Mechanistic Insights into Metabolic Disturbance during Type-2 Diabetes and Obesity Using Qualitative Networks

Antje Beyer[1], Peter Thomason[2], Xinzhong Li[2], James Scott[2], and Jasmin Fisher[3]

[1] Department of Genetics, University of Cambridge, Cambridge, UK
[2] National Heart and Lung Institute, Imperial College London, UK
[3] Microsoft Research, Cambridge, UK
jasmin.fisher@microsoft.com

Abstract. In many complex biological processes quantitative data is scarce, which makes it problematic to create accurate quantitative models of the system under study. In this work, we suggest that the Qualitative Networks (QNs) framework is an appropriate approach for modeling biological networks when only little quantitative data is available. Using QNs we model a metabolic network related to fat metabolism, which plays an important role in type-2 diabetes and obesity. The model is based on gene expression data of the regulatory network of a key transcription factor Mlxipl. Our model reproduces the experimental data and allows *in-silico* testing of new hypotheses. Specifically, the QN framework allows to predict new modes of interactions between components within the network. Furthermore, we demonstrate the value of the QNs approach in directing future experiments and its potential to facilitate our understanding of the modeled system.

Keywords: computational modeling, Qualitative Networks, metabolic pathways, obesity, type-2 diabetes, Mlxipl.

1 Introduction

One of the key objectives of systems biology is to understand the behavior of cellular signaling- and transcription-networks. High-throughput techniques enable the accumulation of large amounts of data relating to the behavior of these networks. However, we need to establish the ways and methodologies to organize this data into a coherent whole that sheds light on the underlying biological system. In recent years, different modeling approaches have been suggested and used to this end. For example, differential equations and Boolean Networks [3,5,8,16,17]. It is clear that for different networks and different kinds of data, different approaches may need to be taken. Many modeling techniques require accurate quantitative data in order to calibrate and set the values of certain "constant rate" parameters. Our main interest is the case where such accurate

C. Priami et al. (Eds.): Trans. on Comput. Syst. Biol. XII, LNBI 5945, pp. 146–162, 2010.

data is missing but still there is a need to create formal executable models that help analyze and understand large amounts of qualitative data.

In this paper we advocate the usage of the *Qualitative Networks* framework [12] as a means for modeling metabolic and transcriptional networks. Qualitative Networks are an extension of Boolean Networks. The network associates variables that range over finite (and usually very small) domains with each of the substances of the network. All variables are updated synchronously in discrete computation steps that are intended to abstract the continuous deterministic evolvement of models described by differential equations. In every step of the system a variable may either increase by one, decrease by one, or remain unchanged. The changes of variables are determined by so called *target functions* that given the current state of the network (i.e., the value assignment to each one of the variables) compute the target value for each substance. This substance then changes in increments / decrements of one until reaching this value (assuming that the value of the target function does not change). Qualitative Networks have several characteristics that arise from their underlying semantics. The system can be in one of a finite number of states (value assignments to all the variables) and as the system is deterministic the execution of the system inevitably terminates in a loop.

There are mainly two analysis techniques that are applied to Qualitative Networks. The first is the execution of a network from a given initial state. In such a simulation, the changes in values of substances are followed over time from a given initial state and mimics specific evolutions of the system. In order to compare such simulations to evolutions of the underlying biological system several measurements of the level of substances in the biological system over a short period of time need to be available. While this information is very useful for debugging and understanding the way the model works, such accurate experimental data is very hard to come by and is almost impossible to find. A second analysis technique is to find the set of stable states (or stable loops) of the model. These are the states in which the model can remain looping forever. Usually, these are computed for a large number of initial conditions simultaneously. These states relate to the stable states of the biological system and their analysis can relate to the values of substances in the stable states of the system. Experimentally, these values are easier to come by with as they correspond to measurements in a single time point and in constant conditions that are easier to maintain. Technically, we use symbolic state-space analysis techniques [1] in order to efficiently and practically compute the set of stable states of large qualitative Networks. We then compare the values of variables in these stable states against data extracted from experimental data.

We illustrate the suitability of Qualitative Networks to modeling of metabolic networks in the case where quantitative data is not available. We show that even with restricted amounts of data, Qualitative Networks can still suggest useful hypotheses that lead to new avenues to explore experimentally and interesting insights. This is demonstrated through an extensive case study of a metabolic and transcriptional interaction network involved in obesity and type-2 diabetes

mellitus (T2D). In both diseases, metabolic and inflammatory pathways play important roles [19] as their dysregulation can lead to insulin resistance which is a characteristic symptom of obesity and T2D [10,11,15]. Nowadays, T2D affects over 110 million people worldwide and, as well as obesity, it highly increases the risk of cardiovascular disease, blindness, amputation and kidney failure [9,10]. With cardiovascular disease being a major cause of mortality, T2D and obesity pose a significant threat to global health [9,19].

Our model is based on unpublished data by Scott and colleagues [15]. The data was obtained in a study looking at macrophage infiltration and gene expression profiles in insulin-responsive tissues of C57Bl6 mice which are susceptible to these metabolic diseases. To determine the onset of inflammation and metabolic disturbance in adipose tissue, mice were fed a high saturated fat diet. The data includes gene expression measurements in white adipose tissue in mice, performed at two days, eight days, three weeks, and 15 weeks after the beginning of a fat-feeding process. We concentrate on the measurements at eight days and 15 weeks as understanding the differences between the gene expression levels in these time points could help understand the difference between processes operating during early onset of T2D and obesity (8 days) and the chronic disease stage (15 weeks) [15]. For the computational model, we used gene expression data obtained from white adipose tissue in the gonadal region of male mice. Scott and colleagues selected genes based on epididymal fat weight after 8 days of fat feeding with fat as 60% of the caloric intake (cf. [15]). Genes were selected if their expression was correlated with epididymal fat weight above a very strict threshold ($r^2 > 0.7$). They hypothesize that these genes might be the bridge between acute fat-feeding and chronic obesity. Most of the genes used in the present computational model are from this selection or very close to this threshold because of their potential importance. The results by Scott and colleagues also suggest that metabolic disturbance leads to inflammation. Therefore, the current model serves also as a basis for future modeling of the inflammatory processes underlying T2D and obesity.

