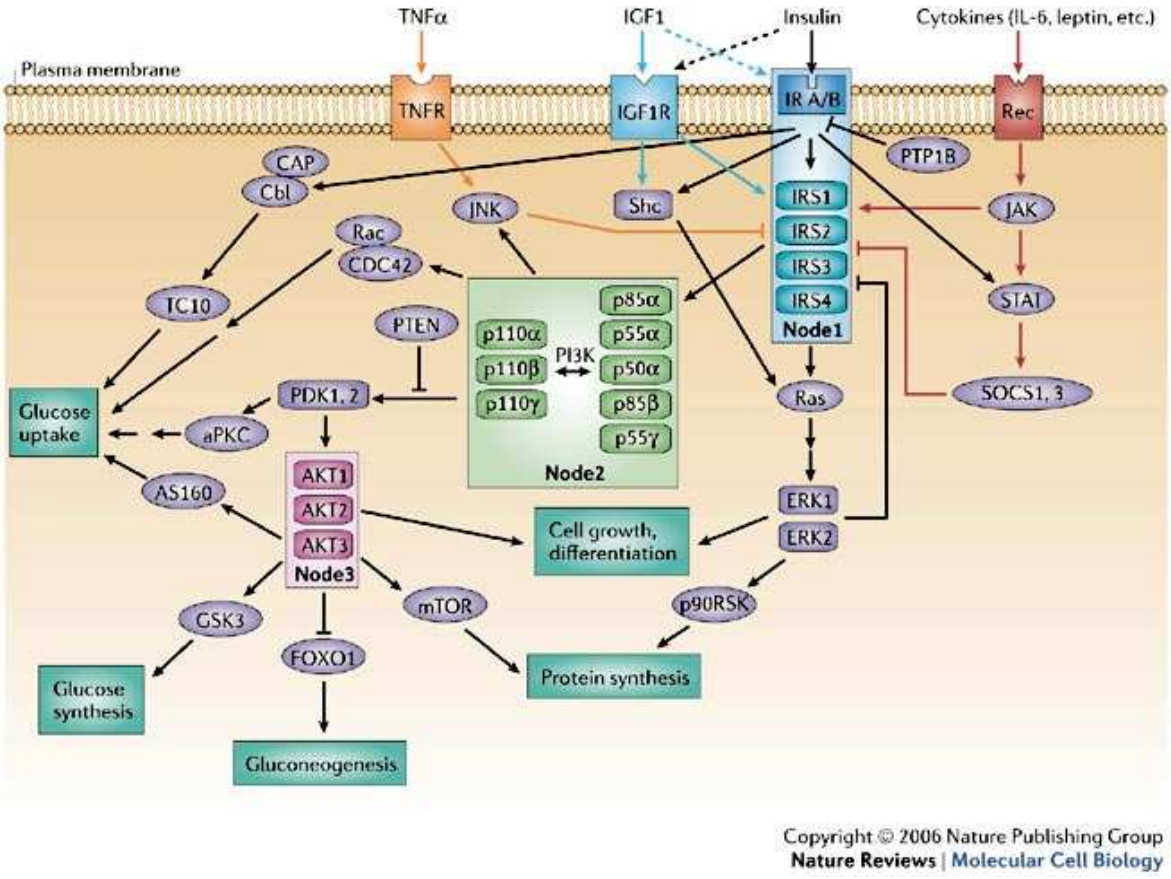


Figure 2.14. A general scheme of insulin signaling network (Taniguchi *et al.*, 2006)

The regulation of this process is only beginning to be understood. On binding insulin, the insulin receptor undergoes receptor autophosphorylation and enhanced tyrosine kinase activity. Subsequently, intracellular substrates (e.g., insulin receptor substrate-1, IRS-1) are phosphorylated on tyrosine residues that serve as docking sites for downstream SH2 domain containing proteins, including the p85 regulatory subunit of PI3K. The p85 binding to phosphorylated IRS-1 results in activation of the p110 catalytic subunit of PI3K that catalyzes production of phosphoinositol lipids including phosphatidylinositol 3,4,5-trisphosphates [PI(3,4,5)P_3] that activate the Ser/Thr kinase 3-phosphoinositide-dependent

protein kinase (PDK)-1. PDK-1 phosphorylates and activates other downstream kinases, including Akt and protein kinase C (PKC)-ζ, that mediate translocation of GLUT4. PTP1B is a protein tyrosine phosphatase (PTPase) that negatively regulates insulin signaling pathways by dephosphorylating the insulin receptor and IRS-1. Interestingly, IRS-1 and PTP1B upstream from Akt and PKC-ζ have recently been identified as substrates for these downstream kinases, suggesting that feedback mechanisms exist (De Fea and Roth, 1997; Paz *et al.*, 1999; Ravichandran *et al.*, 2001) (Figure 2.15). Elements downstream from Akt and PKC-ζ linking insulin signaling pathways with trafficking machinery for GLUT4 are unknown (Pessin *et al.*, 1999). Thus a complete understanding of mechanisms regulating the metabolic actions of insulin has remained elusive.

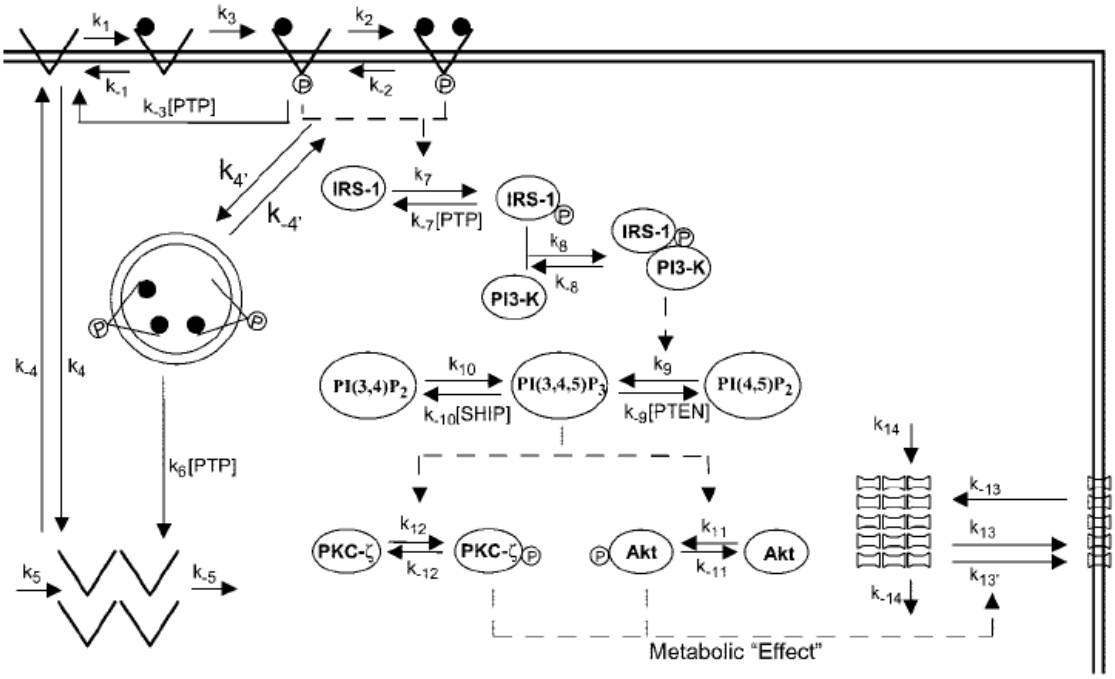


Figure 2.15. Complete model of GLUT4 translocation of metabolic insulin signaling (Sedaghat *et al.*, 2002).

2.2.4 Drug Design through Protein-Protein Interaction Networks

An insidious increase in features of the 'metabolic syndrome' – obesity, insulin resistance and dyslipidaemia – has conspired to produce a worldwide epidemic of Type 2 insulin-resistant diabetes mellitus. The availability of drugs that affect underlying

mechanisms may lead to a new therapeutic paradigm for the prevention of diabetes and its complications. Several mechanistic categories are considered for new therapeutic approaches. First one is on reducing excessive glucose production by the liver; second one is on mechanisms to augment glucose-stimulated insulin secretion; third one is specific molecular targets in the insulin signaling pathway; and fourth one is on developing new approaches for obesity and altered lipid metabolism with prospects of net improvements in insulin action (Moller, 2001).

Insulin resistance can be due to multiple defects in signal transduction (such as impaired activation of insulin receptor-tyrosine kinase and reduced activation of insulin stimulated phosphatidylinositol-3-OH kinase (PI3K)). A number of molecular targets are now being investigated as ways of enhancing insulin-mediated signal transduction (Table 2.2). Non-peptide small molecules that can activate the insulin receptor, or potentiate its activation by insulin, have proved elusive. An alternative approach to targeting the insulin receptor itself would be to inhibit enzymes responsible for deactivation of the receptor or downstream targets in the signaling pathway (for example, IRS proteins). A number of specific tyrosine phosphatases (PTPs) have been identified as candidate targets. PTP1B is an intracellular enzyme specifically implicated in the negative regulation of insulin signaling (Goldstein *et al.*, 1998). Recent results from genetic knockout of PTP1B provide strong validation of this particular PTP as a potential target (Moller, 2001).

Other putative negative regulators of insulin signaling have recently been implicated as independent drug targets. Glycogen synthase kinase-3 (GSK-3) has a role in opposing the effect of insulin, by inhibiting the activation of glycogen synthase and the subsequent accumulation of glycogen in muscle (Weston and Davis, 2001). Furthermore, SH2-domain-containing inositol 5-phosphatase type 2 (SHIP2) enzyme was implicated recently as a diabetes target, as heterozygous null mice has markedly enhanced sensitivity to insulin (Clement *et al.*, 2001). Similarly, PKC-ζ could be an additional drug target, as increased muscle PKC-ζ activity has been observed in the context of fatty-acid-induced insulin resistance (Shulman, 2001; Moller, 2001) (Table 2.2).

Table 2.2. Potential drug targets in the insulin signaling pathway (Moller, 2001)

Target	Validation	Potential mechanism(s)
Insulin recetor	Insulin, small molecule activators/ potentiators	Apparent direct activation of the receptor
PTP1B	Efficacy of vanadium compounds; PTP1B null mice (insulin sensitive and obesity resistant); efficacy of PTP1B antisense oligonucleotide	Mediates dephosphorylation of the insulin receptor (and its lyrosyl-phosphorylated substrates)
SHIP2	SHIP2 mice (insulin sensitive)	Dephosphorylation of phosphoinositides (for example, products of PI3K)
GSK-3	Efficacy of GSK-3 inhibitors in rodent models	Phosphorylation of glucogen synthase leading to inhibition of glycogen synthesis; potential negative regulation of other insulin signaling events
PKC-ζ	Activated in muscle in association with fatty-acid-induced insulin resistance	Negative regulation of insulin signaling, potential serine-threonine phosphorylation of IRS proteins

2.3 Effects of Ceramides on Insulin Signaling

Overweight and obesity has increased dramatically worldwide during the last decades. The International Diabetes Institute in Melbourne estimated that 3% of the world population will have diabetes by the year of 2010. An explanation for this startling number of the obese people is the increased availability and promotion of foods high in fat and sugar and to a large extent by lifestyle changes, where less physical activity is performed due to increased urbanization and changes in transport systems, as well as increased hours spent in front of televisions and computers. Obesity is associated with an increased risk for insulin resistance, a state characterized by impaired responsiveness of liver, muscle and adipose tissue to insulin (Karlsson, 2005).

Researchers have recently proposed two mechanisms by which excess adiposity antagonizes insulin action in peripheral tissues. First, when adipocytes exceed their storage capacity, fat begins to accumulate in tissues not suited for lipid storage, leading to the formation of specific metabolites that inhibit insulin signal transduction. Second, obesity triggers a chronic inflammatory state, and cytokines released from either adipocytes or from macrophages infiltrating adipose tissue antagonize insulin action. The sphingolipid ceramide is a putative intermediate linking both excess nutrients (i.e. saturated fatty acids) and inflammatory cytokines (e.g. tumor necrosis factor-a, TNFa) to the induction of insulin resistance (Figure 2.16).

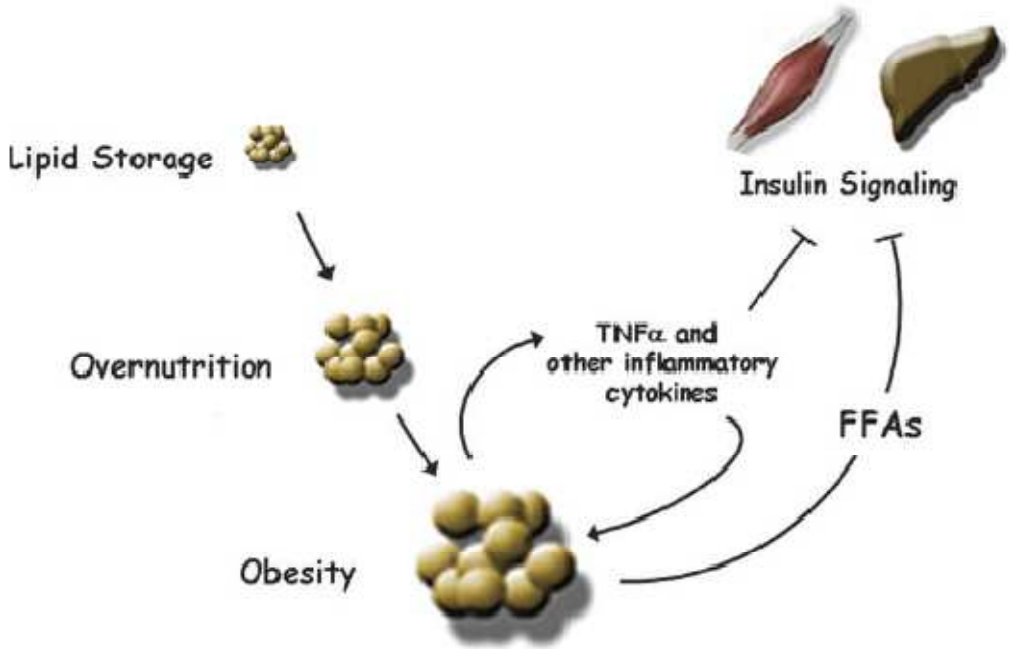


Figure 2.16. Schematic diagram depicting relationships between increased adipose stores and inflammatory cytokines and insulin resistance (Summers, 2006)

Ceramide functions as a mediator in signaling cascades that regulate apoptosis, differentiation and cell cycle arrest (Hannun and Luberto, 2000). An additional role for ceramide in the insulin signaling pathway has more recently been proposed. It inhibits insulin-stimulated glucose uptake, GLUT4 translocation, and/or glycogen synthesis (Summers *et al.*, 1998; Summers *et al.*, 1999). Despite considerable attention, however, a precise understanding of the molecular mechanisms by which this sphingolipid regulates insulin action has remained both elusive and controversial. Specifically, ceramide has been

shown to inhibit several distinct intermediates in the insulin signaling pathway. It does not interfere at the level of IR or IRS-1 phosphorylation, but impairs insulin signaling by inhibition of Akt activation (Stratford *et al.*, 2004). The inhibition of Akt by ceramide is thought to be accomplished by two mechanisms (Figure 2.17 and Figure 2.18). First, ceramide promotes the dephosphorylarion of Akt/PKB by protein phosphatase 2A (PP2A). Ceramides have been shown to directly activate PP2A, which is the primary phosphatase responsible for dephosphorylating Akt/PKB (Summers, 2006).

Ceramide was shown to have no effect on PI3K, but nonetheless blocked the translocation of Akt/PKB from cytoplasmic stores to the plasma membrane. Recent results suggest that PKC-ζ, another enzyme activated by ceramide, inhibits Akt/PKB translocation by phosphorylating Threonine-34, which is present in the enzyme's PH-domain (Stratford *et al.*, 2004; Langeveld and Aerts, 2009; Summers, 2006).

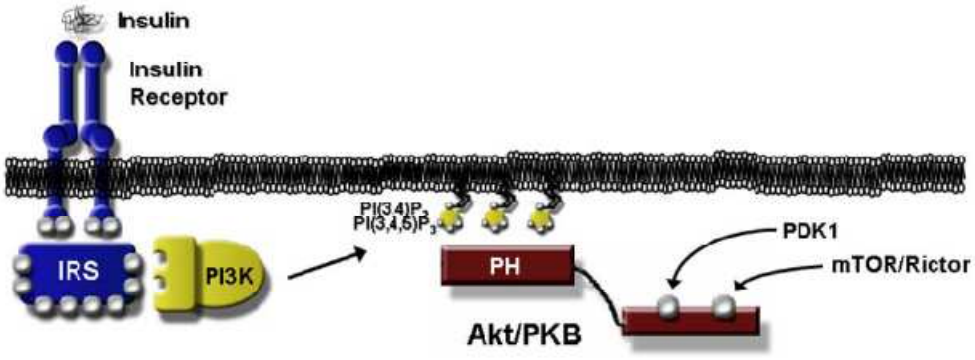


Figure 2.17. Insulin activation of Akt/PKB (Summers, 2006)

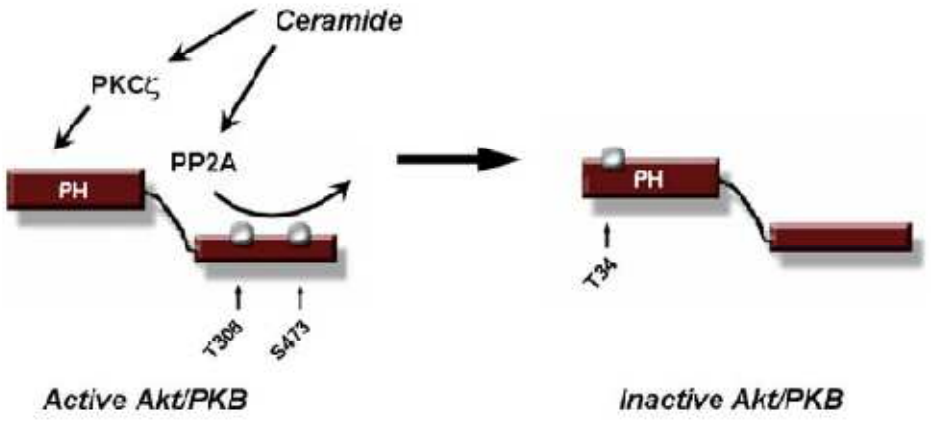


Figure 2.18. Regulation of Akt/PKB by ceramide (Summers, 2006)

3. MATERIALS AND METHODS

3.1 Data

3.1.1 Protein-Protein Interaction Data Sources

A fully comprehensive view of all functionally relevant protein interactions is still not available for any species, not even for relatively simple, single-celled model organisms. However, this information is essential for a systems-level understanding of cellular behavior, and it is needed in order to place the molecular functions of individual proteins into their cellular context (von Mering *et al.*, 2007).

Two public databases, BioGRID (Breitkreutz *et al.*, 2003) and STRING (von Mering *et al.*, 2005) using the results of high throughput experiments and collecting experimentally determined protein-protein interactions from literature as well were applied for the reconstruction of the insulin signaling network.

The Biological General Repository for Interaction Datasets (BioGRID) database (http://www.thebiogrid.org) was developed to house and distribute collections of protein and gene interactions belonging to major model organisms. BioGRID which is a freely accessible database of physical and genetic interactions, currently contains over 198 000 interactions from six different species, as derived from both high-throughput studies and conventional focused studies. The BioGRID provides interaction data with monthly updates to Saccharomyces Genome Database, Flybase and Entrez Gene.

Moreover, the database and online source STRING ('Search Tool for the Retrieval of Interacting Genes/Proteins') generalizes access to protein interaction data, by integrating known and predicted interactions from a variety of sources including existing interaction databases (Alfarano *et al.*, 2005; Salwinski *et al.*, 2004; Zanzoni *et al.*, 2002; Stark *et al.*, 2006; Mishra *et al.*, 2006), PubMed abstracts and several bodies of scientific text. It aims to collect, predict and unify most types of protein-protein associations, including direct and

indirect associations. In order to cover organisms not yet addressed experimentally, STRING runs a set of prediction algorithms (von Mering *et al.*, 2003), and transfers known interactions from model organisms to other species based on predicted ortology of the respective proteins (von Mering *et al.*, 2005). Furthermore, each interaction in the database is annotated with a bench-marked numerical confidence score, which can be used to filter the interaction network at any desired stringency (von Mering *et al.*, 2007).

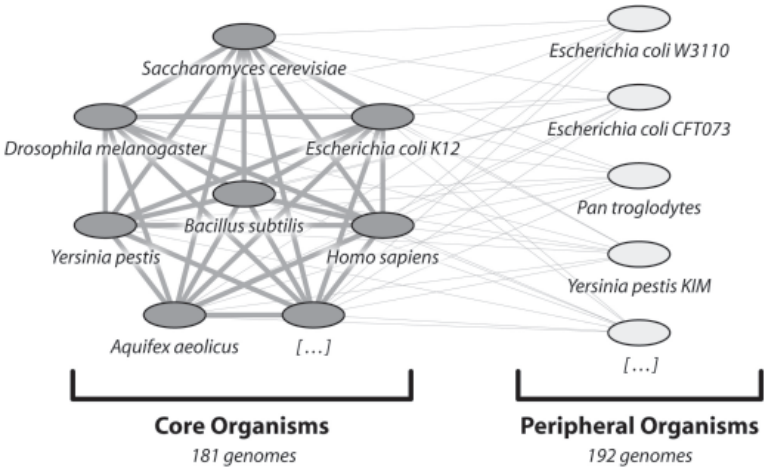


Figure 3.1. Organisms covered by STRING (von Mering *et al.*, 2007)

STRING currently contains 373 fully sequenced organisms (Figure 3.1). These are divided into 'Core Organisms' and 'Peripheral Organisms'. The former include all important model organisms for which experimental data are available, as well as selected representatives for cases of redundant genome sequencing (e.g. when several closely related strains of a bacterial species have been sequenced, only one strain is included). The 'Peripheral Organisms' form the remainder; they tend to be somewhat redundant, and usually have little more than genomic sequence information annotated. For the core organisms, homology relations and interaction transfers are fully computed, whereas the peripheral organisms are only connected to the core but not among themselves (the graphic shows only a small selection of organisms; lines indicate homology searches and interaction transfers). This architecture allows STRING to encompass all sequenced genomes, while still keeping database size and computation time within reasonable limits (von Mering *et al.*, 2007).

3.1.2 Domain-Domain Interaction Data Sources

SUPERFAMILY is a database consisting of structural and functional protein annotations for all completely sequenced organisms. It is freely available at http://supfam.cs.bris.ac.uk/SUPERFAMILY/. The database provides protein domain assignments, at the Structural Classification of Proteins (SCOP), superfamily and family levels, for the predicted protein sequences in over 1000 organisms. Domain assignments are generated using an expert curated set of profile and hidden Markov models. The database includes 46591 total sequences in 1056 unique superfamilies with 41% amino acid coverage for *H. Sapiens*.

3.1.3 Gene Ontology Annotations

Gene Ontology (GO) is a controlled vocabulary used to describe the biology of a gene product in any organism. There are three independent sets of vocabularies, or ontologies, which describe the molecular function of a gene product, the biological process in which the gene product participates, and the cellular component where the gene product can be found. Molecular functions generally correspond to activities that can be performed by individual gene products, but some activities are performed by assembled complexes of gene products. Examples of broad functional terms are ‘catalytic activity’, ‘transporter activity’, or ‘binding’ whereas ‘adenylate cyclase’ or ‘Toll receptor ligand’ are examples of more specific functions. A biological process is series of events accomplished by one or more ordered assemblies of molecular functions. Examples of broad biological process terms are ‘cellular physiological process’ or ‘signal transduction’ whereas ‘pyrimidine metabolism’ or ‘cAMP biosynthesis is an example of more specific biological processes. Cellular component terms help to understand structure of eukaryotic cells. It should be noted that not all the ontologies are applicable to all organisms. ‘Ribosome’, ‘nucleus’ and ‘nuclear membrane’ are examples of cellular component descriptions (The Gene Ontology Consortium, 2000). These ontologies are freely available at http://www.geneontology.org (The Gene Ontology Consortium, 2000).

3.2 Data Visualization Tools

3.2.1 Cytoscape

Cytoscape is an open source bioinformatics software capable of visualizing biomolecular interaction networks and integrating these networks with annotations, gene expression profiles and other state data. It is available at http://www.cytoscape.org/. Core distribution of Cytoscape provides a basic set of features for visualization and data integration. Features of Cytoscape can be extended via plugins. Plugins are available for network and molecular profiling analyses, new layouts, additional file format support, scripting, and connection with databases. Although Cytoscape Core is an open source, plugins are separate softwares which might be protected under any license (Shannon *et al.*, 2003).

3.2.2 Osprey

Osprey is a software platform for visualization of complex interaction networks. Osprey builds data-rich graphical representations from Gene Ontology (GO) annotated interaction data maintained by BioGRID. Osprey visualizes genes/proteins by nodes and interactions between nodes by edges. The software embeds GO categories, experimental evidence and/or data source information in nodes and edges and also offers variety of graphical layouts. Osprey can be downloaded from http://biodata.mshri.on.ca/osprey (Stark *et al.*, 2006).

3.3 Reconstruction of Insulin Signaling Network

The core proteins that are known to have roles in insulin signaling network in *H. sapiens* were first identified by reviewing literature (Taniguchi *et al.*, 2006; Sedaghat *et al.*, 2002; Karlsson, 2007; Myers and White, 1993; Lizcano and Alessi, 2002; Lomberk and Urrutia, 2009; Holgado-Madruga *et al.*, 1996; Rosivatz, 2007; Cheatman and Kahn, 1995). Reviewed literature does not include all the proteins related to insulin signaling network

and all the protein-protein interaction information. Therefore, the Gene Ontology (http://www.geneontology.org/) database was also used to determine the core proteins of the network.

3.3.1 Reconstruction of PPI Network

Using the GO annotations of the core proteins in insulin metabolism, an annotation-collection table was formed to reconstruct the insulin signaling network. The annotation-collection table includes molecular function annotations, biological process annotations and cellular component annotations. The GO annotations of the proteins obtained from BioGRID-Organism-2.0.45 file for *H. sapiens* were compared to those in this annotation-collection table. Proteins were accepted to the network if all three of their GO annotations match with those of core proteins. It means that the chosen protein has a high probability of having characteristics in the insulin signaling. Otherwise, if one or more of the GO annotations do not match, the protein in question is not added to the network.

Some proteins out of core proteins do not have any interaction data either in databases scanned or in literature. To find the interactions of these proteins, sequence homology search was performed using standard protein-protein BLAST (blastp) algorithm which is available at http://blast.ncbi.nlm.nih.gov/Blast.cgi. Blastp is used for both identifying a query amino acid sequence and for finding similar sequences in protein databases and it is designed to find local regions of similarity. When sequence similarity spans the whole sequence, blastp will also report a global alignment, which is the preferred result for protein identification purposes.

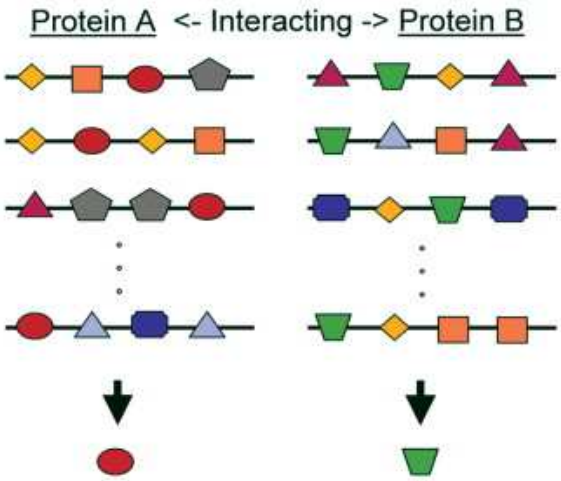
3.3.2 Reconstruction of DDI Network

Proteins typically contain two or more domains. About two-thirds of proteins in prokaryotes and four-fifths in eukaryotes are multidomain proteins (Apic *et al.*, 2001). Interaction between two proteins typically involves binding pairs and it is an important step towards understanding protein interactions and the evolution of protein-protein interaction networks. Many groups have contributed computational methods aimed at

discovering interacting domain pairs (Bock and Gough, 2001; Jothi *et al.*, 2006). They mostly rely on protein-protein interaction networks.

Many domain-domain interaction prediction methods tie the goal of predicting domain interactions to the seemingly related goal of predicting protein-protein interactions (Guimaraes *et al.*, 2006). For example, the Association method (Sprinzak and Margalit, 2001) scores each domain pair by the ratio of the number of occurrences of a given pair in interacting proteins to the number of independent occurrences of those domains. This score can be interpreted as the probability of interaction between the two domains. Several related methods have also been proposed. Deng and colleagues (Deng *et al.*, 2002) extended this idea further and applied a maximum likelihood estimation approach to define the probability of domain-domain interactions. Their expectation maximization algorithm computes domain interaction probabilities that maximize the expectation of observing a given protein-protein interaction network. Other groups proposed alternative methods for this task such as linear programming (Hayashida *et al.*, 2004), support vector machines (Bock and Gough, 2001), and probabilistic network modeling (Gomez and Rzhetsky, 2002; Guimaraes *et al.*, 2006).

In this work, an alternative method which is simpler than the Association method (Sprinzak and Margalit, 2001) was used to score each domain pair. Firstly, the domains of the proteins in the reconstructed pathway were extracted from the whole domain set of *H. sapiens* available at Superfamily 1.73 database. Separating the proteins into domains, protein-protein interactions were converted into domain-domain interactions. Each domain pair was scored by the appearance of a given pair in the whole domain-domain interactions, i.e., the ratio of the number of appearance of a given pair in interacting proteins to the number of all domain-domain pairs was calculated (Figure 3.2). Obtaining the probabilities of the domain interactions enables one to comment about the most probable domain interactions and consequently the most probable protein interactions and functions of the network. It is thus possible to predict interactions for protein pairs that have no interaction data in literature but have one or more domain pairs at high probability of interaction occurrence.



	1							
	3	0						
	2	1	0					
	3	2	**4**	0				
	2	2	2	2	1			
	2	0	1	1	2	0		
	1	1	1	2	0	0	0	
	2	1	2	0	0	0	1	0

Figure 3.2. A schematic representation of the analysis for detecting correlated sequence-signatures in interacting proteins (Sprinzak and Margalit, 2001)

3.4 Structural Analysis Techniques of the Reconstructed Network

The methods for reconstructed signaling network can be reviewed in two main sections, graph theoretic analysis and linear pathway analysis. Graph theoretic analysis investigates the topological properties of the reconstructed network. Linear pathway analysis decomposes the network into fundamental paths from ligands to phenotypes.

