Oana Alina TOMESCU

Master Thesis

Mathematical Models of Adipogenesis

Institute for Genomics and Bioinformatics,
Graz University of Technology
Petersgasse 14, 8010 Graz, Austria
Head: Univ.-Prof. Dipl.-Ing. Dr.techn. Zlatko Trajanoski

Supervisor:
Dipl.-Ing. Dr.techn. Hubert Hackl

Evaluator:
Univ.-Prof. Dipl.-Ing. Dr.techn. Zlatko Trajanoski

Graz, March 2010

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, ...
(date) (signature)

Acknowledgements

I would like to express my gratitude to my supervisor Dipl.-Ing. Dr.techn. Hubert Hackl for his assistance, support, guidance and lots of ideas. I am grateful to Univ.-Prof. Dipl.-Ing. Dr.techn. Zlatko Trajanoski for his help and for allowing me to do my master thesis at the Institute for Genomics and Bioinformatics. I am also indebted to Dipl.-Ing. Dr.techn. Michael Pfeiffer for his support and ideas concerning neural networks as well as to his colleagues Dipl. Ing Helmut Hauser and Dipl. Ing. Gerhard Neumann. In addition, I would like to thank Dipl. Ing. (FH) Maria Fischer. My thanks also go to Univ.-Prof. Dipl.-Ing. Dr. rer. nat. Wolfgang Maass.

I express my deepest gratitude to Gerhart and especially to my mother for making my study and this thesis possible.

Special thanks to Roman for his emotional support, unconditional love and constant encouragement.

Abstract

English

Obesity is nowadays one of the most severe health problems. It causes different diseases that reduce life quality. The existence of a mathematical model that could simplify the testing of new research ideas, before time and costs consuming *in vitro* and *in vivo* experiments are employed, is of major interest and importance.

Here three different models of adipogenesis are presented. One model consists of a system of ordinary differential equations and is able to reproduce the gene expression profiles of some of the well known adipogenesis regulators, inhibitors and markers like Pparg, Cebpa, Gilz, Gata2, Fabp4 and Scd1. Another model is based on an echo state network that models the interactions between the key players of adipogenesis. This model is capable of predicting the gene expression levels in response to a certain adipogenic cocktail out of the gene expression levels in response to other adipogenic cocktails. The third model is based on so called "essential genes". From these genes a subset is chosen to be employed in the parameter estimation of the model. The "impact" of each gene on the dynamic behavior of the model is computed. In this way a list of new genes, sustained by biological research, is proposed for further studies. These genes could play an important role in the process of adipogenesis.

Keywords: adipogenesis, mathematical model,*in silico*, differential equation, neural network, "essential" genes

German

Übergewicht ist ein Gesundheitsproblem welches immer mehr Menschen betrifft. Es verursacht Krankheiten die die Lebensqualität einschränken. Um auf diesem Gebiet schneller und effektiver Fortschritte zu erzielen würde ein matematisches Modell für die Adipogenese vom großen Nutzen sein.

Es werden drei Modelle erarbeitet. Das erste wird durch ein System von Differentialgleichungen dargestellt. Dieses ist in der Lage die Expressionsprofile der wichtigsten Adipogeneseregulatoren, -Inhibitoren und -Marker wie Pparg, Cebpa, Gilz, Gata2, Fabp4 and Scd1 vorherzusagen. Ein weiteres Modell beschreibt die Interaktionen zwischen den bekanntesten Adipogenesefaktoren durch ein Echo State Netzwerk. So ist es möglich die Expressionsprofile der Gene, die von einem bestimmten Adipogenesecocktail verursacht werden, vorherzusagen. Für das dritte Modell wird zuerst eine Liste von so genannten "essentiellen" Genen erstellt. Aus dieser Liste werden dann Gene bestimmt die in das Modell eingesetzt werden. Der "Einfluss" dieser Gene wird berechnet. Somit werden neue Gene vorgeschlagen die eine wichtige Rolle in der Adipogenese spielen könnten.

Stichwörter: Adipogenese, mathematisches Modell,*in silico*, Differentialgleichung, Echo State Netzwerk, "essentielle" Gene

Contents

Chapter 1

Introduction

Nowadays obesity constitutes a severe health problem. Affected people tend to develop cardiovascular diseases and non-insulin-dependent diabetes mellitus (NIDDM). Formation of adipose tissue is due to two processes. New fat cells develop and the total fat cell number increases (hyperplasia). The already existing fat cells increase their storage capacity of triglycerides (hypertrophy). A detailed understanding of the mechanisms governing these processes would be an important step in preventing obesity and the related health problems.

In vitro the process can be studied using different cell lines. The most often used are the 3T3-L1 and 3T3-FA442A lines originally created by Green and Kehinde [11, 12]. Numerous *in vitro* and *in vivo* experiments have discovered some of the transcription factors, adipogenic factors and cell-cycle factors that play a key role during adipogenesis.

The development of preadipocytes into adipocytes is called adipogenesis (see Figure 1.1). For a detailed description see [9]. *In vitro* the process is induced by exposure of the preadipocytes to a cocktail consisting of fetal bovine serum (FBS), dexamethasone (DEX), isobuthymethylxanthine (IBMX) and insulin. As a result, the cells enter the clonal expansion phase followed by terminal differentiation. During the clonal expansion of 3T3-L1 cells at least one cell-cycle is traversed. The terminal differentiation phase is marked by metabolic programs typical for mature fat cells.

A detailed review of the existing information on the key players required for a successful adipogenesis can be found in [32, 8]. The master regulators are considered to be Pparg and Cebpa. Evidence [27] shows that Pparg is capable of launching the process of adipogenesis; Pparg is necessary and sufficient. In [7] it is suggested that Cebpa induces adipogenesis through Pparg. There are evidence that the cross-regulation of Cebpa and Pparg controls the transcriptional pathway of adipogenesis [37]. Two other impor-

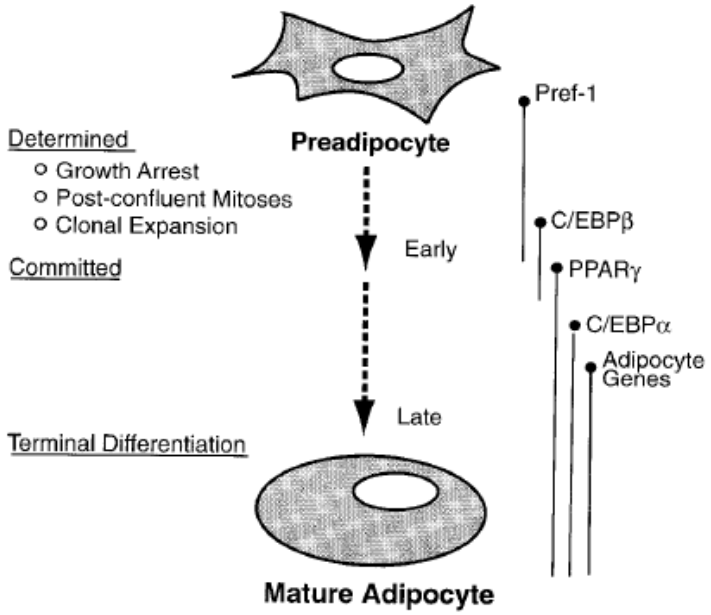


Figure 1.1: The development of a preadipocyte into a mature adipocyte. During this process the preadipocyte traverses different phases and expresses well known regulators like Cebpa and Pparg (adopted from [9]).

tant regulators are Cebpb and Cebpd. They induce the expression of Pparg and Cebpa. Other adipogenic factors are Srebp1c, Klf5, Pcreb, Krox20 and Camp. An overview of the interplay of the mentioned factors is presented in Figure 1.2. Other regulators are cell-cycle-related proteins and some negative factors that inhibit the expression of Pparg, Cebpa, Cebpb and Cebpd. In [32] the negative factors and some of the most studied positive factors that control adipogenesis are summarized into one diagram (see Figure 1.3).

In order to discover the factors that determine a preadipocyte to develop into an adipocyte, high-throughput analysis methods and proteomics can be used. One of the most efficient way to measure gene expression is microarray analysis. It was developed in the 1990s and today is one of the most often used analysis technique. One reason that led to this is the possibility of measuring thousands of genes simultaneously. Proteomics is a large-scale method for analyzing proteins. Protein levels for a few proteins can be determined by western blots.

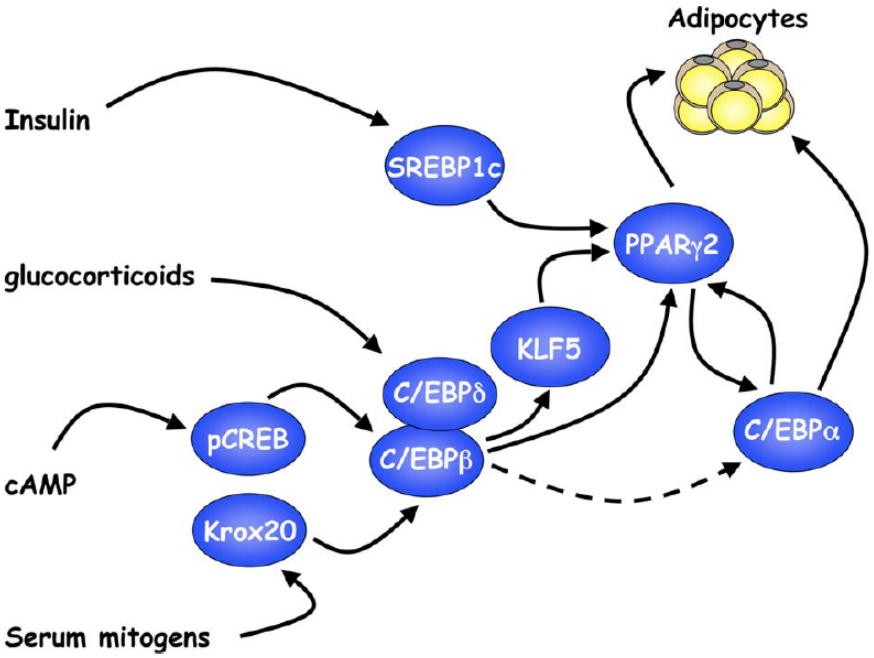


Figure 1.2: Positive transcription factors (adopted from [32]).

