MODELING AND SIMULATION OF INSULIN SIGNALING

AZZURRA CARLON

Department of Information Engineering
Bioengineering
University of Padova
March 2013

ADVISOR: Gianna Maria Toffolo

COADVISOR: Barbara Di Camillo, Federica Eduati

SUMMARY

The purpose of this thesis is to implement a computational model of insulin signaling pathway. Insulin pathway is a complex system that permits, upon insulin binding on its receptor, to trigger a cascades of signals which realize the most typical insulin actions. Due to its high degree of complexity, simple reasoning on experimental data cannot permit a detailed analysis of insulin system, making necessary the introduction of new and more powerful mathematical tools, such as computational models. Computational models allow to describe a complex system populated by many chemical species by means of a network, in which nodes represent these elements and edges (with sign) describe the type of interaction. In particular, we use rule-based modeling as computational approach because it offer a easy way to implement a biological system, respect to common methods using ordinary differential equations (ODEs), with a remarkable saving in time and reducing the risk of error. Thus, insulin model was implemented using BioNetGen language, a software for rule-based modeling, making use of information taken from scientific literature. In particular, we referred to three published models describing different parts of insulin signaling pathway and, after a integration procedure, we obtained the final insulin model. Insulin model was used to predict the outcome of the system. The resulting time series mimicking the concentrations of chemical species populating the insulin signaling pathway were used to characterize the dynamic behavior of the system. Since defects in insulin pathway, often responsible of *insulin resistance* and *diabetes type 2*, are caused by network malfunctioning, we think that to characterize the dynamic behavior of system working under physiological conditions may be a cunning approach to distinguish it from those pathologic. Moreover, model characterization may reveal new insight about complex mechanisms governing insulin system. Hence, we realized a system characterization by classification of the qualitative behaviors of simulated time series, thereby according to their *pattern*. We distinguished four main pattern sets and we related them to their function in insulin system. Using pattern classification and parametric description of each pattern, we validated of insulin model comparing simulations with experimental data according to most significant dynamic features. In future perspective, insulin model characterization may be used to further investigate the origin of network disregulations and to find a drug or combination of drugs able to counterbalance the effect of disease.

SOMMARIO

Lo scopo di questa tesi è quello di implementare un modello computazionale del pathaway dell'insulina. Il pathway dell'insulina è un sistema complesso che permette, in seguito al legame dell'insulina con il suo recettore, di innescare una cascata di segnali biologici che realizzano le azioni tipiche dell'insulina. Dovuto al suo alto grado di complessità, il semplice ragionamento su dati sperimentali non permette un'analisi dettagliata del sistema dell'insulina e rende necessario l'utilizzo di metodi matematici nuovi e più potente, come i modelli computazionali. I modelli computazionali permettono di descrivere un complesso sistema biologico popolato da molte species chimiche per mezzo di una semplice rete, in cui i nodi rappresentano questi elementi e gli archi (dotati di segno) descrivono il tipo di interazione. In particolare, abbiamo usato il rule-based modeling come approcio computazionale perchè offre un metodo semplice per implementazione di sistemi biologici rispetto ai metodi classici che fanno uso di equazioni differenziali ordinarie (ODEs), con un sensibile risparmio di tempo e riducendo la possibilità d'errore. Quindi, il modello dell'insulina è stato implementato usando il linguaggio BioNetGen, un software per il rule-based modeling, utilizzando informazioni prese dalla letteratura scientifica. In particolare, si è fatto uso di tre modelli pubblicati che descrivono diverse parti del pathway dell'insulina e, dopo una procedura d'integrazione, si è ottenuto il modello finale dell'insulina. Il modello dell'insulina è stato usato per realizzare delle predizioni del sistema. Le risultanti serie temporali che descrivono le concentrazioni delle specie chimiche che popolano il pathway dell'insulina sono state usate per caratterizzare il comporatamento dinamico del sistema. Visto che i difetti del pathway dell'insulina, spesso responsabili dell'*insulino resistenza* e del *diabete di tipo 2,* sono causati da malfunzionamenti della rete, riteniamo che caratterizzare il comportamento dinamico del sistema in condizioni fisiologiche possa essere un approcio utile per distinguerlo da quelli patologici. Inoltre, la caratterizzazione del modello può permetterci di comprendere meglio i complessi meccanismi che governano il sistema dell'insulina. Quindi, abbiamo realizzato la caratterizzazione del sistema classificando le serie temporali simulate a seconda dell'andamento qualiticativo, ovvero secondo il loro pattern. Abbiamo distinto quattro principali insiemi di pattern e li abbiamo collegati alla loro funzione all'interno del sistema dell'insulina. Usando la classificazione mediante pattern e la descrizione parametrica di ciascuno di essi, abbiamo realizzato una nuova validazione del modello dell'insulina comparando le simulazioni con i dati sperimentali e valutando le loro caratterische dinamiche più significative. In

futuro, la caratterizzazione del sistema dell'insulina potrebbe essere utile per ulteriori indagini circa le origini delle dis-regolazioni della rete e per trovare un farmaco o una combinazione di farmaci capaci di riportare il sistema in condizioni di normalità.

I'll drown my beliefs
To have you be in peace
I'll dress like your niece
And wash your swollen feet

Just don't leave
Don't leave

And true love waits
In haunted attics
And true love lives
On lollipops and crisps

Just don't leave
Don't leave

I'm not living, I'm just killing time
Your tiny hands, your crazy-kitten smile

Just don't leave
Don't leave

— Radiohead, TRUE LOVE WAITS

ACKNOWLEDGMENTS

A great thanks to everybody support me through these five years of university career, and of life. Thank you Federica, Alessandra, Francesca, Giulia and Giorgia, who taught me friendship meaning and never let me alone. Thank you Giacomo, Stefano, Alessandro, Filippo, Marco and Fabio for the great deal of laughs during and after lectures. Thank you family who taught me the importance of loyalty and hard work and permitted me to start this great adventure, hoping to repay you with this work. Special thanks to who encouraged me more than everyone else, who held me with patience, and who always trusted in my faculties: thank you Davide.

A great thanks also to Gianna Maria Toffolo e Barbara Di Camillo, who gave me the chance to challenge myself and to increase my expertise. Thanks to Federica Eduati, who guided and helped me with competence and patience during this work. Thanks to Leonard Harris from Pittsburg University, always available in solving technical difficulties encountered during model implementation.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

IR	Insulin Receptor
IRS-1	Insulin Receptor Substrate-1
PI3K	Phosphatidyl-inositide 3-Kinase
PI(3,4,5)P3	Phosphatidyl-inositol 3,4,5-trishosphates
PDK	3-Phosphoinositide-dependent Protein Kinase
PKC	Protein Kinase C
Akt	RAC-alpha Serine/Threonine-protein Kinase
PTP	Protein Tyrosine Phosphatase
PTEN	Phosphatase and Tensin Homologue
SHIP	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1
AMPK	Adenosine Monophosphate-dependent Kinase
mTOR	Mammalian Target of Rapamycin
mTORC1	Mammalian Target of Rapamycin Complex 1
mTORC2	Mammalian Target of Rapamycin Complex 2
PRAS40	Proline-rich Akt Substrate of 40 kDa
p70-S6K	S6 Kinase
TSC1-TSC2	Tuberous Sclerosis Complex 1/2 Dimer
Src	Src Avian Sarcoma Viral Oncogene Homolog
SHP2	SH2 Domain-containing Tyrosine Protein Phosphatase 2
RasGAP	Ras GTPase-activating Protein 2
Grb2	Growth Factor Receptor-bound Protein 2
SOS	Son of Sevenless Homolog 1
Raf	RAF Proto-oncogene Serine/Threonine-protein Kinase
MEK	Dual Specificity Mitogen-activated Protein Kinase Kinase 1
ERK	Mitogen-activated Protein Kinase 3

1 INTRODUCTION

1.1 SIGNALING PATHWAYS AND MALFUNCTIONING

Cellular processes of life are controlled by complex regulatory systems including signaling pathways, the transcription network, and specialized circuits for cell cycle control, growth regulation, stress response, and many other cell functions [13].

In particular, signaling pathways are complex, interdependent cascades of signals that sense input stimuli (e. g. extracellular ligands or intracellular metabolites) and transmit, process, and integrate this information to provide output signals that accordingly regulate the activity of transcription factors or other effector proteins. Signaling networks permit cells to communicate with each others and with external environment, such as extracellular matrix, and to undergo phenotypic changes, such as cellular division, differentiation, death and others. Hence, these networks can be considered as information-processing devices that translate input signals into output signals in which information is often coded by concentrations, modifications, and localization of proteins, either in the stationary levels or in temporal patterns.

Malfunctioning of signaling networks may alter physiological processes of cells, potentially leading to severe consequences on the organism. The most common pathologies caused by altered cellular signaling networks concern heart diseases, congenital abnormalities, metabolic disorders and immunological abnormalities. Moreover, it has been recently demonstrated that network malfunctioning is involved with oncogenic properties of cancer cells [26, 1].

Insulin signaling pathway is an essential pathway due to its biological functions and its connection with pathologies which increased their importance in these recent years: *insulin resistance* and *diabetes type 2* (diabetes mellitus). Rates of diabetes increased markedly over the last 50 years in parallel with obesity and, to the present day, this pathology afflicts at least 285 million people [44]. Diabetes involves many long-term complications including heart diseases, strokes, diabetic retinopathy, kidney failure and poor circulation of limbs which may lead to amputations. Even in the absence of diabetes, insulin resistance is often associated with central obesity, hypertension, polycystic ovarian syndrome, dyslipidemia and atherosclerosis.

Thus, detailed analysis of regulatory processes constituting insulin signaling pathway may permit to develop new insights about the origin of these disregulations and, in future perspective, to find a drug

***Insulin Resistance**: physiological condition characterized by inability of tissues to respond to the normal actions of insulin. Cells are not able to take in glucose, amino acids and fatty acids.*

***Diabetes type 2**: chronic metabolic disorder characterized by increased blood glucose caused by insulin resistance and relative insulin deficiency. This is in contrast to diabetes mellitus type 1, in which there is an absolute insulin deficiency due to destruction of islet cells in the pancreas.*

or combination of drugs able to counterbalance the effect of the disease.

1.2 COMPUTATIONAL MODELS OF INSULIN SIGNALING NETWORK

Due to high degree of complexity, the detailed and rational analysis of insulin signaling pathway represents a new challenge that cannot be undertaken by means of simple reasoning on experimental data. During recent years, new and more powerful methods have been introduced to be used in combination with experimental results. Computational methods provide useful information to guide experimental design obtaining more informative experiments, whereas data collected from experiments help implementation and refinement of computational models, providing more accurate predictions of biological systems. Computational models permits to model a signaling pathway as a *network*, in which *nodes* represent the distinct chemical species populating the system, and *edges* (with sign) their interactions.

***Mass action law**: law stating that the rate of any chemical reaction is proportional to the product of the masses of the reacting substances, with each mass raised to a power equal to the coefficient that occurs in the chemical equation.*

There are different approach to construct these network but the most relevant make use of *mass action law* to describe the dynamics of chemical species. There are several *dynamic networks* published in scientific literature and all of them have in common that the rate of change of chemical species is described by a ordinary differential equation (ODEs). These models allows to make assumptions about the behavior of the network. However, difficulties in their implementation are related to the fact that mathematical description of dynamic models requires a great deal of knowledge about concentrations and kinetic rate constants, often hard to obtain.

Several models of insulin control system have been published in last 10 years, providing new insights at three different levels: (i) insulin binding to its receptor, (ii) insulin signaling to cellular responses, and (iii) integration of intracellular insulin signaling with whole-body glucose homeostasis [31].

Insulin signaling pathway may find a description using models belonging to the first two levels. In particular, currently available dynamic models describe:

- insulin binding to its receptor [6],
- insulin receptor autophosphorylation and subsequent phosphorylation of its substrate, together with receptor cycling and endocytosis [30, 29]
- downstream signaling activation focusing on translocation of GLUT4 glucose transporter [22], on mTOR regulation system [33, 34], on dendritic protein synthesis [2], on eukaryotic translation initiation [41], and on breast cancer therapy [10], on joint regulation of insulin and amino acids [42] and on crosstalk

with epidermal growth factor (EGF) signaling and the mitogen-activated protein kinase (MAPK) pathway [12].

1.3 CHAPTERS OVERVIEW

The purpose of this thesis is to realize a computational model of insulin signaling pathway. The model should comprise many of the chemical species populating insulin system and most of their interactions. Such detailed model will permit to simulate concentrations of chemical species. Concentration time series will be used to mainly characterize the dynamic behavior of insulin system and to gain new insights about its regulatory mechanisms. In the following Chapters we will deal with:

CHAPTER2 Insulin signaling pathway. An overview of the main elements constituting insulin network, focusing mainly on kinetics characteristics of processes occurring upon insulin binding, such as molecular interactions and phosphorylation events. Insulin pathway will be presented dividing it in three main subpathways, that are PI3K-PKB/Akt pathway, Ras/MAPK pathway and Cbl/CAP pathway.

CHAPTER3 Computational models of biochemical systems. An overview of the main computational modeling techniques focusing on rule-based modeling approach and on BioNetGen language, the software implementing it.

CHAPTER4 Insulin model implementation. We started from three published models describing different part of insulin pathway to realize a final and more complete model of insulin pathway. We also analyze step by step its implementation using BioNetGen language.

CHAPTER5 Insulin model analysis and validation. We used insulin model to obtain predictions of the system. Using simulated time series describing concentrations of chemical species we characterized the dynamic behavior of the system and we gain new insights about its regulatory processes. Moreover, we realize a further validation of insulin model using some experimental data focusing on most significant dynamic features.

2 INSULIN SIGNALING PATHWAY

2.1 INTRODUCTION

Insulin signaling pathway starts with binding of insulin to its receptor. This event triggers a complex cascade of signals culminating in several biological responses by means of the activation of two major signaling sub-pathways:

- *PI3K-AKT/PKB pathway*, which is responsible for most of the metabolic actions of insulin, and
- *Ras-MAPK pathway*, which regulates expression of some transcription factors and cooperates with the PI3K pathway in controlling cell growth and differentiation.

The PI3K-AKT/PKB pathway is aided in its action by another pathway, the *Cbl/CAP pathway*, and their cooperation regulates the main insulin actions, such as the glycogen synthesis and the glucose transport inside the cell. The Ras-MAPK pathway is a general pathway that can be triggered by different growth factors, including insulin, and that mainly controls DNA, RNA and protein syntheses.

Fig. 1 depicts insulin signaling pathway as a network of several molecules connected to each others by different kind of relationships. For semplicity, the members belonging to the sub-pathways mentioned above are identified by circles of different colours. Red circles indicates elements of PI3K-AKT/PKB pathway, blue circles those belonging to Ras/MAPK pahway and purple circles those constituting Cbl/CAP pathway.

In the following sections we present in more detail the main elements constituting insulin signaling pathway. A section apart is dedicated to insulin receptor due to its key role played in all the abovementioned sub-pathways. Since the purpose of this study is to implement a dynamic model mimicking the main features of insulin network, we mainly focus in catching the key characteristics which determine the kinetics of the processes occurring upon insulin binding, such as molecular interactions and phosphorylation events. This information will be useful in the next chapters where several models of insulin subsystems will be examined and the final model will be presented.

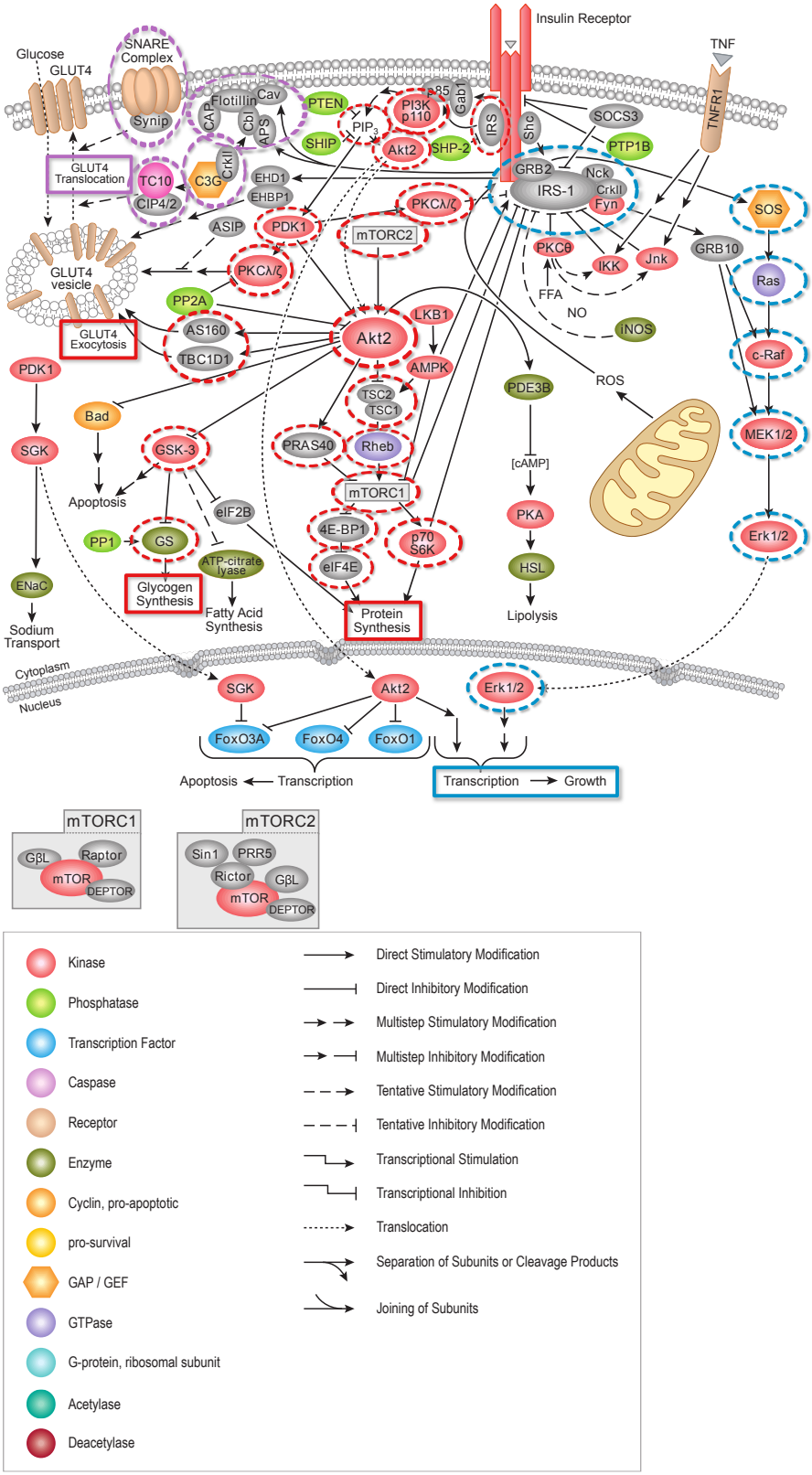

Figure 1: Insulin signaling pathway. Molecules enclosed by circles of different colours belong to dinstict sub-pathways and are, respectively, red for PI3K-AKT/PKB, blue for Ras/MAPK and purple for Cbl/-CAP. Adapted figure from [32]

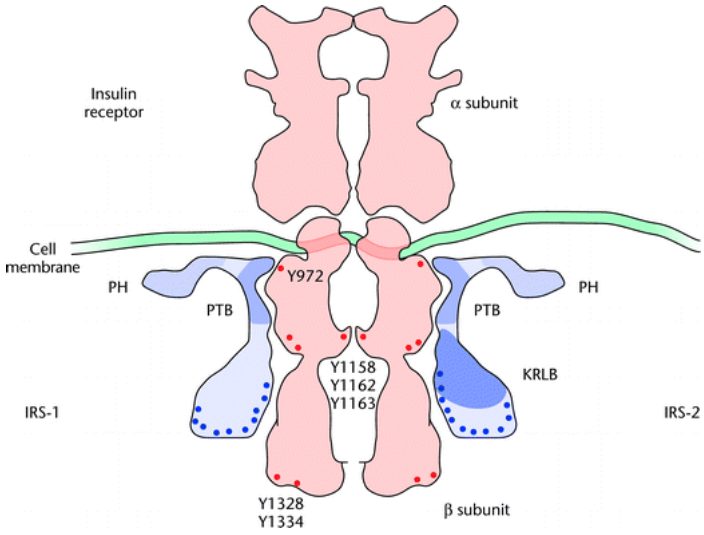

Figure 2: Insulin receptor. Structure and phosphorylation sites. Figure taken from [28]

2.2 INSULIN RECEPTOR: STRUCTURE, FUNCTIONS AND SUBSTRATES

Insulin action is initiated upon insulin binding to its cognate receptor. The insulin receptor (IR), depicted in Fig. 2, belongs to a family of receptors characterized by *intrinsic tyrosine kinase activity* (RTKs) that includes, in addiction to IR, the insulin-like growth factor-1 receptor (IGF-1R), the insulin receptor-related receptor (IRR), the epidermal growth factor receptor (EGFR), the platelet-derived growth factor receptor (PDGFR) and others.

Generally, each receptor is produced starting from two chains, termed α and β, that are linked by a disulfide bond. The α chains contribute to formation of ligand-binding domain, while β chains carry the kinase domain. In the case of insulin, two α chains and two β chains are linked together forming a biologically active receptor heterotetramer ($\alpha^2\beta^2$).

IR can be structurally divided in two main parts, the extracellular and the intracellular regions. Although the chains constituting IR are covalently linked, these two domains function independently.

In the following subsections we treat in more detail all the main features regarding IR including structure, functions, substrates, internalization and degradation processes [15, 32, 4].

2.2.1 *Insulin Receptor Structure*

X-ray crystallography studies revealed that the extracellular domain of IR consists of the entire α-subunit and about one third of the β-subunit. This region is responsible for the insulin binding and, precisely, the α-subunit contains the primary ligand-binding site.

The intracellular region can be divided into several subdomains with different functions and characteristics:the juxtamembrane region, the tyrosine kinase domain and the carboxyl-terminus domain.

The *juxtamembrane* (JM) region is constituted by 50 amino acids and includes at least one autophosphorylation site (Tyr999), which resides in the L-X4-NPXYXSXSD motif. This motif serves as binding site for IR substrates such as Shc and insulin receptor substrate (IRS) proteins. These proteins interact with the NPXY motif through their P-Tyr binding (PTB) domain. The sequences surrounding the phosphorylated NPXY motif contribute differentially to either IRS or Shc proteins recognition. Moreover, the juxtamembrane domain also contains several motifs implicated in regulating IR internalization (see 2.2.4).

The *tyrosine kinase* domain is composed of two lobes that are linked together by a single connection. The N-terminal lobe constitutes the ATP-binding site whereas the C-terminal lobe contains the active site (catalytic loop), three autophosphorylation sites (activation loop) and the kinase-insert region.

The carboxyl-terminus domain contains two autophosphorylation sites and its role is still unresolved.

2.2.2 *Ligand Binding and Insulin Receptor Autophosphorylation*

Receptor Tyr kinases are activated when ligand binds to their extracellular region. In the case of monomeric receptors (e. g., EGFR and PDGFR), the ligand binding includes the receptor dimerization, a necessary step before their activation. However, certain RKTs, including IR, exist in predimerized form even in absence of ligand.

In basal state, the activation loop of IR occludes the catalytic site so that access to ATP and substrates is blocked. A rapid conformational change in the receptor occurs when insulin binds to specific regions on the α-subunit, and this results in activation of the tyrosine kinase domain. When the receptor is activated, the kinase domain in one half of the receptor-dimer phosphorylates cytoplasmic tyrosine residues in the activation loop of the other half of the receptor-dimer. This transautophospohorylation results in a large increase of the kinase activity of the receptor. Upon autophosphorylation, the activation loop swings out of the catalytic site and gives unrestricted access to ATP and substrates.

2.2.3 *Insulin Receptor Substrates*

Binding of insulin leads to phosphorylation of several intracellular substrates, including, insulin receptor substrates (IRS1, 2, 3, 4), Shc, GAB1, Cbl and others (see Fig. 1). The activated insulin receptor kinase (IRK) phosphorylates these substrate proteins on tyrosine residues.

Each of these phosphorylated proteins serve as docking proteins for other signaling proteins that contain Src-homology-2 domains (SH2 domains). Some SH2 proteins are enzymes, such as the P-Tyr phosphatase SHP2 (SH-PTP2). Other SH2 proteins, such as the p85 regulatory subunit of PI 3-kinase, Grb2 and APS, function as adaptor proteins for downstream effectors that further propagate the metabolic and the growth-promoting effects of insulin.

2.2.4 *Insulin Receptor Internalization and Degradation*

Upon insulin binding, IR rapidly internalizes. IR internalization is a multistep process. Stimulation of intrinsic Tyr kinase activity causes the redistribution of receptor-insulin complexes on the cell surface. Subsequently, complexes concentrate in clathrin-coated pits that act as internalization gates. Finally, the internalized receptors undergo sorting that determines whether they will be subjected to degradation in lysosomes or they will recycle back to the membrane surface but how this mechanism occurs is still unclear.

Interestingly, the cellular environment seems to regulate IR internalization. In particular, the extracellular matrix (ECM) to which the cells adhere seems to play an important role due to the interaction of the integrins on the cell surface with the ECM proteins. This interaction may results in the different organization of the cytoskeleton, thereby affecting the rate of endocytosis of IR.

2.3 PI3K-AKT/PKB PATHWAY

The PI3K-AKT/PKB pathway (see Fig.3) is the key mechanism regulating cell metabolism. As said before, its main role is to control glucose metabolism and to stimulate protein and lipid syntheses. In this section, the most important members of the pathway are examined giving particular attention to molecular interactions which govern the dynamics of metabolic response triggered by insulin [15, 32, 4].

The PI3K-AKT/PKB pathway can be briefly resummed as follows. Upon insulin binding and insulin receptor phosphorylation, IRS is phosphorylated on tyrosine residues that serve as docking sites for SH2 domain of PI3K, leading to its activation. Activated PI3K catalyzes the production of PI(3,4,5)P3 that activatse the Ser/Thr kinase of PDK1. PDK1 phosphorylates and activates two main downstream kinases, Akt and PKCζ that mediate translocation of GLUT4 on cell membrane. Moreover, Akt phosphorylates TSC1-TSC2 complex which releases the inhibition of RHEB. RHEB (in GTP-bound state) activates directly mTORC1 which phosphorylates two main downstream effectors, 4E-BP and S6K. 4E-BP and S6K promote ribosome biogenesis and the translation of proteins involved in cell growth and division.

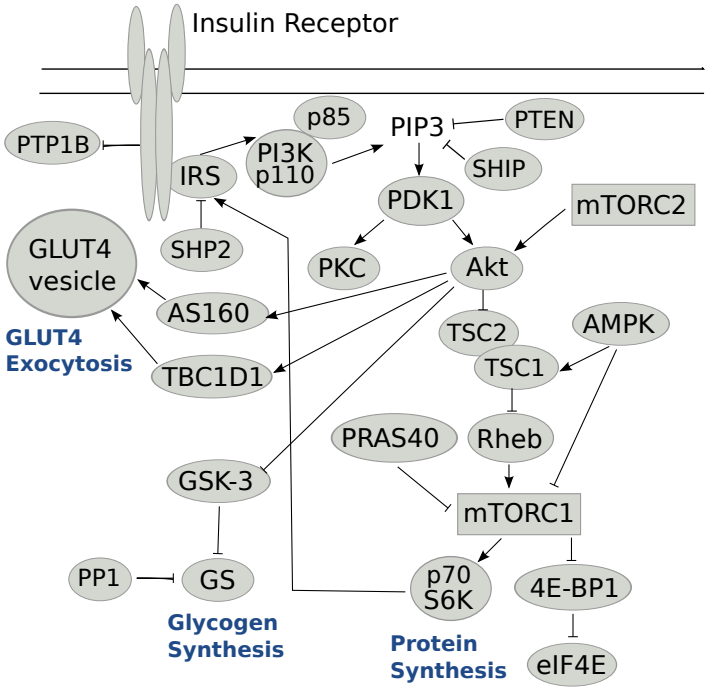

Figure 3: Schematic representation of PI3K-PKB/Akt pathway. Lines with arrows show activation and lines with blunt ends show inhibition.

2.3.1 *Insulin Receptor Substrate (IRS)*

IRS proteins contain a conserved *pleckstrin homology* (PH) domain, located at their amino-terminus, that serves to anchor the IRS proteins to membrane phosphoinositides and helps to localize the IRS proteins in close proximity of the receptor. PH domain of IRS proteins is flanked by a *P-Tyr binding* (PTB) domain. The PTB domain functions as a binding site to the NPXY motif of the juxtamembrane region of IR. C-terminal region of IRS proteins is poorly conserved and contains multiple Tyr phosphorylation motifs that serve as docking sites for SH2 domain-containing proteins like the p85α regulatory subunit of PI 3-kinase, GRB2, SHP-2 and others. In particular, IRS interacts, via phosphorylated YXXM motifs, with p85 subunit of PI 3-kinase thus activating it.