3.4.1 Graph Theoretic Analysis

The aim of interaction proteomics that studies the protein-protein interactions of all expressed proteins is to understand biological processes that are regulated by these interactions. The first step involves the mathematical representation of a protein-protein interaction (PPI) network as a graph, where nodes in the graph represent proteins and the edges that connect them correspond to interactions. The second step is to determine graph properties of the network using graph theoretic techniques, such as the degree or connection of nodes, the number and complexity of highly-connected nodes (hubs), the shortest path length for indirectly connected nodes and the key nodes (lethal proteins) in the network (Przulj *et al.*, 2004).

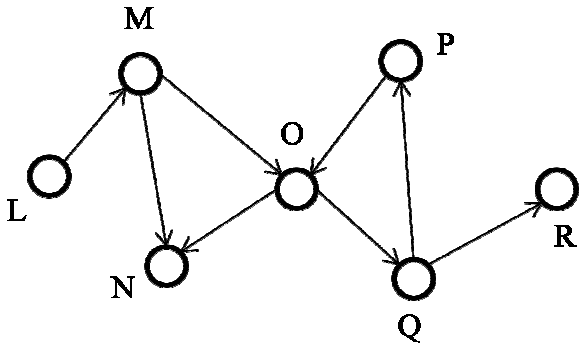
The most elementary characteristic of a node of a graph, k, can be characterized by the number of edges that they have (the number of other nodes to which they are adjacent). This property is called the node degree. In directed networks we distinguish the in-degree, k_{in}, the number of directed edges that point toward the node, and the out-degree, k_{out}, the number of directed edges that start at the node. Whereas node degrees characterize individual nodes, one can define a degree distribution to quantify the diversity of the whole network. The mean degree of a network, $< k >$, is defined as the average number of edges that a node has and can be calculated as

$$< k >= \frac{2L}{N} \tag{3.1}$$

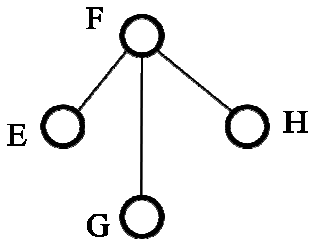
where L is edges and N is nodes (Albert, 2005).

Graph representation and graph analysis reveals regulatory patterns of cellular networks. The number of interactions a component participates in is quantified by its degree, for example node O has a connectivity of 2 (Figure 3.3). The graph distance between two nodes is defined as the number of edges in the shortest path between them. For example, the distance between nodes P and O is 1, and the distance between nodes O

and P is 2 (along the OQP path). The degree distribution $P(k)$ ($P(k_{in})$ and $P(k_{out})$ in directed nodes) quantifies the fraction of nodes with degree k (Albert, 2005).



k	$P(k_{in})$	$P(k_{out})$
0	1/7	2/7
1	4/7	2/7
2	2/7	3/7



k	$P(k)$
1	3/7
2	3/7
3	1/7

Figure 3.3. Graph representation and graph analysis (Albert, 2005)

Until recently, complex networks have been modeled using the classical random network theory (Erdös and Renyi, 1965; Bollobas, 1985). This model assumes that each pair of nodes (constituents) in the network is connected randomly with probability p, leading to a statistically homogeneous network in which, despite the fundamental randomness of the model, most nodes have the same number of links $< k >$ (Figure 3.4a). In particular, the connectivity follows a Poisson distribution that peaks strongly at $< k >$ (Figure 3.4b), implying that the probability of finding highly connected nodes decays exponentially ($P(k) \approx e^{-k}$ for $k >> < k >$). On the other hand, empirical studies on the structure of the World Wide Web (Albert *et al.*, 1999; Faloutsos *et al.*, 1999), social networks (Barabasi and Albert, 1999) and scientific collaboration network (Barabasi *et al.*, 2002) have reported serious deviations from this random structure, showing that these

systems are described by scale-free networks (Barabasi and Albert, 1999) (Figure 3.4c), for which $P(k)$ follows a power-law ($P(k) \approx k^{-\gamma}$) (Figure 3.4d). Here, the degree exponent γ is usually in the range $2 < \gamma < 3$ (Albert and Barabasi, 2002). Unlike exponential networks, scale-free networks are extremely heterogeneous; their topology being dominated by a few highly connected nodes (hubs), which link the rest of the less connected nodes to the system (Figure 3.4c).

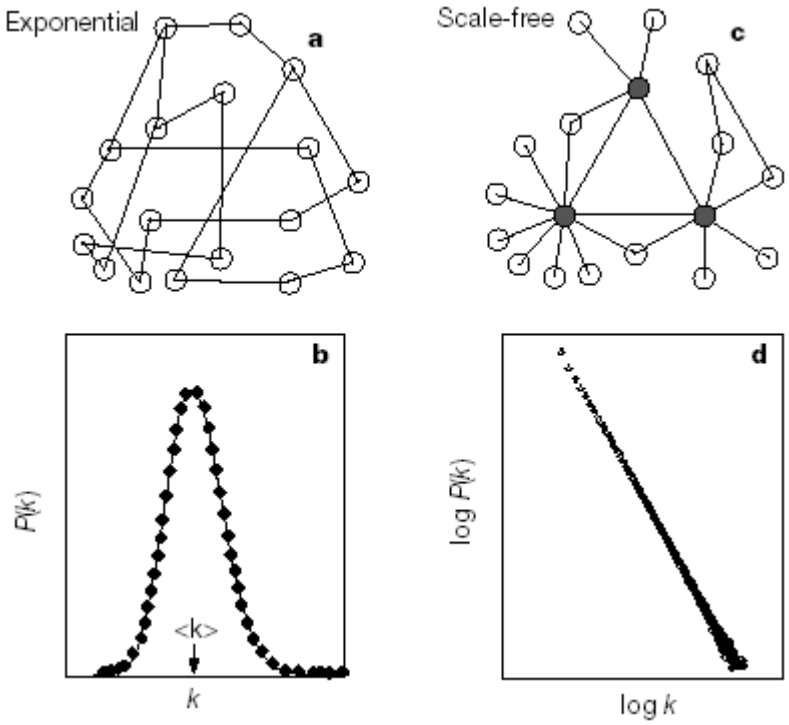


Figure 3.4. Attributes of generic network structures (Jeong *et al.*, 2000)

As the distinction between scale-free and exponential networks emerges as a result of simple dynamic principals (Amaral *et al.*, 2000; Dorogovtsev and Mendes, 2000), understanding the large-scale structure of cellular networks cannot only provide valuable and perhaps universal structural information, but could also lead to a better understanding of the dynamic processes that generated these networks. In this respect the emergence of power-law distribution is intimately linked to the growth of the network in which new nodes are preferentially attached to already established nodes (Barabasi and Albert, 1999), a property that is also thought to characterize the evolution of biological systems (Hartwell *et al.*, 1999).

In this work, the reconstructed insulin signaling network was converted into an undirected interaction graph where nodes are signaling molecules and edges are interactions in the signaling network. The interactions are undirected meaning that when there is a connection between molecules A and B, the signal can flow both from A to B and from B to A in the interaction graph. The graph properties were determined using an algorithm implemented in MATLAB 7.0. The input to the algorithm was the adjacency matrix (S), which is a binary, square matrix representing the edges between nodes (protein-protein interactions). $S(i,j)=1$, if (i,j) is an interacting protein pair, otherwise $S(i,j)=0$. Shortest path lengths were calculated using the fact that the element of the n th power of the adjacency matrix, $S^n(i,j)$, gives the number of paths with path length of n from protein i to protein j (Arga *et al.*, 2007). Shortest path lengths are minimum number of interactions between two indirectly connected nodes and were calculated as the minimum value of n when $S^n(i,j)$ is equal to a nonzero value (Durmuş Tekir *et al.*., 2009). For the insulin signaling network, the constructed adjacency matrix has the dimensions of 416×416 as there are 416 signaling molecules. Moreover, mean path length and network diameter which are the average and the longest shortest path between two nodes, respectively were determined.

One of the most important findings of large-scale proteomics experiments and bioinformatics analyses was the observation that 'hubs' participate in significant numbers of protein interactions and play critical roles in the organization and function of cellular protein interaction networks (Barabasi and Oltvai, 2004; Albert, 2005). It has also been demonstrated that such hub proteins represent very attractive subjects for understanding cellular functions and identifying novel drug targets (Hsing *et al.*, 2008). It is clear that hubs have more interactions than non-hubs, but it is not known exactly how many interactions a hub protein should have. In this work, the hub selection criterion was based on the position of a break point on a protein interaction distribution plot. These hubs tend to share certain relevant functional properties reflected in their Gene Ontology annotations.

3.4.2 Network Decomposition Analysis – Linear Pathway Analysis

Network decomposition analysis includes the decomposition of a network into fundamental paths through linear combination of which the time-invariant topological structure can be captured (Arga *et al.*, 2007). In this project, among the various decomposition techniques, linear path analysis was chosen for the structural investigation of the GLUT4 translocation of the insulin signaling network since the number of fundamental routes for small-world, large scale networks is very high and computationally unmanageable with other pathway analysis (Steffen *et al.*, 2002; Schuster *et al.*, 1996; Durmuş Tekir *et al.*, 2009). As opposed to elementary flux modes and extreme pathways, NetSearch, draws all possible linear paths through the interaction map starting at any membrane protein and ending on any target protein such as DNA-binding protein by specifying the maximum path length in the calculation procedure.

In order to investigate the connection of GLUT4 translocation subnetwork to the reconstructed insulin signaling network, linear paths beginning at membrane protein (INS, insulin protein) and ending at target protein (GLUT4, glucose transporter 4 protein) were determined using NetSearch algorithm (Steffen *et al.*, 2002). The interaction information, inputs (ligands) and outputs (phenotypes as cellular responses) were defined for the linear path calculations (Durmuş Tekir *et al.*, 2009). The final input parameter that required evaluation was the maximum path length allowable for NetSearch paths (Steffen *et al.*, 2002).

3.5 Dynamic Model of Metabolic Insulin Signaling Network

The dynamic model of insulin metabolism (Sedaghat *et al.*, 2002) explained in Section 2.2.3.3 was investigated to enlighten GLUT4 translocation pathway of the insulin signaling network. There are four subsystems in the pathway.

(i) Firstly, in the insulin receptor binding subsystem, on binding the first molecule of insulin, the receptor is rapidly phosphorylated, resulting in receptors that may either bind another molecule of insulin or dissociate from the first molecule of insulin. Binding of a

second molecule of insulin does not affect the phosphorylation state of the receptor, whereas receptor dephosphorylation occurs when insulin diffuses off the receptor, leaving a free receptor. In addition, protein tyrosine phosphatases that dephosphorylate the insulin receptor are explicitly represented as a multiplicative factor $\left(\left[PTP\right]\right)$ that modulates receptor dephosphorylation rate. The differential equations for this subsystem are given in Table 3.1.

Table 3.1. Ordinary differential equations for insulin receptor binding subsystem

x_1 = insulin input

$$\frac{dx_2}{dt} = k_{-1}x_3 + k_{-3}\left[PTP\right]x_5 - k_1x_1x_2 + k_{-4}x_6 - k_4x_2$$

$$\frac{dx_3}{dt} = k_1x_1x_2 - k_{-1}x_3 - k_3x_3$$

$$\frac{dx_4}{dt} = k_2x_1x_5 - k_{-2}x_4 + k_{-4'}x_7 - k_{4'}x_4$$

$$\frac{dx_5}{dt} = k_3x_3 + k_{-2}x_4 - k_2x_1x_5 - k_{-3}\left[PTP\right]x_5 + k_{-4'}x_8 - k_{4'}x_5$$

The parameters are denoted as follows: x_1, free insulin input (system input); x_2, concentration of unbound surface insulin receptors; x_3, concentration of unphosphorylated once-bound surface receptors; x_4, concentration of phosphorylated twice-bound surface receptors; x_5, concentration of phoshorylated once-bound surface receptors; k_1 and k_{-1}, association and dissociation rate constants, respectively, for the first molecule of insulin to bind the receptor; k_2 and k_{-2}, association and dissociation rate constants, respectively, for the second molecule of insulin to bind the receptor; k_3, rate constant for receptor autophosphorylation; k_4, endocytosis rate constant for free receptors; k_{-4}, exocytosis rate constant, $k_{4'}$, endocytosis rate constant for bound receptors; $k_{-4'}$, exocytosis rate for twice-bound and oncebound intracellular phosphorylated receptors (Sedaghat *et al.*, 2002).

(ii) Insulin receptor recycling subsystem represents synthesis, degradation, exocytosis, and both basal and ligand-induced endocytosis of receptors. An additional step representing dephosphorylation of internalized phosphorylated receptors and their incorporation into the intracellular pool is included into the model. Thus, the related differential equations are given in Table 3.2.

Table 3.2. Ordinary differential equations for insulin receptor recycling subsystem

$$\frac{dx_6}{dt} = k_5 - k_{-5}x_6 + k_6\left[PTP\right]\left(x_7 + x_8\right) + k_4 x_2 - k_{-4}x_6$$
$$\frac{dx_7}{dt} = k_{4'}x_4 - k_{-4'}x_7 - k_6\left[PTP\right]x_7$$
$$\frac{dx_8}{dt} = k_{4'}x_5 - k_{-4'}x_8 - k_6\left[PTP\right]x_8$$

The parameters are denoted as follows: x_6, concentration of unbound unphosphorylated intracellular receptors; x_7, concentration of phosphorylated twice-bound intracellular receptors; x_8, concentration of phoshorylated once-bound intracellular receptors; k_5, zero order rate constant for receptor synthesis; k_{-5}, constant for receptor degradation, k_6, dephosphorylation rate constant for intracellular receptors that is modulated by the multiplicative factor [PTP].

(iii) Postreceptor signaling subsystem comprises the elements of the metabolic insulin signaling pathway that are well established (Figure 3.5). It is assumed that this is a closed subsystem so that synthesis and degradation of signaling molecules are not explicitly represented. Differential equations governing phosphorylation of IRS-1 and subsequent formation of phosphorylated IRS-1/activated PI3K, describing the interconversion between these phosphatidylinositides and the activation of downstream kinases of Akt and PKC-ζ is mathematically explained in Table 3.3 (Sedaghat *et al.*, 2002).

Table 3.3. Ordinary differential equations for postreceptor signaling subsystem

$$\frac{dx_9}{dt} = k_{-7}\left[PTP\right]x_{10} - k_7 x_9\left(x_4 + x_5\right)/IRp$$

$$\frac{dx_{10}}{dt} = k_7 x_9\left(x_4 + x_5\right)/IRp + k_{-8}x_{12} - \left(k_{-7}\left[PTP\right] + k_8 x_{11}\right)x_{10}$$

$$\frac{dx_{11}}{dt} = k_{-8}x_{12} - k_8 x_{10}x_{11}$$

$$\frac{dx_{12}}{dt} = k_8 x_{10}x_{11} - k_{-8}x_{12}$$

$$\frac{dx_{13}}{dt} = k_9 x_{14} + k_{10}x_{15} - \left(k_{-9}\left[PTEN\right] + k_{-10}\left[SHIP\right]\right)x_{13}$$

$$\frac{dx_{14}}{dt} = k_{-9}\left[PTEN\right]x_{13} - k_9 x_{14}$$

$$\frac{dx_{15}}{dt} = k_{-10}\left[SHIP\right]x_{13} - k_{10}x_{15}$$

$$\frac{dx_{16}}{dt} = k_{-11}x_{17} - k_{11}x_{16}$$

$$\frac{dx_{17}}{dt} = k_{11}x_{16} - k_{-11}x_{17}$$

$$\frac{dx_{18}}{dt} = k_{-12}x_{19} - k_{12}x_{18}$$

$$\frac{dx_{19}}{dt} = k_{12}x_{18} - k_{-12}x_{19}$$

The parameters are denoted as follows: x_9, concentration of unphosphorylated IRS-1; x_{10}, concentration of tyrosine-phosphorylated IRS-1; x_{11}, concentration of unactivated PI-3-kinase; x_{12}, concentration of tyrosine-phoshorylated IRS-1/activated PI3K complex, x_{13}, percentage of PI(3,4,5)P3 out of the total lipid population, x_{14}, percentage of PI(4,5)P2 out of the total lipid population, x_{15}, percentage of PI(3,4)P2 out of the total lipid population, x_{16}, percentage of unactivated Akt, x_{17}, percentage of activated Akt, x_{18}, percentage of unactivated PKC-ζ, x_{19}, percentage of activated PKC-ζ.

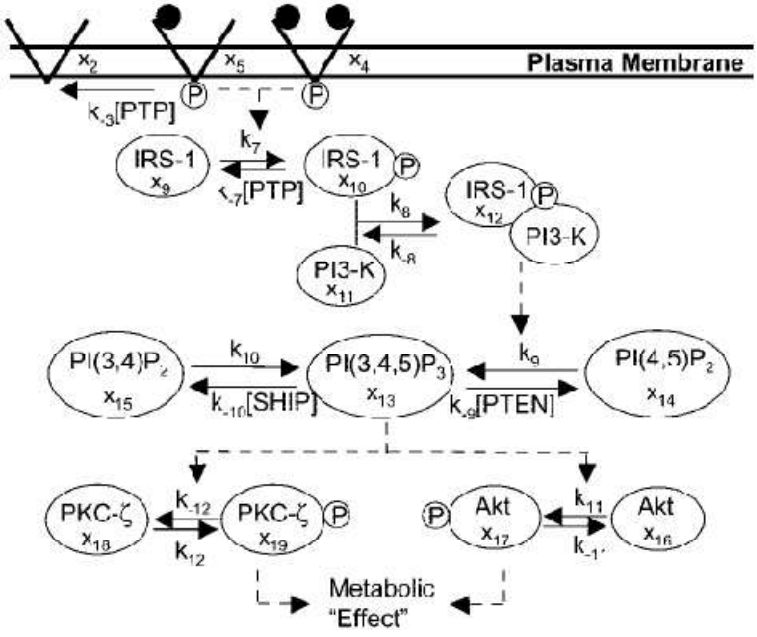


Figure 3.5. Representation of postreceptor signaling subsystem (Sedaghat *et al.*, 2002)

The output of this subsystem is represented as metabolic 'Effect' due to Akt and PKC-ζ activity, with 80% of the metabolic insulin signaling effect attributed to PKC-ζ and 20% of the effect attributed to Akt

$$Effect = \left(0.2x_{17} + 0.8x_{19}\right) / AP_{equil} \tag{3.2}$$

where AP_{equil} is the steady-state level of combined activity for Akt and PKC-ζ after maximal insulin stimulation (normalized to 100%).

(iv) The final subsystem is GLUT4 translocation (Figure 3.6). GLUT4 recycles between an intracellular compartment and the cell surface. To represent the aspect of GLUT4 trafficking, the insulin-stimulated exocytosis rate ($k_{13'}$) is increased to its maximum value as a linear function of the metabolic effect produced by phosphorylated Akt and PKC-ζ (Table 3.4). By assuming that the basal equilibrium distribution of 4% cell surface GLUT4 and 96% GLUT4 in the intracellular pool transitions on maximal insulin stimulation to a new steady state of 40% cell surface GLUT4 and 60% intracellular GLUT4, the equations governing changes in k_{13} and $k_{13'}$ are given in Table 3.5. Thus,

changes in $k_{13'}$ are linearly dependent on the output of the signaling subsystem (Sedaghat *et al.*, 2002).

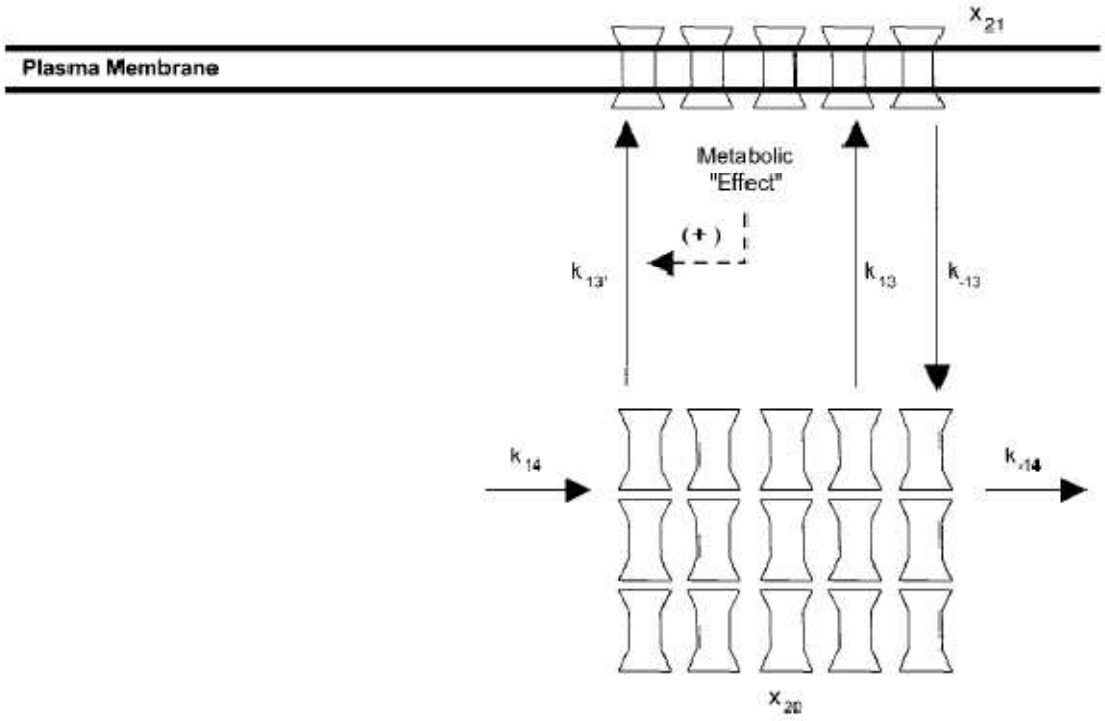


Figure 3.6. Representation of GLUT4 translocation (Sedaghat *et al.*, 2002)

Table 3.4. Ordinary differential equations of GLUT4 translocation

$$\frac{dx_{20}}{dt} = k_{-13}x_{21} - \left(k_{13} + k_{13'}\right)x_{20} + k_{14} - k_{-14}x_{20}$$
$$\frac{dx_{21}}{dt} = \left(k_{13} + k_{13'}\right)x_{20} - k_{-13}x_{21}$$

Table 3.5. The equations for k_{13} and $k_{13'}$

$$k_{13} = \left(4/96\right)k_{-13}$$
$$k_{13'} = \left[\left(40/60\right) - \left(4/96\right)\right]k_{-13}\left(Effect\right)$$

The parameters are as follows: x_{20}, percentage of intracellular GLUT4; x_{21}, percentage of cell surface GLUT4; k_{-13}, rate constant for GLUT4 internalization; k_{13}, rate

constant for translocation of GLUT4 to the cell surface under basal conditions; $k_{13'}$, rate constant for translocation of GLUT4; k_{14} (zero order) and k_{-14}, rate constants for GLUT4 synthesis and degradation, respectively. The metabolic 'Effect' from postreceptor signaling subsystem increases $k_{13'}$.

With four subsystems, this dynamic model was evaluated by generating time courses for all state variables in response to a maximally stimulating step input of 10^{-7} M insulin that was turned off after 15 min. A MATLAB 7.5 (R2007b) code was written for this purpose.

3.6 Analyses of the Dynamic Insulin Signaling Network

In the field of systems biology, much focus has been placed on the development of quantitative tools for analysis of metabolism in living cells, as this part of cellular operation is extremely well annotated (Covert *et al.*, 2001; Palsson, 2000). Despite many studies on the structure of metabolic networks (Klamt and Stelling, 2003; Papin *et al.*, 2004; Schilling *et al.*, 2000), on the operation of these networks (Fernie *et al.*, 2005; Sauer 2004), on correlations between metabolic network operation and transcriptome data (Patil and Nielsen, 2005) and on ways to score modifications of the metabolic network in silico, it is still interesting to identify methods that provide further insight into the operation of metabolic networks.

In this study, the methods used for the analysis of the metabolic signaling network can be reviewed under two main sections: metabolic control analysis and metabolic pathway analysis by elementary flux modes. Metabolic control analysis ranks the genes or proteins or metabolites according to their importance in controlling and regulating cellular metabolic networks. Elementary flux mode and control effective flux analyses allow finding meaningful routes in metabolic networks.

3.6.1 Metabolic Control Analysis

Mathematical modeling allows us to examine an event, process or system that we are unable to observe or understand directly because of its timing, magnitude, location or complexity. Models enable us to view a process or system at different organizational levels (for example, molecular or organismal) 'simultaneously', and to test responses of the system and its components to perturbations. Even incomplete or limited models can pinpoint missing or incorrect pathways or components and can help to ascertain the relative importance of pathways and components in different scenarios (Gershenfeld, 1999). In order to study cellular metabolism and signaling, various modeling approaches have been used. For the analysis of metabolism, such approaches include kinetic simulation, metabolic control analysis, biochemical systems theory, metabolic pathway analysis, and network analysis. Signal transduction pathways and networks have, for the most part, been described qualitatively by the sets of expressed genes associated with the activation of a specific pathway (Wildermuth, 2000).

Metabolic control analysis, a method developed in early seventies by Kacser and Burns (Kacser and Burns, 1973) and Heinrich and Rapoport (Heinrich and Rapoport, 1974), provides a quantitative description of substrate flux and concentrations in response to changes in system parameters of complex enzyme systems. Medical applications of this approach include the following: understanding the threshold effect in the manifestation of metabolic diseases; investigating the gene dose effect of aneuploidy in inducing phenotypic transformation in cancer; correlating the contributions of individual genes and phenotypic characteristics in metabolic disease (e.g., diabetes); identifying candidate enzymes in pathways suitable as targets for cancer therapy; and elucidating the function 'silent' genes by identifying metabolic features shared with genes of known pathways. MCA complements current studies of genomics and proteomics, providing a link between biochemistry and functional genomics and proteomics, providing a link between biochemistry and functional genomics that relates the expression of genes and gene products to cellular biochemical and physiological events. Thus, it is an important tool for the study of genotype – phenotype correlations. It allows the enzymes/metabolites to be ranked according to their importance in controlling and regulating cellular metabolic networks (Cascante, 2002).

A general formalism for Metabolic Control Analysis is derived using general sensitivity analysis and structural information of the metabolic pathway inherent in the stoichiometry matrix. Three types of coefficients are defined; flux control coefficients, concentration control coefficients, and elasticity coefficients (Ehlde and Zacchi, 1996). The flux control coefficients are defined as

$$C_i^{J_j} = \frac{dJ_j}{de_i} \cdot \frac{e_i}{J_j} = \frac{d \ln J_j}{d \ln e_i} \tag{3.3}$$

where J_i is the steady state flux through enzyme E_j and e_i is the concentration of enzyme E_i. The value of a flux control coefficient is a measure of how a change in the concentration of enzyme E_i affects the steady-state flux through enzyme E_j, i.e. the degree of control exerted by enzyme E_i on this steady-state flux (Ehlde and Zacchi, 1996). It is sometimes necessary to assign a metabolic flux a negative value, for example if it is a reversible pathway flowing in the opposite direction to that originally defined. In this case, Eq. 3.3. cannot be used as it is because it is not possible to have the logarithm of a negative number (Negative logarithms represent numbers between 0 and 1). This turns out not to matter because the following modified version works just as well:

$$C_i^{J_j} = \frac{dJ_j}{de_i} \cdot \frac{e_i}{J_j} = \frac{d \left| \ln J_j \right|}{d \ln e_i} \tag{3.4}$$

where $\left| \ln J_j \right|$ signifies that the absolute value of J_j is used, i.e. any negative sign is ignored (Fell, 1997).

The concentration control coefficients are similarly defined as

$$C_i^{X_j} = \frac{dx_j}{de_i} \cdot \frac{e_i}{x_j} = \frac{d \ln x_j}{d \ln e_i} \tag{3.5}$$

where x_j is the steady-state concentration of metabolite X_j. The value of the concentration of control coefficient is a measure of the degree of control exerted by enzyme E_i on the steady-state concentration x_j (Ehlde and Zacchi, 1996).

The third type of coefficients, elasticity coefficients, are defined as

$$\varepsilon_j^i = \frac{\partial \nu_i}{\partial x_j} \cdot \frac{x_j}{\nu_i} = \frac{\partial \ln \nu_i}{\partial \ln x_j} \tag{3.6}$$

where ν_i is the reaction rate of enzyme E_i and x_j is the concentration of metabolite X_j. An elasticity coefficient is a measure of how the reaction rate ν_i will respond to variations in the concentration x_j when all other concentrations are kept constant (Ehlde and Zacchi, 1996).

Both types of control coefficients are global properties in the sense that they reflect the state of the whole pathway. Making a change in a parameter of the pathway (e.g. a K_m value, an enzyme concentration, etc.) will affect all control coefficients of the pathway (Ehlde and Zacchi, 1996). An elasticity coefficient, on the other hand, represents the fractional response of the rate of a step to a fractional change in concentration of a metabolite, effector, or enzyme, and for enzymes is directly derived from the rate laws.

In the early pages of MCA, two types of relationships were derived, called summation theorems and connectivity theorems. The flux summation theorem states that the flux control coefficients of a metabolic pathway always add up to unity and the concentration summation theorem states that the concentration control coefficients add up to zero. The flux summation theorem indicates that the term ‘rate-limiting step’ implies that the flux control coefficient for this step would have a value of 1, while all other flux control coefficients would be 0. This is a very unlikely situation, especially in a complex pathway. The term ‘rate-limiting step’ should therefore be replaced by quantitative measures of the controlling capability of each reaction step, given by the flux control coefficients. The connectivity theorems relate the elasticity coefficients to the control

coefficients and together with the summation theorems they form a linear equation system that makes it possible to calculate the global properties, the control coefficients, from the local properties, the elasticity coefficients (Ehlde and Zacchi, 1996).

In this thesis, a different approach is used in deriving a particular formalism for MCA. Consider an arbitrary metabolic pathway containing *m* internal metabolites and *n* reactions. The stoichiometry of the pathway can be conveniently described by $m \times n$ matrix $\overline{\overline{N}}$, called the stoichiometry matrix. Each row in $\overline{\overline{N}}$ corresponds to an internal metabolite (concentration) and each column corresponds to a reaction (rate). The elements, n_{ij}, of $\overline{\overline{N}}$ are defined as follows (Reder, 1988):

$n_{ij} = +z$ if the reaction j produces z molecules of metabolite i

$n_{ij} = -z$ if the reaction j consumes z molecules of metabolite i

$n_{ij} = 0$ if the reaction j neither produces nor consumes metabolite i

The structural relationships of the pathway can be readily deduced from the stoichiometry matrix. Suppose that $\overline{\overline{N}}$ has the rank m_0, where m_0 is less than or equal to the number of rows, m. This means that $\overline{\overline{N}}$ contains m_0 independent rows and the remaining $m - m_0$ dependent rows can be expressed as linear combinations of the independent rows. Since each row in $\overline{\overline{N}}$ corresponds to a metabolite, this also means that if m_0 is less than m, the pathway will contain dependent metabolites which can be expressed in terms of the independent metabolites. There may be several combinations of metabolites (rows in $\overline{\overline{N}}$) that can be chosen as independent and the choice is arbitrary. To simplify the derivation below, the metabolites are renumbered so that the independent rows are the first m_0 rows of $\overline{\overline{N}}$ (Ehlde and Zacchi, 1996).

