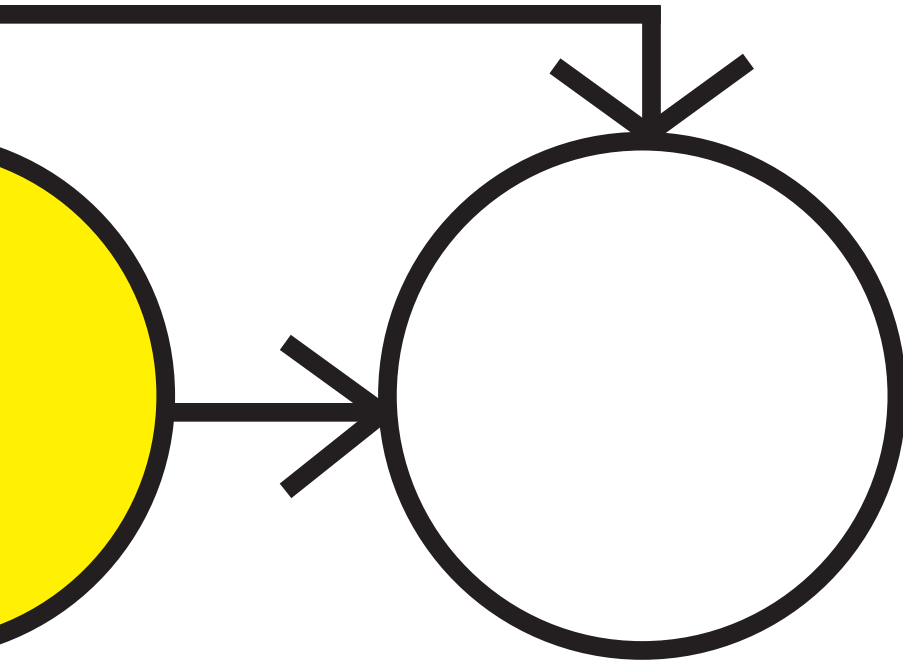


A physiologically based
kinetic model for the
prediction of plasma cholesterol
concentrations in mice and man



Niek C.A. van de Pas

A physiologically based kinetic model
for the prediction of plasma cholesterol
concentrations in mice and man

Niek C.A. van de Pas

Thesis committee

Thesis supervisors

Prof. Dr. Ir. I.M.C.M. Rietjens
Professor of Toxicology
Wageningen University

Prof. Dr. R.A. Woutersen
Senior Scientist; Toxicologic Pathologist TNO, Zeist &
Professor of Toxicology
Wageningen University

Thesis co-supervisor

Dr. ir. A.A. de Graaf
Senior Scientist Fluxomics
TNO, Zeist

Other members

Prof. dr. M.R. Müller, Wageningen University
Prof. dr. L.M. Havekes, Leiden University Medical Center
Prof. dr. A.K. Groen, University of Groningen
Dr. A. P. Freidig, AMT Biopharma, Amsterdam

This research was conducted under the auspices of the graduate school VLAG

A physiologically based kinetic model
for the prediction of plasma cholesterol
concentrations in mice and man

Niek C.A. van de Pas

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University

by the authority of the Rector Magnificus

Prof. dr. M.J. Kropff,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Friday 23 December 2011 at 4 pm in the Aula

Niek C.A. van de Pas

A physiologically based kinetic model for the prediction of plasma cholesterol concentrations in mice and man

Thesis, Wageningen University, Wageningen, NL (2011) with with summary in Dutch

ISBN: 978-94-6173-125-8

For Emmely

Table of contents

| | |
|----------------------------------------------------------------------------------------------------------------------------------|-----|
| Chapter 1. General introduction | 9 |
| Chapter 2. Systematic construction of a conceptual minimal model of plasma cholesterol levels based on knockout mouse phenotypes | 29 |
| Chapter 3. A physiologically-based kinetic model for the prediction of plasma cholesterol concentrations in the mouse | 57 |
| Chapter 4. A physiologically based in silico kinetic model predicting plasma cholesterol concentrations in humans | 87 |
| Chapter 5. Predicting individual responses to pravastatin using a kinetic model for plasma cholesterol concentrations | 115 |
| Chapter 6. Predicting individual response to CETP inhibitors using a kinetic model for cholesterol concentrations | 141 |
| Chapter 7. Summary, general discussion, and future perspectives | 159 |
| Abbreviations | 181 |
| Nederlandse samenvatting | 183 |
| List of publications | 191 |
| Dankwoord | 193 |
| Curriculum Vitae | 195 |
| List of publications | 197 |
| Overview of completed training activities | 200 |



CHAPTER

1

General introduction



Preface

This thesis is entitled “A physiologically based kinetic model for the prediction of plasma cholesterol concentrations in mice and man.” Two main phrases in this title are “plasma cholesterol” and “physiologically based kinetic model”. This chapter introduces both. The paragraph “Cholesterol” introduces cholesterol and its importance as a risk factor in cardiovascular disease (CVD). The general concept of modeling is introduced in the paragraph “Mathematical modeling”. The paragraph “Mathematical modeling of cholesterol” describes the present state of modeling efforts on cholesterol metabolism. The last section of this introduction describes the outline of the thesis.

Cholesterol

Cholesterol is a hydrophobic molecule that plays important roles in the membranes of mammalian cells and the transport of triglycerides (1). Cholesterol attracts attention because of the role of elevated plasma cholesterol concentrations as a risk factor in cardiovascular disease (CVD) (2,3). CVD is a class of diseases that involves the heart or the vascular system. It includes cardiovascular events like myocardial infarction (heart attack) and stroke. In 2010, CVD caused almost 40.000 deaths in the Netherlands (4). This is on average 4.5 deaths per hour.

The main cause of cardiovascular events in Western societies is atherosclerosis (3). This is a disease of lipid accumulations (plaques) in the walls of the main arteries. Plaques result in narrowing of blood vessels and reduced blood flow to important organs, like heart and lung (3,5).

These plaques contain cholesterol derived from lipoproteins in plasma. These lipoproteins are cholesterol-carrying particles that additionally contain proteins and triglycerides. Lipoproteins are discriminated based on their density resulting in definition of Very Low Density Lipoprotein (VLDL), Intermediate Density Lipoprotein (IDL), Low Density Lipoprotein (LDL), and High Density Lipoprotein (HDL). More specifically, plaques result from LDL particles that diffuse through the endothelial layer of arteries (3). The cholesterol in these LDL particles present in plasma will be referred to as LDL-C.

The body has a mechanism to counteract this LDL-C inflow in the arterial wall: Monocytes, a type of white blood cells, migrate through the endothelial layer (2,6), where they differentiate into macrophages (2), take up LDL, and transfer the cholesterol to HDL. Next, HDL particles

transport cholesterol to the liver (7), after which cholesterol is secreted in the bile and excreted in feces. The process of cholesterol transport from the vessel wall to the feces is called reverse cholesterol transport. The cholesterol in the HDL particles present in plasma will be referred to as HDL-C.

12

In healthy individuals, a balance exists between the LDL-C diffusing into the vessel wall and the cholesterol transport out of the wall to HDL. If, however, the concentration of LDL-C in plasma is elevated, the influx of LDL-C will dominate over the outflow and then cholesterol might accumulate into the wall and form a plaque.

Plaques can cause cardiovascular events as follows. Usually, these events are initiated by a thinning of the cap on the plaque, due to erosion by the blood flow. When the cap is thin and unstable, this may result into a cap rupture (3). Platelets in the blood recognize this rupture as an injury and coagulate to form a thrombus inside the vessel. This thrombus can occlude the artery, causing a stroke or cardiac infarction (3,5).

Since treatment of cardiovascular events is not always effective, prevention is important. To select which subjects require preventive measures, predictive biomarkers are needed. The LDL-C concentration is such a biomarker: individuals with high LDL-C have a high risk of getting CVD (2,3). HDL-C is also used as a biomarker for cardiovascular events: individuals with high HDL-C have a low risk of getting CVD (8). For this reason, HDL-C is referred to as “good cholesterol”, while LDL-C is referred to as “bad cholesterol.” Other risk factors for CVD risk include smoking, elevated body weight, and elevated blood pressure.

After the discovery of cholesterol as a biomarker for CVD, pharmaceutical industry responded by developing drugs that target LDL-C as a surrogate biomarker for CVD (9). Several types of these lipoprotein-modifying drugs are on the market, and more are in (pre) clinical development (5). The best known drug class for LDL-C lowering is the class of 3-hydroxy-3-methyl-glutaryl-CoA reductase inhibitors, also called statins (10). These drugs block synthesis of cholesterol in the liver. Upon statin administration, the liver responds by increasing the expression of the LDL-receptor (LDLR) to increase uptake of LDL-C (11). This results in a reduction of LDL-C. Other strategies of LDL-C lowering include inhibition of intestinal cholesterol absorption (e.g. by ezetimibe or dietary plant sterols and stanols) (12-14) and reduction of the production of lipoproteins in the liver (e.g. by niacin) (15). Currently, the pharmaceutical industry is also developing drugs that aim

to increase HDL-C, for example Cholesterol Ester Transport Protein (CETP) inhibitors such as torcetrapib (16). These drugs are often used in combination with a statin to simultaneously affect HDL-C and LDL-C.