Our case study presents an executable model of the transcriptional and post-translational regulations of Mlxipl. The transcription factor Mlxipl is known to increase the transcription of genes involved in several metabolic pathways [6,7,15] and is significantly down-regulated in white adipose tissue during acute fat feeding and obesity [15]. Mlxipl can be linked to many aspects of obesity and T2D [6,7,18,15]. While the post-translational regulation of Mlxipl is fairly well understood, its regulation on the transcriptional and translational level is less clear [18]. Our computational model includes Mlxipl together with four interconnected metabolic pathways regulated by Mlxipl. As our data is taken from a tissue rather than from cells of the same type, we do not know exactly if the data relate to the major cellular constituents of white adipose tissue. But overwhelmingly the data refer to adipocyte functions rather than other cell types and are therefore highly likely to reflect the average fat cell. Hence, we build our computational model as a single cell model. The computational model is based on a metabolic network and a transcription interaction graph.

We analyze the model in conditions that represent the behavior of the network at 8 days and 15 weeks of fat feeding process. We use experimental data to calibrate our model by comparing the stable states of the model with the experimental data. This process has suggested modifications to the interaction graph and metabolic networks that need to be validated experimentally. Once a putative model reproduces the existing experimental data we formulate hypotheses as *in-silico* experiments and test the models as a way to direct actual laboratory experiments.

The computational cost of state-space analysis forced us to break down the network into four modules assuming that the components connecting the modules are constants in the later module. These components get all of their inputs in one module and have some or all of their outputs in the next module. This may cause some of the resulting component values to be too high or too low as it does not fully capture the dynamics of the whole system. However, all regulation loops in the full system are contained in one of the four modules, suggesting that most behaviors that do not exist in the partitioned model are transient. Furthermore, comparison with the experimental data revealed that the variables resulting from the network behave as expected, which emphasizes the validity of the whole model.

Our computational model and its analysis demonstrate the usefulness of Qualitative Networks to investigate cellular networks and their corresponding genomic data. With its individually adjustable target functions and its independence from quantitative data, Qualitative Networks can be used to model interesting biological interaction networks with available genomic data. This is especially emphasized by the Mlxipl case study. Although there is a good amount of knowledge on Mlxipl function, its regulation is not fully understood. The current case study provides insights into this regulatory process as well as highlights the importance of specific regulatory connections. The computational analysis reveals that the model is consistent with the experimental data. Our insights highlight the necessity of further regulators of Mlxipl in addition to the known ones. Furthermore, the analysis predicts various new modes of interactions between components in the network, and demonstrates that acetyl CoA and Mlxipl regulate the level of fatty acid production in a synergistic way. Future work of a more complex computational model of this system will further facilitate a better understanding of metabolic disturbance during obesity and T2D.

2 Methods

2.1 Qualitative Networks

In many cases, lack of quantitative data makes the construction of quantitative models impossible. Boolean Networks were used in recent years to construct abstracted computational models that are based mainly on the types of interactions between molecules. This is obviously a very abstract view, however, it has proven beneficial in numerous occasions [8,16,17]. In some cases, the restriction of Boolean Networks to the values of ON and OFF, seems too restricted and

one would like to extend the modeling framework with a larger range of possible values.

Qualitative Networks are an extension of Boolean Networks; they use discrete variables ranging over finite domains instead of Boolean variables and allow functions to represent different types of interactions on top of activation and inhibition [12].

A *Qualitative Network* $Q(C, T, N)$ comprises a set of components, C, and a list of target functions, T. Each component $c_i \in C$ has a state which can take any integer value between 0 and N representing the qualitative level of the respective component. A target function $target_i \in T$ is a function $target_i : \{0, \ldots, N\}^C \rightarrow \{0, \ldots, N\}$ and it indicates in a given state of the system the specific value between 0 and N towards which each component c_i should move. For example, target functions can be as simple as the identity function, a Boolean function, the maximum or minimum, or a more complicated combination of other functions.

A *state* of the system is an assignment of a value between 0 and N to each of the substances. The system advances in discrete steps and, at each step, the level of a component can only increase or decrease by a single level. The consecutive level of each component c_i is calculated as follows:

$$c_i(t+1) = \begin{cases} c_i(t) - 1 & \text{if } target_i(S(t)) < c_i(t) \\ c_i(t) + 1 & \text{if } target_i(S(t)) > c_i(t) \\ c_i(t) & \text{if } target_i(S(t)) = c_i(t) \end{cases},$$

where $S(t)$ is the state of the network at time t. For example, consider the case of a network with three components c_1, c_2, and c_3. Let their values at time 1 are $c_1(1) = 1, c_2(1) = 2$ and $c_3(1) = 0$, and the target function $target_3(S(t)) := max(c_1(t), c_2(t))$ for the component c_3. Then $c_3(2)$ aspires to $target_3(S(1)) = max(1, 2) = 2$ and as $2 > 0$ the value of $c_3(2)$ becomes 1. The target functions are based on the mechanistic understanding of the biological interactions the model is built from.