Since the column of a rank of a matrix is always equal to its row rank, the matrix $\overline{\overline{N}}$ also contains m_0 independent columns. The columns are renumbered so that the m_0 independent columns are the last columns of $\overline{\overline{N}}$. The number of dependent columns, equal

to $n - m_0$, is denoted n_0 first columns of $\overline{\overline{N}}$. The dependent columns are thus the n_0 first columns of $\overline{\overline{N}}$ (Ehlde and Zacchi, 1996).

Three new matrices are constructed from $\overline{\overline{N}}$ for use in the derivations below. The first consists of the m_0 independent (last) rows of $\overline{\overline{N}}$ and is denoted $\overline{\overline{N_R}}$. The second consists of the m_0 independent (last) columns of $\overline{\overline{N}}$ and is denoted $\overline{\overline{N_C}}$. The third consists of the m_0 independent rows and the m_0 independent columns of $\overline{\overline{N}}$ and is denoted $\overline{\overline{N_{RC}}}$.

$$\overline{\overline{N}} = \begin{pmatrix} n_{11} & \cdots & n_{1n} \\ \cdots & \cdots & \cdots \\ n_{m1} & \cdots & n_{mn} \end{pmatrix}, \qquad \overline{\overline{N_R}} = \begin{pmatrix} n_{11} & \cdots & n_{1n} \\ \cdots & \cdots & \cdots \\ n_{m_0 1} & \cdots & n_{m_0 n} \end{pmatrix}$$

$$\overline{\overline{N_C}} = \begin{pmatrix} n_{1(n_0+1)} & \cdots & n_{1n} \\ \cdots & \cdots & \cdots \\ n_{m(n_0+1)} & \cdots & n_{mn} \end{pmatrix}, \qquad \overline{\overline{N_{RC}}} = \begin{pmatrix} n_{1(n_0+1)} & \cdots & n_{1n} \\ \cdots & \cdots & \cdots \\ n_{m_0(n_0+1)} & \cdots & n_{m_0 n} \end{pmatrix} \tag{3.7}$$

The matrix $\overline{\overline{N_{RC}}}$ is a square and always invertible, since all rows (and columns) are independent (Ehlde and Zacchi, 1996).

3.6.1.1 Relationships Between Metabolites: The m metabolites of the arbitrary pathway are collected in an $m \times 1$ vector denoted $\bar{x}$ and the m_0 independent metabolites are collected in an $m_0 \times 1$ vector, denoted $\overline{x_R}$. The aim is now to express all m metabolites of the pathway in terms of the m_0 independent metabolites.

The stoichiometry matrix $\overline{\overline{N}}$ can be decomposed as (Reder, 1988)

$$\overline{\overline{N}} = \overline{\overline{L^x}} \cdot \overline{\overline{N_R}} \tag{3.8}$$

The $m \times m_0$ matrix $\overline{\overline{L^x}}$ is called the concentration link matrix. If the dependent columns of $\overline{\overline{N}}$ and $\overline{\overline{N_R}}$ are deleted in Eq. 3.8, $\overline{\overline{L^x}}$ can be calculated as

$$\overline{\overline{L^x}} = \overline{\overline{N_C}} \cdot \overline{\overline{N_{RC}^{-1}}} \tag{3.9}$$

Note that if $\overline{\overline{N}}$ has full rank ($m_0 = m$), $\overline{\overline{N_R}}$ will be equal to $\overline{\overline{N}}$ and $\overline{\overline{L^x}}$ will be the $m \times m$ identity matrix.

It can be shown that

$$\frac{d\overline{x}}{dt} = \overline{\overline{L^x}} \cdot \frac{d\overline{x_R}}{dt} \tag{3.10}$$

which can be integrated to give

$$\overline{x(t)} = \overline{\overline{L^x}} \cdot \overline{x_R(t)} + \overline{\overline{A}} \tag{3.11}$$

where $\overline{\overline{A}}$ is an $m \times 1$ vector where the first m_0 elements are zero. The equation specifies the conservation relationships of the metabolites, i.e. how the dependent metabolites can be expressed in terms of the independent ones (Ehlde and Zacchi, 1996).

3.6.1.2 Relationships Between Steady-State Fluxes: The steady-state fluxes of the reactions in a metabolic pathway are related by structural constraints, i.e. not all steady-state fluxes are independent. For instance, in a linear pathway the steady-state fluxes through each reaction are equal and only one can be chosen as independent.

A general relation between the independent and the dependent steady-state fluxes of a metabolic pathway can be derived using the matrices in Eq. (3.7) and a new matrix, $\overline{\overline{N_0}}$, containing the m_0 independent (first) rows and the n_0 dependent (first) columns of $\overline{\overline{N}}$:

$$\overline{\overline{N_0}} = \begin{pmatrix} n_{11} & \cdots & n_{1n_0} \\ \cdots & \cdots & \cdots \\ n_{m_0 1} & \cdots & n_{m_0 n_0} \end{pmatrix} \tag{3.12}$$

The n_0 dependent columns will correspond to the independent fluxes of the pathway. If the steady-state flux through reaction k is denoted as J_k, and all steady-state fluxes are collected in the $n \times 1$ vector $\overline{J}$, the following structural relationship is valid for any pathway:

$$\overline{\overline{N_R}} \cdot \overline{J} = 0 \tag{3.13}$$

This equation can be rearranged in order to obtain the m_0 dependent fluxes, collected in the $m_0 \times 1$ vector $\overline{J_0}$, as functions of the n_0 independent fluxes, collected in the $n_0 \times 1$ vector $\overline{J_R}$:

$$\overline{\overline{N_0}} \cdot \overline{J_R} + \overline{\overline{N_{RC}}} \cdot \overline{J_0} = 0 \tag{3.14}$$

or

$$\overline{J_0} = -\overline{\overline{N_{RC}^{-1}}} \cdot \overline{\overline{N_0}} \cdot \overline{J_R} \tag{3.15}$$

The product $-\overline{\overline{N_{RC}^{-1}}} \cdot \overline{\overline{N_0}}$ will have the dimensions of $m_0 \times n_0$. A new matrix, $\overline{\overline{L^J}}$, is constructed from this product and an n_0 identity matrix as:

$$\overline{\overline{L^J}} = \begin{pmatrix} \overline{\overline{I_{n_0}}} \\ -\overline{\overline{N_{RC}^{-1}}} \cdot \overline{\overline{N_0}} \end{pmatrix} \tag{3.16}$$

The matrix, $\overline{\overline{L^J}}$, which has the dimensions $n \times n_0$, is called the flux link matrix. All n steady-state fluxes, $\overline{J}$, can now be expressed as functions of the n_0 independent steady-state fluxes, $\overline{J_R}$, by

$$\overline{J} = \overline{\overline{L^J}} \cdot \overline{J_R} \tag{3.17}$$

This equation is the basis for metabolic flux analysis, which provides a method of determining all steady-state fluxes in a metabolic network by measuring some external key fluxes.

The concentrations of the m internal metabolites in an arbitrary pathway are denoted $x_1, x_2, \ldots, x_m$ and the n rates of the reactions are denoted $v_1, v_2, \ldots, v_n$. The concentrations and the reaction rates are collected in the $m \times 1$ vector $\overline{x}$ and in the $n \times 1$ vector $\overline{v}$, respectively.

$$\overline{x} = \begin{pmatrix} x_1 & x_2 & \ldots & x_m \end{pmatrix}^T \tag{3.18}$$

$$\overline{v} = \begin{pmatrix} v_1 & v_2 & \ldots & v_m \end{pmatrix}^T \tag{3.19}$$

The pathway also contains r parameters, with values $p_1, p_2, \ldots, p_r$. These parameters, that can be enzyme concentrations, kinetic parameters, external metabolites, etc., are collected in the $r \times 1$ vector $\overline{p}$:

$$\overline{p} = \begin{pmatrix} p_1 & p_2 & \ldots & p_r \end{pmatrix}^T \tag{3.20}$$

Each function rate can be generally expressed as a function of the m concentrations, and of the r parameters, i.e.

$$v_k = f_k\left(\overline{p}, \overline{x}\right) \qquad k = 1, 2, \ldots, n \tag{3.21}$$

If one of these arbitrary functions is differentiated with respect to one enzyme concentration, e_j, the result is

$$\frac{dv_k}{de_j} = \frac{\partial v_k}{\partial p_1} \cdot \frac{dp_1}{de_j} + \ldots + \frac{\partial v_k}{\partial p_r} \cdot \frac{dp_r}{de_j} + \frac{\partial v_k}{\partial x_1} \cdot \frac{dx_1}{de_j} + \ldots + \frac{\partial v_k}{\partial x_m} \cdot \frac{dx_m}{de_j} \tag{3.22}$$

The enzyme concentration, e_j, is present in the parameter vector $\overline{p}$ and if the parameters are not in any way coupled, i.e. if a change in one parameter does not affect the others, it is noted that

$$\frac{dp_i}{de_j} = \begin{cases} 0, & p_i \neq e_j \\ 1, & p_i = e_j \end{cases} \tag{3.23}$$

If the zero terms of Eq. 3.22 are eliminated and the equation is multiplied by e_j / v_k, it can be arranged to give:

$$\frac{dv_k}{de_j} \cdot \frac{e_j}{v_k} = \frac{\partial v_k}{\partial e_j} \cdot \frac{e_j}{v_k} + \sum_{i=1}^{m} \left(\frac{\partial v_k}{\partial x_i} \cdot \frac{x_i}{v_k} \cdot \frac{dx_i}{de_j} \cdot \frac{e_j}{x_i} \right) \tag{3.24}$$

If v_k is chosen to be the steady-state flux through reaction k, i.e. $v_k = \left(v_k\right)_{ss} = J_k$, the elasticity coefficients, ε_i^k, the parameter elasticity coefficients, π_j^k, the flux control coefficients, $C_j^{J_k}$, and the concentration control coefficients, $C_j^{X_i}$, can be recognized in Eq. 3.24 as

$$\varepsilon_i^k = \frac{\partial v_k}{\partial x_i} \cdot \frac{x_i}{v_k}, \qquad \pi_j^k = \frac{\partial v_k}{\partial e_j} \cdot \frac{e_j}{v_k}$$

$$C_j^{J_k} = \frac{dJ_k}{de_j} \cdot \frac{e_j}{J_k}, \qquad C_j^{X_i} = \frac{dx_i}{de_j} \cdot \frac{e_j}{x_i} \tag{3.25}$$

respectively. The parameter elasticity coefficient π_j^k is a measure of the response in the enzymatic reaction rate to variations in an enzyme concentration (Kacser *et al.*, 1990). In traditional MCA these coefficients are assumed to be 0 when $k \neq j$ and 1 when $k = j$. Making the substitutions given by Eq. 3.25, Eq. 3.24 can be written as

$$C_j^{J_k} = \pi_j^k + \sum_{i=1}^{m} \varepsilon_i^k \cdot C_j^{X_i} \tag{3.26}$$

In vector form, this is equal to

$$C_j^{J_k} = \pi_j^k + \begin{pmatrix} \varepsilon_1^k & \varepsilon_2^k & \dots & \varepsilon_m^k \end{pmatrix} \cdot \begin{pmatrix} C_j^{X_1} \\ C_j^{X_2} \\ \vdots \\ C_j^{X_m} \end{pmatrix} \tag{3.27}$$

Extension to all enzymes in the pathway yields:

$$\begin{pmatrix} C_1^{J_k} & \dots & C_n^{J_k} \end{pmatrix} = \begin{pmatrix} \pi_1^k & \dots & \pi_n^k \end{pmatrix} + \begin{pmatrix} \varepsilon_1^k & \dots & \varepsilon_m^k \end{pmatrix} \cdot \begin{pmatrix} C_j^{X_1} \\ C_j^{X_2} \\ \vdots \\ C_j^{X_m} \end{pmatrix} \tag{3.28}$$

Further extension to all steady-state enzyme fluxes yields:

$$\begin{pmatrix} C_1^{J_1} & \dots & C_n^{J_1} \\ \vdots & & \vdots \\ C_1^{J_n} & \dots & C_n^{J_n} \end{pmatrix} = \begin{pmatrix} \pi_1^1 & \dots & \pi_n^1 \\ \vdots & & \vdots \\ \pi_1^n & \cdots & \pi_n^n \end{pmatrix} + \begin{pmatrix} \varepsilon_1^1 & \dots & \varepsilon_m^1 \\ \vdots & & \vdots \\ \varepsilon_1^n & \dots & \varepsilon_n^n \end{pmatrix} \cdot \begin{pmatrix} C_1^{X_1} & \dots & C_n^{X_1} \\ \vdots & & \vdots \\ C_1^{X_m} & \dots & C_n^{X_m} \end{pmatrix} \tag{3.29}$$

If the matrices in Eq. 3.29 are named it can be written as

$$\overline{\overline{C^J}} = \overline{\overline{\pi}} + \overline{\overline{\varepsilon}} \cdot \overline{\overline{C^X}} \tag{3.30}$$

If it is assumed that each reaction rate is directly proportional to the concentration of the corresponding enzyme and that the concentration only affects its own reaction rate, as is assumed in traditional MCA, the matrix $\overline{\overline{\pi}}$ becomes an $n \times n$ identity matrix.

Since the aim is to solve for the control coefficients in terms of the elasticity coefficients, Eq. 3.30 is not sufficient. It is a system of equations with $n \times n + m \times n$ unknowns (the control coefficients) and only $n \times n$ equations. More relations must thus be utilized in order to solve it. The relations between the independent and the dependent metabolites were shown to be described by Eq. 3.11. Row k of this matrix equation can be written as

$$x_k = \sum_{i=1}^{m_0} I_{ki} \cdot x_i + a_k \tag{3.31}$$

where I_{ki} is an element of the concentration link matrix L^X and a_k is an element of the vector $\overline{\overline{A}}$. If x_k and x_i denote steady-state concentrations this equation can be differentiated with respect to an enzyme concentration, e_j. Multiplying the result of this differentiation by e_j / x_k yields

$$\frac{dx_k}{de_j} \cdot \frac{e_j}{x_k} = \frac{e_j}{x_k} \cdot \sum_{i=1}^{m_0} I_{ki} \cdot \frac{dx_i}{de_j} \tag{3.32}$$

Rearrangement and recognition of terms yields

$$C_j^{X_k} = \sum_{i=1}^{m_0} I_{ki} \cdot \frac{x_i}{x_k} \cdot C_j^{X_i} \tag{3.33}$$

If a new matrix, denoted $\overline{\overline{L_F^X}}$, is constructed from $\overline{\overline{L^X}}$ by multiplying each element, I_{ki}, of $\overline{\overline{L^X}}$ by x_i / x_k, and if Eq. 3.33 is extended to all metabolites and enzymes of the pathway, the resulting equation in matrix notation reads:

$$\overline{\overline{C^X}} = \overline{\overline{L_F^X}} . \overline{\overline{C_R^X}} \tag{3.34}$$

The $m \times n$ matrix $\overline{\overline{C^X}}$ contains all concentration control coefficients of the pathway and the $m_0 \times n$ matrix $\overline{\overline{C_R^X}}$ contains the concentration control coefficients for the independent metabolites. If the concentration control coefficients for the independent metabolites have been calculated, the concentration control coefficients for the remaining metabolites can be calculated using Eq. 3.34.

The relation between the independent and the dependent steady-state fluxes was described by Eq. 3.17. This equation can be differentiated as

$$d\overline{J} = \overline{\overline{L^J}} \cdot d\overline{J_R} \tag{3.35}$$

Exactly the same derivation can then be made as when Eq. 3.34 was derived from Eq. 3.31. The result is

$$\overline{\overline{C^J}} = \overline{\overline{L_F^J}}.\overline{\overline{C_R^J}} \tag{3.36}$$

$\overline{\overline{C^J}}$ contains all flux control coefficients and $\overline{\overline{C_R^J}}$ contains the flux control coefficients for the independent steady-state fluxes. $\overline{\overline{L_F^J}}$ is the result when each element, I_{ki}, of $\overline{\overline{L^J}}$ is multiplied by J_i / J_k.

Eqs. 3.34 and 3.36 can be inserted into Eq. 3.30 and the result is

$$\overline{\overline{L_F^J}} \cdot \overline{\overline{C_R^J}} - \overline{\overline{\varepsilon}} \cdot \overline{\overline{L_F^X}} \cdot \overline{\overline{C_R^X}} = \overline{\overline{\pi}} \tag{3.37}$$

The number of unknowns has now been reduced to $n_0 \times n + m_0 \times n = n \times n$ and it is possible to solve the equation system.

Eq. 3.37 can be rearranged as

$$\begin{pmatrix} \overline{\overline{L_F^J}} \\ \overline{\overline{=}} \\ -\varepsilon.\overline{\overline{L_F^X}} \end{pmatrix}^T \cdot \begin{pmatrix} \overline{\overline{C_R^J}} \\ \overline{\overline{C_R^X}} \end{pmatrix} = \overline{\overline{=}} \pi \tag{3.38}$$

and

$$\begin{pmatrix} \overline{\overline{C_R^J}} \\ \overline{\overline{C_R^X}} \end{pmatrix} = \left(\overline{\overline{L_F^J}} - \varepsilon \cdot \overline{\overline{L_F^X}} \right)^{-1} \cdot \overline{\overline{=}} \pi \tag{3.39}$$

The control coefficients for the independent fluxes and metabolites can thus be calculated using Eq. 3.39. If required, all control coefficients can then be calculated using Eqs. 3.34 and 3.36. Implicit in Eq. 3.39 are not only the summation and connectivity theorems, but also the structural relationships of the pathway (Ehlde and Zacchi, 1996).

In the light of this information, stoichiometric matrix was formed with the internal metabolites. To have a better operation, the rows were renumbered in $\overline{\overline{N}}$ to have all independent rows above all dependent rows and all independent columns at the right of the dependent columns in $\overline{\overline{N}}$. Then, the matrices $\overline{\overline{N_R}}$ and $\overline{\overline{N_C}}$, the first m_0 independent rows and the last m_0 independent columns of $\overline{\overline{N}}$, respectively were constructed. $\overline{\overline{L^X}}$ and $\overline{\overline{L^J}}$ matrices were composed using the Eqs. 3.9 and 3.16 as well as the matrices N_0 and N_{RC}. In order to convert $\overline{\overline{L^X}}$ into $\overline{\overline{L_F^X}}$, and $\overline{\overline{L^J}}$ into $\overline{\overline{L_F^J}}$, each element of the matrices were multiplied by x_i / x_k and J_i / J_k, respectively. Furthermore, using the information embedded in ordinary differential equations which represent the changes of each state variable. ODEs were converted into flux distributions of each reaction. Elasticity coefficients were determined by taking the derivative of these flux distributions with respect to each internal metabolite. Finally, a MATLAB code was written to calculate the control coefficients for the independent fluxes and metabolites using Eq. 3.39.

3.6.2 Metabolic Pathway Analysis

In a metabolic network consisting of cellular reactions, the analysis of the fluxes allows one to establish a relationship between cell genotype and phenotype. One of the main approaches for the flux analyses of metabolic networks is a metabolic pathway analysis (MPA), which is used to define the structure of the metabolic network and the overall capabilities of the organism. The method only uses information about the stoichiometry and the reversibility or irreversibility of reactions.

3.6.2.1 Elementary Flux Mode Analysis: An important tool of the metabolic pathway analysis is the elementary flux modes (EFM). A 'flux mode' is defined as a steady-state flux distribution in which the proportions of fluxes are fixed. It is called 'elementary' if it is nondecomposable. In other words, an elementary flux mode is a minimal set of enzymes that can operate at steady state with all irreversible reactions proceeding in the direction prescribed thermodynamically (Schuster *et al.*, 1999). EFM analysis allows the discovery and analysis of the meaningful routes in metabolic networks.

3.6.2.2 Control Effective Flux Analysis: Control-effective flux (CEF) analysis is another tool in assessing a metabolism (Stelling *et al.*, 2002). The CEFs which are directly determined from the set of EFMs, represent the importance of each reaction of a metabolism for efficient and flexible operation of the entire metabolic network, i.e. CEF analysis is performed to analyze EFM results quantitatively and to make comparisons for the activities of reactions in the metabolic pathway (Çakır *et al.*, 2004).

In the present work, the determination of elementary flux modes and control effective fluxes were performed via the simulation tool, CellNetAnalyzer / FluxAnalyzer V. 9.2 (Klamt *et al.*, 2003).

4. RESULTS AND DISCUSSION

4.1 Reconstruction of Insulin Signaling Network

In recent years, genomics (the study of genomes), transciptomics (global analysis of gene expression) and proteomics (the large-scale study of proteins) have produced a large amount of molecular interaction data, contributing to reconstruction of specific cellular networks (Pandey and Mann, 2000).

4.1.1 Core Proteins

The core protein set of the insulin signaling network was formed as explained in Section 3.3. 27 core proteins known as related to insulin signaling network were identified using Gene Ontology (GO) annotations (molecular function, biological process, cellular component). The core proteins and their Entrez Gene IDs are listed in alphabetical order in Table 4.1.

Table 4.1. Core proteins of the insulin signaling network

Approved Symbol	Entrez Gene ID	Approved Symbol	Entrez Gene ID	Approved Symbol	Entrez Gene ID
AHSG	EG197	GRB10	EG2887	MLLT7	EG4303
AKT1	EG207	GRB2	EG2885	PDPK1	EG5170
AP3S1	EG1176	IBP5	EG3488	PHIP	EG55023
BAIAP2	EG10458	IGF1	EG3479	PIK3R1	EG5295
BCAR1	EG9564	IGF1R	EG3480	PIK3R3	EG8503
CILP	EG8483	IGF2	EG3481	SNF1LK2	EG23235
ENPP1	EG5167	INS	EG3630	SOCS1	EG8651
FOXC2	EG2303	INSR	EG3643	SORBS1	EG10580
GAB1	EG2549	IRS-1	EG3667	TSC2	EG7249

8 proteins, reported as critical nodes in literature, form an important part of the insulin signaling network (Taniguchi *et al.*, 2006; Sedaghat *et al.*, 2002; Karlsson, 2007; Myers and White, 2003; Lizcano and Alessi, 2002; Lomberk and Urrutia, 2009; Holgado-Madruga *et al.*, 1996; Rosivatz, 2007; Cheatman and Kahn, 1995) and they are given in Table 4.2. Each of these network components constitutes a critical node for the insulin signaling network, because each is essential for insulin action.

Table 4.2. Important core proteins of the insulin signaling network

Approved Symbol	Entrez Gene ID	Approved Symbol	Entrez Gene ID	Approved Symbol	Entrez Gene ID
AKT1	EG207	INS	EG3630	IRS2	EG8660
GRB2	EG2885	INSR	EG3643	PIK3R1	EG5295
IGF1R	EG3480	IRS-1	EG3667	PIK3R3	EG8503

The binding of insulin receptor substrates 1 and 2 (IRS-1 and IRS-2) to the phosphorylated insulin receptor results in their phosphorylation, that further allows binding and activation of PI3K (Langeveld and Aerts, 2009). Studies with genetic deletion on mouse models with genetic deletion indicate that the IRS proteins have complimentary functions in different tissues as immediate substrates for insulin and IGF1 receptors (Li and Zhang, 2007). The isoforms of PI3K play a pivotal role in the metabolic and mitogenic actions of insulin. Akt has a role in GLUT4 translocation via phosphorylating and regulating components of GLUT4 complex. Furthermore, GRB2 plays a central role in signaling by receptor protein-tyrosine kinases in Ras/MAPK pathway (Taniguchi *et al.*, 2006).

4.1.2 Protein-Protein Interactions

Sources of the GO annotations are experimental, computational or other such as author statements or automatic electronic annotations. The present 27 core proteins have reliable GO annotations with experimental evidence. Using the GO annotations of these 27 core proteins, an annotation-collection table was formed to be used in the reconstruction of insulin signaling network. The annotation-collection table includes 50 molecular function

annotations, 140 biological process annotations and 31 cellular component annotations. The GO annotations of each protein in BioGRID-Organism-2.0.45 file for *H. sapiens* was then compared with the GO terms of 27 core insulin signaling proteins using the above mentioned annotation-collection table. New proteins were accepted to the network if all three of their GO annotations match with the GO annotation collection set of the core proteins, meaning that, it has a high probability of having characteristics in the insulin signaling. Otherwise, if one or more of their GO annotations do not match, they were rejected.

The core proteins FOXC2 and CILP do not have any interaction data either in databases scanned or in literature. The interactions of these proteins were tried to be found by homology search in another organisms using BLAST (Basic Local Alignment Search Tool). As the homology of CILP in other organisms have high E-values and hence are not reliable, the interaction of CILP could not be found. Furthermore, the most trustable homologs of FOXC2 protein were found in *Mus Musculus* as explained in Figure 4.1 and given in Table 4.3. The first neighbors of the homolog proteins were tested whether they were related to insulin signaling or not. Comparing the GO annotations of Bmpr2 with the terms in the present annotation-collection table, it was seen that it has a role in insulin signaling network. Then, the homolog protein of Bmpr2 was found in *H. sapiens* as BMPR2. Finally, the interaction of FOXC2 with BMPR2 was added to the reconstructed network.

Due to the fact that, *H. sapiens* is a complex organism, the lack of GO annotations can be observed frequently. It can be suggested that this general lack of information may cause a decrease in the efficiency of the network reconstruction process. Most of the protein interaction networks include proteins with at least one of the annotations (molecular function annotations, biological process annotations and component annotations) unknown. In most cases, all three of the GO annotations are missing. In the present reconstructed insulin signaling network, all 3 terms of GO annotations are unknown for 1464 proteins. Since the number of unknown GO annotations of the proteins is very high (Figure 4.2), 3012 proteins (726 proteins with 1 unknown, 822 proteins with 2 unknown, 1464 proteins with 3 unknown GO annotations) could not be included in the reconstructed network. Then, to have a robust network, 786 proteins with all known Gene

Ontology terms were selected. The number of interactions between these 786 proteins was found from BioGRID database as 1839. Most of the interactions given in databases were rejected by the algorithm used here either due to the lack of GO terms or the GO terms of the interacting partners do not match with those in annotation-collection table. 169 of these 786 proteins formed an isolated part having no connection to the remaining part of the network.

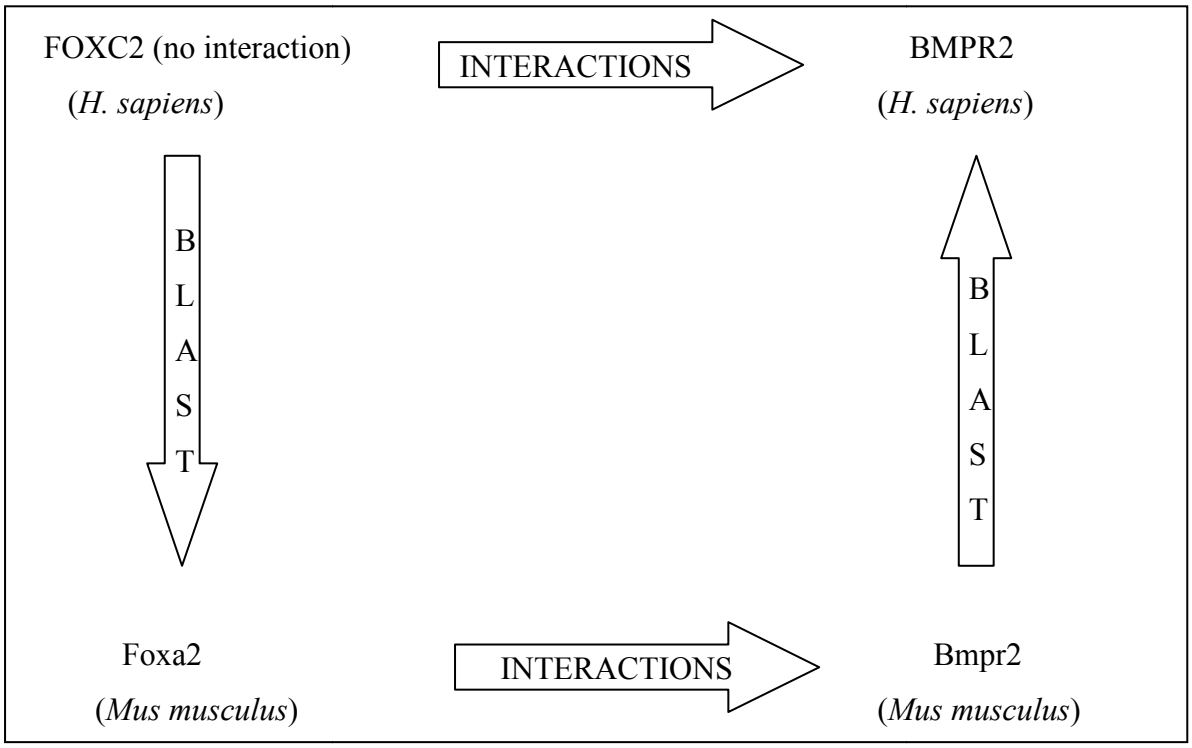


Figure 4.1. Homology search of the interaction of FOXC2

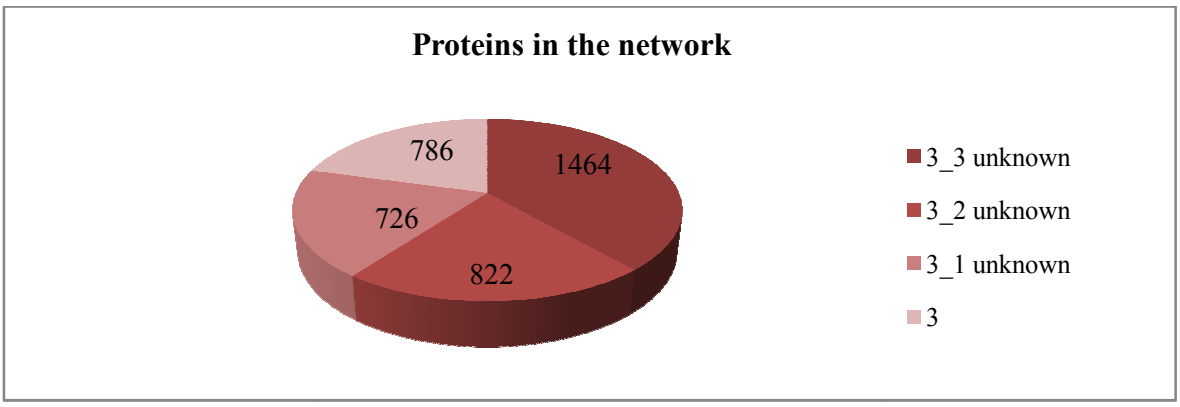


Figure 4.2. The proteins related to insulin signaling network

Table 4.3. The top confident homologies of FOXC2 in *M. Musculus*

Proteins	Organism	E-value	Interactions	Source
Foxc2	*Mus Musculus*	0.0	No interaction	-
Foxc1	*Mus Musculus*	2E-79	No interaction	-
Foxl1	*Mus Musculus*	1E-47	Bmpr2	BIND
Foxa2	*Mus Musculus*	2E-43	Acadm	BIND
			Acadvl	BIND
			Gck	BIND
			Hnf4a	BIND
			Igfbp1	BIND
			Pklr	BIND
			Ucp2	BIND

4.1.3 Confidence Scores by STRING Database

Several studies discussed the successful use of GO annotations to increase the likelihood of true-positive interactions (Bader *et al.*, 2004; Patil and Nakamura, 2005; Wu *et al.*, 2006). STRING database was here utilized to develop a comprehensive map including high-scoring physical connections that link the relevant signal transduction components to each other and to adjacent networks.