Unfortunately, current therapies are often not sufficiently effective. It is known that only 40% of the individuals treated for a high LDL-C concentration, reach their LDL-C targets (17). The lack of treatment success warrants the development of new treatment strategies. More insight into the mechanisms that underlie differences in plasma cholesterol concentrations and statin treatment success might accelerate this development.

It is known that the response to statins is associated with genetic variants in cholesterol-associated genes. This implies that differences in cholesterol metabolism influence the response to statins. Subjects with a different genetic makeup might react to drugs with different targets in cholesterol metabolism in a different way.

Fortunately, much is already known about the metabolism of cholesterol on the molecular and the cellular concentration (11,18-20). It is for example known that cholesterol concentrations are influenced by hereditary factors (21). A large Genome Wide Association Study (GWAS) with more than 100,000 subjects found 95 genetic loci determining plasma lipid concentrations (22). Yet together, these loci only explained 10-12% of the total variance of cholesterol concentrations and 25–30% of the variance attributed to genetic factors (22). This means that a large part of the variation in cholesterol concentrations remains unexplained. More specific, questions to be answered include the following:

- What are the most important biochemical reactions in the body for determining plasma cholesterol concentrations and what is the relation between these reactions?
- What causes the large individual variation in the cholesterol response to cholesterol lowering therapies?
- What is the influence of genetic mutations and pharmacological interventions on cholesterol concentrations in non-plasma compartments?
- What is the effect of combinations of cholesterol lowering drugs as compared to the effect of single drugs?

These questions can be addressed with a predictive simulation tool. As introduced in the next paragraph, mathematical models are such a tool. Therefore, the primary aim of this thesis was to develop a mathematical model, including the most important reactions that determine plasma cholesterol concentrations, that allows the prediction of the effects of genetic, pharmaceutical, and nutritional variations on these concentrations. Secondly, to use that model to address the questions raised above.

Mathematical modeling

One of the dictionary entries for the word “model” is “A simplified description of a system, process, etc., put forward as a basis for theoretical understanding” (23). When we apply this definition to our everyday life, everyone is using models. Not only people who use model trains and other scale models; but also those who use a subway map, talk about left wing political parties, or about a 4-3-3 formation in football. In every day life, the purpose of these models is often brevity. It takes a lot of time to explain all positions of a political party in terms of immigration, protection of the environment, and equality in wealth; instead it is often more convenient to call it a left, or right wing political party.

In science, people are more explicit in their model use, and examples can be found in the field of molecular models (24), *in vitro* models (25), and mouse models (26). The power of modeling is that models allow us to predict scenarios that are hard to measure, for ethical, and/or practical reasons. Mouse models, for example, are used when (invasive) experiments in humans are considered unethical.

Another type of models are mathematical models, which are descriptions of a system in a mathematical language. These models are usually made in situations where the human mind cannot oversee all implications of the assumptions made. A mathematical model can integrate all assumptions and test their consistency, implications, and validity. A model will help to select useful hypotheses, suggest useful experiments, and obtain mechanistic insight.

This thesis will focus on mathematical models of biochemical reaction systems having cholesterol plasma concentrations as model output. These models are referred to as kinetic models. A special type

of kinetic modeling is physiologically based kinetic (PBK) modeling. This type of modeling includes “discrete tissues (...) with appropriate volumes, blood flows, and pathways for metabolism of (...) chemicals (27).” PBK modeling’s close link with physiology allows easier interpretation and, therefore, improved understanding of the system.

In general, PBK modeling (like most scientific modeling) will follow a procedure called the modeling cycle (see Figure 1.1). Other versions of this modeling cycle have been published (e.g. (28,29). Here, the modeling cycle contains 8 steps. In other occasions, steps may be combined or split up, leading to a different number of steps. In general, the procedure itself is consistent.

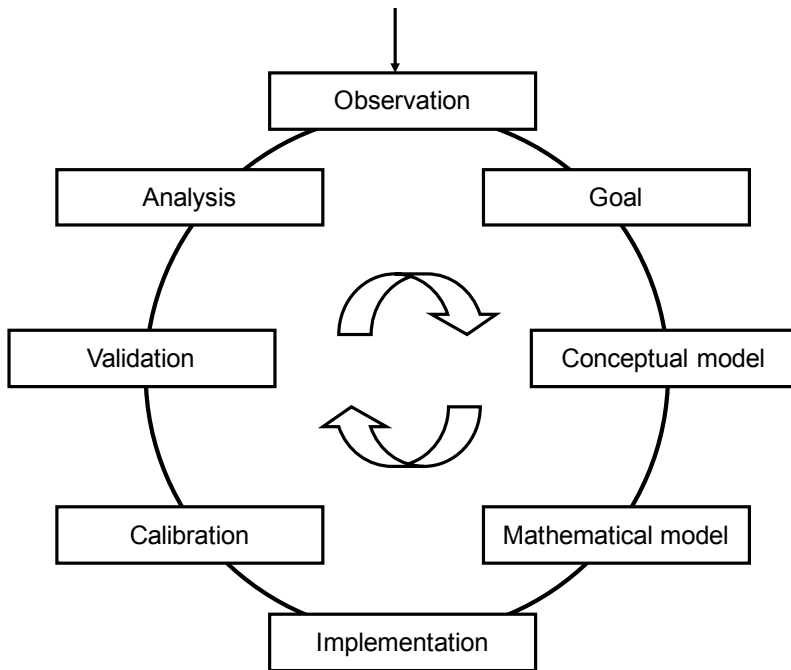


Figure 1.1. Schematic overview of the modeling cycle.

Step 1) Experimental observations.

In the first step, a pattern is observed and identified to be important or interesting, but it is not fully understood or it is so complex that human intuition is not sufficient to see all implications. In our case the paragraph “Cholesterol” in this chapter describes these experimental observations.

Step 2) Setting a goal.

In the second step, the goal of the model is formulated. Ideally, models are not (only) made for fun, but are made with a specific goal in mind, in most cases to answer a specific question. Examples of these questions can be:

- Will there be rain tomorrow?
- What is the maximum velocity of a ball dropped from 1 meter high?
- Will the sea level rise in the next century?

Or in biology:

- Which factors determine penicillin production (30)?
- How does the bioactivation and detoxification of estragole depend on estragole dose (31)?

In our case, the goals are given at the end of the paragraph “Cholesterol” in this chapter.

Step 3) Defining a conceptual model.

After defining the goal, it is time to start building the model. A model that contains every piece of detail requires lots of data to construct and, due to excessive complexity, does not provide any insight. Therefore, for a model to be useful it must be a simplification of reality and thus some factors have to be left out. In this third step it should be decided which factors are important enough to be included in the model. These factors are summarized in a conceptual model. It is most convenient to represent the conceptual model in a graphical representation, often called “model structure”. Figure 1.2 shows an example of a conceptual model. This model integrates pools of chemicals (represented as boxes) and reactions (represented as arrows).

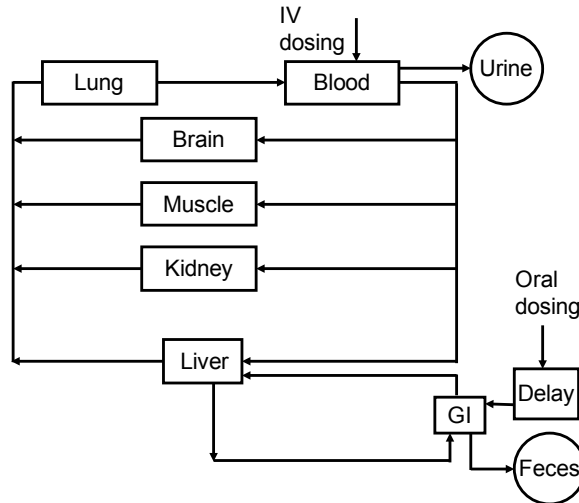


Figure 1.2. Example of a conceptual model, adapted from (32).

Step 4) Mathematical model formulation.

In this step the conceptual model is converted into mathematical equations. All types of equations can be used in this step. Models can be as simple as $F = m \cdot a$ and as complex as a set of thousands of equations (33).

In PBK modeling, a change in the amount of a chemical X_i in time t can be described as follows:

$$\frac{dX_i}{dt} = v_i - v_j \quad (\text{Eqn. 1.1})$$

where v_i and v_j represent the rate of reactions i and j . Equation 1.1 can be read as: in a small time step, the net change of chemical X_i is the difference of the amount of X_i produced, minus the amount of X_i consumed.

If it is assumed that the rate v_i is constant and that the rate v_j is proportional to the amount of the chemical X_j , then equation 1.1 leads to equation 1.2.

$$\frac{dX_i}{dt} = k_i - k_j X_i \quad (\text{Eqn. 1.2})$$

where k_i and k_j are reaction rate constants. The parameter k_j is called a first order rate constant, since it appears in the rate expression multiplied by X_i to the power 1. The parameter k_i is called a zero order rate constant, since it appears in the rate expression multiplied by X_i to the power 0. The form of the rate equations (zero order, first order) is also called the kinetic format of the model.

Sometimes small differences in the model can lead to large differences in model predictions. A method to study this is to produce a library of submodels, each one slightly different from the others. The difference in predictions of the submodels is a measure of the impact of these small changes. This approach is called ensemble modeling, because it combines different submodels. This thesis will use ensemble modeling in a similar approach as outlined by Kuepfer et al. (34).