As all variables range over finite domains, a *Qualitative Network* model has a finite number of states. As all variables are updated simultaneously (synchronously) an execution of such a model proceeds deterministically, i.e., regardless of the full history of an execution, every state proceeds to a fixed next state. It follows that every execution of a *Qualitative Network*[1] eventually ends in a cycle of states that are visited infinitely often. A state is defined as infinitely visited if there is an execution of the model in which the state appears infinitely often. In a biological system these states correspond to the stable states, hence states that are not infinitely visited often are considered as unstable. From these unstable states the system will always evolve towards a loop of infinitely visited states and remain there indefinitely. This suggests the following analysis technique: (a) compute the set of infinitely visited states (b) check properties of this set. For example, check whether in all infinitely visited states the value of a certain variable (substance) is above/below a given threshold. As this analysis aims

[1] As Boolean Networks are a special case, this applies also to Boolean Networks.

to validate the stability of the model, it is usually less important to identify one or few initial states. The analysis is usually applied on most (if not all) states of the system and corresponds to starting execution from an arbitrary state and letting the system stabilize [13].

We use the QNBuilder tool [14] to analyze *Qualitative Networks*. QNBuilder supports a simple and intuitive textual input format that defines the set of substances in the model. Some substances can be defined as constants and others are connected to target functions, that can be defined as look-up tables. Cells can be connected to grids and meshes in a simple way by relating to inputs from neighboring cells according to their locations in the grid. The QNBuilder uses the tool *CrocoPat* [1] for symbolic state analysis. CrocoPat computes a simple fixpoint of the states that can be reached in 1, 2, ... steps. As the number of states is finite, this computation stabilizes on the fixpoint of infinitely visited states. QNBuilder also includes requirements that are then checked on this set of infinitely visited states. In the case that a requirement does not hold, an infinitely visited state that does not conform to it can be extracted.

2.2 The Iterative Improvement Process [12]

When coming to check the accuracy of a putative biological model, two types of requirements have to be met. The first is a static requirement such as, a inhibits b or a activates b. Such requirements are integrated into the target functions and can be checked statically. The second is a dynamic requirement that relates to qualities of infinitely visited states. Such requirements are checked as explained above.

The main value of such models is that they can be contrasted with the experimental data and, hopefully, provide information that the biologist was not aware of initially. Accordingly, if a model fails to reproduce some of the experimental data, it means that our understanding of the process is incomplete, and the model should be refined through changes in the target functions, substances, interactions, or perhaps the range of possible values. Such changes can be justified by consulting the available literature or domain experts. When such information is not available, we may conclude that information is missing. We can try to modify the model so that all the experimental conditions are reproduced. Such modifications then need to be validated experimentally in order to verify that the model may indeed be correct.

Once a model reproduces the experimental data, we view it as a valid putative model and use it to further probe the system. We perform *in-silico* experiments by changing constants or requirements in the model. The results of such experiments can be used to highlight interesting avenues to explore experimentally.

3 Results

3.1 Case Study – Model of the Regulatory Network of Mlxipl

We demonstrate the usefulness of the Qualitative Networks through a case study and the insights that were obtained from its analysis. We construct a Qualitative

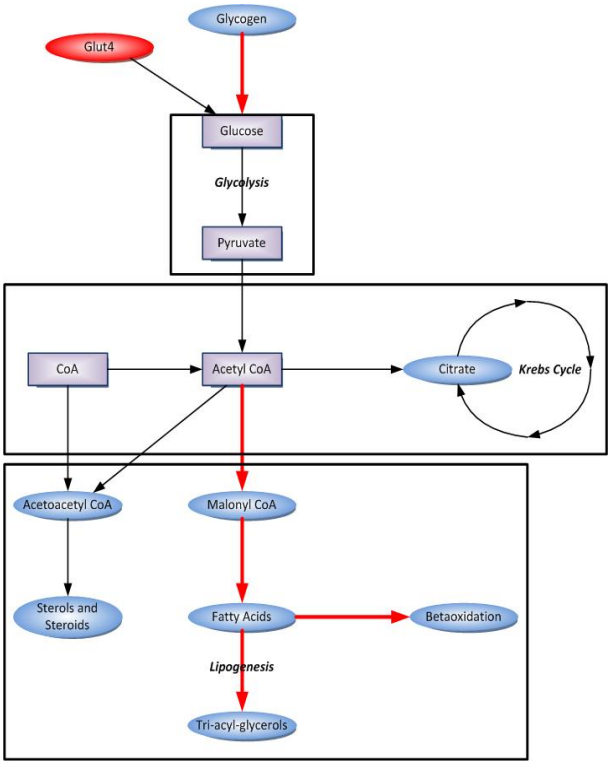

Fig. 1. Coarse representation of the metabolic network used for the computational model. Red arrows and red filling show the parts of the network that Mlxipl has influence on. The boxes indicate three of the separate models that are worked out in more detail.

Networks model of the regulatory network of Mlxipl, a general metabolic network that can be found in all mammalian cells (cf. Figure 1). There are four specific subnetworks making up this big metabolic network which were worked out in more detail as four separate models (due to computational constraints) according to the specific connections that are found in white adipose tissue (cf. Figures 2, 3, 4, and 5 and Figure 6 showing their connectivity). Figure 2 shows the transcription factor Mlxipl with its regulatory connections and some putative regulators of Mlxipl. Figures 3, 4 and 5 show biochemical reactions in the metabolic network underlying the present model including components regulated by Mlxipl.