Out of 1839 interactions, the scores of 1529 interactions were found to be reliable using the STRING database (file protein.aliases.v8.0.txt.gz). The examples of interactions existing in the reconstructed network with the highest confidence score are given in Table 4.4. This analysis revealed an average confidence score of whole interactions as 873. In order to make the signaling network more trustworthy, a threshold value should be selected. According to the Figure 4.3, choosing the interactions whose confidence scores are above 900 did not cause a considerable change in the structure of the network. Although it covered nearly 60 % of all the interactions, it enclosed 75% of all the proteins in the reconstructed network and 95% of the core proteins. Moreover, when the threshold was set to 900, the graph theoretic properties remain the same.

Table 4.4. Example interactions with highest confidence score

Protein A	Protein B	Confidence Score
ACVR1	SMAD1	999
PML	TP53	999
IL6R	STAT3	999
APP	APBB2	999
APP	APBB3	999
APP	APBB1	999
PTPN1	GRB2	999
GRB2	SHC1	999
FLT4	GRB2	999
LAT	GRB2	999
ACVR1	SMAD1	999

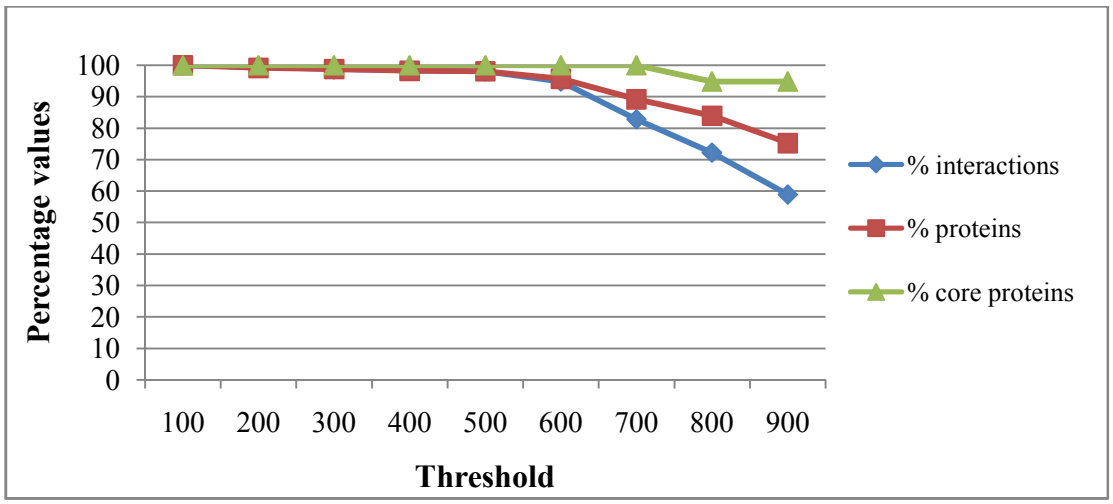


Figure 4.3. The coverage of reconstructed insulin signaling network by changing the threshold of scores of interactions

4.1.4 Reconstruction of Domain-Domain Interaction Network

The domain-domain interactions of the reconstructed network were investigated with the superfamily assignments of *H. sapiens* for all transcripts and longest transcript per gene

using Superfamily 1.73 database. Since the file of all transcripts file comprehends 46598 total sequences with 30713 assignments (Table 4.5), this file was chosen for further work to find the domain-domain interactions of the proteins in the reconstructed insulin signaling and GLUT4 translocation signaling networks. Unfortunately, this file only covers 41% of the amino acid data (Table 4.5). Therefore, the complete domain information of all the proteins could not be obtained and consequently a domain-domain interaction network could not be constructed for insulin signaling.

The frequently observed domain pairs in the reconstructed GLUT4 translocation subnetwork and the reconstructed insulin signaling network were identified as SH2 (Src Homology 2)-SH2 and PH (Pleckstrin-homology domain)-PH domain-domain pairs. This finding is consistent with the information given in literature (Table 4.6). PH domains, which are found in most of the proteins that interact with the insulin receptor, bind to charged headgroups of specific phosphatidylinositides and are thereby targeted preferentially to membrane structures. PH domains in the IRS proteins target the proteins to the membrane adjacent to the insulin receptor. Furthermore, PI3K translocates to the plasma membrane where it converts $PI(4,5)P_2$ into $PI(3,4,5)P_3$ causing the recruitment of PDK1 and Akt to the plasma membrane via their PH domains. Akt is upstream of several kinases involved in proliferation and apoptopic signaling which are often found to be deregulated or mutated in tumors; on the other hand it is also the key enzyme in insulin signaling regulating glucose uptake and cell growth (Rosivatz, 2007).

Most of the intracellular partners of the insulin receptor substrates contain Src homology (SH2) domains. The proteins with SH2 domains such as adaptor molecules, the regulatory subunit of PI3K, or the adaptor molecule GRB2 which play central roles in insulin signaling (Holgado-Madruga, 1996), bind to tyrosine residues on IRS proteins through their specific SH2 domains.

In conclusion, frequently observed domain pairs found in the reconstructed networks lead to critical nodes in the network. Therefore, the important nodes were validated again with this analysis.

Table 4.5. Assignment statistics for domain-domain interactions

	All transcripts	Longest transcript per gene
Sequences		
Total	46591	22997
With assignment	30713	14701
Average length	505	502
Amino acid coverage	41%	40%
Domains		
Total	64236	30312
Unique superfamilies	1056	1048
Unique families	1299	1285
Average length	312	315
Average superfamily size	60.8	28.9
Formed by duplication	98%	97%
Domain combinations		
Domain pairs	1378	1281
Unique architectures	5603	4299

Table 4.6. Classes of domains involved in protein-protein interaction in insulin signaling (Virkamaki *et al.*, 1999)

Domain	Function in insulin signaling
PH domain	Recruitment of IRS proteins to the insulin receptor
PTB domain	Recognition of the insulin receptor by IRS proteins
SH2 domain	Signal transmission from IRS proteins to downstream proteins
SH3 domain	Signal transmission from SH2 proteins to downstream molecules

4.1.5 Overview of Reconstructed Signaling Network

When the proteins and the interactions of the reconstructed signaling network were further investigated, it can be seen that most of the core proteins and the important interactions in the general insulin signaling pathway given in literature were covered. The reconstructed network included critical nodes such as IRS, Akt and the isoforms of PI3K, which form an important part of the insulin signaling. This signaling network is mediated by a group of complex, highly integrated proteins that control several processes through PI3K-Akt pathway, which is responsible for the metabolic actions of insulin, and through the Ras-mitogen activated protein kinase (MAPK) pathway, which regulates some genes and cooperates with the PI3K pathway to control cell growth and differentiation (Taniguchi *et al.*, 2006). Akt, which is included in the reconstructed network, is a serine/threonine kinase that is a downstream target of PI3K signaling, and it also regulates the translocation of the glucose transporter GLUT4 to the plasma membrane. Furthermore, the network has a role in the metabolic and mitogenic actions of insulin through PI3K proteins.

Insulin promotes glucose uptake by muscle and adipose tissues via stimulation of GLUT4 from intracellular sites to the plasma membrane. However, GLUT4 was initially not included in the reconstructed insulin signaling network due to the fact that the GO annotations of the first neighbors (interacting partners) of GLUT4 protein did not match with those present in the annotation-collection table of core proteins. In order to make a connection between GLUT4 subsystem (Sedaghat *et al.*, 2002) and the reconstructed insulin signaling network of the present study, GLUT4 protein and its first neighbors, DAXX (death domain associated protein) and EHD2 (EH domain containing 2) were further investigated in terms of interacting partners. EHD2 could not be added to the reconstructed network, since its first and second neighbors were not present in the insulin network. However, DAXX could be added to the reconstructed network through its first neighbors as dashed lines (Figure 4.4). DAXX protein was actually connected to the network through 11 different proteins (AR, androjen receptor; PML, promyelocytic leukemia protein; DAPK3, death-associated protein kinase 3; NR3C1, nuclear receptor subfamily 3, group C, member 1; HDAC1, histone deacetylase 1; HDAC2, histone

deacetylase 2; HDAC3, histone deacetylase 3; TGFB2, transforming growth factor beta 2; FAS, TNF receptor superfamily, member 6; CDKN2A, cyclin-dependent kinase inhibitor 2A; MDM2, p53 binding protein homolog (mouse)). Finally, the reconstructed network successfully covered most of the insulin signaling cascade in *H. sapiens.*

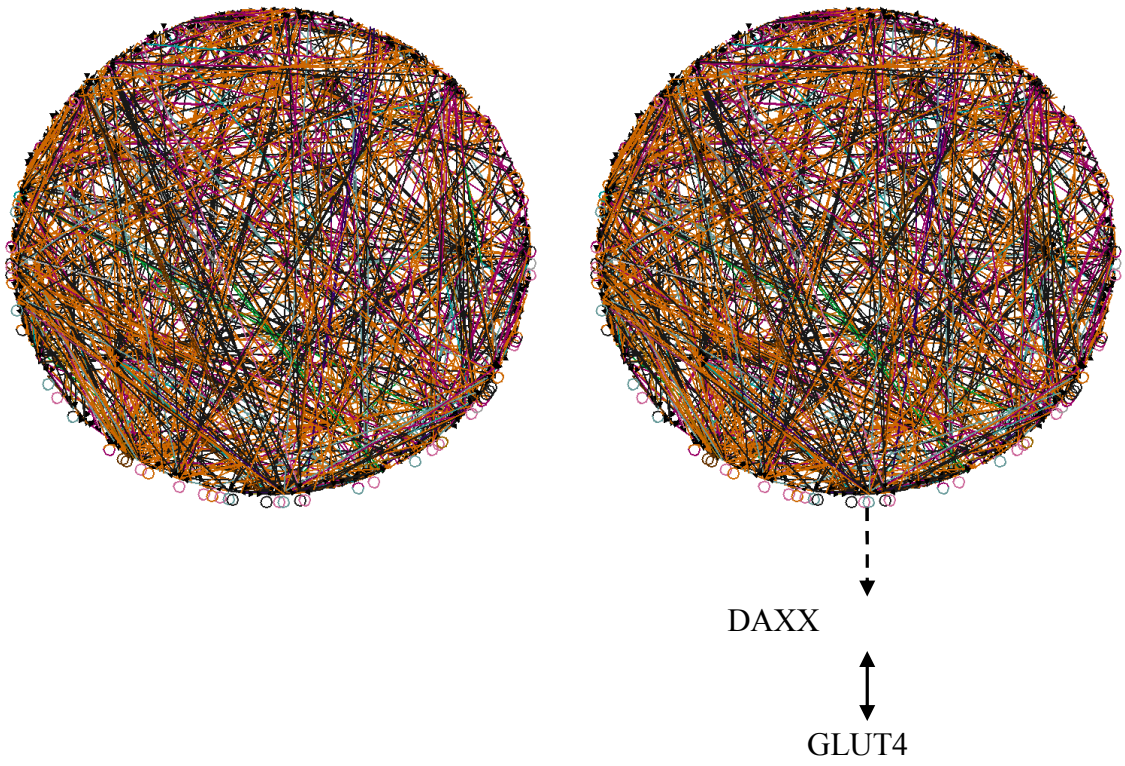


Figure 4.4. The reconstructed insulin signaling network in *H. sapiens* before GLUT4 protein (*left*), after GLUT4 protein (*right*)

4.2 Structural Analysis of Reconstructed Signaling Network

In this study, structural analysis was performed by both graph theoretical calculations and pathway analysis via linear paths. The connectivity distribution of the insulin signaling molecules in the reconstructed network was obtained using the techniques of graph theory. The GLUT4 translocation subnetwork was reconstructed using the linear paths extending from ligands to phenotypes and will be discussed under Section 4.4.1.

4.2.1 Graph Theoretic Analysis

The first goal is usually the identification of the network structure: that is, to establish whether the topology is best described by the power-law model, since the emergence of power-law distribution is intimately linked to characterization of the biological systems. For this purpose, the graph theoretical analysis was performed with the algorithm (Arga *et al.*, 2007) explained in Section 3.1.1. The insulin signaling network was represented by an undirected interaction graph with 416 nodes, 1790 edges. The network diameter and the mean path length were found as 21 and 10.5, respectively.

The degree distribution of the present scale-free network follows the power law model $P(k) \approx k^{-1.95}$, as shown below (Figures 4.5 and 4.6). The degree exponent is usually in the range of $2 < \gamma < 3$ (Albert and Barabasi, 2002). For the reconstructed signaling network, it was found to be 1.95 which is within the acceptable range (Figure 4.6). Furthermore, the continuously decreasing degree distribution (Figure 4.5) indicates that, the majority of the network proteins have a low number of interactions and only a few highly-connected nodes (also called hubs) has a significant number of interacting partners.

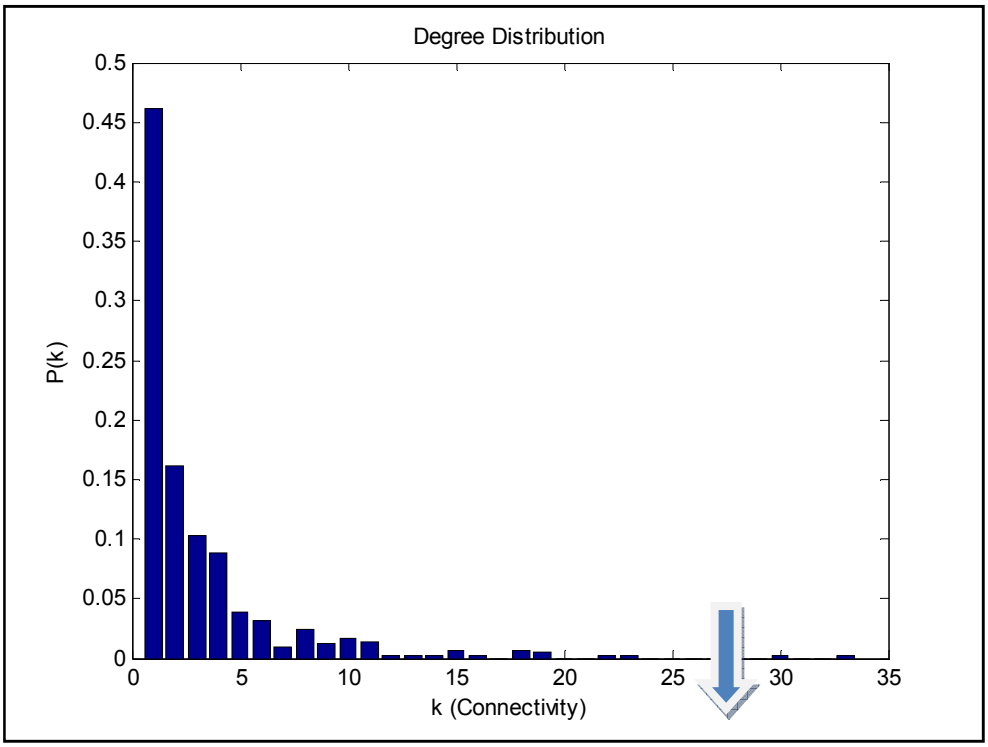


Figure 4.5. Degree distribution plot of the reconstructed signaling network

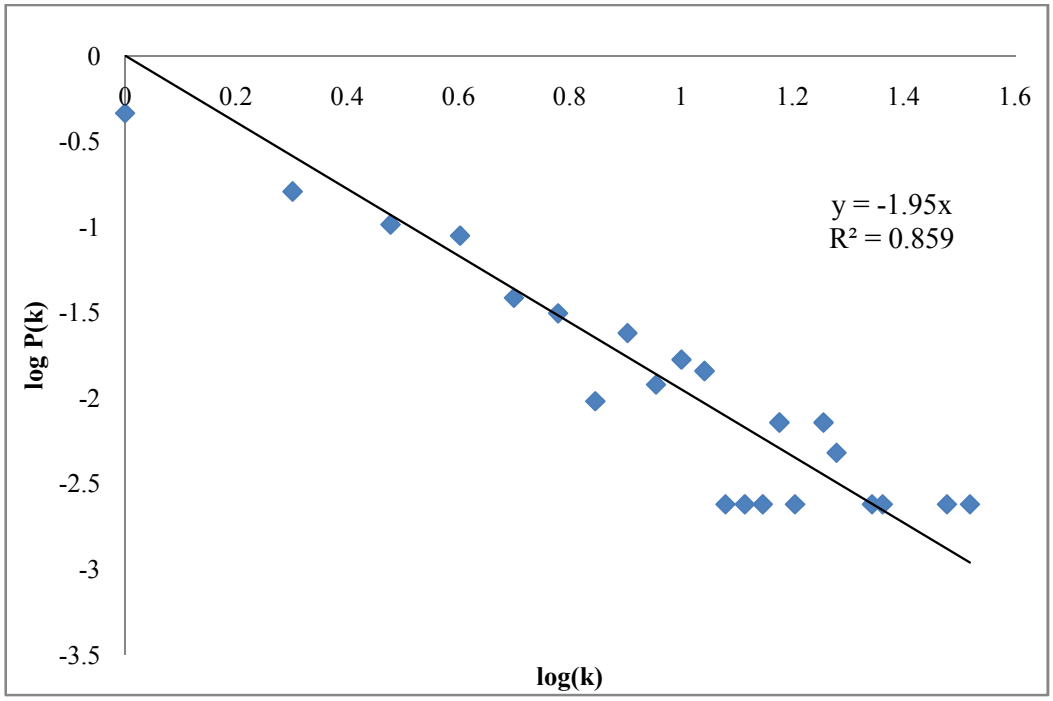


Figure 4.6. Log P(k) vs. log (k) graph

Highly-connected proteins were determined by the selection criterion based on the position of a break point on a protein interaction distribution plot shown in Figure 4.5. or 4.7. 2 hubs were identified, which were GRB2, growth factor receptor-bound protein 2 and TP53, tumor protein p53 having 30 and 33 connectivity, respectively. These results are in agreement with the literature because GRB2 plays a central role in signaling by receptor protein-tyrosine kinases, where its SH2 domain binds to the receptor and its two SH3 domains link to effectors and it has also a role in the insulin signaling pathway through insulin receptor substrate-1 (IRS-1) (Myers and White, 1993; Holgado-Madruga, 1996). Moreover, tumorigenesis is associated with enhanced cellular glucose uptake and increased metabolism. With the mutations of p53 tumor suppressor, it was tested whether p53 regulates expression of the GLUT4 transporter genes. It was concluded that mutations within the DNA-binding domain of p53, which are usually associated with malignancy were found to impair the repressive effect of p53 on transcriptional activity of the GLUT4 gene promoter, thereby resulting in increased glucose metabolism and cell energy supply (Schwartzenberg-Bar-Yoseph *et al.*, 2004).

The graph theoretic analysis was also performed to the final version of the insulin signaling network including GLUT4 protein to determine whether the addition of GLUT4 via DAXX protein is biologically meaningful or not. The new undirected interaction graph contained 418 nodes and 1828 edges. The mean path length and network diameter were found as 10.5 and 21, respectively. The final degree distribution of the scale-free network was not changed and followed power law $P(k) \approx k^{-1.95}$ (Figure 4.8). With a higher R^2 value, the mean path length, the network diameter, and the highly connected nodes remain the same. Moreover, when the graph theoretic properties of the reconstructed insulin signaling network were compared with those of other protein interaction networks reported in literature (Table 4.7), a significant similarity can be seen between the distance measures, especially in network diameter.

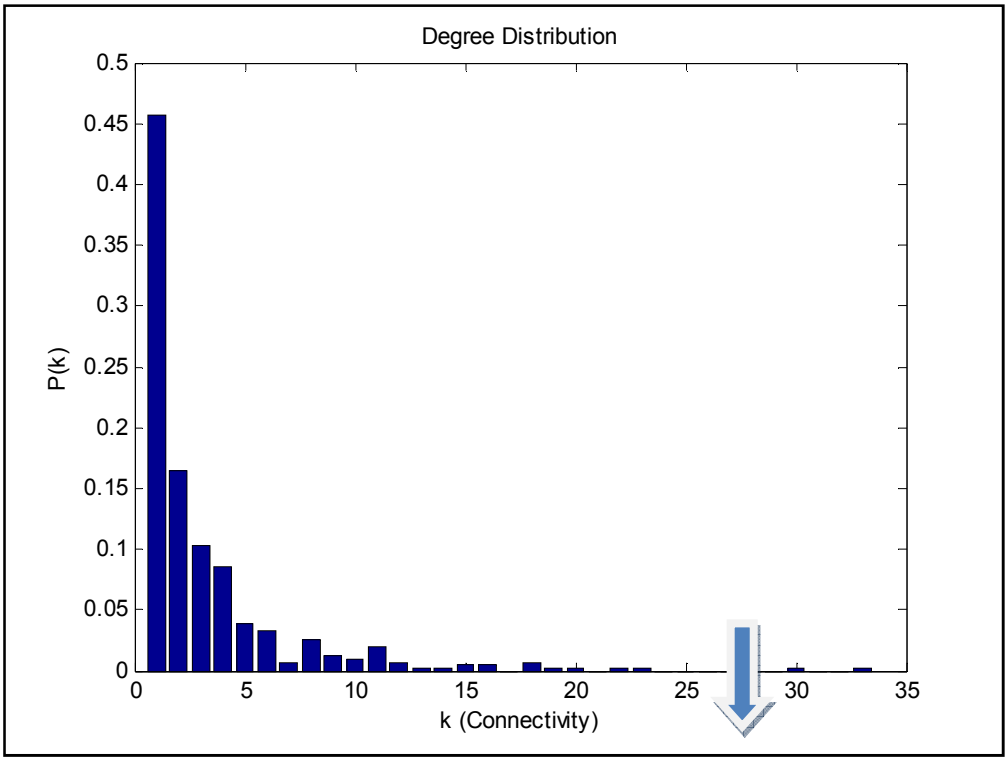


Figure 4.7. Degree distribution plot after GLUT4 adding

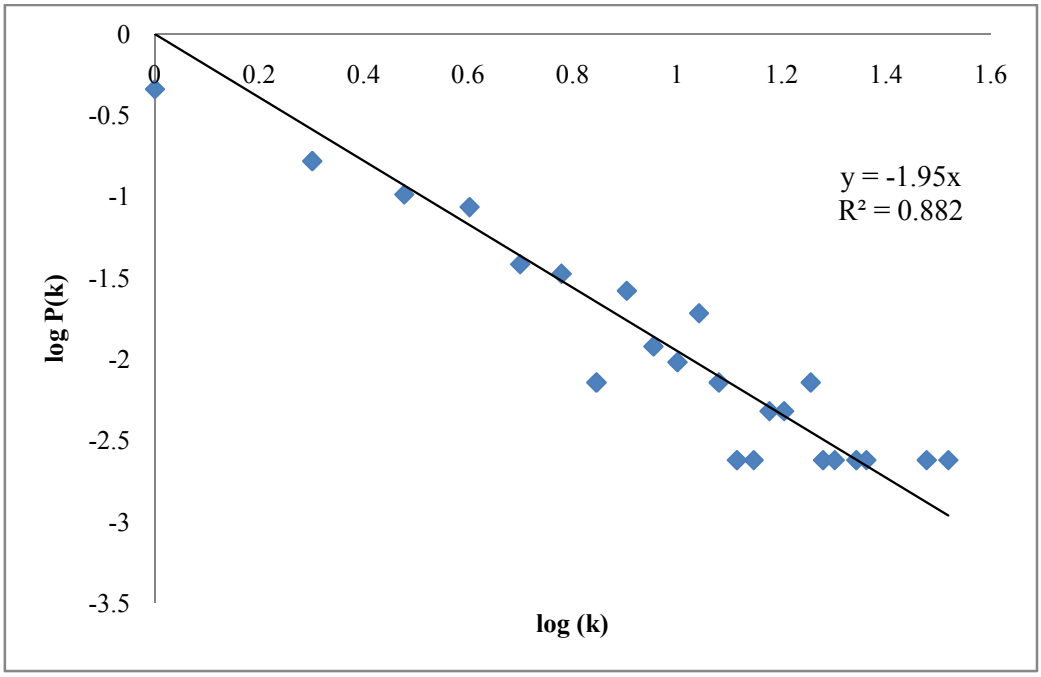


Figure 4.8. Log P(k) vs. log (k) graph

Table 4.7. Graph theoretic properties of protein interaction networks

Model	Number of nodes	Number of interactions	Mean path length	Diameter	Reference
Insulin Signaling	416	895	10.5	21	Present study
Insulin Signaling with GLUT4, DAXX	418	914	10.5	21	Present study
EGFR	329	1795	10.9	4.7	(Durmuş Tekir *et al.*, 2009)
DIP (*S. cerevisiae*)	5798	20098	4.90	12	Yook *et al.*, 2004
DIP(*D. Melanogaster*)	7451	22819	4.40	11	Wu *et al.*, 2005
DIP (*C. elegans*)	2638	4030	4.80	14	Wu *et al.*, 2005
DIP (*H. sapiens*)	1065	1369	6.80	21	Wu *et. al.*, 2005
DIP (*E. coli*)	553	761	5.50	16	Wu *et al.*, 2005
DIP (*M. musculus*)	329	286	3.60	9	Wu *et al.*, 2005

4.3 Dynamic Modeling and Analysis of Insulin Metabolic Network

Dynamic models are generally used to express and model the behaviour of the system over time. In this thesis, a dynamic model of insulin signaling was studied to enlighten the GLUT4 translocation subnetwork. The present dynamic model (Sedaghat *et al.*, 2002) has 21 state variables (12 internal metabolites and 9 external metabolites) and 20 reactions. The methods used for the analysis of the metabolic insulin network can be reviewed under three main sections: dynamic simulations under the stimulation of insulin, metabolic control analysis and metabolic pathway analysis by elementary flux modes. In simulations with the stimulation of insulin, the concentration changes of each state variable can be investigated. Metabolic control analysis ranks the genes or proteins or metabolites according to their importance in controlling and regulating cellular metabolic networks. Elementary flux mode and control effective flux analyses allow finding meaningful routes in metabolic networks.

4.3.1 Dynamic Simulations with the Stimulation of Insulin

A dynamic model of insulin metabolism was studied by generating time courses for all state variables in response to a maximally stimulating step input of 10^{-7} M insulin that was turned off after 15 minutes. The ordinary differential equations describing the insulin metabolism were taken from the work of Sedaghat *et al.*, 2002. It was observed that when insulin was removed after 15 min of stimulation, the concentration of state variables changed immediately. Thus, the insulin uptake and removal reactions were very important for the metabolic pathway.

When Figures 4.9 to 4.11 were investigated, it can be seen that the concentration of free insulin receptors returned to their basal levels with a half-time of 4 minutes after the removal of insulin following 15 min of stimulation. The state variables x_6, x_7, and x_8, which represent intracellular insulin receptors, did not change significantly with acute insulin stimulation (data not shown). These results are in good agreement with both published experimental data and previous results of subsystem models of receptor binding and recycling (Quon and Campfield, 1991; Standaert and Pollet, 1984; Wanant and Quon,

2000). In response to the rise in autophosphorylated surface insulin receptors, unphosphorylated IRS-1 was rapidly converted to tyrosine-phosphorylated IRS-1. Consistent with published experimental data (Madoff *et al.*, 1988), maximal IRS-1 tyrosine phosphorylation was observed within 1 min of the initiation of insulin stimulation. On the removal of insulin, IRS-1 underwent dephosphorylation back to the basal conditions with a half-time of ~8 min (Figure 4.12).

Maximal formation of the phosphorylated IRS-1/activated PI3K complex in response to 10^{-7} M insulin occurred within ~1.5 min (Figure 4.13). This is in good agreement with published data showing that PI3K and tyrosine-phosphorylated IRS-1 molecules associate quickly after insulin stimulation (Backer *et al.*, 1992; Giorgetti *et al.*, 1992). The time course for the disappearance of activated PI3K after the removal of insulin followed the time course for dephosphorylation of IRS-1. PI3K activated in response to insulin stimulation catalyzed the conversion of $PI(4,5)P_2$ to $PI(3,4,5)P_3$, which in turn drove the formation of $PI(3,4)P_2$. On removal of insulin, the phosphatidylinositides returned to basal levels (Figure 4.14). The levels of $PI(3,4,5)P_3$ controlled the formation of phosphorylated, activated PKC-ζ. Maximal PKC-ζ activation occurred within 3 min of insulin stimulation (Figure 4.15). After insulin was removed, the level of activated PKC-ζ declined back to basal levels.

The time course for phosphorylated, activated Akt was identical to that for PKC-ζ (data not shown). Insulin-stimulated activation of PKC-ζ and Akt mediates increased exocytosis of GLUT4 so that 40% of total cellular GLUT4 was at the cell surface and 60% was intracellular after maximal insulin stimulation. The simulations of insulin-stimulated GLUT4 recruitment occurred with a half-time of ~3 min, matching published experimental results (Holman and Cushman, 1996; Karnieli *et al.*, 1981). When insulin was removed, the surface and intracellular GLUT4 levels returned to their basal values (Figure 4.16). Thus the overall response of the complete dynamic model without feedback to an acute insulin input is in good agreement with both a variety of published experimental data and previously validated subsystem models. Therefore, this dynamic model of insulin metabolism is reliable for further studies and analyses.