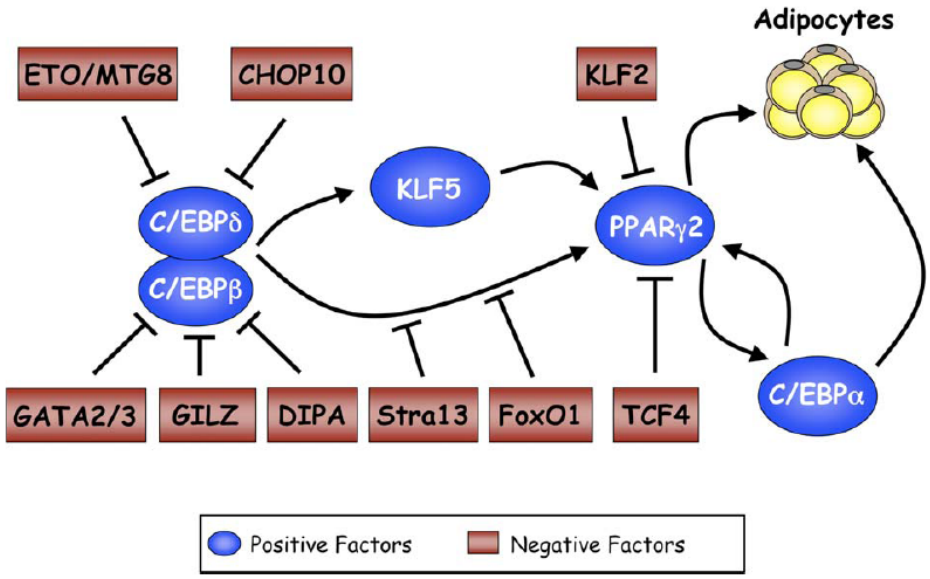


Figure 1.3: Positive and negative regulators of adipogenesis (adopted from [32]).

1.1 Objectives

Although until now a lot of experiments were conducted in order to elucidate the factors and their interactions responsible for the development of preadipocytes into adipocytes, no mathematical models of adipogenesis on molecular level are available. A model would enable *in silico* testing of different hypotheses without the costs and time needed for *in vitro* and *in vivo* testing.

Starting from the available information on adipogenesis, the key factors involved and a number of large scale experiments, three mathematical models were derived.

- One model combined the factors in Figure 1.2 and Figure 1.3. The regulation was described by a system of ordinary differential equations whose coefficients were determined from measured data.
- One model used an echo state network for the interaction of the factors. The input to the echo state network was the time course of the adipogenic cocktail. The output of the model was the time course of two proteins known as markers for adipocytes.
- One model was based on "essential" genes. The interactions between them were specified by differential equations. The parameters of the model were determined from measured data.

Chapter 2

Methods

In the first section of this chapter the concept of the differential equation model is described. The second section shows how an echo state network was used to model the process of adipogenesis. In the last section of this chapter the third model is presented.

2.1 Ordinary differential equation model

Many of the factors involved in the process of adipogenesis are known. Some of them are considered to play a key role because their absence would lead to a stop in the development of preadipocytes into adipocytes. Current knowledge is gathered in the review article [32] and presented in Figure 1.2 and Figure 1.3. For this model the shown factors are extended by two adipocyte markers. Figure 2.1 presents these factors and their dependencies.

2.1.1 Gene expression data

Raw data (.cel files) from an adipocyte differentiation microarray experiment [1, 18] were downloaded from Gene Expression Omnibus (GEO) (GSE6794). 3T3-L1 fibroblasts were cultured *in vitro* and induced to differentiate using standard DMI protocol. At successive time-points (PC, 0h, 6h, 12h, 24h, 48h, 3d, 4d, 7d, 28d) cells were collected, and processed for microarray analysis using Affymetrix Murine 11k A and B arrays. Data was normalized using gcrma (R/Bioconductor) and results from the Mu11kA array were combined with results for complementary probesets of Mu11kB arrays. Relative gene expression levels (log2ratios) from each time point versus the pre-confluent state were determined and averaged over 3 biological replicates.

Since between gene expression levels on day 7 and day 28 of differentiation

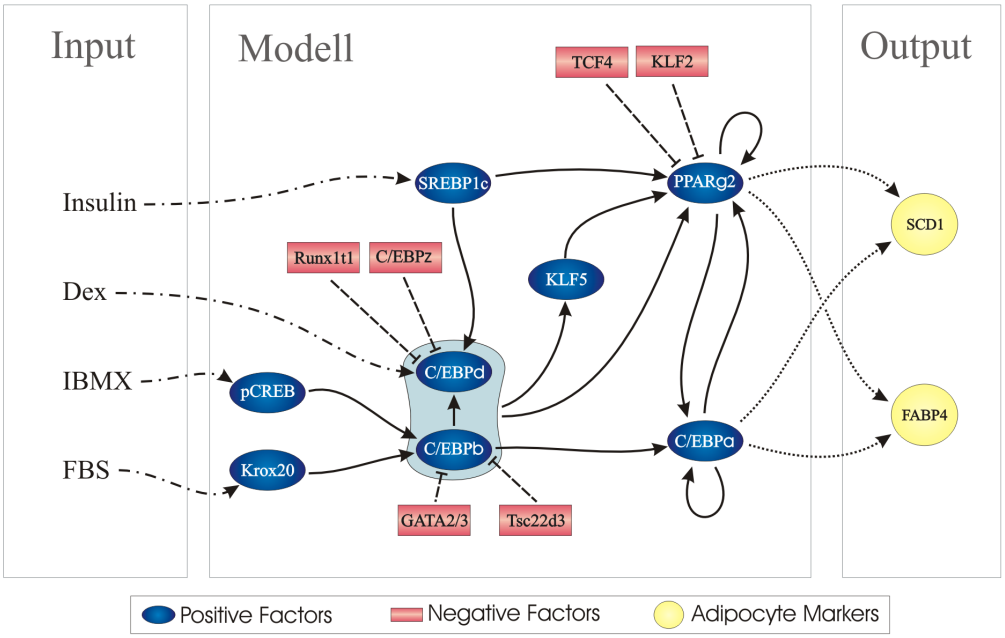


Figure 2.1: Interplay of the key regulators of adipogenesis and adipocyte markers. The regulators are divided into positive factors (blue ellipses) and negative factors (red rectangulars). The yellow circles stand for the adipocyte markers. The influence of the input cocktail is marked by chain dotted arrows. Positive regulation is denoted by regular arrows and the negative regulation by dashed arrows. The markers are determined by the master regulators Cebpa and Pparg (dotted arrows).

no major differences exist, the last time point of the measurements was not considered. In order to obtain a continuous signal the measured data was interpolated using cubic splines with a resolution of one minute. The data processed in this way was used in the differential equation model.

2.1.2 Model description

The model was described by a system of ordinary differential equations [5, 4]. Each factor that is known to exhibit positive regulation was described through an ordinary differential equation (ODE). The factors that are responsible for negative regulation were present in the ODEs of the other factors, but were not be described by an ODE. Because of simplicity reasons these factors had no input.

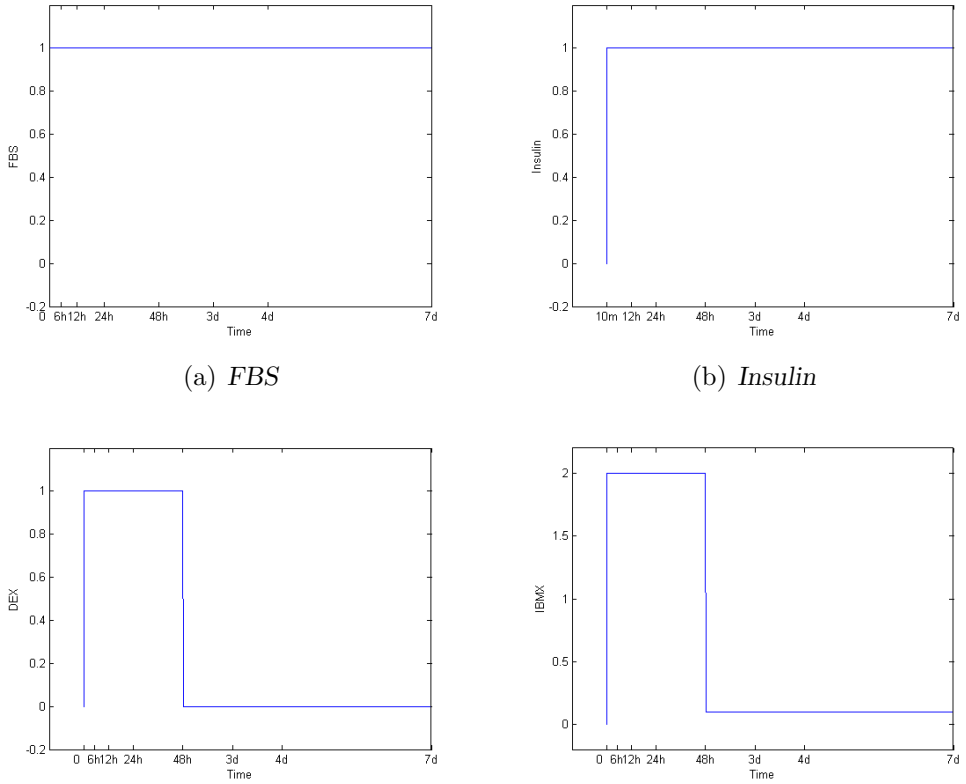
Looking at Figure 2.1, the ODE of Cebpa was written as:

$$\frac{d\ Cebpa(t)}{dt} = a_1 * Pparg(t) + a_2 * Cebpb(t) + a_3 * Cebpa(t)$$

$Cebpa(t)$, $Cebpb(t)$, $Pparg(t)$ were the gene expression levels of Cebpa, Cebpb and Pparg considered as functions of time. In this way each of the regulators was described. a_1, a_2, a_3 were the parameters that scale the influence of each of the involved genes. The system consisting of 10 ODEs and 28 parameters is shown bellow. The parameters were determined based on microarray data.

$$\begin{cases} \frac{d\ Cebpa(t)}{dt} = a_1 * Pparg(t) + a_2 * Cebpb(t) + a_3 * Cebpa(t) \\ \frac{d\ Pparg(t)}{dt} = a_4 * Srebp1c(t) + a_5 * Klf5(t) + a_6 * Cebpa(t) \\ \qquad + a_7 * (Cebpb(t) + Cebpd(t)) + a_8 * Pparg(t) \\ \qquad + a_9 * Klf2(t) + a_{10} * Tcf4(4) \\ \frac{d\ Cebpb(t)}{dt} = a_{11} * Creb(t) + a_{12} * Krox20(t) + a_{13} * Gata23(t) \\ \qquad + a_{14} * Tsc22d3(t) \\ \frac{d\ Cebpd(t)}{dt} = a_{15} * DEX(t) + a_{16} * Srebp1c(t) + a_{17} * Cebpb(t) \\ \qquad + a_{18} * Runx1t1(1) + a_{19} * Cebp\zeta \\ \frac{d\ Srebp1c(t)}{dt} = a_{20} * Insulin \\ \frac{d\ Klf5(t)}{dt} = a_{21} * Cebpb(t) + a_{22} * Cebpd \\ \frac{d\ Krox20(t)}{dt} = a_{23} * FBS(t) \\ \frac{d\ Creb(t)}{dt} = a_{24} * IBMX(t) \\ \frac{d\ Fabp4(t)}{dt} = a_{25} * Pparg(t) + a_{26} * Cebpa(t) \\ \frac{d\ Scd1(t)}{dt} = a_{27} * Pparg(t) + a_{28} * Cebpa(t) \end{cases}$$

The input to the system consisted of the time course of the adipogenic cocktail (see Figure 2.2). The concentration of each substance of the cocktail that was used in the model can not be compared to the real concentration



(a) *FBS* (b) *Insulin*

(c) *DEX* (d) *IBMX*

Figure 2.2: Standard adipogenic cocktail.

in *in vitro* experiments. However, the values used in the model proved to be useful during parameter estimation. The time courses of the gene expression of Fabp4 and Scd1 constituted the output of the model. In this case these two proteins were considered as markers for mature adipocytes.