Kuepfer et al. (34) applied ensemble modeling to the TOR-pathway of *Saccharomyces cerevisiae*. For this pathway, and generally in biology, there are large uncertainties in the network structure (which components do interact and which do not) and kinetics of the reactions. Kuepfer et al. have developed a library of submodels each having a different network structure. Every model was validated independently and as could be expected some submodels better described the experimental data than others did. It is likely that the network structures of the best performing submodels better reflect the network structures of the biological system under study than the network structures of the worst performing submodels.

Generally in biology, not only interactions are unknown, but also the kinetic format. This kinetic format describes how the reaction rate depends on the measured concentrations. Determining the kinetic order of biological processes *in vivo* requires the measurements of reaction rates in various situations which in many cases are not feasible. We used a similar approach to the approach of Kuepfer et al. Whereas Kuepfer et al. used submodels with different network structure, we used submodels with different kinetics. See Chapter 3 for more details.

Step 5) Model implementation.

For a few cases, writing model equations is enough to see all implications just by looking at them. In most cases, however, model equations are too complex to grasp in that way. In these cases, model equations are implemented in specific software to allow more extensive operations and

analysis. In this thesis, all model implementations were performed in MATLAB (35).

In PBK modeling, it is possible to calculate the pool sizes with two different methods: by analytical integration of Eqn. 1.2 or by using so called stepwise integration with a computer. In the latter procedure, one assumes an initial amount of chemical in every pool in the system. With these amounts, reaction rates v_i can be calculated. In a minor time step (Δt), a small amount of chemical ($v_i \Delta t$) is changed in the pools. We now can calculate the amount of chemical after this time step. In the second small time step, reaction rates are slightly different. After many of these time steps, the concentration development in time for each pool is obtained. This simulation can be performed to a specific point in time, or until a situation occurs where none of the concentrations change in time anymore. This situation is referred to as a 'steady state'. In the present thesis we applied the stepwise integration approach to solve the equations.

Step 6) Calibration of the model.

Usually, the implemented model contains several constants: the parameters. In order to reach the goal set in step 2, the values of these parameters must be known. Parameter values can be measured, taken from literature, or fitted to experimental data. In the last case, the parameter combination is used that yield model predictions that best resemble experimental data. The process of setting parameter values is called "model calibration" or "model training". A data set used in the calibration step is called the calibration data set or training data set.

Step 7) Validation of the model.

The model developed needs to include elements of reality. Frequently the question is raised how good the model mimics reality. There are two strategies to answer this question. One strategy is to discuss the validity of all model assumptions. This answer will become very lengthy and is often very difficult to give. Often it is more convenient to answer this question by the other strategy: compare model predictions with reality. If model predictions are similar to available experimental data, then the model might be valid in other situations not yet studied by experiments.

If the comparison is made with the data used for model calibration (step 6), it is possible that the model is only good for these specific data. Consequently, validation with a data set independent of the training set answers the question more convincingly. The independent data set used in

this validation step is called the validation set.

Step 8) Model analysis.

20

In the next step, the validated model can be used to answer the questions raised in the second step of the modeling cycle, and additionally reveal unexpected model characteristics. This step is called model analysis, and can include sensitivity analysis, stability analysis, or time scale analysis. Sensitivity analysis can identify factors that influence model predictions (32). Stability analysis can identify instabilities, oscillations, and switch-like behavior (36,37). Time scale analysis can identify how long it takes before events occur (38). This provides hints for sampling time and frequency in experiments.

Another way to analyze mathematical models is to simulate various scenarios and evaluate model predictions. In this thesis, these scenarios will include genetic mutations and drug treatment.

The next step is then to validate the determined model characteristics and the generated model predictions to experimental data. Usually, this involves the performance of new experiments. This leads to new observations, completing the modeling cycle. Thereafter, the modeling cycle can be repeated.

The next section describes the current state of modeling of cholesterol metabolism.

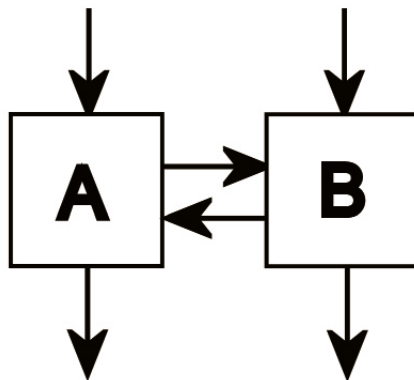


Figure 1.3. Structure of the model of Goodman and Noble, adapted from (48).

Mathematical modeling of cholesterol

The aim of this thesis was to develop a mathematical model, including the most important reactions that determine plasma cholesterol concentrations, that allows the prediction of the effects of genetic, pharmaceutical, and nutritional variations on these concentrations.

From the aim specified, at least three model requirements arise, the model should 1) include a plasma compartment, 2) contain the important genes and targets of pharmaceutical and nutritional interventions, and 3) be simple enough to calculate all model parameters. Below, mathematical models from literature are reviewed and compared with the list of requirements.

Cholesterol metabolism was included in modeling: in mouse models (26,39), in conceptual modeling (5,40-47), but also in mathematical modeling (48-58). Mathematical models were already made in the 1960's. Goodman and Noble (48), for example, used a two compartment model shown in Figure 1.3. This model was able to describe the decay of labeled cholesterol in plasma after an intravenous injection. The model contains a rapidly turning over pool of plasma cholesterol (A) and a slowly turning over pool of plasma cholesterol (B), but the model does not include drug targets and reactions are not linked to genetic factors.

In recent years, more sophisticated models were constructed to analyze isotope tracer data (reviewed by De Graaf and Van Schalkwijk (49)). These models do not focus on cholesterol, but merely on proteins in the lipoproteins that carry cholesterol. This type of models aid interpretation of isotope tracer experiments, as they predict until then inaccessible biological parameters that underlie lipoprotein metabolism (49-52). Van Schalkwijk et al. (52) have found that some of these parameters aid the diagnosis of CVD risk. Pools in these models are directly biologically relevant, but these models still lack known drug targets, such as hepatic cholesterol synthesis. Therefore, these models can predict the effect of genetic, pharmaceutical, and nutritional differences on plasma cholesterol only to a very limited extent.

Hübner et al. (53) have developed a model solely of plasma lipoproteins that considers the entire protein and lipid composition spectrum of individual lipoprotein complexes, of both ApoB-100- and ApoA1- containing lipoproteins. This model also lacks many known drug targets and is consequently not ideal for our goal.

Three models were reported that focus primarily on the molecular

and cellular level, particularly of liver cells (54-56). Figure 1.4 gives the structure of one of these models. As seen in Figure 1.4, the model contains various intracellular reactions, and lacks several reactions that are important for the determination of plasma cholesterol concentrations, like dietary cholesterol absorption.

22

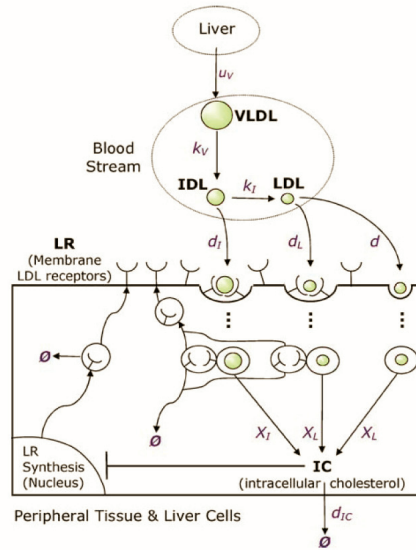


Figure 1.4. Structure of a model for cellular cholesterol. The model includes LDL receptors (LR), intracellular cholesterol IC, extracellular cholesterol (VLDL, IDL, and LDL). Taken from (54) with permission.

The models described in the literature that are including drug targets and plasma cholesterol are rather detailed and tend to be very complex, with many reactions and parameters (57,58). Reich and Knoblauch (57) described an elegant model including various cholesterol fractions in blood (Figure 1.5). Due to their complexity, these models are hard to calibrate and validate. Reich and Knoblauch, for example, did not explain why certain parameter values were chosen (57). Also, despite their complexity, many relevant drug targets are not included in these models.

Thus, to the best of our knowledge, there exists no adequate model that meets all our requirements. Consequently, such a model was developed and validated in this thesis.

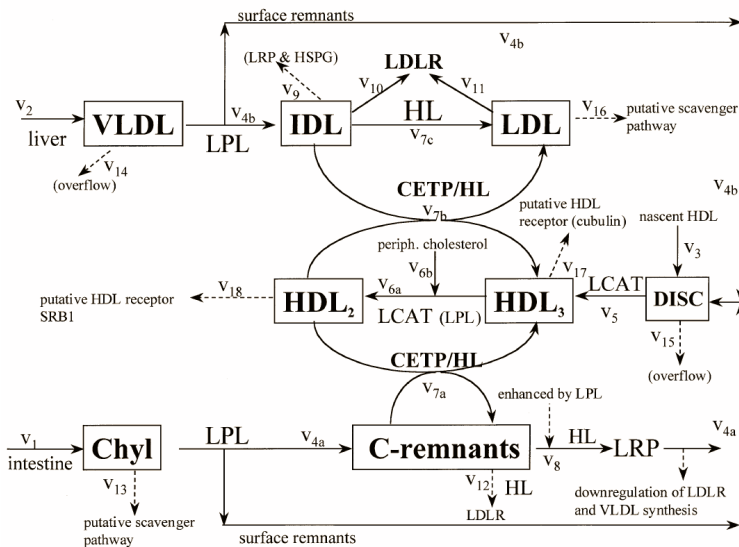


Figure 1.5. Structure of a complex model of plasma cholesterol. Boxes represent pools of the model, in this case various plasma species. Reactions include targets of CETP inhibitors (reaction 7), but not dietary cholesterol absorption, the target of ezetimibe. Taken from (57) with permission.

