University of Padova

Department of Physics and Astronomy "G. Galilei"

Master Thesis in Physics of Data

Mathematical modeling of steatotic liver cell metabolic network

Supervisor
Prof. Marco Baiesi
University of Padova

Co-supervisors
Dr. rer. nat. habil. Dirk Drasdo
Dr. Jieling Zhao
Research centre Inria de Saclay

Master Candidate
Matteo Pedrazzi
2076719

Academic Year
2023-2024

Dedication.

Ringrazio mio fratello ed i miei genitori, oltre che a i miei amici e familiari per il costante supporto e la comprensione.

Senza di voi questo non sarebbe mai stato possibile.

Abstract

Metabolic dysfunction–associated steatotic liver disease (MASLD), previously known as non-alcoholic fatty liver disease (NAFLD), is a chronic liver disease characterized by excessive fat accumulation in the hepatocytes, leading to liver steatosis and potential progression to more severe liver conditions, annually responsible of 1 out of every 25 deaths worldwide. Due to the lack of pharmacological and targeted treatments, this study aims to build a liver numerical digital twin, reproducing the kinetic model of hepatic lipid droplet metabolism from [1], encompassing FFA uptake, TAG esterification, and lipid droplet dynamics, based on experimental *in vitro* and *in vivo* findings. By simulating lipid droplets size distributions in hepatocytes under varying conditions, one of the goals is to highlights the role of regulatory surface proteins (RSPs) in cellular lipid accumulation. Additionally, a novel approach is employed by utilizing sensitivity analysis methods to identify the most significant input parameters. These parameters are then used as key features in the integration of a neural network, enhancing the numerical solution of the model's ordinary differential equations (ODEs) and potentially improving both predictive accuracy and computational efficiency in modeling hepatocyte lipid content.

Contents

Listing of figures

Listing of tables

Listing of acronyms

Compartments

cyt . . . . . cytoplasm

er . . . . . endoplasmatic reticulum

ex . . . . . external space (blood)

LD . . . . . lipid droplet

mit . . . . . mitochondria

nLD . . . . . nascent lipid droplet

VLDL . . . . very low-density lipoprotein

Enzimes & catalyzers

ACS . . . . . acetyl-CoA synthetase

AGPAT . . . acylglycerolphosphate acyltransferase

ATGL . . . . adipose triglyceride lipase

CoA . . . . . coenzime A

CPT . . . . . choline phosphotransferase

CPT1 . . . . carnitine palmitoyltransferase 1

DGAT1 . . diacylglycerol acyltransferase 1

DGAT2 . . diacylglycerol acyltransferase 2

GPAT . . . . glycerol 3-phosphate acyltransferases

HSL . . . . . hormone-sensitive lipase

MGL . . . . . monoacylglyceride lipase

MTP . . . . . microsomal (triglyceride) transfer protein

PAP . . . . . phosphatidic acid phosphatase

PCTP . . . . phosphatidylcholine transfer protein

Metabolites/proteins

ADP . . . . . adenosinediphsphate

ApoB . . . . apolipoprotein B100

ATP . . . . . adenosinetriphsphate

Car . . . . . carnitine

DAG . . . . . diacylglycerol

G3P . . . . . glycerol 3-phosphate

FFA . . . . . free (non-esterified) fatty acids

FA-CoA . . long chain fatty acyl-CoA

LPA . . . . . lysophosphatidic acid

MAG . . . . . monoacylglycerol

malcoa . . . malonyl coenzime A

PA . . . . . phosphatidic acid

PC . . . . . phosphatidylcholine

TAG . . . . . triacylglycerol

Regulatory surface proteins

PLIN1 ... perilipin 1

ADRP ... adipose differentiation-related protein (perilipin 2)

TIP47 .... tail-interacting protein 47 (perilipin 3)

CGI58 ... comparative gene identification-58

FSP27 .... fat-specific protein 27

1 Introduction

Liver diseases rank among the leading causes of mortality in patients: they accounts for two million deaths annually and are responsible for 4% of all deaths (1 out of every 25 deaths worldwide); approximately two-thirds of all liver-related deaths occur in men. Deaths are largely attributable to complications of cirrhosis and hepatocellular carcinoma [2].

This work focuses on the early stage of liver disease progression, known as fatty liver disease, which is the most prevalent liver disorder in modern times. However, recent years have shown a notable increase in its prevalence, as the increase of obesity could be an important contributing factor [3]. Indeed NAFLD, the acronym for Non-Alcoholic Fatty Liver Disease, had prevalence equal to 25.2% in 2022 with the alarming forecast of 33.5% by 2030 [4].

Specifically, fatty liver disease, also known as hepatic steatosis, is defined as a pathological condition in which fat accumulates within liver cells in the form of lipid droplets. Its nomenclature has recently been revised in order to remove the word "alcoholic" from the acronym [5]. Former non-alcoholic fatty liver disease (NAFLD), now metabolic dysfunction–associated steatotic liver disease (MASLD), is a condition in which excess of fat resides in liver cells and at least one metabolic risk factor is present. The definition in literature ([6]) requires:

a) evidence of hepatic steatosis, either by imaging or by histology;

b) no causes for secondary hepatic fat accumulation such as significant alcohol consumption, use of steatogenic medication or hereditary disorders.

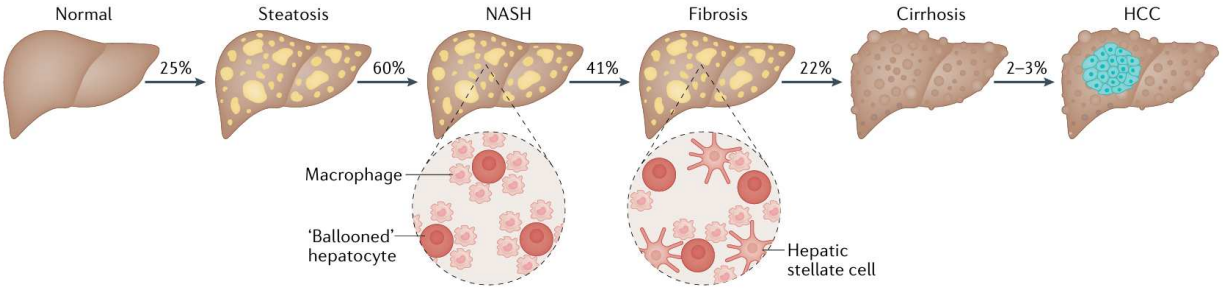


Figure 1.1: The different liver disease progression stages, from steatotic liver to hepatocellular carcinoma (HCC), with the percentages representative of the proportion of the sub-population that will probably present or progress to more advanced stages. Picture taken from [7].

NAFLD includes a constellation of histological findings that goes from steatosis, to necroinflammation, called NASH and progression to advanced fibrosis and cirrhosis [3, 4].

Fatty liver disease can be caused by several different factors, being them either modifiable or not. Primary risk factors for this disease are type 2 diabetes, obesity and dietary reasons like a high fat diet or alcohol and tobacco consumption. Other found modifiable factors are sleep and high blood pressure. Among the non-modifiable ones are the age, gender, ethnicity, genetic and socio-economic status.
The symptoms of steatosis can vary widely, ranging from no noticeable symptoms to mild fatigue and discomfort, and in more severe cases, more pronounced issues. Common symptoms include fatigue, abdominal pain or discomfort, and unexplained weight loss. In more advanced stages, patients may experience jaundice (yellowing of the skin and eyes), swelling in the legs or abdomen (edema or ascites) and mental confusion.

Based on the triggering behaviour of the disease, fatty liver disease can take different names, considering whether or not metabolic and diet reasons are involved in the emergence of the disease. Other than the already mentioned MASLD, we can have MetALD (or ALD) if depending on extent of alcohol intake (or other combination aetiology), when the patient doesn't instead meet any of the cardiometabolic criteria, the name can be either Cryptogenic SLD or other specific aetiology SLD [5].

The accumulation of fat within hepatocytes can progressively disrupt liver function, leading to increasingly severe health consequences. As illustrated in figs. 1.1 and 1.2, metabolic dysfunction-associated steatotic liver disease (MASLD) can advance to a more severe state marked by presence of hepatic steatosis and inflammation with hepatocyte injury (ballooning) with or without fibrosis, known as non-alcoholic steatohepatitis (NASH) [6]. Up to this point, the

disease progression is generally reversible, particularly before fibrosis develops. Liver fibrosis occurs when chronic injury or inflammation triggers the formation of excess connective (scar) tissue in the liver. Over time, fibrosis can advance to cirrhosis, where the scar tissue significantly replaces healthy liver tissue, impairing the liver's structure and function. This stage represents permanent liver damage, especially when combined with NASH, as the liver loses its ability to regenerate. In some cases, cirrhosis may further progress to hepatocellular carcinoma (HCC), a form of liver cancer, which is a leading cause of death, affecting 46% of patients in advanced stages.

The development of steatosis is closely linked to alterations in the metabolic processes of hepatocytes, which result in increased lipid production. The most widely used molecule for storing metabolic energy is triacylglycerol (TAG, or TG), composed by a glycerol head attached to three fatty acid chains. Elevated concentrations of free fatty acids (FFAs) in the blood plasma, due to diet or other reasons, being them either modifiable or not, causes an enhanced uptake of them by the liver, later responsible of their esterification. On the long term this can lead to the accumulation of TAG into lipid droplets (LDs). Lipid droplets are organelles responsible for storing energy in many organs, since virtually all cells have the capacity to synthesize and store TAG in cytosolic lipid droplets (LDs) [8].
Lipid droplet dynamics is further influenced by the presence of enzymes that cooperate for their degradation into smaller droplets, or the regulatory surface proteins (RSPs), which compete to bind to the LD surface, modifying the enzymes activity.
Nevertheless, even though many of the biochemical reactions described remain unclear to the research community, the authors, drawing on experimental $in-vitro$ and $in-vivo$ findings, attempt to develop a model that aligns with the data.

The only therapies available at the moment are the reduction of caloric intake through a healthy diet, physical exercise and medications to reduce the glucose level in the blood. Goal of research is thus try to understand the cellular and molecular processes leading to perturbation of liver lipid metabolism, necessary in order to try to identify novel pharmacological targets and treatment strategies. In my internship project the goal was to try to implement a numerical twin for the hepatocyte metabolic network, based on the model first described in the paper out in 2017 by Wallstab, Berndt *et al.*, explained in detail in Chapter 3, after a brief overview of the biochemistry elements needed to have a possibly good comprehension of it, presented in Chapter 2. Its first section, Sec. 2.1, has two paragraphs on lipids in general and lipid droplets,

followed by some information on enzymes-mediated reaction kinetics and liver cell compartments respectively in Sections 2.2 and 2.3.

In Chapter 3 both the modules of the reproduced model are analyzed in their details, with attached the equations used for each step leading to the synthesis of TAGs inside the hepatocytes endoplasmatic reticulum and then to its effect on the lipid droplets filling. The model consists of a set of time-dependent differential equations. While the structure is not inherently complex, the sheer number and scale of the equations, encompassing various processes and subprocesses, make it highly intricate.

Then, the methods used for the simulations are highlighted in Chapter 4 step by step, followed by the results obtained in Chapter 5. These regard the two approached adopted to solve the problem: first the `Python` script was coded from scratch and used to numerically integrate the ODE system given the initial conditions, then performing a sensitivity analysis revealing that variations in the activity of FFA uptake, diacylglycerol acyltransferase (DGAT) 2, and adipose triglyceride lipase (ATGL) have the strongest influence on the cellular TAG level. A neural network is trained over some randomly generated configurations of these varying parameters to predict the stable state outcome of the system.

Concluding remarks are reported in Chapter 6, highlighting the good performances compared to the results in the paper [1] both for rat and human hepatocytes along with some of the issues encountered and observations that arose during the months of my internship.

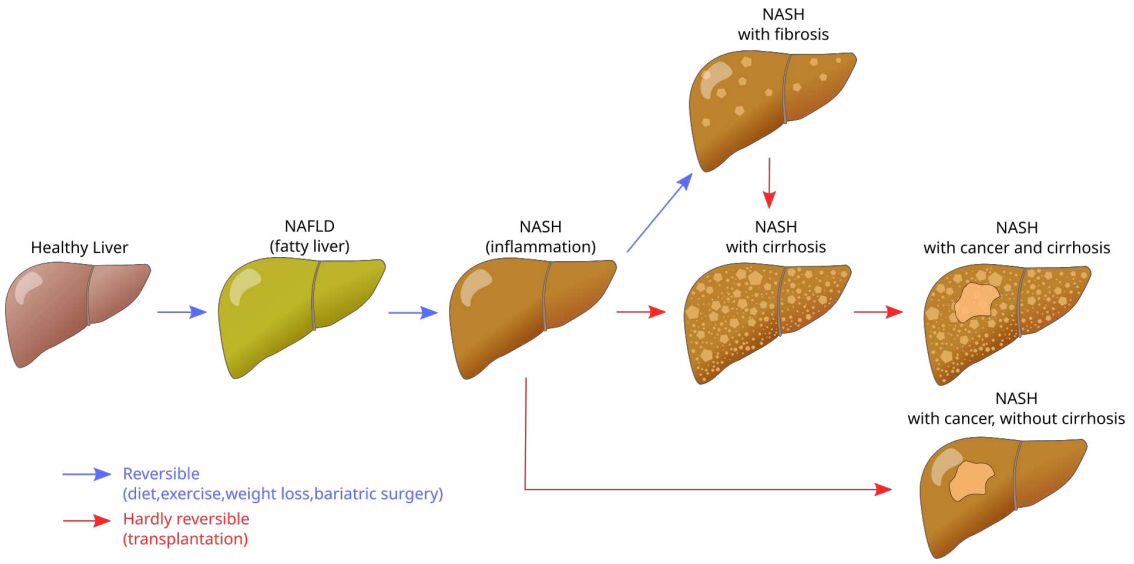


Figure 1.2: All the stages of liver disease progression from NAFLD to HCC. The blue and red arrows indicate whether the transition is generally reversible or difficult to reverse. Image from [9].

2 Elements of Biochemistry

To build a solid understanding of the biochemical reactions and metabolic pathways discussed in the next chapter, this section provides a simplified overview of the topic. The goal is to offer a general introduction from a non-expert perspective, focusing on the basic concepts to help grasp the underlying processes without delving into intricate technical details.
The chapter is divided in four sections: in the first one, Section 2.1, the lipids macro-category is briefly introduced, focusing on the family of storage lipids. Afterwards, Section 2.1.2 resumes the concept of lipid droplets, specialized organelles, sharing some structural and functional patterns with their progenitors, while Section 2.2 introduces the main chemical reaction scheme involved in the liver cell metabolism. Finally the liver cell compartments appearing in the model are announced and their role briefly described in Section 2.3.

2.1 Lipids

Lipids are a diverse and extensive group of organic compounds that play crucial roles in human physiology. Their biological function can be as diverse as their chemistry: while fats and oils are primarily responsible for energy storage in many organisms, phospholipids and sterols are the majors structural elements in membranes. Others play a crucial role as enzymes, catalyzing specific reactions and influencing various biochemical pathways. The classification of lipid families, according to Nelson and Cox [10], is presented below, with example structures illustrated in Fig. 2.1.

- storage lipids (fats, oils), they are neutral;
- membrane structural lipids (phospholipids, sterols) are polar;
- enzymatic lipids

Generally, the storage and structural functional classes of lipids are major cellular components, making most of its part in terms of mass but playing a passive role in the cell. Other groups of lipids, present in much smaller amounts, have active roles in the metabolic traffic as metabolites and messengers. Some serve as signals, as hormones, carried in the blood from one tissue to another, generated in response to an extracellular signal [10].

2.1.1 Storage lipids

This section delves mostly into the first category presented, the storage lipids, starting with an exploration of how they are actually made. Essentially, all naturally occurring fats and oils, independently of their origin, consist of a very high percentage of glycerolipids [11]. They include monoacylglycerols (MAG), diacylglycerols (DAG), and triacylglycerols (TAG), which are central to the energy metabolism. MAG, DAG and TAG are composed by an alcohol, glycerol, attached to either one, two or three fatty acid chains. This characteristic structure, with a "head" attached to multiple "tails," gives these molecules their neutral polarity, distinguishing them from polar membrane lipids. As the polar hydroxyl groups of glycerol and the carboxyl groups of fatty acids are bound by ester linkages, triacylglycerols become non-polar and thus hydrophobic, rendering them essentially insoluble in water.

Lipid metabolism, which will be examined step-by-step in Sections 3.1 and 3.2, encompasses the processes involved in lipid synthesis and degradation within cells, leading to the consump-

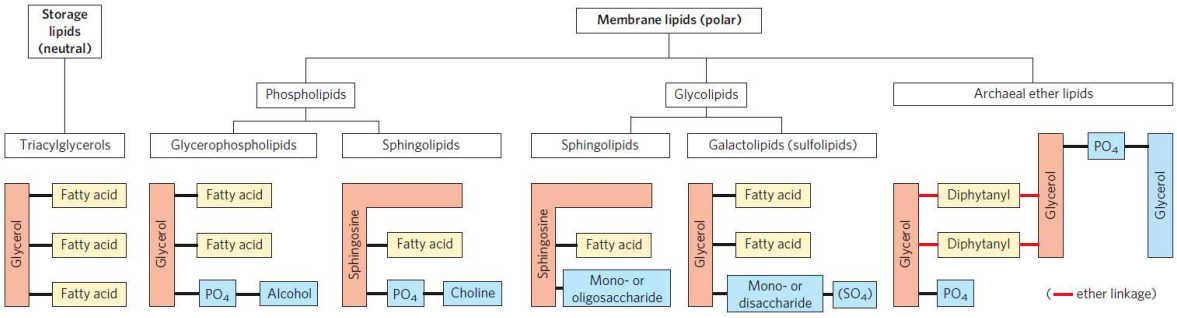
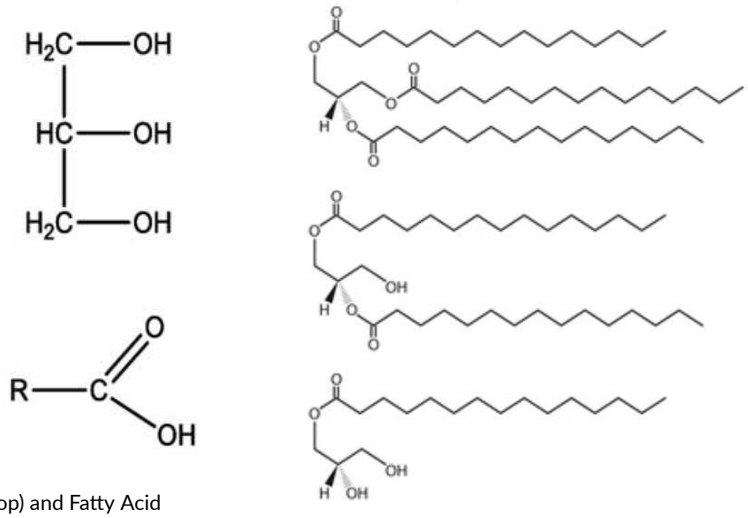


Figure 2.1: Some of the most common types of storage lipids (neutral) and membrane lipids (polar), all having either glycerol or sphingosine as backbone to which are attached one or more fatty acid chains. Original picture from [10].