Moreover, IRS proteins contain several Ser/Thr phosphorylation sites that, if phosphorylated, reduce the ability of this protein to undergo Tyr phosphorylation by the IRK shutting off the insulin signaling. Serine phosphorylation of IRS is reckoned as one of the mechanisms involved in insulin resistance. In particular, Ser312 phosphorylation inhibits insulin action through disruption of IRS interaction with insulin receptor. IRS is also phosphorylated at Ser270, Ser307, Ser636 and Ser1101 by S6K (see 2.3.5) constituting one of the most important negative-feedback loop (NFL) present in the pathway and which inhibits upstream insulin signaling upon mTORC1 activation.

2.3.2 *Phosphatidylinositol 3-Kinase (PI3K) and PI(3,4,5)P3*

PI 3-kinase plays a central role in the metabolic and growth-promoting actions of insulin.

It is constitued of a p110 catalytic subunit and a p85 regulatory subunit. The regulatory subunit, which contains two SH2 domains, maintains the p110 catalytic subunit in a low-activity state. Activation of PI 3-kinase occurs upon direct interaction of the regulatory subunit with Tyr-phosphorylated YMXM and YXXM motifs of activated growth factor receptors, or with adaptor proteins such as IRS proteins. In particular, the association of p85–p110 complex with IRS molecules results in production of phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P3). The most relevant function of PI(3,4,5)P3 is the ability to interact with PH domains of phosphoinositide-dependent kinase 1 (PDK1), protein kinase B (PKB) and other signaling molecules. These interactions result in recruitment of these proteins to the plasma membrane triggering changes in their structure, function and their substrate availability. In the case of PDK1, binding of its PH domain to PI(3,4,5)P3 enables it to phosphorylate several downstream effectors, such as PKC and PKB/Akt and thus to further propagate the metabolic and growth-promoting functions of insulin.

Moreover, two important lipid phosphatases are commonly present in the system and whose mainly operate decreasing levels of PI(3,4,5)P3: PTEN and SHIP2. In particular, the phosphatase and tensin homologue (PTEN) acts removing the phosphate in the D3 position of the inositol ring from phosphatidylinositol PI(3,4,5)P3 to produce PI(4,5)P2. The SH2 domain–containing inositol-5-phosphatase (SHIP2) specifically hydrolyzes the 5-phosphate of PI(3,4,5)P3 to produce PI(3,4)P2. Thus both enzymes cooperate as antagonists of the PI3K-PKB/Akt pathway modulating cell cycle progression and cell survival.

2.3.3 *Atypical Protein Kinase C Isoforms (PKCζ and PKCλ)*

PI 3-kinase and PDK1 trigger the activation of the atypical PKC isoforms (PKCζ and PKCλ) via their N-terminus region. Two specific sites, Thr410 in the activation loop of the kinase domain and Thr560 in the turn motif, need to be phosphorylated for full activation of this molecule.

The main function of PKC concerns the regulation of GLUT4 translocation to the cellular membrane and subsequent induction of glucose transport inside the cell. PKC also plays an important role constituting a negative feedback control mechanism that serves to terminate insulin action. This feedback loop involves phosphorylation of IRS proteins and leads to IRS dissociation from IR, thereby terminating insulin signaling.

2.3.4 *Protein Kinase B (PKB/Akt), GSK3β and GYS*

PKB, also known as Akt, is one of the major substrates of PDK1. PKB is implicated in mediating numerous aspects of insulin action, including the regulation of glucose transport, glycogen synthesis, protein synthesis, the antilipolytic effects of insulin, as well as cell growth and cell survival induced by insulin.

PKB contains a PH domain that allows the binding to phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P3) following phosphatidylinositol 3-kinase (PI3K) activity and its targeting to the plasma membrane. PKB association with PI(3,4,5)P3 brings it to the proximity of plasma membrane facilitating phosphorylation of PKB at Thr309 in the T loop by PDK1. Thr309 phosphorylation is necessary and sufficient for PKB activation. However, maximal activation requires additional phosphorylation in its hydrophobic motif at Ser474 by PDK2, that has been recently identified with mTORC2 complex [Dalle Pezze et al. 2012]. Phosphorylation at these two sites stimulates the catalytic activity of PKB, resulting in the phosphorylation of a series of proteins that affect cell metabolism and growth, cell cycle entry, and cell survival. PKB phosphorylates many of its substrate proteins on Ser residues in a consensus RXRXXS/T site, and most of the known protein targets of PKB become inhibited by this event.

The three most important roles played by PKB are connected to the regulation of glucose uptake, glycogen and protein synthesis.

The first action is achieved by PKB mediating the translocation of GLUT4 glucose transporter to the cell membrane. This mechanism involves RAB GAP AS160 (TBC1D4) and RAL-GAP complex for the GLUT4 vesicles translocation and targeting to plasma membrane, and SNARE regulatory proteins, including Synip and CDP138, for the fusion event [5, 25]. Despite recent studies tried to uncover this mechanism, several steps in GLUT4 trafficking, including endocytosis, sorting and GSV formation are still unclear (see also 2.5).

The second one involves phosphorylation and inactivation by PKB of glycogen synthase kinase 3β (GSK3β) at Ser9. Phosphorylation at this site causes conformational change, preventing access of substrates to the active site. GSK3β acts phosphorylating the glycogen synthases (GYS) at Ser641, Ser645, Ser649 and Ser653. GYS functions transferring glycosyl residue from UDP-Glc to the non-reducing end of alpha-1,4-glucan which catalyses the conversion of glucose in glycogen. Since GSK3β phosphorylates and inactivates glycogen synthase (GYS), inhibition of GSK3β upon PKB phosphorylation promotes GYS activity and glycogen synthesis in response to insulin.

The third action implicates the phosphorylation of TSC2 (in complex with TSC1) (see 2.3.5) at Ser939 and Thr1462 activating mTORC1 signaling and leading to both phosphorylation of 4E-BP1 and p70-S6K.

Finally, PKB phasphatase activity can be prevented by Ser50 phosphorylation of PTP. This results in a negative modulation of PKB action that avoids the dephosphorylation of insulin receptor and the attenuation of insulin signaling.

2.3.5 *Mammalian Target of Rapamycin (mTOR), TCS1-TSC2, S6K and 4E-BP*

As already said, PKB phosphorylates and inhibits the tuberous sclerosis complex 1⁄2 dimer (TSC1–TSC2). TSC1-TSC2 complex acts as a GTPase-activating protein (GAP) for the small GTPase RHEB which is a direct activator of the protein kinase activity of mTOR.

The mammalian target of rapamycin (mTOR) is a central controller of cellular metabolism and also of cellular growth. In particular, mTOR regulates the anabolic and catabolic processes, including translation, ribosome biogenesis and autophagy, in response to hormones, growth factors (insulin), nutrients (amino acids), energy and stress signals [24, 33].

mTOR exists in a multiprotein complex, termed mTORC1. mTORC1 is constituted by DEP domain-containing mTOR-interacting protein (DEPTOR), mammalian lethal with SEC13 protein 8 (mLST8 or GβL), regulatory-associated protein of mTOR (RAPTOR) and 40 kDa Pro-rich Akt substrate (PRAS40) and, as already mentioned, it controls the cellular growth. Another form also exists, termed mTORC2, which is functionally and structurally distinct from mTORC1. mTORC2 has DEPTOR and mLST8 in common with mTORC1 but, differently from it, contains rapamycin-insensitive companion of mTOR (RICTOR) and the mammalian stress–activated map kinase interacting protein 1 (mSIN1). The main function of mTORC2 concerns the control of apoptosis as well as spatial growth via the actin cytoskeleton. mTORC2 is important due to its implication in the fully activation of Akt (see 2.3.4).

mTORC1 activation by growth factors, such as insulin, involves Akt-mediated phosphorylation of TSC1-TSC2 complex, which functions as *GTPase activating protein* (GAP) for the small GTPase Ras homologue enriched in brain (RHEB) , a potent activator of the protein kinase activity of mTORC1. Since GDP-loaded RHEB is unable to activate mTORC1, TSC1–TSC2 effectively shuts off mTORC1 signaling. Akt-mediated TSC2 phosphorylation is likely to inhibit its GAP activity for RHEB, thus promoting mTORC1 activation.

GTPase activating protein: *promotes hydrolysis of GTP to GDP by G proteins, resulting in their disactivation.*

The strongest link between amino acids and mTORC1 is due to the Rag family of small GTPases (RRAGA, RRAGB, RRAGC and RRAGD). The active Rag heterodimer physically interacts with RAPTOR, causing mTORC1 to cluster onto the surface of lysosomes, where the Rag GTPases reside. This relocalization may enable mTORC1 to interact with the small GTPase RHEB (in GTP-bound state) mediating the action of growth factor inputs.

Phosphorylation by both insulin and amino acids is carried out at Ser1261 and promotes autophosphorylation and activation of mTORC1.

mTORC1 pathway indirectly senses low ATP by a mechanism that involves the AMP-activated protein kinase (AMPK). Both AMP and ATP are allosteric regulators of AMPK. When the AMP:ATP ratio increases, AMPK phosphorylates TSC2, possibly stimulating the GAP activity of TSC1–TSC2 towards RHEB to inhibit mTORC1 signaling. Moreover, AMPK phosphorylates RAPTOR, causing it to bind 14-3-3 proteins, which leads to the inhibition of mTORC1 through allosteric mechanisms.

Activated mTORC1 up-regulates protein synthesis by phosphorylating key regulators of mRNA translation and ribosome synthesis. The main mTORC1 substrates are S6-kinase (p70-S6K), the translation initiation regulator 4E-binding protein (4E-BP) and the mTORC1-inhibitor PRAS40.

Activation of p70-S6K requires multiple phosphorylation events on serine/threonine residues. Activation appears to be first mediated by phosphorylation of multiple sites in the autoinhibitory domain which, disrupting the autoinhibitory mechanism, facilitates phosphorylation at Thr412. Phosphorylation at Thr412 is regulated by mTORC1 and maintained by an agonist-dependent autophosphorylation mechanism. When mTORC1 phosphorylates p70-S6K, this promotes its effects by phosphorylating or binding multiple proteins which collectively affect translation initiation and elongation processes. Between the most important p70-S6K substrates there is the ribosomal protein S6K which mediates the translation of mRNAs that have 5´ polypyrimidine tracts. Moreover, active p70-S6K is an important element in insulin pathway due to the phosphorylation and the inhibition it carries out on IRS. Precisely, p70-S6K phosphorylates IRS at multiple serine residues, resulting in accelerated degradation of IRS and thus it constitutes a *negative-feedback loop* (NFL) that inhibits upstream insulin signaling and that may be involved in insulin resistance.

Unphosphorylated 4E-BP1 suppresses mRNA translation. However, when phosphorylated by mTORC1 at Thr37, Thr46, Ser65 and Thr70, 4E-BP1 dissociates from eukaryotic translation initiation factor 4E (eIF4E), allowing eIF4E to recruit the translation initiation factor eIF4G to the 5´ end of most mRNA and thus enhancing its function as elongation initiation factor. eIF4E is involved in several cellular processes including enhanced translational efficiency, splicing, mRNA stability, and RNA nuclear export.

PRAS40 contributes to the inhibition of mTORC1 activity. In response to insulin, mTORC1 phosphorylates PRAS40 at Ser183 and this action causes the release of PRAS40 from the complex and relieves its inhibitory effect on mTORC1 which is allowed to phosphorylates the remaining substrates. Also Akt phosphorylates PRAS40

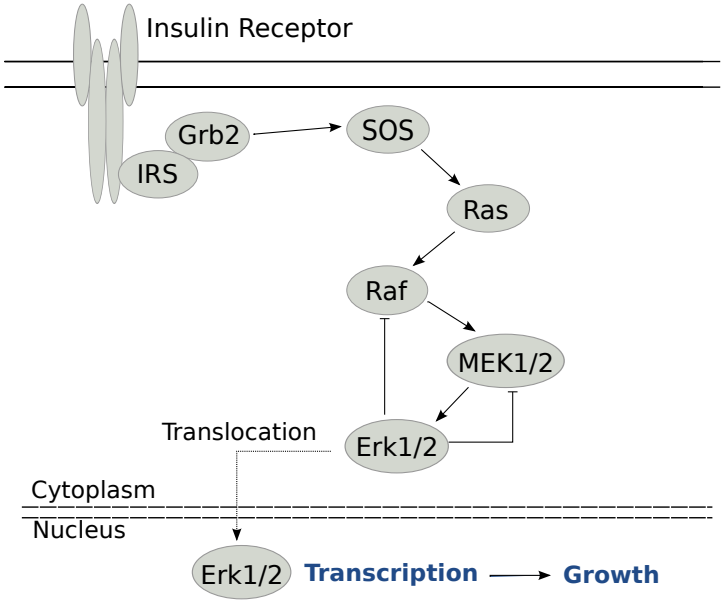

Figure 4: Schematic representation of Ras/MAPK pathway. Solid lines with arrows show activation, solid lines with blunt ends show inhibition and dotted line with arrow shows translocation.

which causes it to bind to 14-3-3 proteins and prevents it from inhibiting mTORC1.

2.4 RAS/MAPK PATHWAY

MAPK pathway (see Fig.4) mainly promotes cell survival, cell division and cell motility. In this section we examine in detail the most important elements contributing to this pathway focusing in their interactions in order to unveil the complexity of this network [15] [Uniprot] [PhosphoSitePlus].

Ras/MAPK pathway can be briefly reassumed as a multistep process startingfrom an initial GTPase-regulated kinase (MAPKKK) that phosphorylates and activates an intermediate kinase (MAPKK) that, in turn, phosphorylates and activates an effector kinase (MAPK). These three steps correspond respectively to Ras GTPase and protein kinases Raf, MEK and ERK.

2.4.1 *Grb2, SOS and Ras*

Insulin triggers Ras/MAPK pathway upon binding of growth factor receptor binding protein 2 (Grb2) to Tyr-phosphorylated IRS (see 2.3.1). Grb2 can also associates directly with activated receptors, such as EGFR and PDGFR, via its SH2 domain. Grb2 binds to mammalian son of sevenless (SOS) by means of SH3 domains in Grb2, forming a complex. Subsequently SOS, which is a guanine nucleotide exchange

protein, catalyzes the exchange of guanosine diphosphate (GDP) for GTP on Ras (a small GTP binding protein), thus resulting in activation of Ras. The prenylated form of Ras binds the inner leaflet of the plasma membrane and, upon SOS activation, binds the N-terminal region of the protein kinase Raf, recruiting Raf to the plasma membrane.

2.4.2 *Raf*

Ras-GTP recruits Raf to the membrane, thereby promoting its activation. The inactive conformation of Raf is maintained by autoinhibitory interactions occurring between the N-terminal regulatory domain and the C-terminal catalytic domain and by the binding of a 14-3-3 protein that contacts two phosphorylation sites, Ser259 and Ser621. Upon mitogenic stimulation, Ras allows the release of autoinhibition and permits the subsequent phosphorylation of activating sites (Ser338, Tyr341, Thr491 and Ser494) which yields a fully active kinase.

Raf is a serine/threonine-protein kinase that acts as a regulatory link between the membrane-associated Ras GTPases and the MAPK/Erk cascade, and this critical link functions as a switch for cell fate decisions including proliferation, differentiation, apoptosis, survival and oncogenic transformation. Raf activation initiates the mitogen-activated protein kinase (MAPK) cascade that comprises the sequential phosphorylation of the dual-specific MAPK kinases (MEK1 and MEK2) and the extracellular signal-regulated kinases (ERK1 and ERK2).

Moreover, Raf is phosphorylated on Ser29, Ser43, Ser289, Ser296, Ser301 and Ser642 by ERK2 leading to kinase inactivation and thereby constituing a negative feedback loop.

2.4.3 *MEK1/2*

As said before, Ras protein mediates the activation of Raf which, in turn, activates ERK kinases MEK1 and MEK2. Activation occurs through phosphorylation of Ser218 and Ser222 on both MEK1 and MEK2. MEK1 and MEK2 function catalyzing the concomitant phosphorylation of a threonine and a tyrosine residue in a Thr-Glu-Tyr sequence located in ERK1 and ERK2 and leading to their activation.

MEK1 is also the target of negative feedback regulation loop by its substrate kinase ERK2. ERK2 phosphorylates MEK1 on Thr292, thereby facilitating dephosphorylation of the activating residues Ser218 and Ser222.

2.4.4 *Erk1/2*

ERK1 (MAPK1) and ERK1 (MAPK3) are the two MAPKs playing the most important role in Ras/MAPK cascade.

ERK1 is phosphorylated by MEK1 and MEK2 on Thr202 and Tyr204 whereas ERK2 on Thr185 and Tyr187. Phosphorylations of both sites are required for ERK1 and ERK2 activation and this event causes dramatic conformational changes that allows their interaction with substrates. ERK2 is also phosphorylated on Ser29 by SGK1 and this results in ERK2 activation by enhancing its interaction with MEK1 and MEK2.

About 160 substrates have already been discovered for ERKs. Many of these substrates are localized in the nucleus confirming ERKs participation in transcript regulation events. However, other substrates are found in the cytosol as well as in some other cellular organelles, and those are responsible for processes such as translation, mitosis and apoptosis.

The phosphorylation sites responsible for the nuclear location are the autophosphorylation sites at Thr207 for ERK1 and at Thr190 for ERK2 (in addiction to phosphorylation sites at Ser246 and Ser248 in the kinase insert domain, KID).

ERK1 and ERK2 nuclear translocation, together with their ability to phosphorylate transcription factors, constitutes an important connection between cytoplasmic and nuclear events.

One of the most important transcription factors phosphorylated and activated by ERKs is Elk-1, which leads to the induction of gene expression.

2.5 CBL/CAP PATHWAY

Cbl/CAP pathway (see Fig. 5) controls the glucose transporter type 4 (GLUT4) traffic to the cell membrane in a PI3K–independent manner and, cooperating with the PI3K-PKB/AKT pathway, ensures the efficient regulation of glucose transport operated by insulin-action [15, 5, 25].

In absence of insulin, the majority of GLUT4 is distributed between endosomes, the trans-Golgi network (TGN) and heterogeneous tubulo–vesicular structures that consist of endosomal sorting intermediates and specialized GLUT4 storage vesicles (GSVs). In basal state, only about 5% of the total GLUT4 transporter pool is found on the cell surface but , in response to insulin or exercise, the glucose uptake is tenfold increased. The failure of GLUT4 translocation to the plasma membrane in response to insulin is an early step in the development of insulin resistance and type 2 diabetes mellitus.

Cbl/CAP pathway is triggered by insulin binding to its cognate receptor. Upon insulin binding, the adaptor protein with pleckstrin ho-

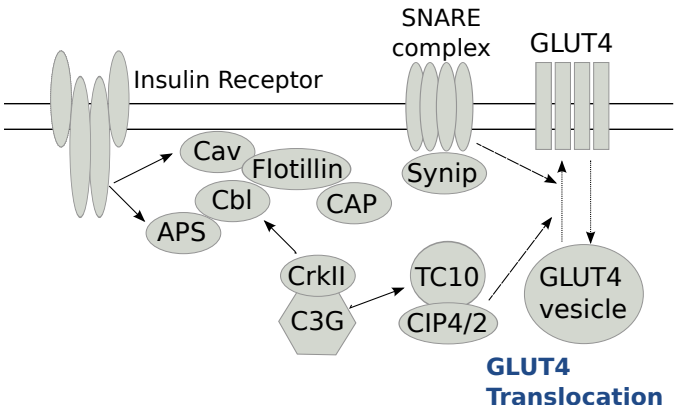

Figure 5: Schematic representation of Cbl/CAP pathway. Solid lines with arrows show activation, dotted lines with arrows show tentative stimulatory action and thin dotted lines with arrows show translocation.

mology (PH) and Src homology 2 (SH2) domains (APS) is recruited and binds with high affinity to the Tyr-phosphorylated insulin receptor. Subsequently, APS undergoes Tyr-phosphorylation and recruits a complex that comprises the protooncogene c-Cbl and cCblassociated protein (CAP) by means of the atypical SH2 domain of Cbl. This triggers insulin receptor-catalyzed Tyr phosphorylation of cCbl. CAP protein is recruited with Cbl to IR by means of the interaction of its third C-terminal SH3 domain with Cbl and another direct interaction with APS through its N- and C-terminal SH3 domains. Upon Tyr phosphorylation of Cbl, the Cbl–CAP complex is released from the receptor and translocates to lipid raft domains in the plasma membrane. This mechanism is mediated by the interaction of the SoHo domain of CAP with the protein flotillin. Phosphorylated Cbl then interacts with the adaptor protein CRK, which is in complex with the guanyl nucleotide-exchange (GEF) protein C3G. Once translocated into lipid rafts, C3G comes into proximity with Rho family GTPase TC10 and catalyses the exchange of GTP for GDP, resulting in activation of TC10. Active TC10 interacts with several effector proteins that regulate GLUT4 vesicle exocytosis. One TC10binding protein is CDC42interacting protein 4 (CIP4) which forms a stable complex with the RAB GEF GAPEX5. This regulates the activity of RAB5 family GTPases that are involved in GLUT4 vesicles retention and translocation. Another TC10 effector is EXO70 which is a subunit of the exocyst tethering complex and has been implicated in GLUT4 vesicle targeting. Together, these molecular targets play an important role within Cbl/CAP signaling in the release of intracellular GSV retention mechanism and in GLUT4 vesicle targeting to the plasma membrane.

Moreover, the movement of GLUT4 transporters along cytoskeletal tracks may also be mediated by molecular motors such as the unconventional myosin Myo1c, which contains a motor domain, three IQ

motifs, and a carboxy-terminal cargo domain. Proper structural organization of plasma membrane caveolin and a functional clathrin may also have a significant role in GLUT4 trafficking regulating GLUT4 endocytosis rate and, thus the overall rate of GLUT4 recycling.

3

RULE-BASED MODELING OF BIOCHEMICAL SYSTEMS

3.1 INTRODUCTION

Signal transduction inside cells is carried out by network of interacting signal mediators and is often a very complex process. Complexity derives not only from the enormous amount of different molecules involved in the process but also from the presence of numerous feedbacks and feedforward loops, both negative or positive, concerning the pathway itself, and ample crosstalks involving distinct pathways. This high degree of complexity issues new challenges in understanding how cellular signaling works in detail and thus new and more powerful tools have to be introduced beyond the simple reasoning on experimental data.

Insulin signaling pathway represents a typical example of complex network due to the numerous molecules constituting it and the several types of interactions between these elements. Hence, in order to analyse in detail the dynamics which characterizes this network it is necessary to make use of mathematical tools able to model the system comprising the majority of its biochemical reactions. The resulting model can be used to make predictions of signaling pathway in physiological state to gain new knowledge about the process or may be used to test the system varying some quantities and parameters which describe the model to obtain information about network malfunctioning and mimicking pathological condictions. The most common approach makes use of a system of ordinary differential equations (ODEs) to model the kinetics of the molecules populating the system. This method works well in case of simple networks comprising just few molecules. When the system under investigation is particularly large and involves a great amount of interactions between molecules, the ODEs approach is no more efficient from both implementational and computational point of view. For this reason the rule-based modeling approach has been introduced providing an easy way to implement the system and new simulation tools.

3.2 COMPUTATIONAL MODELING TECHNIQUES

As already said, experimental biology approach alone does not provide a reasonable strategy to a detailed analysis of complex signaling networks. Therefore, some computational methods have been introduced to be used in combination with experimental results. In this

way computational methods provide useful information to guide experimental design obtaining more informative experiments, whereas data collected from experiments help implementation and refinement of computational models, providing more accurate predictions of biological systems.

Computational methods are typically classified in two main categories: structural and dynamic network analysis [27]. *Structural network analysis* gives information about network connectivity and deduce some properties of the global network as well as some functions of individual proteins. An example of structural model analysis is the boolean network simulation method which uses only signaling network connectivity information to predict the speed of signal transduction and which produced encouraging results in gene regulatory networks [17]. In boolean model the signal transduction is discretized and can be either present or absent, This two-states modeling represents a extreme simplification of the underlying biochemistry and cannot be used to predict fine time courses of protein concentrations. In order to deal with this limitation signaling Petri net-based simulation has been introduced. Petri net-based simulator still represents a non-parametric modeling approach but is a more fine-grained way to model and simulate the dynamics of signaling network. This method was used to study MAPK and AKT signaling network downstream of EGFR in cancer cells showing good results [23].

Dynamic network analysis makes use, in addiction to network connectivity information, of kinetic parameters characterizing the biochemical reactions in order to obtain more detailed information regarding the network behavior. Kinetics parameters, also known as *kinetic rate constants*, give information about the speed of a chemical reaction occurring between species and thus about how fast reactants are transformed into products. In this way dynamic models permits to obtain the time courses of the proteins involved in signal transduction. Even though dynamic models are more informative respect to the structural models, they require additional knowledge of many numerical values describing the kinetics of the reactions. Thus implementation of dynamic models easily encounter problems due to the limited data available from experiments. Moreover, the measurements of kinetics parameters often depend on cell type, experimental conditions and other factors. Thus these values may differ of a order of magnitude from test to test and in some cases they may remain under-determined or non-uniquely estimated. Despite these difficulties dynamic models are widespread and they have been currently used by many research groups with sucessful results.

Nowadays, the most widely used technique to implement dynamic models consists in the use of a set of ordinary differential equations (ODEs) able to mime the system. Concentration changing in time of each chemical species present in the system is described by a differ-

ential equation exploiting the knowledge of kinetics rate constants. Unfortunately model implementation using ODEs presents remarkable problems in describing large and complex networks and thus it makes necessary the introduction of severe simplifications in the system. However in recent years a new approach has been introduced which makes use of simple reaction rules instead of a set of differential equations: the rule-based modelling. Rule-based modeling represents a revolutionary modeling approach which permits to overcome the limitations imposed by ODEs modeling with an outstanding saving of time and reducing the risk of errors during model implementation phase. Rule-based modeling permits also new simulation methods making use of stochastic algorithms. A deeper comparison between these two modeling approaches and all the peculiarities of rule-based modeling are presented in the following Section.

In addiction to scientific literature, a good deal of information useful for model implementation is available from numerous online databases. Generally, databases give lots of heterogeneous information about signaling pathways and about elements constituting those pathways. Some of them focus on networks connectivity, such as KEGG [20], PANTHER [35] and Reactome [40], unveiling the complex connections between different elements and the presence of feedbacks, feedforward loops and crosstalk mechanisms. They usually provide some information about activation and inhibition actions operated by distinct proteins avoiding many details about these interactions. These databases give all the necessary information for the structural modeling approach and represent a starting point for dynamic network implementation. However, more detailed information concerning the structure and the function of each protein can be obtained from other databases, such as PhosphoSitePlus [32] , UniProt [4], HPRD [38], OMIN [37], UCSD signaling gateway [9] and NetPath [18]. In these databases it is possible to find information about the presence of particular molecular domains and motives, phosphorylation and binding sites with all the related interacting molecules. Unfortunately, no database already exists collecting kinetic parameters of the chemical reactions. This makes difficult the research of these important values and implies the examination of a huge amount of scientific literature.

However, in case of incomplete information, it is still possible to infer the unknown parameters training the model against the available experimental data. One way implies the use of measured time courses of phosphorylated proteins as input for a optimization algorithm able to compute the unknown quantities. This approach gives reasonable solutions when the missing parameters are few but introduce higher uncertainties as the quantity of unknown values grows. One example is represented by the stochastic algorithm termed *particle swarm optimization* (PSO) that is used to identify the insulin-like growth factor (IGF-1) signaling network in cancer cell line [19]. However, parame-

ters estimation and optimization tools are out of the purpose of this study in which we decide to rely on already existing models including the kinetic parameters of all reactions and the initial concentrations of all chemical species.

3.3 RULE-BASED MODELING

We mentioned the importance to catch every site-specific detail of protein-protein interactions in order to achieve the complete comprehension of how a signaling pathways works. In addiction to difficulties arising from incomplete experimental data, another impediment is present due to the intrinsic complexity in describing every interaction occurring in the network. Signaling proteins contain multiple functional components and several sites of post-translational modification. As result, interactions among signaling proteins may generate a myriad of protein complexes and post-translational modification states. For instance, a protein containing n phosphorylation sites can be found in up to 2^n distinct states. This feature has been called *combinatorial complexity* and has been recognized as a significant challenge to our understanding about cellular regulation [11].

In conventional model specification using a set of ODEs, each *chemical species* that potentially populates the system and each reaction that can occur must be explicitly specified and this may generate a combinatorial number of coupled differential equations making model implementation tedious, prone to errors or even impossible. Since the same molecule found in a different state represents a different chemical species, to describe completely each molecule as many of differential equations as the number of its dinstict states are necassary (without considering any kind of interaction with other molecules). Taking the example above, a protein containing n phosphorylation sites is described by 2^n differential equations. Moreover, the amount of differential equations increases rapidly when molecules with more than one site bind together forming in a complex which can be found in a myriad of different possible states.

The most common solution adopted to overcome this limitation entails the simplification of the model. Proteins with multiple phosphorylation sites are represented with a single phosphorylation site which reassumes the properties of all sites, termed *virtual phosphorylation site* [12]. The same procedure is adopted when a complex sequential multi-step process is taking place by substituting it with a single step. Obviously these simplifications conflicts with the knowledge of cellular biochemistry and there is no proof that these assumptions could not afflict model predictions in some way.