Outline of the thesis

The aim of this thesis was to develop a PBK model that allows the prediction of the effect of genetic, pharmaceutical, and nutritional variations on plasma cholesterol concentrations. The setup of this thesis will follow the modeling cycle. The first two steps in this modeling cycle, biological observations and setting a goal, are described in **Chapter 1**.

Chapter 2 describes the definition of a conceptual model, based on the function of so-called ‘key genes’ that determine plasma cholesterol concentrations in the mouse.

In **Chapter 3**, this conceptual model was converted to a mathematical model for the mouse. This chapter also includes model calibration, implementation, validation, and analysis. In this chapter, 14 different knockout mouse strains were simulated, including the frequently used *ApoE*^{-/-} and *Ldlr*^{-/-} mouse.

Chapter 4 describes the conversion of the model for the mouse to a model for humans. This chapter also includes model calibration, implementation, validation, and analysis. Validation was done by

comparing the predicted effects of mutations in genes involved in cholesterol metabolism to literature data on plasma cholesterol concentrations.

Chapter 5 reports on the simulation of pravastatin treatment and on an analysis of the variation in response to statin therapy using the human model developed in Chapter 4.

Chapter 6 shows the simulation of treatment with torcetrapib, and with the combination with torcetrapib and pravastatin, and an analysis of the variation in response to torcetrapib therapy, all using the human model developed in the previous chapters.

Finally, results and further opportunities for the use of the model are discussed in **Chapter 7**.

24

Reference List

1. Tabas, I. (2002) *J. Clin. Invest* 110, 583-590
2. Libby, P. (2002) *Nature* 420, 868-874
3. Lusis, A. J. (2000) *Nature* 407, 233-241
4. Vaartjes, I., van Dis, I., Visseren, F. L. J., and Bots, M. L. (2010) *Hart- en vaatziekten in Nederland 2010: Cijfers over leefstijl- en risicofactoren, ziekte en sterfte*
5. Rader, D. J. and Daugherty, A. (2008) *Nature* 451, 904-913
6. Ross, R. (1999) *Am. Heart J.* 138, S419-S420
7. Glassberg, H. and Rader, D. J. (2008) *Annu. Rev. Med.* 59, 79-94
8. Libby, P., Ridker, P. M., and Hansson, G. K. (2011) *Nature* 473, 317-325
9. Psaty, B. M., Weiss, N. S., Furberg, C. D., Koepsell, T. D., Siscovick, D. S., Rosendaal, F. R., Smith, N. L. and others (1999) *JAMA* 282, 786-790
10. Quirk, J., Thornton, M., and Kirkpatrick, P. (2003) *Nat. Rev. Drug Discov.* 2, 769-770
11. Goldstein, J. L. and Brown, M. S. (2009) *Arterioscler. Thromb. Vasc. Biol.* 29, 431-438
12. Davidson, M. H., McGarry, T., Bettis, R., Melani, L., Lipka, L. J., LeBeaut, A. P., Suresh, R. and others (2002) *J. Am. Coll. Cardiol.* 40, 2125-2134
13. Kastelein, J. J., Akdim, F., Stroes, E. S., Zwinderman, A. H., Bots, M. L., Stalenhoef, A. F., Visseren, F. L. and others (2008) *N. Engl. J. Med.* 358, 1431-1443

14. Katan, M. B., Grundy, S. M., Jones, P., Law, M., Miettinen, T., and Paoletti, R. (2003) *Mayo Clin. Proc.* 78, 965-978
15. Kamanna, V. S. and Kashyap, M. L. (2000) *Curr. Atheroscler. Rep.* 2, 36-46
16. Ranalletta, M., Bierilo, K. K., Chen, Y., Milot, D., Chen, Q., Tung, E., Houde, C. and others (2010) *J. Lipid Res.* 51, 2739-2752
17. Pearson, T. A., Laurora, I., Chu, H., and Kafonek, S. (2000) *Arch. Intern. Med.* 160, 459-467
18. Engelking, L. J., Liang, G., Hammer, R. E., Takaishi, K., Kuriyama, H., Evers, B. M., Li, W. P. and others (2005) *J. Clin. Invest* 115, 2489-2498
19. Wu, X., Sakata, N., Dixon, J., and Ginsberg, H. N. (1994) *J. Lipid Res.* 35, 1200-1210
20. Ellsworth, J. L., Erickson, S. K., and Cooper, A. D. (1986) *J. Lipid Res.* 27, 858-874
21. Hamsten, A., Iselius, L., Dahlen, G., and De Faire, U. (1986) *Atherosclerosis* 60, 199-208
22. Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P. and others (2010) *Nature* 466, 707-713
23. *Shorter Oxford English Dictionary on historical principles*, (2002) Oxford University Press.
24. Lopez, C. A., de Vries, A. H., and Marrink, S. J. (2011) *PLoS. Comput. Biol.* 7, e1002020
25. Gomez-Lechon, M. J., Donato, M. T., Martinez-Romero, A., Jimenez, N., Castell, J. V., and O'Connor, J. E. (2007) *Chem. Biol. Interact.* 165, 106-116
26. Zadelaar, S., Kleemann, R., Verschuren, L., de Vries-van der Weij, J., van der Hoorn, J., Princen, H. M., and Kooistra, T. (2007) *Arterioscler. Thromb. Vasc. Biol.* 27, 1706-1721
27. Reddy, M. B., Yang, R. S. H., Clewell, H. J., and Andersen, M. E. (2005) *Physiologically Based Pharmacokinetic Modeling*, John Wiley & Sons, Inc., Hoboken, NJ, USA
28. Van Waveren, R. H., Groot, S., Scholten, H., van Geer, F. C., Wösten, J. H. M., Koeze, R. D., and Noort, J. J. (1999) *Good modeling practice handbook*. STOWA-rapport 99-05
29. Abrams, J. P. (2001) *Mathematical Modeling: Teaching the Open-ended Application of Mathematics*

30. Douma, R. D., Verheijen, P. J., de Laat, W. T., Heijnen, J. J., and van Gulik, W. M. (2010) *Biotechnol. Bioeng.* 106, 608-618
31. Punt, A., Paini, A., Boersma, M. G., Freidig, A. P., Delatour, T., Scholz, G., Schilter, B. and others (2009) *Toxicol. Sci.* 110, 255-269
32. Watanabe, T., Kusuhara, H., Maeda, K., Shitara, Y., and Sugiyama, Y. (2009) *J. Pharmacol. Exp. Ther.* 328, 652-662
33. Herrgard, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., Bluthgen, N. and others (2008) *Nat. Biotechnol.* 26, 1155-1160
34. Kuepfer, L., Peter, M., Sauer, U., and Stelling, J. (2007) *Nat. Biotechnol.* 25, 1001-1006
35. MATLAB. The MathWorks Inc. (2007)
36. Kholodenko, B. N. (2000) *Eur. J. Biochem.* 267, 1583-1588
37. Markevich, N. I., Hoek, J. B., and Kholodenko, B. N. (2004) *J. Cell Biol.* 164, 353-359
38. Nikerel, I. E., van Winden, W. A., Verheijen, P. J., and Heijnen, J. J. (2009) *Metab Eng* 11, 20-30
39. Zhang, Y., Wang, X., Vales, C., Lee, F. Y., Lee, H., Lusic, A. J., and Edwards, P. A. (2006) *Arterioscler. Thromb. Vasc. Biol.* 26, 2316-2321
40. Dietschy, J. M. and Turley, S. D. (2002) *J. Biol. Chem.* 277, 3801-3804
41. Spady, D. K., Woollett, L. A., and Dietschy, J. M. (1993) *Annu. Rev. Nutr.* 13, 355-381
42. Stellaard, F. and Kuipers, F. (2005) *Curr. Drug Targets. Immune. Endocr. Metabol. Disord.* 5, 209-218
43. Lusic, A. J. (1988) *J. Lipid Res.* 29, 397-429
44. Fielding, C. J. and Fielding, P. E. (1995) *J. Lipid Res.* 36, 211-228
45. Rader, D. J. (2001) *Nat. Med.* 7, 1282-1284
46. Shoulders, C. C., Jones, E. L., and Naoumova, R. P. (2004) *Hum. Mol. Genet.* 13 Spec No 1, R149-R160
47. Dietschy, J. M. and Turley, S. D. (2001) *Curr. Opin. Lipidol.* 12, 105-112
48. Goodman, D. S. and Noble, R. P. (1968) *J. Clin. Invest* 47, 231-241