(a) Glycerol (top) and Fatty Acid (bottom) skeletal structures. **(b)** Three categories of neutral lipids. Picture taken from [11]

Figure 2.2: Composition and examples of neutral lipid chains. 2.2b illustrates the three types of neutral lipids that can be formed using glycerol and fatty acids (2.2a): triacylglycerol, diacylglycerol, and monoacylglycerol, containing three, two, and one fatty acid chain, respectively.

tion or production of free fatty acids. These processes include fat storage and breakdown for energy, as well as the synthesis of structural and functional lipids, such as those that form cell membranes. Key components of lipid metabolism include digestion, absorption, transport, storage, catabolism, and lipid biosynthesis. Lipid catabolism, in particular, occurs through beta-oxidation, a crucial process in the mitochondria and peroxisomes that breaks down fatty acids to generate energy.

2.1.2 Lipid droplets

Lipid droplets are dynamic, lipid-rich organelles present in most eukaryotic cells and serve as the main storage sites for neutral lipids. While they are most commonly found in adipose tissue, this work focuses on their presence and role in hepatocytes. The relatively recent discovery of proteins on the lipid droplet surface, which regulate their dynamics and metabolism, has sparked significant interest in studying these organelles, as they play a crucial role in the regulation of metabolic processes.

Lipid droplet structure is characterised by a neutral lipid core consisting mainly of triacylglycerols (TAGs, but also MAGs and DAGs) and sterol esters (SEs) surrounded by a phospho-

lipid unilamellar leaflet (monolayer). The alignment of the phospholipids allows to compose a membrane with inner hydrophobic and outer hydrophilic property, due to their head and tail classical structure derived from the fatty acids [12, 13, 14].

Disregulated LD homeostasis is closely connected to many pathological conditions and the development of various diseases. The central function of LDs is to store TAG and CE if synthesis and/or uptake of FFAs and SEs exceeds the cellular demand and to release these lipids in the opposite case. As saturated FFAs and free SEs are potentially cytotoxic, their conversion into non- toxic TAG and SE and subsequent deposition in LDs are crucial for the maintenance of cell integrity. An overweening delivery of FFAs and/or SE is commonly accompanied by an increase in the synthesis of TAG and SE and thus in the number and size of LDs. Upon energy demand, these FAs are mobilized by specific enzymes, called lipases, which consecutively hydrolyze esterified FAs from the glycerol backbone, thereby generating DAG, MAG, and finally free glycerol. FAs can then enter the circulation and are accessible for other tissues, which can use them for oxidative energy production or lipid synthesis [11, 13, 15].

Among the large array of regulatory surface proteins (RSPs) decorating the LD surface, the most studied are the those belonging to the Perilipins and CIDE families. Perilipins, also known as lipid droplet-associated proteins (PLINs) are mainly five and they all share a certain PAT domain [16]. In the kinetic model of Chapter 3, only the first three will be incorporated into the equations governing lipid droplet surface binding and unbinding. Beside this, also cell death-Inducing DNA Fragmentation Factor Alpha (DFFA)-like Effector (CIDE) proteins have emerged as lipid droplet-associated proteins that regulate fat metabolism. There are three members in the CIDE protein family: CIDEA, CIDEB, and CIDEC. The last one is the only one considered here for its crucial role in the LDs fusion, and it is also known as fat-specific protein 27 (FSP-27) [17, 18].

- Perilipins: PLIN1, PLIN2/ADRP, PLIN3/TIP47, (PIN4, PLIN5)
- CIDE proteins: (CIDEA, CIDEB), CIDEC/FSP-27

While in adipocytes LDs can reach diameters up to $100\mu m$ and considering that size varies tremendously, in liver cells they are smaller, in general ranging from few to $3\mu m$. As we'll see in Chapter 3 their formation is related to the presence of neutral lipids in the endoplasmatic reticulum, with *de novo* genesis of droplets off the ER lumen of filling of pre-formed droplets of different sizes. The size of newly generated droplets can then grow or shrink based on the RSPs reactions happening on their surface [13].

2.2 Enzyme kinetics

The model has an initial module, described in 3.1, in which each step is described by a biochemical equation, and in particular it is the case of enzymes-catalysed reactions of one substrate and one product. The simplest way to model this type of reactions is by using the Michaelis-Menten equation, the kinetics model named after Leonor Michaelis and Maud Menten. It takes the form of a differential equation describing the reaction rate v, equal to the rate of formation of the product P ($v = \frac{d[P]}{dt}$) given the concentration of the substrate S as $[S]$.

In this section, following the approach outlined in Chapter 6.3 of *Lehninger - Principles of Biochemistry* [10], we will examine the assumptions underlying the derivation of the formulas. Subsequently, the reaction rates used by the authors for the metabolic processes will be derived and discussed. This approach, determining the rate of a reaction and how it changes in response to changes in experimental parameters, is a discipline known as enzyme kinetics.

We start by considering a reaction between a catalyst, it can be an enzyme a protein or something else (E), and a single substrate (S). The reaction can be split in three different phases, as we can see from the reaction scheme 2.1: from enzyme+substrate to enzyme-substrate complex in a relatively fast reversible step (k_1, k_{-1}), and then to enzyme+product , where the transitions are governed by specific rates indicated with k:

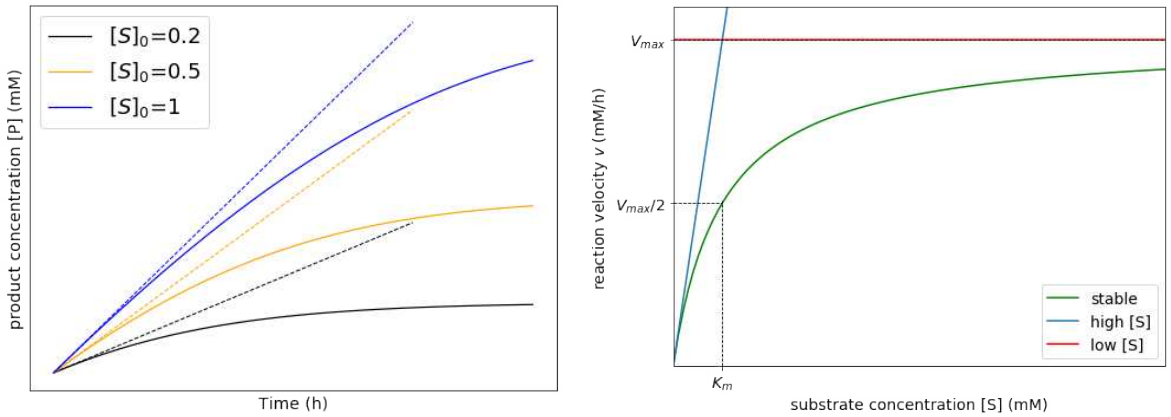
$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{E} + \mathrm{P} \tag{2.1}$$

At any time instants we have the enzyme available both in its free and substrate-bounded forms: the total enzyme concentration catalyzing the reaction will be rewritten then as $[E_{tot}] \coloneqq [E] + [ES]$.

Here we develop the basic logic and the algebraic steps in a modern derivation of the Michaelis-Menten equation, which includes the assumptions below here:

step 1 we can initially neglect the reaction $P \rightarrow S$ described by k_{-2}, since [P] is low at early stages;

step 2 if the concentration of the substrate is in excess in few milliseconds it is concerted to ES and [ES] is approximately constant over time (steady state kinetics). We have so that the rate-limiting step of this type of reaction giving a product (P) is the enzyme-substrate



(a) Initial velocities of enzyme-catalyzed reactions.

(b) Effect of substrate concentration on the velocity of an enzyme-catalyzed reaction.

Figure 2.3: Curves plots for the enzyme-catalyzed reactions.
2.3a A theoretical enzyme catalyzes the reaction S⇌P, and is present at a concentration sufficient to catalyze the reaction at a maximum velocity, V_{max} , of 1 M/h. The Michaelis constant, K_m (explained in the text), is 0.5 mM. Progress curves are shown for substrate concentrations below, at, and above the K_m. The rate of an enzyme-catalyzed reaction declines as substrate is converted to product. A tangent to each curve taken at time equal to 0 (dashed lines) defines the initial velocity of each reaction. The reaction velocity reaches a plateau after an initial pre-steady state period after which [ES] is approximately constant.
2.3b This graph shows the kinetic parameters that define the limits of the curve at high and low [S], with the substrate concentration at which the velocity is half-maximal being K_m, the Michaelis constant. At low [S] ($K_m \gg$[S]) the [S] term in the denominator of the Michaelis-Menten equation 2.5 becomes insignificant, simplifying to $v = V_{max}/K_m$ and it exhibits a linear dependence on [S], as observed here. At high [S] ([S]$\gg K_m$) the K_m term in the denominator of the Michaelis-Menten equation becomes insignificant and the equation simplifies to $v = V_{max}$, consistently with the plateau observed. The concentration of enzyme in an experiment such as this is generally so low that [S] » [E] even when [S] is described as low or relatively low.

(ES) complex breakdown (k_2), thus the overall rate must be proportional to the concentration of the species that reacts in the second step, that is, ES.

Since [ES] is not easily measurable experimentally we can try to rewrite it. Having neglected k_{-2} we can define the ES rates of formation and breakdown, using the variable for the total enzyme concentration previously defined:

$$v_{ES}^{form} = k_1[E][S] = k_1([E_{tot}] - [ES])[S] \tag{2.2}$$

$$v_{ES}^{break} = k_{-1}[ES] - k_2[ES] \tag{2.3}$$

Now according to the reasoning in step 2 on the stable-state one can equal the two rates 2.2 and 2.3, and after some math it gets to the following expression for the concentration of the

enzyme-substrate complex

$$[ES] = \frac{k_1[E_{tot}][S]}{k_1[S] + k_{-1} + k_2} = \frac{[E_{tot}][S]}{K_m + [S]} \tag{2.4}$$

where K_m is the Michaelis constant, defined as $K_m := (k_{-1} + k_2)/k_2$.

We can now show how the equation describing the rate/velocity of these types of reaction, commonly known as Michaelis-Menten equation, is determined by the breakdown of ES to form product, which is determined with [ES] in eq. 2.4:

$$v = k_2[ES] = k_2 \frac{[E_{tot}][S]}{K_m + [S]} = V_{max} \frac{[S]}{K_m + [S]} \tag{2.5}$$

where the maximal reaction velocity is defined as $V_{max} := k_2[E_{tot}]$.

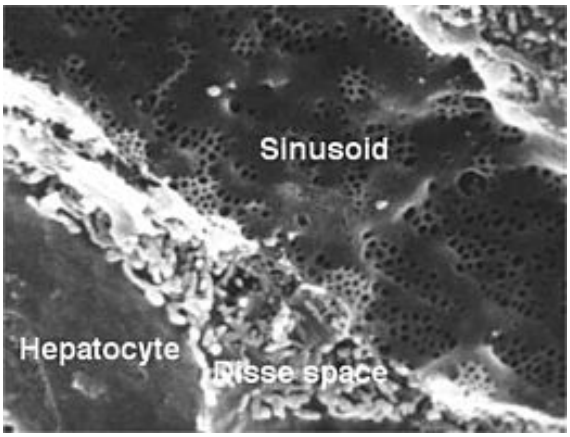
This is finally the Michaelis-Menten equation, the rate equation valid for a single-substrate enzyme-catalyzed reaction, which has the advantage of being easily extendable to describe reactions involving multiple substrates. It is useful to define a more general rate constant, the turnover number k_{cat}, to describe the limiting rate of any enzyme-catalyzed reaction at saturation. When several steps are partially rate-limiting, k_{cat} can become a complex function of several of the rate constants that define each individual reaction step. It is equivalent to the number of substrate molecules converted to product in a given unit of time on a single enzyme molecule when the enzyme is saturated with substrate. In the Michaelis-Menten equation, $k_{cat} = V_{max}/[E_{tot}]$ and equation 2.5 becomes

$$v = \frac{k_{cat}[E_{tot}][S]}{K_m + [S]}. \tag{2.6}$$

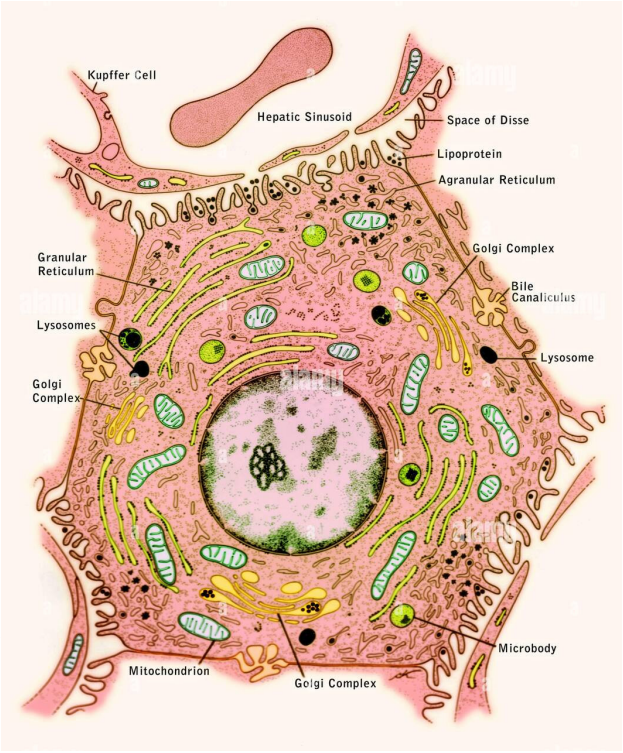
For a more detailed and in-depth analysis of these topics, it is recommended to refer to Chapter 6 of Lehninger's book [10]. This chapter provides additional insights into enzyme mechanisms, including detailed examples of enzyme kinetics.

2.3 Liver cell structure

Hepatocytes, the primary functional cells of the liver, are highly specialized cells constituting approximately 70-80% of the liver's mass and are responsible for a wide array of essential physiological processes. They are typically cuboidal, with dimensions ranging between 20 and 30 μm,



(a) Scanning electron micrograph (SEM) of a rat sectioned hepatocyte and venule sinusoid, the blood provider. Image from public domain (*Wikipedia*).



(b) Labeled and illustrated diagram of the ultrastructure of a liver cell, highlighting the relationships among all previously discussed compartments, along with several additional ones. Image licensed from (*Alamy*).

Figure 2.4: Hepatocytes structure with all the internal organelles and its location in liver tissue, spearated to the sinusoid blood provider by the Disse space.

and taking an input of fatty acids from the sinusoid's blood flow. A section is showed in Figure 2.4a.

They are remarkable for their regenerative capacity, allowing the liver to recover from injury or disease and they have different compartments, but as not to dwell too much on further side topics, here only parts of interest for the model are taken into account here:

- **cytosol**
  The intracellular fluid that houses various organelles and is the site of many metabolic pathways, including glycolysis, fatty acid synthesis, and the initial steps of protein synthesis, and it also contains lipid droplets;

- **endoplasmatic reticulum**
  A network of membranous tubules involved in protein and lipid synthesis. The rough ER, studded with ribosomes, synthesizes proteins destined for secretion or membrane

insertion. The smooth ER is involved in lipid metabolism, detoxification of drugs and toxins, and calcium storage;

- **mitochondria**
  The powerhouse of the cell, where ATP (energy) is produced through oxidative phosphorylation. Mitochondria are also involved in fatty acid oxidation, the urea cycle, and the regulation of apoptosis (programmed cell death), which is crucial in liver disease progression;

- **nucleus**
  Hepatocytes have a large, centrally located spherical nucleus that serves as the control center of the cell, containing DNA that regulates metabolism, growth, and protein synthesis, all of which are essential for liver function;

- **plasma membrane**
  Regulates the exchange of substances between the hepatocyte and its environment, playing a key role in nutrient uptake and waste removal.

In the equations in Appendix A.2, footnotes indicate the compartment in which the metabolites are present or through which are being transported.

3

Wallstab *et al.*'s mathematical model

This work focuses on the reproduction of the Wallstab *et al.*'s model presented first in their paper out in 2017 [1]. The same model was then reused for studying the behaviour of other hepatocyte intracellular pathways, in particular with a focus on hepatocellular carcinoma (HCC) [19] and then also extended to comprise the vast amount of metabolic liver pathways in some more complex models, again under the supervision of Nikolaus Berndt [20].

To gain a clearer understanding of the regulatory influence of individual processes involved in cellular TAG turnover, we reproduced a comprehensive kinetic model that integrates the pathways of fatty acid and triglyceride metabolism, as well as the key molecular mechanisms governing lipid droplet dynamics.
This chapter is divided in two main sections to clarify the two modules of the LDs formation and evolution: in the first one, Section 3.1, the metabolic module of the model is described, containing all the reactions composing the chain leading to the TAGs formation inside the ER, then the lipid droplets undergo filling with TAG and fusion to each other, with proteins binding to their surfaces causing also lipid degradation and consequently the shrinking of the droplets. These steps collectively form what is referred to as the lipid module of the model, presented in Section 3.2.

The molecular processes included into the model developed and implemented by Wallstab *et al.* are depicted in Fig. 3.1 and resumed in Tab. 3.1. In the followings, numbers inside brackets $\langle\rangle$ refer to the indexing of processes, showing also their responsible catalysts in the table.

3.1 The metabolic module

In this first part the free fatty acids uptaken from the external blood flow are converted to the neutral lipids that will fill the droplets in the next module. The quantity of free fatty acids (FFAs) absorbed by the hepatocyte model depends on the glucose level of the blood plasma: elevated blood glucose levels indicate a fed state, while lower levels suggest the rat is in a fasted state. The relations linking the uptakes FFAs are explained in Appendix A.1, with also the insuline and glucagon level computation in attached.

The further processes leading to the LDs formation are considered in all their details, accompanied by the reaction rates associated to the transitions, showed in the Appendix A.2 for the sake of clarity in the main text here.