In order to deal with the issue of combinatorial complexity, a new useful tool has been introduced with the purpose to specify all the reactions arising from molecular interactions in a more efficient and

compact way. *Rule-based modeling* approach [21] is based on the key assumption that the characteristics of molecular interactions depend on local properties of the protein involved and this *modularity* mostly determines the network dynamics. According to this assumption, structure of a reaction occurring in the network can easily be defined by means of a rule. A rule represents a class of reactions involving reactants with common components and component properties. The important simplification of the rule-based modeling approach is that all the reactions within a class are assigned the same rate law. An impressive example offered by rule-based modeling approch is the one described in [14] in which EGF receptor signalling is implemented using 70 rules and which generate over 10^{23} dinstict molecular species and thus 10^{23} differential equations would be necessary using ODEs modeling.

In conclusion, both rule-based model and ODEs model provide a representation of chemical kinetics but they are different in the model specification procedure. While in ODEs model the modeler must state *explicitly* which chemical species populate the system and how these are connected and influence each other, in a rule-based model, the modeler must state only the interactions occurring in the system and their contextual dependencies [21].

Rule-based modeling approach is implemented by means of two major languages, termed Kappa [link] and BioNetGen [link]. Due to its facility of use we chose BioNetGen language to implement our models but the differences compared with Kappa language are neglegible.

3.4 BIONETGEN LANGUAGE

BioNetGen[16] is a set of software tools for rule-based modeling and is a mnemonic for "Biological Network Generator". The software not only generates reaction networks starting from reaction rules, but also simulates such networks using a variety of methods (see 3.4.1).

In order to understand how this software practically implements a biological system using rule-based modeling, we report a simple example taken from [16] that will clear the basic concepts lying beneath this approach.

BIONETGEN LANGUAGE (BNGL) ENCODING EXAMPLE Fig. 6 depicts the implementation of a simple network in all its essential parts. Molecules are implemented as structured objects that can be constituted by different components. These components represent functional elements of proteins and may have associated states representing covalent modifications or conformations (e.g. phosphorylated/unsphosphorylated state, active/inactive state etc.). The expression reported in Fig.6(A) indicates the definition of two *molecules*

types. `A(a)` represents the presence of molecules type `A` having only the componet `a`. `B(b1,b2~U~P)` represents molecules type `B` having two components, `b1` and `b2`. In particular, the expression `b2~U~P` indicates that component `b2` may be found in two possible states, `~U` and `~P`, which can be used to indicates respectively the unphophorylated and the phosphorylated states.

Components of distinct molecules can link together forming a bond, thereby building complexes of assembled molecules. *Patterns* can be used to select particular molecular attributes. In particular, the pattern `B(b1)` shown in Fig. 6(B) selects molecules `B` having the binding site `b1` completely free despite of the phosphorylation and binding status of `b2` component.

Rules are needed to specify the biochemical transformations that can potentially take place in the system. The term "transformation" is used instead of "reaction" to indicate that the same reaction rule is applied to a selected set of reactant species and not necessarily to only one chemical species. This approach is worthy if, as already said, the chemical reactions mainly depend on the local properties of protein components. Modularity feature implies that the same reaction rule can be used to describe the transformation of different chemical species sharing common components and component properties. This selection method of the chemical species permits to write many complicated chemical reactions in a set of few rules making more compact and efficient the implementation of the model. In rules, the "pattern matching" is accomplished specifying two essential parts:

- the protein components directly modified by the transformation (*reaction center*) and
- the components and components states needed for the selection of chemical species (*reaction context*).

In Fig. 6(C) are reported 3 different rules and their reaction centers are underlined. Rule `1 A(a) + B(b1) <-> A(a!1).B(b1!1) kp1,km1` represents a common reaction of binding and unbinding between molecules `A` and `B`. Bond formation involves the components `a` of molecule `A` and component `b1` of molecules B. Since nothing is specifyed about component `b2`, the reaction between the two molecules takes place indipendently from the phosphorylation state of `b2`. Since this reaction is reversible (`<->`) its kinetics is described by means of two parameters, `kp1` and `km1`, which respectively determine the speed of the forward and backward reactions. Rule `2 B(b1!+,b2~U) -> B(b1!+,b2~P) k2` describes a typical phosphorylation event. Component `b2` of molecule `B` changes its state from `~U` to `~P` according to kinetics parameter `k2`. The expression `b1!+` specifies that the reaction takes place only if site `b1` is bound to another molecule. Finally, rule `3 B(b2~P) -> B(b2~U) k3` describes a desphosphorylation event. Component `b2` of molecules `B` changes its state from `~P` to `~U` according to

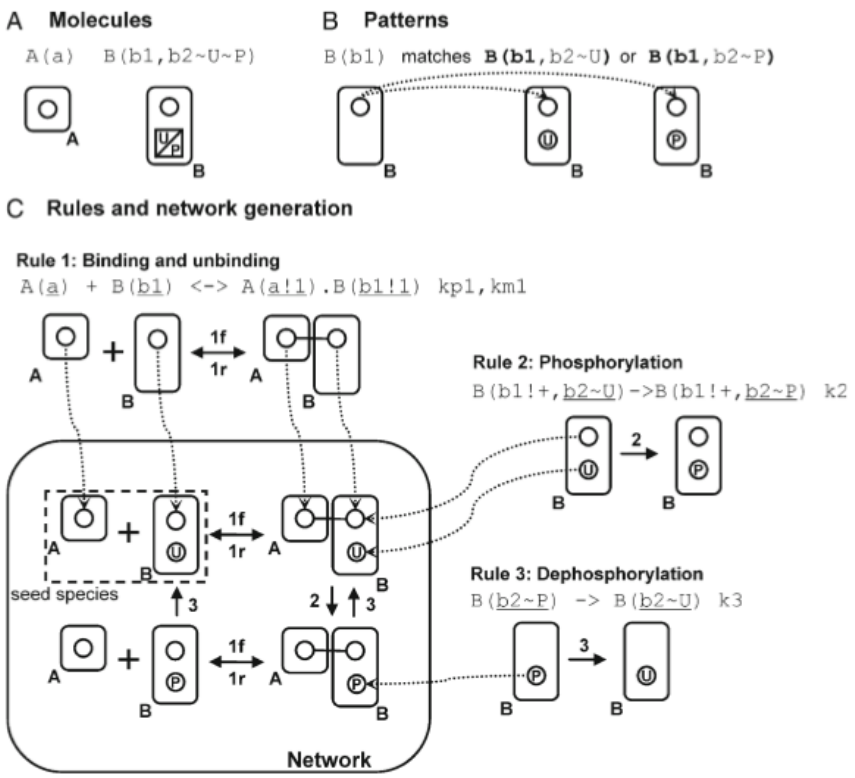

Figure 6: BioNetGen language encoding example. Figure taken from[16]

rate `k3`. Since nothing is specifyed about `b1`, the reaction takes place indipendently from presence of bond in component `b1`.

Starting from the seed species the rules are applied generating other new species and thus new reactions. The process continue iteratively until no new reactions are found or some other stopping criteria are satisfied creating the complete network model. Fig. 6(C) depicts this phase starting from the seed species `A(a)` and `B(b1,b2~U)` up to the creation of a complete network containing all the possible species which may populate the system according to rules 1, 2 and 3.

This simple but explicative example depicts some of the typical events occurring in a biological pathway. Phosphorylation and dephosphorylation of a molecular site and binding and unbinding of two proteins are fundamental events taking place in a signaling pathway and governing its dynamics. However, this set of few rules can be extended using more molecules with different components and components properties opening the door to implementation of very complex systems.

A BioNetGen input file is mainly constituted by six sections including all the information about the biological system and they are briefly presented here:

- `(parameters)` Define the parameters that govern the dynamics of the system such as rate constants, the values for initial concentrations of chemical species, compartment volumes and physical constants used in unit conversions. The synthax of a line in the `parameters` block is

```
[index] parameter [=] value
```

  where square brackets indicate optional elements, `parameter` is a string consisting of only alphanumeric characters plus the underscore character and containing at least one nonnumeric character. `value` may be either a number in integer, decimal or exponential notation or a formula involving numbers and other parameters in C-style math synthax;

- `(molecule types)` Define molecules, including components and allowed component states. As already said, molecules are structured objects composed of components able of binding to each other. Components typically represent physical part of proteins such as domains and motives and may also be associated with a list of state labels, which represent states or properties of the components. Examples of component states modeled using state labels may concern conformation (e. g. open or closed), phosphorylation status and location (e. g. extracellular space, membrane, cytoplasm). The synthax of a line in the `molecule types` block is

```
[index] moleculeTypes
```

  where `moleculeTypes` has the synthax for a BioNetGen species;

- `(seed species)` Define the initial state of system such as initial chemical species to which rules are applied. The synthax of a line in the `seed species` block is

```
[index] species [initialPopulation]
```

  where `species` has the synthax for a BioNetGen species described above and `initialPopulation` is a number or formula that specifies the amount of the species present at the start of the first simulation (default is zero);

- `(observables)` Define model outputs, which are functions of the population levels of multiple chemical species sharing a set of properties. The synthax of a line in the `observables` block is

```
[index] [observableType] observableName pattern1[,pattern2]
```

  where `observableType` is either `Molecules` or `Species` (default is `Molecules`), `observableName` is a valid name for a BioNetGen observable, and each `pattern` is a valid BioNetGen pattern;

- `(functions)` Define the functions describing the kinetic rate constants which cannot be described by a simple constant value but which show a kind of dependencies from concentrations of some chemical species present in the system. The synthax of a line in the `functions` block is

```
f() = formula
```

  where `f()` is the name of the function and `formula` is an mathematical expression involving the concentration of one or more chemical species which have to be necessarily listed in the `observables` block;

- `(reaction rules)` Define rules describing which chemical species are involved in a transformation and which are the consequences of this transformation. Each rule is similar to a standard chemical reaction notation having four basic elements: reactant patterns, an arrow, product patterns and a rate law specification. The synthax of a line in the `reaction rules` block is

```
[index] rPattern1 [+rPattern2] ... arrow pPattern1 [+
    pPattern2] ... rateLaw1[,rateLaw2] [command1] ...
```

  where each `Pattern` is a valid BioNetGen pattern, `arrow` is one of "->" or "<->", each `rateLaw` is a parameter or a rate law function. Rules may transform a selected set of reactant species by adding or deleting molecules or bonds and by changing component state labels. An example for each case is reported below

```
begin reaction rules
        #Add a bond
        A(a) + B(b) -> A(a!1).B(b!1) k_a
        #Delete a bond
        A(a!1).B(b!1) -> A(a) + B(b) k_d
        #Change a component state label
        A(Y~P) -> A(Y~U) km3
        #Add a molecule
        I() -> I() + A(a,Y~U) ksynth
        #Delete a molecule
        A() -> Trash() kdeg
end reaction rules
```

- `(actions)` Perform two basic types of actions, generate the chemical reaction network implied by the model specification and simulate this network using different methods (see 3.4.1). An example of actions commonly used are reported below

```
#actions
generate_network({overwrite=>1});

#kinetics
```

```
writeSBML({});
simulate_ode({t_end=>120,n_steps=>120});
resetConcentrations();
simulate_ssa({suffix=>ssa,t_end=>120,n_steps=>120});
```

`generate_network` command directs BioNetGen to generate a network of species and reactions through iterative application of the rules starting from the set of seed species. This behaviour can be overridden setting the option `overwrite=>1`. `writeSBML` indicates that the network is written to an SBML file. `simulate_ode` command initiates a simulation of the dynamics implementing and numerically solving a set of ODEs. `t_end` and `n_steps` specify respectively the end time for the simulation and the number of steps at which the results are written to the output files. `resetConcentrations` command restores the concentrations to the initial values. `simulate_ssa` command initiates a stochastic simulation. `suffix=>ssa` command appends "_ssa" to the basename for output files of the simulation.

All the information about BioNetGen software and syntax are fully described in [16].

3.4.1 *BioNetGen Simulation Tools*

It is worthy to clear some aspects concerning the simulation tools offered by BioNetGen software.

One possible method (the one presented above) entails the iterative application of rules to the set of defined seed species in order to generate a network before the simulation starts. Subsequently, the simulation may be carried out either by numerically solving ODEs or by implementing a stochastic simulation algorithm (SSA). SSA implies a kinetic Monte Carlo simulation using the Gillespie algorithm [8] and produces stochastic trajectories representing the concentrations of observables species.

Alternatively, the rules may be applied during the simulation as the set of populated species grows using a procedure that has been called "on-the-fly" network generation and simulation. As a first step only the reactions involving seed species are initially generated. Then, BioNetGen detects when a reaction event occurs that populates one or more species to which rules have not been previously applied and automatically expands the network through rule application. This method is useful when a network is potentially unrestricted such as in polimerization processes which allows aggregates formation of any size. For this reason to generate a complete network before prediction starts may be dangerous unless some stopping criteria are specified. An example of this simulation method is reported below

```
#simulation on-the-fly
```

```
generate_network({overwrite=>1,max_iter=>1});
simulate_ssa({t_end=>50,n_steps=>20});
```

where `max_iter` set to 1 indicates that only the reactions involving seed species are initially generated.

For all the mentioned methods the simulation cost scales with the network size, hence simulation of large-scale reaction networks may become impractical. In particular, the CPU time required for model simulation increases exponentially as the number of network reactions grows. In order to overcome this computational limit, *network-free methods* have been introduced such as NFsim. NFsim [7] guarantees a constant cost of simulation per reaction event and thereby implies a linear increase of the CPU time with the number of reaction events in the system. NFsim generalized the rule-based version of Gillespie's stochastic simulation algorithm (SSA) and guarantees very similar results compared with SSA but with a outstanding saving in time. An example of NFsim is reported below

```
#network-free simulation
simulate_nf({suffix=>nf,complex=>1,t_end=>50,n_steps=>20})
```

4

COMPUTATIONAL MODEL OF INSULIN SIGNALING PATHWAY

4.1 INTRODUCTION

The purpose of this thesis is to implement a computational model describing the insulin signaling pathway in its essential parts in order to simulate the most important features of insulin action.

As seen in previous Chapters, insulin signaling pathway is constituted by great deal of molecules having distinct functions and interacting with each other in different ways. Due to this complexity, the knowledge concerning how insulin signaling pathway really works cannot be achieved by means of simple reasoning on data obtained from experiment. Thus, new and more powerful computational tools have been recently introduced in order to be used in combination with experimental data. Insulin pathway can be viewed as a complex network of interacting elements that can be described by means of a mathematical model. In order to overcome combinatorial complexity that often limits the implementation of large models, the rule-based modeling approach may be used. As said in the previous chapter, BioNetGen language is a software that permits to implement this new modeling approach and to encode the model by means of a simple set of rules.

We decided to realize the model of insulin signaling pathway starting from three models published in literature during the past years. These models consider only some parts of the entire insulin pathway but their combination is sufficient to cover the essential elements responsible for the most of the insulin action. By means of the information derived from these models we entirely reconstructed the two major sub-systems of the insulin signaling pathway, the PI3K-PKB/Akt and the Ras/MAPK pathways.. Unfortunately, no information were available for Cbl/CAP pathway due to the lack of knowledge in some essential steps involved in GLUT4 translocation.

Models taken from literature were originally implemented using sets of ordinary differential equations (ODEs) and thus they needed to be translated using the rule-based modeling approach. First, we realized the models separately and we verified that the resulting simulations matched those showed in the papers (see Appendix). Subsequently, we integrate them creating a single complete model including as many elements as possible of the insulin signaling pathway.

The flow chart depicted in Fig.7 describes all the elements constituting the complete model and mostly define their relationships.

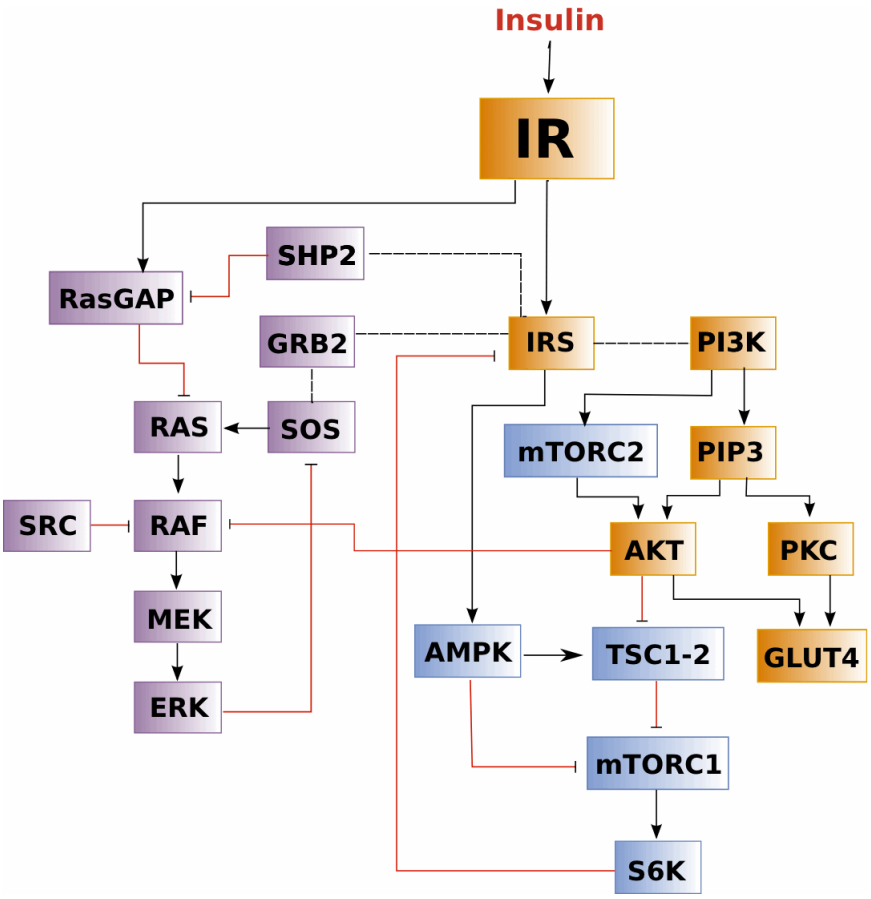

Figure 7: Flow chart of complete model of insulin signaling pathway. The colors indicates which subsystem elements belong to and are orange for PI3K-Akt/PKB subsystem, blue for mTOR subsystem and purple for Ras/MAPK subsystem. Solid lines with arrows show the activation or tyrosine phosphorylation of proteins and lipids. Dotted lines represent protein-protein complex formation. Red lines with blunt ends show inhibition.

The colors of the blocks indicate the sub-system which elements belong to and thus the main reference papers, and are orange for PI3K-PKB/Akt subsystem (Quon's paper [22]), blue for mTOR subsystem (Thedieck's paper [34]) and purple for Ras/MAPK subsystem (Kholodenko's paper [12]). In particular, complete insulin model makes use of

- 23 differential equations and 55 parameters (42 kinetics parameters and 13 initial conditions) taken from Quon's paper;
- 17 differential equations and 26 parameters (17 kinetics parameters and 9 initial conditions) taken from Thedieck's paper, and
- 19 differential equations and 54 parameters (45 kinetics parameters and 9 initial conditions) taken from Kholodenko's paper.

The resulting model is constituted by total 156 parameters and 52 reactions rules involving 67 dinstict chemical species. Since in traditional ODEs approach concentration of each chemical species is decribed using a ordinary differential equations, model implementation using ODEs would required the use a set of 67 equations (instead of 52 rules using rule-based modeling). Hence, rule-based modeling approach permits to describe the model making use of a lower number of rules respect to that needed using traditional approach. This is an example of the advantages carried by rule-based modeling but, in some particular cases, the number of rules respect to ODEs may be more significantly reduced leading to a remarkable saving in time and a lower risk of errors during whole the implementation procedure.

In the following sections, we present the most important details concerning complete model implementation using BioNetGen language. The sections will treat separately three main portions in which we divide the system and which correspond to information mostly taken from the same paper. In addiction, for each section we collected useful information about the referential paper including all the basic assumptions and simplifications introduced by research groups.

4.2 PI3K-PKB/AKT MODEL

To represent insulin receptor and PI3K-PKB/Akt signaling pathway discussed in Chapter 2 (Sections 2.1 and 2.2) we used the model presented in the paper "A Mathematical model of metabolic insulin signaling pathways" by Quon M. J et al. in 2002 [22]. The purpose of the study was to realize a comprehensive model ables to represent the knowledge available from experimental data and to gain new insight regards to molecular mechanisms underlying the insulin signal transduction pathway. In particular, the model was realized using concentrations and kinetics parameters taken from literature mostly

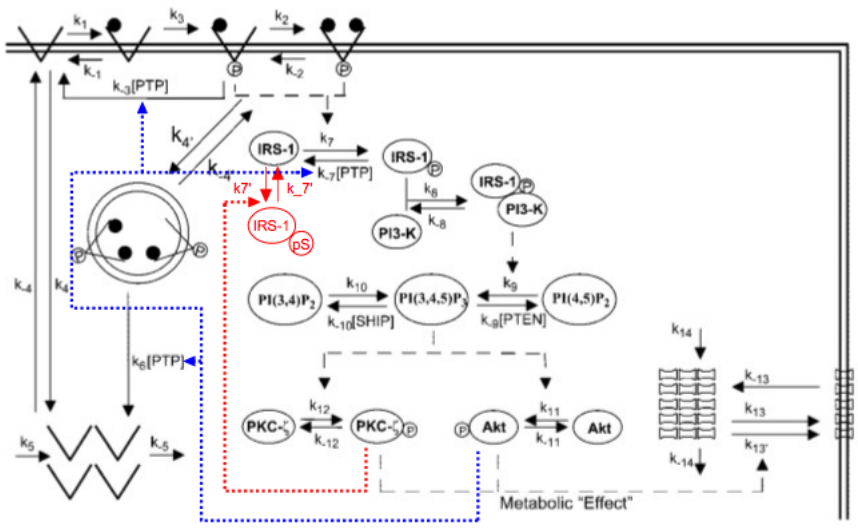

Figure 8: Complete model of metabolic insulin signaling pathway with feedbacks. Positive and negative feedback loops are indicated by dotted lines. PKC-ζ serine phosphorylates IRS-1 to create a negative feedback (red line) and Akt phosphorylates PTP1B to create a positive feedback (blue line). Figure adapted from [22].

refering to experiments on 3T3-L1 adypocytes, and then used to run predictions of the system under different condictions. The model is depicted in Fig.8 where the compartments represent the molecules and the molecule states considered in the model and the arrows indicates the interactions between these elements. Despite it is dated 2002, the model contains several of the most accepted mechanisms and intermediates in the downstream signaling controlling the glucose uptake, such as activation of phosphoinositide 3-kinase (PI3K), phosphorylation of protein kinase B (PKB), and translocation of GLUT4 to the plasma membrane. This model has reigned for several years and has been used by many researcher groups. Here we present a brief description of the model followed by some details about its implementation in the original paper and using BioNetGen language.

4.2.1 *Model description*

Upon insulin binding, insulin receptor (IR) undergoes autophosphorylation and enhances its tyrosine kinase activity. Subsequently, insulin receptor substrate-1 (IRS-1) is phosphorylated on tyrosine residues. Tyrosine residues of IRS-1 function as docking sites for downstream SH2 domain containing proteins, such as the p85 regulatory subunit of phosphatidylinositide 3-kinase (PI3K). The p85 binding to phosphorylated IRS-1 results in activation of the p110 catalytic subunit of PI3K that catalyzes production of phosphatidylinositol 3,4,5-trisphosphates [PI(3,4,5)P3] that activates the Ser/Thr kinase 3-phosphoinositide-

dependent protein kinase (PDK)-1. PDK-1 phosphorylates and activates two major downstream kinases, Akt and protein kinase C (PKC)-γ, that mediate the translocation of GLUT4. The protein tyrosine phosphatase (PTPase) PTP1B is also included and it negatively regulates insulin signaling pathway by dephosphorylating IR and IRS-1.

4.2.2 *Model implementation in original paper*

In Quon's paper two mathematical models of metabolic insulin pathway were presented. The work started implementing a complete model without feedback mechanisms. This first model was then extended in order to incorporate informations about feedback loops resulting in a second and more accurate model. We decided to reassumed here the main features of the complete model with feedback loops which is the one we refer in our work.

Quon's model uses previously validated models of insulin receptor binding kinetics, receptor recycling and GLUT4 translocation. These were integrated with new subsystems generating a complete model counting 23 state variables. The majority of model parameters and rate constants characterizing the steps of the metabolic insulin signaling pathway were based on experimental data taken from literature. In the cases where mechanisms regulating interactions between signaling elements were not fully understood, they modeled these interactions as simple linear functions. This simplification was introduced for the rate of IRS-1 phosphorylation in response to activated insulin receptors, the rate of PI(3,4,5)P3 generation in response to activated PI3K, the rate of PKC-ζ and Akt phosphorylation in response to increased levels of PI(3,4,5)P3, and the rate of exocytosis for GLUT4 in response to phosphorylated PKC-ζ and Akt.

Despite the effort in including as many elements as possible in the model, the complete mechanism regulating the metabolic action of insulin remains still unclear and requires the introduction of marked simplifications. Thus insulin controlling glucose uptake is simply represented by a direct control of Akt and PKC-ζ on the trafficking machinery for GLUT4.

However, the model includes a detailed description of insulin receptor response upon insulin binding. Insulin receptor represents a critical point within the signaling cascade because its kinetics markedly influences the following downstream events. Nowadays many theories and mathematical models are available in literature which provide mechanistic and quantitative explanations to experimental data. Models focus mainly on modeling the binding event of insulin to its receptor, the internalization mechanism of insulin receptor and some of them offer interesting theories about receptor aggregation and cooperation phenomena.

Quon's group performed several tests on their insulin model. First they generate model time courses and insulin dose-response curves observing a good overall match between experimental data and model simulations. In particular, PKC-ζ activation time course obtained by model with feedbacks shows the typical biphasic feature, absent in the simulation of the model without feedbacks. Moreover, to exploit the model ability to represent pathological conditions, they ran model simulations examining the effects of increasing level of PTPase on GLUT4 translocation.

All the details concerning PI3K-PKB/Akt model implementation using BioNetGen language are presented in the following section.

4.2.3 *Model implementation using BioNetGen language*

First, in BioNetGen code we define in the `molecule types` block all the molecules which populate this part of the system, and that are listed below

```
begin molecules types
        I(bs)
        IR(bs,alpha,alpha,Y~u~p,loc~ex~c)
        IR_Gen()
        IRS1(bs,YXXM,Y~u~p,S~u~p)
        PI3K(SH2)
        PI45()
        PI345()
        PI34()
        PKC(T~u~p) #T410_T560
        Akt(T308~u~p,S473~u~p)
        GLUT4(loc~c~ex)
        GLUT4_Gen()
        Trash()
```

For matter of clearness, we present PI3K-PKB/Akt pathway according to the method suggested in Quon's paper dividing the system in four main parts. These subsystems are discussed in the following paragraphs and treat the binding and recycling mechanisms of insulin receptor, the signaling of all the downstream molecules and the translocation of GLUT4 transporters to the cell membrane.

INSULIN RECEPTOR BINDING SUBSYSTEM. Referring to the molecules listed above, `I(bs)` indicates insulin molecules having the binding site `bs` for insulin receptor. `IR(bs,alpha,alpha,Y~u~p,loc~ex~c)` represents insulin receptor having the binding site `bs` for RasGAP (see Section 4.4.3) and two `alpha` subunits for the binding of insulin molecules. IR can also be found in unphosphorylated (`Y~u`) or phosphorylated (`Y~p`) state on tyrosine residues and can change its location from cell surface (`loc~ex`) to cytoplasm (`loc~c`).

The reaction rules presented below describe exactly insulin receptor binding event. When a first molecule of insulin binds to it (rule 1), IR is rapidly phosphorylated (rule 2). Subsequently, IR may either binds another molecule of insulin (rule 3) or dissociate from it returning to the free receptor state (rule 4). Binding of a second molecule of insulin does not affect the phosphorylation state of IR, whereas IR dephosphorylation occurs when insulin diffuses off leaving IR completely free. Since IR occupancy and phosphorylation are tightly coupled, IR in unbound phosphorylated state is not included in the model.

```
begin reaction rules
        # Receptor binding 1st insulin molecule
        1 IR(bs,alpha,alpha,Y~u,loc~ex) + I(bs) <-> IR(bs,alpha
            !1,alpha,Y~u,loc~ex).I(bs!1) k1,k_1
        # Receptor phosphorylation
        2 IR(bs,alpha!+,alpha,Y~u,loc~ex) -> IR(bs,alpha!+,alpha,
            Y~p,loc~ex) k3
        # Receptor binding 2nd insulin molecule
        3 IR(alpha!+,alpha,Y~p,loc~ex) + I(bs) <-> IR(alpha!+,
            alpha!1,Y~p,loc~ex).I(bs!1) k2,k_2
        # Receptor unbinding and dephosphorylation (on the cell
            membrane)
        4 IR(bs!?,alpha!1,alpha,Y~p,loc~ex).I(bs!1) -> IR(bs,
            alpha,alpha,Y~u,loc~ex) + I(bs) ptp1()
```

In addiction, protein tyrosine phosphatases which dephosphorylate insulin receptor are inserted in the model. Phosphatases effect is explicitly represented using the function `ptp1()` defined in the `functions` block, that modulates receptor dephosphorylation rate and whose action depends on activated Akt, as shown below

```
ptp1()=(k_3)*(1-0.25*((100-(Akt_pT308+Akt_pS473+Akt_phosph)*(100/
    Akt_init))/(100/11)))
```

where `Akt_pT308` represents the observable quantity `Akt(T308~p,S473~u)`, `Akt_pS473` represents `Akt(T308~u,S473~p)` and `Akt_phosph` represents `Akt(T308~p,S473~p)` (for details see section).