2.1.3 Implementation

The parameters of the model were determined so that the gene expression level of the proteins matched the measured data. This task was done using the SBToolbox2. This is a toolbox for systems biology that was developed for Matlab (Mathworks Inc., Natick, USA).

The SBToolbox2 [15] provides functionality for modeling, simulation and analysis of biochemical systems. This toolbox requires Matlab version 7.1 or higher and the SBPD project for the parameter estimation functionality.

The system of ODEs was transfered to Matlab and the parameter esti-

mation tool was used. Since for the estimation of the 29 parameters only one microarray experiment was available, the estimation was limited. The computed values were used as a starting point. The system was simulated with these values and the resulted expression levels of the genes were compared to the measurements. The SBToolbox2 allows the use of events. Events define changes in the parameters or ODEs of the system. With the help of events, the values of the parameters were changed at certain time points during simulation so that the measured data could be reproduced by the model.

2.2 Echo state network model

In the previous section, the interaction between the regulators of adipogenesis was modeled using ordinary differential equations. In this section the interactions were modeled using an echo state network [14].

2.2.1 Principles of echo state networks

Neural networks [25, 36] can be divided into two types. There are feedforward networks and recurrent networks. The typical structure of these two types of networks is presented in Figure 2.3. As shown, the major difference is the "direction" of information processing. In the feedforward neural network the activation travels from the input through the hidden units (if present) and then reaches the output. In the recurrent neural network there is at least on cyclic path that the activation passes through. An echo state network is a special recurrent neural network in which only certain connections are trained while others remain fixed.

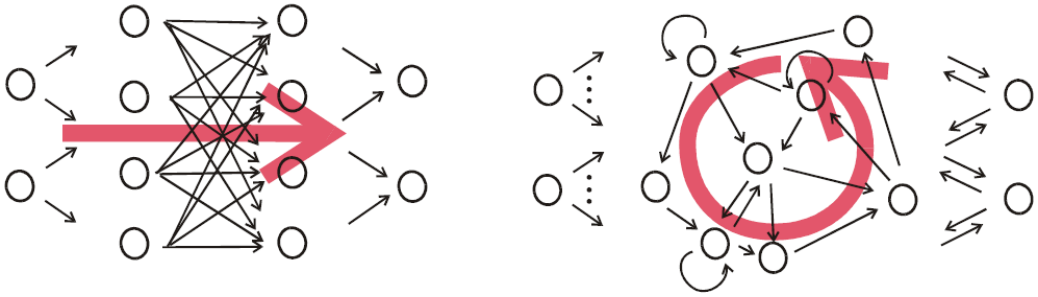


Figure 2.3: Typical structure of a feedforward network and a recurrent network (adopted from [14]). The direction of the information processing is marked with red arrows.

	DMI	DM	DI	MI	D	M	I
Cebpa	100%	130%	0%	0%	0%	0%	0%
Fabp4	100%	80%	90%	0%	90%	0%	0%

Table 2.1: Western bolts derived protein levels in response to different input cocktails (derived from Figure 2.4).

2.2.2 Gene expression and protein level data

For this model the same gene expression data (see subsection 2.1.1) as for the ODE model was used, as well as protein levels of differentiating 3T3-L1 preadipocytes presented in [16]. The measurements of the protein levels were done by western blots on day five of adipogenesis. The results are shown in Figure 2.4. This experiment was used in an unconventional way.

The protein levels in response to DMI were considered to be 100%. In addition, these protein levels were considered to be equivalent to the corresponding gene expression measured on day five of the microarray experiment. The protein levels measured in response to the other adipogenic cocktails (Table 3.1) were determined as percentage of the response to DMI. Table 2.1 was derived from Figure 2.4. 0% percent was used when the protein levels in response to the adipogenic cocktails were equal to the NCS case. The measured gene expression levels in response to DMI were multiplied with the percentage values from Table 2.1 to obtain the corresponding gene expression levels in response to the other adipogenic cocktails.

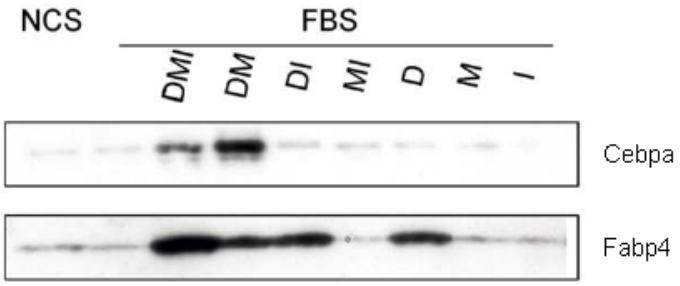


Figure 2.4: Protein levels of Cebpa and Fabp4 determined by western blots (adopted from [16]).

2.2.3 Model description

The echo state network (ESN) used in this approach was shown in Figure 2.5. The choice of an ESN for this task was based on the fact that all biological networks are recurrent networks.

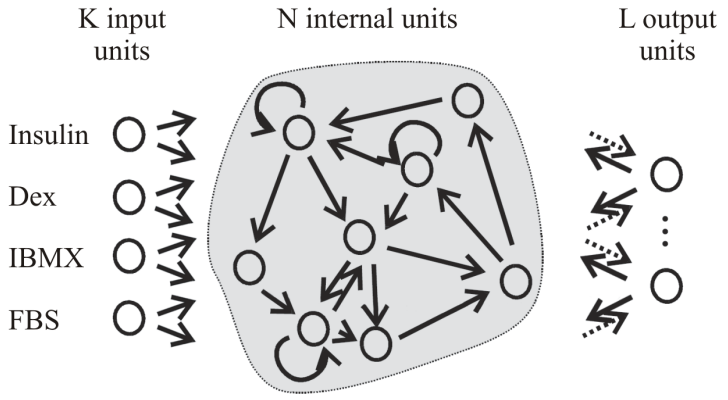


Figure 2.5: Structure of the echo state network modeling the differentiation process (adopted from [14]). The dotted arrows stand for connections that are trained. All other connections are fixed. The internal units (on the gray background) form the ***dynamic reservoir*** *of the echo state network.*

The most important mathematical characteristics [14] of the ESN model are summarized here. The input to the ESN network was referred to as $\mathbf{u}(n)$. The vector describing the state of the internal units was called $\mathbf{x}(n)$. $\mathbf{y}(n)$ was the output of the network. n denoted the time points of the simulation. The ESN model possessed K input units, N internal units and L output units. The input was connected to the dynamic reservoir through the matrix of weights $\mathbf{W}^{in} \in \mathbb{R}^{\mathbb{N}\times\mathbb{K}}$. The weights of the dynamic reservoir were gathered in the matrix $\mathbf{W} \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$. The matrix $\mathbf{W}^{out} \in \mathbb{R}^{\mathbb{L}\times(\mathbb{K}+\mathbb{N})}$ contained the weights from the dynamic reservoir to the output units. Since there also existed connections from the output units to the dynamic reservoir there was also the matrix of these weights $\mathbf{W}^{back} \in \mathbb{R}^{\mathbb{N}\times\mathbb{L}}$.

The activation of the internal units was computed using the following equation:

$$\mathbf{x}(n+1) = \mathbf{f}(W^{in}\mathbf{u}(n+1) + \mathbf{W}\mathbf{u}(n) + \mathbf{W}^{back}\mathbf{y}(n)), \tag{2.1}$$

where $\mathbf{f}$ was the component wise transfer function of the internal units, $\mathbf{u}(n+1)$ was the external input. The output of the ESN was computed with the following equation

$$\mathbf{y}(n+1) = \mathbf{f}^{out}(\mathbf{W}^{out}(\mathbf{u}(n+1), \mathbf{x}(n+1), \mathbf{y}(n))), \tag{2.2}$$

where $(\mathbf{u}(n+1), \mathbf{x}(n+1), \mathbf{y}(n))$ denoted a vector resulted from the concatenation of the input vector, internal state vector and output vector. $\mathbf{f}^{out}$ was the transfer function of the output units.

The model had four input units, each one receiving the time course of an ingredient of the adipogenic cocktail. Through testing, the following parameters of the ESN model proved to be the best choice:

- the dynamic reservoir had 30 units,
- 20% of the reservoir units received the input signals,
- the inputs were scaled with 0.9,
- the transfer function of the reservoir units was tanh,
- 10% of the reservoir received feedback from the output,
- the feedback from the output was scaled with 0.1,
- the output units used a linear transfer function,
- the output weights were learned using Bayesian regularization in a Levenberg-Marquardt algorithm.

2.2.4 Implementation

The ESN model was implemented in Matlab. The mechanisms of the ESN were implemented in concordance with the formulas presented earlier. The Bayesian regularization method used for the computation of the output weights was provided by Matlab. When leave-one-out cross-validation was used, the training presented some subtleties. Since six different input/output experiments were used for training, it was important that between the different experiments, the ESN had enough time to prepare for the next experiment. The solution was to feed the network with noise between the different experiments.

In order to test the model, two scenarios were created. In the first scenario the ESN model was simulated using DMI as input cocktail. In this case the time course of the genes expression levels of the adipogenesis regulators constituted the output of the network (16 output units). The same regulators as in the ODE model were used. The ESN model was tested on the data derived from the western blots.

In the second scenario, the ESN was used in a slightly different manner. The concept of leave-one-out cross-validation was applied. For this, the

I/O cases	Train	Test
Case I	DMI, DM, DI, MI, D, M	I
Case M	DMI, DM, DI, MI, D, I	M
Case D	DMI, DM, DI, MI, M, I	D
Case MI	DMI, DM, DI, D, M, I	MI
Case DI	DMI, DM, MI, D, M, I	DI
Case DM	DMI, MI, DI, D, M, I	DM
Case DMI	DM, MI, DI, D, M, I	DMI

Table 2.2: ESN model - input/output cases. Here the exact train and test data is presented. The name of the cases corresponds to the adipogenic cocktail that was used for testing.

microarray analysis data and the western blot data were used. Out of the measured protein levels, proportional gene expression levels were computed. In this way seven different input/output cases were available. These were used in the following way. Six of them were employed in the training phase and the seventh was used in the testing phase. This was done for each of the resulted input/output cases during seven simulations of the model. During each simulation a new set of parameters, that corresponded to a new ESN model, were computed. The exact training and testing data were presented in Table 2.2. The ESN model had 2 output units because during western blotting only Cebpa and Fabp4 were measured. The Matlab code for this scenario is provided in the Appendix.

2.3 "Essential" genes model

In the previous two sections, the process of adipogenesis was modeled using two different models. In the beginning, a system of ordinary differential equations was used. Afterwards an echo state network modeled the interactions between the regulators of the process. In both approaches, the used regulators are well known key players of adipogenesis (see [32] and [8]). The third way of describing adipogenesis is not restricted to the well known regulators. First, a set of "essential" genes was determined. Then, these genes were used to model adipogenesis. For this model the name essential genes model (EG model) will be used.