We use granularity of three for all components in the network. The three possible levels correspond to: 0 representing down-regulation, 2 representing up-regulation, and 1 representing normal level of expression. The network includes two different kinds of interactions between components: regulatory connections on the gene or protein level (especially in the Mlxipl network), and biochemical reactions. These different interactions are represented by different target functions. For regulatory connections of the first kind (e.g., the control of PPARγ

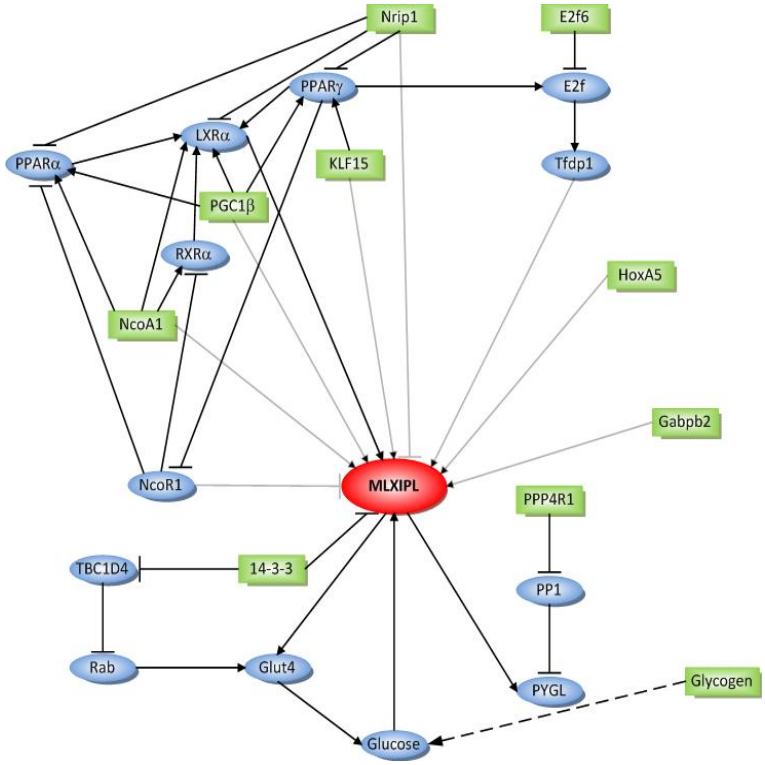

Fig. 2. Network of Mlxipl regulatory connections as used in the computational model. The gray arrows show putative inputs that are not verified experimentally. The green components are constant in the model, the blue ones are not constant and used in the requirement, the red one is key for consistency of the model. The dashed arrow indicates a biochemical reaction, the attached component is the enzyme of the reaction.

in Figure 2) target functions compare the amount of activating inputs with the amount of inhibiting inputs. For biochemical reaction of the second kind (e.g., the control of Fructose-6-P in Figure 3), the amount of the output depends on whichever input component has a lower level in this reaction, substrate or enzyme. The actual target functions are determined as explained below [12].

Consider a protein c_i of the network. An edge e_{ji} represents an interaction from c_j to c_i and has an associated weight α_{ji}. The edge is activating when $\alpha_{ji} > 0$ and inhibiting if $\alpha_{ji} < 0$. The target function of c_i includes the computation of the sum of activations and inhibitions that are scaled so that they always fall in the range 0 to N. Formally, the sum of activation on c_i is

$$act_i(s) = \frac{\sum\limits_{\alpha_{ji}>0} \alpha_{ji} c_j}{\sum\limits_{\alpha_{ji}>0} \alpha_{ji}},$$

and the sum of inhibition on c_i is

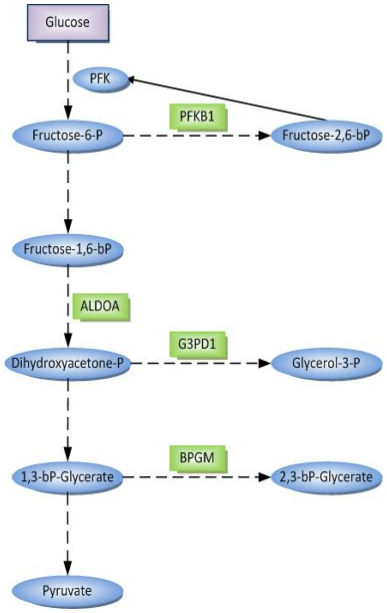

Fig. 3. Network representation of glycolysis used for the computational model. The purple component is a constant in this module, however it is an output in the Mlx-ipl module. Green components are constants in this module, the blue components are not constant and hence used in the requirement. Dashed arrows indicate biochemical reactions, the attached component is the enzyme of the reaction. If there is no component attached to the dashed arrow, the enzyme was not included in the model as its expression level did not significantly change from normal.

$$inh_i(s) = \frac{\sum\limits_{\alpha_{ji}<0} -\alpha_{ji} c_j}{\sum\limits_{\alpha_{ji}<0} \alpha_{ji}},$$

where s represents the state of the model. The target function for c_i is determined according to the difference between these two values. In the case that a component c_i has only inhibiting inputs, lack of inhibition is interpreted as activation. Formally,

$$target_i(s) = \begin{cases} \max(0, act_i(s) - inh_i(s)) & \text{If } \max(\alpha_{ji} > 0) \\ N - inh_i(s) & \text{If } \max(\alpha_{ji}) \leq 0 \end{cases}.$$

For example, the target function of PPARγ is the sum of its two activating components, PGC1β and KLF15 divided by 2, minus its inhibiting component, Nrip1.

If component c_i is controlled by a biochemical reaction, its target function is set to the minimum of the levels of its substrate and its enzyme. For example, Fructose-6-P tends towards the level of the level of the lowest between Glucose and PFK.