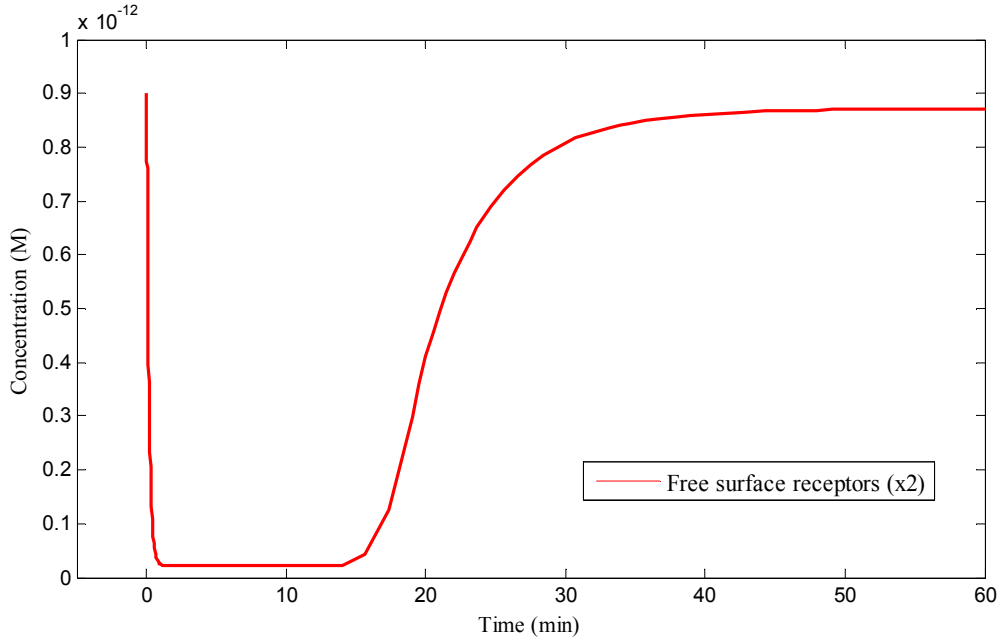


Figure 4.9. Time courses for unbound receptors

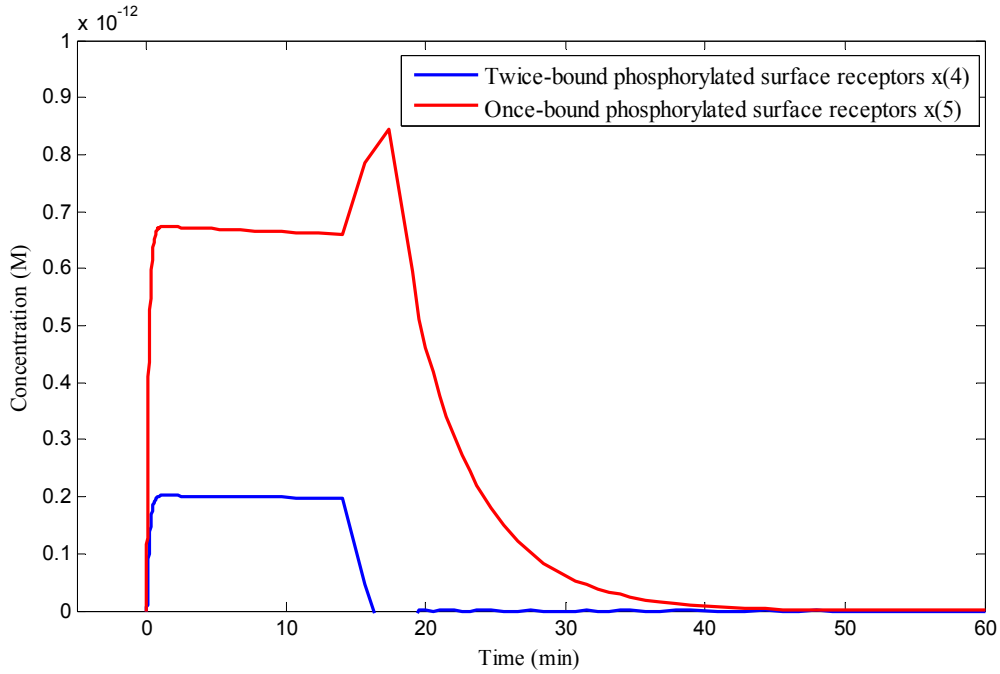


Figure 4.10. Time courses for once- and twice-bound phosphorylated surface receptors

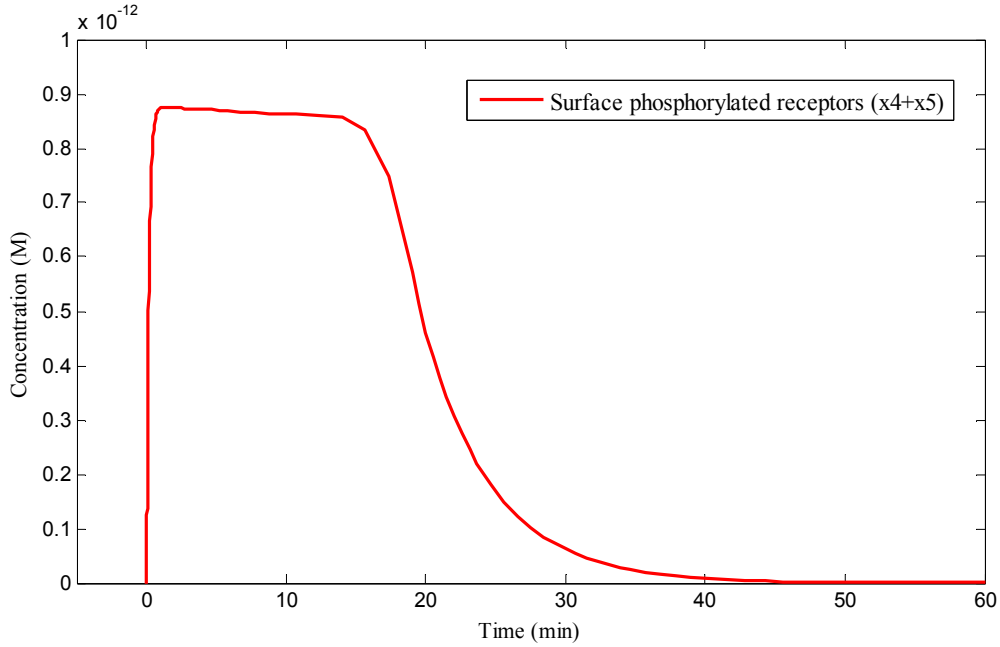


Figure 4.11. Time courses for total phosphorylated surface receptors

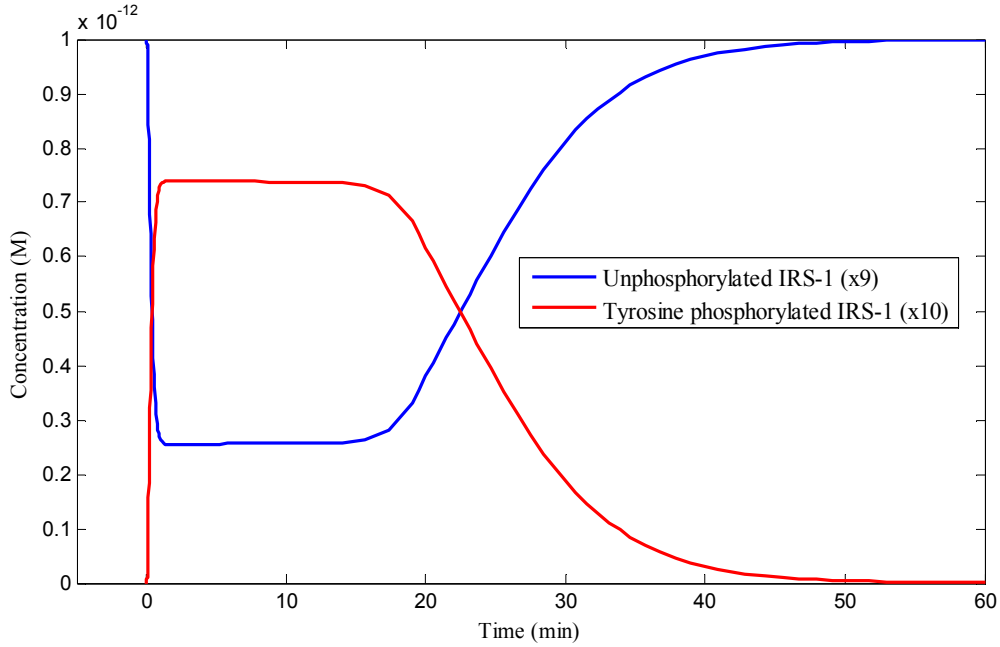


Figure 4.12. Time courses for unphosphorylated and tyrosine-phosphorylated IRS-1

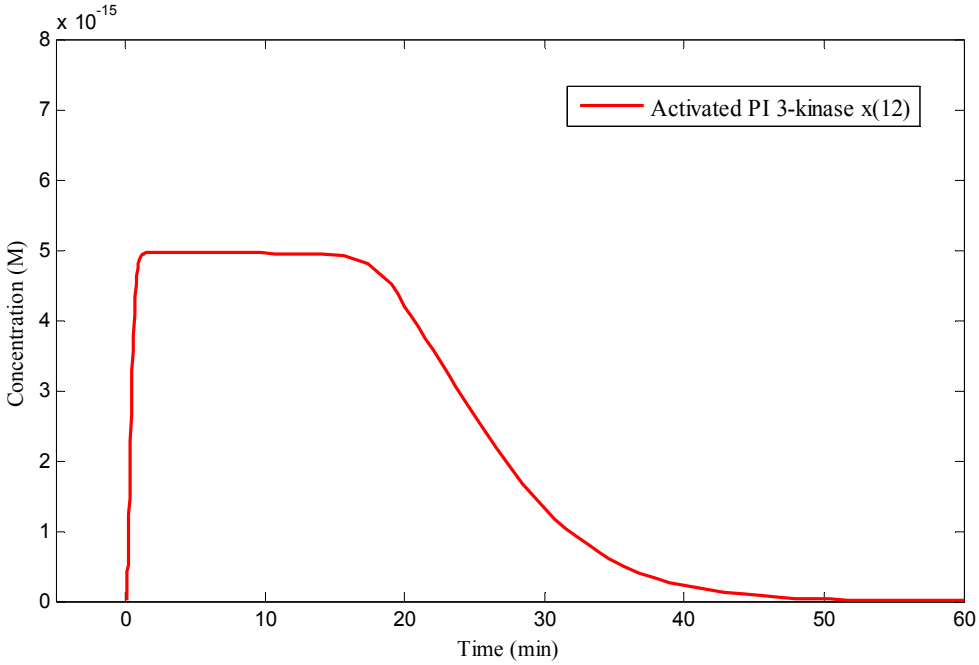


Figure 4.13. Time courses for activated PI3K

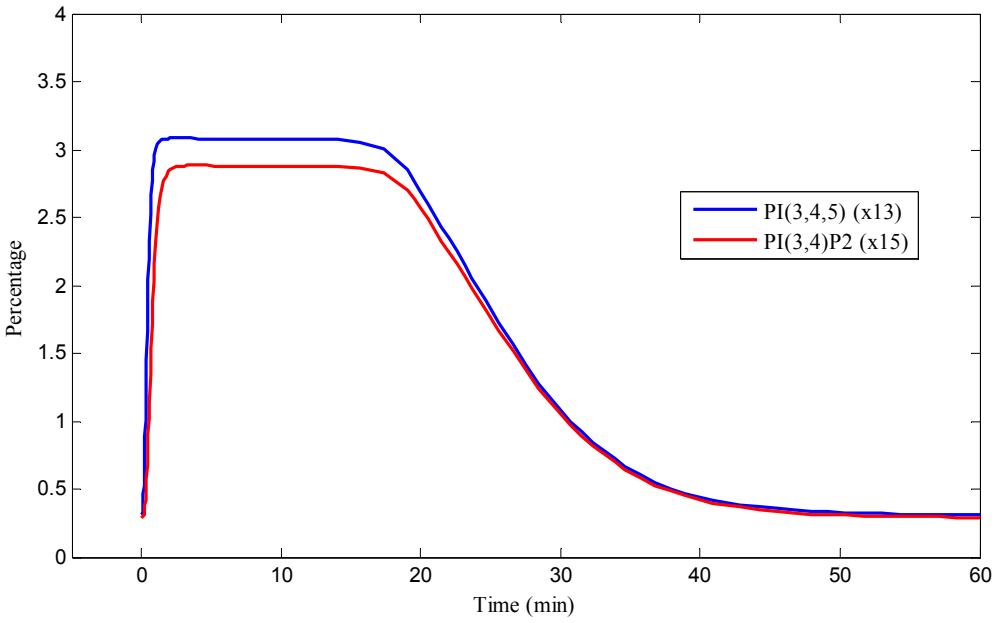


Figure 4.14. Time courses for levels of phosphatidylinositol3,4,5-triphosphate [$PI(3,4,5)P_3$] and phosphatidylinositol 3,4-bisphosphate [$PI(3,4)P_2$]

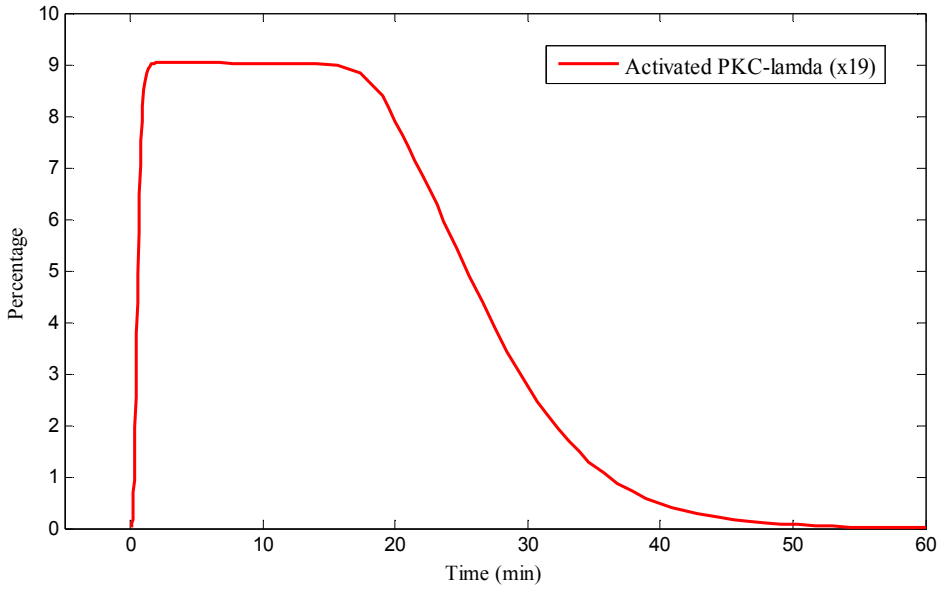


Figure 4.15. Time courses for activated PKC-ζ

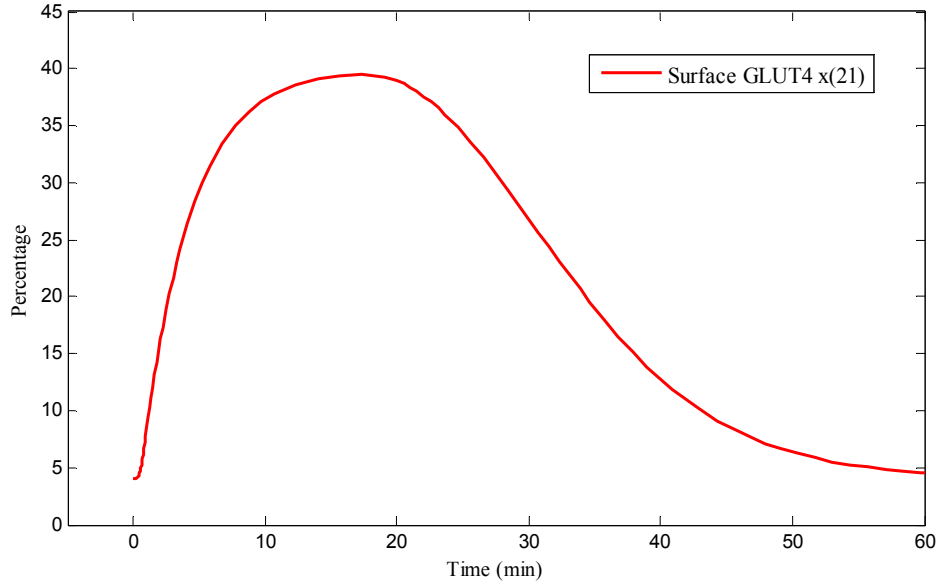


Figure 4.16. Time courses for cell surface GLUT4

4.3.1.1. Sensitivity Analysis: Sensitivity analysis was also performed by changing the rate constants of each reaction. Each rate constant of 20 reactions was multiplied by 1000 and divided by 1000, and the results were evaluated as follows. The reversible phosphorylation reaction of IRS-1 was found to have more control on the metabolic pathway. When k_7 and k_{m7}, which are the forward and backward rate constants of IRS-1 phosphorylation,

respectively, were changed, the concentration profile of each metabolite was changed (Figures 4.17-4.20).

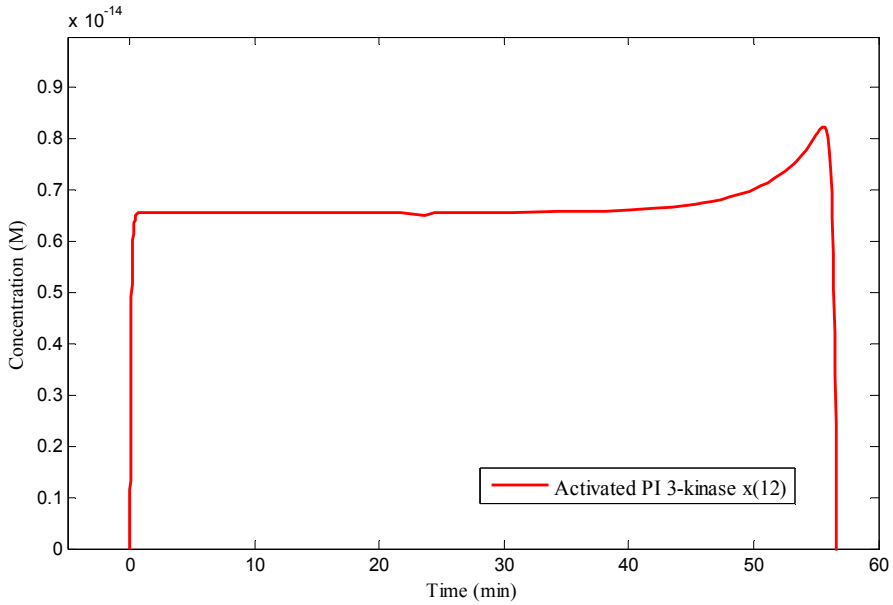


Figure 4.17. Time courses for activated PI3K after changing k_7

Consistent with literature, the phosphorylation reaction of IRS-1 has an important role in insulin signaling and IRS-1 is a critical node in the insulin signaling network having a role in insulin resistant, hyperinsulinaemic states, including obesity and Type-2 diabetes (Taniguchi, 2006).

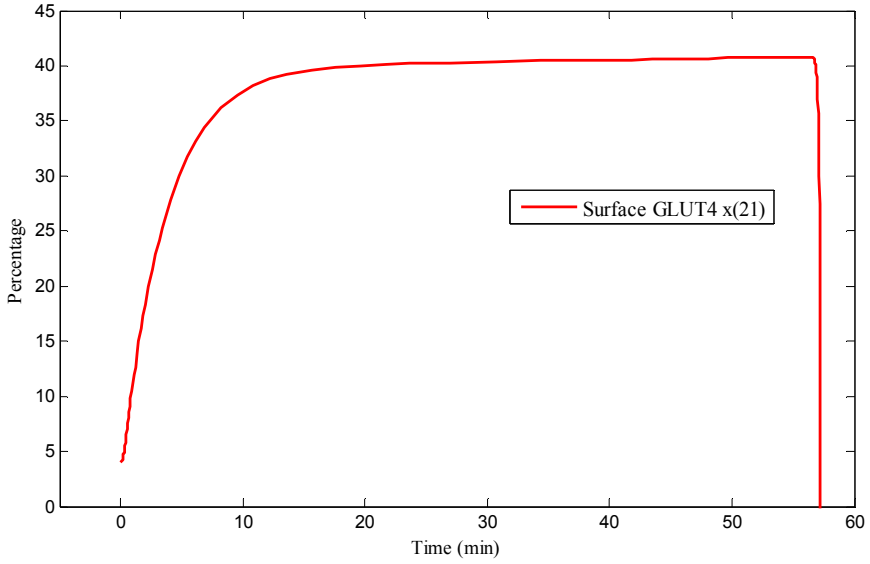


Figure 4.18. Time courses for cell surface GLUT4 after changing k_7

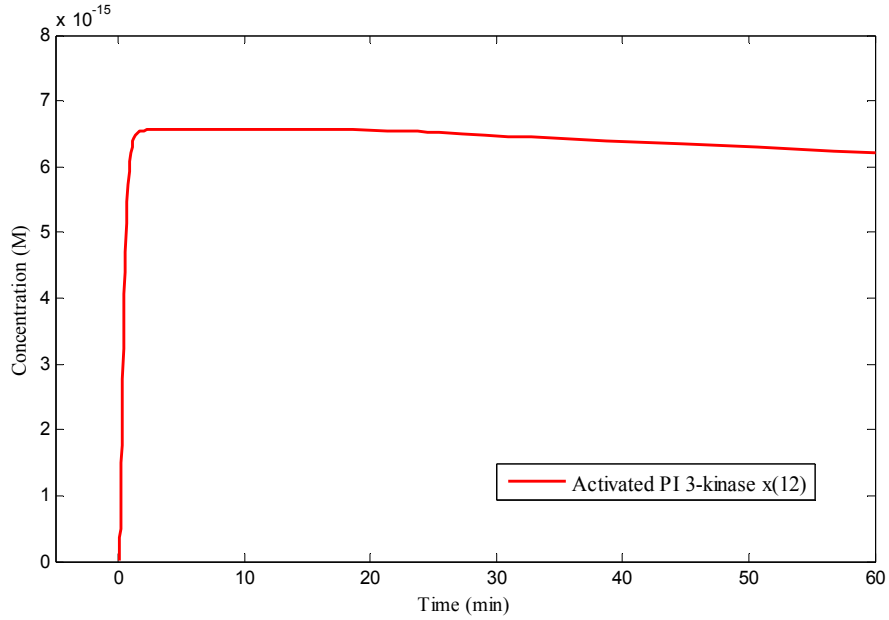


Figure 4.19. Time courses for activated PI3K after changing k_{m7}

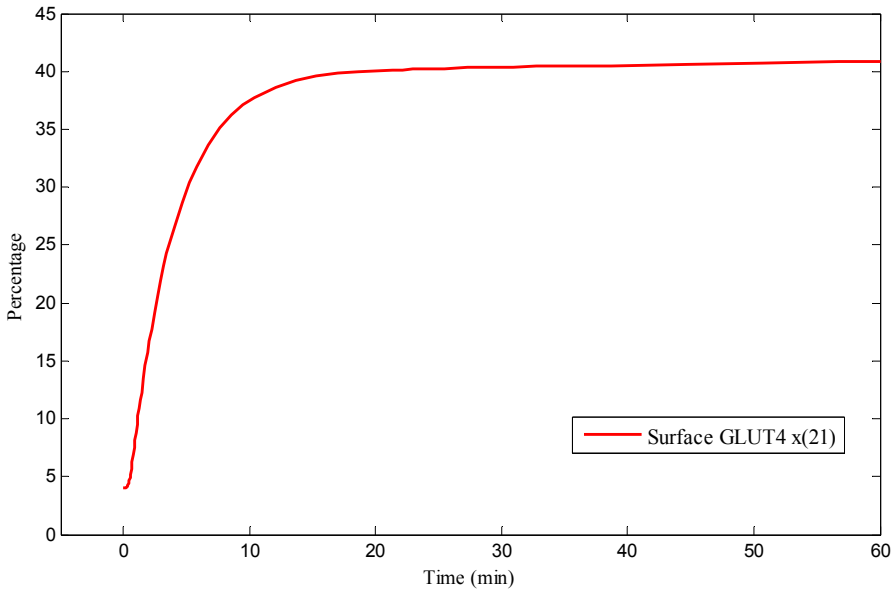


Figure 4.20. Time courses for cell surface GLUT4 after changing k_{m7}

4.3.1.2. Effects of Ceramides: Ceramide was known to block the translocation of Akt from cytoplasmic stores to the plasma membrane. The effect of the inhibition of Akt on GLUT4 translocation was reflected through the concentration profiles of activated Akt and cell surface GLUT4. When the inhibition of Akt phosphorylation reaction occurred, a decrease

in the maximum level of Akt was observed. Furthermore, a decrease in the GLUT4 concentration was also seen in the time interval of 10 and 30 min (Figure 4.22).

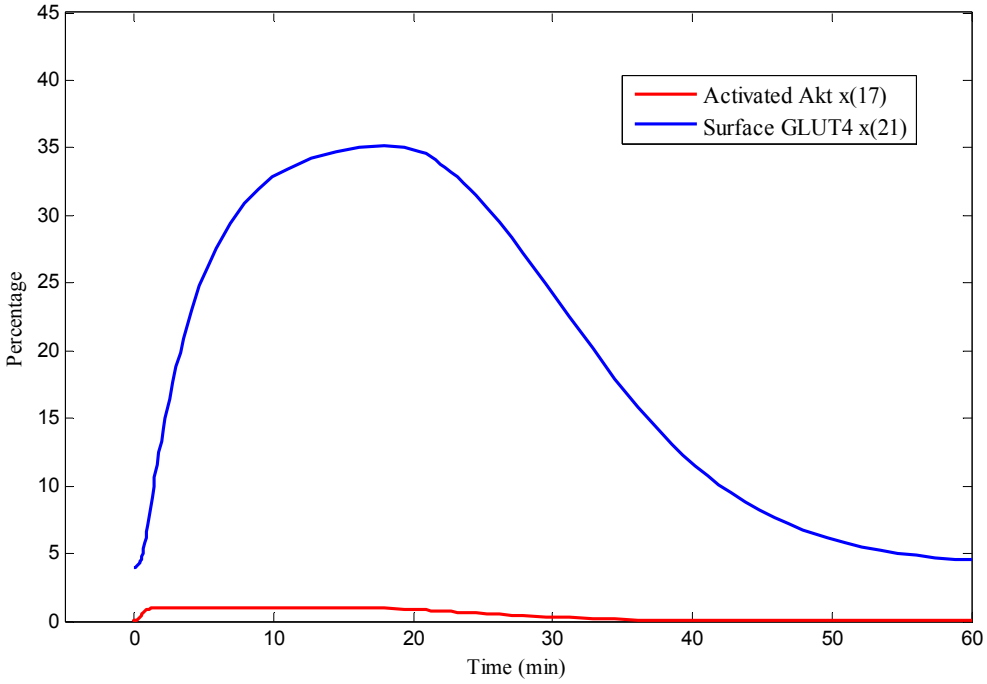


Figure 4.21. Time courses for surface GLUT4 and activated Akt

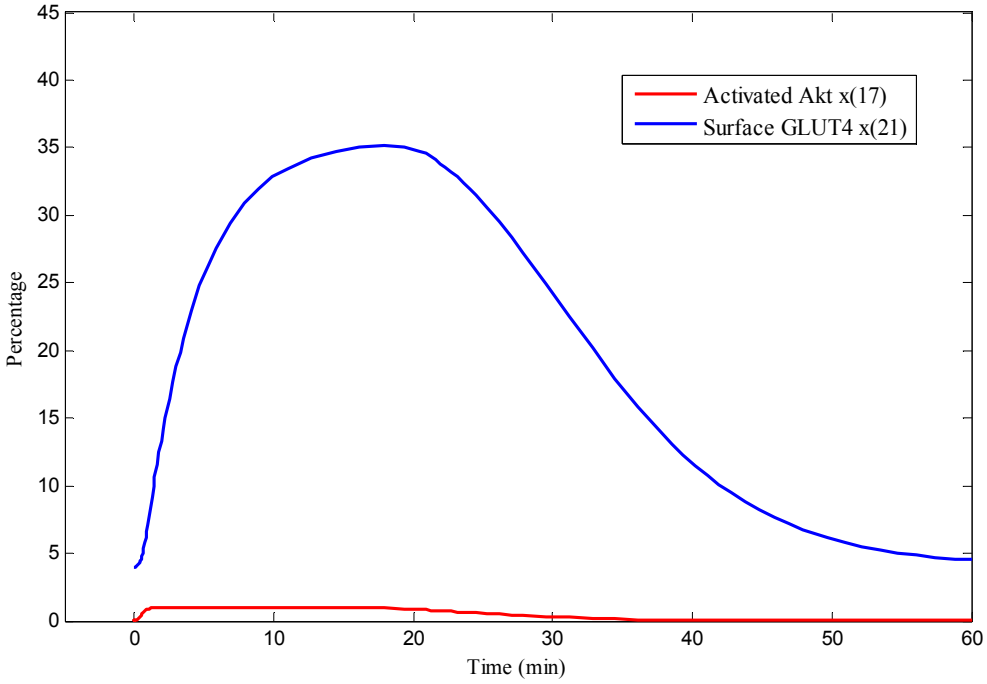


Figure 4.22. Time courses for surface GLUT4 and activated Akt after 90% inhibition

4.3.2 Metabolic Control Analysis

The metabolic pathway of insulin (Sedaghat *et al.*, 2002) contains 12 internal metabolites, 9 external metabolites and 20 enzymatic reactions. The stoichiometry of the pathway, $\overline{\overline{N}}$, was described by $m \times n$, 12×20 matrix. m_0 is the number of independent rows of $\overline{\overline{N}}$, and n_0 is the number of dependent columns, found by $n - m_0$, they were determined as 12 and 8, respectively. To have a better operation, the rows and the columns were renumbered in $\overline{\overline{N}}$. All independent rows and all independent columns were set above all the dependent rows, and at the right of dependent columns in $\overline{\overline{N}}$, respectively.

Then, the matrices $\overline{\overline{N_R}}$ and $\overline{\overline{N_C}}$, the first m_0 independent rows and the last m_0 independent columns of $\overline{\overline{N}}$, respectively, were formed. Moreover, $\overline{\overline{N_{RC}}}$, the intersection matrix of $\overline{\overline{N_R}}$ and $\overline{\overline{N_C}}$, was a 12×12 matrix of $\overline{\overline{N}}$. $\overline{\overline{L^X}}$ and $\overline{\overline{L^J}}$ matrices were composed by Eqs. 3.9 and 3.16. Then $\overline{\overline{L^X}}$ was converted into $\overline{\overline{L_F^X}}$, and $\overline{\overline{L^J}}$ was converted into $\overline{\overline{L_F^J}}$ when each element of the matrices were multiplied by x_i / x_k and J_i / J_k, respectively, Finally using Eq. 3.39, the control coefficients for the independent fluxes and metabolites were calculated.