2.3.1 ”Essential“ genes

Out of different experiments a set of new genes with common characteristics were identified. In [16] a list of IBMX regulated genes was determined. For this, 3T3L1 cells were treated with DMI or DI. Out of this data, the IBMX regulated genes were identified. A list of DEX-regulated genes is presented in [6]. In order to identify (see Figure 2.6) the ”essential“ genes, first a list of IBMX and DEX regulated genes was set up. These genes were than compared to results form three different microarray experiments. A gene is considered to be an ”essential“ gene if in at least one of these experiments a log fold change (two fold change) of at least one was measured.

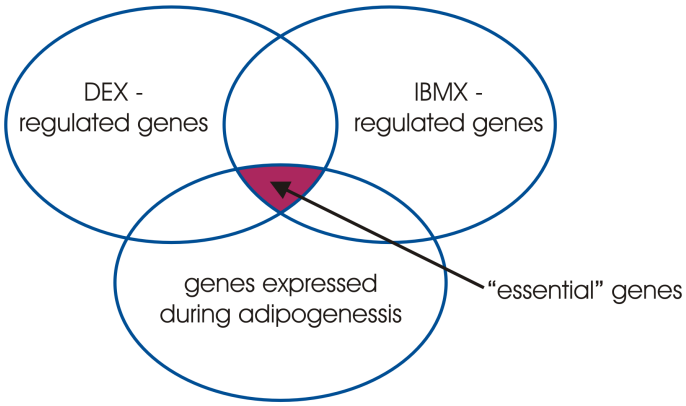


Figure 2.6: Selection of the ”essential“ genes.

Possible connections between the identified genes were determined using STRING 8 [22] from different sources like experimental repositories, computational prediction methods and public text collections. In this way, an association network was generated. The genes that presented at least four connections were chosen for modeling. These could correspond to hubs. Hubs tend to be more important than the other nodes of a network. Hubs also have the tendency to be far away from each other and often present no direct link. The association network can be visualized using Cytoscape [26]. Cytoscape is an open source software project that allows the integration and visualization of biomolecular interaction networks with expression data in a unified framework.

The time courses of the gene expression levels were computed by cubic interpolation. The used data included microarray data used in the ODE and ESN model [1]. 3T3L1 and MEF cells were measured during induced adipogenesis. The results from the 3T3-L1 [13] and MEF cells are available

under ArrayExpress E-MARS-2 and ArrayExpress E-MARS-13. The heat map of the gene expression levels was plotted using Genesis [3, 2].

2.3.2 Model description

Inspired by the model described in [10, 30] for keratinocyte migration, a similar model for adipogenesis was formulated. The "essential" genes were employed in a neural network model. The dynamics of the model were defined as a system of differential equations. Such a differential equation is presented by equation 2.3.

$$\frac{dg_i(t)}{dt} = \frac{1}{\tau_i}\left(g_i(t-1) + \sum_{j=1}^{N} w_{ij} * A(g_j(t-1) - \delta_j)\right) + I_i * \exp(k_i * t) \tag{2.3}$$

The functions and variables used in 2.3 had different meanings. The change in the gene expression level of gene g_i depended on the other N genes. The time constant of the change was τ_i. The activation function of the genes was called A and was defined as $A(x) = \frac{1}{1+\exp(-x)}$. δ_j was an offset term. The term $I_i * \exp(k_i * t)$ accounted for the external input to each gene. The weights w_{ij} described the strength of the interaction between gene g_i and gene g_j.

For each of the N genes g_i the following parameter were determined:

- 1 time constant τ_i,
- N weights w_{ij} connecting gene g_i to the other N genes,
- 1 offset term δ_i,
- 1 amplitude I_i of external input,
- 1 argument k_i of the external input.

For this model N*(N+4) parameters were determined.

2.3.3 Implementation

The identification of the "essential" genes was implemented using the programming languages C and C++. The parameters of the model were determined using Matlab and the provided Genetic Algorithm Toolbox. The parameters were determined so that the mean squared error between the

measurements and the simulation was minimized. The mean squared error (MSE) between the measurements M and the simulation S for N genes and T time points was defined as:

$$MSE = \frac{1}{TN} \sum_{i=1}^{N} \sum_{t=1}^{T} (M(i,t) - S(i,t))^2 \tag{2.4}$$

The parameters of the model were determined by a genetic algorithm [23, 24, 28] employing 5000 generations, each with a population of N*(N+4)*10 individuals.

Chapter 3

Results

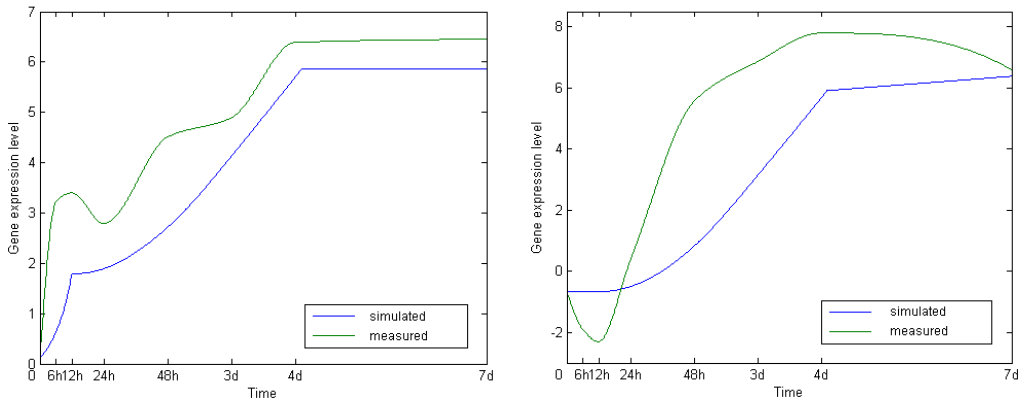
In the first subsection the results of the ODE model are shown. The results of the second model follow in the second subsection. The thirst subsection is dedicated to the results obtained from the third model. These results are discussed in the next chapter.

3.1 ODE model can compute gene expression profiles of well known regulators of adipogenesis

The model described in subsection 2.1.2 on page 6 was simulated using the SBToolbox2. The computed time courses of the employed regulators matched the measured data. Gene expression levels of the two adipocyte markers (Figure 3.1) were shown. The gene expression levels of Pparg and Cebpa (Figure 6.1) which are considered to be the master regulators of adipogenesis were presented in the Appendix. These results were obtained if the model's input was the well-known adipogenic cocktail (Insulin, DEX, IBMX and FBS). The time courses of the other regulators exhibited the same similarity to the measured data.

The model was also simulated for other different inputs. Three simulations were performed in which either Insulin, DEX or IBMX was not used. In other three simulations FBS together with either Insulin, DEX or IBMX formed the input. These six different input cocktails together with the initial input cocktail were described in Table 3.1.

The simulation results for the different input cocktails were shown in Figure 3.2 and Figure 6.2. Figure 6.2 can be found in the Appendix. Only the gene expression levels of Cebpa were plotted.



(a) *Scd1* (b) *Fabp4*

Figure 3.1: ODE model - gene expression levels of adipocyte markers.

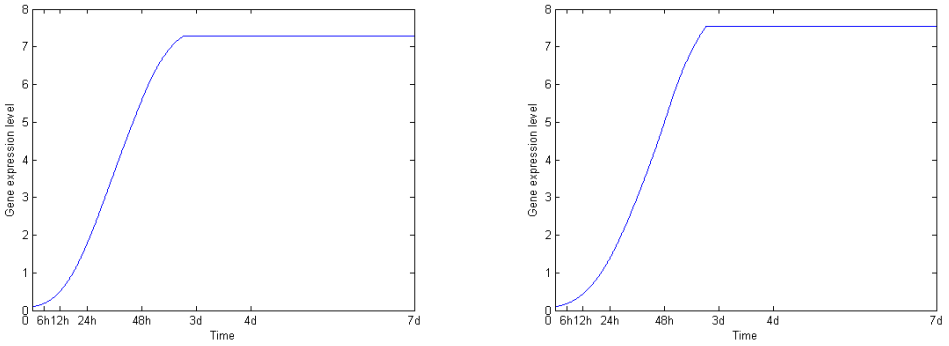
Input	FBS	Insulin (I)	DEX (D)	IBMX (M)
DMI	x	x	x	x
DM	x		x	x
DI	x	x	x	
MI	x	x		x
D	x		x	
M	x			x
I	x	x		

Table 3.1: Ingredients of the different input cocktails.

3.2 ESN model predicts higher gene expression of Cebpa in the absence of Insulin

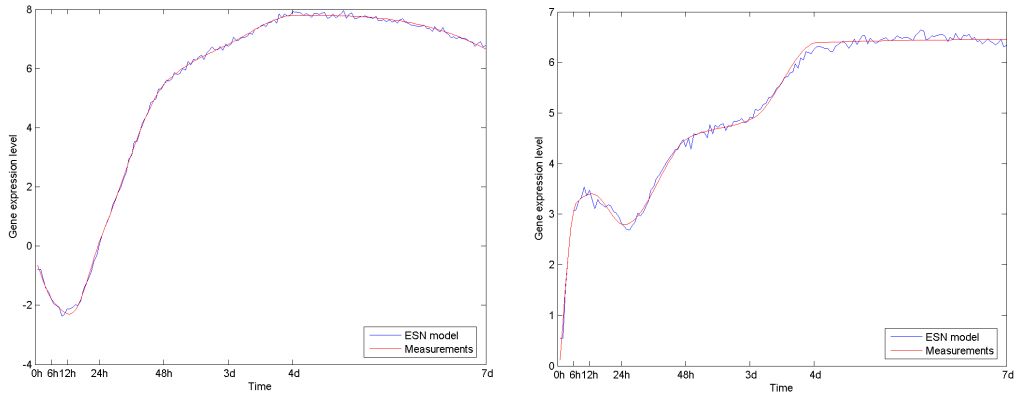
The ESN model described in section 2.2 on page 9 was employed in two scenarios. First the model was trained using DMI as input cocktail. The output units computed the time courses of the gene expression levels of the regulators of adipogenesis. The gene expression levels simulated during the training phase, together with the measured gene expression levels, were shown in Figure 3.3. Here only the gene expression levels of the 2 markers were shown. The quality of the simulated gene expression levels of the other regulators was similar.

Then the network was tested using different adipogenic cocktails (Table 3.1). Figure 3.4 showed the simulated gene expression levels for the case in which DM was used as input cocktail.



(a) *Input cocktail DM* (b) *Input cocktail DI*

Figure 3.2: ODE model - gene expression level of Cebpa for different input cocktails.



(a) *Fabp4* (b) *Scd1*

Figure 3.3: ESN model - simulated and measured gene expression levels of Fabp4 and Scd1 during the training phase.

3.3 ESN model is able to predict correct gene expression levels of Cebpa and Fabp4 in response to Insulin

In the second scenario the concept of leave-one-out cross-validation was used. The results in response to I were shown in Figure 3.5. The results in response to DMI and D were presented in Figure 6.3 and 6.4 in the Appendix.