INSULIN RECEPTOR RECYCLING SUBSYSTEM. Receptor life cycle is represented as a multi-step process including synthesis, degradation, ligand-induced endocytosis and exocytosis to cell membrane. The reaction rules characterizing this subsystem are shown below

```
# Phosphorylated receptor internalization/externalization
5 IR(bs!?,alpha!+,Y~p,loc~ex) <-> IR(bs,alpha!+,Y~p,loc~c) k4p,k
    _4p
# Receptor unbinding and dephosphorylation (inside the cell)
6 IR(bs,alpha!+,Y~p,loc~c) -> IR(bs,alpha,alpha,Y~u,loc~c) ptp2()
# Free receptor internalization/externalization
7 IR(bs,alpha,alpha,Y~u,loc~ex) <-> IR(bs,alpha,alpha,Y~u,loc~c)
    k4,k_4
```

```
# Receptor synthesis
8 IR_Gen() -> IR_Gen() + IR(bs,alpha,alpha,Y~u,loc~c) k5
# Receptor degradation
9 IR(bs,alpha,alpha,Y~u,loc~c) -> Trash() k_5
```

Ligand-induced endocytosis is applied to IR in phosphorylated state located on cell surface, binding one or two insulin molecules (rule 5). Internalized IR are then desphosphorylated and incorporated into intracellular pool (rule 6) which is continuously repopulated by new synthesized insulin receptors (rule 8). From intracellular pool, free and dephosphorylated receptors can be either degraded (rule 9) or translocated back to cell membrane (rule 7). In order to mime the production and degradation of intracellular IR, two fictitious molecules, termed `IR_Gen()` and `Trash()`, where defined in the `molecule types` block. `IR_Gen()` was created with initial concentration equal to 1 and retaining the rates of synthesis of the original model. In order to keep fixed the concentration of this fictitious molecule, the transformation imposes the presence of `IR_Gen()` as among the reactants as among the products (rule 8). `Trash()`, instead, was created with initial concentration of 0 and it functions as fictitious compartment collecting all the molecules (not only intracellular IR) degraded from the system (rule 9). In order to mime the phosphatases effect leading to the dephosphorylation of intracellular IR, the function `ptp2()` was defined having similar expression to `ptp1()`.

POSTRECEPTOR SIGNALING SUBSYSTEM. `IRS1(bs,YXXM,Y~u~p,S~u~p)` indicates IRS-1 molecules which contain the binding site `bs` for Grb2 and SHP2 (see Section 4.4.3) and the `YXXM` motif for PI3K. IRS1 can be phosphorylated either on tyrosine residues (`Y~u~p`) or on serine residues (`S~u~p`). `PI3K(SH2)` represents PI3K molecules having the binding site `SH2` for IRS-1. Then, `PI45()`, `PI345()` and `PI34()` respectively represent the lipids PI(4,5)P2, PI(3,4,5)P3 and PI(3,4)P2. The production of PI(3,4,5)P3 regulates two major downstream kinase, PKC-ζ and Akt. `PKC(T~u~p)` represents PKC-ζ molecules which may be phosphorylated on threonine residues (`T~u~p`). `Akt(T308~u~p,S473~u~p)` indicates Akt molecules which may be phosphorylated in Thr308 residue by PDK1 and in Ser473 by mTORC2 (see Section 4.3.3).

All the reaction rules involved in the postreceptor signaling subsystem are presented below

```
# IRS-1 phosphorylation/dephosphorylation in Tyr
10 IRS1(bs,YXXM,Y~u,S~u) <-> IRS1(bs,YXXM,Y~p,S~u) f7(),ptp3()
# IRS-1_PI3-K complex formation (PI3-K activation)
11 IRS1(YXXM,Y~p,S~u) + PI3K(SH2) <-> IRS1(YXXM!1,Y~p,S~u).PI3K(
    SH2!1) k8,k_8
# IRS-1 phosphorylation/dephosphorylation in Ser
12 IRS1(bs,YXXM,Y~u,S~u) <-> IRS1(bs,YXXM,Y~u,S~p) k7p,k_7p
# lipids formation
13 PI45() <-> PI345() f9(),k_9*PTEN
```

```
# lipids formation
14 PI34() <-> PI345() k10,k_10*SHIP
# Akt phosphorylation/dephosphorylation
15 Akt(T308~u) <-> Akt(T308~p) f11(),k_11
# PKC phosphorylation/dephosphorylation
16 PKC(T~u) <-> PKC(T~p) f12(),k_12
```

Upon insulin stimulation, activated insulin receptors phosphorylate IRS-1 on tyrosine residues (rule `10`), which then binds and activates PI3K (rule `11`). IRS may also be phosphorylated on serine residues (rule `12`). The rate constant for IRS-1 phosphorylation (`f7()`) depends on the concentration of phosphorylated insulin receptors on the cell surface, and is shown below

```
f7()=k7*IR_bound
```

where `IR_bound` represents the observable quantity `IR(alpha!+,Y~p,loc~ex)` which selects both once- and twice-bound phosphorylated surface receptors.

Activated PI3K converts the substrate PI(4,5)P2 to the product PI(3,4,5)P3 (rule `13`). The rate constant for PI(3,4,5)P3 (`f9()`) conversion is modeled as a linear function depending on the amount of activated PI3K using the same procedure adopted for `f7()`. PI(3,4,5)P3 is also produced starting from PI(3,4)P2 (rule `14`). We included in the model the action of the lipid phosphatases PTEN and SHP2 which convert PI(3,4,5)P3 respectively to PI(4,5)P2 and PI(3,4)P2. These lipid phosphatases are modeled using factors `PTEN` and `SHIP` which correspond to the their relative activity inside the cell and are assigned a value of 1 under physiological conditions.

The level of PI(3,4,5)P3 regulates the activation of downstream kinases Akt (rule `15`) and PKC-ζ (rule `16`). In particular, Akt is phosphorylated and thus activated only on Thr308 residue.

GLUT4 TRANSLOCATION SUBSYSTEM. `GLUT4(loc~c~ex)` indicates GLUT4 transporters which may be located either in cytoplasm (`loc~c`) or translocate to cell membrane (`loc~ex`). The reaction rules governing the subsystem are presented below

```
# GLUT4 translocation
17 GLUT4(loc~c) <-> GLUT4(loc~ex)  f13(),k_13
# GLUT4 synthesis
18 GLUT4_Gen() -> GLUT4_Gen() + GLUT4(loc~c) k14
# GLUT4 degradation
19 GLUT4(loc~c) -> Trash() k_14
```

Under basal conditions, GLUT4 transporters recycle between an intracellular compartment and the cell surface. On insulin stimulation, a separate pool of intracellular GLUT4 is recruited to the cell surface. Thus, the insulin-stimulated exocytosis rate may be modeled as a sum of two contributes, the effect on basal conditions and the effect

depending on Akt and PKC-ζ activation, thus due to insulin stimulation. However, Akt and PKC-ζ contribute to GLUT4 translocation in different way. Precisely, 80% of the metabolic insulin signaling effect is attributed to PKC-ζ and 20% to Akt. This effect is implemented in BioNetGen by means of function `f13()` which describes the rate of GLUT4 translocation (rule `19`) according to the following expression

```
f13()=k13+k13a*(100-(Akt_pT308+Akt_pS473+Akt_phosph)*(100/Akt_
    init))+k13b*PKC_phosph
```

where `100-(Akt_pT308+Akt_pS473+Akt_phosph)*(100/Akt_init)` indicates the concentration of once- and twice- phosphorylated Akt (for details see Section 4.5) and `PKC_phosph` indicates the concentration of phosphorylated PKC. The contribute of these two main quantities is weighted according to previous assumptions using parameters `k13a` and `k13b` and the overall effect deriving from insulin is summed to that one under basal condition (`k13`).

GLUT4 transporters constitutes a cytoplasmatic pool which is continuously repopulated by new synthesized molecules and reduced by their degradation. In order to model this mechanism, we introduce the fictitious molecule `GLUT4_Gen()` which, likewise intracellular IR, permits a continuous synthesis of GLUT4 transporters. To delete from the system the degraded GLUT4 transporters we used `Trash()` compartment introduced before in same way we adopted for degraded IR in intracellular pool.

4.3 MTOR MODEL

Quon's model may be extended in order to include mammalian target of rapamycin (mTOR) and some other aspects of insulin signaling network not considered before that were discussed in Chapter 2 (Section 2.2.5). The paper "A modelling-experimental approach reveals insulin receptor substrate (IRS)-dependent regulation of adenosine monosphosphate-dependent kinase (AMPK) by insulin" of Thedieck K. et al. published in 2012 [34] modeled mTOR network, covering both mTOR complexes and insulin and nutrient inputs. As already said in Chapter 1, mTOR kinase responds to growth factors, nutrients and cellular energy status and is a central controller of cellular growth. mTOR exists in two multiprotein complexes (mTORC1 and mTORC2) that are embedded into a complex network.

The model was implemented starting from a previous model published by Dalle Pezze P. et al. the same year [33]. The model was extended including another essential element which was recently discovered to respond not only to energy deprivation but also to insulin stimulation, AMPK. Adenosine monophosphate-dependent kinase (AMPK) is activated by energy deprivation and shuts off adenosine 5'-triphosphate (ATP)-consuming anabolic processes, in part via

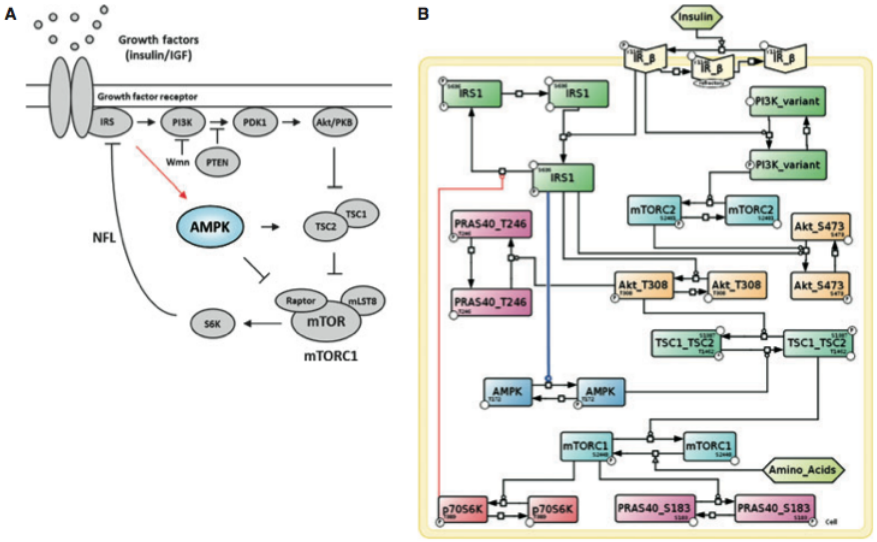

Figure 9: **(A)** Schematic representation of the insulin-induced mTOR pathway, including IRS-dependent AMPK induction. **(B)** Graphical representation of insulin-mTOR-AMPK network including all the variable states and the interactions present in the implemented model. Figure taken from [34]

the inactivation of mTORC1. AMPK module was included to generate the to date most comprehensive data-driven dynamic AMPK-mTOR network model. In order to define which was the interaction connecting AMPK to insulin network, Thedieck's group compared simulations generated using different hypothesis which suggest the insulin receptor substrate (IRS) as the most probable linking element of the network. A simple and intuitive representation of the model is depicted in Fig.9(A) whereas Fig.9(B) shows in details all the molecules and molecule states included in the model and their relations.

The model included the description of the majority of the elements constituting this pathway. However, mTORC1 activation is a complex mechanism which is also controlled by other two molecules, Rag GTPases and Ras homologue enriched in brain (RHEB). When activated by amino acids, Rags physically interact with RAPTOR (a component od mTORC1 complex), causing mTORC1 to cluser onto the surface of endosomes and lysosomes. This relocalization may enable mTORC1 interaction with RHEB which is an essential activator of mTORC1 and that is controlled by growth factors. This mechanism is described in detail by Zoncu R. et al. [24] but has not already found a reasonable mathematical description.

Here we present a brief description of the model followed by its implementation in the original paper and in BioNetGen language.

4.3.1 *Model description*

Insulin signaling induces a kinase cascade through the insulin receptor (IR), IR substrate (IRS), class I phosphoinositide 3-kinases (PI3Ks), phosphoinositide-dependent protein kinase 1 (PDK1), and the AGC kinase Akt (also known as PKB). Akt inhibits the tuberous sclerosis complex 1/2 (TSC1-TSC2) dimer, which is the inhibitory GTPase-activating protein (GAP) for Rheb. Through this cascade, Akt stimulates mTORC1 activity. The major mTORC1 substrates are the AGC kinase p70 ribosomal protein S6 kinase (p70S6K), the translation initiation regulator 4E binding protein (4E-BP), and the proline-rich Akt substrate PRAS40, which is an inhibitor of mTORC1. By binding mTORC1, PRAS40 contributes to the inhibition of mTORC1 activity. In response to insulin, Ser183 of PRAS40 is phosphorylated by mTORC1, which releases PRAS40 from the complex and relieves its inhibitory effect on mTORC1, allowing mTORC1 to phosphorylate its downstream substrates p70S6K and 4E-BP and promote cellular growth. Furthermore, there is a negative feedback loop (NFL) that inhibits upstream insulin signaling upon mTORC1 activation. Indeed active p70S6K phosphorylates and inhibits IRS, which prevents activation of PI3K in response to insulin. Akt is also phosphorylated by mTORC2 at Ser473 that, in combination with PDK1 action, permits to Akt to be twice phosphorylated and thus to be fully active. Moreover, to inhibit mTORC1 signaling, AMPK multiply phosphorylates and activates TSC2 when cellular energy is low. In addiction, AMPK also directly phosphorylates the essential mTORC1 component Raptor on two serine residues. This phosphorylation induces 14-3-3 to bind to Raptor and is required for mTORC1 inhibition by energy deprivation.

4.3.2 *Model implementation in original paper*

To parameterize the network model, Thedieck's group generated semi-quantitative dynamic phosphorylation immunoblot data for network components along the signaling cascade. They analyzed HeLa cells under starvation conditions (cell deprived of amino acids and growth factors for 16 hours) and then they stimulated them with amino acids and insulin monitoring the dynamics of network components from 1 minute up to 2 hours.

They previously decided to minimize the amount of detail of the model since a precise dynamics of each interacting elements could not be defined due to the high number of parameters and the difficulty in obtaining sufficient experimental data. Thus they chose the essential regulation mechanisms governing the dynamic behavior of the system. They also selected molecules and interactions that could be detected by reliable measurements. Finally, they calibrated the model parameters by means of the experimental time courses obtain-

ing a total of 26 reaction rate constants. In particular, the monitored chemical species were Tyr1146-phosphorylated IR, Ser636-phosphorylated IRS1, Ser473- and Thr308-phosphorylated Akt, Ser1387- and Thr1462-phosphorylated TSC1-TSC2, Ser2448- and Ser2481-phosphorylated mTOR, Thr246- and Ser183-phosphorylated PRAS40, Thr389-phosphorylated p70S6K and Thr172-phosphorylated AMPK.

The resulting model present some critical simplifications of the underlying biochemistry governing insulin signaling pathway. The first one concerns the modeling of IR which can be found only in three possible states: unphosphorylated receptor, phosphorylated receptor and internalized receptor (refractory state). The second simplification concerns the modeling of IRS1-PI3K complex activation that, in order to minimize the amount of state variables of the model, was implemented by phosphorylated IR controlling independently tyrosine phosphorylation of IRS1 and PI3K activation. Moreover, the possible states assumed by IRS1 are simplified. Unphosphorylated IRS1 can be phosphorylated only on tyrosine residues, tyrosine phosphorylated IRS1 is then phosphorylated on serine residues and finally only serine phosphorylated IRS1 is subject to complete dephosphorylation. However, the above mentioned simplifications do not concern the part of the system that we are going to use in our model, that starts with Akt phosphorylation and, through TSC1-TSC2 complex and mTORC1, ends with IRS phosphorylation on serine residues by S6K. A particular issue represents Akt molecule defined in the previous section. Differently from Quon's group, Thedieck's model takes into account not only Akt phosphorylation on Thr308 by PDK1 but also its phosphorylation on Ser437 by PDK2, recently identified with mTORC2. This additional information about Akt phosphorylation state represents an improvement respect to Quon's model but the way in which Thedieck's group implement it presents a critical simplifications. Probably for sake of simplicity, Thedieck's group decided to distinguished between the molecules of Akt which may be potentially phosphorylated in Thr308 from those which may be phosphorylates in Ser437. In other words, there are four possible states that Akt may assumes, that are:

- Thr308-unphosphorylated Akt,
- Thr308-phosphorylated Akt,
- Ser437-unphosphorylated Akt, and
- Ser437-phosphorylated Akt.

To deal with this issue we simply created only one Akt molecule (the one defined in the previous section) which may potentially be phosphorylated either in Thr308 or in Ser437 or in both sites. The same procedure was adopted for PRAS40 which may be phosphorylated in Ser183 and in Thr246.

4.3.3 *Model implementation using BioNetGen language*

First of all we determine all the molecules populating mTOR subsystem, that are listed below

```
mTORC2(S2481~u~p)
AMPK(T172~u~p)
TSC1_TSC2(T1462~u~p,S1387~u~p)
Amino_Acids()
mTORC1(S2448~u~p)
p70S6K(T389~u~p)
PRAS40(T246~u~p,S183~u~p)
```

`mTORC2(S2481~u~p)` represents mTORC2 molecules having the phosphorylation site on Ser2481. mTORC2 phosphorylation is controlled by PI3K activation, or rather by IRS1-PI3K complex formation (rule `20`) and it also presents spontaneous dephosphorylation (rule `21`). As already said, Akt may be phosphorylated on Thr308 and on Ser437 respectively by PDK1 and PDK2. Akt phosphorylation on Thr308 was presented in the previous section where PDK1 activation was substituted by directed control of PI(3,4,5)P3 lipids production. Whereas Akt phosphorylation on Ser437 is accomplished by mTORC2 complex (PDK2) (rule `22`) and it presents spontaneous dephosphorylation (rule `23`).

```
#mTORC2
20 mTORC2(S2481~u) + IRS1(YXXM!1,Y~p,S~u).PI3K(SH2!1) -> mTORC2(S
    2481~p) + IRS1(YXXM!1,Y~p,S~u).PI3K(SH2!1) p16
21 mTORC2(S2481~p) -> mTORC2(S2481~u) p15
#Akt
22 Akt(S473~u) + mTORC2(S2481~p) -> Akt(S473~p) + mTORC2(S2481~p)
     p11
23 Akt(S473~p) -> Akt(S473~u) p10
```

`AMPK(T172~u~p)` indicates AMPK molecules having phosphorylation site on Thr172. The phosphorylation event is modulated by the concentration of phosphorylated IRS-1 on tyrosine residues (rule `24`) and it presents spontaneous dephosphorylation (rule `25`).

`TSC1_TSC2(T1462~u~p,S1387~u~p)` represents TSC1-TSC2 complex having two phosphorylation sites on Thr1462 and on Ser1387. TSC1-TSC2 can be either phosphorylated on threonine residue or on serine residue and the last one site determines the activation of the complex. Hence, Thr308- phosphorylated Akt determines the phosphorylation TSC1-TSC2 on Thr1462 (rule `26`), thereby inhibiting its action. AMPK, instead, modulates the phosphorylation of TSC1-TSC2 complex on Ser1387, thereby enhancing its action (rule `27`).

```
#AMPK
24 AMPK(T172~u) + IRS1(YXXM,Y~p,S~u) -> AMPK(T172~p) + IRS1(YXXM,
    Y~p,S~u) p7
25 AMPK(T172~p) -> AMPK(T172~u) p8
```

```
#TSC1-TSC2
26 TSC1_TSC2(T1462~u,S1387~p) + Akt(T308~p) -> TSC1_TSC2(T1462~p,
    S1387~u) + Akt(T308~p) p24
27 TSC1_TSC2(T1462~p,S1387~u) + AMPK(T172~p) -> TSC1_TSC2(T1462~u
    ,S1387~p) + AMPK(T172~p) p23
```

`mTORC1(S2448~u~p)` represents mTORC1 complex having phosphorylation site on Ser2448. As already said, mTORC1 activation depends on nutrients input. This effect is modeled defining `Amino_Acids()` molecules having fixed concentration which positively regulate mTORC1 complex activation (rule 28). mTORC1 action is also controlled by growth factors, thus by insulin. This effect is implemented by direct TSC1-TSC2 complex regulating mTORC1 dephosphorylation, thus acting on its inhibition (rule 29).

```
#mTORC1
28 mTORC1(S2448~u) + Amino_Acids() -> mTORC1(S2448~p) + Amino_
    Acids() p14
29 mTORC1(S2448~p) + TSC1_TSC2(T1462~u,S1387~p) -> mTORC1(S2448~u
    ) + TSC1_TSC2(T1462~u,S1387~p) p13
```

The last element constituting this negative feedback loop (NFL) is S6K which is indicated as `p70S6K(T389~u~p)`. S6K has one phosphorylation site on Thr389 and its activation is positively controlled by mTORC1 complex (rule 30). S6K presents also spontaneous dephosphorylation (rule 31). S6K exerts its action on a critical upstream element which belongs to the PI3K/Akt subsystem presented before, IRS-1. Indeed, S6K regulates IRS-1 phosphorylation on serine residues (rule 32) thus impeding the phosphorylation on tyrosine residues which is necessary to IRS1-PI3K complex formation and PI3K activation.

```
#p70S6K
30 p70S6K(T389~u) + mTORC1(S2448~p) -> p70S6K(T389~p) + mTORC1(S
    2448~p) p18
31 p70S6K(T389~p) -> p70S6K(T389~u) p17
#IRS-1
32 IRS1(YXXM,Y~u,S~u) + p70S6K(T389~p) -> IRS1(YXXM,Y~u,S~p) + p
    70S6K(T389~p) p5
```

Another elements is also included in the model, PRAS40.

`PRAS40(T246~u~p,S183~u~p)` presents two phosphorylation sites on Thr246 and on Ser183, and their activation depends respectively on Thr308- phosphorylated Akt (rule 33) and on Ser2448- phosphorylated mTORC1 complex (rule 35). However, PRAS40 has no control on other reactions in the mTOR subsystem but it has been used by Thedieck's group as simple readout during the experiments.

```
#PRAS40
33 PRAS40(T246~u) + Akt(T308~p) -> PRAS40(T246~p) + Akt(T308~p) p
    22
34 PRAS40(T246~p) -> PRAS40(T246~u) p20
```

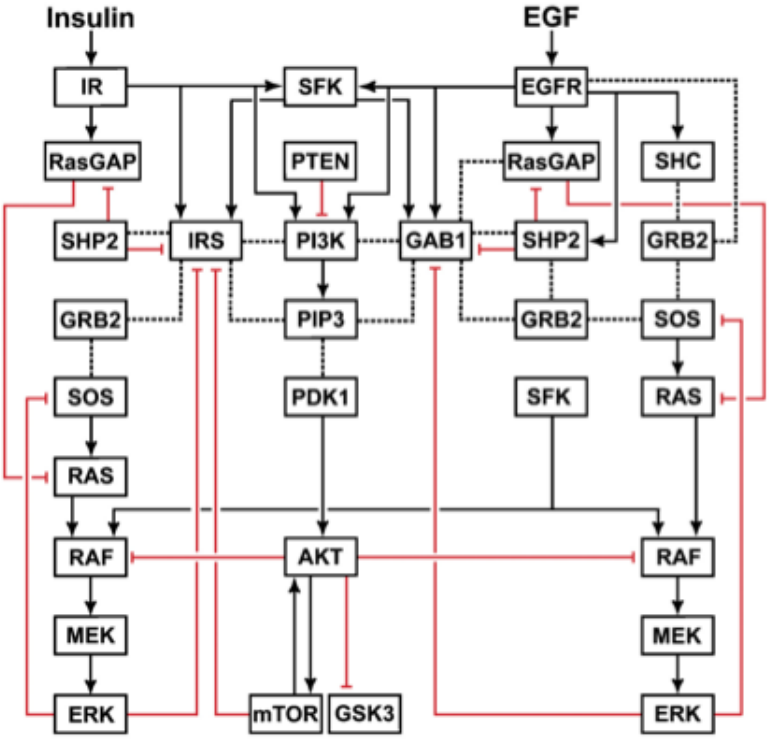

Figure 10: Schematics representing the signal propagation through EGFR and IR signaling networks. Tyrosine phosphorylation of proteins and lipids are represented by solid lines with arrows, protein-protein and protein-lipid interactions with dotted line, and inhibitions with red lines with blunt ends. Figure taken from [12].

```
35 PRAS40(S183~u) + mTORC1(S2448~p) -> PRAS40(S183~p) + mTORC1(S
    2448~p) p21
36 PRAS40(S183~p) -> PRAS40(S183~u) p19
```

4.4 RAS/MAPK MODEL

One of the most studied signaling network is the epidermal growth factor (EGF) signaling and, in particular, the activation of the Ras/MAPK pathway. The Ras/MAPK pathway is also controlled by insulin and has been discussed in Chapter 2 (Section 2.3). EGF and insulin signaling pathways are thus sub-networks in a larger network of overlapping and crosstalking signaling pathways. One of the most interesting problems that has not been disclosed yet about signaling mechanisms regards crosstalk versus pathway specificity, in other words, how downstream targets distinguish between different inputs despite sharing the same intermediary signals. An attempt in this direction was realized by Zielinski R. et al. [3] which implemented a Boolean network model that combines three major pro-survival signaling path-

ways: epidermal growth factor receptor (EGFR), insulin-like growth factor-1 receptor (IGF-1R), and insulin receptor (IR). Despite the importance of the issue, analysis of crosstalking pathways is out of the purpose of this work but helps in acquiring awareness about the complexity of cellular signaling mechanisms.

In the paper "System-level interaction between insulin-EGF networks amplify mitogenic signaling" by Kholodenko B. N. et al. in 2009 [12] a first larger model was realized including both EGF and insulin signaling. Insulin-specific aspects were developed by selecting key intermediates and using some previously modeled interactions. The purpose of the study was to investigate the crosstalk mechanisms between insulin and EGF signaling pathways. They exploited a combination of experimental and computational approaches revealing a concordant interplay between insulin and EGF networks in potentiating mitogenic signaling. The flow chart describing the EGF-insulin model is depicted in Fig.10 where the blocks indicates the elements included in the model and the lines represent the type of interaction.

As said in Chapter 1, insulin receptor (IR) and epidermal growth factor (EGF) receptor belong to the family of receptors with intrinsic tyrosine kinase activity (referred to as receptors tyrosine kinases, RTKs), which regulates crucial cellular processes, such as cell proliferation, differentiation, metabolism, survival and apoptosis. However, the major physiological function of insulin signaling is metabolic whereas EGF mainly induces proliferative responses. Kholodenko's group investigated how EGF and insulin inputs are processed and integrated in order to result in physiological cellular response. However, since we are interested in insulin signaling pathway and we have already explored PI3K-PKB/Akt signaling cascade in Quon's model, we decide to ignore EGF signaling mechanism and to consider only MAPK pathway under insulin control.

In the following Sections we present a brief description of the Ras/MAPK model followed by some details about its implementation in the original paper and using BioNetGen language.

4.4.1 *Model description*

Upon insulin binding, insulin receptor autophosphorylates on tyrosine residues and phosphorylates IRS proteins. These are linked to the activation of Ras/MAPK pathway through binding to Grb2-SOS complex. Grb-SOS complex activation leads to a cascade of phosphorylation events involving Ras, Raf, Mek and Erk. IR can also bind the GTPase-activating protein RasGAP, which catalyzes Ras deactivation. SHP2 positively influences ERK activity by dephosphorylation of the specific sites involved in RasGAP binding.

4.4.2 *Model development in original paper*

The current model derived from previously developed EGFR network models that were based on in vitro and in vivo measurements of signaling kinetics. EGFR model was then extended to incorporate IR signaling and regulatory processes concerning EGFR-IR crosstalk mechanisms. Since one of the main limitation in implementing large network model is combinatorial complexity, Kholodenko's group decided to construct a minimal model considering only its basic features. The resulting model involves 78 variables for different molecular species, 111 chemical reactions and more than 200 parameters. For many reactions, the parameters values were assumed from previously published studies. For newly introduced processes and parameters, the values were computed and optimized using a training set of data. The signaling dynamics was analyzed comparing experimentally and simulated in silico time courses of activation of the complete model stimulated by step changes using different concentrations of EGF and insulin.

In the model, signal transduction is initiated by insulin binding to their cognate receptor. This event causes allosteric transition and autophosphorylation of the kinase activation loop of the predimerized IR, which leads to activation of the IR kinase and autophosphorylation of its cytoplasmic domain. In the model, phosphorylated IR can directly associate with IRS, PI3K and RasGAP.