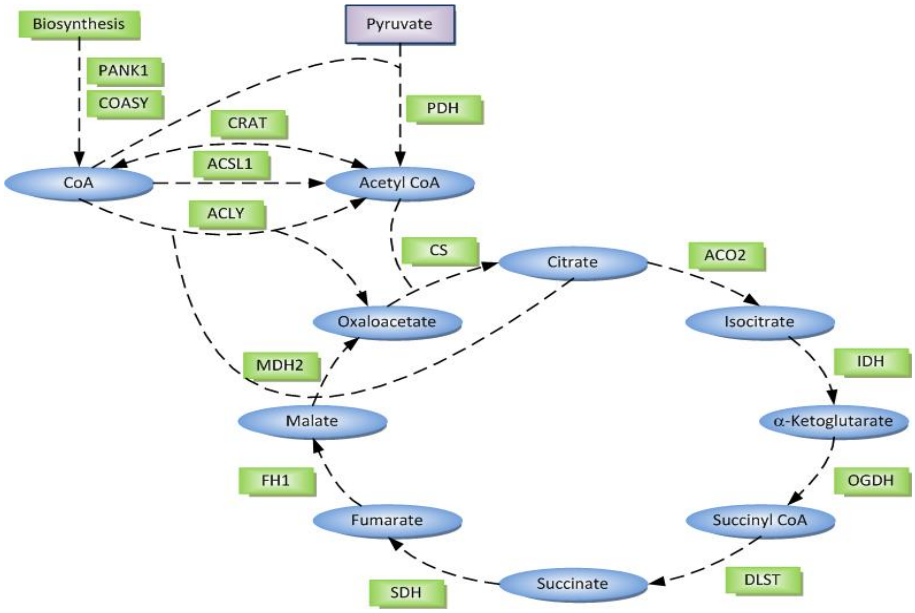

Fig. 4. Network surrounding acetyl CoA as used in the computational model. The purple component is constant in this module, however is an output in the glycolysis module. Green components are constants in this module, the blue components are not constant and used in the requirement. Dashed arrows indicate biochemical reactions, the attached component is the enzyme of the reaction.

For constant components, i.e. components that were not getting input from other components in the network, these values were extracted from the gene expression data (Table 1, data supplied by Scott and colleagues [15]). This data relates to two different time points after 8 days and 15 weeks of fat feeding process. These give rise to two versions of each model, differing in the values of the constant components, where there are differences in gene expression. For example, PP4R1 is constantly down-regulated in the 8 days model and constantly up-regulated in the 15 weeks model. We stress that these two version of the models use the same network structures and same target functions and differ only in these constant values. We used a threshold of 1 ± 0.1 to represent normal expression. A level above 1.1 is treated as up-regulated and a level below 0.9 is treated as down regulated[2]. In Table 1, green represents the value 0, white represents 1 and red represents 2.

[2] Setting the threshold for the different levels is an interesting biological question. From discussions with colleagues, it seems that a change of 10% in expression level is considered significant. It would be interesting to study models with a lower threshold for down- and up-regulation. Another interesting option is to allow more expression levels in the model and use extra levels for expressing down- and up-regulation in a more accurate way. Increasing the level of granularity of substances produces models that are too large to be analyzed. This is one of the topics of our future work.

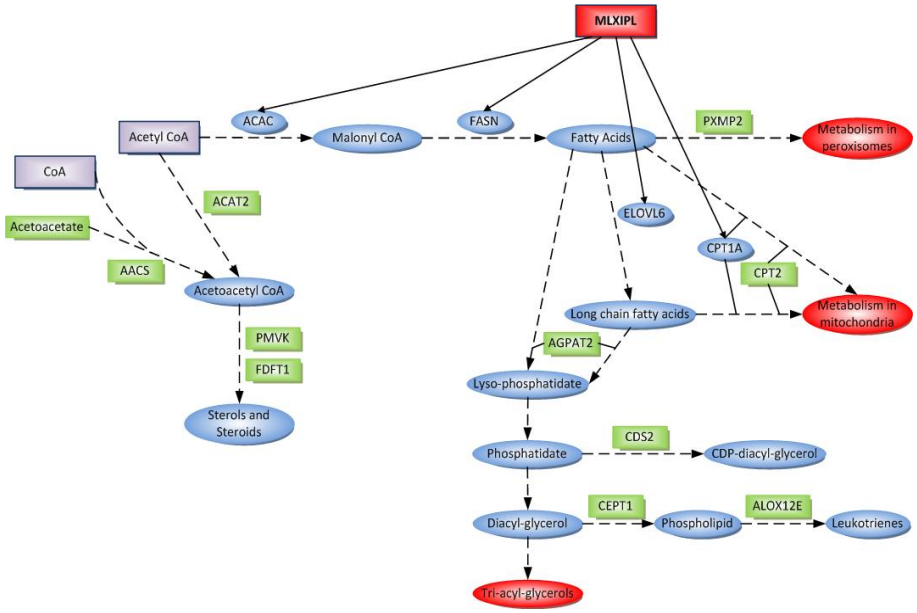

Fig. 5. Network representation of fatty acid metabolism used for the computational model with the exception of the components between Lyso-phosphatidate and Tri-acyl-glycerols. The purple components are constants in this module, however are outputs in the acetyl CoA module. Green components are constant in this module, the blue components are not constant and used in the requirement, and the red components are the key for consistency of the model. Dashed arrows indicate biochemical reactions, the attached component is the enzyme of the reaction.