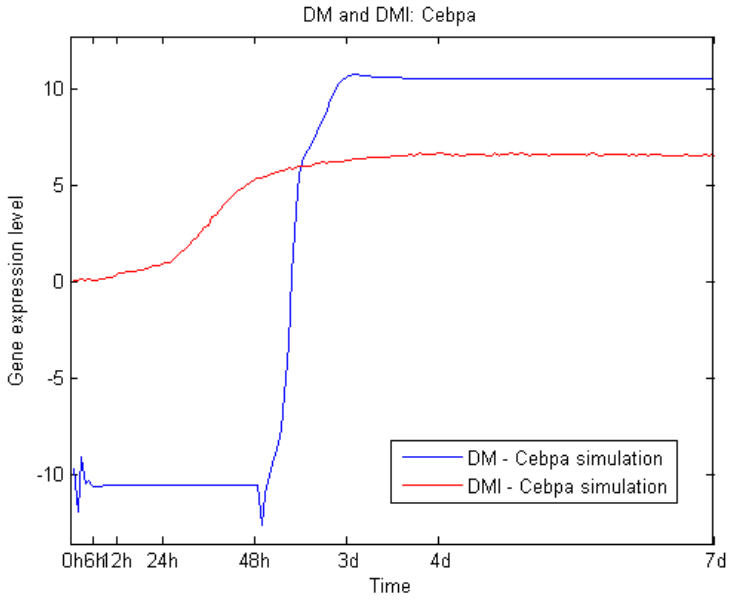


Figure 3.4: ESN model - case DM: Time courses of Cebpa.

	Bdnf	Bhlhb	Cd36	Cebpa	Egr1	Id2	Myc	Pparg
τ	-2,611	-1,668	-12,67	-4,410	-0,791	-1,384	-1,151	-3,316
δ	1,554	1,373	1,664	1,916	0,037	-1,588	1,261	2,132
I	1,229	0,556	0,104	0,507	1,033	0,438	1,851	0,701
k	-0,035	-0,011	-0,005	0,001	0,003	0.006	-0,001	-0,003

Table 3.2: EG model - values of the parameters tau, δ, I and k.

3.4 EG model proposes a list of new genes that could play important roles during adipogenesis

This section is dedicated to the results obtained by the EG model described in section 2.3 on page 13. First a set of "essential" genes was determined. These genes were shown in Figure 3.11. The heat map of the genes was presented in Figure 3.6. Out of these genes a subset was chosen for modeling. The selected genes were shown in red circles in Figure 3.11.

These selected genes were connected by parameters in a system of differential equations. The heat map of the weights was shown in Figure 3.7. The other parameters were presented in Table 3.2.

Based on the computed weights of the EG model a diagram of the network

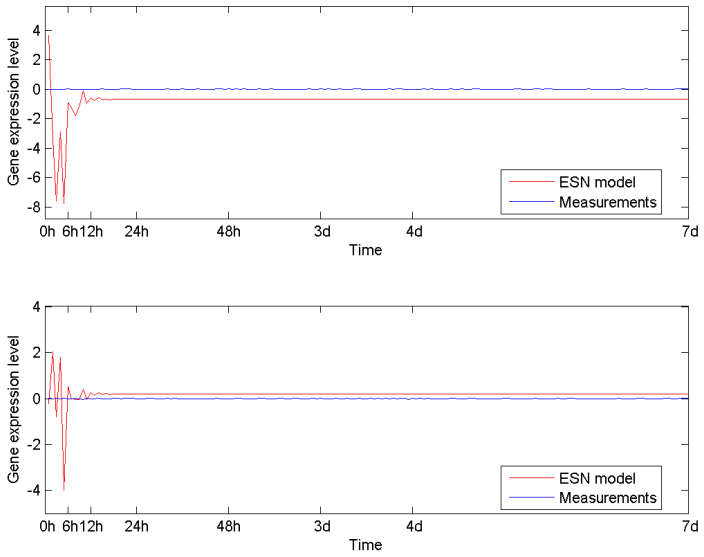


Figure 3.5: ESN model - case I using leave-one-out cross-validation: gene expression level of Cebpa (upper plot) and Fabp4 (lower plot) during testing phase.

was drawn. Only weights having an absolute value higher than four were considered. The strength of the drawn connections was proportional to the weights. The sign of the weights was shown using different types of arrows. The resulting network can be seen in Figure 3.8.

The parameters were computed so that the mean square error between the simulated and measured gene expression levels was minimized. With the determined parameters, the mean squared error is equal to 0.15. The simulated and measured gene expression profiles are presented in Figure 3.9.

After the parameters of the model were determined, the "impact" of each gene on the dynamic behavior of the model was determined. The impact was measured as the mean squared change in the gene expression levels. The results are shown in Figure 3.10.

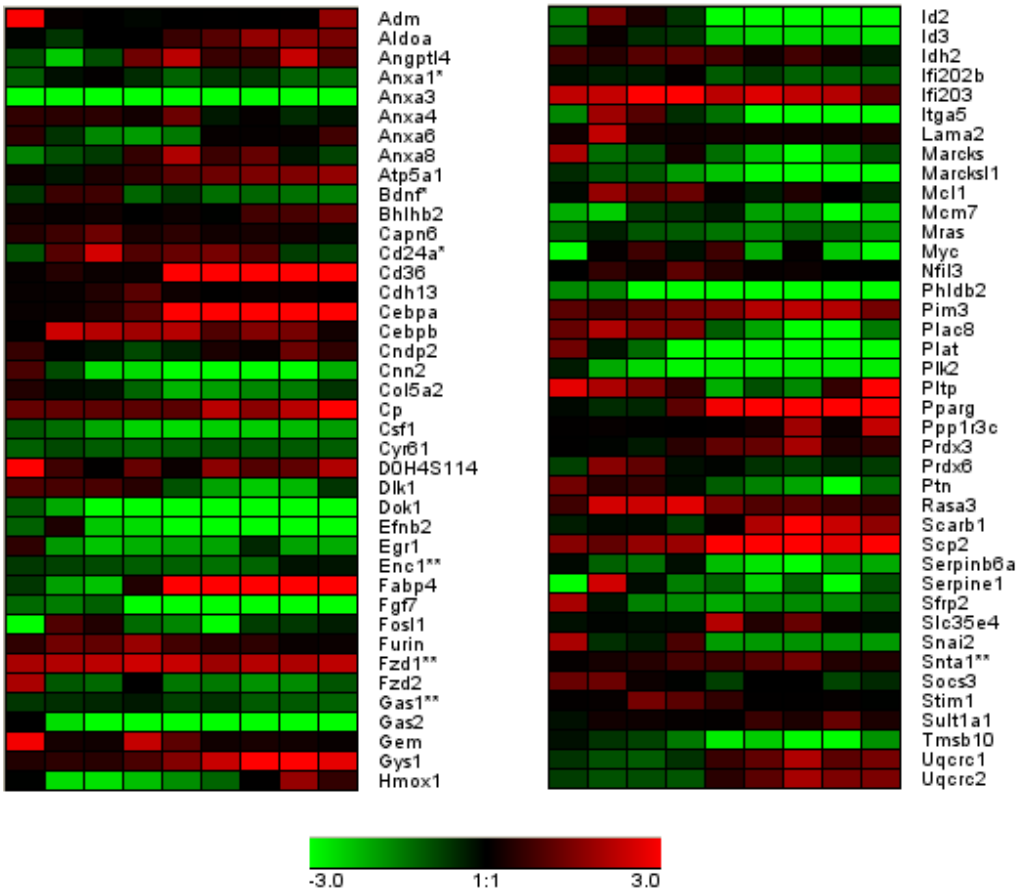


Figure 3.6: EG model - heat map of the 80 identified essential genes.

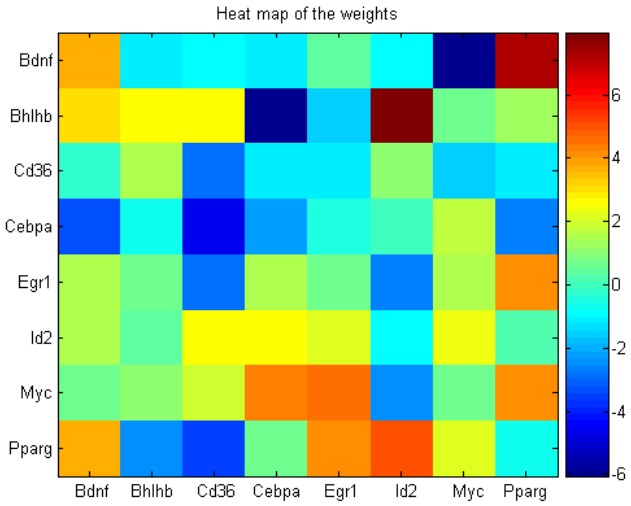


Figure 3.7: EG model - heat map of the estimated weights.

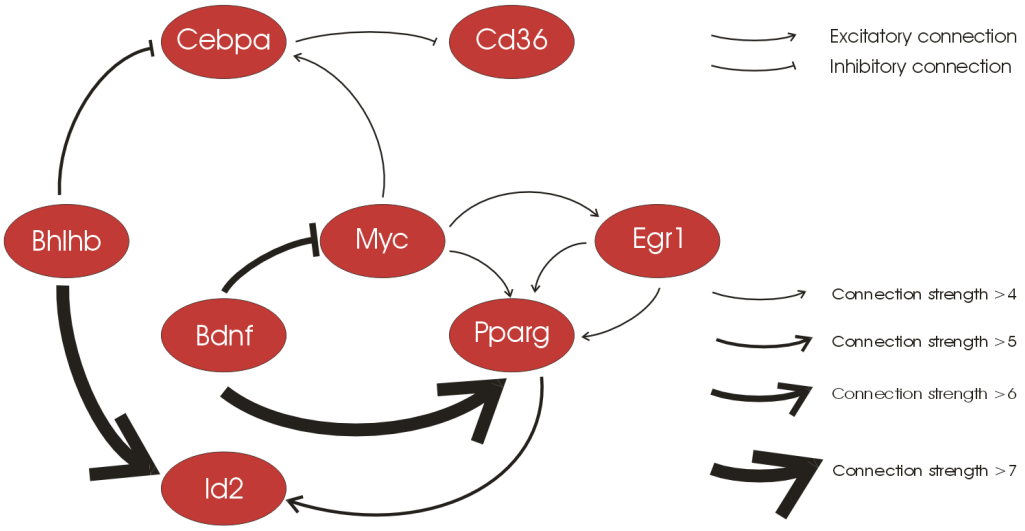
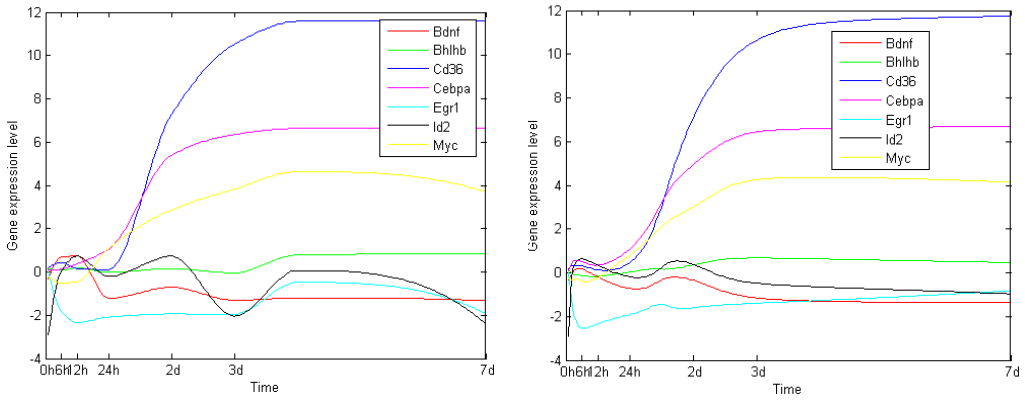


Figure 3.8: EG model - network determined from the essential genes and the estimated weights.