As it can be seen from Figure 4.23, the reactions v(14), v(15), v(16) and v(17) were found to exert more control on the metabolic insulin network. These reactions correspond to PI3K and Akt phosphorylation and they were found important for the metabolic pathway. These results are in good agreement with literature. The finding that inhibition of PI3K, or the overexpression of dominant negative mutants of this enzyme, causes also an inhibition of most of the cellular responses to insulin, including stimulation of glucose transport as well as glycogen and protein synthesis, established that $PI(3,4,5)P_3$ is a key second messenger in the insulin signaling pathway. A key effector of $PI(3,4,5)P_3$ in insulin signaling is Akt. The interaction of Akt with $PI(3,4,5)P_3$ at the membrane does not activate Akt directly. However, it activates phosphorylation of Akt. Once activated, Akt dissociates from the plasma membrane and phosphorylates numerous substrates in both the cytoplasm and the nucleus, which play important roles in regulating insulin dependent processes

(Lizcano and Alessi, 2002). Moreover, v(11) is the reaction of phosphorylation of IRS-1 (insulin receptor substrate-1). Consistent with literature, insulin receptor substrate proteins are linked to the activation of PI3K-Akt pathway, which is responsible for most of the metabolic actions of insulin (Taniguchi, 2006). The enzymes involved in this PI3K-Akt pathway may be suitable as targets for treatment of Type-2 diabetes. As explained in Section 2.2.4., PKC-ζ which has a role in this pathway, is a drug target, as increased muscle PKC-ζ activity has been observed in the context of fatty-acid-induced insulin resistance (Shulman, 2001; Moller, 2001). The flux control coefficient of each enzyme catalyzing the reactions in insulin metabolism is given in the Appendix part.

4.3.3 Metabolic Pathway Analysis: Elementary Flux Modes and Control Effective Fluxes

In the present work, the elementary flux modes (EFM) and control effective fluxes were determined for the insulin metabolic network (Sedaghat *et al.*, 2002) via using the simulation tool, CellNetAnalyzer / FluxAnalyzer V. 9.2 (Klamt *et al.*, 2003). Having 12 internal and 9 external metabolites, 21 elementary flux modes were found. Each elementary mode represents a biochemical function (Çakır *et al.*, 2004). Important EFMs sorted with respect to their end products are given in Table 4.9. 5 of the 21 elementary modes are associated with GLUT4 translocation. They start with INS and end at GLUT4. The other EFMs differ in terms of the membrane protein and the end product.

Control effective flux (CEF) analysis was also performed to represent the importance of each reaction in insulin metabolism through these modes. Insulin uptake reactions were found to be important by observing the changes on the metabolic pathway after insulin stimulation and GLUT4 reactions are the main reactions for this metabolic pathway. Therefore, in order to calculate CEF values, insulin uptake and GLUT4 reactions were selected as objective reactions.

According to CEF analysis, the reactions v(3), v(5) and v(14) were found to be important for the metabolic pathway (Table 4.8). The reactions of v(3) and v(5) correspond to the reactions of insulin receptors. This finding is consistent with the literature. Insulin

binding leads to a rapid cascade of autophosphorylation of the receptor itself. The insulin receptor provides an innovative and simple way to think about mediating control of various cellular processes by insulin (Myers and White, 1993). As explained in Section 4.3.2, the reaction of v(14) which is the phosphorylation reaction of $PI(4,5)P_2$ to generate $PI(3,4,5)P_3$. This event occurs within the first minute of insulin binding to its receptor and result in a significant increase in the concentration of $PI(3,4,5)P_3$ in cells. This reaction has a central role in insulin signaling. These intracellular reactions are pivotal in transmitting the signal from the receptor to the final cellular effect, GLUT4 translocation.

Table 4.8. CEF values of the important reactions in the metabolic pathway

Reactions	CEF values
$v(16)$	1.69
$v(17)$	1.69
$v(19)$	1.69
$v(11)$	1.76
$v(12)$	2.83
$v(14)$	3.89
$v(5)$	4.00
$v(3)$	4.10

Table 4.9. Elementary flux modes for insulin metabolic network

EFM number	Important EFMs sorted with respect to their end products
	INS → Cell surface GLUT4
1	2 v(1) 2 v(3) -2 v(5) 2 v(12) 2 v(13) 2 v(14) 1 v(16) 1 v(17) -1 v(18) 1 v(19)
2	2 v(1) 2 v(2) 2 v(3) -2 v(5) 2 v(11) 2 v(13) 2 v(14) 1 v(16) 1 v(17) -1 v(18) 1 v(19)
3	2 v(1) 2 v(2) 2 v(3) -2 v(5) 2 v(11) 2 v(13) 2 v(14) 1 v(16) 1 v(17) -1 v(18) 1 v(19)
	INS → Intracellular GLUT4
4	2 v(1) 2 v(2) 2 v(3) -2 v(5) 2 v(11) 2 v(13) 2 v(14) 1 v(16) 1 v(17) 1 v(19) 1 v(20)
5	2 v(1) 2 v(3) -2 v(5) 2 v(12) 2 v(13) 2 v(14) 1 v(16) 1 v(17) 1 v(19) 1 v(20)
	INS → Unbound surface insulin receptors
6	1 v(1) 1 v(3) 1 v(4)
	INS → Tyrosine phosphorylates IRS-1
7	1 v(2) 1 v(11) -1 v(12)
8	1 v(1) 1 v(3) -1 v(5) 1 v(6) 1 v(9) -1 v(11) 1 v(12)
	INS → $PI(3,4)P_2$
9	1 v(1) 1 v(2) 1 v(3) -1 v(5) 1 v(11) 1 v(13) 1 v(14) -1 v(15)
10	1 v(1) 1 v(3) -1 v(5) 1 v(12) 1 v(13) 1 v(14) -1 v(15)
11	1 v(2) 1 v(6) 1 v(9) -1 v(12) -1 v(13) -1 v(14) 1 v(15)
12	-1 v(2) 1 v(7) 1 v(10) -1 v(11) -1 v(13) -1 v(14) 1 v(15)
13	-1 v(2) 1 v(4) 1 v(5) -1 v(11) -1 v(13) -1 v(14) 1 v(15)
	$PI(3,4)P_2$ → Intracellular GLUT4
14	2 v(15) 1 v(16) 1 v(17) 1 v(19) 1 v(20)

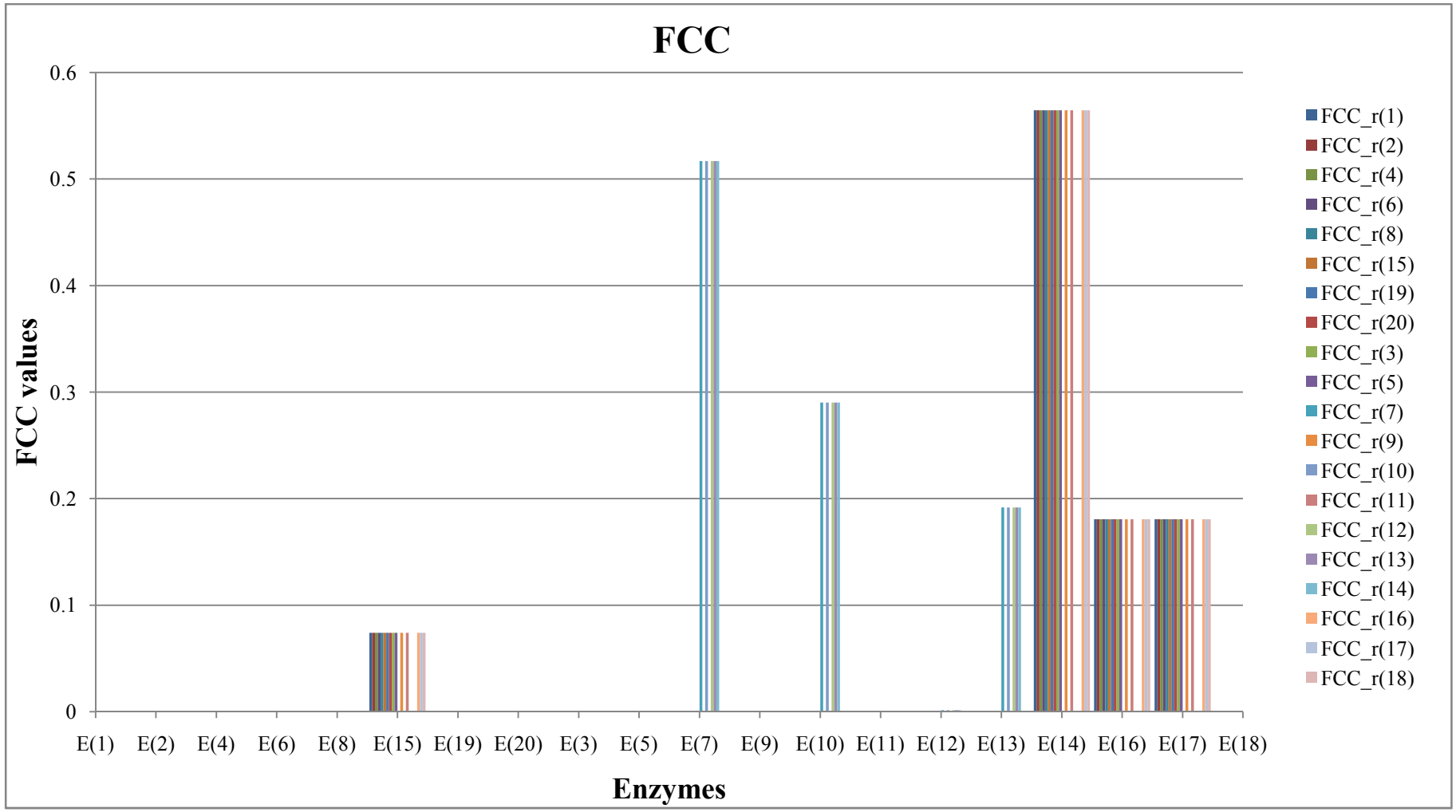


Figure 4.23. Flux control coefficients for metabolic insulin signaling network

4.4 General Aspects on Insulin Signaling and Metabolic Networks

4.4.1 Linear Path Analysis

In order to make a connection between the reconstructed insulin signaling network including 418 proteins and 914 interactions and the dynamic insulin metabolic network (Sedaghat *et al.*, 2002), the linear paths of the reconstructed network were found using NetSearch algorithm (Steffen *et al.*, 2002). The final version of the insulin signaling network with INS (insulin) and GLUT4 (glucose transporter 4) proteins as the ligand and the target, respectively, was used for the linear path calculations. Furthermore, as a first step, the optimal maximum path length (Steffen *et al.*, 2002) was determined examining the lengths connecting every possible pair of 418 proteins in the interaction dataset. The connectivity of each protein was analyzed and the average path length between any two proteins was found as 10.5. Therefore, in the NetSearch algorithm, the maximum path length was determined as 11. Using the algorithm, the reconstructed interaction network generated 2507 linear paths at eleven steps that begin at INS and end at GLUT4.

Figure 4.24 shows the exponential increase in the number of linear paths with maximum path length for the phenotype of GLUT4 translocation.

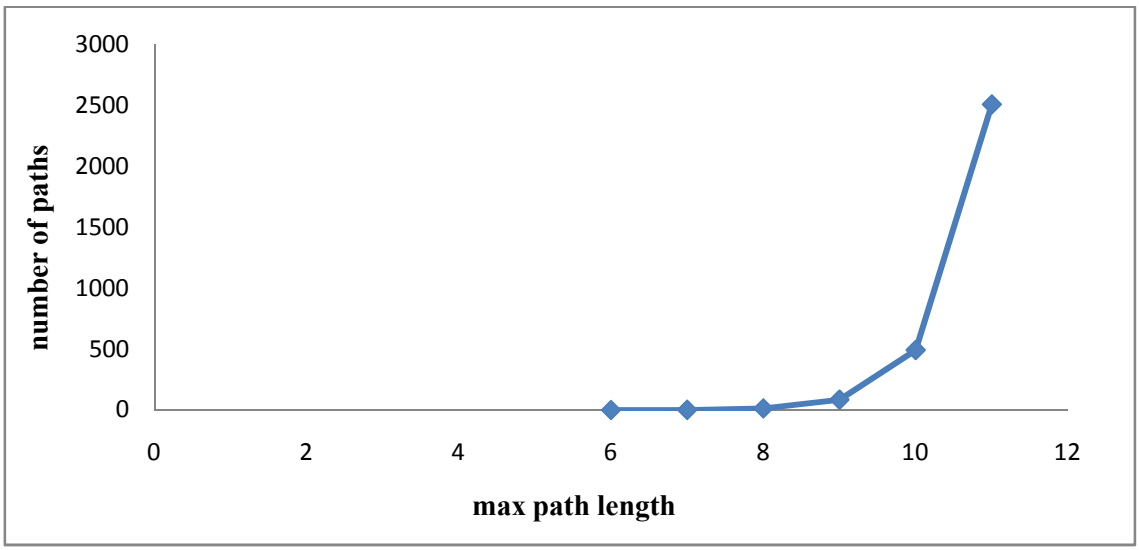


Figure 4.24. The exponential increase in maximum path length with respect to the number of paths

Using only the proteins involved in the linear paths from insulin (INS) to glucose transporter gene (GLUT4), the network was reduced to 102 proteins and 354 physical protein-protein interactions (Figure 4.25). In order to investigate whether the new smaller reconstructed network for GLUT4 translocation is biologically meaningful or not, graph theoretic analysis was again performed. It was concluded that this network yielded the common topological features of the biological networks, i.e. it has small world characteristics with scale free topology, and it has a power law degree distribution. Moreover, highly connected nodes were found, which are Rb1 and HDAC1, having connectivity 15 and 14, respectively. These findings are consistent with the information given in literature; Rb1 was reported to increase the phosphorylation of IRS-1 and Akt, and to stimulate PI3K activity in the absence of the activation of the insulin receptor. Rb1-induced glucose uptake as well as GLUT1 and GLUT4 translocation was inhibited by the PI3K inhibitor. These results suggest that Rb1 stimulates glucose transport in insulin-sensitive cells by promoting translocations of GLUT1 and GLUT4 by partially activating insulin signaling pathway (Shang *et al.*, 2008). Inhibition of HDAC was reported to induce cell cycle arrest, cell differentiation, and apoptotic cell death of transformed cells *in vivo* and *in vitro* which renders HDAC inhibitors as candidate pharmaceuticals in cancer treatment. The isoform HDAC1 caused GLUT4 content to increase and its translocation in muscle cells, resulting in stimulation of glucose uptake, independent of insulin action. The stimulation of glucose uptake and GLUT4 translocation is significantly greater than that of insulin, implying a novel potential for blood glucose control (Takigawa-Imamura *et al.*, 2003, Butler *et al.*, 2001, Su *et al.*, 2000).

Table 4.10. The number of linear paths with respect to path lengths

Path length	Passes threshold
6	0
7	1
8	13
9	85
10	492
11	2507

The proteins included in the linear paths with 7 and 8 steps were further investigated. The linear paths included IGF1R (insulin-like growth factor 1 receptor), Grb10 (growth factor receptor-bound protein 10), EGFR (epidermal growth factor receptor) different from highly connected nodes related to GLUT4 translocation network. The main signaling pathway for IGF1R-mediated protection from apoptosis rests on the activation of PI3K and Akt. PI3K activation then leads to the transduction of the functional effects of IGFs, such as enhanced glucose transport, enhanced cardiomyocyte contractility, and the inhibition of apoptosis by activating several downstream proteins and molecules. Grb10 is a Pleckstrin homology and Src homology 2 (SH2) domain-containing protein that binds to the tyrosine-phosphorylated insulin receptor in response to insulin stimulation (Langlais, 2004). Furthermore, overexpression of EGFR in adipocytes has demonstrated that multiple signaling pathways lead to GLUT4-mediated glucose uptake. These findings implicated pathways which function independently of IRS-1 phosphorylation and PI3K activation. Signaling molecules that are dependent on EGFR auto-phosphorylation may lie in the signaling pathway to glucose transport (Van Epps-Fung, 1998).

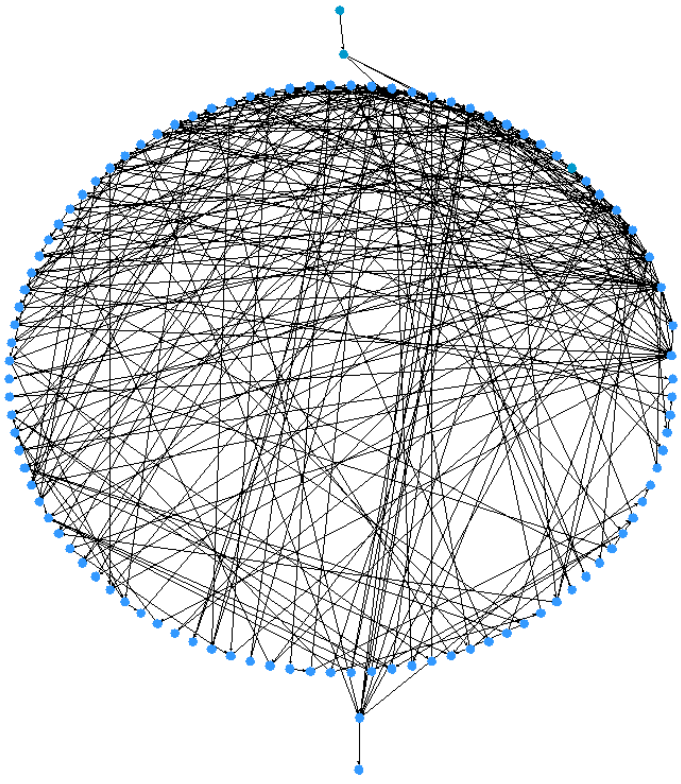


Figure 4.25. The reduced network for GLUT4 translocation drawn by Cytoscape

4.5 Common Species in Both Insulin Signaling Networks

The common molecules in both reconstructed insulin signaling network with GLUT4 translocation subnetwork (present study) and the insulin metabolic model (Sedaghat *et al.*, 2002) are shown in Figure 4.26 and explained with their functions in Table 4.11. The common molecules are actually critical nodes for the insulin signaling network. Each critical protein is essential for insulin action and has several isoforms with unique functions. IRS proteins mediate the binding of intracellular effectors. They share common mechanisms of regulation: they are activated by tyrosine phosphorylation, and they are negatively regulated by protein tyrosine phosphatases (PTPs), serine phosphorylation and ligand-induced downregulation. The isoform IRS-1 is found to regulate glucose uptake and GLUT4 translocation. Therefore, decreased hepatic IRS-1 correlates with the increased expression of genes that are involved in the gluconeogenesis (Taniguchi *et al.*, 2006). In addition to these, IRS-1 is an important member of the reconstructed subnetwork, as many of the linear paths overpass through IRS-1.

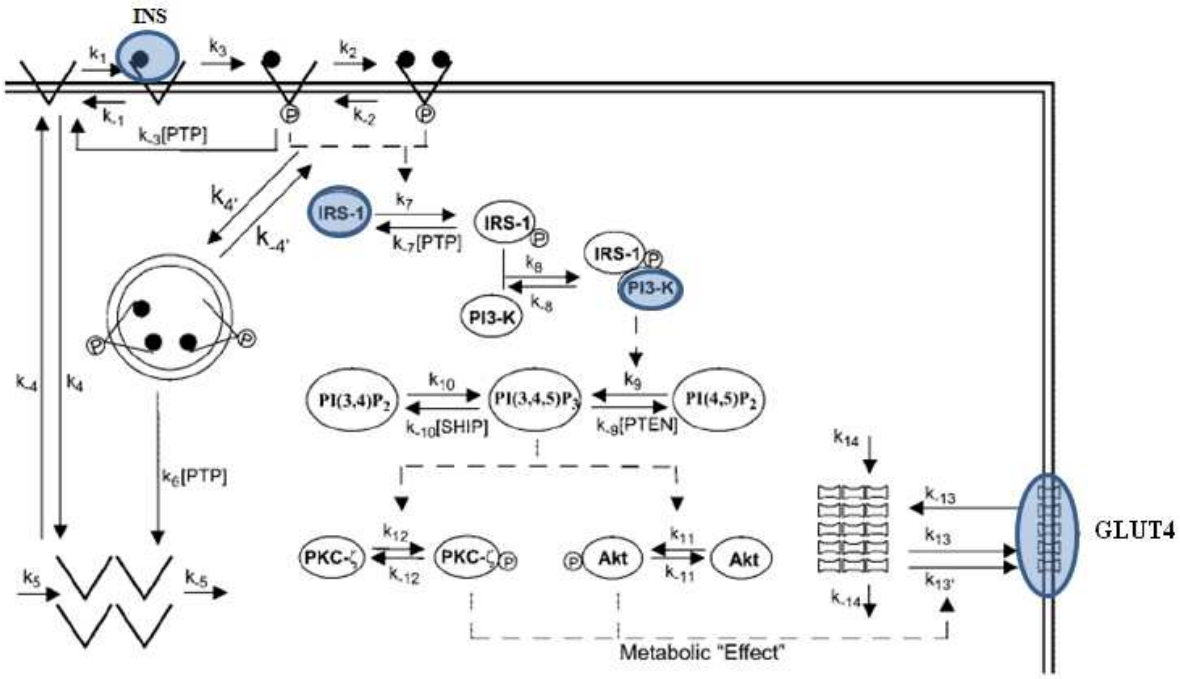


Figure 4.26. Common molecules in both signaling networks

Glucose uptake into cells is stimulated via insulin induction of GLUT4, i.e., translocation from intracellular storage to the plasma membrane. Furthermore, PI3K has

been shown to play a role in this GLUT4 translocation (Lizcano and Alessi, 2002). PI3K activates critical regulators of insulin signaling (Taniguchi *et al.*, 2006). The isoforms of PI3K are frequently observed to take place in the protein-protein interactions of the reconstructed subnetwork.

Table 4.11. Common molecules in both signaling networks

Name	Description	Function
INS	Insulin	➢ hormone activity ➢ insulin receptor binding ➢ insulin-like growth factor receptor binding
IRS-1	Insulin receptor substrate-1	➢ phosphoinositide 3-kinase binding ➢ signal transducer activity ➢ SH2 domain binding ➢ insulin-like growth factor receptor binding ➢ insulin receptor binding
PI3K	Phosphoinositide 3-kinase	➢ insulin receptor binding ➢ phosphoinositide 3-kinase regulator activity ➢ protein phosphatase binding ➢ insulin binding ➢ insulin-like growth factor receptor binding ➢ insulin receptor binding ➢ 1-phosphatidylinositol-3-kinase activity
GLUT4	Glucose transporter type 4	➢ glucose transmembrane transporter activity

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

- A dynamic model of insulin metabolism with 21 state variables (12 internal metabolites and 9 external metabolites) and 20 reactions was reconstructed using the ordinary differential equations given by Sedaghat *et al.*, 2002. This metabolic insulin network was further investigated by metabolic control and pathway analyses. In the simulations with insulin stimulation, it was observed that the concentration of state variables changed immediately after the removal of insulin. Thus, the insulin uptake and removal reactions were very important for the metabolic pathway.

- Through metabolic control analysis, PI3K and Akt phosphorylation were found to exert more control on the metabolic pathway. These results are also in good agreement with the information given in literature. Employing metabolic pathway analysis, that is elementary flux modes, the reactions of insulin receptor were found to be important for the metabolic insulin network.

- A protein-protein interaction network of insulin signaling was then reconstructed and analyzed in order to identify potential drug targets for the epidemic diabetes mellitus. The frequently observed domain pairs in the reconstructed insulin signaling network were SH2 (Src Homology 2)-SH2 and PH (Pleckstrin-homology domain)-PH domain-domain pairs.

- Comprehensive structural analysis of this reconstructed large scale network of insulin was performed using graph theoretical and pathway analysis techniques. The undirected interaction graph of insulin signaling network with 416 nodes (proteins) and 895 edges (protein-protein interactions) shows a small-world characteristics with a network diameter of 21 and a mean path length of 10.5.

- The connectivity distribution of the nodes follows nearly a power law model with γ=1.95 ($P(k) \approx k^{-\gamma}$) due to the scale-free nature of the insulin signaling network.

- In order to enlighten GLUT4 translocation subnetwork and connect the reconstructed signaling network with the metabolic model, the first interacting pair of GLUT4, i.e. DAXX, was added to the reconstructed network. The network consequently covered 418 proteins and 914 interactions between these proteins.

- The undirected interaction graph of the present reconstructed network with 418 nodes and 1828 edges yielded the common topological features of the biological networks, i.e. it has small world characteristics with scale free topology and has a power law degree distribution. The graph theoretic properties remain the same.

- Linear path analysis was performed through ligand to target protein in the reconstructed signaling network. Using the connectivity distribution of insulin signaling network, the network diameter of the directed graph was obtained as 11 which is the maximum value of the shortest path lengths from INS to GLUT4 protein. 2507 linear paths were found at eleven steps.

- The network was reduced to 102 proteins and 354 protein-protein interactions. In order to investigate whether the new smaller reconstructed network for GLUT4 translocation is biologically meaningful or not, the graph theoretic analysis was again performed. It was seen that this small network also yielded the common topological features of biological networks. Consistent with literature, two highly connected nodes were found, which are Rb1 and HDAC1.

- INS, GLUT4, PI3K and IRS-1 were identified as the common species in both insulin networks, the reconstructed signaling network with GLUT4 translocation subnetwork and the dynamic metabolic network.

- An important role in the insulin signaling pathway has recently been proposed for ceramide. It inhibits insulin-stimulated glucose uptake, GLUT4 translocation, and/or glycogen synthesis; however, it does not interfere at the level of IR or IRS-1

phosphorylation, but impairs insulin signaling by inhibition of Akt activation. The results of simulations (the time courses of activated and notactivated Akt) also confirmed that Akt is affected by ceramide.

- One of the approaches for the identification of putative drug targets for diabetes is to investigate the components of insulin signaling pathway. Accepted models are (i) to target the insulin receptor itself by inhibiting either the enzymes responsible for its deactivation or the downstream targets in the signaling pathway, e.g. specific tyrosine phosphatases such as PTP1B; (ii) to inhibit the enzyme glycogen synthase kinase-3 which has a negative effect on insulin and (iii) to inhibit the enzyme PKC-ζ which shows increased activity in insulin resistance. In agreement with these literature reports, the present study also indicates that the insulin receptor, PKC-ζ, PI3K and Akt have important roles in regulating insulin signaling. In addition to these molecules, the interacting protein partners GRB2 and PTPN1 are found to play critical roles as GRB2 protein is determined to be a hub with high connectivity and must therefore be assessed with special care in drug targeting.

- Moreover, the important reactions of the dynamic model were found as IRS-1, PI3K and Akt phosphorylation via metabolic control analysis and control effective flux analysis. The enzymes involved in the reactions are considered as suitable drug targets for the treatment of Type-2 diabetes.

5.2 Recommendations

- One of the aims of this study is to clarify the unknown topological features of the insulin signaling and GLUT4 translocation subnetworks. The results indicate the key points (crucial signaling molecules) in the signal transduction mechanisms. These may be considered as starting points for further experimental research.

- The protein DAXX with unknown function, but having physical interaction with GLUT4, is found to participate in GLUT4 translocation network. This result can be used as a starting point for annotating a function to this protein. Furthermore, EHD2 protein with

unknown function should also be investigated whether it has a role in GLUT4 translocation or not.

- The important reactions of the dynamic model were found via metabolic control analysis and control effective flux analysis. The enzymes involved in these reactions are suitable as targets for the treatment of Type-2 diabetes. Experimental work has to be carried out with these enzymes to confirm the computational results.

APPENDIX A: ANNOTATION-COLLECTION TABLES

Annotation-collection table is a table of Gene Ontology (GO) terms of insulin signaling network. It includes 231 GO terms (50 molecular function terms, 140 biological process terms and 31 cellular component terms).