The experimental data is used, in addition, to provide requirements on the network. For the present model, the requirements were derived from the gene expression data for eight days and 15 weeks of fat-feeding (cf. Table 1, [15]) for all non-constant components corresponding to the three levels of expression, as explained above. For example, in the model corresponding to 8 days, the stable states of the Mlxipl module (cf. Figure 2) have to satisfy the following requirement. The values of PPARγ, PPARα, LXRα, RXRα, TBC1D4, Glut4, PYGL, and Mlxipl must be 0 and the value of Ncor1 must be high. Similar requirements were derived for all modules and for both versions of the model, corresponding to 8 days and 15 weeks. The objective of the formal analysis is to verify that these requirements generally hold for the model. In the process of formal verification, all infinitely visited states were computed first and the requirements were then tested on *all* of these states.

3.2 New Biological Insights Revealed by Model Analysis

in-silico experiments were performed to test further assumptions. This was done in two ways, either by changing the values of constant components in the network

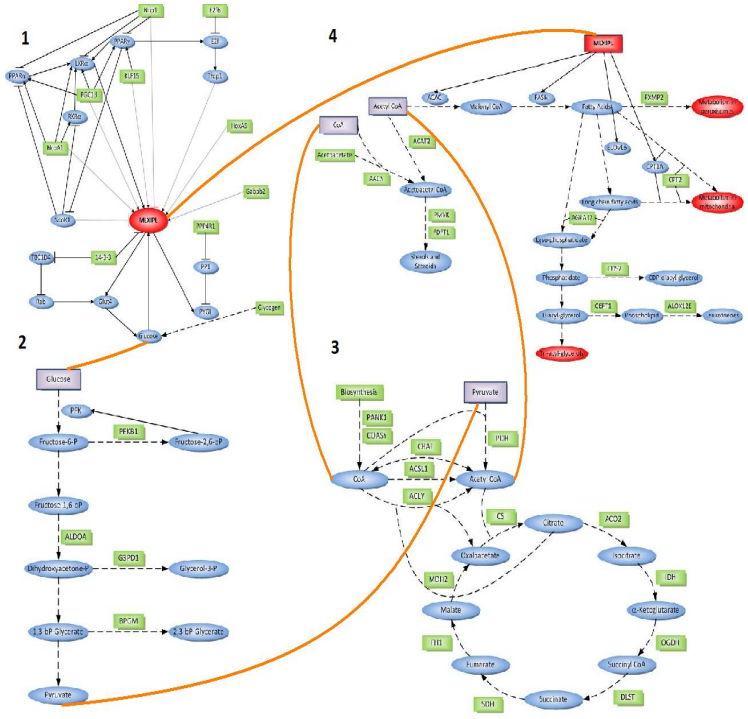

Fig. 6. Figure showing the four main parts of the metabolic network with their interactions as used in the computational model. (1) network of Mlxipl regulatory connections, (2) glycolysis, (3) network surrounding acetyl CoA and Krebs cycle, (4) fatty acid metabolism.

or by removing interactions (and components) from the network. While changing the values of the components, the values of other interesting components were monitored so that the effect of the changed components on the other components could be investigated. By removing interactions we tested if the model was still consistent with the data and hence if the existing connections are important or necessary to maintain the behavior of the system.

Ywhab is likely to be involved in Mlxipl down-regulation. It is not clear which one of the proteins of the 14-3-3 family is involved in the inhibition of Mlxipl, but Ywhab and Ywhag are the most likely candidates (cf. Table 1). During the consistency checking, we tried to find out which of these two proteins is most probable to inhibit Mlxipl in the present network. Considering all required levels of the other components in the Mlxipl network shows that, while both proteins are in accordance with the major requirement of Mlxipl being at a low level, the inclusion of Ywhag into the network was less consistent with the experimental data than including Ywhab. This indicates that Ywhab is the most likely

Table 1. Tables of gene expression data relative to gene expression with normal diet (represented by 1 in the model) sorted alphabetically and with regard to their model affiliation. Green indicates genes that are down-regulated (represented by 0 in the model), red shows genes that are up-regulated (represented by 2 in the model). Genes marked in yellow were not used as constants in the model.

MLXIPL Regulation

Gene	*8 days*	*15 weeks*
14-3-3		
Ywhab	0.86	0.60
Ywhag	0.93	0.92
E2f6	0.84	0.83
Gabpb2	0.84	0.82
Glut4	0.48	0.73
HoxA5	0.74	0.57
KLF15	0.63	0.55
LXRa	0.80	0.56
MLXIPL	0.45	0.69
NcoA1	0.79	0.73
NcoR1	1.15	0.96
Nrip1	1.33	1.30
PGC1b	0.58	0.84
PP1	0.83	1.08
PPARa	0.72	0.94
PPARg	0.76	0.71
PPP4R1	0.80	1.11
PYGL	0.68	0.78
RAB40B	0.43	0.51
RXRa	0.89	0.92
TBC1D4	0.65	0.83
TFDP1	0.80	0.92

Glycolysis

Gene	*8 days*	*15 weeks*
ALDOA	0.58	0.90
BPGM	2.08	1.21
G3PD1	0.38	0.83
PFK	0.76	0.92
PFKFB1	0.63	0.69

Acetyl CoA Network and Krebs Cycle

Gene	*8 days*	*15 weeks*
AACS	0.56	0.59
ACAC	0.57	0.36
ACAT2	0.54	0.78
ACLY	0.71	1.31
ACO2	0.51	0.94
ACSL1	0.65	0.58
COASY	0.52	0.71
CRAT	0.70	0.50
CS	0.66	0.88
DLST	0.83	0.72
FASN	0.11	1.05
FDFT1	0.62	1.13
FH1	0.62	0.79
IDH	1.15	1.15
MDH2	0.68	1.03
OGDH	0.77	0.97
PANK1	0.37	0.86
PDH	0.58	0.82
PMVK	0.34	0.81
SDH	0.74	0.88