(a) *Measurements* (b) *Simulation*

Figure 3.9: EG model - measured and simulated gene expression levels.

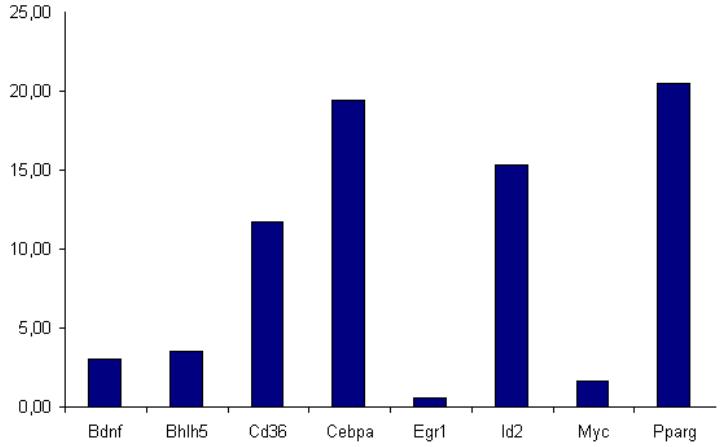


Figure 3.10: EG model - impact of the essential genes measured as the resulted mean squared change in the gene expression profiles.

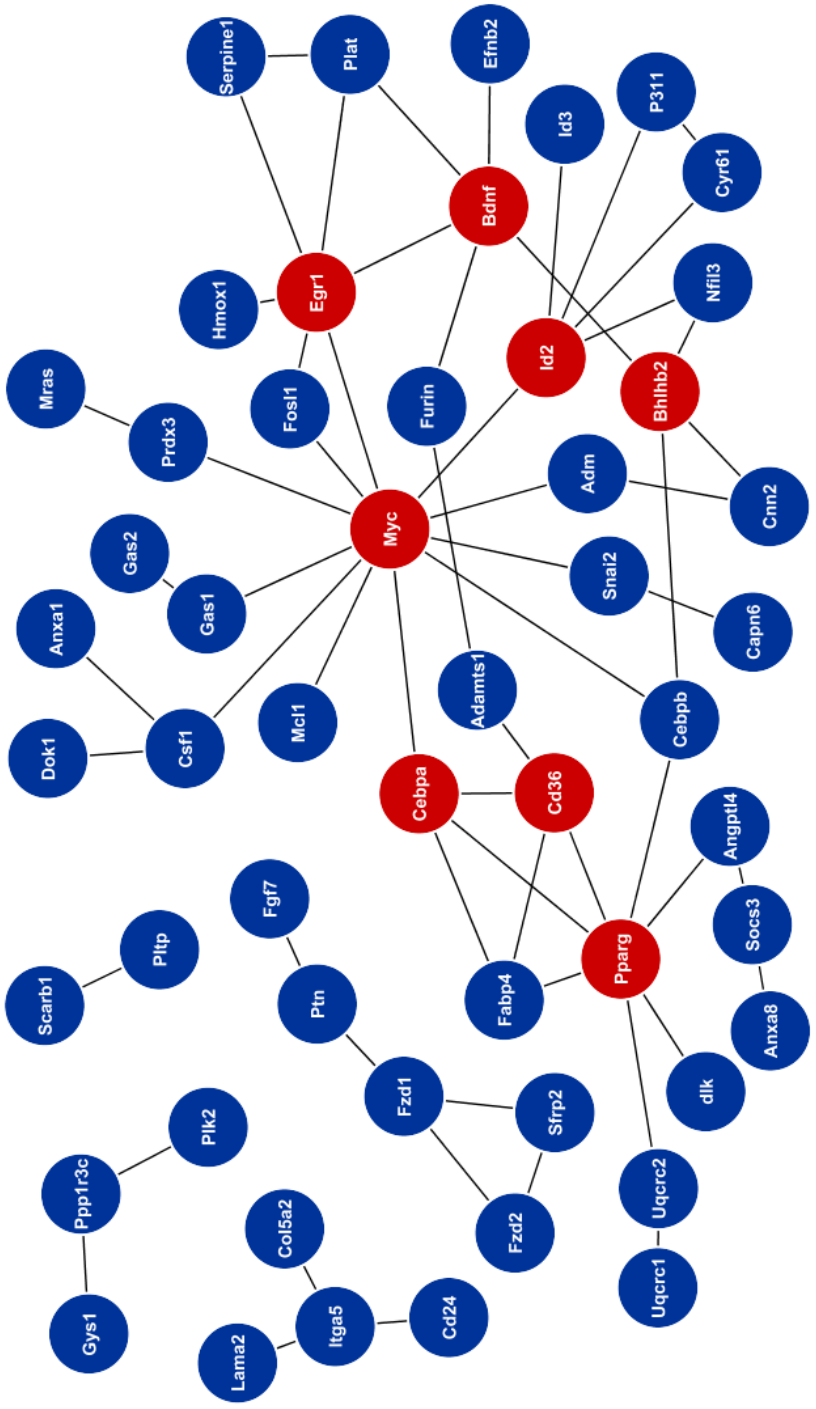


Figure 3.11: EG model - association network of the identified essential genes. The genes selected for modeling are shown in red circles.

Chapter 4

Discussion

4.1 Ordinary differential equation model

Three mathematical models of adipogenesis were developed. For each model miscellaneous approaches were used leading to a characteristic prediction behavior.

The ODE model was simulated using the standard adipogenic cocktail DMI and the computed approximation of the measured data was good. The simulated gene expressions did not perfectly match the measured data, but it was obvious that they shared some characteristics. One of them was that the gene expression reached a saturated value that did not drastically change after day four of adipogenesis. In general the forms of the curves were similar to each other.

The gene expressions of Scd1 and Fabp4 correlated with the biological understanding of adipocyte markers. Starting with about day four of differentiation, the markers exhibited very high gene expression levels.

The simulated gene expressions of the master regulators Pparg and Cebpa and adipogenesis markers Scd1 and Fabp4 matched the measured data. The simulation results for the other factors were also good, but these were not shown for clarity reasons.

The network was also simulated using different input cocktails (DM, DI, MI, D, M, I). Cebpa was chosen for comparison because the protein levels of it were determined in [16] using the same input cocktails. It is clear that the ODE model computes gene expression levels and not protein levels. But here a general comparison is done, not a comparison of the exact numbers. The gene expression levels on day five of differentiation were considered for comparison, since western blots were done on day five.

The computed gene expression profiles did not match the results of the

western blots. The model predicted high gene expression levels for all the used input cocktails. The western blots experiments showed that the protein levels of Cebpa were high for DMI and DM, but low for all other input cocktails (no difference compared to the control case). In case of DM, the simulation results matched the western blots results (high gene expression levels and high protein levels). In all other cases the simulation results did not match the western blots results.

The response of Cebpa to different input cocktails was almost equal. Only the gene expression levels in the end of the differentiation process were, in some cases, different (see y-axis of the plots). This revealed that the model did not show a good generalization capacity. Only on one input cocktail, different to the one used to determine the model's parameters, the model computed a result similar to results obtained through other techniques. This is probably due to the fact that the parameters of the model were computed from only one experiment. These parameters are capable of reproducing the data they were computed from. If more measurements were available that could be used for parameter estimation, the generalization capacity of the model would increase.

4.2 Echo state network model

The ESN model was trained on the gene expression data obtained during the microarray experiment. All adipogenesis regulators known from the ODE model were considered. In this way, gene expression levels of 14 regulators were used. The input to the ESN model was the time course of the standard adipogenic cocktail DMI. During the training phase, the gene expression levels of the 14 regulators were almost equal to the measured gene expression levels.

The network trained in this way was used to simulate the gene expression levels of the adipogenesis regulators for different input cocktails. The predicted gene expression level of Cebpa when DM was used as input cocktail did not match the measurements but was higher than in the DMI case. This fact corresponds to the current knowledge that that adipogenic cocktail DM (lacking Insulin) induces higher gene expression levels of Cebpa compared to the adipogenic cocktail DMI.

In the other cases the performance of the model was not good. It seems that the ESN model adjusted to well to the training data.

The ESN was also trained using leave-one-out crossvalidation. In this case the output of the ESN model comprised of the gene expression levels of Cebpa and Fabp4. The best results were obtained when the ESN model was

tested on the input cocktail I. This means that the parameters of the ESN model were computed from the input/output data of DMI, DM, DI, MI, D and M and that the gene expression levels of Cebpa and Fabp4 were simulated in response to the input cocktail I. In this case the exhibited generalization quality was very good. The model is suited to be used in other biologically interesting scenarios.

The ESN model exhibited also a very good generalization quality when it was trained on DMI, DM, DI, MI, M, I and tested on the adipogenic cocktail D. The simulated gene expression levels of Fabp4 were almost equal to the measured data, but this was not the case for Cebpa. In the other cases, the generalization performance was limited.

Even if only a small amount of data was available, a model was derived that was able to predict the gene expression levels of Cebpa and Fabp4 in response to an adipogenic cocktail that was never "seen" by the model.

4.3 "Essential" genes model

"Essential" genes were identified from different experiments and their expression profiles were used in the parameter estimation of the EG model. The computed mean squared error between the measurements and the simulation results was very small. Then the impact of each gene on the dynamical behavior of the model was computed. This impact could be an indicator for the importance of each gene in the process of adipogenesis. It is not a surprise that Cebpa and Pparg had the highest impact. The model confirmed once again the important role played by this two regulators. But more interesting are the other genes that also exhibited a high impact. These could be important regulators of adipogenesis.

In the computed network of the "essential" genes one of the strongest predicted connections was between Bdnf and Pparg. In [33] evidence was presented that shows sensitivity of Bdnf to overexpression of Cebpb. The connection between Cebpa and Pparg is well known. In [31] a study was presented in which gene transfer of Bdnf in mice led to weight loss and to alleviation of obesity related insulin resistance. Another article [20] shows a direct link between eating behavior and Bdnf activity.

According to the EG model another important regulator of adipogenesis was Bhlhb. In [38] and [29] it was shown that inhibiting of Pparg is possible through Bhlhb.

A connection between Pparg and Id2 is studied in [19]. Overexpression of Id2 induced expression of Pparg and mice lacking Id2 expression exhibited reduced adiposity. Furthermore the article presented results confirming the

role of Id2 in the modulation of Pparg expression.