Phosphorylated IRS recruits cytoplasmic proteins PI3K, Grb2 and SHP2 to the plasma membrane, which results in additional PIP3 production and both activatory and inhibitory regulation activity of Ras. Although IRS has multiple tyrosine phosphorylation sites, they were simply represented in the minimal model as a single, virtual phosphorylation site. This kind of simplification, already discussed in Chapter 2 (Section 2.3), is distant from the reality depicted by current chemistry knowledge and does not guarantee the same results respect to models including all the details about the phosphorylation sites.

Another simplification concerns the description of complex sequential multi-step processes as a single, semi-mechanistic step. This approach is commonly used to reduce the model retaining the original network topology and, in case of not distributed processes, does not critically influence the model dynamics. For instance, the activation of Raf by Ras includes a conformational change in Raf caused by binding to Ras-GTP, followed by the dissociation of 14-3-3 protein, dephosphorylation of inhibitory Ser259 and phosphorylation of activatory Ser338 site. In the model, all these processes are condensed in a single step of Raf activation.

Finally, some feedback circuits are also included in the model, such as the activated ERK inhibiting SOS.

The population of different chemical species present in the network are detected by means of protein immunoprecipitation in HEK293 cells. In particular, the monitored events are Ser217- and S221-phosphorylated MEK, Tyr1162- and Tyr1163-phosphorylated IR, Tyr612-phosphorylated IRS-1, Thr202-, Thr185-, Tyr204- and Tyr187-phosphorylated ERK1/2, and Ser473-phosphorylated Akt.

4.4.3 *Model implementation using BioNetGen language*

All the molecules populating Ras/MAPK system are defined in the `molecule types` block, and are listed below

```
RasGAP(bs)
Src(state~i~a)
GS(bs,state~i~a)
SHP2(bs)
Ras(state~d~t)
Raf(state~i~a~aa)
Mek(state~u~pp)
Erk(state~u~p~pp)
```

`RasGAP(bs)` represents RasGAP molecules having binding site `bs` for insulin receptor. In particular, RasGAP binds reversibly to insulin receptor (rule `37`) when is phosphorylated and before it undergoes internalization process (`IR(bs,Y~p,loc~ex)`). Insulin receptor is subject to degradation process despite RasGAP binding, leaving RasGAP molecules in unbound free state (rule `38`).

```
#RasGAP module
37 IR(bs,Y~p,loc~ex) + RasGAP(bs) <-> IR(bs!1,Y~p,loc~ex).RasGAP(
    bs!1) kon27,kd27
38 IR(bs!1,alpha!+,Y~p,loc~ex).RasGAP(bs!1) -> RasGAP(bs)  kf34
```

`Src(state~i~a)` represents Src molecules which may be found in inactive (`state~i`) or active (`state~a`) state. Activation (`v40()`) and inactivation (`v41()`) rates functions are implemented by means of functions depending on the concentration of active and inactive Src itself. In addiction, the activation rate function is also modulated by the amount of phosphorylated insulin receptors located on the cell membrane (observable `IR_phosph`).

```
#Src module
39 Src(state~i) <-> Src(state~a) v40(),v41()
#functions
v40()=k_Src*kcat40*(alpha40*IR_phosph)/(Km40+iSrc)
v41()=V41/(Km41+aSrc)
```

As we have briefly mentioned in PI3K-PKB/Akt Section, IRS-1 is subject to complex formation involving not only PI3K but also other two molecules, Grb2 and SHP2. Due to mutual affinity, Grb2 is assumed in a pre-formed complex with SOS and thus we directly refer

to it as GS. `GS(bs,state~i~a)` has a binding site `bs` for IRS-1 and may be found in inactive (`state~i`) or in active state (`state~a`). In particular, Grb2-SOS complex inactivation (rule `44`) is modulated by twice phosphorylated Erk (`v79()`) constituting a negative feedback loop in Ras/MAPK subsystem. The same site involved in IRS1-GS complex formation is responsible for SHP2 binding establishing a competition between these two molecules. Both GS (rule `40`) and SHP2 (rule `41`) bind to IRS-1 when phosphorylated on tyrosine residues independently from PI3K binding (`IRS1(bs,Y~p,S~u)`) and their dissociation (only when PI3K is unbound) causes IRS-1 dephosphorylation (rules `42` and `43`). Moreover, the complex involving IRS-1 and SHP2 modulates RasGAP dissociation from insulin receptor (`v115()`).

```
#IRS1-GS and IRS1-SHP2 complex formation
40 IRS1(bs,Y~p,S~u) + GS(bs,state~a) <-> IRS1(bs!1,Y~p,S~u).GS(bs
    !1,state~a) kon45,kd45
41 IRS1(bs,Y~p,S~u) + SHP2(bs) <-> IRS1(bs!1,Y~p,S~u).SHP2(bs!1)
    kon47,kd47
#IRS1-GS and IRS1-SHP2 complex distruption and IRS
    dephosphorylation
42 IRS1(bs!1,YXXM,Y~p,S~u).GS(bs!1,state~a) -> IRS1(bs,YXXM,Y~u,S
    ~u) + GS(bs,state~a) v100()
43 IRS1(bs!1,YXXM,Y~p,S~u).SHP2(bs!1) -> IRS1(bs,YXXM,Y~u,S~u) +
    SHP2(bs) kf101
#GS inhibition
44 GS(bs,state~a) <-> GS(bs,state~i) v79(),k_79
#SHP2 activity against RasGAP
45 IR(bs!1,Y~p,loc~ex).RasGAP(bs!1) -> IR(bs,Y~p,loc~ex) + RasGAP
    (bs) v115()
#functions
v100()=V100/(Km100+IRSp_GS)
v79()=kcat79*ppErk/(Km79+GS)
v115()=k115*IRSp_SHP2
```

Grb2-SOS complex binding to IRS-1 starts a multistep kinase process. The first element involved this cascade is Ras, indicated with `Ras(state~d~t)`, which may be found in GDP- (`state~d`) and GTP-bound (`state~t`) state. Ras-GTP formation (`v62()`) is positively modulated by GS-IRS1 complex formation (observable `IRSp_GS`) whereas Ras-GDP formation (`v63()`) is controlled by major contributes, the basal RasGAP action (parameter `bRasGAP`) and that one given by the concentration of RasGAP bound to insulin receptor (observable `IRp_RasGAP`).

```
#Ras module
46 Ras(state~d) <-> Ras(state~t) v62(),v63()
#functions
v62()=kcat62*(IRSp_GS)/(Km62+dRas)
v63()=kcat63*(bRasGAP+IRp_RasGAP)/(Km63+tRas)
```

Ras in GTP- bound state induce Raf activation. `Raf(state~i~a~aa)` indicates Ras molecules that may be found in inactive (`state~i`), ac-

tive (`state~a`) and double active (`state~aa`) state. The first activation of Raf (rule 47) is controlled by Ras-GTP (observable `tRas`) whereas its second activation (rule `47`) by active Src (observable `aSrc`). Raf inactivation (rule `49`) depends on two main contributes, one given by PKA (defined in `parameters` block with fixed quantity) and by Thr308-phosphorylated Akt (observables `Akt_pT308` and `Akt_phosph`). Since Akt belongs to PI3K-PKB/Akt subsystem, Raf inactivation constitutes a crosstalk mechanism between these two important pathways where PI3K-PKB/Akt subsystem negatively modulates Ras/MAPK subsystem action.

```
#Raf module
47 Raf(state~i) -> Raf(state~a) v65()
48 Raf(state~a) -> Raf(state~aa) v66()
49 Raf(state~aa) -> Raf(state~i) v67()
#functions
v65()=kcat65*tRas/(Km65+Raf)
v66()=kcat66*aSrc/(Km66+aRaf)
v67()=kcat67*PKA/(Km67+aaRaf)+alpha67*(Akt_pT308+beta67*Akt_
    phosph)
```

Raf in double active state (observable `aaRaf`) modulates the phosphorylation event on Mek (rule `50`) which is indicated as `Mek(state~u~pp)`. Finally, phosphorylated Mek promotes Erk activation. Erk molecules are defined in the system as `Erk(state~u~p~pp)` which can be found in unphosphorylated (`state~u`), phosphorylated (`state~p`) and twice phosphorylated (`state~pp`) state. Phosphorylated Mek (observable `ppMek`) positive modulates both once- (`v70()`) and twice- (`v71()`) Erk phosphorylation.

```
#Mek module
50 Mek(state~u) <-> Mek(state~pp) v68(),v69()
#functions
v68()=kcat68*aaRaf/(Km68+Mek)
v69()=V69/(Km69+ppMek)

#ERK module
51 Erk(state~u) <-> Erk(state~p) v70(),v73()
52 Erk(state~p) <-> Erk(state~pp) v71(),v72()
#functions
v70()=kcat70*ppMek/(Km70+Erk+pErk*Km70/Km71)
v71()=kcat71*ppMek/(Km71+pErk+Erk*Km71/Km70)
v72()=V72/(Km72+ppErk+pErk*Km72/Km73)
v73()=V73/(Km73+pErk+ppErk*Km73/Km72)
```

4.5 MODELS INTEGRATION

The three models we used to realize the complete system of insulin signaling pathway contain several overlapping parts. Shared elements have been modeled differently by each research group basing their

definitions on slightly different assumptions. Hence, these elements taken from distinct models were carefully analyzed and compared with each other and with the current knowledges of chemistry. The most critical elements are presented in the following paragraphs including some important details regards to the corresponding interactions with other elements.

IR. In Quon's paper, a detailed model of insulin receptor was presented including complex insulin receptor binding and recycling subsystems. According to Quon's group, insulin receptor binds two insulin molecules and subsequently undergoes internalization process constituting a cytoplasmatic pool of unbound insulin receptors where it may be either degraded or translocated back to cell membrane. Kholodenko's paper, presented a insulin receptor module having less complex binding subsystem and where the recycling subsystem is simply replaced by a reversible inactivation of the insulin receptor. However, Kholodenko's group included in the model the ability of insulin receptor to bind several molecules (IRS-1, PI3K and RasGAP) whereas Quon's group preferred to introduce insulin receptor modulation on IRS phosphorylation by means of a kinetic rate function. Thedieck's paper presented a very simple system where insulin receptor may be found in only three possible states, that are unbound (or unphosphorylated) receptor state, bound (or phosphorylated) receptor state and internalized (or inhibited) receptor state. Preferring the insulin receptor system that more closely matches the current chemical knowledges on the insulin signaling pathway, we chose to base our insulin receptor subsystem mainly on Quon's model. However, we decide to include the information given in Kholodenko's model about insulin receptor binding to multiple molecules. Since IRS phosphorylation by insulin receptor has been already included in Quon's model, we focus on PI3K and RasGAP binding. PI3K binding to insulin receptor mainly controls PI(3,4,5)P3 lipids production whereas RasGAP binding has a critical role in Ras deactivation process, thus modulating Ras/MAPK signaling pathway. Since PI(3,4,5)P3 production has been modeled according to Quon's paper we decide to ignore PI3K binding and to create only one binding site for RasGAP molecules. However, the regions of insulin receptor involved in RasGAP binding are probably different from those one implicated in PI3K binding, thereby a binding site for both RasGAP and PI3K would have introduced a false competition between molecules.

IRS-1. In Quon's paper, IRS-1 is introduced in the system in completely unphosphorylated state and may be phosphorylated either on tyrosine or on serine residues. Kholodenko's model presents a complex IRS-1 module in which they distinguished IRS-1 located in the cytoplasm from that bound to PI(3,4,5)P3 lipids (or PIP3) translocated in

proximity of cell membrane. According to Kholodenko's group, IRS-1 is available for binding other molecules (PI3K, Grb2 and SHP2) only in proximity of the cell membrane. Despite the great improvement in modeling IRS-1 cellular location we prefer to retain IRS-1 definition made by Quon's group, not only because the introduction of this information in the model would be problematic but also because in Kholodenko's model IRS-1 phosphorylation on serine is replaced by a inhibition state which would be difficult to integrate with mTOR subsystem. Kholodenko's group also modeled IRS-1 complex formation with three molecules, PI3K, Grb2-SOS and SHP2. PI3K binding modulates PI(3,4,5)P3 lipids production which have two major roles, to translocate IRS-1 to cell membrane and to activate PDK1. Grb2-SOS binding critically controls Ras activation process, thus starting Ras/MAPK signaling pathway. SHP2 binding promoted RasGAP unbinding mechanism from insulin receptor. However, Kholodenko's group consider the formation of only three distinct complexes, hence without considering any other possible combination in which IRS-1 binds more then one molecule at the same time. This assumption introduces a competition between molecules in binding to IRS-1. To unravel this issue we checked if the same binding site in IRS-1 was responsible for the binding of all three molecules. We found that nine motives YXXM are present for PI3K binding but they do not comprise the region between aa 896-898 which is responsible for Grb2 binding. Hence, excluding steric volume issues, we assume that PI3K and Grb2 can bind to IRS-1 independently and thus that Kholodenko's model introduced false competition between these two molecules. Since we couldn't find any information about SHP2 we assume as IRS-1 binding site the same responsible for Grb2 binding.

Thedieck's paper introduces a simple system in which IRS-1 changes its phosphorylation state according to three consecutive states, that are unphosphorylated IRS-1, tyrosine phosphorylated IRS-1 and serine phosphorylated IRS-1. According to Thedieck's group, S6K constitutes the negative feedback loop (NFL) modulating the transition of IRS-1 from tyrosine phosphorylated to serine phosphorylated state. We thought more reasonable to model S6K action modulating IRS-1 phosphorylation on serine starting from completely unphosphorylated IRS-1.

AKT. Quon's paper defines Akt phosphorylation on threonine residue alone directly modulated by PI(3,4,5)P3 lipids production. Kholodenko's paper distinguishes between once- and twice- phosphorylated Akt. The first phosphorylation is controlled by PDK1 and the second one by mTOR. Thedieck's paper, instead, presents Akt having two phosphorylation sites, one on threonine and one on serine. Threonine phosphorylation is controlled by PI3K activation and tyrosine phosphorylated IRS-1 whereas serine phosphorylation is modulated by

mTORC2. We decide to mix all these informations defining Akt molecules having both phosphorylation sites where threonine phosphorylation is induced by PI(3,4,5)P3 lipids production and serine phosphorylation in controlled by mTORC2. The distinction between single and double phosphorylation of Akt molecules is interesting because of change in its kinase action. Hence, we take into account this information in implementing the crosstalk mechanism between the PI3K-PKB/Akt and the Ras/MAPK subsystems that, as already said, requires Akt contribute in Raf deactivation process.

Due to lack of information about molecules concentrations, Quon's group expresses Akt in percentage instead of mole. We decide to use the initial concentration suggested in Kholodenko's paper. Since GLUT4 translocation on the cell membrane is controlled both by PKC-ζ and Akt activation we computed the percentage of phosphorylated Akt to be used as contribute in metabolic "Effect".

MTOR. Thedieck's paper define mTORC1 and mTORC2 complexes inside the system. mTORC1 activation is controlled by Akt through TSC1-TSC2 complex and modulates the upstream element IRS-1 through S6K. mTORC2 activation is controlled by PI3K activation and induces serine phosphorylation on Akt. Kholodenko's model does not distinguish between the two complexes but define a mTOR molecule which is activated by Akt and in turn twice- phosphorylates Akt. According to the current knowledge of chemistry, we decide to choose the dinstiction between mTORC1 and mTORC2 complexes introduced in Thedieck's paper.

Morever, the negative feedback loop (NFL) is modelled in Kholodenko's paper by a simple inhibition on IRS-1 induced by mTOR, whereas Quon's group presents a model in which Akt controls IRS-1 tyrosine dephosphorylation and PKC-ζ modulates IRS-1 serine phosphorylation. According to the current knowledges, we decide to reasonably replace the entire negative feedback loop using the one suggested in Thedieck's paper which includes TSC1-TSC2, mTORC1, S6K and also AMPK.

CORRECTIVE FACTORS. First of all, homogeneity in time units describing kinetics parameters must be guarantee. Quon's and Thedieck's models express the parameters in $[min^{-1}]$ whereas Kholodenko's model uses $[sec^{-1}]$. Hence, all the parameter expressed in $[sec^{-1}]$ should be multiplied for 60 to transform them in $[min^{-1}]$.

It is worthy to discuss an essential aspect linked to BioNetGen implementation concerning the fact that concentrations should be expressed in units of *copies per cell* and thus bimolecular rate constants in *per molecule per cell*. This simply means that every time a quantity is expressed in `M` or in `/M` (unless time units) this has to be multiplied or divided by the product `(NA*V)`, where `NA` is the Avogadro's num-

ber and V indicates cellular volume. Since many biological systems involve ligands present in extracellular space, receptors spread out on the cell surface and molecules situated in the cytoplasm, the value of V changes depending on molecule location. For this reason it is important to identify where molecules are located, such as IR that binds to insulin on the plasma membrane and subsequently undergoes internalization mechanism moving in intracellular space. Therefore, the computed values of volume, expressed in (L), for the three most important cellular locations are:

- $V_o = 1.0 \cdot 10^{-10}$ for extracellular volume,
- $V_m = 3.0 \cdot 10^{-13}$ for plasmatic membrane volume, and
- $V = 3.0 \cdot 10^{-12}$ for cytoplasmic volume.

However, Thedieck's model expresses its concentration in Arbitrary Units (AU) instead of M, suggesting that the resulting quantities are proportional to light intensity level obtained during immunoblot experiments. Also Kholodenko's group obtain its data from immunoblot experiments but the knowledge of initial concentrations permits to obtain simulation expressed in M. Unfortunately no information about initial concentrations are available for those elements constituting mTOR subsystem. Hence, waiting for these essential information we decide to integrate the data expressed in *copies per cell* with those expressed in AU. In order to do this, some factors were introduced for those parameters connecting the elements expressed in different units. These factors were used to correct the IRS-1 tyrosine- phosphorylation by S6K, AMPK phosphorylation by IRS-1, Akt serine phosphorylation by mTORC2, mTORC2 phosphorylation by IRS1-PI3K complex, TSC1-TSC2 complex inactivation by Akt. Some factors were also introduced to compensate gaps due to slightly different assumptions made in the complete model respects to the original ones. These so called "corrective" factors were manually tuned comparing the results obtained from simulations of the complete model with those of the original models, and were used to multiply the corresponding kinetic parameters. In particular, these factors were used to compensate effects due to false competition between molecules. For instance, if a molecule forms a bond with multiple molecules it is important to decide if the sites involved in the binding process are distinct or not. In some cases, it would be useful to check the adjacency of these sites to avoid problems connected to the steric volume of the binding molecules. Hence, binding kinetics for each molecule may differ a lot depending on these information. For sake of simplicity, many research groups introduce a single binding site for all the molecules and do not consider all the possible combinations that a complex may form. In our model, we prefer to more closely match the current knowledges of chemistry allowing the formation of complex that

are different from those presented in the original models. Hence, to compensate the gap we introduced some corrective factors for the molecules RasGAP, GS and SHP2.

Finally, the initial quantities of all the molecular species were also compared and, excluding those expressed in AU, we realized that Quon's concentrations were unreal. In fact, multiplying the proposed quantities for `(NA*V)` we obtain less the one molecule per cell. This fact was also confirmed by Nyman E. et al. [31] in explaining the major limits connected to this model. Thus, we decided to scale all the initial concentrations taken from Quon's paper multiplying them by 10^5 in order to obtain quantities comparable with those showed in Kholodenko's paper.

5 INSULIN MODEL ANALYSIS AND VALIDATION

5.1 INTRODUCTION

In this Chapter, we use the insulin model described in Chapter 4 to simulate the temporal series of most significant chemical species in order to analyse the dynamic behavior of the system. Concentrations time series permit to carry out a detailed analysis of dynamic behavior of the system. This may be useful to develop a deeper understanding of network regulation processes and of the main role played by each chemical species. Moreover, we can consider the dynamic behavior of each chemical species as a kind of fingerprint of the system working under physiological conditions.

Network structure and type of interactions between different species determine characteristic response of the system upon stimulation. If some of these characteristic features are lost for different reasons, typical dynamics may change and then result in inappropriate actions of the system. System disregulations cause defects in signal transduction that can be pathologic. In case of insulin signaling, for instance, *insulin resistance* and *diabetes type 2* originate by disregulation of that process involving IRS-1 phosphorylation of serine residues, thereby of the negative feedback loop involving S6K (see Chapter 2). Hence, to characterize physiological working conditions of a signaling pathway may be a useful method to distinguish it from those in pathologic conditions, to investigate the origin of disregulations and, in future perspective, to find a drug or combination of drugs able to bring the system back to normal conditions.

In order to do this, a detailed and rational analysis of collected data from insulin model simulations have to be done. Since the main purpose is to characterize system response, we decide to classify the obtained time series according to their dynamic behavior. In other words, we distinguish them depending on their qualitative behavior, or pattern.

In next sections, we present the details about classification of main patterns (Section 5.2), results and analysis of insulin model simulations (Section 5.3), and comparison with experimental data (Section 5.4).

5.2 PATTERN CLASSIFICATION

Classification of resulting time series according to their dynamic behavior may be a cunning approach to realize a useful characterization

of the system and it may reveal important information about network working mechanisms.

System dynamics is strictly related to network structure. It has been observed that characteristic dynamic behaviors originate often by particular network topology features. Networks commonly contain subsystems characterized by particular configurations of causal relationships. Some of these subsystems appear more frequently and such subnetworks are called *network motifs* [13, 39]. Typical motifs are, for instance, negative feedback loops and feedforward loops, which will be briefly discussed in the following Subsection.

5.2.1 *Network motifs*

Network motifs were originally recognized and formalized during the analysis of complex gene transcription networks. However, same criteria introduced for the study of this sub- systems may be used in transduction networks.

The most common network motif observed in transcription networks is the *feedforward loop*, that has been thoroughly studied by Uri Alon [39]. Extended scientific literature suggests that feedforward loops may serve in transcription networks as filters, sign-sensitive delays, and pulse generators. However, differrently from transcription networks where it was rarely observed, transduction networks frequently reveal the presence of *negative feedback loops,* which remarkably determine dynamic behavior of the system. This is probably due to the fact that signal transduction networks represent faster information processing networks (acting on timescale of minutes or seconds) respect to transcription networks (timescale of hours), thus particular responsivity requirements of system must be guaranteed.

If we consider a chain of reactions in which the first reaction is inhibited by one of the downstream elements, we are observing a negative feedback loop. In general, without such feedback, the concentration of each element reaches steady state after a certain transition time. If the second element of the chain acts as inhibitor, the level of the first element shows an overshooting response. With a long-ranging feedback involving a longer time delay, this effect becomes more pronounced and may leads to damped oscillation of the last element of the chain. The result is that response time ($\tau_{\frac{1}{2}}$) decreases from the situation without feedback to that with feedback. In particular, chains containing long-ranging feedback are faster compared to those having short-ranging feedback. Hence, presence of negative feedback loops may be a trick, accurately selected through evolution, to *speed up system response* to external changes.

***Response time**: time at which the last element of the chain reach its half-maximal level.*

Negative feedback loops may entail various effects on cellular dynamics, as: stabilization of the state of cellular network, reduction of the variance of fluctuations and the variability of steady states, ro-

bustness to boundary conditions, production of pulse-like overshoots, and induction of sustained oscillations.

5.2.2 *Pattern definitions*

Considering the remarkable importance of negative feedback loops in transduction networks and that this particular network motif usually entails the presence of overshooting response, it could be helpful to distinguish the simulated concentrations according this particular feature. Hence, we can divide them in two sets, one for those concentrations characterized by overshoot, and one for those without overshoot.

To classify dynamic behaviours according to their speed may also give useful information about how signaling system works and help in its characterization.

Thus, we decide to classify predicted time series obtained from insulin model according to these two criteria, obtaining four main pattern sets, shown in Fig.11:

- slow (not-overshooting) response,
- rapid (not-overshooting) response,
- slow overshooting response, and
- rapid overshooting response.

Each pattern may be characterized by means of specific parameters. We already mentioned the response time $\tau_{\frac{1}{2}}$ (or $\tau_{50\%}$) [min] that indicates the period of time the concentration takes to reach its half-maximum level. Similarly, we can define $\tau_{90\%}$ [min] which indicates the time to reach the 90% of maximum level, and which may help to further describe concentration dynamics. These two parameters may be particularly useful for characterization of increasing (not-overshooting) responses and for classification of their dynamics according to the speed.

For concentrations characterized by overshooting response, we can easily obtain information about τ_{peak} [min], which indicates the period of time the concentration takes to reach its peak value. This parameter permits to distinguish concentration dynamics according to their speed. Another parameter which can be easily computed from simulations is $\frac{x_{st}}{x_{peak}}$ [%], which gives indication about steady-state levels respect to the peak value, and which may serve for further considerations on concentration dynamics.

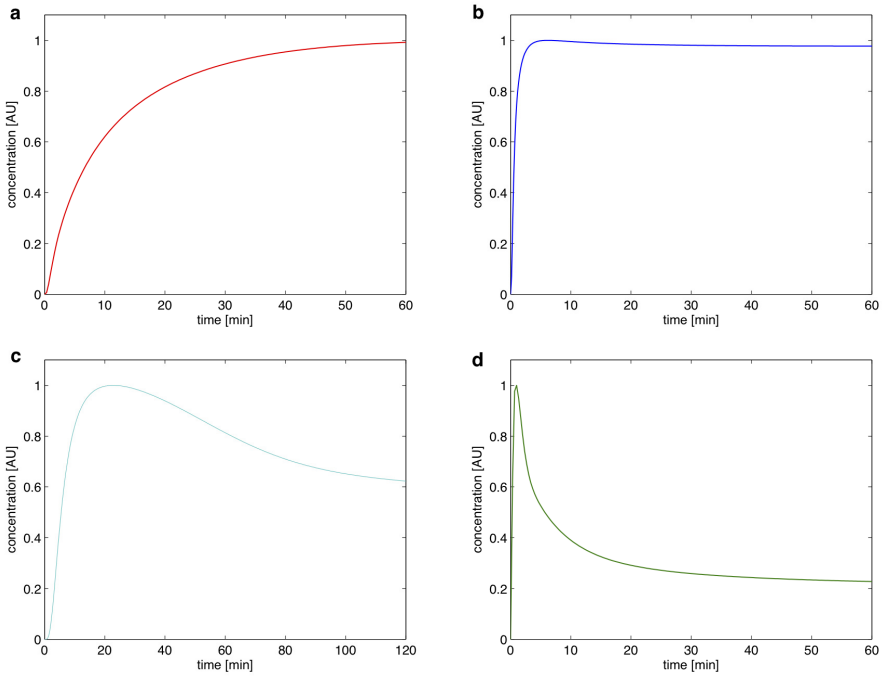

Figure 11: Representative dynamic behaviour for each pattern: slow (not-overshooting) response **(a)**, rapid (not-overshooting) response **(b)**, slow overshooting response **(c)** and rapid overshooting repsonse **(d)**.

5.3 INSULIN MODEL SIMULATIONS: RESULTS AND ANALYSIS

5.3.1 *Model simulations*

Simulations of insulin model were realized using a step function of 100 nM insulin as input. Insulin stimulation started at 0 time and was held for 60 minutes. Using BioNetGen language

```
#actions
generate_network({overwrite=>1});
# Kinetics
simulate_ode({t_end=>60,n_steps=>200});
```

where, as presented in Chapter 3, `generate_network` command permits to generate the complete reactions network according to an iterative method which terminates when no more new reactions are found. `simulate_ode` command indicates that simulation of dynamics is carried out implementing and numerically solving a set of ODEs. Simulation lasts for 60 minutes (`t_end=>60`) and shows in output 200 time samples for each observable species (`n_steps=>200`). Time series of all the observables species are reported in Appendix.

5.3.2 *Simulation results and classification*

According to discussion in Section 5.2.2, we decide to first distinguish the dynamic behavior of observable species according to the presence or not of overshooting response. Thus, chemical species characterized by overshoot are: Tyr- phophorylated IRS1, IRS1-PI3K complex formation, PI(3,4,5)P3 complex production, Thr308- phosphorylated Akt, GLUT4 translocation on cell membrane, IRS1-GS and IRS1-SHP2 complex formations, GTP- bound Ras, single- and double- activated Raf, Ser221- and Ser217- phosphorylated Mek and Thr202/185- phosphorylated Erk, Thr202/185- and Tyr204/187- phosphorylated Erk. The remaining time series depict increasing concentrations which reach the steady state level without any overshoot, and are: Tyr- phosphorylated insulin receptor, Ser- phosphorylated IRS1, Ser473- phosphorylated Akt, Ser2481- phosphorylated mTORC2, Thr172- phosphorylated AMPK, Ser1387- phosphorylated TSC1-TSC2 complex, Ser2448- phosphorylated mTORC1, Thr389- phosphorylated S6K, IR-RasGAP complex formation, active Src, and inhibited GS.

We decide to further distinguish concentration dynamics according to their speed using the parameters introduced in Section 5.2.2. In particular, among the species characterized by increasing (not-overshooting) response, we classify:

- as rapid those having $\tau_{50\%} < 1$ min, and
- as slow, all the others.