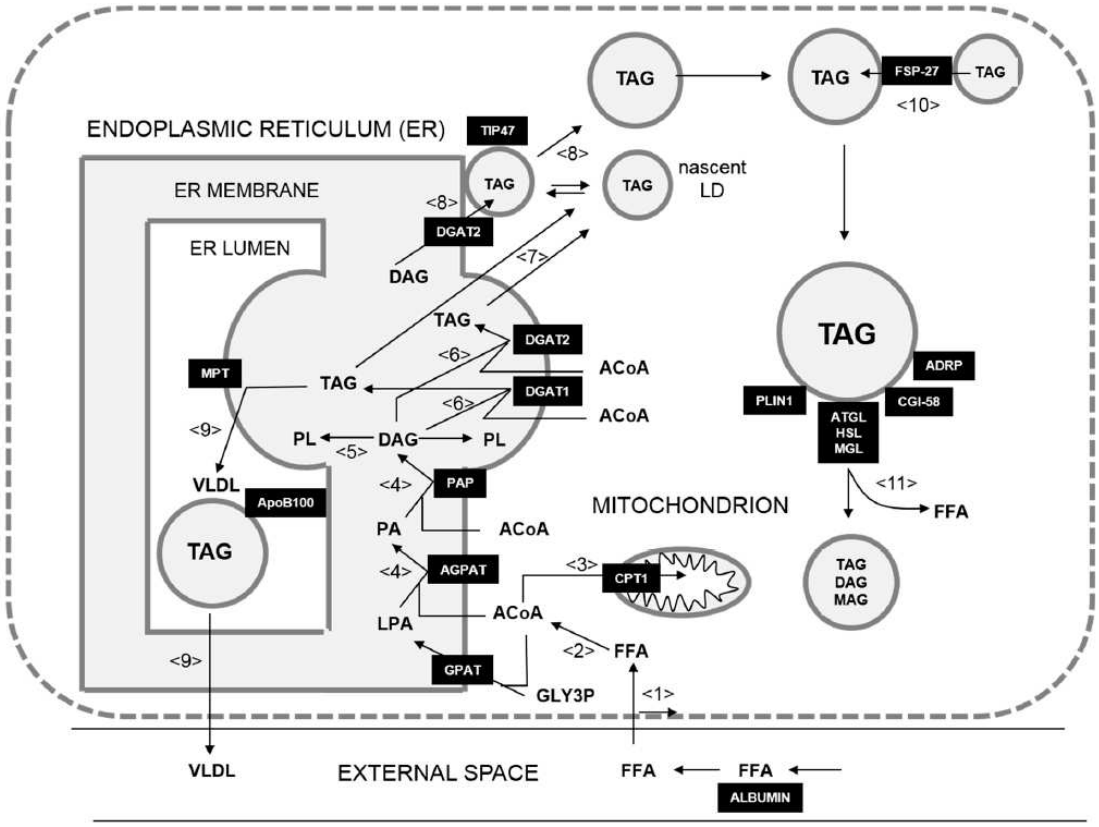


Figure 3.1: Model originally introduced in [1], representing the main reactions of the metabolic and lipid modules, from the uptake of FFA from the blood to the formation and finally degradation of LDs. The numbered reactions in the brackets are defined in Tab 3.1. Picture taken from [19].

In this module, namely above the horizontal line in Table 3.1, all the necessary steps of the TAG production are broken down to single biochemical reactions. As foretold, the free fatty acids are uptaken into the cytoplasm of the hepatocyte $\langle 1 \rangle$ through a combination of diffusive and active (carrier-mediated, by transport protein CD36) transport. FFAs are then activated to acyl-CoA $\langle 2 \rangle$ (in figure as ACoA), which is a composition of FFA and CoA obtained through interaction. The membrane carrier CPT1 is responsible for the acyl-CoA entering the mitochondria $\langle 3 \rangle$, using a 3-step reaction where FFA first binds to Carnitine allowing taken up into mitochondria and then reforming acyl-CoA inside the mitochondria and bringing back Car outside. The mitochondrial β-oxidation is responsible for the energy production inside cells, through the citric acid cycle.
The metabolic pathway leading to the formation of TAG starts with the esterification of acyl-CoA with G3P by GPAT producing LPA, whose is consequently esterificated by AGPAT $\langle 4 \rangle$ newly forming the phosphatidic acid PA inside the endoplasmatic reticulum. Now PAP acts in order to favour PA dephosphorilation (hydrolysis) to produce DAG. The DAG synthesised in the ER is responsible for different pathways, our interest focuses on two of them here. First, PC phospholipids synthesis $\langle 5 \rangle$ by CPT, later used for the formation of lenses (nascent droplets) in the ER membrane. Second, but most importantly, DAG is used to synthesize TAGs inside the ER through the interaction with the fatty acid of acyl-CoA, having the reactions catalyzed by two isoforms of the diacylglycerol acyltransferases, DGAT1 and DGAT2 $\langle 6 \rangle$. Even though the pictorical representation in Figure 3.1 shows TAG forming in the cytosol outside the ER, other sources ([12, 13]) tell us that this is a result achieved in a series of step but whose exact dynamics is not fully understood yet. For the sake of simplicity and not to add other variables to the model, this will be considered as a simplified model for the actual process which involves lenses formation and irregular growing of extroflexions on the ER membrane. At this point we have two "storages" of TAG inside the endoplasmatic reticulum: $\text{TAG}_{er(i)}$, with $i = 1, 2$ both serving as a TAG source for nascent LDs formation, nevertheless $\text{TAG}_{er(1)}$ is also used in the VLDL synthesis, mediated by the presence of ApoB, whose availability is insulin dependent. VLDLs are later synthesised and expelled back to the blood $\langle 9 \rangle$.

3.2 The Lipid Droplet module

This module comprises central processes to control the dynamics of lipid droplets (LDs) and the export of the lipoprotein VLDL $\langle 9 \rangle$. All the different physical-biochemical processes involved in the LD modeling from this point on are presented: the level of TAG contained in-

	process	involved enzime/protein
$\langle 1 \rangle$	FFAs uptake into cytosol	passive diffusion, CD36
$\langle 2 \rangle$	FFAs activation	ACS
$\langle 3 \rangle$	FFA-CoAs uptake into mitochondria	CPT1
$\langle 4 \rangle$	FFA-CoAs esterification	GPAT, AGPAT, PAP
$\langle 5 \rangle$	PC synthesis	CPT
$\langle 6 \rangle$	TAG synthesis	DGAT1, DGAT2
$\langle 7 \rangle$	nLD formation	TIP47
$\langle 8 \rangle$	LDs filling	DGAT2
$\langle 9 \rangle$	VLDL synthesis and secretion	MTP
$\langle 10 \rangle$	LDs fusion	FSP27
$\langle 11 \rangle$	LDs degradation	ATGL, HSL, MGL

Table 3.1: Main processes and involved enziymes appearing in the model. All the equation rates for these processes, along with some others, are showed in Appendix A.2.

side each lipid droplets has five different contributes that are the LD genesis $\langle 7 \rangle$ and filling $\langle 8 \rangle$, fusion $\langle 10 \rangle$, RSP binding and degradation processes $\langle 11 \rangle$, respectively in paragraphs sections 3.2.1 to 3.2.4. Finally also the LD dynamic change in number is presented in 3.2.5, explaining how the number of droplets is computed.

Here is where for the first time modeling with LDs is encountered. When dealing with LDs, the approach followed by the authors is to subdivide LDs originally into $N_c = 30$ discrete size classes based on the radius of the droplets ranging from 0.1 to 3.0 μm. Taking this into account, both the concentrations of lipids in the droplets and the fractions of LD covered by RSPs or enzymes are described by vectors with the dimension equal to N_c.

3.2.1 Genesis and filling

The first process contributing to the growth of lipid droplets is related to the genesis of new droplets. This process can happen both in a *de novo* scenario, where lipid droplets are newly generated, or adding TAGs to already existing droplets. We'll refer to the first case as a *de novo* synthesis of LDs $\langle 7 \rangle$, called nascent Lipid Droplet (nLD), while in the latter the process is known as a filling process, where some already existing lipid droplets formed in the cytosol and binding to the ER, are filled with the produced neutral lipids again through the enzyme DGAT2 $\langle 8 \rangle$ and esters.

The equations describing the filling of LDs with TAG are responsible for the partition of TAG

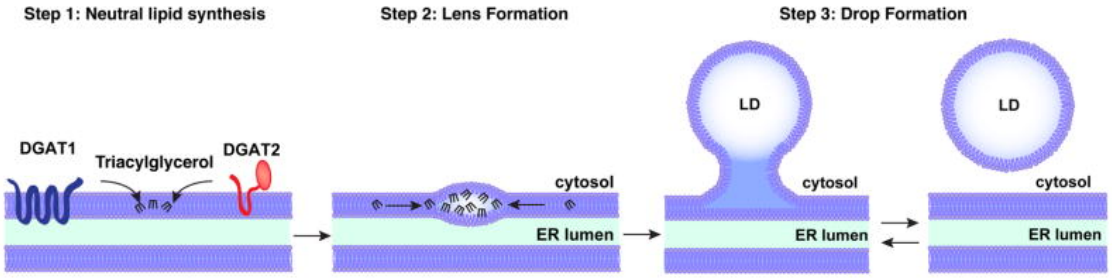


Figure 3.2: The synthesis of nascent Lipid Droplets at the interface of cytosol with ER. Picture taken from [14].

content among ER(2) and LDs: the higher is the number of droplets the higher is the rate of LDs filling with TAG and the lower is the nLD formation (eqs. (A.44) and (A.45)). The equation A.45 clarify the concept that larger droplets are less likely to associate with ER an be filled with TAG [14, 15, 21].

The synthesis of nascent LDs is catalyzed by the protein TIP47 (PLIN3) binding to the ER and LD interacting surface. Studies show that both DGAT1 and DGAT2-produced TAG is involved in *de novo* nLDs formation, with neutral lipids coolescence in the leaflet (lenses) formation in between the ER bilayer, as can be seen in Figure 3.2. This is due to the asymmetry in the phpspholipid composition, protein recruitment and binding [12, 14]. All the nLDs produced by $de\ novo$ synthesis belong to the first lipid droplet class, namely with radius $r_1 = 0.1\mu m$, and they are not going to be covered by any regulatory surface protein (RSP).

3.2.2 Fusion

On the top of that, LDs can interact one with the others through a fusion process $\langle 10 \rangle$ mediated by the fat-specific protein 27, FSP27. The rates of this process are in the Appendix A.2, having that the larger particle is absorbing the smaller one and the rate of their fusion is dependent on the difference in ratio between the two particles. The rate of lipid transfer (eq. 3.1) depends on the size difference between lipid droplets and the presence of FSP27 on the surface of the fusing droplets [22]. The net effect of this term on the total LDs volume is null, it has only an effect on their size distribution, shifting it towards a configuration with more bigger droplets in case fusion is favored, e.g. for the high availability of FSP27 [14, 23].

$$fr_{ij} = V_{max}^{fus} \cdot [FSP27]_{cyt} \cdot num_{LD_i} \cdot num_{LD_j} \cdot \left| \frac{1}{r_{LD_j}} - \frac{1}{r_{LD_i}} \right| \tag{3.1}$$

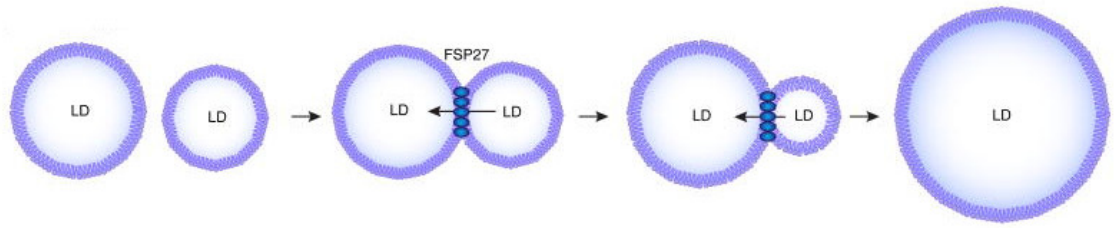


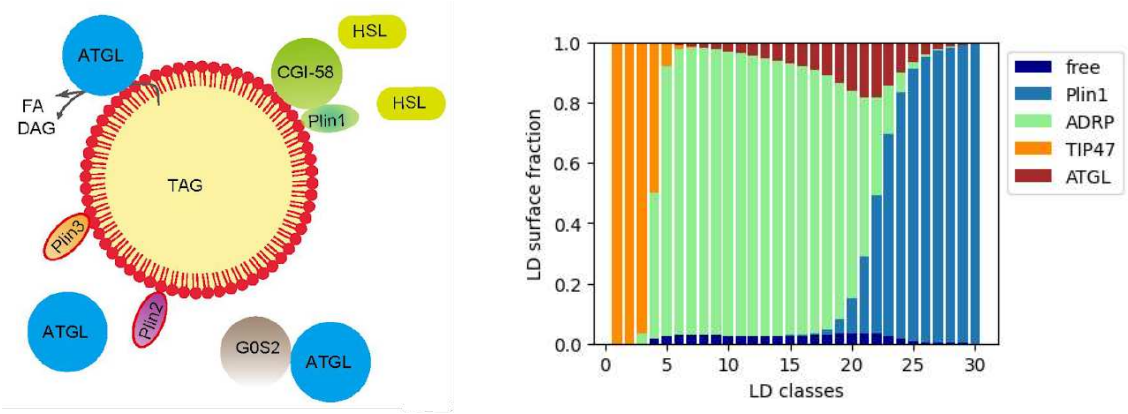
Figure 3.3: All the steps required in the fusion of two Lipid Droplets mediated by FSP27. Picture taken from [14].

3.2.3 RSPs on LDs surface

The surface of lipid droplets is a place of hectic binding and unbinding for different families of proteins and enzymes. These have an effect on the dynamics of the droplets, affecting its shape, and possibly favouring either their shrinking or their growth [13, 16, 24, 25].
Reversible binding of regulatory surface proteins (RSPs) to LDs is modeled as time-dependent change in the LD surface fraction, X_{LD}, covered by the protein X, with X=PLIN1, ADRP, TIP47 or ATGL. Equation 3.2 here below is responsible for the rates modeling of this binding-unbinding process, where f_{LD}^{free} denotes the non occupied free LD surface, while the rates k_{ON} and k_{OFF} are dependent on the cytosolic concentration of the protein X.

$$\frac{dX_{LD}}{dt} = k_{ON}([X]_{cyt}) f_{LD}^{free} - k_{OFF}([X]_{cyt}) X_{LD} \tag{3.2}$$

Different proteins of the PAT family (PLIN1, ADRP, TIP47) have different binding and unbinding rates depending on some factors defined basing on radius, constants and exponent of the LDs (see Appendix A.2). This is resulting in what can be observed in Fig. 3.4b, showing how PLINs 1, 2 and 3 are likely to bind respectively on the surface of large, medium and small LDs, while ATGL has rates which are depending on the number of LDs in each discrete class, and its steady basal activity is significantly increased due phosphorylation of PLIN1, liberating CGI58 that then binding to ATGL to activate it [16].



(a) Binding of various species on the lipid droplet surface, including regulatory surface proteins (PLIN1, PLIN2, PLIN3, CGI58) and enzymes (HSL, ATGL). Image from [16].

(b) Distributions of the percentage of lipid droplet surface occupied by different RSPs: PLIN1, ADRP, TIP47, and ATGL. In each class, their presence sums to 1, representing the total available surface of the droplet.

Figure 3.4: Lipid droplets and their associated binding proteins: the various species (3.4a) and their expected distributions across different size classes (3.4b) according to the data provided in the *Supplementary Information* of the reference paper [1].

Since the cytosolic concentration of RSPs has proven to be very difficult to determine and define in time, lot of simulations have been run using 1 as reference value for these concentrations, varying then the value in simulations for testing the effect of over-expression and under-expression on the final LDs distribution.

PLINs are sort of competitors for binding to the LD unoccupied surface, while CGI58 is treated as a co-regulator and it binds to the LD surface through mediation of PLIN1 and ADRP. Also enzymes responsible for TAG degradation inside the LDs have to bind to the droplet surface to start their activity: this is the case for ATGL. A deficiency in RSPs necessary for the de novo synthesis of lipid droplets, or the overexpression of RSPs that promote lipid droplet hydrolysis, can both reduce the number and size of lipid droplets.

3.2.4 Shrinking

LD-associated enzymes (ATGL, HSL, MGL) are responsible for the disruption of LDs content into simpler molecules, namely free fatty acids. The main process responsible for the shrinking of LDs is the removal of fatty acids from TAG, DAG and MAG inside the core of the lipid dropplets. This is known as the LDs degradation ⟨11⟩ and is divided in different steps: first ATGL, and HSL also with a minor role, performs the first step from TAG to DAG, then HSL breaks down the DAG into MAG and the process in finalized by MGL leaving FFA in the

cytosol out of the MAGs [13, 15].

These degradation steps are catalyzed by enzymes again binding on the surface of the lipid droplets and thus competing with the RPSs for the available free surface in the game. The degradation process are enlisted in the Appendix A.2 in eqs. (A.104), (A.107) and (A.114) and they consist of 3 main steps through which the TAG is freed into free FAs:

1. TAG $\xrightarrow{ATGL}$ DAG + FA
2. DAG $\xrightarrow{HSL}$ MAG + FA
3. MAG $\xrightarrow{MGL}$ FA

HSL is also in little percentage responsible for the first step of the degradation, from TAG to DAG.

3.2.5 LDs content dynamics

Each lipid class, by radius, is represented by the total amount of lipids contained in that class. To obtain the number the approach followed is to consider the total amount divided by the droplet volumes. The formula 3.3 was developed in the absence of well-established equations on the subject in the original paper [1].

$$LD_n = \sum_{sp} Vol_{sp_n}/Vol_{LD_n} \in \mathrm{R}, \text{ where } sp = \mathrm{MAG, DAG, TAG} \tag{3.3}$$

On the top of that the number of droplets in each class can vary, and was supposed to be modeled by a lipid transfer to the closest bigger LD_{n+1} or smaller LD_{n-1} classes. About this changes in the amount and the composition of LDs, the formula provided by the authors seems to be ineffective, see Appendix A.3. Probably some details are missing and it was thus not included finally in the model simulated for out purpose. There is some inconsistency in the notation here as well as the dynamics here represented seems to overwhelm the natural evolution by providing an either excessive or negligible contribute to the number of LDs that are transferred from/to each class at each iteration.

4 Methods

In this chapter, the various methods employed to solve the system of ordinary differential equations (ODEs) and conduct further analysis are introduced, detailing the techniques and packages or libraries used, all of which were coded using `Python` scripts.
Section 4.1 outlines the packages and the pipeline implemented for the numerical integration scheme, then Section 4.2 presents the idea of sensitivity analysis, performed to find the key system parameters. Lastly, Section 4.3 discusses the development of a neural network model designed to learn the input-output relationship of the ODE system.

4.1 Model simulations

Initially the set of $13 + 12 \cdot N_c$ coupled ODEs (see *Stoichiometric Matrix* A.2.3), have been solved using the `scpiy.integrate.solve_ivp` package, using as the initial conditions the values expected to be close to the reference stable state concentrations of the metabolites (see *Supplementary Data S4* [1]). All the values used for the maximal activities as well as the constants are reported here in the final Appendix A.2.
The resulting variable consist thus of a number of arrays equal to the number of variables used containing 24 values each, given that the stored concentration values are 1 at each hour and the system is simulated for a duration of 24 h, getting to a stable state. The code have been organized using classes having as methods the systems that needs to be solved or other further operations on it and as attributes its constant parameters. Changing parameters are taken as

an external input. This approach allows for easily attempting simulations by modifying one or multiple parameters, while ensuring consistency and flexibility within class method calls.

Many attempts have been made in order to find the optimal solver going from easier integration methods to more complicate ones and also trying to provide the Jacobian of the stoichiometric matrix in order to save computational time. Finally *backward differentiation formula* (BDF) proved to be the most efficient method for stiff-problems, meaning that given the plenitude of variables in the system they can evolve on different time-scales and this adaptive size-step method, defined in order to overcome this kind of possible issues. Beyond this, the system is also very sensitive to the initial conditions, especially the ones regarding the neutral lipid concentration and the number of droplets in each class, so a careful and well-considered initialization is essential to ensure that the simulated dynamics are meaningful and accurate.