49. De Graaf, A. A. and Van Schalkwijk, D. B. (2011) *Clinical Lipidology* 6, 25-33
50. Adiels, M., Packard, C., Caslake, M. J., Stewart, P., Soro, A., Westerbacka, J., Wennberg, B. and others (2005) *J. Lipid Res.* 46, 58-67
51. Packard, C. J., Gaw, A., Demant, T., and Shepherd, J. (1995) *J. Lipid Res.* 36, 172-187
52. van Schalkwijk, D. B., de Graaf, A. A., van Ommen, B., van Bochove, K., Rensen, P. C., Havekes, L. M., van de Pas, N. C. and others (2009) *J. Lipid Res.* 50, 2398-2411
53. Hubner, K., Schwager, T., Winkler, K., Reich, J. G., and Holzhutter, H. G. (2008) *PLoS. Comput. Biol.* 4, e1000079
54. August, E., Parker, K. H., and Barahona, M. (2007) *Bull. Math. Biol.* 69, 1233-1254
55. Tindall, M. J., Wattis, J. A., O'Malley, B. J., Pickersgill, L., and Jackson, K. G. (2009) *J. Theor. Biol.* 257, 371-384
56. Pearson, T., Wattis, J. A., O'Malley, B., Pickersgill, L., Blackburn, H., Jackson, K. G., and Byrne, H. M. (2009) *J. Math. Biol.* 58, 845-880
57. Knoblauch, H., Schuster, H., Luft, F. C., and Reich, J. (2000) *J. Mol. Med.* 78, 507-515
58. Balgi, G., Kadambi, A., and Paterson, T. S. (2006) *Methods and models for cholesterol metabolism.*



CHAPTER

2

Systematic construction of a conceptual minimal model of plasma cholesterol levels based on knockout mouse phenotypes.

Niek C.A. van de Pas, Ans E.M.F. Soffers, Andreas P. Freidig, Ben van Ommen, Ruud A. Woutersen, Ivonne M.C.M. Rietjens, and Albert A. de Graaf

The chapter is based on *Biochim Biophys Acta*. 1801:646-654, (2010)

Supplementary material can be found at www.vdpastreffers.nl/cholesterol



Abstract

Elevated plasma cholesterol, a well known risk factor for cardiovascular diseases, is the result of the activity of many genes and their encoded proteins in a complex physiological network. We aim to develop a minimal kinetic computational model for predicting plasma cholesterol levels. To define the scope of this model, it is essential to discriminate between important and less important processes influencing plasma cholesterol levels. To this end, we performed a systematic review of mouse knockout strains and used the resulting dataset, named KOMDIP, for the identification of key genes that determine plasma cholesterol levels. Based on the described phenotype of mouse knockout models 36 of the 120 evaluated genes were marked as key genes that have a pronounced effect on the plasma cholesterol concentration. The key genes include well-known genes, e.g. *ApoE* and *Ldlr*, as well as genes hardly linked to cholesterol metabolism so far, e.g. *Plagl2* and *Slc37a4*. Based on the catalytic function of the genes a minimal conceptual model was defined. A comparison with nine conceptual models from literature revealed that each of the individual published models is less complete than our model. Concluding: we have developed a conceptual model that can be used to develop a physiologically based kinetic model to quantitatively predict plasma cholesterol levels.

Introduction

Cholesterol is an important molecule in fat metabolism, a precursor for vitamin D and is involved in maintaining cellular integrity (1). Elevated plasma concentrations of LDL-cholesterol (LDL-C) have been correlated with the risk of atherosclerosis and cardiovascular diseases (2), which are leading causes of death in Western societies (3). A large part of cholesterol research uses the mouse as a model organism (4) to benefit from many practical advantages such as short generation times and reduced genetic variation (inbred strains). Additionally, several modern biological techniques are applicable in mice, but not in humans. A very powerful one is the gene knockout technology (5) which allows direct study of the effect of individual genes *in vivo*. Using knockout mouse models, the involvement of many genes in cholesterol metabolism and plasma cholesterol levels has been studied. From these studies, it has become clear that a complex system of biochemical processes is involved in the enzymatic conversions and transport of cholesterol in the body. Unfortunately, the quantitative interplay of these processes *in vivo* is not fully understood. As a consequence, it is generally difficult to predict the effect of interventions on plasma cholesterol.

Physiologically based kinetic (PBK) models can be effective tools to assess this issue. PBK models allow to predict the effect of system perturbations (e.g. genetic defects, therapies), help to gain quantitative insight, and can play a role in translational research (6-8). These computational models are generally built from conceptual models that present knowledge in an integrated fashion (9).

Several conceptual models of whole body cholesterol metabolism have been published (4,10-12). These models are graphical representations of the set of organs, metabolite pools and fluxes that together comprise the most important enzymatic conversions in cholesterol metabolism and pathways of cholesterol transport between different organs. However, the conceptual models on cholesterol metabolism defined so far (4,10,12-18) do not explicitly focus on plasma cholesterol.

Therefore, the aim of the present study was to construct a comprehensive minimal conceptual model of cholesterol metabolism in the mouse, explicitly focusing on processes that will affect plasma cholesterol. This will subsequently allow the development of a quantitative PBK model for prediction of plasma cholesterol concentrations, and of the effect of cholesterol lowering therapies in preclinical research. In order to avoid overparameterization, this model should contain as few processes

and parameters as possible (i.e. it should represent a minimal conceptual model). The choice of the processes to be included in the model was made on the basis of key genes. To discriminate key genes in determining plasma cholesterol concentrations from less important ones, we systematically reviewed the cholesterol phenotype of available knockout mouse models. These key genes were defined in two different ways: 1) genes were marked key genes type A if the cholesterol plasma level of the knockout mouse of the gene is highly affected compared to the wild type level and 2) genes that lead to lethality early in embryogenesis when knocked out (i.e. no plasma cholesterol data available for the knock out) were labeled key genes type B, if the primary function of the gene could be assumed to directly affect plasma cholesterol concentrations.

The outcomes were used to construct a minimal conceptual model for plasma cholesterol levels that was subsequently compared with conceptual models described in literature (4,10,12-18).

Methods

Data set construction

To rank the impact of genes on cholesterol plasma levels and to define key genes type A (highly affecting plasma cholesterol levels in knockout strains compared to the wild type), an inventory of all relevant knockout mouse models was made using the Mouse Genome Database (available via <http://www.informatics.jax.org>) (19). The database was searched for alleles that correspond to the phenotypic categories 1) ‘abnormal cholesterol homeostasis’ and 2) ‘abnormal bile salt homeostasis’. The second search term was included since bile salts are a degradation product of cholesterol and bile salt metabolism is closely related to cholesterol metabolism (4). These two categories are the most comprehensive phenotypic categories having a direct link with cholesterol. They comprise the following 16 different daughter categories: 1) ‘decreased cholesterol level’, 2) ‘increased cholesterol level’, 3) ‘abnormal circulating cholesterol level’, 4) ‘decreased circulating cholesterol level’, 5) ‘increased circulating cholesterol level’, 6) ‘abnormal circulating HDL cholesterol level’, 7) ‘decreased circulating HDL cholesterol level’, 8) ‘increased circulating HDL cholesterol level’, 9) ‘abnormal circulating LDL cholesterol level’, 10) ‘decreased circulating LDL cholesterol level’, 11) ‘increased circulating LDL cholesterol level’, 12) ‘abnormal circulating VLDL cholesterol level’, 13) ‘decreased circulating VLDL cholesterol level’, 14) ‘increased circulating VLDL

cholesterol level', 15) 'abnormal bile salt level' and 16) 'increased bile salt level'. The final database search was performed on August 28th 2008. To increase the homogeneity of the data set, knockin and transgenic mouse strains were not included and this also holds for mouse strains that carried alleles with non-targeted mutations. The latter included Quantitative Trait Loci, chemically induced and spontaneous mutations. For all alleles, data on the wild type strain and the reference to the literature source were obtained from the Mouse Genome Database. This information was linked to data on diet, gender, total plasma cholesterol for both wild type (wt) and knockout (ko), plasma LDL-C levels (wt + ko) and plasma HDL-C levels (wt + ko) extracted from the original publications. This resulted into a manually checked KnockOut MouseData Inventory of cholesterol Phenotype of models (KOMDIP) containing plasma cholesterol phenotype data together with data on the experimental design.

Key genes type A identified from plasma data in KOMDIP

To distinguish the most important genes affecting plasma cholesterol concentrations, a gene effect (E) was calculated for every experiment with a given knockout strain present in KOMDIP.

$$E = 2 \log \frac{[TC]_{ko}}{[TC]_{wt}} \quad (\text{Eqn. 2.1})$$

where, $[TC]_{ko}$ stands for the reported average total plasma cholesterol concentration of the knockout mouse and $[TC]_{wt}$ for the average total plasma cholesterol concentration of the wild type counterpart. E is based on a 2-log ratio as commonly used in gene expression analysis, because it produces a continuous spectrum of values for increased and decreased cholesterol concentrations as compared to wild type (20). Genes corresponding to absolute values of E larger than a cut off value are henceforth referred to as key genes type A. Since plasma cholesterol levels typically vary approximately a factor of two between background strains (21), we chose the cut off value to be 1 or -1. An E value larger than +1 means that the knockout mouse has more than twofold higher cholesterol levels than the wild type and an E value smaller than -1 means that the knockout mouse has a more than twofold lower cholesterol concentration than the wild type.