For the concentrations characacterized by overshooting response, we defined

- as rapid those with $\tau_{peak} < 5$ min, and
- as slow, all the others.

All the detected concentrations classified according to their pattern are reported in Tab. 1.

Resulting pattern classification is also presented in Fig.12 and 13. These schematics depict respectively the two main sub-systems of insulin signaling pathway, the PI3K-PKB/Akt and Ras/MAPK subpathways. Having only few elements in common, we prefer to present them separately. Fig.12 and 13 show the most important chemical species considered in insulin model which are represented as distinct compartments. Compartments show also some colored labels indicating the type of residue involved in phosphorylation event, which are red for tyrosine, blue for serine and green for threonine. Solid lines connecting different compartments indicate the transformation between a chemical species to another. Dotted lines joining two different species indicate interactions with complex formation. Dotted lines

not-overshooting response		overshooting response	
rapid	**slow**	**rapid**	**slow**
IR(pY)	IRS1(pS)	IRS1(pY)	a_Raf
IR-RasGAP	Akt(pS473)	IRS1-PI3K	aa_Raf
a_Src	mTORC2(pS2481)	PI(3,4,5)P3	Mek(pS221, pS217)
	AMPK(pT172)	PKC(pT)	Erk(pT202/185)
	TSC1-TSC2(pS1387)	Akt(pT308)	Erk(pT202/185, pY204/187)
	mTORC1(pS2448)	Akt(pT308, pS473)	
	S6K(pT389)	GLUT4	
	i_GS	IRS1-SHP2	
		IRS1-GS	
		Ras-GTP	

Table 1: Detected concentrations in insulin model classified according to their pattern.

are also used to indicate chemical species controlling different reactions. Detected quantities are indicated in red and, in their proximity, are present plots depicting their qualitative dynamic behaviours. Plots present colored frames which indicate under which pattern the detected concentration is classified, and are blue for rapid (not-overshooting) response, red for slow (not-overshooting) response, green for rapid overshooting response, and cyan blue for slow overshooting response.

5.3.3 *Classification analysis and discussion*

By means of pattern classification we were able to characterize dynamic behaviour of some of the most important elements constituting insulin signaling system. As already said, this characterization may be useful to distinguish insulin signaling pathway working under normal condictions from those affected by disregulations, such as in *insulin resistance* and *diabetes type 2*.

As already said, pattern classification may also unveil complex mechanisms involved in insulin signaling transduction. From a first analysis, we observe that rapid (not-overshooting) response characterizes insulin receptor phosphorylation and those processes strictly related to this, such as IR-RasGAP complex formation and Src activation. Tyr- phosphorylated IRS1 is also controlled by phosphorylated insulin receptor inducing a rapid dynamic. However, Tyr- phosphorylated IRS1 concentration presents a overshooting response which remarkbly determines the complex formation dynamics with its binding proteins, such as PI3K, Grb2-SOS (GS) and SHP2. By IRS1-PI3K

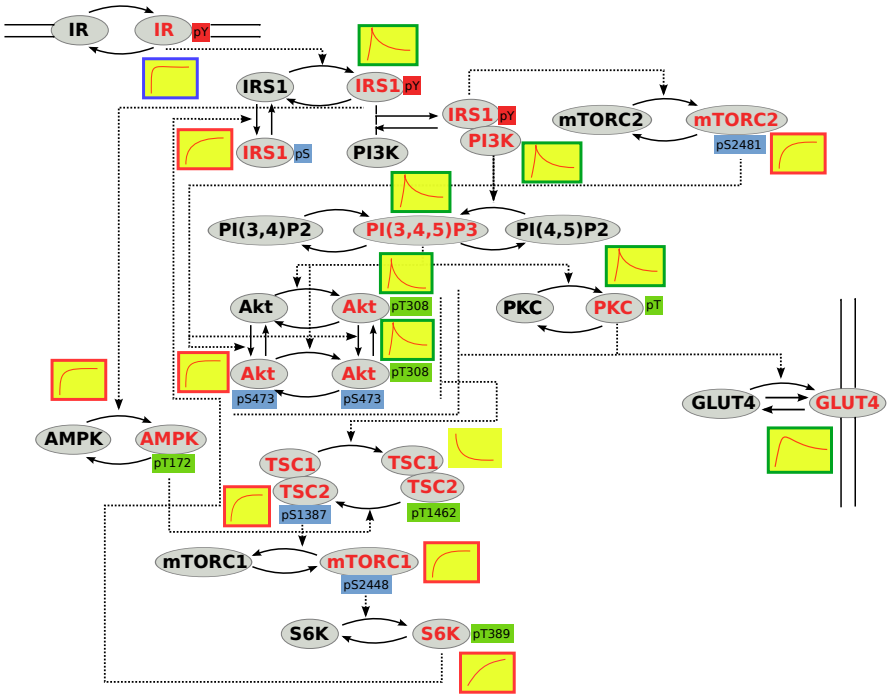

Figure 12: Schematics of PI3K-PKB/Akt sub- system.

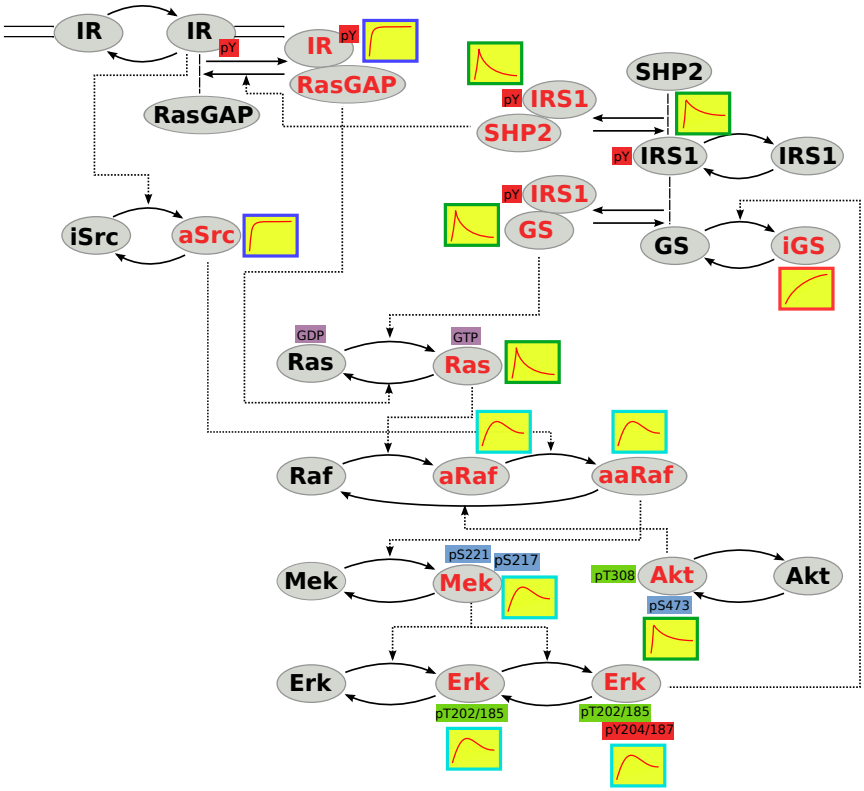

Figure 13: Schematics of Ras/MAPK sub- system.

complex, rapid overshooting response extends to PI(3,4,5)P3 lipids production, Thr308- phosphorylated Akt, Thr- phosphorylated PKC-ζ, up to GLUT4 translocation on cell membrane. However, pattern changes from rapid overshooting response to slow (not-overshooting) response in Ser2481- phosphorylated mTORC2 and Thr172- phosphorylated AMPK. Phosphorylated mTORC2 affects dynamics of Ser473-phosphorylated Akt which presents slow (not-overshooting) response. Phosphorylated AMPK influences together with Thr308- phosphorylated Akt, TSC1-TSC2 complex dynamics which is still characterized by a slow (not-overshooting) response. By TSC1-TSC2 complex, slow (not-overshooting) dynamics extends to Ser2448- phosphorylated mTORC1, Thr389- phosphorylated S6K, up to Ser- phosphorylated IRS1.

By IRS1-GS complex, instead, rapid overshooting response affects dynamics of GTP- bound Ras. Raf activation mechanism is controlled by GTP- bound Ras, active Src and phosphorylated Akt, which implements a crosstalk between the two sub- systems. Complex Raf regulation determines the pattern changing from rapid overshooting response to slow overshooting response, which affects the dynamics of subsequent kinases cascade, constituted by double- phosphorylated Mek, and single- and double- phosphorylated Erk. Double- phosphorylated Erk implements another negative feedback loop acting on GS, whose inhibited form assumes a slow (not-overshooting) response.

Hence, pattern classification reveals that, generally:

- rapid (not-overshooting) response characterizes those mechanisms strictly related to insulin receptor,
- slow (not-overshooting) response characterizes those mechanisms related to negative feedback loops,
- rapid overshooting response characterizes those mechanisms related to Tyr- phosphorylatead IRS1,
- slow overshooting reponse characterizes those mechanisms related to MAPK kinases cascade.

We briefly discussed in Section 5.2.1 the meaning of network motifs and their importance in determining dynamic behavior of the system. These results are in agrrement with the previous observation that negative feedback loops are often related to the presence of overshooting response. This analysis performed using pattern classification confirms the important link between network structure and its dynamic behavior. Indeed, overshooting response affecting all the elements described above results from the negative feedback loop involving IRS1 phosphorylation on serine residues operated by S6K. Precise temporal balance of IRS1 phosphorylation mechanism on tyrosine and serine residues critically determines the overall dynamic behaviour of the system. Insulin receptor induces a strong and rapid activation of

Rapid (not-overshooting) response			
	x_{st} [nM]	$\tau_{50\%}$ [min]	$\tau_{90\%}$ [min]
IR(pY)	9.37	<0,5	<1
IR-RasGAP	4.01e-05	<0,5	<1
a_Src	0.0097	<0,5	<1

Table 2: Classified concentrations.

all downstream elements immediately after insulin stimulation. Activation rapidly propagates until it reaches TSC1-TSC2 complex, whose regulation is doubly controlled by Thr308- phosphorylated Akt and Thr172- phosphorylated AMPK. Probably due to its kinetics properties, AMPK assumes a slow (not-overshooting) response even if controlled by Tyr- phosphorylated IRS1. However, mechanism regulating TSC1-TSC2 complex seems essential in modeling the subsequent dynamics involving mTORC1 and especially S6K. S6K mediates the negative feedback loop enhancing Ser- phosphorylation on IRS1, thus causing a decrease in the concentration of Tyr- phosphorylated IRS1 which determines the characteristic overshooting response.

This complex and precise mechanism is probably the result of evolutive phenomena which, through selection, made possible the perfect optimization of insulin signaling network. We previously mentioned all the effects that negative feedback loop entails on cellular dynamics. To speed up response time may be particularly important in determining the responsiveness of cellular transduction networks. In this case, insulin secretion by β-cells should guarantee all those mechanisms which allow to reduce blood glucose level in the shortest time possible. GLUT4 translocation is one among these mechanisms. Upon insulin stimulation, GLUT4 transporters rapidly move to cellular membrane allowing glucose to enter into the cell where it undergoes glycogen conversion. Hence, this negative feedback loop permits to accelerate this mechanism making the insulin system more effecient.

Negative feedback loop is present also in Ras/MAPK sub- system and involves GS inhibition by double- phosphorylated Erk action, which reduces GS ability in binding IRS1. However, rapid overshooting response describing IRS1-GS complex formation is mostly due to Tyr- phosphorylated IRS1, whose dynamics is affected by the negative feedback we discussed before. Hence, the negative feedback action involving GS is weaker if compared to that one involving S6K and does not determine remakable changes in system dynamics. However, further considerations may be derived from parameters characterizing each pattern, and which are presented and discussed in the next Subsection.

Slow (not-overshooting) response			
	x_{st} [nM]	$\tau_{50\%}$ [min]	$\tau_{90\%}$ [min]
IRS1(pS)	65,43	8	31
Akt(pS473)	93,27	2	4
mTORC2(pS2481)	8,05 (AU)	7	29
AMPK(pT172)	19,37 (AU)	>1	3.5
TSC1-TSC2(pS1387)	10,29 (AU)	2.5	15
mTORC1(pS2448)	7,21 (AU)	4	12
S6K(pT389)	7,20 (AU)	39	80
i_GS	31,80	38	87

Table 3: Classified concentrations.

Rapid overshooting response				
	x_{peak} [nM]	x_{st} [nM]	τ_{peak} [min]	$\frac{x_{st}}{x_{peak}}$ [%]
IRS1(pY)	109,09	25.50	~1	23
IRS1-PI3K	0,72	0,18	~1	23
PI(3,4,5)P3	7,56 (%)	2,30 (%)	~1	30
PKC	20,67 (%)	6,62 (%)	~1	32
Akt(pT308)	16,05	0,01	~1	~0
Akt(pT308, pS473)	14,29	6,61	~3	46
GLUT4	46,30 (%)	29,67 (%)	~4	64
IRS1-SHP2	1,31	0,31	~1	23
IRS1-GS	0,02	0,0035	~1	20
Ras-GTP	2,96	0,70	~2.5	24

Table 4: Classified concentrations.

Slow overshooting response				
	x_{peak} [nM]	x_{st} [nM]	τ_{peak} [min]	$\frac{x_{st}}{x_{peak}}$ [%]
a_Raf	16,27	6.22	17	38
aa_Raf	2,85	1,92	22	67
Mek(pS221, pS217)	12,38	8.37	22	68
Erk(pT202/185)	38,27	26,18	23	68
Erk(pT202/185, pY204/187)	5,77	2,59	23	45

Table 5: Classified concentrations.

5.3.4 *Further analysis and discussion on quantitative results*

Tab.2-5 present all the observable species classified according to the mentioned criteria and report the most significant parameters describing each pattern.

As previously detailed, double- phosphorylated Erk plays a role in controlling an upstream mechanism in insulin signaling system, IRS1-GS complex formation. In particular, Erk inhibits the ability of GS to bind IRS1 thus realizing a negative feedback loop. Although the major contribute in shaping the overshooting response of IRS1-GS is given by Tyr- phosphorylated IRS1, the analysis of the parameters collected is Tab.4 reveals that GS inhibition induces a decrease of the steady state level of IRS1-GS complex ($\frac{x_{st}}{x_{peak}}$=20%) respect to Tyr- phosphorylated IRS1 ($\frac{x_{st}}{x_{peak}}$ =23%). Consequently, GTP- bound Ras has a lower steady state and this may be critical in regulating the following kinases cascade, determining overshoot durations and steady state levels of Raf, Mek and Erk (see Tab.5). In particular, fine regulation of Erk may be particularly important due to its role in acting directly on genetic transcription network.

Further discussion may be done on Akt phosphorylation mechanism observing the values in Tab.3 and 4. Akt phosphorylation on Thr308 is controlled by PI(3,4,5)P3 lipids production. Akt is also phosphorylated on Ser473 by phosphorylated mTORC2. mTORC2 phosphorylation is controlled by active PI3K, thus by IRS1-PI3K complex formation. Due to its kinetics properties, phosphorylated mTORC2 results in increasing (not-overshooting) response affecting also Ser473 Akt phosphorylation. Simulations indicate that the majority of Akt is phosphorylated on Ser473 whereas Thr308 phosphorylation increases rapidly during the first period and then reduces its contribute. In particular, concentration of Akt phosphorylated uniquely on Thr308 (and not on Ser473) extinguishes after a brief period of time (τ_{max}~1 min, $\frac{x_{st}}{x_{peak}}$~0 %) due to gradual phosphorylation on Ser473 ($\tau_{50\%}$=2 min, $\tau_{90\%}$=31 min). This mechanism results in a double- phosphorylated Akt state characterized by a delayed peak (τ_{max} ~3 min) and by a higher steady state level ($\frac{x_{st}}{x_{peak}}$=46%) if compared to the upstream species analyzed before ($\frac{x_{st}}{x_{peak}}$~23-30%). Phosphorylated Akt (in any form) together with phosphorylated PKC-ζ contributes to GLUT4 translocation which, after an overshoot, settles to a significantly high steady state level (τ_{max}~4, $\frac{x_{st}}{x_{peak}}$=64 %).

It is also worthy to observe that S6K phosphorylation increases very slowly respect to upstream elements ($\tau_{90\%}$~12-15 min) reaching a steady state level after a long period of time ($\tau_{90\%}$=80 min). The same dynamics characterized GS in inhibited state ($\tau_{90\%}$=87 min) revealing that both elements involved in negative feedback loop presents a dynamics which is almost linear during the first 60 minutes. This feature

may reveal a high sensitivity characterizing these elements, thus may be responsible of critical disregulations in insulin signaling pathway.

5.4 INSULIN MODEL VALIDATION

Insulin model presented in Chapter 4 is the result of integration of three different models. Although the difficulties to deal with the overlapping parts of these networks, we were able to obtain a final model which is a counterfeit between state of art chemistry knowledge and simplicity in implementation procedure. Then, we run simulations and we compared them with those obtained from original models verifying a good agreement. Figures showing the comparison between models are presented in Appendix.

However, we also decided to make a direct comparison of the predictions obtained from insulin model with the available experimental data. Hence, we decided to use pattern classification showed in Section 1.3 to simplify the validation procedure of collected times series. We used the parameters characterizing each pattern to realize a comparison between the simulated time series and the data available from literature. During this procedure, we used experimental data taken from the same models we referred during model implementation [34, 12], but also data collected by other research groups [36, 43]. Since experimental data are always expressed in arbitrary units (AU), parametric characterization of each pattern represents a cunning approach in data comparison that otherwise would be difficult.

In next Subsections we present the details of experimental data we refer to during validation procedure (Subsection 5.4.1) and the comparison with predictions of insulin model (Subsection 5.4.2).

5.4.1 *Experimental data*

For validation procedure of insulin model we refer to experimental data collected by four research groups:

- Thedieck's group [34]- collected experimental data on HeLa cells starved for 16 hours and restimulated with amino acids and 100 nM insulin for 3, 20, 45 and 100 minutes. Measured time series we refer to are IR(pY1146), IRS1(pS636), Akt(pT308,pS473), mTOR(pS2448), p70S6K(pT389).

- Kholodenko's group [12]- collected experimental data on HEK293 cells starved from 12 to 16 hours and then stimulated with 100 nM insulin. Number of time samples collected during experiments depends on detected concentration but is generally about 5 or 10 time samples in a time span of 16 or 30 minutes. Measured time series we refer to are MEK (pS217,pS221),

IR(pY1162,pY1163), IRS1 (pY612), ERK1/2 (pT202/185,pY204/187) and Akt(pS473).

- Toffolo's group [36]- collected experimental data on human skeletal muscle cells (SkMCs), after starvation, exposed to 100 nM insulin for 0, 2, 5, 10, 30, and 60 minutes. Measured time series we refer to are Akt (pS473), ERK1/2 (pT202, pY204), mTOR (pS2448) and p70S6K (pT389).
- White's group [43]- collected experimental data on 3T3-L1 adipocytes stimulated with 150 nM insulin for 0, 5, 15 and 45 minutes. Measured time series we refer to are IR(pY1175,pY1179 and pY1180 in kinase domain, pY1345 and pY1351 in COOH-terminal), IRS1(pY460, pY935 and pY983 in YXXM motif), ERK1(pT203, pY205) and ERK2(pT183,pY185).

We also refer to simulations presented in Quon's model, which mostly compared their results with experiments conducted on 3T3-L1 adypocytes, and which may help in validating those processes involved in PI3K-PKB/Akt sub- system.

5.4.2 *Validation results*

In the next paragraphs, we compare the data reported in Tab.2-5 with experimental data, thus providing a further model validation.

INSULIN RECEPTOR. Insulin receptor dynamics matches exactly with Quon's model simulations and shows a rapid (not-overshooting) response with $\tau_{90\%}$ <1 min. Strong and rapid activation of insulin receptor is also confirmed by Thedieck's , Kholodenko's and White's groups. In particular, White's data shows an activation increase of more than 10- fold after 5 min from insulin stimulation which, apart from a slight decrease to 45 min, confirmed the dynamics of our insulin model.

IRS1. IRS1 dynamics of insulin model do not present oscillatory behavior as simulations reported by Quon's group. This is probably due to the replacement of the original modulation on IRS1 phosphorylation on serine carried out by phosphorylated PKC-ζ, with a more detailed negative feedback loop involving Akt, TSC1-TSC2 complex, mTORC1 and S6K. However, both steady states levels of Tyr- and Ser-phosphorylated IRS1 are very close to those reported by Quon's paper (x_{st}= 27.86 nM for Tyr- phosphorylation and x_{st}= 64.79 nM for Ser- phosphorylation) and also the peak time of Tyr- phosphorylated IRS1 (τ_{max} = 2 min).

Kholodenko's group also reports concentration of Tyr- phosphorylated IRS1 increasing during the first minute, then settling to 15%

of its maximum value. White's group shows maximal phosphorylation after 5 min (more than 3- fold increase) followed by a marked decrease for longer insulin stimulation.

The slower dynamics of Ser- phosphorylated IRS1 was confirmed by Thedieck's group which reports that concentration reaches half-maximum level between 5 and 10 min after insulin stimulation.

AKT. T308- phosphorylated Akt dynamics does not present oscillatory behavior as reported in Quon's simulations but peak time and steady state level compared to peak value are really close to those reported in the paper (τ_{max} = 2 min, $\frac{x_{st}}{x_{peak}}$ = 42%). Also Thedieck's group confirmed a rapid dynamics and marked decrease (τ_{max} = 2 - 3 min, $\frac{x_{st}}{x_{peak}}$ = 56%).

S473- phosphorylated Akt dynamics is in good agreement with data reported by Thedieck's (τ_{max} = 3 - 4 min) and Kholodenko's groups (τ_{max} = 2- 3 min). Toffolo's data shows that S473- phosphorylated Akt reaches maximum value in 10 min and settles to steady state level without decrease as we reported in our insulin model predictions.

MTOR. Insulin model simulations show that phosphorylated mTORC1 increases gradually reaching its half-maximum value in 4 min and this is confirmed by Thedieck's group in its experiments ($\tau_{50\%}$= 4 - 5 min). Also Toffolo's data reports an increasing trend which reaches its steady state level in 10 min after stimulation ,without any decrease.

S6K. S6K phosphorylation dynamics matches with that one reported by Thedieck's group in its experiment. Toffolo's data also shows an increasing trend which is extremely slower respect those one reported before and which reaches steady state level after 30 min.

RAS. Insulin model predicts a GTP- bound Ras concentration having a peak after 2.5 min and which reaches a steady state of 24% respect to its peak value. This is confirmed by experimental data reported by Kholodenko's group which shows a peak around 2 min and a decreasing behavior.

MEK. Phosphorylated Mek concentration presents a slower dynamics respect to Ras and this is confirmed by Kholodenko's group experiments, where Mek reaches its peak after 5 min and it slightly decreases until it reaches its steady state level ($\frac{x_{st}}{x_{peak}}$ =50%).

ERK. Insulin model predictions show a single- phosphorylated Erk concentration similar to Mek and a double- phosphorylated state with a lower steady state level. Kholodenko's data reported double-phosphorylated Erk dynamics with a peak at 5 min and marked de-

crease. Toffolo's data shows slower dynamics with a peak after 10 min and a marked decreases. White's group classifies Erk among those proteins which increase their phosphorylation activity more than 10-folds (after 5 min from stimulation). Reported concentrations have a peak around 5 min and then decrease, having a higher steady state for single- phosphorylated Erk and a lower one for double- phosphorylated Erk, thus confirming our predictions.

6

CONCLUSIONS

In this thesis we implemented a computational model of insulin signaling pathway. We made use of three models already published in literature that were integrated to realize a final and more complete insulin model. Model was implemented using rule-based modeling approach which permitted a simple system description by a set of reaction rules. In particular, we used BioNetGen software and, by means of its tools, we generated predictions of chemical species populating insulin system. Simulated time series describing the concentrations were used to characterize dynamic behavior of the system. First, we distinguished concentrations depending on overshoot presence. Then ,we evaluate response time and we further distinguished between rapid and slow dynamics. Hence, we classified dynamics according to four main pattern sets: rapid (non-overshooting) response, slow (not-overshooting) response, rapid overshooting response, and slow overshooting response. By this analysis we realized that rapid (not-overshooting) response characterizes those process related to insulin receptor phosphorylation, such as IR-RasGAP binding formation and Src activation. Rapid overshooting response characterizes those mechanisms related to Tyr- phosphorylated IRS1, such as IR-PI3K, IR-SHP2, and IR-GS complex formation, PI(3,4,5)P3 lipids production, Thr308- phosphorylated Akt, translocated GLUT4, and GTP- bound Ras. Slow overshooting response characterizes MAPK kinases cascade constituted by once- and double- active Raf, double- phosphorylated Mek and once- and double- phosphorylated Erk. Last, slow (not-overshooting) response characterizes those species implementing negative feedback loops in insulin system, such as Ser473- phosphorylated Akt, phosphorylated AMPK, Ser1387- phosphorylated TSC1-TSC2 complex, phosphorylated mTORC1, phosphorylated mTORC2, phosphorylated S6k, Ser- phosphorylated IRS1, and inhibited GS. We used these results to clear some fundamental regulation processes. In particular, we observed that system dynamics in markedly determined by IRS1 phosphorylation mechanism. Insulin receptor induces rapid phosphorylation of IRS1 on tyrosine residues. Subsequently, activation cascade of downstream elements is induced up to S6K, which controls IRS1 phosphorylation on serine residues. This causes a decrease in Tyr- phophorylated IRS1 which describes a overshooting response. We hypothesized this mechanism responsible in speeding up system response. System response is a important requirement in transduction networks which act in a period of time of minutes or seconds. In particular, insulin signaling pathway, upon stimulation,

must assure all those mechanisms which allow blood glucose level to return under physiologic conditions. Glucose transporter GLUT4 is involved in these processes and, as it is possible to observed from simulations, it translocates to cell membrane very rapidly. Moreover, we used parametric characterization of each pattern to conduct further analysis on system dynamics. Finally, we discussed model validation comparing predictions with some experimental data obtained from immunoblot. We mainly focused on description of significant dynamic features, showing good agreement.

Hence, computational model of insulin signaling pathway permitted dynamic characterization of the system and of involved regulation processes, and to unveil complex mechanisms and function of chemical species. This can be useful to distinguish pathway in pathologic conditions and to recognize the presence of pathologies, such as *insulin resistance* and *diabetes type 2*. Moreover, this model may be used to analyze cross-talk action with other signaling pathways, in particular that one induced by EGF which is particularly studied due to its connection with cancer.