Following the reference paper [1], the simulation have been refined using the rat cell parameters, and then extended to the human hepatocyte to try to reproduce some experiments findings, changing some of the identified crucial parameters and constants in this latter case. For this reason most of the results will show two plots, one for the rat refinement simulations, the other for the human observatory results. In doing this it was decided, after simulating the rat metabolites evolution using a number of class $N_c = 30$, to reduce it to only 10. This will give a more coarse-grained solution, but still with a good enough resolution to observe the distribution shape of the droplets size in the human case, where the total volume of lipids produced is greater. The number of variables whose associated ODEs are numerically integrated after this change reduces from 373 to only 133. A further reduction in this number could be attained by fixing some of the less impactful variables. However, this approach proved challenging, as altering a single step in the process could trigger an unexpected cascade effect on the remaining variables.

# eqs	variable	dim.
13	metabolites	1
3	neutral lipids	N_c
8	RSPs	N_c
1	number of LDs	N_c
tot.	$13 + 12 \cdot N_c$	

Table 4.1: Number and dimensionality of the ordinary differential equations in the model.

4.2 Sensitivity Analysis

Sensitivity analysis is the study of how the uncertainty in the output of a mathematical model (numerical or otherwise) can be divided and allocated to different sources of uncertainty in its inputs. A key milestone in the field was the landmark paper by Morris [26], published in 1991. For a detailed timeline of the evolution of sensitivity analysis, refer to the recent work by Tarantola [27]. For a more in-depth exploration of the subject, the well-regarded book by Saltelli *et al.* [28] offers a comprehensive overview of various methods, covering both local and global approaches.

There are different categories of sensitivity analysis methods and different approaches to use them. In here only two methods are presented, belonging the first, Morris, to the family of the screening techniques 4.2.1, while the second, Sobol', to the variance-based methods 4.2.2. Derivative-based ones are also quite diffused, aiming in particular at understanding how small variations in the parameters affect the model output. These methods are very useful in inverse problems and in parameter estimation, but will not be covered here.

The goals of sensitivity analysis are many, some of them are:

- identify the most important parameters (those that contribute the most to output variability) and insignificant parameters (can be eliminated from the model);
- determine if and which parameters interact with each other;
- simplify the model (by fixing model inputs that have no effect on the output);
- test the robustness of a model in the presence of uncertainty;
- detect errors in the model (by encountering unexpected relationships between inputs and outputs);
- reduce uncertainty through the identification of model inputs that cause significant uncertainty in the output;
- determine the optimal regions within the parameters space (for use in optimisation studies).

Sensitivity analysis are applied here as a novel approach in order to confirm, after the paper test varying some constants by $\pm 10\%$, which are the most significant parameters among the huge quantity of the ones retained in the model.

The first step consists of a global or screening sensitivity analysis, qualitatively inspecting and selecting some of the most important parameters, followed by a more qualitative one to catch their real effect through the Sobol'indices.

For the mathematical formulation of how these indices work, we start by considering a model with K parameters, calling them $\vec{x} = (x_1, ..., x_K) \in [0, 1]^K$ independently and uniformly distributed within the unit hypercube. They are mapped onto the univariate output model variable Y by the model function f, such therefore that $f : \vec{x} \in [0, 1]^K \mapsto y = f(\vec{x}) \in \mathbb{R}$.
The code implementation is done using the Sensitivity Analysis library in Python `SALib` [29].

4.2.1 One-step-At-a-Time (OAT) global Morris method

The most used design for screening technique is a One-step-At-a-Time (OAT) method, where each input is varied while fixing the others, giving thus a new value to only one parameter in each run. This is also known as the Morris method [26, 30].
It is fast since not many model execution are required, but this comes at the cost of not being able to differentiate non-linearities with interactions. For this reason it is commonly used as a screening method to qualitatively identify the non-influential parameters of a model, which then could be safely excluded from further detailed analysis, like Sobol' one, presented in the next subsection 4.2.2. The quantities and parameters defined and utilized for this method are outlined and explained here to provide a comprehensive overview of the possibilities it offers.

First of all, the elementary effect EE of the parameter k on the output is defined as follows:

$$EE_k = \frac{y(x_1, ..., x_k + \Delta, ..., x_K) - y(x_1, ..., x_K)}{\Delta} \tag{4.1}$$

where Δ is a grid jump selected in a way such that $\vec{x} + \Delta$ is still in the specified domain of parameter space. Δ is a value in $\frac{1}{p-1}, ..., 1 - \frac{1}{p-1}$, where p is the number of levels that partitions the model parameter space into a uniform grid of points at which the model can be evaluated. The grid constructs a finite distribution of size $p^{K-1}[p - \Delta(p - 1)]$ elementary effects per input parameters [26].
The key idea of Morris method is to evaluate the model from various randomly sampled points $\vec{x}$ and then gradually advancing one grid jump at a time between each model evaluation (one-at-a-time), along a different dimension of the parameter space selected randomly.

At the end of this process let's say n_R elementary effects will have been sampled for the k-th parameter from the EE_k finite distribution. The arithmetic mean and the standard deviation of the elementary effects sampled from all trajectories defines

$$\mu_k = \frac{1}{n_R} \sum_{i=1}^{n_R} EE_k^i \tag{4.2}$$

$$\sigma_k = \sqrt{\frac{1}{n_R} \sum_{i=1}^{n_R} (EE_k^i - \mu_k)^2} \tag{4.3}$$

While 4.2 describes the average elementary effect computing the arithmetic mean, 4.3 gives and indication of the presence of non-linearity and interactions between k-th parameter and the others. As a change in a parameter value might have a changing sign on the output and thus result in a cancellation effect, Campolongo *et al.* [31] proposed the use of the mean of the absolute elementary effect to circumvent this issue. It is defined as

$$\mu_k^* = \frac{1}{n_R} \sum_{i=1}^{n_R} |EE_k^i| \tag{4.4}$$

On a large enough number of trajectories n_R, Morris states that there are three different types of categories to which a parameters could fit into:

- non-influential parameter (small μ_k, or μ_k^*, and small σ_k) - Δ is null:
  the parameter has a negligible overall effect on the model output;
- linear influencing parameters (large μ_k, or μ_k^*, and small σ_k) - Δ is non-null and constant:
  the variation of elementary effects is small while the magnitude of the effect itself is consistently large for the perturbations in the parameter space
- non-linear influencing parameters (large μ_k, or μ_k^*, and large σ_k) - Δ varying with k:
  indicates that the aggregate effect of perturbations is seemingly small while a large value of σ_k indicates that the variation of the effect is large; the effect can be large or negligibly small depending on the other values of parameters at which the model is evaluated. Such large variation is a symptom of nonlinear effects and/or parameter interaction.

This classification enables parameter importance ranking, allowing for the identification and screening of non-influential parameters.

4.2.2 Variance-based global Sobol' method

Variance-based sensitivity analysis methods exploit the variance decomposition in non-linear and non-monotonic cases. The main tools is the Sobol' method (or Sobol' indices, named after Ilya M. Sobol'), decomposing the variance of a model's output into portions that can be attributed to individual inputs or combinations of inputs. The method for using certain indices to estimate the influence of individual variables or groups of variables on model output was first introduced by Sobol' in [32].

The output y of the model can be decomposed as a linear combination of functions with increasing dimensionality:

$$y = f(\vec{x}) = f_0 + \sum_{i=1}^{K} f_i(x_i) + \sum_{i<j}^{K} f_{i,j}(x_i, x_j) + ... + f_{1,2,...,K}(x_1, x_2, ..., x_K) \tag{4.5}$$

where f_0 is a constant and the condition necessary for this decomposition is expressed as

$$\int_0^1 f_{i_1,i_2,...,i_s}(x_1, x_2, ..., x_s)dx_m = 0, \text{ for } x_m = (i_1, ..., i_s) \tag{4.6}$$

$$\text{and } 1 \leq i_1 < i_2 < ... < i_s \leq K (s \in 1, ..., K) \tag{4.7}$$

meaning that the functional decomposition is made such a way to have all the terms orthogonal, expressing the effect of varying x_i alone in f_i, x_i together with x_j in f_{ij} (second-order interaction), can be re-expressed through the equation 4.7 as terms of conditional expected values:

$$f_0 = \mathbb{E}(Y) \tag{4.8}$$

$$f_i(x_i) = \mathbb{E}_{\sim i}[Y|X_i] - \mathbb{E}(Y) \tag{4.9}$$

$$f_{ij}(x_i, x_j) = \mathbb{E}_{\sim i,j}[Y|X_i, X_j] - \mathbb{E}_{\sim i}[Y|X_i] - \mathbb{E}_{\sim j}[Y|X_j] - \mathbb{E}(Y) \tag{4.10}$$

and so on for the other higher-order terms of the decomposition. In eqs. (4.9) and (4.10), $\mathbb{E}_{\sim i}[Y|X_i]$ corresponds to the conditional expectation operator, and the $\sim \circ$ in the subscript means that the integration over the parameter space is carried out over all parameters except the specified parameter in the subscript.

Applying then the variance operator over Y, assuming that f is square-integrable, we get:

$$\mathbb{V}[Y] = \sum_{i=1}^{K} \mathbb{V}[f_i(x_i)] + \sum_{i<j} \mathbb{V}[f_{i,j}(x_i, x_j)] + ... + \mathbb{V}[f_{1,2,...,k}(x_1, x_2, ..., x_K)] \quad (4.11)$$

First-order indices A direct variance-based measure of sensitivity S_i, called the "first-order sensitivity index", or "Sobol' main-effect sensitivity index" is defined as follows:

$$S_i = \frac{\mathbb{V}[\mathbb{E}_{\sim i}[Y|X_i]]}{\mathbb{V}[Y]} \quad (4.12)$$

measuring the contribution to the output variance of the specific input X_i, by averaging the effect of the other inputs on the output. Using this definition in 4.12 and dividing both sides of the previous formula 4.11 by $\mathbb{V}[Y]$ we get that

$$\sum_{i=1}^{K} S_i + \sum_{i<j}^{K} S_{ij} + ... + S_{1,2,...,k} = 1 \quad (4.13)$$

Total-effect index Another measure, proposed by Saltelli [33] is the "Sobol' total-effect index", which measures the contribution to the output variance of x_i, including all variance caused by its interactions, of any order, with any other input variables.

$$S_{Ti} = \frac{\mathbb{E}_{\sim i}[\mathbb{V}_i[Y|\mathbf{X}_{\sim i}]]}{\mathbb{V}[Y]} \quad (4.14)$$

The index, also a global sensitivity measure, can be interpreted as the amount of variance left in the output if the values of all input parameters, except x_i, can be fixed.

These two sensitivity measures can be related to the objectives of global sensitivity analysis for model assessment, having as main advantage the exploration in the space of input and the efficiency in "analyst" time (the method only needs Monte Carlo/quasi Monte-Carlo evaluations).
The main-effect index is relevant to parameter prioritization in the context of identifying the most influential parameter since fixing a parameter with the highest index value would, on average, lead to the greatest reduction in the output variation.
The total-effect index, on the other hand, is relevant to parameter fixing (or screening) in the

context of identifying the least influential set of parameters since fixing any parameter that has a very small total-effect index value would not lead to significant reduction in the output variation. The use of total-effect index to identify which parameter can be fixed or excluded is similar to that of the elementary effect statistics of the Morris method, albeit more exact but also more computationally expensive to compute. And finally, the difference between the two indices of a given parameter (eqs. (4.12) and (4.14)) is used to quantify the amount of all interactions involving that parameters in the model output.

4.3 Learning the input-output relation

Once the model is confirmed to be running smoothly and yielding satisfactory results, the focus can shift to assessing whether the performance achieved with a specific parameter configuration can be generalized across different dietary regimes, whether in a fed or fasted state. Among the variety of available possibility one approach that can be attempted in this case is the implementation of an artificial neural network (ANN, or also simply as NN), able to learn the input-output relation encoded in the set of differential equation that the numerical integration scheme is solving. While the output we are interested in is the lipid droplet final size distribution, the input can vary. The idea is to use a set of relevant parameters for the system, previously selected during the performed sensitivity analysis tests.
Using then as the loss the mean square error between the solver-predicted and the NN-predicted distribution, our NN will become a black-box method to solve the ODE system, hopefully in an efficient way.
The neural network is implemented and all the operation performed in this section and in the next 4.3.1, are performed using the `PyTorch` library [34].

This approach offers the advantage that, once the time-intensive training process is completed, the model can efficiently and quickly predict the LD output distribution for various cases and scenarios, requiring only the key parameters involved. On the other hand, one possible drawback for this approach could be definitely the lack of experimental patient data on which the network could be trained. To overcome this issue many configurations of the significant system parameters have been sampled from a normal distribution using the reference value as mean and $1/10$ of it as variance.

4.3.1 Inverse problem: from the distribution to the parameters

Gradient tracking methods in `PyTorch` provide a wide range of operations and optimizations when training neural networks. This is accomplished by setting to `requires_grad=True` the argument of PyTorch's `Tensor`, which instructs PyTorch's automatic differentiation engine, `autograd`, to track operations performed on the tensor.

An application that is quite straightforward at this point is the implementation of an approach to solve the inverse problem, aiming to recover the set of parameters responsible for the generating a particular output LD size distribution, when knowing only this output distribution.

This inverse analysis, when properly optimized, could be particularly useful in clinical settings where direct measurements are either too costly or invasive for patients. By measuring only a single cell's lipid droplet size distribution, it may be possible to infer other meaningful concentrations, since the inputs of the problem we are going to optimize have been pre-identified through sensitivity analysis as key variables.

In order to do this, the easiest way to adapt the parameters to match the target distribution y_{target} is applying a simple gradient descent (GD) approach, updating at each iteration the set of parameters $\vec{x'}$ minimizing a loss function between the network prediction and the target measured distribution, where the loss L is computed as the mean square error (MSE):

$$\vec{x'} \leftarrow \vec{x'} - \eta \cdot \nabla_{\vec{x'}} L(\vec{x'}) \tag{4.15}$$

$$L(\vec{x'}) = \frac{1}{2} ||NN(\vec{x'}) - y_{target}||^2 \tag{4.16}$$

with $NN(\cdot)$ representing the neural network prediction function for the output.

This approach encounters challenges when increasing the number of significant input parameters for the neural network, likely because a larger parameter space may lead to multiple combinations of input values producing similar output distributions. To address this issue, the proposed solution is to run the optimization problem multiple times, starting from different initial points within the parameter space, and then analyze whether the resulting solutions converge.

Another technique used to maintain positive parameter values is to work with their logarithms, and then apply the exponential function at the final stage. This approach ensures that concentrations remain non-negative, producing the desired output distribution.

5

Results

In this chapter, the results of the methods outlined in Chapter 4 are presented and thoroughly analyzed. Special attention is given to the key decisions made during the simulation process, as well as any challenges or limitations encountered at various stages of the pipeline. The discussion also highlights the implications of these results and evaluates their impact on the overall objectives of the study.

Results here analyzed are concerning initially the rat hepatocyte scenario. The study will also attempt to model human hepatocytes; the parameter adjustments required to align with the experimental findings on human cell cultures will be summarized. In this latter case, a measure that have been adopted to reduce the computational demand of the single runs, is the decrease of the number of lipid droplets size classes N_c from the initial 30 to 10.
All tests have been performed using a simulation time of 24 hours, after which all the metabolites concentrations are expected to reach a stable state.

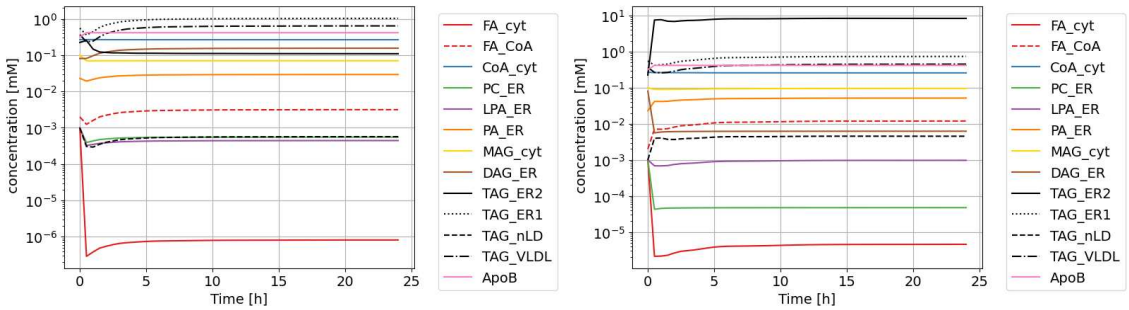
This chapter is organized as follows: Section 5.1 presents the dynamics of the simulated system, along with the droplet size distribution obtained under different input source and cell type; Section 5.2 covers the results of the Morris and Sobol' sensitivity analyses, which inform the selection of input parameters for the neural network. Finally, Section 5.3 provides an explanation of the neural network setup.

5.1 Dynamics of the system

The simulation described in 4.1 gives the evolution of the concentrations of all the considered variables. The following paragraphs of this section simulate various scenarios by altering both cell parameters (mouse or human) and the input FFA source (blood glucose or FFA challenge).

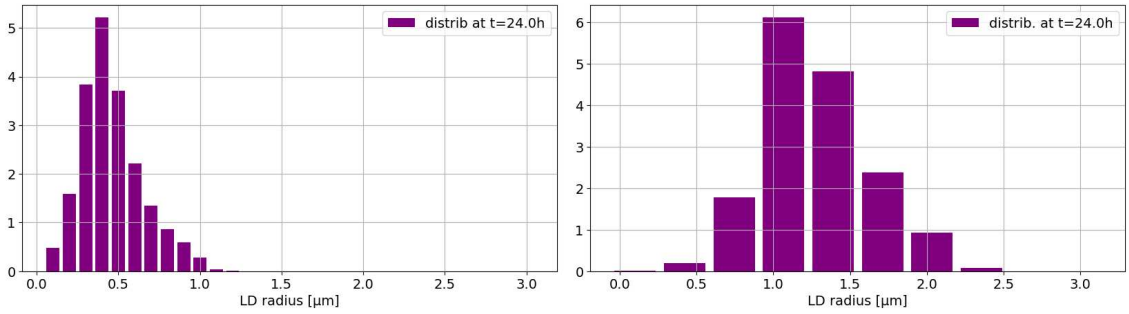
Rat and human liver cell in fed state The results presented in this section are based on simulations conducted under a fed-state scenario, with a constant blood glucose concentration of $7.637\ mM$. Although this glucose level naturally fluctuates over time, it is treated as a constant within the model for simplicity. The primary focus of the simulation is to examine the temporal evolution of metabolite concentrations, starting from an initial fed-state condition.

As expected the concentrations reach in a very good approximation the stable state in figs. 5.1a



(a) Rat hepatocyte. **(b)** Human hepatocyte.

Figure 5.1: Evolution of the metabolites residing either in cytoplasm or endoplasmatic reticulum. These results have been simulated using a blood glucose concentration equal to $7.637\ mM$, corresponding to the mice diurnal fed state. Logarithmic scales are applied to the y-axis to improve the clarity of the plots.



(a) Rat hepatocyte. **(b)** Human hepatocyte.