Impact of gender and diet

To judge the importance of gender and dietary effects, wild type total plasma cholesterol data (knockout data were not considered since these would focus on the genetic factors) were distributed in different groups according to the reported experimental design. Differences in average total plasma concentrations between the groups were tested for significance using the Wilcoxon rank sum test as implemented in the Statistics toolbox of MATLAB (version 7.5 R2007b). Tests for normality were performed using the Lillietest as implemented in the same toolbox. Differences with $p < 0.05$ were considered significant.

Key genes type B identified from non viable knockout mice

An obvious reason why genes might be missed in the previous analysis is that the corresponding knockout mouse is not viable. To find the genes which lead to a non viable mouse strain and are tightly linked to plasma cholesterol levels, the Mouse Genome Database was searched for targeted alleles that were grouped in the categories ‘embryonic lethality at implantation’, ‘embryonic lethality before implantation’, ‘embryonic lethality before turning of embryo’, ‘embryonic lethality before somite formation’ and ‘embryonic lethality during organogenesis’. References to the corresponding literature sources were obtained from the Mouse Genome Database. To verify whether these genes can be considered a key gene type B, Pubmed was used to search the titles and abstracts of these publications for the term ‘cholesterol’. If the term cholesterol is present and the gene was not integrated in the KOMDIP data set and the hypothetical function of the gene is directly related to (plasma) cholesterol, the gene was marked as a key player type B.

Conceptual model

The key genes of both types which resulted from the aforementioned analyses were used to construct a conceptual model including the processes which influence plasma cholesterol. The choices on which processes are relevant enough to include in the model were made based on the (suggested) functions of the key genes. Since we aim to construct a minimal conceptual model, we chose to include only those genes that can be directly appointed to a specific metabolic process in cholesterol metabolism or to a cholesterol transport process. As regulatory genes often affect different processes, they were not included. Also, genes that affect cholesterol metabolism indirectly, for example by affecting triglyceride metabolism, were not included.

A representation of the conceptual model was made by drawing boxes that represent the organs wherein the key genes that code for metabolic enzymes or transport proteins are known to be active. Within these organs, separate pools of free cholesterol and/or cholesterol esters were defined as necessary. Arrows that represent the transport processes of cholesterol or cholesterol esters from one organ to another interconnect the boxes (organs).

36 A major concern while constructing the conceptual model was to avoid so called dead-end pools i.e. pools that are connected to either only processes that decrease the pool size, or to only processes that increase the pool size. This would mean that in the corresponding computational model, the pool size would decrease to zero or increase to infinity, which does not occur in real life (4). Therefore, in cases where such pools occurred in the model resulting from the procedure described so far, the model was extended by including the missing processes.

RESULTS

KOMDIP data set construction

The manually checked data set KOMDIP containing information on genes, alleles, background strains, diet, sex and plasma cholesterol concentrations is supplied as supplementary information. The data set is based on 168 publications and contains information on 122 different targeted genes, and includes over 600 total cholesterol concentration data entries. In total 311 pairs of total cholesterol concentrations of the knockout and corresponding wild type are present. As not all publications consistently list standard deviations and group sizes, only reported average values were used for further analysis.

Impact of gender and diet

The results of the quantification of the effects of gender and diet on cholesterol concentrations are given in Table 2.1.

Table 2.1. Total plasma cholesterol concentrations of the wild type counterparts of knockout mouse models considered in the KOMDIP data set. The experiments were sorted in diet and gender groups according to their experimental design. Data are expressed as average \pm standard deviation, of the concentration values reported in different original publications. The chow diet group includes data obtained from publications indicating diets as chow, standard chow, standard lab diet, and data from studies that do not indicate the diet. The non-chow group contains data on different diets including, high fat diet, western type diet and fasting.

| Diet | Gender | Number data points | Total cholesterol concentration (mM) |
|-----------------------|---------------------|--------------------|--------------------------------------|
| All | Both ^e | 313 | 2.94 \pm 1.45 |
| | Male | 113 | 3.19 \pm 1.56 |
| | Female ^b | 88 | 2.63 \pm 1.19 |
| Chow | Both ^a | 194 | 2.40 \pm 0.88 |
| | Male | 70 | 2.57 \pm 0.81 |
| | Female ^b | 59 | 2.18 \pm 0.72 |
| Non-chow ^d | Both ^c | 119 | 3.82 \pm 1.75 |
| | Male | 43 | 4.22 \pm 1.92 |
| | Female ^d | 29 | 3.56 \pm 1.41 |

^a $P = 6.5 \cdot 10^{-17}$ versus chow, ^b $P = 0.0013$ versus male, ^c $P = 0.0045$ versus male, ^d $P = 0.1157$ versus male, ^e Sum of the number of data points in the males and females groups are smaller than the number of data points in the groups with both males and females, since some publications do not mention gender or mention mixed gender groups.

Differences between the averages of the chow and non chow groups are larger than 50% of the average cholesterol concentration in the chow group, while differences between the averages of the male and female groups are smaller than 25% of the average cholesterol concentration in males. This indicates that male mice have higher total plasma cholesterol concentrations than females, but also that this gender effect is smaller than the diet effect (Table 2.1).

Key genes type A identified from plasma data in KOMDIP

In order to discriminate the processes associated with key genes from the others, key genes were identified using the KOMDIP data set. In the KOMDIP data set, 311 pairs of total cholesterol concentrations of the knockout and corresponding wild type are present. This is slightly less than

the total of wild type entries (Table 2.1), since some publications only give data for wild type and heterozygote and not for the full knockout. The 311 calculated E values (Eqn 2.1) are presented in Figure 2.1. It can be seen that the E values are distributed in bell shape (average is 0.2, standard deviation is 1.3), although the distribution did not pass the test for normality ($p < 0.001$).

Gene effects vary from -5.8 to 6.0, while 71% of all E values are between -1 and 1, which shows that many of the genes only have a minor effect compared to the high impact genes (Figure 2.1). Figure 2.2 shows results for the key genes identified.

38

Data corresponding to 31 genes were observed to lead to an E value higher than 1 or lower than -1. This set includes several genes that are directly involved in lipoprotein metabolism, like apolipoproteins (*Apoa1*, *ApoE*), lipoprotein receptors (*Ldlr* and *Scarb1*), as well as some genes that have hardly been linked to cholesterol metabolism so far like *G6pc* and *Slc37a3*.

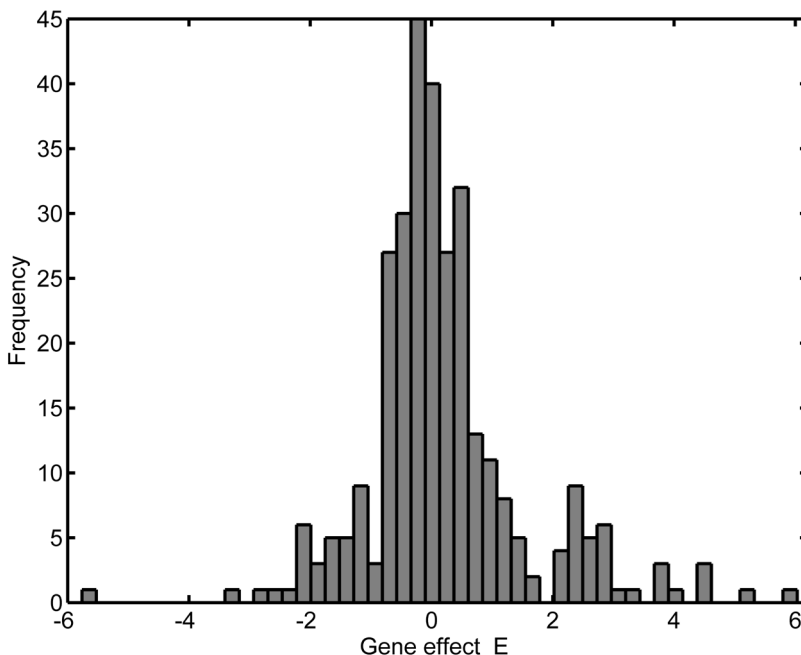


Figure 2.1. Histogram of the frequency distribution of calculated gene effects (E , see Eqn. 2.1) for the 311 experiments (pairs of total cholesterol levels of the knockout and corresponding wild type) present in the KOMDIP data set ($n = 311$).

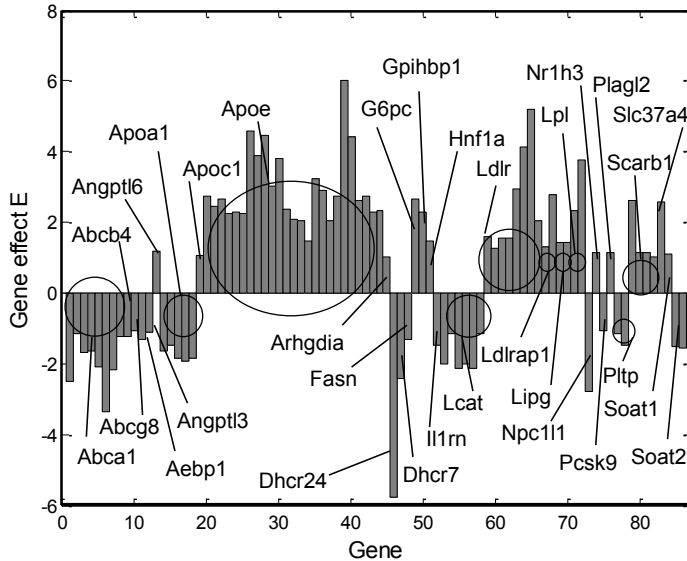


Figure 2.2. Calculated gene effects (E) that are larger than the cut off value ($E < -1$ and $E > 1$) for mouse knockout experiments. The effects are presented in alphabetical order of the corresponding gene names. Multiple bars can correspond to a single gene (e.g. *Apoe*). These bars refer to separate experiments, for example differing in diet, gender or age of the mice used.