Myc was able to inhibit adipocyte differentiation through blocking Cebpb and Cebpd caused activation of Cebpa and Pparg [34, 35]. It was shown that Egr1 and Egr2 (Krox20) have different roles in adipogenesis. While Krox20 is known to play a positive role (see ODE model and [32]), adipogenesis is inhibited by ectopic expression of Egr1 and potentiated by knockdown of Egr1 [17].

Cd36 is implicated in the transport of long chain fatty acids. It was also shown that mice in which the corresponding gene was deleted were save from high fat diet induced obesity. In [21] it was demonstrated that Cebpa, one of the key factors of adipogenesis, regulates the expression of Cd36.

The list of proposed genes is reasonable from a biological point of view.

Chapter 5

Conclusion

There exists no mathematical model of adipogenesis on molecular level and the complicated mechanisms governing this process are far from being elucidated. Here three different mathematical models were proposed.

A mathematical model provides an *in silico* test environment that is faster and cheaper than the *in vitro* or *in vivo* equivalent. In this way it is easier to conduct a preliminary examination of new research ideas.

Nevertheless a mathematical model is always an abstraction of a biological process. The simplicity of the model is opposed to the complexity of the process. A model never reflects the hole biological truth but it is able to make predictions that can be used to determine new insight to the biological process.

The more complicated a model is (the more factors it considers) the more conditions are needed for reliable parameter estimation. The performance and the generalization ability of each model increases with the number of different conditions that were considered for data measuring.

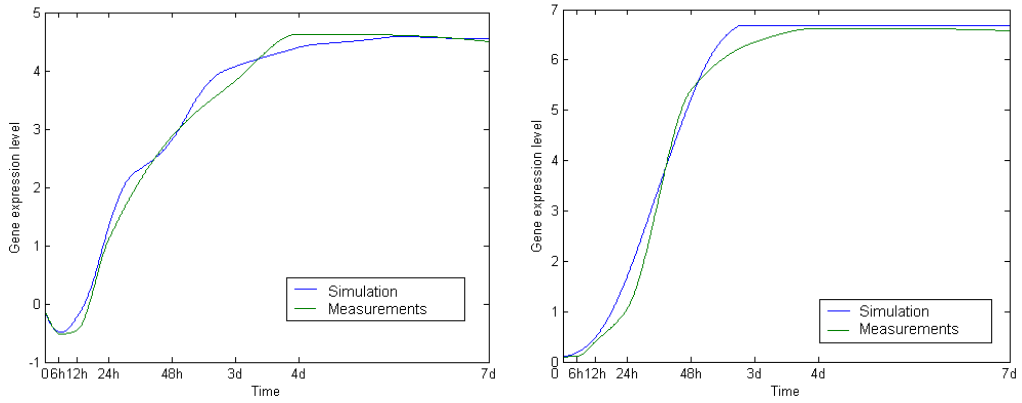
The use of a mathematical model also allows target ranking. In this way the impact on the dynamical behavior of the model of each of the considered factors can be computed.

A model offers the possibility of *in silico* testing of biological hypothesis before investing time and money in *in vitro* and *in vivo* testing. In this way testable hypothesis can be formulated and the results could be helpful in deciding over further research.

Chapter 6

Appendix

6.1 ODE Model - parameters and additional results



(a) *Pparg* (b) *Cebpa*

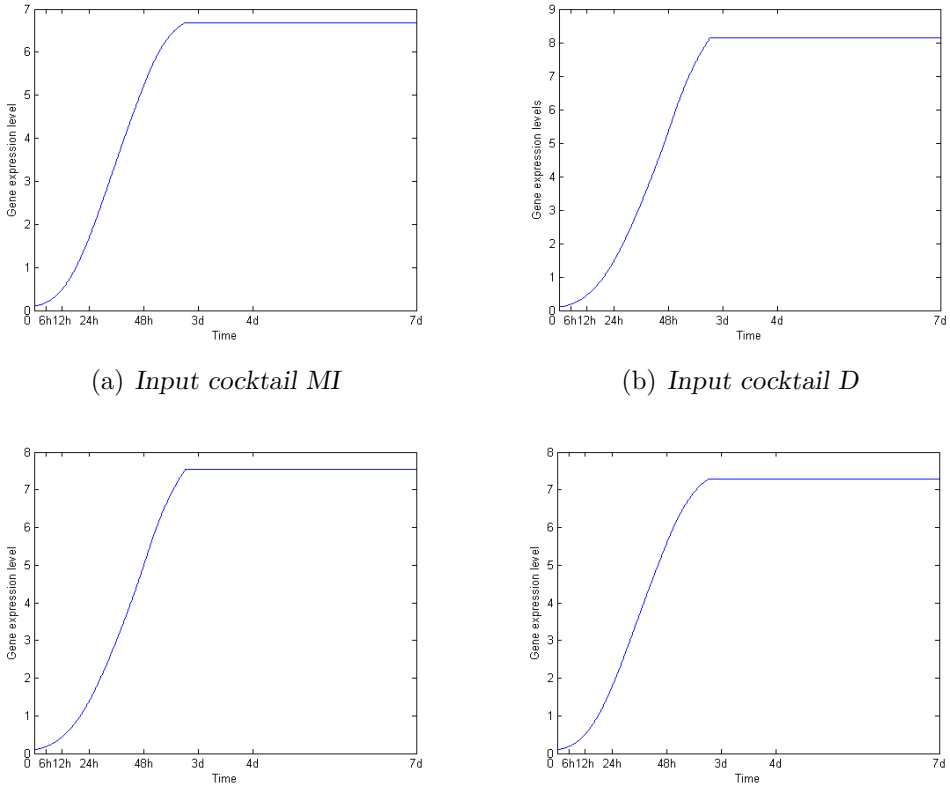
Figure 6.1: ODE model - gene expression levels of master regulators.

Events	E1	E2	E3	E4	E5	E6	E7
Simulation time	0	720	2000	3000	4000	6000	8000

Table 6.1: Events at which the ODE parameters change. The total simulation time was of 10080 which corresponds to 7 days of differentiation with a resolution of 1 minute. The corresponding parameter values were listed in Table 6.2.

	E1	E2	E3	E4	E5	E6	E7
a_1	3e-4				0		
a_2	18e-4				0		
a_3	-3e-4				0		
a_4	-54e-4						
a_5	-4e-4						
a_6	32e-4						
a_7	12e-4						
a_8	-39e-4						
a_9	63e-4						
a_{10}	-42e-4						
a_{11}	4e-3			1e-9	0		
a_{12}	1e-8			1e-5	0		
a_{13}	-6e-5				0		
a_{14}	-6e-5				0		
a_{15}	1e-9						
a_{16}	1e-9		2e-3	0			
a_{17}	15e-4		0				
a_{18}	-9e-9		-5e-3	0			
a_{19}	-16e-4		-21e-3	4e-4			
a_{20}	3e-4			-15e-5			
a_{21}	14e-4		-24e-4		-42e-4		0
a_{22}	14e-4						0
a_{23}	-15e-4	15e-4		-13e-4	16e-4	-1e-4	-15e-4
a_{24}	6e-5	-7e-5		15e-4		0	
a_{25}	1e-4					1e-5	
a_{26}	1e-5	2e-4				1e-5	
a_{27}	1e-5	1e-4			0	1e-7	
a_{28}	2e-2	1e-4				1e-7	

Table 6.2: Computed values of the ODE model parameters. The parameter values change at different events E1 to E7. The events are described in Table 6.1.



(a) *Input cocktail MI*

(b) *Input cocktail D*

(c) *Input cocktail I*

(d) *Input cocktail M*

Figure 6.2: ODE model - simulated gene expression level of Cebpa for different input cocktails.

6.2 ESN model - sample Matlab code for the implementation of the model using leave-one-out cross-validation

```
%% Echo state network model using leave one
%% out cross validation
%% Trains the model and than computes the
%% mean squared error of the model
%% during training phase
%%
close all
clear all
clc
```

```
%load data for the different input cocktails
load('DMI_LOO.mat');
load('DI_LOO.mat');
load('DM_LOO.mat');
load('MI_LOO.mat');
load('I_LOO.mat');
load('M_LOO.mat');
load('D_LOO.mat');

% set initial state of the random generator
% reproducibility of results
rand('state', 1);
randn('state',1);

% define random sequence between each
% input/ouput data

%time steps of random input
rt = 1000;
%random iput
ri = randn(size(inputD,1),rt);
%random output
ro = randn(size(outputD,1),rt);
%train time
tt = size(inputD,2);

% total input and output
input = [inputDMI ri inputDI ri inputDM ri ...
        inputMI ri inputI ri inputM ri ...
        inputD ri];
output = [outputDMI ro outputDI ro outputDM ...
         ro outputMI ro outputI ro outputM ...
         ro outputD ro];

%iterating through the 7 input/output cases
for k = 1:1:7
    %training phase

    %preparing input/output data
    [train_input train_output ...
```

```
 test_input  test_output ...
 text data_index fig] = create_in_out_LOO ...
                       (input,output,k,rt,tt);
N = 30;             %units in dynamic reservoir
L = size(train_output,1);         %output units
K = size(train_input,1);          % input units

%weights from the input to the dynamic reservoir
w_scale = 0.9;
connect = 0.1;
C = rand(N,K)<connect  ;
C=C.*(2*randint(N,K)-1);
Win = w_scale .* C;

% weights in the dynamic reservoir
W0 = randn(N,N);
[V,D] = eig(W0);
lambda_max = max(max(abs(D)));
W1 = W0./abs(lambda_max);
alpha = 0.9;
W = alpha.*W1;

% weights from the output units back to the
% dynamic reservoir
connect_back = 0.01;
C_back = rand(N,L)<connect_back  ;
C_back = C_back.*(2*randint(N,L)-1);
wb_scale = 0.1;
Wback = wb_scale*C_back;

%training with teacher forcing
net_delay = 2;
noise_scale = 0.01;
X = zeros(size(train_input,2),N);

%for each time step
for i = net_delay+1:1:size(train_input,2)
    noise = randn(1,N)*noise_scale;
    X(i,:) = tanh(Win *train_input(:,i-...
              (net_delay))+W*X(i-1,:)'+...
              +Wback*train_output(:,i - ...
```

```
                net_delay)+noise');
    end

    %computing Wout with Bayesian regulation
    startM = 1;
    m = X(data_index,:);
    a = ones(1,size(m,1))';
    M = [a m]';
    T = output(:,data_index);
    net = newff(minmax(M),size(T,1),{'purelin'}...
                ,'trainbr');
    net.trainParam.epochs = 50;
    randn('seed',11);
    net = init(net);
    [net tr] = train(net,M,T);
    Y = sim(net,M);

    %mean squared error
    mse_train_cebpa(k) = sum((Y(1,:)-...
    train_output(1,data_index)).^(2))/(length(Y(:,1)))
    mse_train_fabp4(k) = sum((Y(2,:)- ...
    train_output(2,data_index)).^(2))/(length(Y(:,2)))

    %test phase
    [mse_test_cebpa(k), mse_test_fabp4(k)] = ...
    test_ESN_LOO_reg(test_input,test_output,...
    net,X,Win,W,Wback,1000,text,fig);
end

%% Computes the mean squared error
%% of the network
%% RETURN VALUES:
%% ec = mean squared error of Cebpa
%% ef = mean squared error of Fabp4
%% INPUT PARAMETERS:
%% u = random input to the ESN
%% d = random output of the ESN;
%% net = network computing the output
%% of the ESN
%% Win = weights from the input units
%% to the dynammic reservoir
```

```
%% W = weights of the dynamic reservoir
%% Wback = weights from the output
%% units back to the dynamic reservoir
%% r = delay
%% text, fig = used for ploting

function [ec,ef] = test_ESN_LOO_reg(u,...
d,net,X,Win,W,Wback,r,text,fig)

random_input = randn(size(u,1),r);
random_output = randn(size(d,1),r);
u = [random_input u];
d = [random_output d];

for j = 2:1:size(u,2)
    Y(j-1,:) = sim(net,[1 X(j-1,:)]');
    %no teacher forcing is used anymore
    X(j,:) = tanh(Win *u(:,j-1)+W*...
    X(j-1,:)'+Wback*Y(j-1,:)');

end

day5 = 0;
cebpa_esn = Y(r+day5+1:end,1);
cebpa_meas = d(1,r+day5+1:end-1)';
fabp4_esn = Y(r+day5+1:end,2);
fabp4_meas = d(2,r+day5+1:end-1)';

ec = sum((cebpa_esn - cebpa_meas).^(2))...
         /(length(cebpa_esn));
ef = sum((fabp4_esn - fabp4_meas).^(2))...
         /(length(fabp4_esn));
```