7 APPENDIX

7.1 INSULIN MODEL IMPLEMENTED USING BIONETGEN

```
begin parameters
   NA 6.02e23 #Avogadro constant
   Vo 1.0e-10 #Volume outside the cell
   Vm 3.0e-13 #Volume of the cellular membrane
   V 3.0e-12 #Volume inside the cell

   PTP 1.0
   PTEN 1.0
   SHIP 1.0
   PKC 1.0
   IRp 8.97e-13*NA*Vm*(1.0e5)
   PI3K 2.54e-15*NA*V*(1.0e5)
   APequil 100/11

   nM 1.0e-9
   u NA*V*nM
   um NA*Vm*nM
   uo NA*Vo*nM
   s 1/60

   # Initial conditions
   I_init 1.0e-7*NA*Vo
   IR_ext_free_init 9.0e-13*NA*Vm*(1.0e5)
   IR_int_free_init 1.0e-13*NA*V*(1.0e5)
   IR_gen 35*1.67e-18*NA*V*(1.0e5)
   IRS1_init 1.0e-12*NA*V*(1.0e5)
   PI3K_init 1.0e-13*NA*V*(1.0e5)
   PI45_init 99.4 #%
   PI345_init 0.31 #%
   PI34_init 0.29 #%
   PKC_init 100 #%
   GLUT4_in_init 96 #%
   GLUT4_ex_init 4 #%
   GLUT4_gen 96*0.001155 #%

   Amino_Acids_input 10
   AMPK_init 20.50644836
   mTORC1_init 25.14
   mTORC2_init 18.79587754
   p70S6K_init 14.30096315
   PRAS40_S183_init 13.56128255
   PRAS40_T246_init 17.55000883
   PRAS40_init 13.56128255+17.55000883
   TSC1_TSC2_pT1462_init 14.9175

   GS_init 200*u
   RasGAP_init 50*u
   SHP2_init 300*u
   iSrc_init 518*u
   dRas_init 150*u
   Raf_init 100*u
   Mek_init 200*u
   Erk_init 400*u
   Akt_init 100*u
```

```
  # Correction factors
  k_p70S6K_IRS1 0.00003
  k_IRS1_AMPK 0.0000004
  k_mTORC2_Akt 0.05
  k_Complex_mTORC2 0.00012
  k_Akt_TSC1_TSC2 0.0015

  k_RasGAP 0.18
  k_Src 0.1
  k_GS 0.08
  k_SHP2 0.08

  # Rate constants min-1 (CONSISTENT UNITS)
  k1 0.5*6e7/(NA*Vo) #1st insuline binding
  k_1 0.20 #1st insuline unbinding
  k2 6e7/(NA*Vo) #2nd insuline binding
  k_2 0.5*20 #2nd insuline unbinding
  k3 2500 #IR phosphorylation
  k_3 0.20 #IR dephosphorylation [PTP]
  k4 3.333e-4 #IR free internalization
  k_4 0.003 #IR free externalization
  k4p 2.1e-3 #IR internalization
  k_4p 2.1e-4 #IR externalization
  k5 1 #IR internalizated production
  k_5 1.67e-18 #IR internalizated degradation
  k6 0.461 #IR internalizated unphosphorylation [PTP]
  k7 4.16/IRp #IRS-1 phosphorylation in Tyr
  k_7 (2.5/7.45)*4.16 #IRS-1 dephosphorylation in Tyr
  k7p 0.69/2 #IRS-1 phosphorylation in Ser
  k_7p (0.69/2)*((2.5/7.45)*3.7e-13)/(6.27e-13-(2.5/7.45)*3.7e-13) #IRS-1 dephosphorylation
in Ser
  k8 10.0*(5.0/70.775)*(1.0e12/(NA*V*1.0e5)) #IRS-1_PI3-k complex activation
  k_8 10.0 #IRS-1_PI3-k complex deactivation
  k9a (1.39-(0.31/99.4)*(94/3.1)*1.39)*(1/PI3K) #PI(4,5)P2 -> PI(3,4,5)P3
  k9b (0.31/99.4)*(94/3.1)*1.39
  k_9 (94/3.1)*1.39 #PI(3,4,5)P3 -> PI(4,5)P2 [PTEN]
  k10 (3.1/2.9)*2.77 #PI(3,4)P2 -> PI(3,4,5)P3
  k_10 2.77 #PI(3,4,5)P3 -> PI(3,4)P2 [SHIP]
  k11a (0.1*10*0.69)/(3.1-0.31) #Akt phosphorylation
  k11b (0.1*10*0.69*0.31)/(0.31-3.1)
  k_11 10*0.69 #Akt dephosphorylation
  k12a (0.1*10*0.69)/(3.1-0.31) #PKC phosphorylation
  k12b (0.1*10*0.69*0.31)/(0.31-3.1)
  k_12 10*0.69 #PKC dephosphorylation
  k13 (4/96)*0.167 #GLUT4 basal translocation to cell membrane
  k_13 0.167 #GLUT4 internalization
  k13a ((40/60)-(4/96))*0.167*(0.2/APequil) #Akt dependent - GLUT4 translocation
  k13b ((40/60)-(4/96))*0.167*(0.8/APequil) #PKC dependent - GLUT4 translocation
  k14 1 #GLUT4 production
  k_14 0.001155 #GLUT4 degradation

  p5 k_p70S6K_IRS1*1682.75 #IRS1_p_phosphorylation_by_p70S6K_pT389
  p7 k_IRS1_AMPK*9.79766 #AMPK_T172_phosphorylation
  p8 0.0107215 #AMPK_pT172_dephosphorylation
  p10 0.00640216 #Akt_pS473_dephosphorylation
  p11 k_mTORC2_Akt*13.1442 #Akt_S473_phosphorylation_by_mTORC2_pS2481_n_IRS1_p
  p13 0.0106652 #mTORC1_pS2448_dephosphorylation_by_TSC1_TSC2_pS1387
  p14 0.00438916 #mTORC1_S2448_activation_by_Amino_Acids
  p15 0.0183735 #mTORC2_pS2481_dephosphorylation
  p16 k_Complex_mTORC2*0.375353 #mTORC2_S2481_phosphorylation_by_PI3K_variant_p
  p17 0.0113512 #p70S6K_pT389_dephosphorylation
  p18 0.00184043 #p70S6K_T389_phosphorylation_by_mTORC1_pS2448
  p19 2.33014 #PRAS40_pS183_dephosphorylation
  p20 1.60513 #PRAS40_pT246_dephosphorylation
  p21 0.187621 #PRAS40_S183_phosphorylation_by_mTORC1_pS2448
  p22 0.137729 #PRAS40_T246_phosphorylation_by_Akt_pT308
```

```
p23 0.0365589 #TSC1_TSC2_S1387_phosphorylation_by_AMPK_pT172
p24 k_Akt_TSC1_TSC2*0.0177562 #TSC1_TSC2_T1462_phosphorylation_by_Akt_pT308

kon27 k_RasGAP*6.66e-8/(u*s)
kd27 2e6*u*kon27/k_RasGAP
kf34 1.85e-4/(s)
kcat40 6.66/(s)
alpha40 2.5e-4
Km40 110*u
V41 6.66*u/(s)
Km41 50*u
kon45 k_GS*6.66e-4/(u*s)
kd45 1e5*u*kon45/k_GS
kon47 k_SHP2*6.66e-4/(u*s)
kd47 1000*u*kon47/k_SHP2
kcat62 5.33/(s)
Km62 50*u
bRasGAP 1e-5*u
kcat63 2e4/(s)
Km63 50*u
kcat65 0.1/(s)
Km65 400*u
kcat66 3.33/(s)
Km66 10*u
kcat67 0.666/(s)
PKA 100*u
Km67 1e4*u
alpha67 (1e-6)/(u*s)
beta67 2
kcat68 0.133/(s)
Km68 50*u
V69 16.6*u/(s)
Km69 675.299*u
kcat70 0.333/(s)
Km70 500*u
Km71 500*u
kcat71 0.666/(s)
V72 33.3*u/(s)
Km72 500*u
Km73 500*u
V73 23.33*u/(s)
kcat79 0.0466/(s)
Km79 5000*u
k_79 6.66e-5/(s)
V100 333*u/(s)
Km100 143.3*u
kf101 0.666/(s)
k115 0.0133/(u*s)
end parameters
begin molecule types
I(bs)
IR(bs,alpha,alpha,Y~u~p,loc~ex~c)
IR_Gen()
IRS1(bs,YXXM,Y~u~p,S~u~p)
PI3K(SH2)
PI45()
PI345()
PI34()
PKC(T~u~p)
GLUT4(loc~c~ex)
GLUT4_Gen()
Trash()

Amino_Acids()
Akt(T308~u~p,S473~u~p)
PRAS40(T246~u~p,S183~u~p)
```

```
TSC1_TSC2(T1462~u~p,S1387~u~p)
mTORC1(S2448~u~p)
mTORC2(S2481~u~p)
p70S6K(T389~u~p)
AMPK(T172~u~p)

GS(bs,state~i~a)
Src(state~i~a)
RasGAP(bs)
SHP2(bs)
Ras(state~d~t)
Raf(state~i~a~aa)
Mek(state~u~pp)
Erk(state~u~p~pp)
end molecule types
begin seed species
I(bs) I_init
IR(bs,alpha,alpha,Y~u,loc~ex) IR_ext_free_init
IR(bs,alpha,alpha,Y~u,loc~c) IR_int_free_init
IR_Gen() IR_gen
IRS1(bs,YXXM,Y~u,S~u) IRS1_init
PI3K(SH2) PI3K_init
PI45() PI45_init
PI345() PI345_init
PI34() PI34_init
PKC(T~u) PKC_init
GLUT4(loc~c) GLUT4_in_init
GLUT4(loc~ex) GLUT4_ex_init
GLUT4_Gen() GLUT4_gen
Amino_Acids() Amino_Acids_input
AMPK(T172~u) AMPK_init
Akt(T308~u,S473~u) Akt_init
mTORC1(S2448~u) mTORC1_init
mTORC2(S2481~u) mTORC2_init
p70S6K(T389~u) p70S6K_init
PRAS40(T246~u,S183~u) PRAS40_init
TSC1_TSC2(T1462~p,S1387~u) TSC1_TSC2_pT1462_init
GS(bs,state~a) GS_init
RasGAP(bs) RasGAP_init
SHP2(bs) SHP2_init
Src(state~i) iSrc_init
Ras(state~d) dRas_init
Raf(state~i) Raf_init
Mek(state~u) Mek_init
Erk(state~u) Erk_init
end seed species
begin observables
Molecules IR_free IR(alpha,alpha,Y~u,loc~ex)
Molecules IR_1bound IR(alpha,alpha!+,Y~p,loc~ex)
Species IR_2bound IR(alpha!+,alpha!+,Y~p,loc~ex)
Molecules IR_phosph IR(bs,Y~p,loc~ex)
Molecules Insulin I(bs)
Molecules IRS1_unphosph IRS1(bs,YXXM,Y~u,S~u)
Molecules IRS1_phosph_tyr IRS1(YXXM,Y~p,S~u)
Molecules IRS1_phosph_ser IRS1(YXXM,Y~u,S~p)
Molecules Complex IRS1(YXXM!1,Y~p).PI3K(SH2!1)
Molecules PI345 PI345()
Molecules PI34 PI34()
Molecules PKC_phosph PKC(T~p)
Molecules Akt_phosph Akt(T308~p,S473~p)
Molecules GLUT4_ex GLUT4(loc~ex)
Molecules IR_bound IR(alpha!+,Y~p,loc~ex)
Molecules AMPK_pT172 AMPK(T172~p)
Molecules Akt_pT308 Akt(T308~p,S473~u)
Molecules Akt_pS473 Akt(T308~u,S473~p)
Molecules Akt_pT Akt(T308~p)
```

```
  Molecules Akt_pS Akt(S473~p)
  Molecules TSC1_TSC2_pS1387 TSC1_TSC2(T1462~u,S1387~p)
  Molecules TSC1_TSC2_pT1462 TSC1_TSC2(T1462~p,S1387~u)
  Molecules mTOR_pS2448 mTORC1(S2448~p)
  Molecules mTOR_pS2481 mTORC2(S2481~p)
  Molecules p70S6K_pT389 p70S6K(T389~p)
  Molecules PRAS40_pT246 PRAS40(T246~p)
  Molecules PRAS40_pS183 PRAS40(S183~p)
  Molecules iSrc Src(state~i)
  Molecules aSrc Src(state~a)
  Molecules dRas Ras(state~d)
  Molecules IRSp_GS IRS1(bs!1,Y~p,S~u).GS(bs!1,state~a)
  Molecules tRas Ras(state~t)
  Molecules IRp_RasGAP IR(bs!1,Y~p,loc~ex).RasGAP(bs!1)
  Molecules Raf Raf(state~i)
  Molecules aRaf Raf(state~a)
  Molecules aaRaf Raf(state~aa)
  Molecules Mek Mek(state~u)
  Molecules ppMek Mek(state~pp)
  Molecules Erk Erk(state~u)
  Molecules pErk Erk(state~p)
  Molecules ppErk Erk(state~pp)
  Molecules GS GS(bs,state~a)
  Molecules iGS GS(bs,state~i)
  Molecules IRSp_SHP2 IRS1(bs!1,Y~p,S~u).SHP2(bs!1)
  end observables
  begin functions
  f7()=k7*IR_bound
  f9()=k9a*Complex+k9b
  f11()=k11a*PI345+k11b
  f12()=k12a*PI345+k12b
  f13()=k13+k13a*(100-(Akt_pT308+Akt_pS473+Akt_phosph)*(100/Akt_init))+k13b*PKC_phosph
  #feedback loops
  ptp1()=(k_3)*(1-0.25*((100-(Akt_pT308+Akt_pS473+Akt_phosph)*(100/Akt_init))/(100/11)))
  ptp2()=(k6)*(1-0.25*((100-(Akt_pT308+Akt_pS473+Akt_phosph)*(100/Akt_init))/(100/11)))
  ptp3()=(k_7)*(1-0.25*((100-(Akt_pT308+Akt_pS473+Akt_phosph)*(100/Akt_init))/(100/11)))
  pkc()=(k7p)#*V_max*PKC_del^n/(Kd^n+PKC_del^n)
  v40()=k_Src*kcat40*(alpha40*IR_phosph)/(Km40+iSrc)
  v41()=V41/(Km41+aSrc)
  v62()=kcat62*(IRSp_GS)/(Km62+dRas)
  v63()=kcat63*(bRasGAP+IRp_RasGAP)/(Km63+tRas)
  v65()=kcat65*tRas/(Km65+Raf)
  v66()=kcat66*aSrc/(Km66+aRaf)
  v67()=kcat67*PKA/(Km67+aaRaf)+alpha67*(Akt_pT308+beta67*Akt_phosph)
  v68()=kcat68*aaRaf/(Km68+Mek)
  v69()=V69/(Km69+ppMek)
  v70()=kcat70*ppMek/(Km70+Erk+pErk*Km70/Km71)
  v71()=kcat71*ppMek/(Km71+pErk+Erk*Km71/Km70)
  v72()=V72/(Km72+ppErk+pErk*Km72/Km73)
  v73()=V73/(Km73+pErk+ppErk*Km73/Km72)
  v79()=kcat79*ppErk/(Km79+GS)
  v_79()=k_79
  v100()=V100/(Km100+IRSp_GS)
  v115()=k115*IRSp_SHP2
  end functions
  begin reaction rules
  # Receptor binding 1st insulin molecule
  1 IR(bs,alpha,alpha,Y~u,loc~ex) + I(bs) <-> IR(bs,alpha!1,alpha,Y~u,loc~ex).I(bs!1)
k1,k_1
  # Receptor phosphorylation
  2 IR(bs,alpha!+,alpha,Y~u,loc~ex) -> IR(bs,alpha!+,alpha,Y~p,loc~ex) k3
  # Receptor binding 2nd insulin molecule
  3 IR(alpha!+,alpha,Y~p,loc~ex) + I(bs) <-> IR(alpha!+,alpha!1,Y~p,loc~ex).I(bs!1)
k2,k_2
  # Receptor unbinding and dephosphorylation (on the cell membrane)
```

```
   4 IR(bs!?,alpha!1,alpha,Y~p,loc~ex).I(bs!1) -> IR(bs,alpha,alpha,Y~u,loc~ex) + I(bs)
ptp1()
   # Phosphorylated receptor internalization/externalization
   5 IR(bs!?,alpha!+,Y~p,loc~ex) <-> IR(bs,alpha!+,Y~p,loc~c) k4p, k_4p
   # Receptor unbinding and dephosphorylation (inside the cell)
   6 IR(bs,alpha!+,Y~p,loc~c) -> IR(bs,alpha,alpha,Y~u,loc~c) ptp2()
   # Free receptor internalization/externalization
   7 IR(bs,alpha,alpha,Y~u,loc~ex) <-> IR(bs,alpha,alpha,Y~u,loc~c) k4, k_4
   # Receptor synthesis
   8 IR_Gen() -> IR_Gen() + IR(bs,alpha,alpha,Y~u,loc~c) k5
   # Receptor degradation
   9 IR(bs,alpha,alpha,Y~u,loc~c) -> Trash() k_5
   # IRS-1 phosphorylation/dephosphorylation in Tyr
   10 IRS1(bs,YXXM,Y~u,S~u) <-> IRS1(bs,YXXM,Y~p,S~u) f7(),ptp3()
   # IRS-1_PI3-K complex formation (PI3-K activation)
   11 IRS1(YXXM,Y~p,S~u) + PI3K(SH2) <-> IRS1(YXXM!1,Y~p,S~u).PI3K(SH2!1) k8,k_8
   # IRS-1 phosphorylation/dephosphorylation in Ser
   12 IRS1(bs,YXXM,Y~u,S~u) <-> IRS1(bs,YXXM,Y~u,S~p) pkc(), k_7p
   # lipids formation
   13 PI45() <-> PI345() f9(),k_9*PTEN
   # lipids formation
   14 PI34() <-> PI345() k10,k_10*SHIP
   # Akt phosphorylation/dephosphorylation
   15 Akt(T308~u) <-> Akt(T308~p) f11(),k_11
   # PKC phosphorylation /dephosphorylation
   16 PKC(T~u) <-> PKC(T~p) f12(), k_12
   # GLUT4 translocation
   17 GLUT4(loc~c) <-> GLUT4(loc~ex) f13(), k_13
   # GLUT4 synthesis
   18 GLUT4_Gen() -> GLUT4_Gen() + GLUT4(loc~c) k14
   # GLUT4 degradation
   19 GLUT4(loc~c) -> Trash() k_14
   # mTORC2
   20 mTORC2(S2481~u) + IRS1(YXXM!1,Y~p,S~u).PI3K(SH2!1) -> mTORC2(S2481~p)
   + IRS1(YXXM!1,Y~p,S~u).PI3K(SH2!1) p16
   21 mTORC2(S2481~p) -> mTORC2(S2481~u) p15
   # Akt
   22 Akt(S473~u) + mTORC2(S2481~p) -> Akt(S473~p) + mTORC2(S2481~p) p11
   23 Akt(S473~p) -> Akt(S473~u) p10
   # AMPK
   24 AMPK(T172~u) + IRS1(YXXM,Y~p,S~u) -> AMPK(T172~p) + IRS1(YXXM,Y~p,S~u) p7
   25 AMPK(T172~p) -> AMPK(T172~u) p8
   # TSC1-TSC2
   26 TSC1_TSC2(T1462~u,S1387~p) + Akt(T308~p) -> TSC1_TSC2(T1462~p,S1387~u) + Akt(T308~p)
p24
   27 TSC1_TSC2(T1462~p,S1387~u) + AMPK(T172~p) -> TSC1_TSC2(T1462~u,S1387~p) + AMPK(T172~p)
p23
   # mTORC1
   28 mTORC1(S2448~p) + TSC1_TSC2(T1462~u,S1387~p) -> mTORC1(S2448~u)
   + TSC1_TSC2(T1462~u,S1387~p) p13
   29 mTORC1(S2448~u) + Amino_Acids() -> mTORC1(S2448~p) + Amino_Acids() p14
   # p70S6K
   30 p70S6K(T389~u) + mTORC1(S2448~p) -> p70S6K(T389~p) + mTORC1(S2448~p) p18
   31 p70S6K(T389~p) -> p70S6K(T389~u) p17
   # IRS1
   32 IRS1(YXXM,Y~u,S~u) + p70S6K(T389~p) -> IRS1(YXXM,Y~u,S~p) + p70S6K(T389~p) p5
   # PRAS40
   33 PRAS40(T246~u) + Akt(T308~p) -> PRAS40(T246~p) + Akt(T308~p) p22
   34 PRAS40(T246~p) -> PRAS40(T246~u) p20
   35 PRAS40(S183~u) + mTORC1(S2448~p) -> PRAS40(S183~p) + mTORC1(S2448~p) p21
   36 PRAS40(S183~p) -> PRAS40(S183~u) p19
   # RasGAP module
   37 IR(bs,Y~p,loc~ex) + RasGAP(bs) <-> IR(bs!1,Y~p,loc~ex).RasGAP(bs!1) kon27,kd27
   38 IR(bs!1,alpha!+,Y~p,loc~ex).RasGAP(bs!1) -> RasGAP(bs) kf34 DeleteMolecules
   # Src module
   39 Src(state~i) <-> Src(state~a) v40(), v41()
```

```
  # IRS1-GS and IRS1-SHP2 complex formation
  40 IRS1(bs,Y~p,S~u) + GS(bs,state~a) <-> IRS1(bs!1,Y~p,S~u).GS(bs!1,state~a) kon45,kd45
  41 IRS1(bs,Y~p,S~u) + SHP2(bs) <-> IRS1(bs!1,Y~p,S~u).SHP2(bs!1) kon47,kd47
  # IRS1-GS and IRS1-SHP2 complex disruption and IRS1 dephosphorylation
  42 IRS1(bs!1,YXXM,Y~p,S~u).GS(bs!1,state~a) -> IRS1(bs,YXXM,Y~u,S~u) + GS(bs,state~a)
v100()
  43 IRS1(bs!1,YXXM,Y~p,S~u).SHP2(bs!1) -> IRS1(bs,YXXM,Y~u,S~u) + SHP2(bs) kf101
  # GS inhibition
  44 GS(bs,state~a) <-> GS(bs,state~i) v79(),v_79()
  # SHP2 activity against RasGAP
  45 IR(bs!1,Y~p,loc~ex).RasGAP(bs!1) -> IR(bs,Y~p,loc~ex) + RasGAP(bs) v115()
  # Ras module
  46 Ras(state~d) <-> Ras(state~t) v62(), v63()
  # Raf module
  47 Raf(state~i) -> Raf(state~a) v65()
  48 Raf(state~a) -> Raf(state~aa) v66()
  49 Raf(state~aa) -> Raf(state~i) v67()
  # Mek module
  50 Mek(state~u) <-> Mek(state~pp) v68(), v69()
  # ERK module
  51 Erk(state~u) <-> Erk(state~p) v70(), v73()
  52 Erk(state~p) <-> Erk(state~pp) v71(), v72()
  end reaction rules
  #actions
  generate_network({overwrite=>1});
  # Kinetics
  simulate_ode({t_end=>60,n_steps=>200});
```

7.2 INSULIN MODEL SIMULATIONS

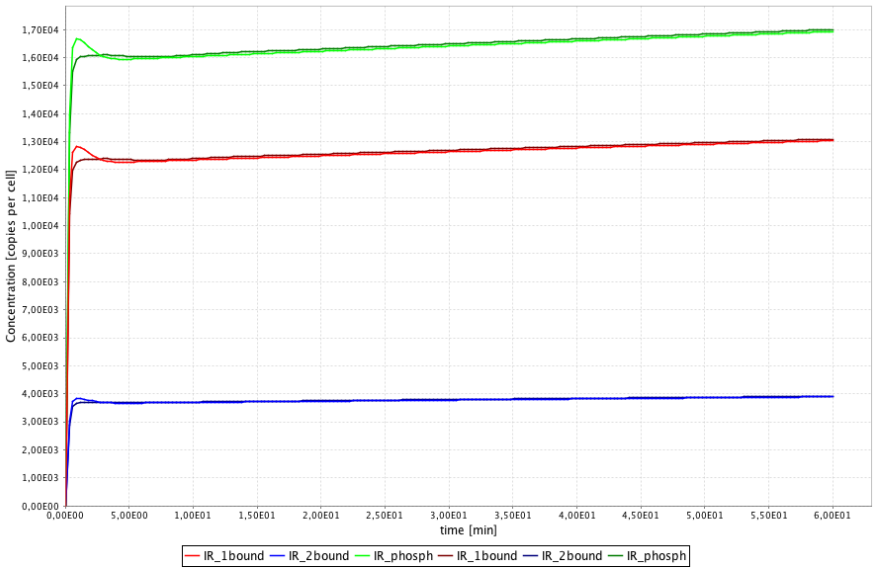

Figure 14: Temporal series of insulin receptor binding one (IR_1bound) or two (IR_2bound) insulin molecules and of phosphorylated insulin receptor (IR_phosph). Comparison between insulin model and Quon's model.

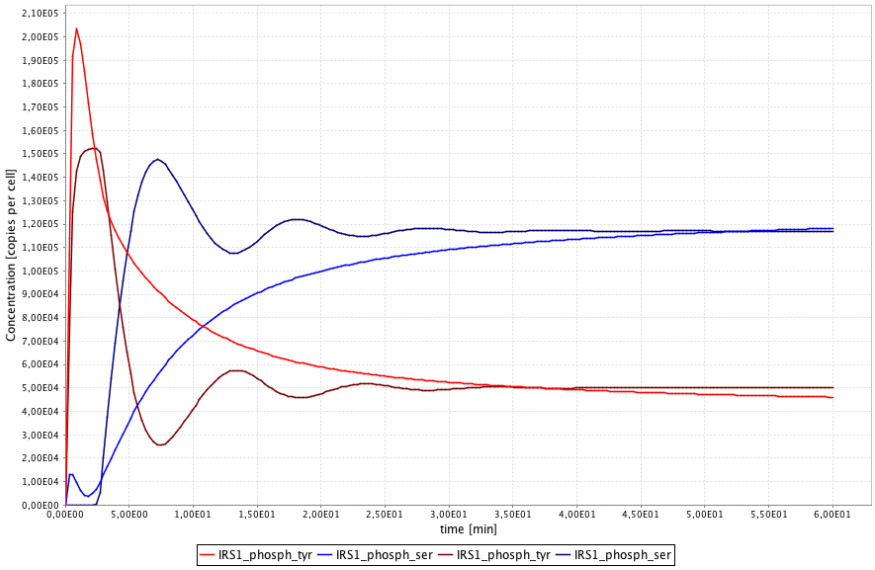

Figure 15: Temporal series of IRS-1 phosphorylated on tyrosine (IRS1_phosph_tyr) and on serine (IRS1_phosph_ser) residues. Comparison, respectively, between insulin model and Quon's model.

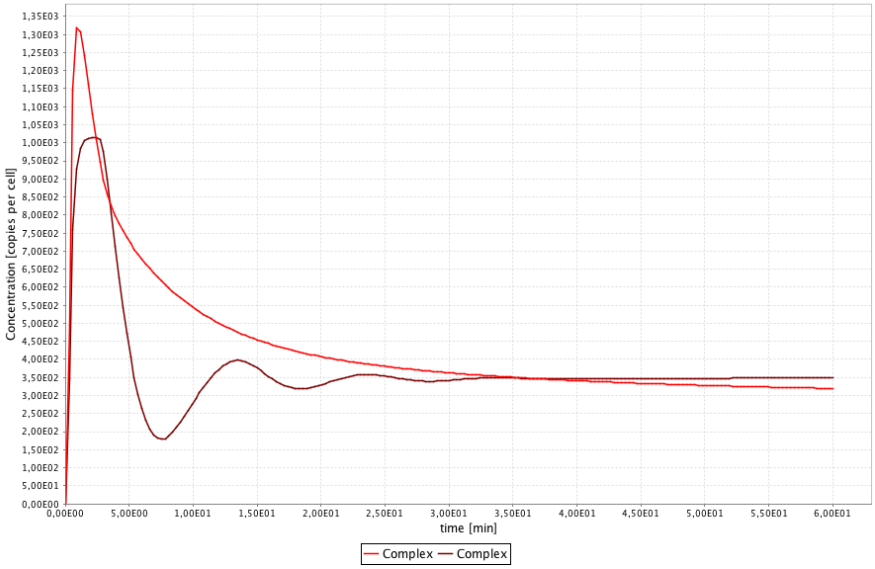

Figure 16: Temporal series of IRS1-PI3K complex formation. Comparison, respectively, between insulin model and Quon's model.

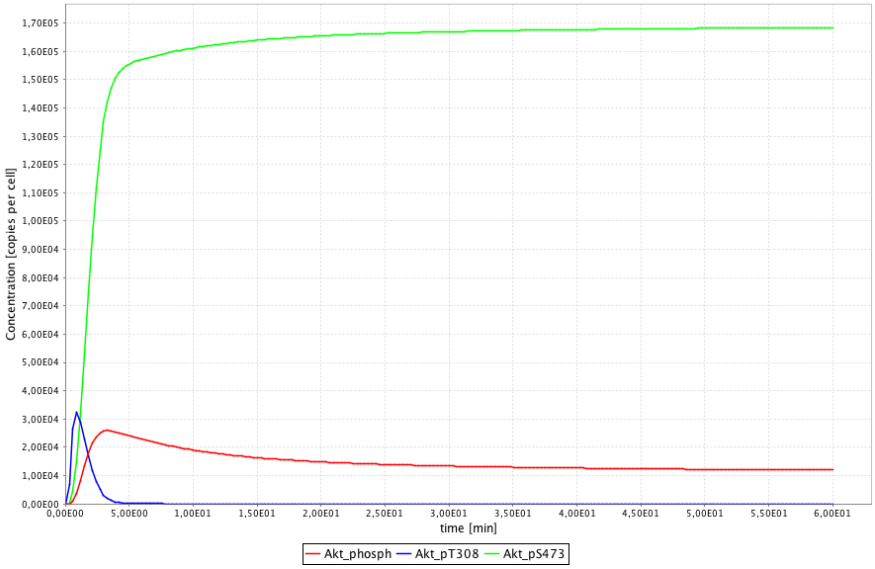

Figure 17: Temporal series of Akt phosphorylated on both residues (Akt_phosph), uniquely on Thr308 (Akt_pT308) and uniquely on Ser473 (Akt_pS473).

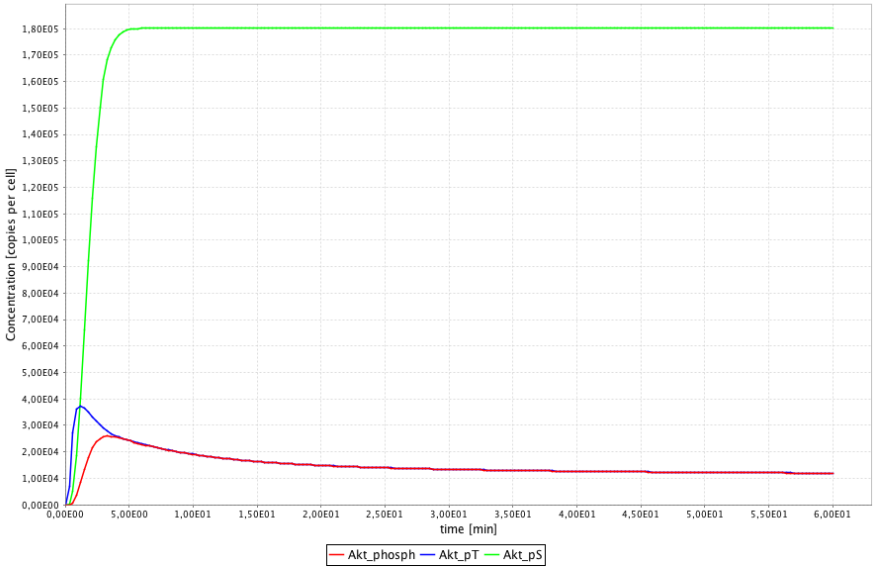

Figure 18: Temporal series of Akt phosphorylated on both residues (Akt_phosph), on Thr308 (Akt_pT) and on Ser473 (Akt_pS).