Fatty Acid Metabolism

Gene	*8 days*	*15 weeks*
AGPAT2	0.52	0.82
ALOX12E	0.66	0.57
CDS2	0.72	1.08
CEPT1	0.78	0.80
CPT1A	1.50	1.42
CPT2	0.65	0.67
ELOVL6	0.31	1.73
PXMP2	0.64	0.63

inhibitor and, in addition, down-regulation of Mlxipl shows the consistency of the Mlxipl model with the experimental data. This consistency implies that new hypotheses can be tested using this model. With the model adjusted to our results regarding the proteins of the 14-3-3 family, we tested different hypotheses with respect to the number of regulatory inputs to Mlxipl. Our tests have shown that the three known regulators in the network are not sufficient to explain the behavior of Mlxipl. Further testing has shown that adding eight putative inputs to these three is, however, sufficient to have Mlxipl down-regulated. With regard to these putative inputs the model indicates that the three known inputs together with one arbitrary activating or inhibiting input of the eight putative ones are sufficient to explain Mlxipl down-regulation. However, we were not able to highlight specific inputs that are necessary for Mlxipl down-regulation. Nonetheless, the model has highlighted that all of the putative inputs are good candidates to be regulators of Mlxipl.

Table 2. Table showing the synergistic effect of acetyl CoA and Mlxipl on fatty acid production in the model. The numbers show the values of the components, green represents downregulation, white normal regulation and red upregulation.

Acetyl CoA	MLXIPL	Fatty Acids
0	0	0
0	1	0
0	2	0
1	0	0
1	1	1
1	2	1
2	0	0
2	1	1
2	2	2

Several genes must have additional inputs in the network. After checking all of the sub-network models for their consistency, we found that in the final model of lipid metabolism, betaoxidation of fatty acids in the mitochondria and peroxisomes, as well as fat production, were down-regulated as expected. This means that the model we derived from the complete network is consistent with the experimental data and allows new hypotheses about the network to be tested. In this process of state-space analysis, we were also able to highlight the genes TBC1D4, PP1 and CPT1A whose behavior was not explained by their inputs in the present network. Hence, we can assume that these genes must have additional regulatory inputs.

Also, the genes NcoR1, RXRα, PPARα and Tfdp1 from the Mlxipl module (cf. Figure 2), PFK from the glycolysis module (cf. Figure 3) and FASN and ELOVL6 in the acetyl CoA module (cf. Figure 4) have a different expression level at 8 days compared to 15 weeks (cf. Table 1). Comparing the models for these two time points showed that for these genes the changed level was not explained at 15 weeks. Further tests showed that the complete model was still consistent when applying these changed values to the 15 weeks model. Hence, we conclude that while these genes must have other regulatory inputs at both time points or only at 15 weeks, the rest of the network with all its connections represents the data at both time points.

Mlxipl and acetyl CoA have a synergistic effect on fatty acids production. To test the cooperative influence of acetyl CoA and Mlxipl on betaoxidation and fat production, we changed the value of one of these components while keeping the other at its normal level. In all cases the levels of betaoxidation and fat production remained low. Tests with both of the components being at an elevated level also did not change the levels of betaoxidation and fat production. This suggests that the other genes working on these pathways restrict the influence of acetyl CoA and Mlxipl on this part of fat metabolism. A similar series of tests were conducted to explore the effect of acetyl CoA and Mlxipl on the amount of fatty acids production. Our test have shown that neither acetyl CoA nor Mlxipl can be regarded as the main regulator of fatty acid production,

but rather they seem to be acting in a synergistic way on the level of fatty acid production (cf. Table 2).

4 Discussion

We have presented the Qualitative Networks framework as an appropriate framework to modeling biological systems where little quantitative data is available. Qualitative Networks can be regarded as an abstraction of continuous models described by differential equations. Qualitative Networks do not require exact rates of reactions and are suitable for the kind of genomic data that is often available (as in our case study). They offer additional flexibility over Boolean Networks by allowing general target functions and multiple possible values of components. Overall, the Qualitative Networks framework is a very suitable approach for networks with no or only little quantitative data available and with complex biochemical interactions. Previously, Qualitative Networks were also used to construct a model of the crosstalk between the Wnt and Notch pathways in the Keratinocytes [12], and more recently to the modeling of cell fate specification during *C. elegans* vulval development [2].

Here we demonstrate the usefulness of the Qualitative Networks framework through a model of Mlxipl regulation. This model is based solely on gene expression data, and does not require exact reaction rates. The flexibility of target functions and multiple values enables to accurately represent enzymatic reactions and protein interactions in the same model. In accordance with several biological studies [6,15,18], the model reproduces the down-regulation of Mlxipl. Analysis of the model showed a down-regulation of the whole fat metabolism, which has also been demonstrated under similar experimental conditions [15]. This observation is especially important in terms of T2D and obesity as it is one of the major features of these diseases [15].

In addition, the model sheds light on aspects of the network that are not yet fully understood. For example, the model highlights the most likely candidate of the genes in the 14-3-3 family to inhibit Mlxipl [6,18]. Furthermore our model shows that Mlxipl must have other regulators in addition to well established ones and indicates which are the possible regulators. Additionally, analysis of the model suggests that Mlxipl and acetyl CoA have a synergistic effect on fatty acid production. Overall, this shows the applicability of the Qualitative Networks framework to the modeling of metabolic networks and its potential to highlight questions with implications for medical research.