Recently, Su et al. (22) have demonstrated that the observed effect in the *Soat1* knockout mouse (23) was due to differences between different background strains used to construct the *Soat1*^{-/-} mouse and not to the impact of *Soat1* itself, consequently *Soat1* is not treated as key player type A.

Key genes type B identified from non viable knockout mice

Since the analysis of the KOMDIP data base described above would not find key genes that would lead to non viable knock-out mouse strains, an additional analysis of non viable knockout mouse strains was performed. The search of the Mouse Genome Database for non viable knockout mice yielded 3246 publications containing information on 4626 different genotypes. In total, 46 of those publications have the term ‘cholesterol’ in their title and/or abstract, indicating that the gene might be a key gene in determining plasma cholesterol concentrations. Again, to ensure specificity of our analysis, only the 28 genes leading to non viable mice carrying a single, non transgenic targeted mutation as given in Table 2.2 were included in further analysis.

Table 2.2. The genes corresponding to a non viable knockout mouse strain that are described in publications that include “cholesterol” in their abstract. The second column indicates whether the gene is present in the KOMDIP data set and the third indicates whether it was marked as a key player of type A or B. The numbers in the last column correspond with the numbers of a motivation presented in the Results section Y, Yes; N, No; A, key player type A; B, key player type B.

| Gene name | In KOMDIP? | Key player? | Motivation |
|----------------|------------|-------------|------------|
| <i>Abca1</i> | Y | A | 1 |
| <i>Abca3</i> | N | N | 3 |
| <i>Apob</i> | Y | B | 1, 2 |
| <i>Atxn2</i> | Y | N | 1 |
| <i>Cyp11a1</i> | N | N | 4 |
| <i>Cyp7a1</i> | Y | N | 1 |
| <i>Dhcr24</i> | Y | A | 1 |
| <i>Dhcr7</i> | Y | A | 1 |
| <i>Disp1</i> | N | N | 5 |
| <i>Fasn</i> | Y | A | 1 |
| <i>Fdft1</i> | N | B | 6 |
| <i>Hmgcr</i> | N | B | 6 |
| <i>Hnf4a</i> | N | B | 7 |
| <i>Hsd17b7</i> | N | B | 6 |
| <i>Lpl</i> | Y | A | 1 |
| <i>Lrp2</i> | N | N | 1 |
| <i>Mttp</i> | N | B | 8 |
| <i>Nr5a2</i> | Y | N | 1 |
| <i>Por</i> | Y | N | 1 |
| <i>Psen1</i> | N | N | 9 |
| <i>Sc5d</i> | Y | N | 1 |
| <i>Scarb1</i> | Y | A | 1 |
| <i>Shh</i> | N | N | 5 |
| <i>Sptlc1</i> | N | N | 10 |
| <i>Sptlc2</i> | N | N | 10 |
| <i>Srebf1</i> | Y | N | 1 |
| <i>Star</i> | N | N | 4 |
| <i>Vcam1</i> | N | N | 10 |

The motivation to mark a gene as key player type B is based on arguments specific to each case or small set of cases. The sections below presenting those motivations are numbered and the last column of Table 2.2

refers to these numbers for the various genes listed.

1. Out of the 28 genes in Table 2.2, 13 were already in the KOMDIP data set (see 2nd column) and were, therefore, already included in the key gene type A analysis. Data on the effect of these genes have often been obtained from plasma analysis of newborn mice. In total, 6 out of these 13 genes were already identified as a key player type A (Figures 2.2 and 2.3), namely *Abca1*, *Fasn*, *Dhcr24*, *Dhcr7*, *Lpl*, and *Scarb1*.

2. The gene *Apob* is present in the KOMDIP data set and was not marked as key gene type A. The *Apob* gene codes for 2 different proteins, which are splice variants. The *Apob* knockout models in the KOMDIP data set are specifically modified to retain one of the functional Apob proteins. When both splice variants are disabled, the mouse is not viable and the plasma cholesterol concentration of the heterozygote is highly affected (24). Since Apob is the main structural protein in LDL, chylomicrons and VLDL (25), we consider *Apob* a key gene (type B) of cholesterol metabolism.

3. Both *Lrp2* and *Abca3* knockout mice died shortly after birth due to respiration complications (26,27). It has been claimed that this might be due to a defect of lipids in the alveolar fluid. Since no link with cholesterol plasma concentrations is given and plasma cholesterol concentrations are not reported in both articles, the genes were assumed not to be a type B key player.

4. Two of the genes in Table 2.2 play a role in the conversion of cholesterol to steroid hormones (*Cyp11a1* and *Star*). Since both mouse strains can only stay alive when injected with steroids, these strains are assumed to be non viable due to defects in the hormone system rather than in cholesterol metabolism (28,29). The genes were, therefore, not marked as a key player type B.

5. Two genes in Table 2.2 include genes that are involved in hedgehog signaling (*Shh* and *Disp1*), an important mechanism in embryonic development. Cholesterol is required for Shh processing. There is, however, very little evidence that the hedgehog pathway affects plasma cholesterol concentrations in the adult mouse (30,31). The genes *Shh* and *Disp1* were therefore not marked as key player type B.

6. Three genes are enzymes in cholesterol biosynthesis: *Hmgcr* (32), *Fdft1* (33), and *Hsd17b7* (34). Since the principal function of these genes is related to cholesterol, it can be assumed that the observed lethality is indeed due to defects in cholesterol metabolism. The genes were therefore labeled as key player type B.

7. *Hnf4a* is central to the maintenance of hepatocyte differentiation and is a major *in vivo* regulator of genes involved in the control of lipid homeostasis (35). It is pointed out (35) that the role of *Hnf4a* in lipid homeostasis includes a role in regulating plasma cholesterol concentrations. *Hnf4a* was therefore marked as a key player type B.

8. *Mttp* is a gene involved in the secretion of lipoproteins. Mutations in the *Mttp* gene are the cause of the disease abetalipoproteinemia, an inherited disease of extremely low plasma concentrations of cholesterol (36). Therefore *Mttp* was marked as key player type B.

9. Cholesterol is mentioned in the article on *Psen1* (37) as an analogy only. No argument for the affection of cholesterol metabolism is given. Consequently, this gene was not marked as key player type B.

10. In the article about *Sptlc1* and *Sptlc2* (38), plasma cholesterol is mentioned as a control parameter in which no change was detected. This is also the case for the publication on *Vcam1* (39). These genes were therefore not labeled as key player.

In summary, 6 additional genes were marked as key player (type B): *Apob*, *Mttp*, *Hnf4a*, *Hmgcr*, *Fdft1*, and *Hsd17b7*. Since 30 key genes type A were defined by analysis of the KOMDIP data set, in total 36 different genes were marked as key genes.

Functions of the key genes.

Table 2.3 presents an overview of the biological functions associated with the key genes thus identified. The key processes include several transport proteins, like (*Scarb1* and *Abca1*), apolipoproteins (*Apoa1*, *Apoc1* and *Apob*), several metabolic enzymes (*Fasn*, *Hmgcr*) as well as regulation factors (*Nr1h3*, *Arhgdia*). More details on the key gene functions are given in the discussion section.

Table 2.3: Short summary of the function of the key genes. TF, Transcription factor; G6P glucose-6-phosphate, CM chylomicron, TG triglyceride, PL phospholipid.