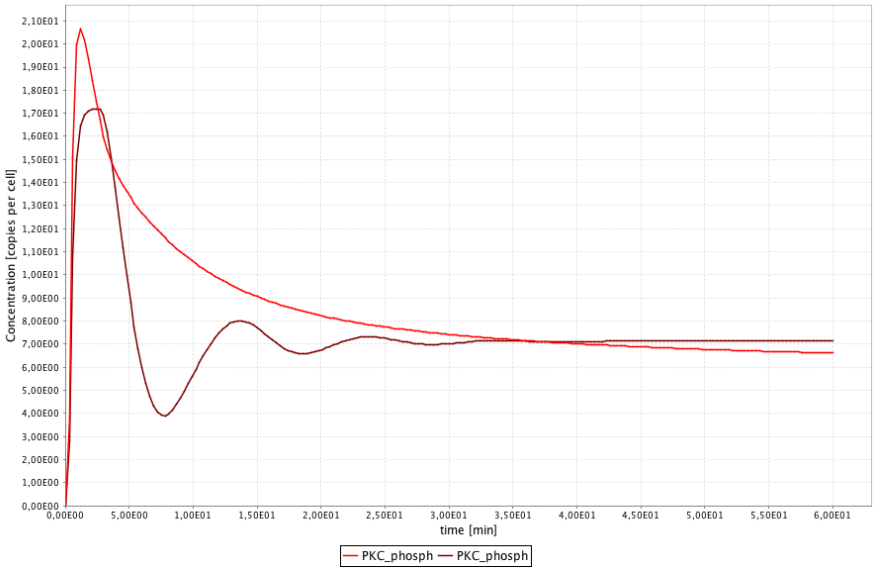

Figure 19: Temporal series of phosphorylated PKC (PKC_phosph). Comparison between insulin model and Quon's model.

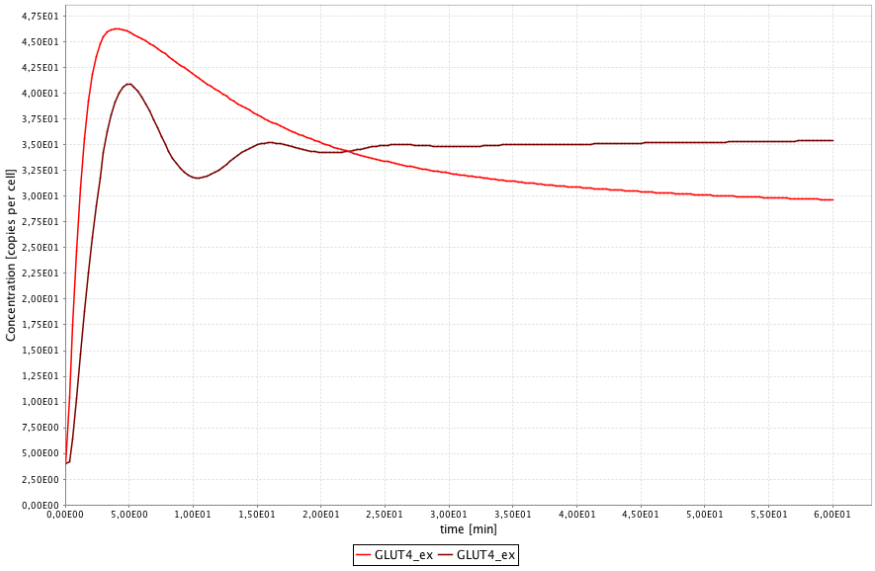

Figure 20: Temporal series of translocated GLUT4 (GLUT4_ex). Comparison between insulin model and Quon's model.

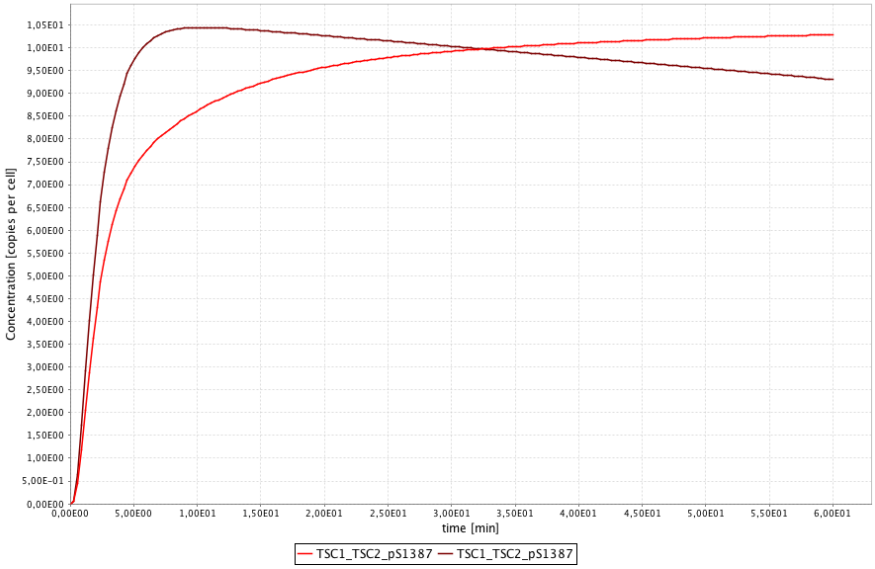

Figure 21: Temporal series of TSC1-TSC2 complex phosphorylated on S1387 (TSC1_TSC2_pS1387). Comparison between insulin model and Thedieck's model.

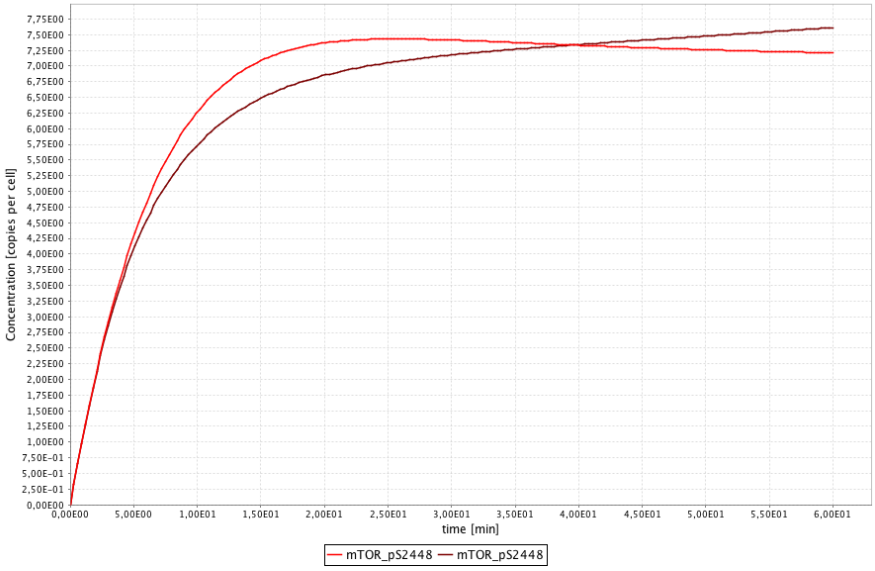

Figure 22: Temporal series of phosphorylated mTORC1 (mTOR_pS2448). Comparison between insulin model and Thedieck's model.

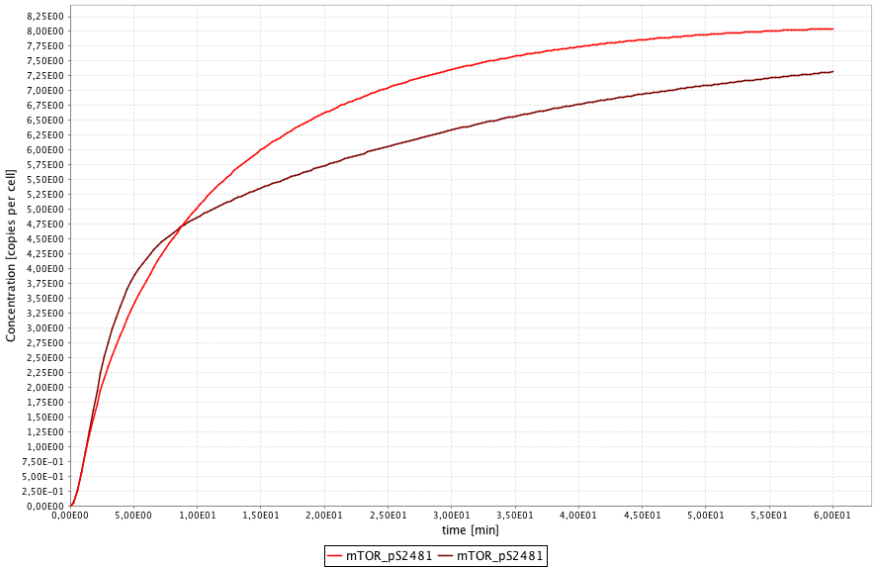

Figure 23: Temporal series of phosphorylated mTORC2 (mTOR_pS2481). Comparison between insulin model and Thedieck's model.

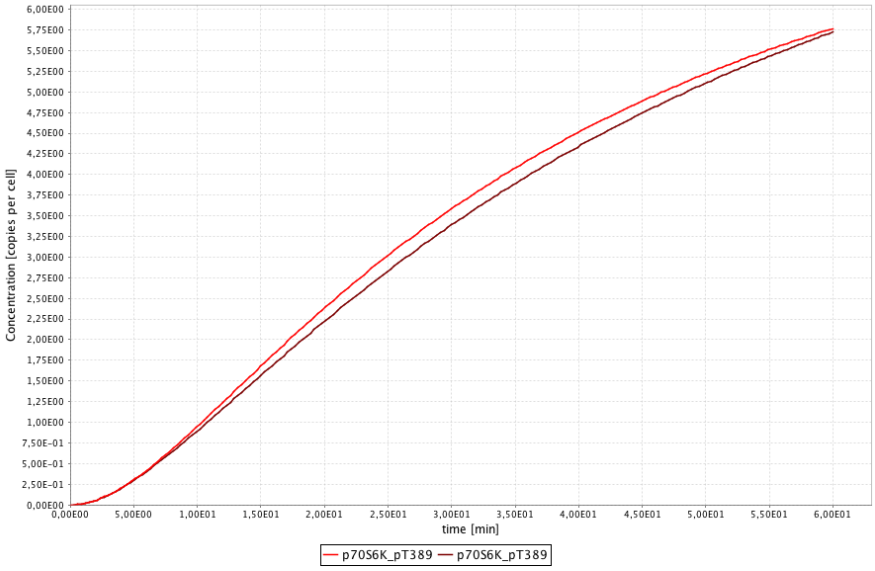

Figure 24: Temporal series of phosphorylated S6K (p70S6K_pT389). Comparison between insulin model and Thedieck's model.

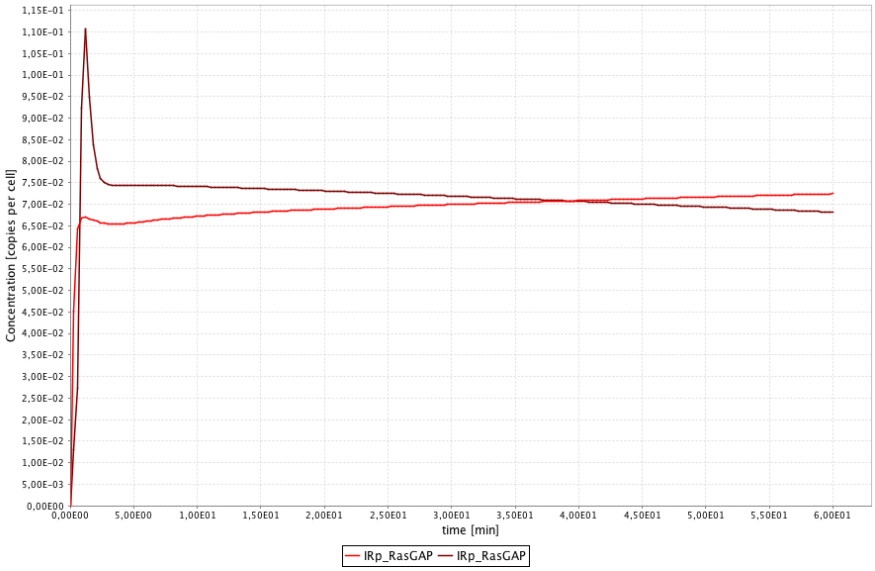

Figure 25: Temporal series of IR-RasGAP complex formation (IRp_RasGAP). Comparison between insulin model and Kholodenko's model.

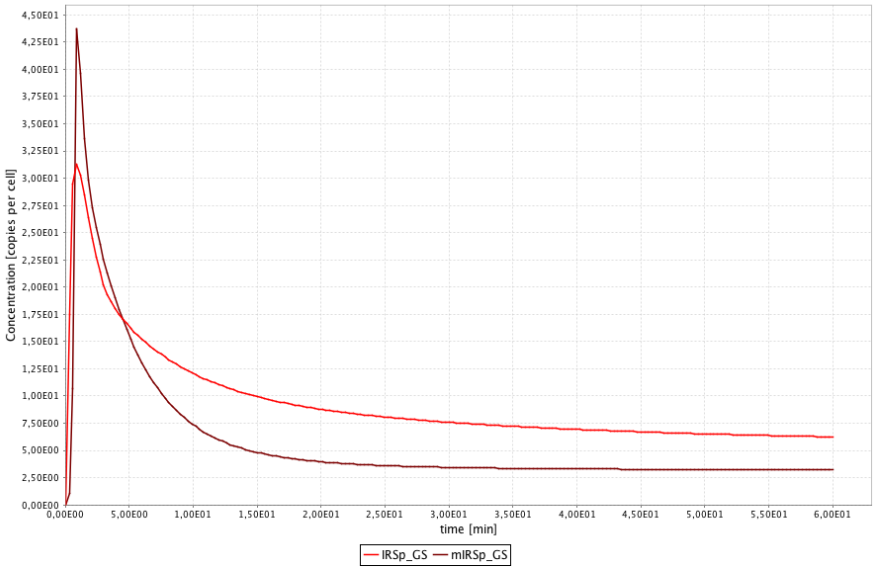

Figure 26: Temporal series of IRS1-GS complex formation (IRSp_GS). Comparison between insulin model and Kholodenko's model.

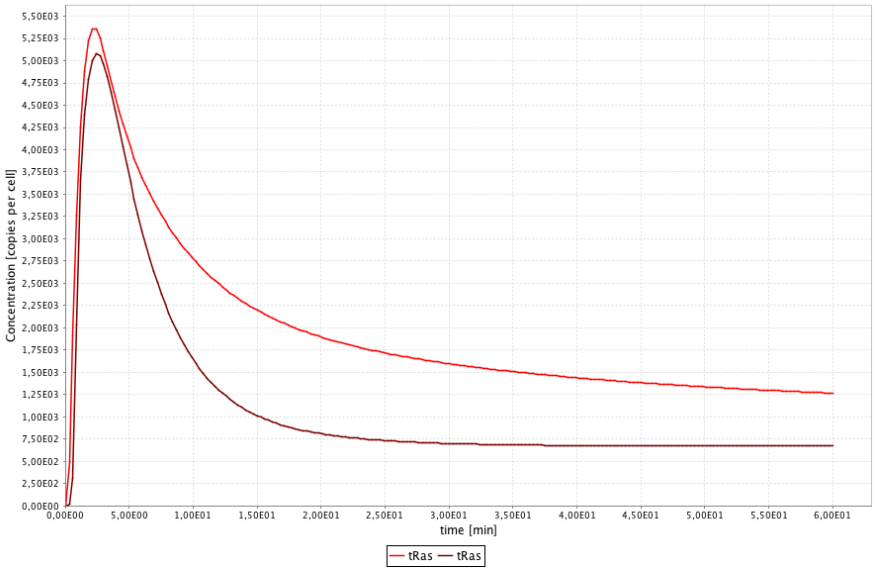

Figure 27: Temporal series of GTP- bound Ras (tRas). Comparison between insulin model and Kholodenko's model.

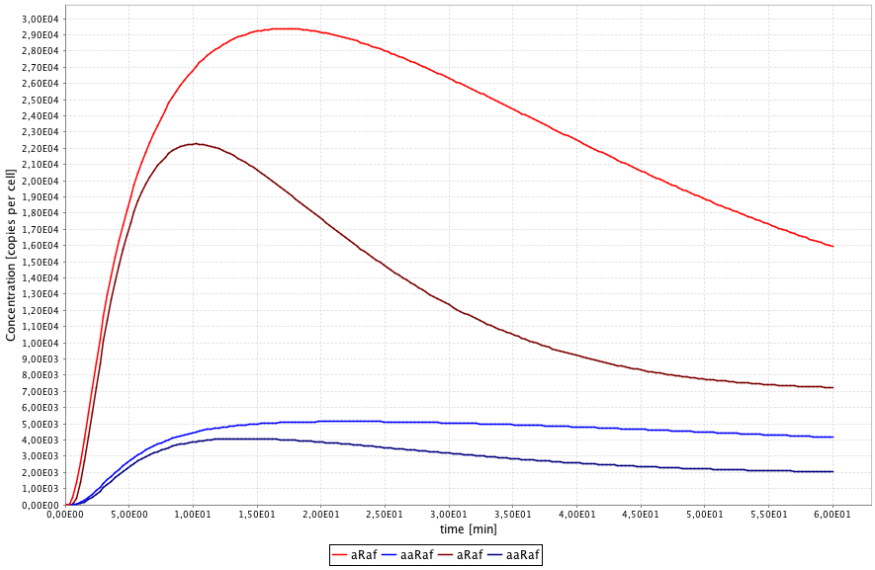

Figure 28: Temporal series of once- (aRaf) and double- (aaRaf) activated Ras. Comparison between insulin model and Kholodenko's model.

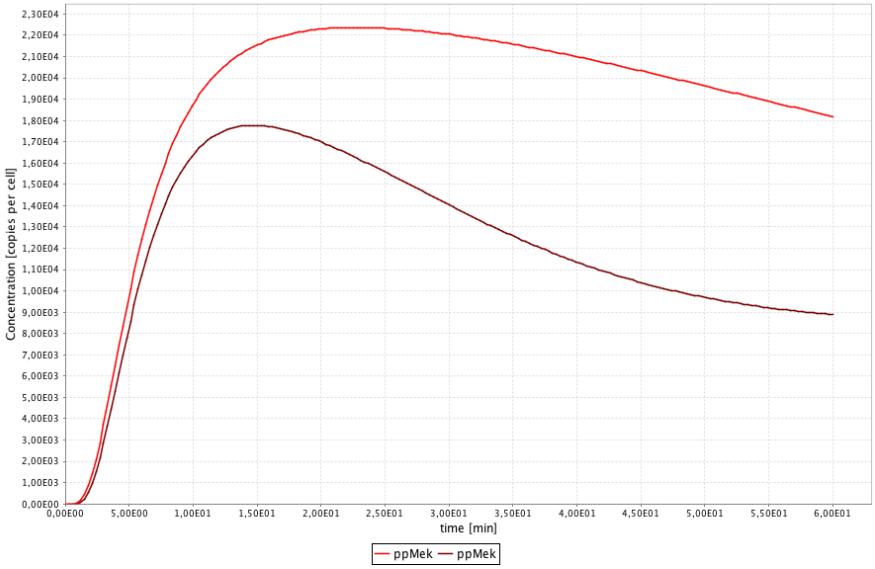

Figure 29: Temporal series of double- phosphorylated Mek (ppMek). Comparison between insulin model and Kholodenko's model.

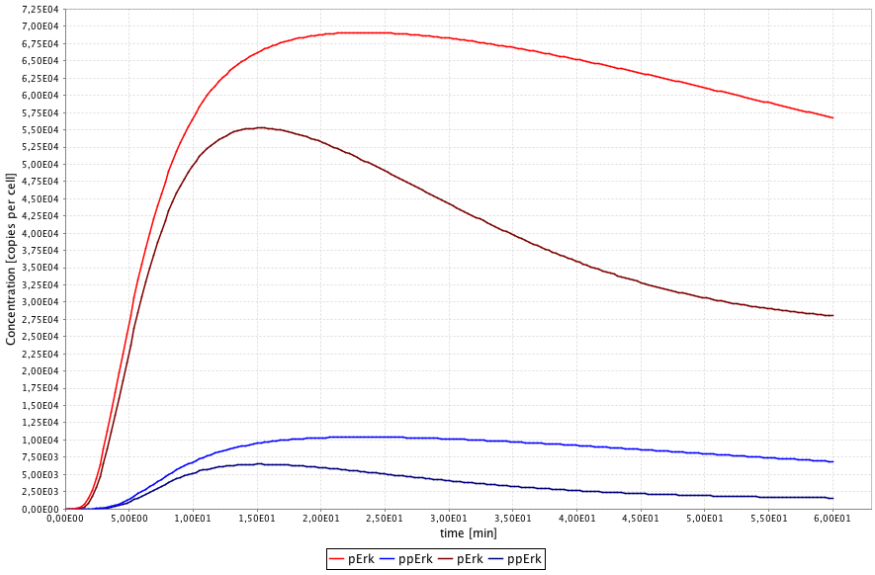

Figure 30: Temporal series of once- (pErk) and double- (ppErk) phosphorylated Erk. Comparison between insulin model and Kholodenko's model.

BIBLIOGRAPHY

[1] L. M. A. Aparicio. State of the art in silico tools for the study of signaling pathways in cancer. *Int. J. Mol. Sci.*, 13:6561–6581, May 2012. (Cited on page 1.)

[2] U. S. Bhalla. Signaling logic of activity-triggered dendritic protein synthesis: An mTOR gate but not a feedback switch. *PLoS Comput. Biol.*, 5, 2009. (Cited on page 2.)

[3] J. Capala. The crosstalk between EGF, IGF, and insulin cell signaling pathways-computational and experimental analysis. *BMC Systems Biology*, 3:88, September 2009. (Cited on page 48.)

[4] UniProt Consortium. UniProt. http://www.uniprot.org, 2013. (Cited on pages 7, 9, and 23.)

[5] E. J. David. Mapping insulin/GLUT4 circuitry. *Traffic*, 12:672–681, 2011. (Cited on pages 12 and 17.)

[6] P. De Meyts. Harmonic oscillator model of the insulin and igf1 receptorsÕ allosteric binding and activation. *Mol. Syst. Biol.*, 4 (243), 2009. (Cited on page 2.)

[7] T. Emonet. Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nature Methods*, 8(2):177–183, February 2011. (Cited on page 31.)

[8] D. T. Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, October 2007. (Cited on page 30.)

[9] NIH/NIGMS Grant 1 R01 GM078005-01. UCSD-Signaling Gateway. http://www.signaling-gateway.org/search/, 2013. (Cited on page 23.)

[10] D. J. Harrison. Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of pten in resistance to trastuzumab. *Cancer Res.*, 69:6713–6720, 2009. (Cited on page 2.)

[11] W. S. Hlavacek. A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. *BioSystems*, 83:136–151, June 2005. (Cited on page 24.)

[12] B. N. Kholodenko. Systems-level interactions between insulin-EGF networks amplify mitogenic signaling. *Molecular Systems Biology*, 5(256):1–15, February 2009. (Cited on pages xii, 3, 24, 35, 48, 49, and 70.)

[13] E. Klipp. *Systems Biology: A Textbook*. Wiley-Blackwell, 2009. (Cited on pages 1 and 60.)

[14] J Krivine. Scalable simulation of cellular signaling networks. *APLAS*, 4807:139–157, 2007. (Cited on page 25.)

[15] Olefsky Jerrold M. LeRoith Derek, Taylor Simeon I. *Diabetes Mellitus: A Fundamental and Clinical Text*. Lippincott Williams and Wilkins, 3rd edition, 2004. (Cited on pages 7, 9, 15, and 17.)

[16] Ivan V. Maly. *Methods in Molecular Biology, Systems Biology*. Humana Press, vol. 500 edition, 2009. (Cited on pages xi, 25, 27, and 30.)

[17] M. Matsui. Biopathways representation and simulation on hybrid functional petri net. *In Silico Biology*, 3:389–404, 2003. (Cited on page 22.)

[18] S. S. Mohan. NetPath: A public resource of curated signal transduction pathways. *Genome Biology*, 11(R3), 2010. (Cited on page 23.)

[19] F. C. Morales. Identification of optimal drug combinations targeting cellular networks: Integrating phospho-proteomics and computational network analysis. *Cancer Research*, 70:6704–6714, July 2010. (Cited on page 23.)

[20] Kyoto Encyclopedia of Genes and Genomes. KEGG PATHWAY database. `http://www.genome.jp/kegg/pathway.html`, 2013. (Cited on page 23.)

[21] R. G. Posner. Specification, annotation, visualization and simulation of a large rule-based model for ERBB receptor signaling. *BMC Systems Biology*, 6(107):1–15, August 2012. (Cited on page 25.)

[22] M. J. Quon. A mathematical model of metabolic insulin signaling pathways. *Am J Physiol Endocrinol Metab*, 283:1084–1101, July 2002. (Cited on pages xi, 2, 35, and 36.)

[23] P. T. Ram. The signaling petri net-based simulator: A non-parametric strategy for characterizing the dynamics of cell-specific signaling networks. *PLoS Computational Biology*, 4(2):1–15, February 2008. (Cited on page 22.)

[24] D. M. Sabatini. mTOR: from growth signal integration to cancer, diabetes and ageing. *Nature Reviews, Molecular Cell Biology*, 12: 21–35, January 2011. (Cited on pages 13 and 43.)

[25] A. R. Saltiel. Regulation of glucose transport by insulin: traffic control of GLUT4. *Nature Reviews, Molecular Cell Biology*, 13:383–396, June 2012. (Cited on pages 12 and 17.)

[26] C. Sawyers. Targeted Cancer Therapy. *Nature*, 432:294–297, November 2004. (Cited on page 1.)

[27] S. Shankar. Reconstruction of cellular signalling networks and analysis of their properties. *Nature Reviews, Molecular Cell Biology*, 6:99–111, February 2005. (Cited on page 22.)

[28] N. Sonia. Insulin action: Molecular basis of diabetes. *eLS*, January 2003. (Cited on pages xi and 7.)

[29] P. Stralfors. Model-based hypothesis testing of key mechanisms in initial phase of insulin signaling. *PLoS Comput. Biol.*, 4, 2008. (Cited on page 2.)

[30] P. Stralfors. Mass and information feedbacks through receptor endocytosis govern insulin signaling as revealed using a parameter-free modeling framework. *J. Biol. Chem.*, 285:20171–20179, 2010. (Cited on page 2.)

[31] P. Stralfors. Insulin signaling-mathematical modeling comes of age. *Trends in Endocrinology and Metabolism*, 23(3):107–115, March 2012. (Cited on pages 2 and 58.)

[32] Cell Signaling Technology. Phosphositeplus. http://www.phosphosite.org/homeAction.do, 2013. (Cited on pages xi, 6, 7, 9, and 23.)

[33] K. Thedieck. A dynamic network model of mTOR signaling reveals TSC-independent mTORC2 regulation. *Science Signaling*, 5 (217):1–17, March 2012. (Cited on pages 2, 13, and 42.)

[34] K. Thedieck. A modelling-experimental approach reveals insulin receptor substrate (IRS)-dependent regulation of adenosine monosphosphate-dependent kinase (AMPK) by insulin. *FEBS Journal*, 279:3314–3328, March 2012. (Cited on pages xii, 2, 35, 42, 43, and 70.)

[35] P. Thomas. PANTHER classification system. http://www.pantherdb.org/, 2013. (Cited on page 23.)

[36] G. M. Toffolo. Leucine modulates insulin signaling in a human skeletal muscle cell line and enhances insulin-dependent glycogen synthesis. *submitted*, 2013. (Cited on pages 70 and 71.)

[37] Johns Hopkins University. Online Mendelian Inheritance in Man. http://omim.org/, 2013. (Cited on page 23.)

[38] Johns Hopkins University and the Institute of Bioinformatics. Human Protein Reference Database. http://www.hprd.org/, 2013. (Cited on page 23.)

[39] Alon Uri. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC, 2007. (Cited on page 60.)

[40] Ontario Research Fund US National Institutes of Health, EU grant LSHG-CT-2005-518254 "ENFIN" and the EBI Industry Programme. Reactome. http://www.reactome.org/ReactomeGWT/entrypoint.html, 2013. (Cited on page 23.)

[41] J. D. Varner. Computational modeling and analysis of insulin induced eukaryotic translation initiation. *Mol. Biosyst.*, 7(11), November 2011. (Cited on page 2.)

[42] K. V. Venkatesh. Quantification of the effect of amino acids on an integrated mTOR and insulin signaling pathway. *Mol. Biosyst.*, 5: 1163–1173, 2009. (Cited on page 2.)

[43] F. M. White. Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes*, 55(9):2171–2179, August 2006. (Cited on pages 70 and 71.)

[44] Wikipedia. Diabetes mellitus type 2. http://en.wikipedia.org/wiki/Diabetes_mellitus_type_2, 2013. (Cited on page 1.)