4.1 Future Prospects

Several aspects of modeling were dictated by computational limitations that we hope to overcome in the future. Our main aim is to remove the partitioning of the network and analyze it as one model. This would prevent possible inconsistencies due to components being set to their stabilization values too early. Another goal is to increase the granularity of the target functions from 3 to a higher order, as

having only three levels of expression cannot always capture the complexity of a system. As an intermediate goal, we could use different granularities for different components according to their roles and importance. As previously discussed, we would also like to consider different cut-offs for the expression level, for example 1 ± 0.5. In such a case, 1 would represent all genes with a medium level of expression and 0 and 2 would represent all genes with an expression deviating significantly from normal. The comparison of the results for different cut-offs could perhaps replace the usage of target functions with higher granularity.

We would also like to extend this model with additional biological data. For example, the control of the inputs that are currently modeled as constants. Considering the bigger picture, a major goal for the future would be to add the inflammatory pathways to this model, as this would help to get a deeper and more fundamental understanding of how these processes in adipose tissue contribute to T2D and obesity.

Acknowledgments

We would like to thanks Nir Piterman for critical reading of the manuscript and valuable comments and discussions. Part of this work was done while Antje Beyer was an intern at Microsoft Research, Cambridge, UK.

References

1. Beyer, D.: Relational Programming with CrocoPat. In: Proceedings of the 28th International Conference on Software Engineering (ICSE 2006), Shanghai, May 20-28, pp. 807–810. ACM Press, New York (2006), `http://www.cs.sfu.ca/~dbeyer/CrocoPat/`
2. Beyer, A., Fisher, J.: Unpublished results (2009)
3. von Dassow, G., Meir, E., Munro, E.M., Odell, G.M.: The segment polarity network is a robust developmental module. Nature 406, 188–192 (2000)
4. Fisher, J., Piterman, N., Hajnal, A., Henzinger, T.A.: Predictive modeling of signaling crosstalk during C. elegans vulval development. PLoS Comput. Biol. 3, e92 (2007)
5. Giurumescu, C.A., Sternberg, P.W., Asthagiri, A.R.: Intercellular coupling amplifies fate segregation during *Caenorhabditis elegans* vulval development. PNAS 103, 1331–1336 (2006)
6. Iizuka, Y., Horikawa, K.: ChREBP: a Glucose-activated Transcription Factor Involved in the Development of Metabolic Syndrome. Endocr. J. 55(4), 617–624 (2008)
7. Kooner, J.S., Chambers, J.C., Aquilar-Salinas, C.A., Hinds, D.A., Hyde, C.L., Warnes, G.R., Gómez Pérez, F.J., Frazer, K.A., Elliot, P., Scott, J., Milos, P.M., Cox, D.R., Thompson, J.F.: Genome-wide scan identifies variation in Mlxipl associated with plasma triglycerides. Nature Genetics 40, 149–151 (2008)
8. Li, F., Long, T., Lu, Y., Ouyang, Q., Tang, C.: The yeast cell-cycle network is robustly designed. Proc. Natl. Acad. Sci. USA 101, 4781–4786 (2004)

9. Lin, J., Yang, R., Tarr, P.T., Wu, P., Handschin, C., Li, S., Yang, W., Pei, L., Uldry, M., Tontonoz, P., Newgard, C.B., Spiegelman, B.M.: Hyperlipidemic Effects of Dietary Saturated Fats Mediated through PGC-1β Coactivation of SREBP. Cell 120, 261–273 (2005)
10. Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M.J., Patterson, N., Mesirov, J.P., Golub, T.R., Tamayo, P., Spiegelman, B.M., Lander, E.S., Hirschhorn, J.N., Altshuler, D., Groop, L.C.: PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature Genetics 34, 267–273 (2003)
11. Neels, J.G., Olefsky, J.M.: Inflamed fat: what starts the fire? J. Clin. Invest. 116, 33–35 (2006)
12. Schaub, M.A., Henzinger, T.A., Fisher, J.: Qualitative networks: a symbolic approach to analyze biological signaling networks. BMC Systems Biology 1, 4 (2007), `http://www.biomedcentral.com/1752-0509/1/4/`
13. Schaub, M.A.: QNBuilder v0.1 User Manual, Draft (2008)
14. Schaub, M.A.: QNBuilder v0.1.c (2008), `http://cs.stanford.edu/people/mschaub/QNBuilderBeta/`
15. Scott, J.: Oxidative stress in adipose tissue is a unifying trigger for inflammation and insulin resistance (manuscript and personal communication)
16. Shmulevich, I., Zhang, W.: Binary analysis and optimization-based normalization of gene expression data. Bioinformatics 18, 555–565 (2002)
17. Shmulevich, I., Lahdesmaki, H., Dougherty, E.R., Astola, J., Zhang, W.: The role of certain Post classes in Boolean Network models of genetic networks. Proc. Natl. Acad. Sci. USA 100, 10734–10739 (2003)
18. Uyeda, K., Repa, J.J.: Carbohydrate hydrate response element binding protein, ChREBP, a transcription factor coupling hepatic glucose utilization and lipid synthesis. Cell Metabolism 4, 107–110 (2006)
19. Wellen, K.E., Fucho, R., Gregor, M.F., Furuhashi, M., Morgan, C., Lindstad, T., Vaillancourt, E., Gorgun, C.Z., Saatcioglu, F., Hotamisligil, G.S.: Coordinated Regulation of Nutrient and Inflammatory Responses by STAMP2 is Essential for Metabolic Homeostasis. Cell 129, 537–548 (2007)