A.1 Function

Table A.1. Molecular function terms of the annotation-collection table

1-phosphatidylinositol-3-kinase activity
3-phosphoadenosine 5-phosphosulfate binding
3-phosphoinositide-dependent protein kinase activity
actin binding
alkaline phosphatase activity
ATP binding
chromatin DNA binding
cytoskeletal adaptor activity
enzyme binding
epidermal growth factor receptor binding
ErbB-3 class receptor binding
GTP binding
GTPase activator activity
hormone activity
identical protein binding
insulin binding
insulin receptor activity
insulin receptor binding
insulin receptor substrate binding
insulin-like growth factor I binding

Table A.1. Molecular function terms of the annotation-collection table (*Continued*)

insulin-like growth factor II binding
insulin-like growth factor binding
insulin-like growth factor receptor activity
insulin-like growth factor receptor binding
kinase inhibitor activity
magnesium ion binding
nucleoside-triphosphate diphosphatase activity
nucleotide diphosphatase activity
phosphatidylinositol binding
phosphoinositide 3-kinase binding
phosphoinositide 3-kinase regulator activity
protein binding
protein C-terminus binding
protein homodimerization activity
protein kinase binding
protein kinase inhibitor activity
protein phosphatase binding
protein serine/threonine kinase activity
PTB domain binding
receptor signaling protein tyrosine kinase activity
sequence-specific DNA binding
SH2 domain binding
SH3/SH2 adaptor activity
signal transducer activity
transcription activator activity
transcription factor activity
transcription factor binding
transmembrane receptor protein tyrosine kinase docking protein activity

Table A.1. Molecular function terms of the annotation-collection table (*Continued*)

transporter activity

A.2 Process

Table A.2. Biological process terms of the annotation-collection table

3-phosphoadenosine 5-phosphosulfate metabolic process
actin cytoskeleton organization and biogenesis
actin filament organization
activated T cell apoptosis
activation of MAPK activity
activation of pro-apoptotic gene products
acute-phase response
alpha-beta T cell activation
anti-apoptosis
axonogenesis
B cell receptor signaling pathway
cell cycle arrest
cell death
cell division
cell proliferation
cell-cell signaling
cellular phosphate ion homeostasis
cellular response to insulin stimulus
cytokine and chemokine mediated signaling pathway
epidermal growth factor receptor signaling pathway

Table A.2. Biological process terms of the annotation-collection table (*Continued*)

fatty acid homeostasis
focal adhesion formation
G1 phase of mitotic cell cycle
generation of precursor metabolites and energy
genetic imprinting
glucose homeostasis
glucose transport
G-protein coupled receptor protein signaling pathway
heart development
heart morphogenesis
inorganic diphosphate transport
insulin receptor signaling pathway
insulin-like growth factor receptor signaling pathway
integrin-mediated signaling pathway
intracellular signaling cascade
lymphangiogenesis
mesoderm development
multicellular organismal development
muscle hypertrophy
myoblast differentiation
myoblast proliferation
myotube cell development
negative regulation of acute inflammatory response
negative regulation of angiogenesis
negative regulation of bone mineralization
negative regulation of cell growth
negative regulation of cell proliferation
negative regulation of fat cell differentiation
negative regulation of fatty acid metabolic process
negative regulation of glucose import
negative regulation of glycogen biosynthetic process

Table A.2. Biological process terms of the annotation-collection table (*Continued*)

negative regulation of glycogen catabolic process
negative regulation of insulin receptor signaling pathway
negative regulation of insulin-like growth factor receptor signaling pathway
negative regulation of JAK-STAT cascade
negative regulation of metabolic process
negative regulation of NAD(P)H oxidase activity
negative regulation of phosphoinositide 3-kinase cascade
negative regulation of protein amino acid autophosphorylation
negative regulation of protein catabolic process
negative regulation of protein kinase activity
negative regulation of protein kinase B signaling cascade
negative regulation of protein secretion
negative regulation of proteolysis
negative regulation of smooth muscle cell apoptosis
negative regulation of smooth muscle cell differentiation
negative regulation of smooth muscle cell migration
negative regulation of smooth muscle cell proliferation
negative regulation of tyrosine phosphorylation of Stat3 protein
negative regulation of vasolidation
nerve growth factor receptor signaling pathway
neural tube closure
nitric oxide biosynthetic process
nucleoside triphosphate catabolic process
peptidyl-serine phosphorylation
phosphate metabolic process
phosphoinositide 3-kinase cascade
phosphoinositide mediated signaling
phosphoinositide phosphorylation
pinocytosis
platelet-derived growth factor receptor signaling pathway
positive chemotaxis

Table A.2. Biological process terms of the annotation-collection table (*Continued*)

positive regulation of cell migration
positive regulation of cell poliferation
positive regulation of cellular protein metabolic process
positive regulation of cytokine secretion
positive regulation of developmental growth
positive regulation of DNA replication
positive regulation of epithelial cell proliferation
positive regulation of fibroblast proliferation
positive regulation of gene-specific transcription
positive regulation of glucose import
positive regulation of glycolysis
positive regulation of insulin receptor signaling pathway
positive regulation of insulin-like growth factor receptor signaling pathway
positive regulation of metabolic process
positive regulation of mitosis
positive regulation of nitric oxide biosynthetic process
positive regulation of nitric oxide synthase activity
positive regulation of phagocytosis
positive regulation of Ras protein signal transduction
positive regulation of smooth muscle cell migration
positive regulation of smooth muscle cell proliferation
positive regulation of tyrosine phosphorylation of Stat5 protein
positive regulation of vasodilation
protein amino acid autophosphorylation
protein amino acid phosphorylation
protein heterotetramerization
protein import into nucleus
protein kinase cascade
protein localization
protein modification process
protein tetramerization

Table A.2. Biological process terms of the annotation-collection table (*Continued*)

Ras protein signal transduction
regulation of amino acid metabolic process
regulation of apoptosis
regulation of bone mineralization
regulation of cell growth
regulation of embryonic development
regulation of endocytosis
regulation of gene-specific transcription
regulation of inflammatory response
regulation of insulin receptor signaling pathway
regulation of protein amino acid phosphorylation
regulation of protein localization
regulation of transcription, DNA-dependent
regulation of transmembrane transporter activity
respiratory burst
response to heat
response to hormone stimulus
satellite cell maintenance involved in skeletal muscle regeneration
sequestering of triacylglycerol
signal transduction
skeletal development
stress fiber formation
T cell receptor signaling pathway
transcription from RNA polymerase II promoter
transformation of host cell by virus
wound healing

A.3 Component

Table A.3. Cellular component terms of the annotation-collection table

1-phosphatidylinositol-4-phosphate 3-kinase, class IA complex
AP-type membrane coat adaptor complex
basolateral plasma membrane
caveola
cell surface
cell-substrate adherens junction
cytoplasm
cytosol
endosome membrane
extracellular matrix part
extracellular region
extracellular space
focal adhesion
Golgi apparatus
insulin-like growth factor binding protein complex
integral to membrane
integral to plasma membrane
membrane fraction
membrane raft
microsome
nucleoplasm
nucleus
perinuclear region of cytoplasm
plasma membrane
platelet alpha granule lumen
ruffle

APPENDIX B: METABOLIC INSULIN NETWORK MODEL

B.1 Flux Distributions

The ordinary differential equations of the dynamic model (Sedaghat *et al.*, 2002) explained in 3.5 were converted to flux distributions of each reaction for analyses.

Table B.1. Flux distributions of each reaction in the dynamic model

$$v(1) = k_1 x_1 x_2 - k_{m1} x_3$$
$$v(2) = k_2 x_1 x_5 - k_{m2} x_4$$
$$v(3) = k_3 x_3$$
$$v(4) = k_{m3} [PTP] x_5$$
$$v(5) = k_4 x_2 - k_{m4} x_6$$
$$v(6) = k_{4p} x_4 - k_{m4p} x_7$$
$$v(7) = k_{4p} x_5 - k_{m4p} x_8$$
$$v(8) = k_5 - k_{m5} x_6$$
$$v(9) = k_6 [PTP] x_7$$
$$v(10) = k_6 [PTP] x_8$$
$$v(11) = \frac{k_7}{IRp} x_4 x_9 - k_{m7} [PTP] x_{10}$$
$$v(12) = \frac{k_7}{IRp} x_5 x_9 - k_{m7} [PTP] x_{10}$$
$$v(13) = k_8 x_{11} x_{10} - k_{m8} x_{12}$$
$$v(14) = k_9 x_{14} - k_{m9} [PTEN] x_{13}$$
$$v(15) = k_{10} x_{15} - k_{m10} [SHIP] x_{13}$$
$$v(16) = k_{11} x_{16} - k_{m11} x_{17}$$
$$v(17) = k_{12} x_{18} - k_{m12} x_{19}$$
$$v(18) = k_{13} x_{20} - k_{m13} x_{21}$$
$$v(19) = k_{13p} x_{20}$$
$$v(20) = k_{14} - k_{m14} x_{20}$$

B.2 Model Parameters

Model parameters and initial conditions of each state variable were taken from the work of Sedaghat *et al.*, 2002 (Table B.2 and B.3). The other initial conditions which are not given in Table B.3 are 0.

Table B.2. Model parameters of the metabolic insulin network

$k_1 = 6\times10^7\ M^{-1}\ \min^{-1}$
$k_{-1} = 0.20\ \min^{-1}$
$k_2 = k_1$
$k_{-2} = 100k_{-1}$
$k_3 = 2500\ \min^{-1}$
$k_{-3} = k_{-1}$
$k_4 = k_{-4}/9$
$k_{-4} = 0.003\ \min^{-1}$
$k_{4'} = 2.1\times10^{-3}\ \min^{-1}$
$k_{-4'} = 2.1\times10^{-4}\ \min^{-1}$
$k_5 = 10k_{-5}M\ \min^{-1}\ if\ (x_6+x_7+x_8)<1\times10^{-13}$
$\quad = 60k_{-5}M\ \min^{-1}\ if\ (x_6+x_7+x_8)\le1\times10^{-13}$
$k_{-5} = 1.67\times10^{-18}$
$k_6 = 0.461\min^{-1}$
$k_7 = 4.16\min^{-1}$
$k_{-7} = (2.5/7.45)k_7$
$k_8 = k_{-8}(5/70.775)\times10_{12}$
$k_8 = 10\min^{-1}$
$k_9 = (k_{9(stimulated)} - k_{9basal})(x_{12}/PI3K) + k_{9basal}$
$k_{9(stimulated)} = 1.39\min^{-1}$
$k_{-9} = (94/3.1)k_{9(stimulated)}$
$k_{9(basal)} = (0.31/99.4)k_{-9}$
$k_{10} = (3.1/2.9)k_{-10}$
$k_{-10} = 2.77\min^{-1}$
$k_{11} = (0.1k_{-11})(x_{13}-0.31)/(3.10-0.31)$
$k_{-11} = 10\ln(2)\min^{-1}$
$k_{12} = (0.1k_{-12})(x_{13}-0.31)/(3.10-0.31)$
$k_{-12} = 10\ln(2)\min^{-1}$
$k_{-13} = 0.167\min^{-1}$
$k_{13} = (4/96)k_{-13}$
$k_{13'} = [(40/60)-(4/96)]k_{-13}(Effect)$
$k_{14} = 96k_{-14}$
$k_{-14} = 0.001155\min^{-1}$
$Effect = (0.2x_{17}+0.8x_{19})/(AP_{equil})$
$IRp = 8.97\times10^{-13}\ M$
$[SHIP] = 1.00,\ [PTEN] = 1.00,\ [PTP] = 1.00$
$AP_{equil} = 100/11,\ PI3K = 5\times10^{-15}\ M$

Table B.3. Initial conditions of state variables

$x_2(0) = 9\times10^{-13}\ M$
$x_6(0) = 1\times10^{-13}\ M$
$x_9(0) = 1\times10^{-12}\ M$
$x_{11}(0) = 1\times10^{-13}\ M$
$x_{13}(0) = 0.31\%$
$x_{14}(0) = 99.24\%$
$x_{15}(0) = 0.29\%$
$x_{16}(0) = 100\%$
$x_{18}(0) = 100\%$
$x_{20}(0) = 96\%$
$x_{21}(0) = 4\%$

APPENDIX C: MATLAB CODES

A Matlab code was written for the analysis of dynamic model, stimulation of insulin and metabolic control analysis.

C.1 MATLAB Code for Insulin Stimulation

insulin

```
clear all
clc

tspan=[0 60];
 x0=[9E-13 0 0 0 ...
    1E-13 0 0 1E-12 0 ...
    1E-13 0 0.31 99.4 0.29 ...
    100 0 100 0 96 4];
[t,x] = ode15s('insulinsub',tspan,x0);

 % Figures

figure
plot(t,x(:,1),'r-')
xlabel('Time (min)');
ylabel('Concentration (M)')
title('Concentration of free surface receptors (x2) vs time')
axis([-5 60 0 1E-12])

figure
plot(t,x(:,3))
xlabel('Time (min)');
```

```
ylabel('Concentration (M)')
title('Conc. of once and twice bound phosphorylated receptors (x5&x4) vs time','FontSize',8);
hold on
plot(t,x(:,4),'-r');
axis([-5 60 0 1E-12])
 figure
a(:,1)=x(:,3)+x(:,4);
plot(t,a(:,1),'r-')
title('Concentration of surface phosphorylated receptors (x4+x5) vs time','FontSize',8)
xlabel('Time (min)')
ylabel('Concentration (M)')
axis([-5 60 0 1E-12])

figure
plot(t,x(:,8))
hold on
plot(t,x(:,9),'r-')
title('Concentration of unphosphorylated IRS-1 (x9) and tyrosine phosphorylated IRS-1 (x10) vs time')
xlabel('Time (min)')
ylabel('Concentration (M)')
axis([-5 60 0 1e-12])

figure
plot(t,x(:,11));
title('x12')
title('Concentration of activated PI3K (x12) vs time')
xlabel('Time (min)')
ylabel('Concentration (M)')
axis([-5 60 0 6E-15])

figure
```

```
plot(t,x(:,12));
hold on
plot(t,x(:,14),'r-');
title('x13 and x15')
title('Percentage of activated PI(3,4,5) (x13) and PI(3,4)P2 (x15) vs time')
xlabel('Time (min)')
ylabel('Percentage')
axis([-5 60 0 4])

figure
plot(t,x(:,18),'r-')
title('Percentage of activated PKC-lamda (x19) vs time')
xlabel('Time (min)')
ylabel('Percentage')
axis([-5 60 0 10])


figure
plot(t,x(:,20),'r-')
title('Percentage of surface GLUT4 (x21) vs time')
xlabel('Time (min)')
ylabel('Percentage')
axis([-5 100 0 45])
```

insulin.sub

```
function  der=der(t,x)

% Model parameters
% b: backward
% i: prime
% s: stimulated
% ba: basal
```

```
k1=6E7;
km1=0.20;
k2=6E7;
km2=100*0.20;
k3=2500.0;
km3=0.20;
km4=0.003;
k4=0.003/9;
k4p=2.1E-3;
km4p=2.1E-4;
km5=1.66667E-18;
if (x(5)+x(6)+x(7)>1E-13)
   k5=10*1.66667E-18;
else
   k5=60*1.66667E-18;
end

PI3K=5.0E-15;
k6=0.461;
k7=4.16;
km7=(2.5/7.45)*4.16;
km8=10;
k8=5E-15/(7.45E-13*9.5E-14)*10;
k9stim=1.39;
km9=(94/3.1)*1.39;
k9basal=(0.31/99.4)*(94/3.1)*1.39;
k9=((k9stim-k9basal)*(x(11)/(PI3K))+k9basal);
km10=2.77;
k10=(3.1/2.9)*2.77;
km11=6.93;
k11=(0.1*6.93)*((x(12)-0.31)/2.79);
km12=6.93;
k12=(0.1*6.93)*((x(12)-0.31)/2.79);
```

```
km13=0.167;
k13=(4/96)*0.167;

APequil=100/11;
effect=min((0.2*x(16)+0.8*x(18))/(APequil),1);

IRp=8.97E-13;
SHIP=1;
PTEN=1;
PTP=1;
k13p=(40/60 - 4/96)*0.167*effect;
km14=0.001155;
k14=96*0.001155;

if(t<15)
   z=1E-7;
else
   z=0;
end


% State variables

% x1: insulin input
% x2: concentration of unbound surface insulin receptors
% x3: concentrattion of unphosphorylated once-bound surface receptors
% x4: concentration of phosphorylated twice-bound surface receptors
% x5: concentration of phoshorylated once-bound surface receptors
% x6: concentration of unbound unphosphorylated intracellular receptors
% x7: concentration of phosphorylated twice-bound intracellular receptors
% x8: concentration of phoshorylated once-bound intracellular receptors
% x9: concentration of unphosphorylated IRS-1
% x10: concentration of tyrosine-phosphorylated IRS-1
```

```
% x11: concentration of unactivated PI-3-kinase
% x12: concentration of tyrosine-phoshorylated IRS-1/activated PI 3-
% kinase complex
% x13: percentage of PI(3,4,5)P3 out of the total lipid population
% x14: percentage of PI(4,5)P2 out of the total lipid population
% x15: percentage of PI(3,4)P2 out of the total lipid population
% x16: percentage of unactivated Akt
% x17: percentage of activated Akt
% x18: percentage of unactivated PKC-lamda
% x19: percentage of activated PKC-lamda
% x20: percentage of intracellular GLUT4
% x21: percentage of cell surface GLUT4
  % Differential Equations

der(1)=km1*x(2)+km3*(PTP)*x(4)-k1*z*x(1)+km4*x(5)-k4*x(1);
der(2)=k1*z*x(1)-km1*x(2)-k3*x(2);
der(3)=k2*z*x(4)-km2*x(3)+km4p*x(6)-k4p*x(3);
der(4)=k3*x(2)+km2*x(3)-k2*z*x(4)-km3*(PTP)*x(4)+km4p*x(7)-k4p*x(4);
der(5)=k5-km5*x(5)+k6*(PTP)*(x(6)+x(7))+k4*x(1)-km4*x(5);
der(6)=k4p*x(3)-km4p*x(6)-k6*(PTP)*x(6);
der(7)=k4p*x(4)-km4p*x(7)-k6*(PTP)*x(7);
der(8)=km7*(PTP)*x(9)-k7*x(8)*(x(3)+x(4))/(IRp);
der(9)=k7*x(8)*(x(3)+x(4))/(IRp)+km8*x(11)-(km7*(PTP)+k8*x(10))*x(9);
der(10)=km8*x(11)-k8*x(9)*x(10);
der(11)=k8*x(9)*x(10)-km8*x(11);
der(12)=k9*x(13)+k10*x(14)-(km9*(PTEN)+km10*(SHIP))*x(12);
der(13)=km9*(PTEN)*x(12)-k9*x(13);
der(14)=km10*(SHIP)*x(12)-k10*x(14);
der(15)=km11*x(16)-k11*x(15);
der(16)=k11*x(15)-km11*x(16);
der(17)=km12*x(18)-k12*x(17);
der(18)=k12*x(17)-km12*x(18);
der(19)=km13*x(20)-(k13+k13p)*x(19)+k14-km14*x(19);
```

```
der(20)=(k13+k13p)*x(19)-km13*x(20);

der=der';
```

C.2 MATLAB Code for Metabolic Control Analysis

```
clear all
clc

tspan=(0:+10:60);
x0=[9E-13 0 0 0 ...
    1E-13 0 0 1E-12 0 ...
    1E-13 0 0.31 99.4 0.29 ...
    100 0 100 0 96 4];
[t,x] = ode15s('insulinsub',tspan,x0);

x1=x(7,:);

k1=6E7;
km1=0.20;
k2=6E7;
km2=100*0.20;
k3=2500.0;
km3=0.20;
km4=0.003;
k4=0.003/9;
k4p=2.1E-3;
km4p=2.1E-4;
km5=1.66667E-18;
if (x1(5)+x1(6)+x1(7)>1E-13)
    k5=10*1.66667E-18;
else
```

```
   k5=60*1.66667E-18;
end

PI3K=5.0E-15;
k6=0.461;
k7=4.16;
km7=(2.5/7.45)*4.16;
km8=10;
k8=5E-15/(7.45E-13*9.5E-14)*10;
k9stim=1.39;
km9=(94/3.1)*1.39;
k9basal=(0.31/99.4)*(94/3.1)*1.39;

k9=((k9stim-k9basal)*(x1(11)/(PI3K))+k9basal);

km10=2.77;
k10=(3.1/2.9)*2.77;
km11=6.93;
k11=(0.1*6.93)*((x1(12)-0.31)/2.79);
km12=6.93;
k12=(0.1*6.93)*((x1(12)-0.31)/2.79);
km13=0.167;
k13=(4/96)*0.167;
APequil=100/11;
effect=min((0.2*x1(16)+0.8*x1(18))/(APequil),1);
IRp=8.97E-13;
SHIP=1;
PTEN=1;
PTP=1;
k13p=(40/60 - 4/96)*0.167*effect;
km14=0.001155;
k14=96*0.001155;
if(t<15)
```

```
    z=1E-7;
else
    z=0;
end

v(1)=k1*z*x1(1)-km1*x1(2);
v(2)=k2*z*x1(4)-km2*x1(3);
v(3)=k3*x1(2);
v(4)=km3*(PTP)*x1(4);
v(5)=k4*x1(1)-km4*x1(5);
v(6)=k4p*x1(3)-km4p*x1(6);
v(7)=k4p*x1(4)-km4p*x1(7);
v(8)=k5-km5*x1(5);
v(9)=k6*(PTP)*x1(6);
v(10)=k6*(PTP)*x1(7);
v(11)=k7*x1(8)*x1(3)/(IRp)-km7*(PTP)*x1(9);
v(12)=k7*x1(8)*x1(4)/(IRp)-km7*(PTP)*x1(9);
v(13)=k8*x1(10)*x1(9)-km8*x1(11);
v(14)=k9*x1(13)-km9*(PTEN)*x1(12);
v(15)=k10*x1(14)-km10*(SHIP)*x1(12);
v(16)=k11*x1(15)-km11*x1(16);
v(17)=k12*x1(17)-km12*x1(18);
v(18)=k13*x1(19)-km13*x1(20);
v(19)=k13p*x1(19);
v(20)=k14-km14*x1(19);

e=zeros(20,12);

e(1,1)=z*x1(1)*k1/v(1);
e(1,2)=-x1(2)*km1/v(1);
e(2,3)=-x1(3)*km2/v(2);
e(2,4)=z*x1(4)*k2/v(2);
e(3,4)=1;
```

```
e(4,3)=k4p*x1(3)/v(6);
e(4,5)=-km4p*x1(6)/v(6);
e(6,9)=-km10*(SHIP)*x1(12)/v(15);
e(7,12)=1;
e(8,12)=-km14*x1(19)/v(20);
e(9,2)=1;
e(10,1)=k4*x1(1)/v(5);
e(11,4)=k4p*x1(4)/v(7);
e(11,6)=-km4p*x1(7)/v(7);
e(12,5)=1;
e(13,6)=1;
e(14,3)=k7*x1(8)*x1(3)/(IRp)/v(11);
e(14,7)=-km7*(PTP)*x1(9)/v(11);
e(15,4)=k7*x1(8)*x1(4)/(IRp)/v(12);
e(15,7)=-km7*(PTP)*x1(9)/v(12);
e(16,7)=k8*x1(10)*x1(9)/v(13);
e(16,8)=-km8*x1(11)/v(13);
e(17,9)=-km9*(PTEN)*x1(12)/v(14);
e(18,10)=-km11*x1(16)/v(16);
e(18,9)=0.1*6.93/2.79*x1(12)*x1(15)/v(16);
e(19,9)=0.1*6.93/2.79*x1(12)*x1(17)/v(17);
e(19,11)=-km12*x1(18)/v(17);
e(20,12)=k13*x1(19)/v(18);

if (0.2*x1(16)+0.8*x1(18))/(APequil)<= 1
   e(7,10)=(40/60 - 4/96)*0.167*0.2/APequil*x1(16)*x1(19)/v(19);
   e(7,11)=(40/60 - 4/96)*0.167*0.8/APequil*x1(18)*x1(19)/v(19);
else
   e(7,10)=0;
   e(7,11)=0;
end
 e(17,8)=((k9stim-k9basal)/(PI3K))*x1(13)*x1(11)/v(14);
N=[-1   0   1   0   0   0   0   0   0   -1  0   0   0   0   0   0   0   0   0   0
```

```
1  0  0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0  0  0  0
0  1  0  -1 0  0  0  0  0  0  0  0  0  -1 0  0  0  0  0  0
0  -1 -1 0  0  0  0  0  1  0  -1 0  0  0  -1 0  0  0  0  0
0  0  0  1  0  0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  1  0  -1 0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0  1  1  -1 0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  -1 0  0  0
0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  1  -1 -1 0
0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0  0  0  1  0  0
0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0  0  0  0  1  0
0  0  0  0  0  0  -1 1  0  0  0  0  0  0  0  0  0  0  0  -1
];

Nr=[-1 0  1  0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0  0  0
1  0  0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0  0  0  0
0  1  0  -1 0  0  0  0  0  0  0  0  0  -1 0  0  0  0  0  0
0  -1 -1 0  0  0  0  0  1  0  -1 0  0  0  -1 0  0  0  0  0
0  0  0  1  0  0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  1  0  -1 0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0  1  1  -1 0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  -1 0  0  0
0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  1  -1 -1 0
0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0  0  0  1  0  0
0  0  0  0  0  0  -1 0  0  0  0  0  0  0  0  0  0  0  1  0
0  0  0  0  0  0  -1 1  0  0  0  0  0  0  0  0  0  0  0  -1
];

Nc=[0  -1 0  0  0  0  0  0  0  0  0  0
-1 0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  -1 0  0  0  0  0  0
1  0  -1 0  0  0  -1 0  0  0  0  0
0  0  0  -1 0  0  0  0  0  0  0  0
0  0  1  0  -1 0  0  0  0  0  0  0
```

```
0  0  0  0  0  1  1  -1 0  0  0  0
0  0  0  0  0  0  0  1  -1 0  0  0
0  0  0  0  0  0  0  0  1  -1 -1 0
0  0  0  0  0  0  0  0  0  1  0  0
0  0  0  0  0  0  0  0  0  0  1  0
0  0  0  0  0  0  0  0  0  0  0  -1
];

Nrc=[0  -1  0  0  0  0  0  0  0  0  0  0
-1  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  -1 0  0  0  0  0  0
1  0  -1 0  0  0  -1 0  0  0  0  0
0  0  0  -1 0  0  0  0  0  0  0  0
0  0  1  0  -1 0  0  0  0  0  0  0
0  0  0  0  0  1  1  -1 0  0  0  0
0  0  0  0  0  0  0  1  -1 0  0  0
0  0  0  0  0  0  0  0  1  -1 -1 0
0  0  0  0  0  0  0  0  0  1  0  0
0  0  0  0  0  0  0  0  0  0  1  0
0  0  0  0  0  0  0  0  0  0  0  -1
];

No=[-1  0  1  0  0  0  0  0
1  0  0  0  0  0  0  0
0  1  0  -1 0  0  0  0
0  -1 -1 0  0  0  0  0
0  0  0  1  0  0  0  0
0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0
0  0  0  0  0  1  0  0
0  0  0  0  0  0  -1 0
0  0  0  0  0  0  -1 0
```

```
0   0   0   0   0   0   -1  1
];
 Xr=[x(1);x(2);x(3);x(4);x(6);x(7);x(9);x(11);x(12);x(16);x(18);x(19)];

Lx=Nc*inv(Nrc);
Lj=[eye(8); -inv(Nrc)*No];

 for i=1:12
    for j=1:12
       LFx(i,j)=Lx(i,j)*Xr(j)/Xr(i);
    end
 end

Jr=[v(1);  v(2);  v(4);  v(6);  v(8);  v(15);  v(19);  v(20)];
Jo=[v(3);  v(5);  v(7);  v(9);  v(10);  v(11);  v(12);  v(13);  v(14);  v(16);  v(17);  v(18)];
J=[Jr ; Jo];

 for i=1:20
    for j=1:8
       LFj(i,j)=Lj(i,j)*J(j)/J(i);
    end
 end

CLK=pinv([LFj, -e*LFx]);

CRj=abs([CLK(1,:);CLK(2,:);CLK(3,:);CLK(4,:);CLK(5,:);CLK(6,:);CLK(7,:);CLK(8,:)])
;
CRx=[CLK(9,:);CLK(10,:);CLK(11,:);CLK(12,:);CLK(13,:);CLK(14,:);CLK(15,:);
   CLK(16,:);CLK(17,:);CLK(18,:);CLK(19,:);CLK(20,:)];

Cj=abs(LFj*CRj);
Cx=LFx*CRx;
for i=1:12
```

```
    D(i)=0;
    for j=1:20
        D(i)=D(i)+Cx(i,j);
    end
end

for i=1:20
    B(i)=0;
    for j=1:20
        B(i)=B(i)+Cj(i,j);
    end
    Cj_final(i,:)=Cj(i,:)/B(i);
end


format('long','e')
```

APPENDIX D: FLUX CONTROL COEFFICIENTS

Table D.1. Flux control coefficients of the metabolic pathway

	$E(1)$	$E(2)$	$E(4)$	$E(6)$	$E(8)$	$E(15)$	$E(19)$
$v(1)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(2)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(4)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(6)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(8)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(15)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(19)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(20)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(3)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(5)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08

Table D.1. Flux control coefficients of the metabolic pathway (*Continued*)

	$E(20)$	$E(3)$	$E(5)$	$E(7)$	$E(9)$	$E(10)$	$E(11)$
$v(1)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(2)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(4)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(6)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(8)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(15)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(19)$	1.86E-14	1.71E-12	5.47E-13	4.78E-06	5.39E-26	2.69E-06	2.49E-12
$v(20)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(3)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
$v(5)$	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12

Table D.1. Flux control coefficients of the metabolic pathway (*Continued*)

	$E(12)$	$E(13)$	$E(14)$	$E(16)$	$E(17)$	$E(18)$	sum
$v(1)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(2)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(4)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(6)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(8)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(15)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(19)$	1.27E-08	1.11E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(20)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(3)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
$v(5)$	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00

Table D.1. Flux control coefficients of the metabolic pathway (*Continued*)

	$E(1)$	$E(2)$	$E(4)$	$E(6)$	$E(8)$	$E(15)$	$E(19)$
$v(7)$	2.16E-18	1.08E-17	1.16E-17	1.70E-24	6.56E-22	4.67E-07	1.73E-13
$v(9)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(10)$	2.16E-18	1.08E-17	1.16E-17	1.70E-24	6.56E-22	4.67E-07	1.73E-13
$v(11)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(12)$	2.16E-18	1.08E-17	1.16E-17	1.70E-24	6.56E-22	4.67E-07	1.73E-13
$v(13)$	2.16E-18	1.08E-17	1.16E-17	1.70E-24	6.56E-22	4.67E-07	1.73E-13
$v(14)$	2.16E-18	1.08E-17	1.16E-17	1.70E-24	6.56E-22	4.67E-07	1.73E-13
$v(16)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(17)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08
$v(18)$	1.70E-16	1.72E-12	1.83E-12	2.69E-19	1.04E-16	7.41E-02	2.74E-08

Table D.1. Flux control coefficients of the metabolic pathway (*Continued*)

	E(20)	*E*(3)	*E*(5)	*E*(7)	*E*(9)	*E*(10)	*E*(11)
v(7)	1.18E-19	8.64E-18	3.45E-18	5.17E-01	3.40E-31	2.90E-01	1.57E-17
v(9)	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
v(10)	1.18E-19	8.64E-18	3.45E-18	5.17E-01	3.40E-31	2.90E-01	1.57E-17
v(11)	1.86E-14	1.71E-12	5.47E-13	3.27E-06	5.39E-26	1.84E-06	2.49E-12
v(12)	1.18E-19	8.64E-18	3.45E-18	5.17E-01	3.40E-31	2.90E-01	1.57E-17
v(13)	1.18E-19	8.64E-18	3.45E-18	5.17E-01	3.40E-31	2.90E-01	1.57E-17
v(14)	1.18E-19	8.64E-18	3.45E-18	5.17E-01	3.40E-31	2.90E-01	1.57E-17
v(16)	1.86E-14	1.71E-12	5.47E-13	4.78E-06	5.39E-26	2.69E-06	2.49E-12
v(17)	1.86E-14	1.71E-12	5.47E-13	4.78E-06	5.39E-26	2.69E-06	2.49E-12
v(18)	1.86E-14	1.71E-12	5.47E-13	4.78E-06	5.39E-26	2.69E-06	2.49E-12

Table D.1. Flux control coefficients of the metabolic pathway (*Continued*)

	E(12)	*E*(13)	*E*(14)	*E*(16)	*E*(17)	*E*(18)	sum
v(7)	1.37E-03	1.92E-01	3.56E-06	1.14E-06	1.14E-06	2.72E-18	1.00E+00
v(9)	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
v(10)	1.37E-03	1.92E-01	3.56E-06	1.14E-06	1.14E-06	2.72E-18	1.00E+00
v(11)	8.70E-09	1.16E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
v(12)	1.37E-03	1.92E-01	3.56E-06	1.14E-06	1.14E-06	2.72E-18	1.00E+00
v(13)	1.37E-03	1.92E-01	3.56E-06	1.14E-06	1.14E-06	2.72E-18	1.00E+00
v(14)	1.37E-03	1.92E-01	3.56E-06	1.14E-06	1.14E-06	2.72E-18	1.00E+00
v(16)	1.27E-08	1.11E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
v(17)	1.27E-08	1.11E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00
v(18)	1.27E-08	1.11E-05	5.65E-01	1.81E-01	1.81E-01	2.95E-13	1.00E+00

REFERENCES

Albert, R. and A. L. Barabasi, 2002, "Statistical mechanistic of complex networks", *Review of Modern Physics*, Vol. 74, pp. 47-97.