Figure 5.2: Resulting size distribution for the lipid droplets inside an hepatocyte after 24 h, considering as before an initial condition of diurnal fed average glucose concentration.

parameter	**rat**	**human**	correlation
CD36	1	10	positive
k_{DGAT2}	1	0.5	positive
DGAT1	1	1	negative
DGAT2	1	10	positive
FSP27	1	0.5	positive
ATGL	1	0.005	negative

Table 5.1: Recap of the parameters that are adjusted in the model when transitioning from rat hepatocytes to human hepatocytes. The value are expressed as fold changes (factors), with measurement units omitted. The "correlation" column indicates the long-term impact of each parameter change on lipid production, where a positive correlation signifies an increase in lipid production, and a negative correlation indicates a decrease.

and 5.1b. Given the different setting of constants and parameters governing the distributions of TAG stored among ER and LDs, the results here shows a significantly larger amount of TAG stored in ER$(1, 2)$ in the human hepatocyte. Consequently also the FFA level is higher, since it is produced at the end by TAG degradation and is also required in the chain cycle to start its synthesis.
The distributions in Figure 5.2 are in good approximation stable distributions, reached simulating the metabolic system for an entire day duration. This is proved also looking at the rates of the processes, which are all stable after 24 h. The comparison of the fluxes (i.e. the reaction rates) for the rat cell with the authors' results are reported in A.4, showing a generally good agreement with the paper in terms of relative error.
As can be seen in 5.2a, the distribution is shifted more to the left compared to 5.2b, indicating a higher production of smaller lipid droplets in rats. In contrast, the human distribution is broader, with larger droplets being produced. This significantly impacts the total lipid production, as the volume of the droplets, which are considered to be spherical, scales as the cube of the radius.

Human liver cell in a FFA challenge From here on, the analysis will focus exclusively on the scenario using human parameters, with the number of lipid droplet size classes reduced to 10. The fatty acid input will no longer be sourced from the blood; instead, it will replicate an *in-vitro* solution consisting of a $1 : 1$ mixture of different fatty acids at a concentration of $0.5\ mM$, resulting in a total fatty acid concentration of $1\ mM$ [1]. Based on the model equations reported in A.2, this FFA challenge is expected to cause a switch from the *de novo* synthesis of LDs to enforced lipid loading to already existing LDs, through the filling mecha-

nism, allowing to generate massive lipid quantities.

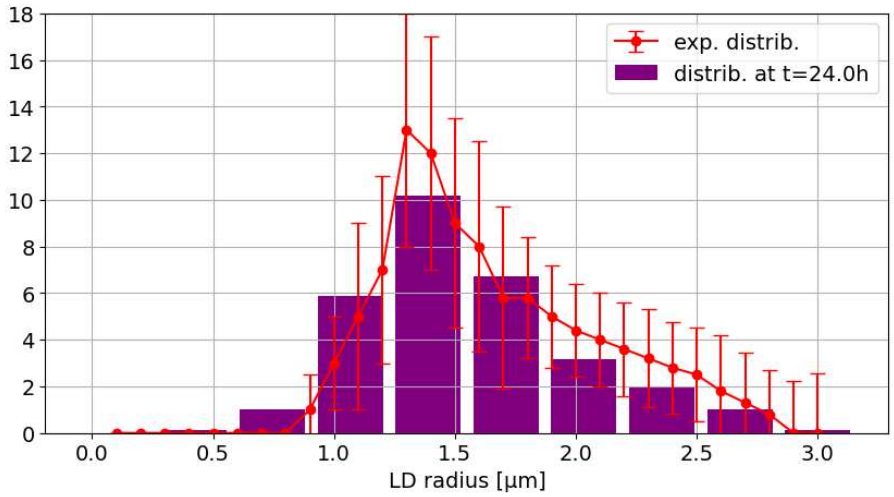


Figure 5.3: Obtained LDs distribution for the human hepatocyte in a FFA challenge after 24h (violet) and median values with errors of the 27 single cell LDs size distribution experimentally measured in patients (red), from [1].

Figure 5.3 demonstrates good compatibility between the reproduced model and the experimental results from the original study [1]. The model parameters in the human case were fine-tuned based on this distribution, so it is important that the model qualitatively aligns with the expected outcomes.

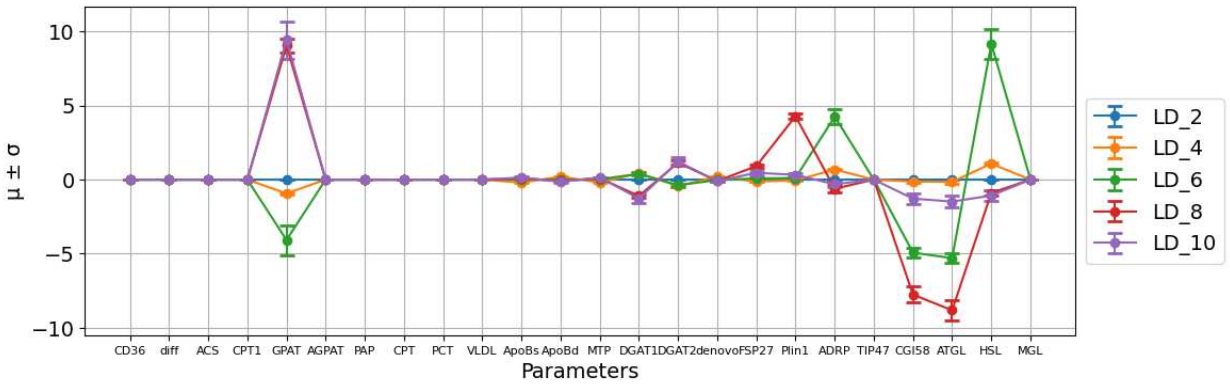
5.2 Sensitivity analysis

At this stage, as previously mentioned, it is time to conduct a sensitivity analysis to identify the most critical parameters in the model: those that have the greatest influence on the quantity of TAG produced in each class.
It is recommended to begin with a qualitative screening method to identify the most impactful parameters. Once these key parameters are identified, more precise quantitative indices, such as Sobol' indices, can be computed for further analysis.

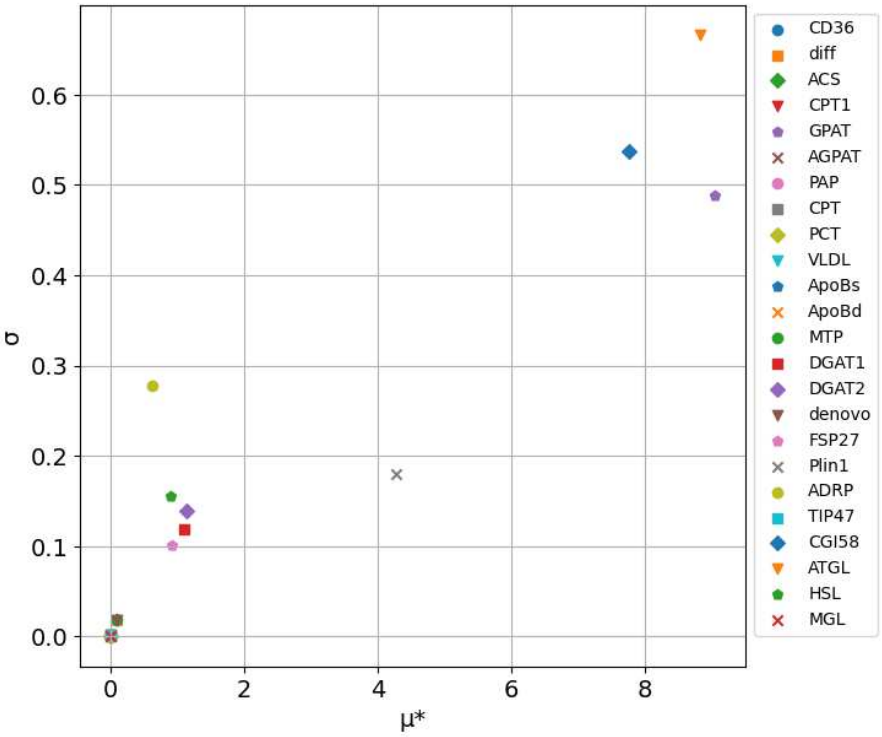
5.2.1 Morris method

The preliminary qualitative analysis have been conducted using Morris method. It focuses on the parameters of the system describing the maximal activities of the 24 fluxes of the system, telling which are the crucial reactions retaining for the major part the system behaviour and

without them this behaviour would be lost. Values are used as change factors multiplying the maximal rates V_{max} and their associated ranges, necessary for the sampling algorithm, are de-



(a) Effect on the even classes of the maximal activities fold changes. Values on the y-axis are the average elementary effects μ (4.2), while errorbars represent the standard deviation σ (4.3).



(b) Effect on $8-$th TAG class alone of the maximal activities changes. The scatterplot, σ vs. $\mu*$, allows to identify region of interest for non-influencing, linearly and non-linearly influencing parameters, as previously explained in Methods 4.2.1.

Figure 5.4: Results of the Morris OAT-type analysis performed on all the maximal activities of the fluxes, kept in the interval $[0.9, 1.1]$.

fined using fold factors of $[0.9, 1.1]$, thus replicating the original idea of a $\pm 10\%$ variation applied to the reaction rates.

Figure 5.4 confirms fold changes applied to the reaction rates that have the major influence on the TAG production, here measured for even classes in 5.4a and for only the $8-$th class in 5.4b, are the uptake rate (sum of v_{CD36} and V_{diff}) and the ones catalyzed by GPAT, DGAT1, DGAT2 and ATGL, as well as all the other rates regulating the degradation of lipids and RSPs binding/unbinding to the droplets surface.

5.2.2 Sobol method

After the crucial fluxes have been selected in 5.2.1 using Morris method, it's time to conduct a more quantitative analysis. This analysis will also include the cytosolic concentrations of all RSPs, which, while sparsely documented in the literature, are known to significantly influence the dynamics of LDs. This impact arises from binding processes on the LD surfaces, which are directly affected by these concentrations.
The sampling procedure here relies on `SALib.sample.sobol`, using Saltelli's extension of the

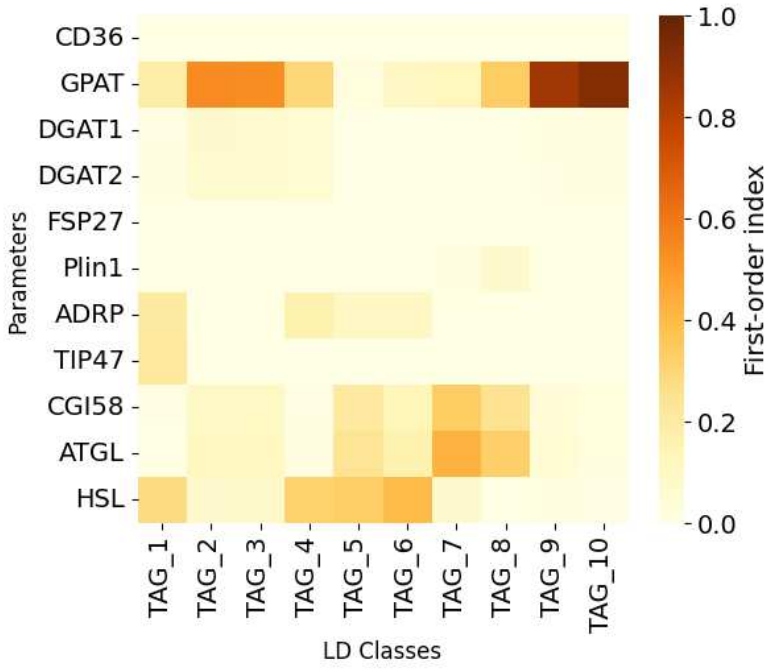


Figure 5.5: The Sobol analysis results are presented as a matrix illustrating the effects of the 11 selected parameters on the TAG content across each of the 10 analyzed classes, measured in terms of first-order index (4.2.2).

Sobol' sequence and returning a `NumPy` matrix containing the model input.

The results showed in Figure 5.5 are obtained setting a number of samples high enough to allow the Mote Carlo sampling procedure to converge, without resulting in negative first-order indices. They suggest that GPAT is a key enzyme involved in TAG production, which may seem unexpected at first but is logical. If the FA-CoA compounds are not efficiently utilized in this pathway, they may be diverted for oxidation in the mitochondria. Thus, GPAT activity can be seen as a bottleneck, limiting the progression of the reaction chain in case it is limited. The computation performed also enables the analysis of parameter correlations through second-order Sobol' indices; however, this is not the primary focus of the current work at the moment.

5.3 Learning the input-output relation with a Neural Network

When designing neural networks for tasks that are not excessively complicated, it is generally beneficial to avoid using overly complex architectures, as they can introduce unnecessary computational demands and increase the likelihood of overfitting without offering substantial improvements in performance. In this case the neural networks only aims to get to the same solution previously computed by solving the ODEs system. The architecture chosen for this quite simple task is a shallow 3-layers NN, with a 50-neurons hidden layer and having the accuracy of setting equal to 11 and 10 the number of neurons in the input and output layers respectively, equaling the number of selected input parameters and the number of real values for the LDs clsses occupation. The loss function is the mean squared error, in PyTorch called `MSELoss` criterion and using `Adam` as optimizer, setting the learning rate to 0.001.

Figure 5.6 shows results that, as expected, are very similar to those obtained by numerical integration of the ODEs. The key advantage of this approach lies in its potentially faster computational time. Although training the neural network is time-consuming, it only needs to be done once. After training, the model can quickly load the learned weights and make predictions much more efficiently.
In this relatively straightforward scenario, testing the two approaches on a single prediction, the NN-based method is slightly faster than the traditional one. This advantage might become particularly significant when scaling to multicellular problems or when iterating the solution multiple times, as will be demonstrated in the next application of the model (5.3.1).

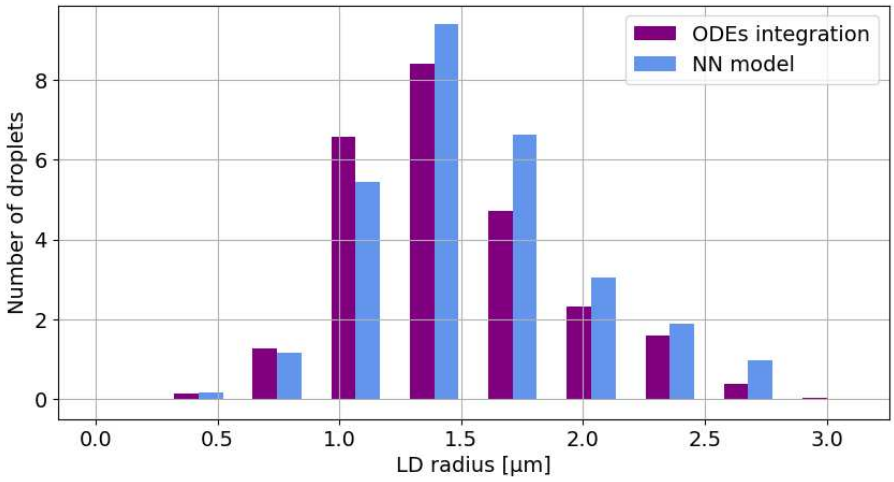


Figure 5.6: The result achieved applying the trained model to a new parameters configuration (blue) compared to the classical one by ODEs solution (violet). The synthetic dataset used for this purpose has been split into three parts: 70% for training and 15% each for validation and testing.

5.3.1 Inverse problem with GD

The goal here is to infer the parameters generating a particular LDs output distribution by adjusting the input values through gradient descent (GD) to minimize the difference between the predicted lipid droplet size distribution from the neural network and the desired LDs size distribution, which has to be the input of this kind of inverse problem.

The optimized parameters, displayed in Figure 5.7, are promising but still not fully accurate. There appears to be a systematic deviation in the parameters, particularly in the ones between PLIN1 and HSL. Further refinement could be achieved by performing multiple runs with varying initial conditions or by reducing the number of parameters.
As it stands, these results should be regarded as preliminary and not yet definitive, though this section underscores the potential of gradient descent-based inverse problem techniques to bridge the gap between observable clinical phenomena, such as LD distribution, and underlying biological mechanisms, while also highlighting the need for further validation with actual patient datasets.

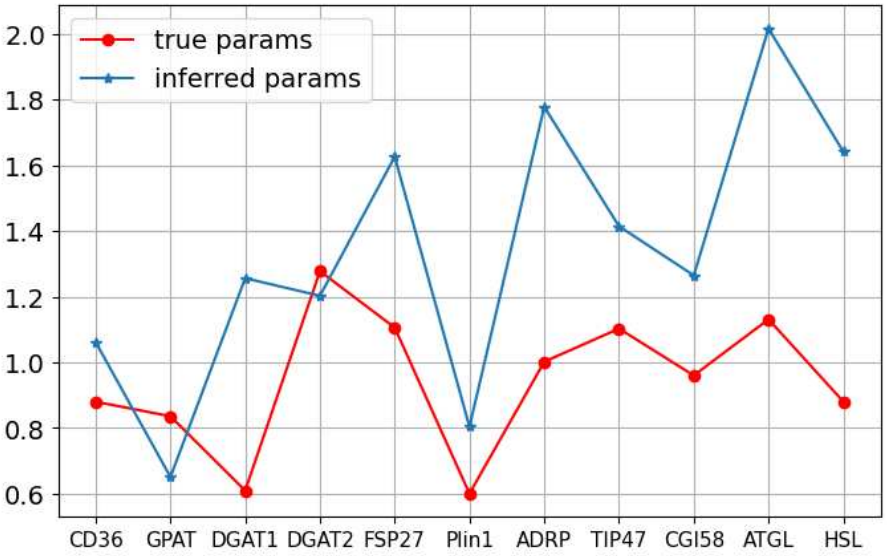


Figure 5.7: The original parameters (red) and the inferred ones (blue), obtained minimizing the difference between the NN-predicted and the target LDs distribution.

6 Conclusion

In this work the goal was to implement a liver cell digital twin for the lipid metabolism, comprising a simulation for the lipid droplets genesis, degradation and all the other intermediate steps, building the model on the mathematical framework introduced in [1].

Even though some difficulties have been encountered during the implementation of the model, the results are in good approximation satisfactory when compared to the ones of the reference paper. The issues encountered are of different kinds and can be summed up as:

- notation issues, mathematically the notation used in the reference paper's equations is often inaccurate or not precise, requiring a careful analysis of the measures and typos in order to recover the original model with an accuracy as high as possible;
- the equation 3.3 have been introduced to model the content of neutral lipids in each LD size class. Moreover, it was challenging for an external reader to implement the shift between neighbouring classes (A.3) that was mentioned in the original paper [1];
- the mathematical model is very sensitive to small variation not only of the metabolites initial concentrations, but also Michaelis constants K_m, maximal enzymes activities V_{max} and other parameters. The values have been selected with caution, as small variations in parameters for this kind of stiff problems can significantly alter the dynamics of the system;
- starting from scratch with the entire code is a very time demanding process that involves numerous trial-and-error steps before achieving a smoothly running model simulation.

There is certainly room for improvement in various aspects of the implementation and the overall pipeline.

Based on the original mechanistic model for the hepatocyte metabolic network, the model simulates a good approximation of the stable lipid droplet distribution, providing an estimation of the variability in the abundance of central enzymes and RSPs accounting for the observed inter-cellular heterogeneity of LD size distributions.

For a more accurate simulation of the disease over time, the cytosolic concentration of regulatory surface proteins must be properly modeled to reflect dynamic changes rather than being treated as fixed constants. This is particularly important for capturing the evolution of key parameters that significantly influence the model. Currently, the model has only been tested and applied to relatively short time periods, and its effectiveness beyond these duration remains unknown.