6.3 ESN Model - additional results

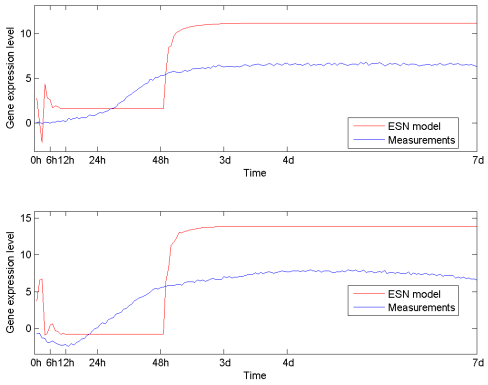


Figure 6.3: ESN model - case DMI using leave-one-out cross-validation: gene expression level of Cebpa (upper part) and Fabp4 (lower part) during testing phase.

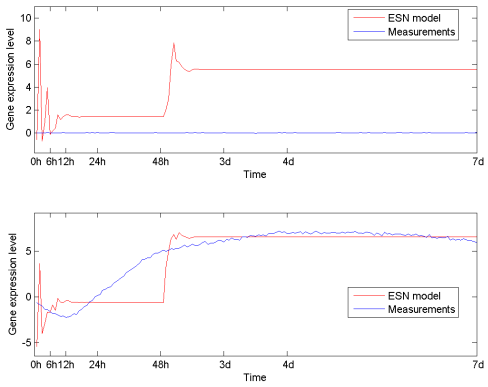


Figure 6.4: ESN model - case D using leave-one-out cross-validation: gene expression level of Cebpa (upper part) and Fabp4 (lower part) during testing phase.

List of Figures

List of Tables

Bibliography

[1] Soukas A, Socci ND, Saatkamp BD, Novelli S, and Friedman JM. Distinct transcriptional profiles of adipogenesis in vivo and in vitro. *J Biol Chem*, 276:34167–34174, 2001.

[2] Sturn A, Mlecnik B, Pieler R, Rainer J, Truskaller T, and Trajanoski Z. Client-Server environment for high-performance gene expression data analysis. *Bioinformatics*, 19:772–773, 2003.

[3] Sturn A, Quackenbush J, and Trajanoski Z. Genesis: Cluster analysis of microarray data. *Bioinformatics*, 18:207–208, 2002.

[4] Sedaghat AR, Sherman A, and Quon MJ. A mathematical model of metabolic insulin signaling pathways. *J Physiol Endocrinol Metab*, 283:E1084–E1101, 2002.

[5] Cobelli C and Carson E. *Introduction to modeling in physiology and medicine.* Academic Press, 2008.

[6] Pantoja C, Huff JT, and Yamamoto KR. Glucocorticoid Signaling Defines a Novel Commitment State during Adipogenesis *In vitro*. *MolBiol Cell*, 19:4032–4041, 2008.

[7] Rosen ED, Hsu CH, Wang X, Sakai S, Freeman MW Gonzalez FJ, and Spiegelman BM. C/EBPα induces adipogenesis through PPARγ: a unified pathway. *Genes Dev*, 16:22–26, 2002.

[8] Rosen ED and MacDougald OA. Adipocyte differentiation from the inside out. *Nat Rev Mol Cell Biol*, 7:885–897, 2006.

[9] Gregoire FM, Smas CM, and Sul HS. Understanding adipocyte differentiation. *Phys Rev*, 78:783–809, 1998.

[10] Busch H, Camacho-Truilio D, Rogon Z, Breuhahn K, Angel P, Eilis R, and Szabowski A. Gene network dynamics controlling keratinocyte migration. *Mol Syst Biol*, 4, 2008.

[11] Green H and Kehinde O. An established preadipose cell line and its differentiation in culture. II. Factors affecting the adipose conversion. *Cell*, 5:19–27, 1975.

[12] Green H and Kehinde O. Spontaneous heritable changes leading to increased adipose conversion in 3T3 cells. *Cell*, 7:105–113, 1976.

[13] Hackl H, Burkard TR, Sturn A, Rubio R, Schleiffer A, Tian S, Quackenbush J, Eisenhaber F, and Trajanoski Z. Molecular processes during fat cell development revealed by gene expression profiling and functional annotation. *Genome Biol*, 6:R108, 2005.

[14] Jaeger H. Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach. GMD Report 159, German National Research Center for Information Technology, 2002.

[15] Schmidt H and Jirstrand M. Systems biology toolbox for matlab: a computational platform for research in systems biology. *Bioinformatics*, 22:514–515, 2006. `http://www.sbtoolbox2.org`.

[16] Liu J, DeYoung SM, Zhang Ming, Zhang Mei, Cheng A, and Saltiel AR. Changes in the integrin expression during adipocyte differentiation. *Cell Metab*, 2:165–177, 2005.

[17] Boyle KB, Hadaschik D, Virtue S, Cawthorn WP, Ridley SH, O'Rahilly S, and Siddle K. The transcription factors Egr1 and Egr2 have opposing influences on adipocyte differentiation. *Cell Death Differ*, 16:782–789, 2009.

[18] Cheung KJ, Tzameli I, Pissios P, Rovira I, Gavrilova O, Ohtsubo T, Chen Z, Finkel T, Flier JS, and Friedman JM. Xanthine oxidoreductase is a regulator of adipogenesis and PPARγ activity. *Cell Metab*, 5:115–128, 2007.

[19] Park KW, Waki H, Villeanueva CJ, Monticelli LA, Hong C, Kang S, MacDougald OA, Goldrath AW, and Tontonoz P. Inhibitor of DNA Binding 2 is a small molecule-inducible modulator of Pparγ expression and adipocyte differentiation. *Mol Endocrinol*, 22:2038–2048, 2008.

[20] Cao L, Lin EJD, Cahill MC, Wang C, Liu X, and Curing MJ. Molecular therapy of obesity and diabetes by a physiological autoregulation approach. *Nat Med*, 15:447–455, 2009.

[21] Qiao L, Zou C, ShaoP, Schaack J, Johnson PF, and Shao J. Transcriptional regulation of fatty acid translocase/Cd36 expression by Cebpa. *J Biol Chem*, 284:8788–895, 2008.

[22] Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P, and von Mering C. STRING 8 - a global view on proteins and their functional interactions in 630 organisms. http://string-db.org/. *Nucl Acids Res*, 37, 2009.

[23] Wahde M and Hertz J. Modeling Genetic Regulatory Dynamics in Neural Development. *J Comput Biol*, 8:429–442, 2001.

[24] Wahde M and Szallasi Z. A survey of Methods for Classification of Gene Expression Data Using Evolutionary Algorithms. *Expert Rev mol Diag*, 6:101–110, 2006.

[25] Arbib MA, editor. *The Handbook of Brain Theory and Neural Networks*. The MIT Press, 1995.

[26] Shannon P, Markiel A, Ozier O, Baliga SN, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res*, 13:2498–2504, 2003.

[27] Tontonoz P, Graves RA, Budavari AI, Erdjument-Bromage H, Lui M, Hu E, Tempst P, and Spiegelman BM. Adipocyte specific transcription factor ARF6 is a heterodimetric complex of two nuclear hormone receptors, PPARγ and RXR alpha. *Nucl Acids Res*, 22:5628–5634, 1994.

[28] Willett P. Genetic algorithms in molecular recognition and design. *Trends biotechnol*, 13:516–521, 1995.

[29] Tan Q. Inhibition of cholesterol biosynthesis under hypoxia. Master's thesis, A&M University of Texas, 2005.

[30] Beer RD. On the dynamics of Small Continuous-Time Recurrent Neural Networks. *Adaptive Behavior*, 3:469–509, 1995.

[31] Saha RN, Liu X, and Pahan K. Up-regulation of BDNF in Astrocytes by TNF-a: A Case for the Neuroprotective Role of Cytokine. *J Neuroimmune Pharmacol*, 1:212–220, 2006.

[32] Farmer S. Transcriptional control of adipocyte formation. *Cell Metab*, 4:263–273, 2006.

[33] Kernie SG, Liebl DJ, and Parada LF. BDNF regulates eating behavior and locomotoractivity in mice. *Embo J*, 19:1290–1300, 2000.

[34] Health VJ, Gillespie ADF, and Crouch DH. Inhibition of adipocyte differentiation by cMyc is not accompanied by alterations in cell cycle control. *Biochem Biophys Res Commun*, 269:438–443, 2000.

[35] Health VJ, Gillespie ADF, and Crouch DH. Inhibition of the terminal stages of adipocyte differentiation by cMyc. *Exp Cell Res*, 254:91–98, 2000.

[36] Steeb WH. *The Nonlinear Workbook: Chaos, Fractals, Neural Networks, Genetic Algorithms, Gene Expression Programming, Support Vector Machine, Wavelets, Hidden Markov Models, Fuzzy Logic with C++, Java and SymbolicC++ Programs.* World Scientific Publishing, 2008.

[37] Wu Z, Rosen ED, Brun R, Hauser S, Adelmont G, Troy AE, McKeon C, Darlington GJ, and Spiegelman BM. Cross-regulation of C/EBPα and PPARγ controls the transcriptional pathway of adipogenesis and insulin sensitivity. *Mol Cell*, 3:151–158, 1999.

[38] Yun Z, Maecker HL, Johnson RS, and Giaccia AJ. Inhibition of PPARγ2 gene expression by the HIF-1-regulated gene DEC1/Stra13: a mechanism for regulating adipogenesis by hypoxia. *Dev Cell*, 3:331–341, 2002.