Albert, R., 2005, "Scale-free networks in cell biology", *Journal of Cell Science*, Vol. 118, pp. 4947-4957.

Albert, R., H. Jeong, A. L. Barabasi, 1999, "Diameter of the world-wide web", *Nature*, Vol. 401, pp. 130-131.

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, 2002, *Molecular Biology of the Cell*, 4th edition, Ch.15, Garland Science, New York.

Alfarano, C., C. E. Andrade, K. Anthony, N. Bahroos, M. Bajec, K. Bantoft, D. Betel, B. Bobechko, K. Boutilier, E. Burgess, 2005, " The biomolecular interaction network database and related tools 2005 update", *Nucleic Acids Res*, Vol. 33, pp. 418-424.

Amaral, L., A. N. Scala, M. Barthelemy, H. E. Stanley, 2000, "Classes of behavior of small-world networks", *PNAS*, Vol. 97, pp. 11149–11152, October.

Apic, G., J. Gough, S. A. Teichmann, 2001, "Domain combinations in archael, eubacterial and eukaryotic proteomes, *J Mol Biol*, Vol. 310, pp. 311-325.

Arga, K. Y., Z. İ. Önsan, B. Kırdar, K. Ö. Ülgen, J. Nielsen, 2007, "Understanding signaling in yeast: Insights from network analysis", *Biotechnology and Bioengineering*, Vol. 97, pp. 1246-1258, August.

Backer, J. M., M. G. Myers, S. E. Shoelson, D. J. Chin, X. J. Sun, M. Miralpeix, B. Margolis, E. Y. Skolnik, J. Schlessinger, M. F. White, 1992, "Phosphatidylinositol 3'-kinase is activated by association with IRS-1 during insulin stimulation", *EMBO J*, Vol. 11, pp. 3469-3479.

Bader, J., A. Chaudhuri, J. Chant, 2004, "Gaining confidence in high-throughput protein interaction networks", *Nat Biotechnol*, Vol. 22, pp. 78-85.

Bailey, J. E., 2001, "Complex biology with no parameters", *Nature Biotechnol.*, Vol. 19, pp. 503–504.

Barabasi A. L., H. Jeong, Z. Neda, E. Revasz, A. Schubert, T. Vicsek, 2002, "Evolution of the social network of scientific collaborations" *Physica A.* Vol. 311, pp. 590-614.

Barabasi, A. and R. Albert, 1999, "Emergence of scaling in random networks", *Science*, Vol. 286, pp. 509-512.

Barabasi, A. L. and Z. N. Oltvai, 2004, "Network biology: Understanding the cell's functional organization", *Nature Reviews Genetics*, Vol. 5, pp. 101-113, February.

Bhalla, U. S. and R. Iyengar, 1999, "Emergent properties of networks of biological signaling pathways" *Science*, Vol. 283, No. 5400, pp. 381-387, January.

Birnhaum, M. J., 1989, "Identification of a novel gene encoding an insulin-responsive glucose transporter protein", *Cell*, Vol. 57, pp. 305-315.

Bock, J. R. and D. A. Gough, 2001, "Predicting protein-protein interactions from primary structure", *Bioinformatics*, Vol. 17, pp. 455-460.

Bollobas, B., 1985, *Random Graphs*, Academic, London.

Bomsztyk, K., T. H. Stanton, L. L. Smith, N. A. Rachie, S. K Dower, 1989, "Properties of interleukin-1 and interferon-γ receptors in B lymphoid cell line", *J. Biol. Chem.*, Vol. 264, 6052–6057.

Bork, P., T. Dandekar, Y. Diaz-Lazcoz, F. Eisenhaber, M. Huynen, Y. Yuan, 1998, "Predicting function: From genes to genomes and back", *J. Mol. Biol.* , Vol. 283, pp. 707-725.

Breitkreutz, B., C. Stark, M. Tyers, 2003, "The GRID: the general repository for interaction datasets", *Genome Biol*, Vol. 4, pp. R23.

Bu, D., Y. Zhao, L. Cai, H. Xue, X. Zhu, H. Lu, J. Zhang, S. Sun, L. Ling, N. Zhang, G. Li, R. Chen, 2003, "Topological structure analysis of the protein–protein interaction network in budding yeast", *Nucleic Acids Res.*, Vol. 31, pp. 2443–2450.

Butler, L. M., Y. Webb, D. B. Agus *et al.*, "Inhibition of transformed cell growth and induction of cellular differentiation by pyroxamide, an inhibitor of histone deacetylase", *Clin Cancer Res*, Vol. 7, pp. 962-970.

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

Chang, D. Z., Z. Wu, T. L. Ciardelli, 1996, "A point mutation in interleukin-2 that alters ligand internalization", *J. Biol. Chem.*, Vol. 271, pp. 13349–13355.

Cheatman, B. and R. C. Kahn, 1995, "Insulin action and the signaling network", *Endocrine Reviews*, Vol 16, pp. 117-142.

Clement, S., U. Krause, F. Desmedt *et al.*, 2001, "The lipid phosphatase SHIP2 controls insulin sensitivity", *Nature*, Vol. 409, pp. 92-97.

Cline, G. W., K. F. Petersen, M. Krssak, J. Shen, R. S. Hundal, Z. Trajanoski, S. Inzucchi, A. Dresner, D. L. Rothman, G. I. Shulman, 1999, "Impaired glucose transport as a cause of decreased insulin-stimulated muscle glycogen synthesis in Type 2 diabetes", *N Engl J Med*, Vol. 341, pp. 240-246.

Cohen, P., R. A. Dario, A. E. Darren, 1997, "PDK1, one of the missing links in signal transduction?", *FEBS Letters*, Vol. 410, pp. 3-10.

Covert, M. W., C. H. Schilling, I. Famili, J. S. Edward, I. I. Goryanin, E. Selkow, B. O. Palsson, 2001, "Metabolic modeling of microbial strains in silico", *Trends Biochem Sci*, Vol. 26, pp.179-186.

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

De Fea, K. and R. Roth, 1997, "Protein kinase C modulation of insulin receptor substrate-1 tyrosine phosphorylation requires serine 612", *Biochemistry*, Vol. 36, pp. 12939-12947.

Deane, C., L. Salwinski, I. Xenarios, D. Eisenberg, 2002, "Protein interactions: two methods for assessment of the reliability of high throughput observations", *Moll Cell Proteomics*, Vol 1, pp. 349-356.

Deng, M., S. Mehta, F. Sun, T. Chen, 2002, "Inferring domain-domain interactions from protein-protein interactions", *Genome Res*, Vol 12, pp. 1540-1548.

Dorogovtsev, S. N. and J. F. F. Mendes, 2000 "Evolution of reference networks with aging", arXiv:cond-mat/0001419v1, January.

Downward, J., 2001, "The ins and outs of signaling", *Nature*, Vol. 411, No. 14, pp. 759-762, June.

Durmuş, Tekir, S., K. Y. Arga, K. Ö. Ülgen, 2009, "Drug targets for tumorigenesis: Insights from structural analysis of EGFR signaling network", *Journal of Biomedical Informatics*, Vol. 42, pp. 228-236.

Ehlde, M. and G. Zacchi, 1996, "A general formalism for metabolic control analysis", *Chemical Engineering Science*, Vol. 52, No. 15, pp. 2599-2606.

Erdös, P. and A. Renyi, 1960, "On the evolution of random graphs", *Publ. Math. Inst. Hung. Acad. Sci.*, Vol. 5, pp. 17-61.

Faloutsos, M., P. Faloutsos, C. Faloutsos, 1999, "On power-law relationships of the internet topology", *Comp. Comm. Rev.*, Vol. 29, pp. 251.

Fasshauer M., J. Klein, K. Ueki, K. M. Kriaucinuas, M. Benito, M. F. White, C. R. Kahn, 2000, " Essential role of IRS-2 in insulin stimulation of Glut4 translocation and glucose uptake in brown adipocytes", *The Journal of Biological Chemistry*, Vol. 25, No. 33, pp. 25494-25501.

Fell, D., 1997, "Understanding the control of metabolism", *Metabolic Control Analysis*, pp. 101-115, Portland Press, London.

Fernie A. R., P. Geigenberger, M. Stitt, 2005, "Flux an important, but neglected, component of functional genomics", *Curr Opin Plant Biol*, Vol. 8, pp: 174-182.

Finkel, T. and J. S. Gutkind, 2003, "Signal transduction and human disease", *Wiley–Liss*, Hoboken, New Jersey, USA.

Fukumoto, H., T. Kayano, J. B. Buse, Y. Edwards, P. F. Pilch, G. I. Bell, S. Seino, 1989, "Cloning and characterization of the major insulin-responsive glucose transporter expressed in human skeletal muscle and other insulin-responsive tissues", *J Biol Chem* Vol. 264, pp. 7776-7779.

Gershenfeld, N., 1999, "The nature of mathematical modeling", *Cambridge University Press*.

Giorgetti, S., R. Ballotti, A. Kowalski-Chauvel, M. Cormont, E. Van Obberghen, 1992, "Insulin stimulates phosphatidylinositol-3-kinase activity in rat adipocytes", *Eur J Biochem*, Vol. 207, pp. 599-606.

Goldstein, B. J., P. M. Li, W. D. Ding, F. Ahmad, W. R. Zhang, 1998, "In Vitamins and Hormones - Advances in Research and Applications", *Academic*, Vol. 54, pp. 67-96.

Gomez, S. and A. Rzhetsky, 2002, "Towards the prediction of complete protein- protein interaction networks", *Pac Symp Biocomput*, Vol. 7, pp. 413-424.

Goodman, O. B. Jr., J. G. Krupnick, F. Santini, V. V. Gurevich, R. B. Penn, A. W. Gagnon, J. H. Keen, J. L. Benovic, 1998, "Role of arrestins in G-protein-coupled receptor endocytosis", *Adv. Pharmacol.*, Vol. 42, pp. 429–433.

Guimaraes, K. S., R. Jothi, E. Zotenko, T. Przytycka, 2006, "Predicting domain-domain interactions using a parsimony approach, *Genome Biology*, Vol. 7, R104.

Hartwell, L., J. J. Hopfield, S. Leibler, A. W. Murray, 1999, "From molecular to modular cell biology", Nature, Vol. 402, pp. 47-52.

Hayashida, M., N. Ueda, T. Akutsu, 2004, "A simple method for inferring strengths of protein-protein interactions", *Genome Inform*, Vol. 15, pp. 56-68.

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

Hirshman, M. F., L. J. Goodyear, L. J. Wardzala, E. D. Horton, E. S. Horton, 1990, "Identification of an intracellular pool of glucose transporters from basal and insulin-stimulated rat skeletal muscle", *J Biol Chem*, Vol. 265, pp. 987-991.

Hoffmann, A., A. Levchenko, M. L. Scott, D. Baltimore, 2002, "The IκB–NFκB signaling module: temporal control and selective gene activation" *Science*, Vol. 298, pp. 1241–1245.

Holgado-Madruga, M., D. R. Emlet, D. K. Moscatello, A. K. Godwin, A. J. Wong, 1996, "A Grb2-associated docking protein in EGF- and insulin receptor signaling", *Nature*, Vol. 379.

Holman, G. D., and S. W. Cushman, 1996, "Subcellular trafficking of GLUT4 in insulin target cells", *Seminars Cell Dev Biol*, Vol. 7, pp. 259-268.

Hsing, M., G. B. Kendall, A. Cherkasov, 2008, "The use of GO terms for predicting highly-connected 'hub' nodes in protein-protein interaction networks", *BMC Systems Biology*, Vol. 2, No. 80.

Intawat, N., A. Meechai, C. Thammarongtham, K. Laoteng, V. Ruanglek, S. Cheevadhanarak, J. Nielsen, S. Bhumiratana, 2007, "Identification of flux regulation coefficients from elementary flux modes: a systems biology tool for analysis of metabolic networks", *Biotechnology and Bioengineering*, Vol. 97, pp. 1535-1549, August.

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

Jeong, H., S. P. Mason, A. L. Barabasi, Z. N. Oltvai, 2001, "Lethality and centrality in protein networks", *Nature*, Vol. 411, pp. 41–42.

Joost, H. G., G. I. Bell, J. D. Best *et al.*, 2002, "Nomenclature of the GLUT/SLC2A family of sugar/polyol transport facilitators", *Am J Physiol Endocrinol Metab*, Vol. 282, pp. 974-976.

Jothi, R., P. Cherukuri, A. Tasneem, T. Przytycka, 2006, "Co-evolutionary analysis of domains in interacting proteins reveals insights into domain-domain interactions mediating protein-protein interactions", *J Mol Biol*, Vol. 362, pp. 861-875.

Jullien, J., V. Guili, L. F. Reichardt, B. B. Rudkin, 2002, "Molecular kinetics of nerve growth factor receptor trafficking and activation", *J. Biol. Chem.*, Vol. 277, pp. 38700–38708, June.

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

Kacser, H., H. M. Sauro, L. Acerenza, 1990, "Enzyme-enzyme interactions and control analysis: The case of non-additivity: monomer-oligomer associations", *Eur J Biochem*, Vol. 187, pp. 481-491.

Karlsson, H. K. R., 2007, "Insulin signaling and glucose transport in insulin resistant human skeletal muscle", *Cell Biochem Biophys*, Vol. 48, pp. 103-113.

Karlsson, H. K. R., 2005, *Insulin signaling and glucose transport in insulin resistant human skeletal muscle*, Ph.D Thesis, Kongl Carolinska Medico Chirurgiska Institute, Stockholm.

Karnieli, E., M. J. Zarnowski, P. J. Hissin, I. A. Simpson, L. B. Salans, S. W. Cushman, 1981, "Insulin-stimulated translocation of glucose transport systems in the isolated rat adipose cell. Time course, reversal, insulin concentration dependency, and relationship to glucose transport activity", *J Biol Chem*, Vol. 256, pp. 4772-4777.

Klamt, S. and J. Stelling, 2003, "Two approaches for metabolic pathway analysis?", *Trends Biotechnol*, Vol. 21, pp. 64-69.

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

Kristiansen, S., M. Hargreaves, E. A. Richter, 1996, "Exercise-induced increase in glucose transport, GLUT-4 and VAMP-2 in plasma membrane from human muscle", *Am J Physiol Endocrinol Metab*, Vol. 270, pp. 197-201.

Langlais, P., L. Q. Dong, F. J. Ramos, D. Hu, Y. Li, M. J. Quon, F. Liu, 2004, "Negative regulation of insulin-stimulated mitogen activated protein kinase signaling by Grb10", *Molecular Endocrinology*, Vol. 18, No. 12, pp. 350-358.

Langeveld, M. and J. M. F. G. Aerts, 1997, "Glycosphingolipids and insulin resistance", *Progress in Lipid Research*, Vol. 48, pp. 196-205.

Lee, E., A. Salic, R. Kruger, R. Heinrich, M. W. Kirschner, 2003, "The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway", *PLoS Biol.*, Vol. 1, pp. 116–132.

Levchenko, A., 2003, "Dynamical and integrative cell signaling: challenges for the new biology", *Biotechnol. Bioeng.*, Vol. 84, pp. 773–782.

Li, C., Zhang, B.B., 2007, "Insulin signaling and action: glucose, lipids, protein", http://www.endotext.org/diabetes/diabetes4/diabetesframe4.htm.

Lizcano, J. M. and D. R. Alessi, 2002, "The insulin signaling pathway", *Current Biology*, Vol 12., No 7.

Lomberk, G. and R. Urrutia, 2009, "Primers on molecular pathways- The insulin pathway", *Pancreatology*, Vol. 9, pp. 203-205.

Madoff, D. H., T. M. Martensen, M. D. Lane, 1988, "Insulin and insulin-like growth factor 1 stimulate the phosphorylation on tyrosine of a 160 kDa cytosolic protein in 3T3-L1 adipocytes", *Biochem J*, Vol. 252, pp. 7-15.

McAdams, H. H. and A. Arkin, 1998, "Simulation of prokaryotic genetic circuits", *Annu. Rev. Biophys. Biomol. Struct.*, Vol. 27, pp. 199–224.

Mishra, G. R., M. Suresh, K. Kumaran, N. Kannabiran, S. Suresh, P. Bala, K. Shivakumar, N. Anuradra, R. Reddy, T. M. Raghavan, 2006, "Human protein reference database – 2006 update", *Nucleic Acids Res*, Vol. 34, pp. 411-414.

Moller, D. E., 2001, " New drug targets for type 2 diabetes and the metabolic syndrome", *Nature*, Vol. 414, pp. 821-827.

Mrowka, R., A. Patzak, H. Herzel, 2001, "Is there a bias in proteome research?", *Genome Res*, Vol. 11, pp. 1971-1973.

Myers, M. G. and M. F. White, 1993 “The new elements of insulin signaling. Insulin receptor substrate-1 and proteins with SH2 domains”, *Diabetes*, Vol. 42, pp. 643-650.

Nelson, D. L. and M. M. Cox, 2004, “Lehninger’s Principles of Biochemistry”, Molecular Mechanisms of Signal Transduction, 4ed. draft, pp. 422-423.

Neves, S. R., P. T. Ram, R. G. Iyengar, 2002, “Protein pathways”, *Science*, Vol. 296, pp. 1636–1639.

Oda, K. and H. Kitano, 2006, “A comprehensive map of the toll-like receptor signaling network”, *Molecular Systems Biology*, 4100057.

Oda, K., H. Moriya, Y. Matsuoka, H. Kitano, 2005b, *A Comprehensive Molecular Interaction Map of Budding Yeast*,http://celldesigner.org/download/YeastMapPoster.pdf

Oda, K., Y. Matsuoka, A. Funashashi, H. Kitano, 2005a, “Comprehensive pathway map of epidermal growth factor receptor signaling”, *Molecular Systems Biology*, 4100014-E1-E17, May.

Palsson, B., 2000, “The challenges of in silico biology”, *Nat Biotechnol*, Vol. 18, pp.1147-1150.

Pandey, A. and M. Mann, 2000, “Proteomics to study genes and genomes”, *Nature*, Vol. 406, pp. 837-846.

Papin, J. A. and B. O. Palsson, 2004a, “Topological analysis of mass-balanced signaling networks: a framework to obtain network properties including crosstalk”, *Journal of Theoretical Biology*, Vol. 227, No. 2, pp. 283-297, March.

Papin, J. A. and B. O. Palsson, 2004b, "The JAK–STAT signaling network in the human B-cell: an extreme signaling pathway analysis", *Biophys. J.*, Vol. 87, No. 1, pp. 37–46, July.

Papin, J. A., T. Hunter, B. O. Palsson, S. Subramaniam, 2005, "Reconstruction of signaling networks and analysis of their properties", *Nature Rewievs Molecular Cell Biology*, Vol. 6, No. 2, pp. 99-111, February.

Papin, K. R., J. Stelling, N. D. Price, S. Klamt, S. Schuster, B. O. Palsson, 2004, "Comparison of network-based pathway analysis methods", *Trends Biotechnol*, Vol. 22, pp. 400-405.

Patil, A. and H. Nakamura, 2005, "Filtering high–throughput protein-protein interaction data using a combination of genomic features", *BMC Bioinformatics*, Vol. 6, pp. 100.

Patil, K. R., I. Rocha, J. Forster, J. Nielsen, 2005, "Evolutionary programming as a platform for in silico metabolic engineering", *BMC Bioinformatics*, Vol. 6, pp. 308.

Paz, K., Y. F. Liu, H. Shorer, R. Hemi, D. LeRoith, M. Quon, H. Kanety, R. Seger, Y. Zick, 1999, "Phosphorylation of insulin receptor substrate-1 (IRS-1) by protein kinase B positively regulates IRS-1 function", *J Biol Chem*, Vol. 274, pp. 28816–28822.

Pessin, J. E, D. C. Thurmond, J. S. Elmendorf, K. J. Coker, S. Okada, 1999, "Molecular basis of insulin-stimulated GLUT4 vesicle trafficking. Location! Location! Location!", *J Biol Chem*, Vol. 274, pp. 2593–2596.

Przulj, N., D. A. Wigle, I. Jurisica, 2004, "Functional topology in a network of protein interactions", *Bioinformatics*, Vol. 20, pp. 340-348.

Quon, M. J., L. A. Campfield, 1991, "A mathematical model and computer simulation study of insulin receptor regulation", *J Theor Biol*, Vol. 150, pp. 59-72.

Ravichandran, L. V., D. L. Esposito, J. Chen, M. J. Quon, 2001, "Protein kinase C-ζ phosphorylates insulin receptor substrate-1 and impairs its ability to activate phosphatidylinositol 3-kinase in response to insulin", *J Biol Chem*, Vol. 276, pp. 3543–3549.

Reder, C., 1988, "Metabolic control theory: a structural approach", *J Theor Biol*, Vol.135, pp. 175-201.

Rives, A. W. and T. Galitski, 2003, "Modular organization of cellular networks", *Proc. Natl Acad. Sci.*, Vol. 100, pp. 1128–1133.

Rosivatz, E., "Inhibiting PTEN", *3rd Focused Meeting on PI3K Signaling and Disease*, 2007.

Ruffner, H., A. Bauer, T. Bouwmeester, 2007, "Human protein-protein interaction networks and the value for drug discovery, *Drug Discovery Today*, Vol. 12, pp. 709-716, September.

Ryder, J., J. Yang, D. Galuska *et al.*, 2000, "Use of a novel impermeable biotinylated photolabeling reagent to assess insulin- and hypoxia-stimulated cell surface GLUT4 content in skeletal muscle from Type 2 diabetic patients", *Diabetes*, Vol. 49, pp. 647-654.

Salwinski, L., C. S. Miller, A. J. Smith, F. K. Pettit, J. U. Bowie, D. Eisenberg, 2004, "The database of interacting proteins: 2004 update", *Nucleic Acids Res*, Vol. 32, pp. 449, 451.

Sauer U., 2004, "High throughput phenomics: experimental methods for mapping fluxomes", *Curr Opin Biotechnol*, Vol. 15, pp. 58-63.

Scheepers, A., H. G. Joost, A. Schurmann, A., 2004, "The glucose transporter families SGLT and GLUT: molecular basis of normal and aberrant function", *J Parenter Enternal Nutr*, Vol. 28, pp. 364-371.

Schilling, C. H., D. Letscher, B. O. Palsson, 2000, "Theory for systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective, *J Theor Biol*, Vol. 203, pp. 229-248.

Schuster S., T. Dandekar, 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.

Schuster, S., B. N. Kholodenko, H. V. Westerhoff, H., 2001, "Cellular information transfer regarded from a stoichiometry and control analysis perspective", *Biosystems*, Vol. 55, pp. 73–81.

Schuster, S., C. Hilgetag, J. H. Woods, and D. A. Fell, 1996, "Elementary modes of functioning in biochemical reaction networks" In Cuthbertson, R., Holcome, M., & Paton, R., eds., *Computation in Cellular and Molecular Biological Systems*, pp.151-165, World Scientific, Singapore.

Schuster, S., Kholodenko, B. N., Westerhoff, H. V., 2000, "Cellular information transfer regarded from a stoichiometry and control analysis perspective", *Biosystems*, Vol. 55, pp. 73-81.

Schwartzenberg-Bar-Yoseph, F., M. Armoni, E. Karnieli, 2004, "The tumor suppressor p53 down-regulates glucose transporters GLUT1 and GLUT4 gene expression", *Cancer Research* Vol. 64, pp. 2627-2633.

Sedaghat, A. R., A. Sherman, M. J. Quon, 2002, "A mathematical model of insulin signaling pathways", *J Physiol Endocrinol Metab*, Vol. 283, pp. 1084-1101.

Shang, W., Y. Yang, L. Zhou, B. Jiang, H. Jin, M. Chen, 2008, "Ginsenoside Rb1 stimulates glucose uptake through insulin-like signaling pathway in 3T3-L1 adipocytes", Journal of Endocrinology, DOI: 10.1677/JOE-08-0104.

Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, 2003, "Cytoscape: a software environment for integrated models of biomolecular interaction networks", *Genome Res,* Vol. 13, pp. 2498-2504.

Shulman, G. I., 2000, "Cellular mechanisms of insulin resistance", *J Clin Invest*, Vol. 106, pp. 171-176.

Sprinzak, E. and H. Margalit, 2002, "Correlated sequence-signatures as markers of protein-protein interaction", *J Mol Biol*, Vol. 311, pp. 681-692.

Standaert, M. L. and R. J. Pollet, 1984, "Equilibrium model for insulin-induced receptor down-regulation. Regulation of insulin receptors in differentiated BC3H-1 myocytes", *J Biol Chem*, Vol. 259, pp. 2346-2354.

Stark, C., B. J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, M. Tyers, 2006, "BioGRID: a general repository for interaction datasets", *Nucleic Acids Res*, Vol. 34, pp. 535-539.

Steffen, H., A. Petti, J. Aach, P. D'haeseleer, G. Church, 2002, "Automated modeling of signal transduction networks", BMC Bioinformatics, Vol. 3, pp. 34.

Stelling, J., S. Klamt, K. Bettenbrock, S. Schuster, E. D. Gilles, 2002, "Metabolic network structure determines key aspects of functionality and regulation", *Nature*, Vol. 420, pp. 190-193.

Stratford, S., K. L. Hoehn, S. A. Summers, 2004, "Regulation of insulin action by ceramide: dual mechanisms linking ceramide accumulation of Akt/protein kinase B", *J Biol Chem*, Vol. 279, pp. 36608-36615.

Stryer, L., 1995, *Biochemistry*, W. H. Freeman and Company, New York.

Su, G. H., T. A. Sohn, B. Ryu, S. E. Kern, 2000, "A novel histone deacetylase inhibitor identified by high throughput transcriptional screening of a compound library", *Cancer Res*, Vol. 60, pp. 3137-3142.

Summers, S. A, 2006, "Ceramides in insulin resistance and lipotoxicity", *Progress in Lipid Search*, Vol. 45, pp. 42-72.

Summers, S. A, A. W. Kao, A. D. Kohn, G. S. Backus, R. A. Roth, J. E. Pessin, 1999, "The role of glycogen synthase kinase 3beta in insulin stimulated glucose metabolism", *J Biol Chem*, Vol. 274, pp. 34-40.

Summers, S. A., L. A. Garza, H. Zhou, M. J. Birnbaum, 1998, "Regulation of insulin-stimulated glucose transporter GLUT4 translocation and Akt kinase activity by ceramide", *Mol Cell Biol*, Vol. 18, pp. 57-64.

Takigawa-Imamura, H. Sekine, T. Murata, M. Takayama, K. Nakazawa, K. Nakagawa, J. Nakagawa, 2003, "Stimulation of glucose uptake in muscle cells by prolonged treatment with scriptide, a histone deacetylase inhibitor", *Biosci Biotechnol Biochem,* Vol. 67, pp. 1499-1506.

Taniguchi, M. C., B. Emanuelli, C. R. Kahn, 2006, "Critical nodes in signaling pathways: insights into insulin action", *Nature Reviews,* Vol. 7, pp. 85-96.

Teruel, M. N. and T. Meyer, 2000, "Translocation and reversible localization of signaling proteins: a dynamic future for signal transduction", *Cell*, Vol. 103, pp. 181–184.

The Gene Ontology Consortium, 2000, "Gene ontology: tool for the unification of biology", *Nature Genetics,* Vol. 25, pp. 25-29.

Theurkauf, W. E., 1994, "Premature microtubule-dependent cytoplasmic streaming in cappuccino and spire mutant oocytes", *Science*, Vol. 265, pp. 2093–2096.

Van Epps-Fung, M., K. Gupta, R. W. Hardy, A. Wells, "A role for phospholipase C activity in GLUT4 mediated glucose transport", *Endocrinology*, Vol. 138, No. 12, pp. 5170-5175.

Virkamaki, A., K. Ueki, C. R. Kahn, 1999, "Protein-protein interaction in insulin signaling and the molecular mechanisms of insulin resistance", Journal of Clinical Investigation, Vol. 103, No.7, pp. 931-943.

von Mering, C., L. J. Jensen, B. Snel, S. D. Hooper, M. Krupp, M. Foglierini, N. Joufrre, M. A. Huynen, P. Bork, 2005, "STRING: known and predicted protein-protein associations, integrated and transferred across organisms", *Nucleic Acids Res*, Vol. 33, pp. 433-437.

von Mering, C., L. J. Jensen, M. Kuhn, S. Chaffron, T. Doerks, K. Beate, B. Snel, P. Bork, 2007, "STRING 7 - recent developments in the integration and prediction of protein interactions", *Nucleic Acids Research*, Vol. 35, pp. 358-362.

von Mering, C., M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, B. Snel, 2003, "String: a database of predicted functional associations between proteins, *Nucleic Acids Res*, Vol. 34, pp. 354-357.

von Mering, C., R. Krause, B. Snel, M. Cornell, S. Oliver, S. Fields, P. Bork, 2002, "Comparative assessment of large-scale data sets of protein-protein interactions", *Nature*, Vol. 417, pp. 399-403.

Vuong, T. M. and M. Chabre, 1991, "Deactivation kinetics of the transduction cascade of vision", *Proc. Natl Acad. Sci.*, Vol. 88, pp. 9813–9817.

Wanant, S. and M. J. Quon, 2000, "Insulin receptor binding kinetics: modeling and simulation studies", *J Theor Biol*, Vol. 205, pp. 355-364.

Weng, G., U. S. Bhalla, R. Iyengar, 1999, "Complexity in biological signaling systems", *Science*, Vol. 284, pp. 92–96.

Weston, C. R. and R. J. Davis, 2001, "Signaling specifity – a complex affair", *Science*, Vol. 292, pp. 2439-2440.

Widermuth, M. C., 2000, "Metabolic control analysis: biological applications and insights", Genome Biology, Vol. 1, No. 6.

Wu, X., L. Zhu, J. Guo, D. Zhang, K. Lin, 2006, "Prediction of yeast protein-protein interaction network: insights from the gene ontology and annotations", *Nucleic Acids Res*, Vol. 34, pp. 2137-2150.

Zanzoni, A., L. Montecchi-Palazzi, M. Quondam, G. Ausiello, M. Helmer-Citterich, G. Cesareni, 2002, "MINT: a molecular interaction database", *FEBS Lett*, Vol. 513, pp. 135-140.

Zubay, G., 1973, "In vitro synthesis of protein in microbial systems", *Annu. Rev. Genet.*, Vol. 7, pp. 267–287.