The contributions of this work are significant, especially in terms of potential clinical applications. The use of neural networks offers a novel approach for identifying key parameters with precision, while maintaining accuracy. This is particularly relevant given the challenges associated with measuring various concentrations, which are often either difficult or prohibitively expensive to obtain. Moreover, the inverse problem framework allows for the identification of parameter ranges that can lead to steatosis, offering valuable insights for understanding and potentially mitigating lipid-related disorders.
However, a potential drawback of this method is that since it is trained on synthetic data, it may lose its generalization capability when applied to real patient data, especially given that such data is unfortunately not currently available.

In conclusion, potential future developments could build upon this work in several ways, starting from the extension of the model to a multicellular scenario as an hepatic lobule, having approximately 10^8 cells/g, to see how the actual implementation scales for problems on a much larger scenario. Additionally, there is potential for extending this approach to other tissues, broadening its applicability and enhancing our understanding of lipid dynamics in various physiological contexts.

A

Appendix

This Appendix A is used to provide the mathematical details omitted in the main text for the sake of simplicity, ensuring clarity while preventing the inclusion of complex formulas in the narrative.

A.1 Hormones and metabolites plasma profiles

The glucose level in the blood is the proxy used for the feeding state of the patient/mice, high glucose level means we are measuring short after a meal and this cause an increase in the TAG synthesis in hepatocytes.

A.1.1 Insulin and glucagon

Concentration of the insulin and glucagon hormones is controlled by the glucose concentration of the blood. While insuline enhances TAG synthesis and storage in hepatocytes, here directly influencing the $ApoB$ degradation rate (A.67), glucagon counteracts the effects of insulin and leads to the breakdown of TAGs. It promotes lipolysis by activating hormone-sensitive lipase (HSL), which breaks down stored TAGs into free fatty acids and glycerol.

To compute their level, the following glucose hormone transfer function (GHT) is used.

$$\text{Ins} = 2 \cdot \left(1.55\text{nM} \cdot \frac{(\text{Glc}_{\text{ex}})^{5.7}}{(\text{Glc}_{\text{ex}})^{5.7} + (7.7\text{mM})^{5.7}} \right) \tag{A.1}$$

$$\text{Glucagon} = 2 \cdot \left(0.253\text{nM} \cdot \left(1 - \frac{(\text{Glc}_{\text{ex}})^{5.65}}{(\text{Glc}_{\text{ex}})^{5.65} + (4.7\text{mM})^{5.65}} \right) + 0.02\text{nM} \right) \tag{A.2}$$

The concentration of these two hormones determine the phosphorylation state γ of the interconvertible enzymes.

$$\gamma = \frac{1}{2} \cdot \left(1 - \frac{\text{Ins}^{1.75}}{\text{Ins}^{1.75} + (0.70\text{nM})^{1.75}} + \frac{\text{Glucagon}^{3.6}}{\text{Glucagon}^{3.6} + (0.08\text{nM})^{3.6}} \right) \tag{A.3}$$

A.1.2 Total and free fatty acids

The relationship between plasma total and free fatty acids (TFA and FFA) and plasma glucose (Glc_{ex}) levels can be described using the glucose free fatty acid transfer (GFT) functions. These functions were developed to capture the dependence of plasma FFA concentrations on plasma glucose levels, using experimental data from several studies to fit Hill-type transfer functions.

$$\text{TFA}_{\text{ex}} = 1.2\text{mM} - 1.1\text{mM} \frac{\text{Glc}_{\text{ex}}^4}{\text{Glc}_{\text{ex}} + (6.5\text{mM})^4} \tag{A.4}$$

Plasma fatty acids are primarily bound to plasma albumin, with only free fatty acids available for liver uptake. To estimate the free fatty acid (FFA) concentration, we use a fourth-order polynomial equation to calculate FFAs from the total fatty acids (TFA) in the blood plasma (ex).

$$\frac{\text{FFA}_{\text{ex}}}{[\text{nM}]} = 27.86 \left(\frac{\text{TFA}_{\text{ex}}}{[\text{mM}]} \right)^4 - 18.29 \left(\frac{\text{TFA}_{\text{ex}}}{[\text{mM}]} \right)^3 + 30.88 \left(\frac{\text{TFA}_{\text{ex}}}{[\text{mM}]} \right)^2 + 17.83 \left(\frac{\text{TFA}_{\text{ex}}}{[\text{mM}]} \right) \tag{A.5}$$

A.2 Reaction equations

List of the rate equations and kinetic parameters from the model [1]. The square parenthesis is the notation commonly used to indicate the metabolite concentration, while **bold** style is used to indicate vectors, namely array of values containing a number of elements equal to the chosen number of LD classes. The brackets with no number inside, $\langle - \rangle$, refer to processes that are not graphically depicted in Fig.3.1. Letters are used, e.g. $6a, 6b, 6c$ where multiple processes are involved in the main reaction.

A.2.1 Metabolic module

$\langle 1 \rangle$ FA uptake by a combination of diffusion and an active transport process with CD36.

$$[FA]_{ex} \xrightarrow{diff,CD36} [FA]_{cyt} \tag{A.6}$$

$$v_{diff} = V_{max}^{diff}([FA]_{ex} - [FA]_{cyt}) \tag{A.7}$$

$$v_{CD36} = V_{max}^{CD36} \frac{[FA]_{ex} - [FA]_{cyt}}{1 + \frac{[FA]_{ex}}{K_m^{FA_{ex}}} + \frac{[FA]_{cyt}}{K_m^{FA_{cyt}}}} \tag{A.8}$$

$$K_m^{FA_{ex}} = 0.000083\, mM,\ K_m^{FA_{cyt}} = 0.004\, mM \tag{A.9}$$

$$V_{max}^{diff} = 1.2442 \cdot 10^5\, h^{-1},\ V_{max}^{CD36} = 1.14 \cdot 10^1\, h^{-1} \tag{A.10}$$

$\langle 2 \rangle$ Activation of FFA by ACS.

$$[FA]_{cyt} + [CoA]_{cyt} + [ATP] \xrightarrow{ACS} [FA\text{-}CoA]_{cyt} + [ADP] \tag{A.11}$$

$$v_{ACS} = V_{max}^{ACS} \left(\frac{[FA]_{cyt}}{[FA]_{cyt} + K_m^{FA_{cyt}}} \right) \left(\frac{[ATP]}{[ATP] + K_m^{ATP}} \right) \left(\frac{[CoA]_{cyt}}{[CoA]_{cyt} + K_m^{CoA_{cyt}}} \right) \tag{A.12}$$

$$K_m^{FA_{cyt}} = 0.05\, mM,\ K_m^{ATP} = 0.65\, mM,\ K_m^{CoA_{cyt}} = 0.0064\, mM \tag{A.13}$$

$$V_{max}^{ACS} = 3.1104 \cdot 10^6\, mM/h \tag{A.14}$$

⟨3⟩ Mitochondrial β-oxidation by CPT1.

$$[FA\text{-}CoA]_{cyt} + [Car]_{cyt} \xrightarrow{CPT1} [FA\text{-}Car]_{cyt} + [CoA]_{cyt} \tag{A.15}$$

$$v_{CPT1} = V_{max}^{CPT1} \left(\frac{[FA\text{-}CoA]_{cyt}^{2.47}}{[FA\text{-}CoA]_{cyt}^{2.47} + (K_m^{FA\text{-}CoA})^{2.47}} \right) \left(\frac{[Car]_{cyt}}{[Car]_{cyt} + K_m^{Car}} \right) \tag{A.16}$$

$$K_m^{FA\text{-}CoA} = K_0^{FA\text{-}CoA} \left(1 + \frac{[malcoa]_{cyt}}{K_i^{malcoa}} \right) \tag{A.17}$$

$$K_0^{FA_{cyt}} = 0.03\, mM,\ K_m^{Car} = 0.0032\, mM,\ K_i^{malcoa} = 0.0025\, mM \tag{A.18}$$

$$V_{max}^{CPT1} = 1.8000 \cdot 10^3\, mM/h \tag{A.19}$$

⟨4a⟩ Esterification of G3P by GPAT.

$$[G3P]_{cyt} + [FA\text{-}CoA]_{cyt} \xrightarrow{GPAT} [LPA]_{er} + [CoA]_{cyt} \tag{A.20}$$

$$v_{GPAT} = V_{max}^{GPAT} \left(\frac{[G3P]_{cyt}}{[G3P] + K_m^{G3P}} \right) \left(\frac{[FA\text{-}CoA]_{cyt}}{[FA\text{-}CoA]_{cyt} + K_m^{FA\text{-}CoA}} \right) \tag{A.21}$$

$$K_m^{G3P} = 0.2\, mM,\ K_m^{FA\text{-}CoA} = 0.09\, mM \tag{A.22}$$

$$V_{max}^{GPAT} = 5.4000 \cdot 10^2\, mM/h \tag{A.23}$$

⟨4b⟩ Esterification of LPA with a long-chain acyl-CoA by AGPAT.

$$[LPA]_{er} + [FA\text{-}CoA]_{cyt} \xrightarrow{AGPAT} [PA]_{er} + [CoA]_{cyt} \tag{A.24}$$

$$v_{AGPAT} = V_{max}^{AGPAT} \left(\frac{[LPA]_{er}}{[LPA]_{er} + K_m^{LPA}} \right) \left(\frac{[FA\text{-}CoA]_{cyt}}{[FA\text{-}CoA]_{cyt} + K_m^{FA\text{-}CoA}} \right) \tag{A.25}$$

$$K_m^{LPA} = 0.0065\, mM,\ K_m^{FA\text{-}CoA} = 0.004\, mM \tag{A.26}$$

$$V_{max}^{AGPAT} = 5.4000 \cdot 10^2\, mM/h \tag{A.27}$$

⟨4c⟩ Hydrolysis of PA by PAP.

$$[PA]_{er} \xrightarrow{PAP} [DAG]_{er} + [Pi]_{cyt} \tag{A.28}$$

$$v_{PAP} = V_{max}^{PAP} \left(\frac{[PA]_{er}^{2.2}}{[PA]_{er}^{2.2} + (K_m^{PA})^{2.2}} \right) \tag{A.29}$$

$$K_m^{PA} = 0.35\, mM \tag{A.30}$$

$$V_{max}^{PAP} = 3.6000 \cdot 10^3\, mM/h \tag{A.31}$$

⟨5⟩ PC synthesis by CPT.

$$[DAG]_{er} \xrightarrow{CPT} [PC]_{er} \tag{A.32}$$

$$v_{CPT} = V_{max}^{CPT} \left(\frac{[DAG]_{er}}{[DAG]_{er} + K_m^{DAG_{er}}} \right) \tag{A.33}$$

$$K_m^{DAG} = 0.125\, mM \tag{A.34}$$

$$V_{max}^{CPT} = 3.6\, mM/h \tag{A.35}$$

⟨5*b*⟩ PC export.

$$[PC]_{er} \xrightarrow{PCTP} [PC]_{ex} \tag{A.36}$$

$$v_{PCTP} = V_{max}^{PCTP} [PC]_{er} \tag{A.37}$$

$$V_{max}^{PCTP} = 3.6000 \cdot 10^3\, h^{-1} \tag{A.38}$$

⟨6⟩ Synthesis of TAG by DGAT1 and DGAT2.

⟨6*a*⟩ TAG synthesis by DGAT1.
DGAT1 resides in the ER membrane and synthesizes $TAG_{er(1)}$ used for VLDL synthesis and nLD formation.

$$[DAG]_{er} + [FA\text{-}CoA]_{cyt} \xrightarrow{DGAT1} [TAG]_{er(1)} + [CoA]_{cyt} \tag{A.39}$$

$$v_{TAG_{er(1)}}^{DGAT1} = V_{max}^{DGAT1} \left(\frac{[FA\text{-}CoA]_{cyt}}{[FA\text{-}CoA]_{cyt} + K_m^{FA\text{-}CoA}} \right) \left(\frac{[DAG]_{er}}{[DAG]_{er} + K_m^{DAG_{er}}} \right) \tag{A.40}$$

$$K_m^{FA\text{-}CoA} = 0.1\, mM,\ K_m^{DAG} = 0.03\, mM \tag{A.41}$$

$$V_{max}^{DGAT1} = 2.8800 \cdot 10^2\, mM/h \tag{A.42}$$

⟨6*b*⟩ TAG synthesis by DGAT2.
DGAT2 resides either in the ER membrane and synthesizes $[TAG]_{er(2)}$ used for nascent LD formation or distributes to LD surface synthesizing $[\boldsymbol{TAG}]_{\boldsymbol{LD}}$ directly filling LDs ⟨10⟩. The distribution of DGAT2 between the er and the different LD classes, LD_n, depends on the

number of LDs in each class, $\boldsymbol{num_{LD}}$, and the radius of the LDs, $\boldsymbol{r_{LD}}$.

$$[DAG]_{er} + [FA\text{-}CoA]_{cyt} \xrightarrow{DGAT2} [TAG]_{er(2)} + [CoA]_{cyt} \tag{A.43}$$

$$K_{er}^{DGAT2} = \frac{10}{\sum \boldsymbol{num_{LD}}} \tag{A.44}$$

$$\boldsymbol{K_{LD}^{DGAT2}} = \boldsymbol{num_{LD}} \cdot \left(1 - \frac{\boldsymbol{r_{LD}}^3}{\boldsymbol{r_{LD}}^3 + k_{r_{LD}}^3}\right) \tag{A.45}$$

$$V_{max_{er}}^{DGAT2} = V_{max}^{DGAT2} \frac{K_{er}^{DGAT2}}{K_{er}^{DGAT2} + \sum \boldsymbol{K_{LD}^{DGAT2}}} \tag{A.46}$$

$$v_{TAG_{er(2)}}^{DGAT2} = V_{max_{er}}^{DGAT2} \left(\frac{[FA\text{-}CoA]_{cyt}}{[FA\text{-}CoA]_{cyt} + K_m^{FA\text{-}CoA}}\right) \left(\frac{[DAG]_{er}}{[DAG]_{er} + K_m^{DAG_{er}}}\right) \tag{A.47}$$

$$k_{r_{LD}} = 0.15\,\mu m,\ K_m^{FA\text{-}CoA} = 0.1\,mM,\ K_m^{DAG} = 0.03\,mM \tag{A.48}$$

$$V_{max}^{DGAT2} = 3.0240 \cdot 10^2\,mM/h \tag{A.49}$$

A.2.2 Lipid module

⟨7⟩ *De novo* synthesis of nLD.
The produced nascent lipid droplets belong to the first LD class, both for the ones synthesized from $TAG_{er(1)}$ and $TAG_{er(2)}$, and they are not concerned by protein coating to their surface.

$$[TAG]_{er(1,2)} \xrightarrow{de\ novo} [TAG]_{nLD} \tag{A.50}$$

$$v_{LD(1,2)}^{de\ novo} = V_{max_{LD}}^{de\ novo} \cdot [TAG]_{er(1,2)} \tag{A.51}$$

$$V_{max_{LD}}^{de\ novo} = 3.6\,h^{-1} \tag{A.52}$$

⟨−⟩ nLD coating with TIP47. Coating with TIP47 convert the nascent lipid droplets to actual lipid droplets, detaching from the ER.

$$[TAG]_{nLD} + [TIP47]_{cyt} \xrightarrow{TIP47} \delta_{n,1}[\boldsymbol{TAG}]_{\boldsymbol{LD}} \tag{A.53}$$

$$v_{nLD}^{TIP47} = k_{nLD}^{TIP47} \cdot [TAG]_{nLD} \cdot [TIP47]_{cyt} \tag{A.54}$$

$$k_{nLD}^{TIP47} = 3.6 \cdot 10^3\,h^{-1} \tag{A.55}$$

$\langle - \rangle$ nLD degradation by ATGL.

$$[TAG]_{nLD} + [ATGL]_{cyt} \xrightarrow{ATGL} \tag{A.56}$$

$$v_{nLD}^{ATGL} = k_{nLD}^{ATGL} \cdot [TAG]_{nLD} \cdot [ATGL]_{cyt} \tag{A.57}$$

$$k_{nLD}^{ATGL} = 7.92\, h^{-1} \tag{A.58}$$

$\langle 8 \rangle$ TAG filling of LDs by DGAT2.

The distribution of DGAT2 between ER and LDs have been already defined in A.44 and A.45.

$$[DAG]_{er} + [FA\text{-}CoA]_{cyt} \xrightarrow{DGAT2} [\boldsymbol{TAG}]_{\boldsymbol{LD}} + [CoA]_{cyt} \tag{A.59}$$

$$\boldsymbol{V}_{\boldsymbol{max_{LD}}}^{\boldsymbol{DGAT2}} = V_{max}^{DGAT2} \frac{\boldsymbol{K}_{\boldsymbol{LD}}^{\boldsymbol{DGAT2}}}{K_{er}^{DGAT2} + \sum \boldsymbol{K}_{\boldsymbol{LD}}^{\boldsymbol{DGAT2}}} \tag{A.60}$$

$$\boldsymbol{v}_{\boldsymbol{TAG_{LD}}}^{\boldsymbol{DGAT2}} = \boldsymbol{V}_{\boldsymbol{max_{LD}}}^{\boldsymbol{DGAT2}} \left(\frac{[FA\text{-}CoA]_{cyt}}{[FA\text{-}CoA]_{cyt} + K_m^{FA\text{-}CoA}} \right) \left(\frac{[DAG]_{er}}{[DAG]_{er} + K_m^{DAG_{er}}} \right) \tag{A.61}$$

The filling of LDs of class n diminishes the cellular TAG content $[TAG]_{LD_n}$ in the respective class and increases the content $[TAG]_{LD_{n+1}}$ in class $n+1$ by shifting TAG from the smaller to the bigger droplet. The shifting rate depends on the relative volume ΔVol_{LD_n} between classes n and $n-1$.

$$\Delta Vol_{LD_n} = \frac{Vol_{LD_n}}{Vol_{LD_{n-1}}} - 1 \tag{A.62}$$

$$v_{shift_{LD_n}}^{DGAT2} = v_{TAG_{LD_{n-1}}}^{DGAT2} \cdot \Delta Vol_{LD_n} \tag{A.63}$$

The change of $[TAG]_{LD_n}$ is the sum of filling this class accompanied by the shifting of $[TAG]_{LD_{n-1}}$ into this class, and shifting from this class in the next bigger class

$$v_{TAG_{LD_n}}^{fill_{DGAT2}} = v_{TAG_{LD_{n-1}}}^{DGAT2} + v_{shift_{LD_n}}^{DGAT2} - v_{shift_{LD_{n+1}}}^{DGAT2} \tag{A.64}$$

$\langle 9a \rangle$ Synthesis and degradation of ApoB.

The degradation rate of ApoB, v_{ApoB}^{deg}, is the sum of a basal rate k_1 and a insulin dependent rate

k_2, while the synthesis rate, v^{syn}_{ApoB}, is considered to be a constant.

$$\xrightarrow{v^{syn}_{ApoB}} [ApoB]_{er} \xrightarrow{v^{deg}_{ApoB}} \tag{A.65}$$

$$v^{syn}_{ApoB} = \text{const.} = 3.6 \cdot 10^3\, mM/h \tag{A.66}$$

$$v^{deg}_{ApoB} = V^{ApoB_{deg}}_{max} \cdot [ApoB]_{er} \cdot \left(k_1 + k_2 \frac{[Ins]_{ex}}{[Ins]_{ex} + K^{Ins_{ex}}_m} \right) \tag{A.67}$$

$$k_1 = 0.3,\ k_2 = 0.7,\ K^{Ins_{ex}}_m = 10^3\, pM \tag{A.68}$$

$$V^{ApoB_{deg}}_{max} = 2.8804 \cdot 10^4\, h^{-1} \tag{A.69}$$

⟨9*b*⟩ Synthesis and secretion of VLDL.

$$[ApoB]_{er} + [TAG]_{er(1)} \xrightarrow{MTP} [TAG\text{-}VLDL]_{er} \xrightarrow{VLDL_{ex}} [VLDL]_{ex} \tag{A.70}$$

$$v_{MTP} = V^{MTP}_{max} \cdot [TAG]_{er(1)} \cdot [ApoB]_{er} \tag{A.71}$$

$$v_{VLDL_{ex}} = V^{VLDL_{ex}}_{max} \left(\frac{[TAG\text{-}VLDL]_{er}}{[TAG\text{-}VLDL]_{er} + K^{[TAG\text{-}VLDL]_{er}}_m} \right) \tag{A.72}$$

$$V^{MTP}_{max} = 3.440\, mM^{-1}h^{-1},\ V^{VLDL_{ex}}_{max} = 5.4000 \cdot 10^1\, mM/h \tag{A.73}$$

⟨10⟩ LD fusion regulated by FSP27.

The fusion rate fr_{ij} between two droplets belonging to size classes i and j depends on their radii r_i and r_j and their droplet number num_{LD_i} and num_{LD_j}, as well as on the availability of the regulating protein FSP27 in the cytoplasm.

$$fr_{ij} = V^{fus}_{max} \cdot [FSP27]_{cyt} \cdot num_{LD_i} \cdot num_{LD_j} \cdot \left| \frac{1}{r_{LD_j}} - \frac{1}{r_{LD_i}} \right| \tag{A.74}$$

$$V^{fus}_{max} = 1.8000 \cdot 10^{-2}\, mM\mu m^{-2}h^{-1} \tag{A.75}$$

Considering the droplets indices i and j with $i > j$, the model is implemented in such a way that $[TAG]_{LD_i}$ and $[TAG]_{LD_j}$ decrease and $[TAG]_{LD_{j+1}}$ increases, simulating the absorption of the droplet i by j and the filling of $[TAG]_{LD_j}$ with $[TAG]_{LD_i}$. Analogously to the LDs filling process, there is a shift of TAG from $[TAG]_{LD_j}$ to $[TAG]_{LD_{j+1}}$, with a rate proportional to the relative volume ΔVol_{LD_j}.

Droplets of the maximum size (radius equal to $3\mu m$) cannot fuse.

$$v^{shrink}_{TAG_{LD_j}} = \sum_{i=j}^{N_c-1} fr_{ij} \cdot Vol_{LD_j} \tag{A.76}$$

$$v^{shrink}_{shift_{LD_j}} = v^{shrink}_{TAG_{LD_j}} \cdot \Delta Vol_{LD_j} \tag{A.77}$$

$$v^{growth}_{TAG_{LD_j}} = \sum_{i=i}^{j-1} fr_{ij} \cdot Vol_{LD_i} \tag{A.78}$$

$$v^{growth}_{shift_{LD_j}} = v^{growth}_{TAG_{LD_j}} \cdot \Delta Vol_{LD_j} \tag{A.79}$$

$$v^{fus}_{TAG_{LD_j}} = v^{growth}_{TAG_{LD_j}} - v^{shrink}_{TAG_{LD_j}} + v^{growth}_{shift_{LD_j}} - v^{shrink}_{shift_{LD_{j+1}}} \tag{A.80}$$

$\langle - \rangle$ Reversible binding of RSPs to LD surface.

Binding rate depends both on a function of the radius of the LD, $\boldsymbol{k_r^X}$, and the free fraction of its surface, $\boldsymbol{f_{LD}^{free}}$. The vector $\boldsymbol{X_{LD}}$ describes the fraction of the LD surface occupied by the RSPs X, where $X = PLIN1, ADRP, TIP47, CGI58, ATGL$.

$$\xrightarrow{v^X_{on}} \boldsymbol{X_{LD}} \xrightarrow{v^X_{off}} \tag{A.81}$$

$$\boldsymbol{v^{PLIN1}_{on}} = k^{PLIN1}_{on} \cdot \boldsymbol{k_r^{PLIN1}} \cdot [PLIN1]_{cyt} \cdot \boldsymbol{f_{LD}^{free}} \tag{A.82}$$

$$\boldsymbol{v^{PLIN1}_{off}} = k^{PLIN1}_{off} \cdot (1 - \boldsymbol{k_r^{PLIN1}}) \cdot \boldsymbol{PLIN1_{LD}} \tag{A.83}$$

$$\text{where } \boldsymbol{k_r^{PLIN1}} = \frac{\boldsymbol{r_{LD}^{15}}}{\boldsymbol{r_{LD}^{15}} + k^{15}_{r_{LD}}} \text{ and } k_{r_{LD}} = 2\mu m \tag{A.84}$$

$$\boldsymbol{v^{ADRP}_{on}} = k^{ADRP}_{on} \cdot (1 - \boldsymbol{k_r^{ADRP}}) \cdot [ADRP]_{cyt} \cdot \boldsymbol{f_{LD}^{free}} \tag{A.85}$$

$$\boldsymbol{v^{ADRP}_{off}} = k^{ADRP}_{off} \boldsymbol{ADRP_{LD}} \tag{A.86}$$

$$\text{where } \boldsymbol{k_r^{ADRP}} = \frac{\boldsymbol{r_{LD}^{12}}}{\boldsymbol{r_{LD}^{12}} + k^{12}_{r_{LD}}} \text{ and } k_{r_{LD}} = 2\mu m \tag{A.87}$$

$$\boldsymbol{v^{TIP47}_{on}} = k^{TIP47}_{on} \cdot \boldsymbol{k_r^{TIP47}} \cdot [TIP47]_{cyt} \cdot \boldsymbol{f_{LD}^{free}} \tag{A.88}$$

$$\boldsymbol{v^{TIP47}_{off}} = k^{TIP47}_{off} \cdot (1 - \boldsymbol{k_r^{TIP47}}) \cdot \boldsymbol{TIP47_{LD}} \tag{A.89}$$

$$\text{where } \boldsymbol{k_r^{TIP47}} = 1 - \frac{\boldsymbol{r_{LD}^{10}}}{\boldsymbol{r_{LD}^{10}} + k^{10}_{r_{LD}}} \text{ and } k_{r_{LD}} = 0.5\mu m \tag{A.90}$$

$$k^{PLIN1}_{on} = 1.8000 \cdot 10^4\, h^{-1},\ k^{PLIN1}_{off} = 3.6000 \cdot 10^3\, h^{-1}, \tag{A.91}$$

$$k^{ADRP}_{on} = 1.2600 \cdot 10^5\, h^{-1},\ k^{ADRP}_{off} = 3.6000 \cdot 10^3\, h^{-1}, \tag{A.92}$$

$$k^{TIP47}_{on} = k^{TIP47}_{off} = 3.6000 \cdot 10^3\, h^{-1} \tag{A.93}$$

The binding of $CGI58$ to the LD is mediated by $PLIN1$ and $ADRP$, and its phosphorylation in response to hormonal stimulation leads to binding and therefore activation of $ATGL$.

$$\boldsymbol{v_{on}^{CGI58}} = k_{on}^{CGI58} \cdot [CGI58]_{cyt} \cdot (k_1 \cdot \boldsymbol{PLIN1_{LD}} + k_2 \cdot \boldsymbol{ADRP_{LD}}) \tag{A.94}$$

$$\boldsymbol{v_{off}^{CGI58}} = k_{off}^{CGI58} \cdot \boldsymbol{CGI58_{LD}} \tag{A.95}$$

$$\text{where } k_1 = 0.1 \text{ and } k_2 = 0.2 \tag{A.96}$$

$$\boldsymbol{v_{on}^{ATGL}} = k_{on}^{ATGL} \cdot \boldsymbol{num_{LD}} \cdot [ATGL]_{cyt} \cdot \boldsymbol{f_{LD}^{free}} \tag{A.97}$$

$$\boldsymbol{v_{off}^{ATGL}} = k_{off}^{ATGL} \cdot \boldsymbol{ATGL_{LD}} \tag{A.98}$$

$$\boldsymbol{v_{ATGL_{LD}}^{CGI58_{on}}} = k_{ATGL_{LD}}^{CGI58_{on}} \cdot \boldsymbol{num_{LD}} \cdot \boldsymbol{Sur_{LD}} \cdot (\boldsymbol{CGI58_{LD}} \cdot (1-\gamma)) \cdot \boldsymbol{ATGL_{LD}} \tag{A.99}$$

$$\boldsymbol{v_{ATGL_{LD}}^{CGI58_{off}}} = k_{ATGL_{LD}}^{CGI58_{off}} \cdot \boldsymbol{ATGL\text{-}CGI58_{LD}} \tag{A.100}$$

$$k_{on}^{CGI58} = 3.6000 \cdot 10^3\, h^{-1},\ k_{off}^{CGI58} = 3.6000 \cdot 10^4\, h^{-1}, \tag{A.101}$$

$$k_{on}^{ATGL} = 3.6\, \mu m^{-2} h^{-1},\ k_{off}^{ATGL} = 3.6000 \cdot 10^1\, h^{-1}, \tag{A.102}$$

$$k_{ATGL_{LD}}^{CGI58_{on}} = 3.6\, \mu m^{-2} h^{-1},\ k_{ATGL_{LD}}^{CGI58_{off}} = 33.6000 \cdot 10^1\, h^{-1} \tag{A.103}$$

⟨11a⟩ LD's TAG degradation by ATGL.

$$[\boldsymbol{TAG}]_{\boldsymbol{LD}} \xrightarrow{ATGL} [\boldsymbol{DAG}]_{\boldsymbol{LD}} + [FA]_{cyt} \tag{A.104}$$

$$\boldsymbol{v_{ATGL}} = V_{max}^{ATGL} \cdot \boldsymbol{num_{LD}} \cdot (\boldsymbol{ATGL_{LD}} + 20 \cdot [\boldsymbol{ATGL\text{-}CGI58}]_{\boldsymbol{LD}}) \cdot [\boldsymbol{TAG}]_{\boldsymbol{LD}} \tag{A.105}$$

$$V_{max}^{ATGL} = 7.92\, h^{-1} \tag{A.106}$$

⟨11b⟩ LD's TAG and DAG degradation by HSL.

HSL can hydrolyze TAG, although its main activity is DAG hydrolyzation. HSL is activated

by phosphorylation of $\boldsymbol{PLIN1_{LD}} + \boldsymbol{ADRP_{LD}}$.

$$[\boldsymbol{TAG}]_{\boldsymbol{LD}} \xrightarrow{HSL} [\boldsymbol{DAG}]_{\boldsymbol{LD}} + [FA]_{cyt} \tag{A.107}$$

$$\boldsymbol{v}_{\boldsymbol{HSL}}^{\boldsymbol{TAG}} = V_{max}^{HSL_{TAG}} \cdot (1-\gamma) \cdot \boldsymbol{num_{LD}} \cdot \boldsymbol{Sur_{LD}} \cdot \tag{A.108}$$

$$\cdot (\boldsymbol{f}_{\boldsymbol{LD}}^{\boldsymbol{free}} + \boldsymbol{PLIN1_{LD}} + \boldsymbol{ADRP_{LD}}) \cdot \frac{[\boldsymbol{TAG}]_{\boldsymbol{LD}}}{[\boldsymbol{MAG}]_{\boldsymbol{LD}} + [\boldsymbol{DAG}]_{\boldsymbol{LD}} + [\boldsymbol{TAG}]_{\boldsymbol{LD}}} \tag{A.109}$$

$$[\boldsymbol{DAG}]_{\boldsymbol{LD}} \xrightarrow{HSL} [\boldsymbol{MAG}]_{\boldsymbol{LD}} + [FA]_{cyt} \tag{A.110}$$

$$\boldsymbol{v}_{\boldsymbol{HSL}}^{\boldsymbol{DAG}} = V_{max}^{HSL_{DAG}} \cdot (1-\gamma) \cdot \boldsymbol{num_{LD}} \cdot \boldsymbol{Sur_{LD}} \cdot \tag{A.111}$$

$$\cdot (\boldsymbol{f}_{\boldsymbol{LD}}^{\boldsymbol{free}} + \boldsymbol{PLIN1_{LD}} + \boldsymbol{ADRP_{LD}}) \cdot \frac{[\boldsymbol{DAG}]_{\boldsymbol{LD}}}{[\boldsymbol{MAG}]_{\boldsymbol{LD}} + [\boldsymbol{DAG}]_{\boldsymbol{LD}} + [\boldsymbol{TAG}]_{\boldsymbol{LD}}} \tag{A.112}$$

$$V_{max}^{HSL_{TAG}} = 3.1700 \cdot 10^{-3}\, mM\mu m^{-2}h^{-1},\ V_{max}^{HSL_{DAG}} = 3.17\, mM\mu m^{-2}h^{-1} \tag{A.113}$$

$\langle 11c \rangle$ LD's MAG degradation by MGL.

$$[\boldsymbol{MAG}]_{\boldsymbol{LD}} \xrightarrow{MGL} [FA]_{cyt} \tag{A.114}$$

$$\boldsymbol{v_{MGL}} = V_{max}^{MGL} \cdot \boldsymbol{num_{LD}} \cdot \boldsymbol{Sur_{LD}} \cdot \frac{[\boldsymbol{MAG}]_{\boldsymbol{LD}}}{[\boldsymbol{MAG}]_{\boldsymbol{LD}} + [\boldsymbol{DAG}]_{\boldsymbol{LD}} + [\boldsymbol{TAG}]_{\boldsymbol{LD}}} \tag{A.115}$$

$$V_{max}^{MGL} = 3.6000 \cdot 10^{3}\, mM\mu m^{-2}h^{-1} \tag{A.116}$$

A.2.3 Stoichiometric matrix

The set of coupled ordinary differential equations that govern the system evolution. When indexing over the N_c different size classes it is used the index n.

$$\frac{d[FA]_{cyt}}{dt} = v_{CD36} + v_{diff} - v_{ACS} + \sum \boldsymbol{v_{ATGL}} + \sum \boldsymbol{v_{HSL}^{TAG}} + \\ + \sum \boldsymbol{v_{HSL}^{DAG}} + \sum \boldsymbol{v_{MGL}} \tag{A.117}$$

$$\frac{d[FA\text{-}CoA]_{cyt}}{dt} = v_{ACS} - (v_{GPAT} + v_{AGPAT}) \cdot \frac{Vol_{er}}{Vol_{cyt}} - v_{CPT1} - \\ - (v_{er(1)}^{DGAT1} + v_{er(2)}^{DGAT2}) \cdot \frac{Vol_{er}}{Vol_{cyt}} - \sum \boldsymbol{v_{TAG_{LD}}^{DGAT2}} \tag{A.118}$$

$$\frac{d[CoA]_{cyt}}{dt} = -v_{ACS} + (v_{GPAT} + v_{AGPAT}) \cdot \frac{Vol_{er}}{Vol_{cyt}} + v_{CPT1} + \\ + (v_{er(1)}^{DGAT1} + v_{er(2)}^{DGAT2}) \cdot \frac{Vol_{er}}{Vol_{cyt}} + \sum \boldsymbol{v_{TAG_{LD}}^{DGAT2}} \tag{A.119}$$

$$\frac{d[PC]_{er}}{dt} = v_{CPT} - v_{PCTP} \tag{A.120}$$

$$\frac{d[PA]_{er}}{dt} = v_{AGPAT} - v_{PAP} \tag{A.121}$$

$$\frac{d[LPA]_{er}}{dt} = v_{GPAT} - v_{AGPAT} \tag{A.122}$$

$$\frac{d[MAG]_{cyt}}{dt} = \sum (\boldsymbol{v_{HSL}^{DAG}} - \boldsymbol{v_{MGL}}) \tag{A.123}$$

$$\frac{d[DAG]_{er}}{dt} = v_{PAP} - v_{CPT} - v_{er(1)}^{DGAT1} - v_{er(2)}^{DGAT2} - \left(\sum \boldsymbol{v_{TAG_{LD}}^{DGAT2}}\right) \frac{Vol_{er}}{Vol_{cyt}} \tag{A.124}$$

$$\frac{d[\boldsymbol{MAG}]_{\boldsymbol{LD}}}{dt} = \boldsymbol{v_{HSL}^{DAG}} - \boldsymbol{v_{MGL}} \tag{A.125}$$

$$\frac{d[\boldsymbol{DAG}]_{\boldsymbol{LD}}}{dt} = \boldsymbol{v_{ATGL}} + \boldsymbol{v_{HSL}^{TAG}} - \boldsymbol{v_{HSL}^{DAG}} \tag{A.126}$$

$$\frac{d[TAG]_{er(1)}}{dt} = v_{er(1)}^{DGAT1} - v_{MTP} - v_{LD(1)}^{de\ novo} \frac{Vol_{ER}}{Vol_{cyt}} \tag{A.127}$$

$$\frac{d[TAG]_{er(2)}}{dt} = v_{er(2)}^{DGAT2} - v_{LD(2)}^{de\ novo} \frac{Vol_{ER}}{Vol_{cyt}} \tag{A.128}$$

$$\frac{d[\boldsymbol{TAG}]_{\boldsymbol{LD}}}{dt} = \boldsymbol{v_{TAG_{LD}}^{fill_{DGAT2}}} - \boldsymbol{v_{ATGL}} - \boldsymbol{v_{HSL}^{TAG}} + \boldsymbol{v_{TAG_{LD}}^{fus_{FSP27}}} + \delta_{n,1} v_{LD}^{TIP47} \tag{A.129}$$

$$\frac{d[TAG]_{nLD}}{dt} = v_{LD(1)}^{de\ novo} + v_{LD(2)}^{de\ novo} - v_{LD_1}^{TIP47} - v_{ATGL}^{TAG_{nLD}} \tag{A.130}$$

$$\frac{d[TAG\text{-}VLDL]_{er}}{dt} = v_{MTP}\frac{Vol_{er}}{Vol_{cyt}} - v_{VLDL_{ex}} \tag{A.131}$$

$$\frac{d[ApoB]_{er}}{dt} = v^{syn}_{ApoB} - v^{deg}_{ApoB} - v_{MTP} \tag{A.132}$$

$$\frac{d\boldsymbol{X_{LD}}}{dt} = \boldsymbol{v^{X}_{on}} - \boldsymbol{v^{X}_{off}}, \text{ (for } X = PLIN1,\ ADRP,\ TIP47) \tag{A.133}$$

$$\frac{d\boldsymbol{CGI58_{LD}}}{dt} = \boldsymbol{v^{CGI58}_{on}} - \boldsymbol{v^{CGI58}_{off}} - \boldsymbol{v^{CGI58_{on}}_{ATGL_{LD}}} + \boldsymbol{v^{CGI58_{off}}_{ATGL_{LD}}} \tag{A.134}$$

$$\frac{d\boldsymbol{ATGL_{LD}}}{dt} = \boldsymbol{v^{ATGL}_{on}} - \boldsymbol{v^{ATGL}_{off}} - \boldsymbol{v^{CGI58_{on}}_{ATGL_{LD}}} + \boldsymbol{v^{CGI58_{off}}_{ATGL_{LD}}} \tag{A.135}$$

$$\frac{d\boldsymbol{CGI58\text{-}ATGL_{LD}}}{dt} = \frac{d\boldsymbol{ATGL\text{-}CGI58_{LD}}}{dt} = \boldsymbol{v^{CGI58_{on}}_{ATGL_{LD}}} - \boldsymbol{v^{CGI58_{off}}_{ATGL_{LD}}} \tag{A.136}$$

$$\begin{aligned}\frac{d\boldsymbol{f^{free}_{LD}}}{dt} = &-\frac{d\boldsymbol{PLIN1_{LD}}}{dt} - \frac{d\boldsymbol{ADRP_{LD}}}{dt} - \\ &- \frac{d\boldsymbol{ADRP_{LD}}}{dt} - \frac{d\boldsymbol{TIP47_{LD}}}{dt} - \frac{d\boldsymbol{ATGL_{LD}}}{dt}\end{aligned} \tag{A.137}$$

where the compartment volumes multiplied and divided in order to conserve the mass are fixed constants

$$Vol_{cyt} = 2.81 \cdot 10^{-12}\ L \tag{A.138}$$

$$Vol_{er} = 2.1 \cdot 10^{-12}\ L \tag{A.139}$$

A.3 Change in number of LDs

This section will present calculations related to the changes in the number of lipid droplets (LDs) as they grow or shrink in response to variations in their lipid content, briefly explained in section 3.2.5.

Uptake of a given lipid amount $v_{TAG} \cdot \Delta t = \Delta TAG > 0$ into a LD owing to lipid (TAG) flow v_{TAG} within the time span Δt would actually increase the size of a LD. Since we are working with classes, both the uptake and release of TAG into or from lipid droplets (LDs) of size class n increase the total TAG content and the number of LDs in neighboring size classes ($n+1$ and $n-1$), while simultaneously decreasing the total TAG content and the number of LDs in size class n. This process is described by:

if $\Delta TAG_n \geq 0$:

$$LD_n \to LD_n \cdot (1 - \Delta LD_n) \tag{A.140}$$
$$LD_{n+1} \to LD_{n+1} \cdot (1 + \Delta LD_n) \tag{A.141}$$
$$\text{where } \Delta LD_n = \Delta TAG_n / (TAG_{n+1} - TAG_n) \tag{A.142}$$

if $\Delta TAG_n < 0$:

$$LD_n \to LD_n \cdot (1 - \Delta LD_n) \tag{A.143}$$
$$LD_{n-1} \to LD_{n-1} \cdot (1 + \Delta LD_n) \tag{A.144}$$
$$\text{where } \Delta LD_n = |\Delta TAG_n| / (TAG_n - TAG_{n-1}) \tag{A.145}$$

Working on the first case, when $\Delta TAG_n \geq 0$, we can reformulate the equations to include explicit time dependence, allowing us to derive an ODE that describes the evolution of the number of LDs.

$$LD_n(t + \Delta t) = LD_n(t) \cdot (1 - \Delta LD_n) \tag{A.146}$$
$$\Rightarrow \frac{LD_n(t + \Delta t) - LD_n(t)}{\Delta t} = -LD_n(t) \frac{\Delta LD_n}{\Delta t} \tag{A.147}$$
$$\Rightarrow \frac{\partial LD_n(t)}{\partial t} = -LD_n(t) \frac{v_{TAG}}{TAG_{n+1} - TAG_n} \tag{A.148}$$
$$LD_{n+1}(t + \Delta t) = LD_{n+1}(t) \cdot (1 + \Delta LD_n) \tag{A.149}$$
$$\Rightarrow \frac{LD_{n+1}(t + \Delta t) - LD_{n+1}(t)}{\Delta t} = LD_{n+1}(t) \frac{\Delta LD_{n+1}}{\Delta t} \tag{A.150}$$
$$\Rightarrow \frac{\partial LD_{n+1}(t)}{\partial t} = LD_{n+1}(t) \frac{v_{TAG}}{TAG_{n+1} - TAG_n} \tag{A.151}$$

Similarly, the same applies to the case of $\Delta TAG_n < 0$ leading to

$$\frac{\partial LD_n(t)}{\partial t} = -LD_n(t) \frac{|v_{TAG}|}{TAG_n - TAG_{n-1}} \tag{A.152}$$
$$\frac{\partial LD_{n-1}(t)}{\partial t} = LD_{n-1}(t) \frac{|v_{TAG}|}{TAG_n - TAG_{n-1}} \tag{A.153}$$

A.4 Results comparison

In this Appendix section some metrics are shown as comparison for the obtained results with the ones of the original paper [1].

A.4.1 Rates comparison

List of the steady state flux values. The metabolites and proteins residing in lipid droplets, whose fluxes are reported in bold, have been summed up over all the size classes. Fluxes expected here by the authors may be wrong considering the volume corrections to ensure mass conservation between compartments seem not to have been considered.

flux	obtained	expected	rel. diff.
v_{diff}	1.77	1.8	0.016
v_{CD36}	11.3	11.4	0.013
v_{ACS}	39.7	27.3	0.31
v_{CPT1}	3.56	0.82	0.77
v_{GPAT}	15.3	9.24	0.40
v_{AGPAT}	15.3	9.24	0.40
v_{PAP}	15.3	9.24	0.40
v_{CPT}	2.00	1.26	0.37
v_{PCTP}	2.00	1.26	0.37
v_{ApoB}^{syn}	3600	3600	-
v_{ApoB}^{deg}	3595	3600	0.001
v_{MTP}	4.64	1.90	0.59
$v_{VLDL_{ex}}$	3.47	1.90	0.45
$v_{TAG_{er(1)}}^{DGAT1}$	7.41	3.89	0.48
$v_{TAG_{er(2)}}^{DGAT2}$	0.30	0.81	1.73
$\boldsymbol{v_{TAG_{LD}}^{DGAT2}}$	7.49	3.28	0.56
$\boldsymbol{v_{TAG_{LD}}^{fill_{DGAT2}}}$	7.49	4.68	0.38
$v_{LD(1)}^{de\ novo}$	3.72	1.99	0.46
$v_{LD(2)}^{de\ novo}$	0.39	0.81	1.04
$\boldsymbol{v_{on}^{PLIN1}}$	2585	2220	0.14
$\boldsymbol{v_{off}^{PLIN1}}$	2585	2220	0.14
$\boldsymbol{v_{on}^{ADRP}}$	60918	57400	0.058
$\boldsymbol{v_{off}^{ADRP}}$	60918	57400	0.058
$\boldsymbol{v_{on}^{TIP47}}$	153	157	0.023
$\boldsymbol{v_{off}^{TIP47}}$	153	157	0.023
v_{nLD}^{TIP47}	2.06	1.40	0.32
$\boldsymbol{v_{on}^{CGI58}}$	15512	14400	0.072
$\boldsymbol{v_{off}^{CGI58}}$	15512	14400	0.072
$\boldsymbol{v_{on}^{ATGL}}$	5.54	62.3	10.2
$\boldsymbol{v_{off}^{ATGL}}$	5.54	62.30	10.2
$\boldsymbol{v_{ATGL_{LD}}^{CGI58_{on}}}$	0.26	0.401	0.53
$\boldsymbol{v_{ATGL_{LD}}^{CGI58_{off}}}$	0.26	0.401	0.53
$\boldsymbol{v_{ATGL}}$	8.81	4.52	0.49
v_{nLD}^{ATGL}	2.06	1.4	0.32
$\boldsymbol{v_{HSK}^{TAG}}$	0.08	0.154	0.85
$\boldsymbol{v_{HSL}^{DAG}}$	8.89	4.68	0.47
$\boldsymbol{v_{MGL}}$	8.89	4.68	0.47
$\boldsymbol{v_{TAG_{LD}}^{fus}}$	-0.52	1e-08	1.00

Table A.1: Relative difference between the obtained and expected (from [1]) fluxes stable values for the rat hepatocyte parameters. The fluxes in bold, namely vectors, are summation over all the LD classes.

References

[1] C. Wallstab, D. Eleftheriadou, T. Schulz, G. Damm, D. Seehofer, J. Borlak, H.-G. Holzhütter, and N. Berndt, "A unifying mathematical model of lipid droplet metabolism reveals key molecular players in the development of hepatic steatosis," *The FEBS Journal*, vol. 284, no. 19, pp. 3245–3261, 2017. [Online]. Available: https://febs.onlinelibrary.wiley.com/doi/abs/10.1111/febs.14189

[2] H. Devarbhavi, S. K. Asrani, J. P. Arab, B. Niu, D. Aguirre, I. K. de Araujo, I. Graupera, and H. Pham, "Global burden of liver disease: 2023 update," *Journal of Hepatology*, vol. 79, no. 2, pp. 516–537, 2023.

[3] J. Cabezas, M. Mayorga, and J. Crespo, "Nonalcoholic fatty liver disease: A pathological view," in *Liver Biopsy*, N. Tagaya, Ed. Rijeka: IntechOpen, 2012, ch. 8. [Online]. Available: https://doi.org/10.5772/52622

[4] R. Basu, M. Noureddin, and J. M. Clark, "Nonalcoholic fatty liver disease: Review of management for primary care providers," *Mayo Clinic Proceedings*, vol. 97, no. 9, pp. 1700–1716, 2022.

[5] M. E. Rinella, J. V. Lazarus, V. Ratziu, S. M. Francque, A. J. Sanyal, F. Kanwal, D. R. Romero, M. F. Abdelmalek, Q. M. Anstee, J. P. Arab, M. Arrese, R. Bataller, U. H. W. Beuers, J. Boursier, E. Bugianesi, C. D. Byrne, G. E. C. Narro, A. Chowdhury, H. Cortez-Pinto, D. R. Cryer, K. Cusi, M. El-Kassas, S. Klein, W. Eskridge, J.-G. Fan, S. Gawrieh, C. D. Guy, S. A. Harrison, S.-U. Kim, B. G. P. Koot, M. Korenjak, K. V. Kowdley, F. Lacaille, R. S. Loomba, R. Mitchell-Thain, T. R. Morgan, E. Powell, M. Roden, M. Romero-Gomez, M. Silva, S. P. Singh, S. Sookoian, C. W. Spearman, D. G. Tiniakos, L. Valenti, M. B. Vos, V. W. Wong, S. A. Xanthakos, Y. Yılmaz, Z. Younossi, A. Hobbs, M. Villota-Rivas, and P. N. Newsome, "A multisociety delphi consensus statement on new fatty liver disease nomenclature," *Hepatology (Baltimore, Md.)*, vol. 78, pp. 1966 – 1986, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:259260747

[6] N. Chalasani, Z. Younossi, J. E. Lavine, A. M. Diehl, E. M. Brunt, K. Cusi, M. Charlton, and A. J. Sanyal, "The diagnosis and management of non-alcoholic fatty liver disease: Practice guideline by the american association for the study of liver diseases, american college of gastroenterology, and the american gastroenterological association," *Hepatology*, vol. 55, no. 6, pp. 2005–2023, 2012. [Online]. Available: https://aasldpubs.onlinelibrary.wiley.com/doi/abs/10.1002/hep.25762

[7] D. Ferguson and B. N. Finck, "Emerging therapeutic approaches for the treatment of nafld and type 2 diabetes mellitus," *Nature Reviews Endocrinology*, vol. 17, no. 8, pp. 484–495, 2021. [Online]. Available: https://doi.org/10.1038/s41574-021-00507-z

[8] D. L. Brasaemle and N. E. Wolins, "Packaging of fat: An evolving model of lipid droplet assembly and expansion*," *Journal of Biological Chemistry*, vol. 287, no. 4, pp. 2273–2279, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021925820531888

[9] B. Kubickova and M. N. Jacobs, "Development of a reference and proficiency chemical list for human steatosis endpoints in vitro," *Frontiers in Endocrinology*, vol. 14, 2023. [Online]. Available: https://www.frontiersin.org/journals/endocrinology/articles/10.3389/fendo.2023.1126880

[10] D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 6th ed. W. H. Freeman, 2013.

[11] T. O. Eichmann and O. L. Knittelfelder, *Glycerolipids: Tri-, Di-, and Monoacylglycerols*. Dordrecht: Springer Netherlands, 2015, pp. 1–4. [Online]. Available: https://doi.org/10.1007/978-94-007-7864-1_57-1

[12] J. A. Olzmann and P. Carvalho, "Dynamics and functions of lipid droplets," *Nature Reviews Molecular Cell Biology*, vol. 20, no. 3, pp. 137–155, 2019. [Online]. Available: https://doi.org/10.1038/s41580-018-0085-z

[13] J. Yu and P. Li, "The size matters: regulation of lipid storage by lipid droplet dynamics," *Science China Life Sciences*, vol. 60, no. 1, pp. 46–56, 2017.

[14] F. Wilfling, J. T. Haas, T. C. Walther, and R. V. J. Farese, "Lipid droplet biogenesis," *Current Opinion in Cell Biology*, vol. 29, pp. 39–45, 2014. [Online]. Available: https://doi.org/10.1016/j.ceb.2014.03.008

[15] D. G. Mashek, "Hepatic lipid droplets: A balancing act between energy storage and metabolic dysfunction in nafld," *Molecular Metabolism*, vol. 50, p. 101115, 2021, special issue on Non-alcoholic fatty liver disease. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212877820301897

[16] H. Itabe, T. Yamaguchi, S. Nimura, and N. Sasabe, "Perilipins: a diversity of intracellular lipid droplet proteins," *Lipids in Health and Disease*, vol. 16, p. 83, 2017. [Online]. Available: https://doi.org/10.1186/s12944-017-0473-y

[17] M. Slayton, A. Gupta, B. Balakrishnan, and V. Puri, "Cide proteins in human health and disease," *Cells*, vol. 8, no. 3, p. 238, 2019.

[18] L. Xu, L. Li, L. Wu, P. Li, and F.-J. Chen, "Cide proteins and their regulatory mechanisms in lipid droplet fusion and growth," *FEBS Letters*, vol. 598, no. 10, pp. 1154–1169, 2024. [Online]. Available: https://febs.onlinelibrary.wiley.com/doi/abs/10.1002/1873-3468.14823

[19] N. Berndt, J. Eckstein, N. Heucke, R. Gajowski, M. Stockmann, D. Meierhofer, and H.-G. Holzhütter, "Characterization of lipid and lipid droplet metabolism in human hcc," *Cells*, vol. 8, no. 5, 2019. [Online]. Available: https://www.mdpi.com/2073-4409/8/5/512

[20] N. Berndt, S. Bulik, I. Wallach, T. Wünsch, M. König, M. Stockmann, D. Meierhofer, and H.-G. Holzhütter, "Hepatokin1 is a biochemistry-based model of liver metabolism for applications in medicine and pharmacology," *Nature Communications*, vol. 9, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:49311152

[21] H. Wang, M. V. Airola, and K. Reue, "How lipid droplets "tag" along: Glycerolipid synthetic enzymes and lipid storage," *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, vol. 1862, no. 10, Part B, pp. 1131–1145, 2017, recent Advances in Lipid Droplet Biology. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1388198117301142

[22] J. Gong, Z. Sun, L. Wu, W. Xu, N. Schieber, D. Xu, G. Shui, H. Yang, R. G. Parton, and P. Li, "Fsp27 promotes lipid droplet growth by lipid exchange and transfer at lipid droplet contact sites," *The Journal of Cell Biology*, vol. 195, no. 6, pp. 953–963, 2011. [Online]. Available: https://doi.org/10.1083/jcb.201104142

[23] I. Ganeva, K. Lim, J. Boulanger, P. C. Hoffmann, D. B. Savage, and W. Kukulski, "The architecture of cidec-mediated interfaces between lipid droplets," *bioRxiv*, 2021. [Online]. Available: https://www.biorxiv.org/content/early/2021/03/23/2021.03.23.436555

[24] A. Lass, R. Zimmermann, G. Haemmerle, M. Riederer, G. Schoiswohl, M. Schweiger, P. Kienesberger, J. G. Strauss, G. Gorkiewicz, and R. Zechner, "Adipose triglyceride lipase-mediated lipolysis of cellular fat stores is activated by cgi-58 and defective in chanarin-dorfman syndrome," *Cell Metabolism*, vol. 3, no. 5, pp. 309–319, 2006.

[25] A. Boeszoermenyi, H. M. Nagy, H. Arthanari, C. J. Pillip, H. Lindermuth, R. E. Luna, G. Wagner, R. Zechner, K. Zangger, and M. Oberer, "Structure of a cgi-58 motif provides the molecular basis of lipid droplet anchoring," *The Journal of Biological Chemistry*, vol. 290, no. 44, pp. 26 361–26 372, 2015.

[26] M. D. Morris, "Factorial sampling plans for preliminary computational experiments," *Technometrics*, vol. 33, no. 2, p. 161–174, apr 1991. [Online]. Available: https://doi.org/10.2307/1269043

[27] S. Tarantola, F. Ferretti, S. Lo Piano, M. Kozlova, A. Lachi, R. Rosati, A. Puy, P. Roy, G. Vannucci, M. Kuc-Czarnecka, and A. Saltelli, "An annotated timeline of sensitivity analysis," *Environmental Modelling & Software*, vol. 174, p. 105977, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364815224000380

[28] A. Saltelli, K. Chan, and E. M. Scott, *Sensitivity Analysis*. Wiley, 2009.

[29] J. Herman and W. Usher, "Salib: An open-source python library for sensitivity analysis," *Journal of Open Source Software*, vol. 2, no. 9, p. 97, 2017. [Online]. Available: https://doi.org/10.21105/joss.00097

[30] F. Campolongo, J. Cariboni, and A. Saltelli, "An effective screening design for sensitivity analysis of large models," *Environmental Modelling & Software*, vol. 22, no. 10, pp. 1509–1518, 2007, modelling, computer-assisted simulations, and mapping of dangerous phenomena for hazard assessment. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364815206002805

[31] F. Campolongo, A. Saltelli, and J. Cariboni, "From screening to quantitative sensitivity analysis. a unified approach," *Computer Physics Communications*, vol. 182, no. 4, pp.

978–988, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010465510005321

[32] I. Sobol′, "Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates," *Mathematics and Computers in Simulation*, vol. 55, no. 1, pp. 271–280, 2001, the Second IMACS Seminar on Monte Carlo Methods. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378475400002706

[33] T. Homma and A. Saltelli, "Importance measures in global sensitivity analysis of nonlinear models," *Reliability Engineering & System Safety*, vol. 52, no. 1, pp. 1–17, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0951832096000026

[34] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," 2019. [Online]. Available: https://arxiv.org/abs/1912.01703

Acknowledgments

I would like to express my gratitude to the SimbioTX team of Inria Centre de Saclay for having me in months of my internship there, and in particular to my supervisors Dirk Drasdo and Jieling Zhao for having trust in me and for providing valuable resources and support during this internship. I am equally grateful to my internal university supervisor, Dr. Marco Baiesi for his advice and assistance during this process.

Finally, I am also grateful to the staff and colleagues at the research center for their technical assistance and collaboration, which greatly contributed to the success of this project.