| Nr | Gene name | Function of gene product | Ref |
|----|----------------|-------------------------------------------------------------------------|------|
| 1 | <i>Abca1</i> | Cholesterol transport into HDL | (40) |
| 2 | <i>Abcb4</i> | Biliary phospholipid transport | (41) |
| 3 | <i>Abcg8</i> | Biliary cholesterol transport | (42) |
| 4 | <i>Aebp1</i> | Regulator of adipose tissue homeostasis | (43) |
| 5 | <i>Angptl3</i> | Inhibition of Lpl activity | (44) |
| 6 | <i>Angptl6</i> | Circulating peptide induces angiogenesis | (45) |
| 7 | <i>Apoa1</i> | HDL building block | (46) |
| 8 | <i>Apoc1</i> | Regulation of lipoprotein metabolism | (47) |
| 9 | <i>ApoE</i> | Regulation of LDL and CM uptake | (48) |
| 10 | <i>Arhgdia</i> | Regulation of cytoskeleton dependent functions | (49) |
| 11 | <i>Dhcr24</i> | Enzyme in cholesterol synthesis | (50) |
| 12 | <i>Dhcr7</i> | Enzyme in cholesterol synthesis | (51) |
| 13 | <i>Fasn</i> | Enzyme in the synthesis of fatty acids | (52) |
| 14 | <i>G6pc</i> | Catalyzing conversion of G6P into glucose and phosphate | (53) |
| 15 | <i>Gpihbp1</i> | Lipolytic processing of CM | (54) |
| 16 | <i>Hnf1a</i> | TF controlling genes expressed preferentially in liver | (55) |
| 17 | <i>Il1rn</i> | Agonist of the interleukin 1 receptor | (56) |
| 18 | <i>Lcat</i> | Catalyzing HDL associated cholesteryl esterification | (57) |
| 19 | <i>Ldlr</i> | Uptake of Apob containing lipoproteins | (58) |
| 20 | <i>Ldlrap1</i> | Increasing catabolism of LDL | (58) |
| 21 | <i>Lipg</i> | Involved in lipoprotein metabolism, function unknown | (59) |
| 22 | <i>Lpl</i> | Catalyses the hydrolysis of TG in VLDL and CM | (60) |
| 23 | <i>Npc1l1</i> | Intestinal cholesterol absorption | (61) |
| 24 | <i>Nr1h3</i> | Oxysterol sensor that regulates several genes in cholesterol metabolism | (62) |
| 25 | <i>Pcsk9</i> | Targets Ldlr for degradation | (63) |
| 26 | <i>Plagl2</i> | TF in adipocyte metabolism | (64) |
| 27 | <i>Pltp</i> | Transfer of PL between different plasma lipoproteins | (65) |
| 28 | <i>Scarb1</i> | Transport for the hepatic uptake of HDL C and CE | (66) |
| 29 | <i>Slc37a4</i> | Transport of G6P | (67) |
| 30 | <i>Soat2</i> | Cellular cholesterol esterification enzyme | (68) |
| 31 | <i>Apob</i> | Building block of (V)LDL and CM | (25) |
| 32 | <i>Fdft1</i> | Enzyme in cholesterol synthesis | (33) |
| 33 | <i>Hmgcr</i> | Enzyme in cholesterol synthesis | (32) |
| 34 | <i>Hnf4a</i> | TF in regulation of lipid metabolism | (35) |
| 35 | <i>Hsd17b7</i> | Enzyme in cholesterol synthesis | (34) |
| 36 | <i>Mttp</i> | Loads Apob with cholesterol and TG to form VLDL and CM | (36) |

Conceptual model

Based on the key genes, a conceptual model for describing cholesterol plasma concentrations with a minimum amount of essential parameters was defined to serve as a basis for a future kinetic model. Therefore, we chose to include only those genes in our model that can be directly appointed to a specific metabolic or transport process in the cholesterol metabolism. These are the genes: *Dhcr7*, *Dhcr24*, *Fdft1*, *Hmgcr*, *Hsd17b7*, *Ldlr*, *Scarb1*, *Abca1*, *Mttp*, *Soat2*, *Abcg8*, and *Npc1l1*. To make the model, we first review the functions and locations of these key genes.

44

The gene *Npc1l1* is only expressed in the intestine (69) and plays a role in the uptake of dietary and biliary cholesterol. This biliary cholesterol is transported into bile by the protein that is coded for by gene *Abcg8* that is expressed in the liver and the intestine (69). Cholesterol biosynthesis (genes: *Dhcr7*, *Dhcr24*, *Hsd17b7*, *Fdft1*, and *Hmgcr*) occurs in every cell in the body (70). The key gene for esterification of this cholesterol (*Soat2*) is only expressed in the liver and the intestine (69). The esterified cholesterol is packed into Apob containing lipoproteins by the product of gene *Mttp*. The liver and the intestine are the tissues that primarily express this gene (69). Cholesterol in Apob containing lipoproteins is taken up (gene *Ldlr*) by the liver and peripheral tissues, while Apoal containing lipoproteins are primarily taken up (gene: *Scarb1*) by the liver (70). The cholesterol in these particles originates from hepatic, intestinal and peripheral sources (related gene: *Abca1*) (71). The enzyme *Lcat* associated with HDL in the plasma esterifies this cholesterol.

Currently there is an ongoing debate on the origin of HDL cholesterol (HDL-C). The classical picture is that HDL-C originates from peripheral tissues (4). Contrasting evidence from organ specific knockout mouse strains suggests however that HDL-C mainly originates from the liver and the intestine (71). Data in the mouse gene atlas (69) on the location of gene expression of *Abca1* strongly supports the hepatic and peripheral origin of HDL-C. To be in agreement with all literature, we have included all three sources in our model. Based on the locations of these genes, plasma, intestine, and liver were included as important compartments in our model. The other tissues were lumped into the peripheral compartment. Although the key genes were defined on total plasma cholesterol data, we have found genes that are related to either LDL cholesterol (i.e. *Ldlr*) or HDL cholesterol (i.e. *Lcat*) and therefore, the plasma compartment was split up into two compartments (HDL and LDL). The model based on these biological processes is represented with solid arrows in Figure 2.3.

Pools that are only connected to processes that increase the pool

size and not to processes that decrease the pool size are undesired in a good kinetic model (see Methods for motivation). In the conceptual model defined so far (Figure 2.3, solid arrows), the total body cholesterol pool is increased by dietary absorption and *de novo* synthesis while a reaction that removes cholesterol from the body is missing.

In mammals, cholesterol is lost due to various processes including a) fecal excretion, b) bile acid formation from cholesterol in the liver and c) skin shredding (4). These processes were not identified in our analysis, since 1) the final removal of cholesterol from the intestine via the feces is not catalyzed by a single transport protein, but is rather dependent on peristaltic movement, a multifactorial process i.e. no specific gene can be linked to this process, 2) several alternative pathways for bile acid formation (76) exist and 3) skin shredding is not catalyzed by a single transport protein.

To allow for a realistic description of the total body pool balance we therefore included three different reactions: fecal cholesterol excretion, hepatic cholesterol loss, and loss of cholesterol from the peripheral compartment. These additional processes are indicated with dashed arrows in Figure 2.3. The full model including this extension is given by the total of solid and dashed arrows in Figure 2.3.

DISCUSSION

In this work, we have collected data that allowed identification of the most important processes determining total plasma cholesterol concentrations. We have chosen to assemble data on knockout mouse models, since these mouse models, unlike transgenic mutations or (human) SNPs, carry a defined biological perturbation, allowing a better comparison between different genes. The constructed data set KOMDIP is a manually checked integration of quantitative mouse cholesterol phenotype data and as such provides information on genes likely to be related to cardiovascular diseases. Since not for every mouse gene a knockout phenotype is available, the screening of the genes cannot be fully comprehensive at the present stage but new information can be added as additional knockout strains become available. Based on this new database, ranking of the genes and associated processes having a large effect on plasma cholesterol concentrations became feasible.

Analysis of the KOMDIP data set assigned 30 genes with the label “key player type A” (Figure 2.2) including some genes well-known to be involved in modulating cholesterol plasma concentrations (like *ApoE* and *Ldlr*, see (48,58)). Surprisingly the data set of key player type A genes

also included some genes not so prominently linked to cholesterol plasma concentrations (like *Gpihbp1* (54) and *Plagl2* (64)). The gene *Soat1* was excluded from this list, since the cholesterol phenotype of the knockout mouse was due to experimental artifacts (22). Currently, there is no evidence that this also holds for other genes identified in this analysis to be of importance for regulating cholesterol plasma levels.

46 Since plasma cholesterol in the mouse is mainly localized in HDL, it can be expected that the identified genes play a role in HDL metabolism. While our analysis indeed identified genes related to HDL as key player (like *Apoa1*, *Lipg*, *Abca1*, *Scarb1*, *Lcat*), several genes that are considered to be related to VLDL and LDL metabolism (like *Apoe*, *Ldlr*, *Gpihbp1*, *Lpl*) were also marked as key player type A. Therefore, our analysis indicates that lipoprotein metabolism in its entirety must be considered for a correct understanding of the factors that influence total plasma cholesterol concentrations.

Since plasma cholesterol in the mouse is mainly localized in HDL, it can be expected that the identified genes play a role in HDL metabolism. While our analysis indeed identified genes related to HDL as key player (like *Apoa1*, *Lipg*, *Abca1*, *Scarb1*, *Lcat*), several genes that are considered to be related to LDL metabolism (like *Apoe*, *Ldlr*, *Gpihbp1*, *Lpl*) were also marked as key player type A. Therefore, our analysis indicates that lipoprotein metabolism in its entirety must be considered for a correct understanding of the factors that influence total plasma cholesterol levels.

Nonviability of knockout mice can obviously lead to mice not being included in KOMDIP and hence to missing potential high impact genes. After case-specific analysis of the non viable knockout models present in the Mouse Genome Database, six additional genes were marked as key player. Three of them are enzymes in the cholesterol biosynthesis pathway, including *Hmgcr*, the target of statins. The cholesterol biosynthesis pathway comprises over a dozen enzymes (72), seven of the genes coding for these enzymes have been a target of a knockout strategy and five of these genes are marked as key player. They are *Hmgcr* (type B), *Dhcr7* (type A), *Dhcr24* (type A), *Fdft1* (type B), and *Hsd17b7* (type B). Surprisingly, the knockout mice for the two other cholesterol biosynthesis genes, *Tm7sf2* and *Sc5d*, did not display severely altered plasma cholesterol concentrations. Plasma levels of the *Tm7sf2*^{-/-} mouse were normal (73) and the *Sc5d*^{-/-} mouse, that died shortly after birth, displayed only a relatively mild change in cholesterol

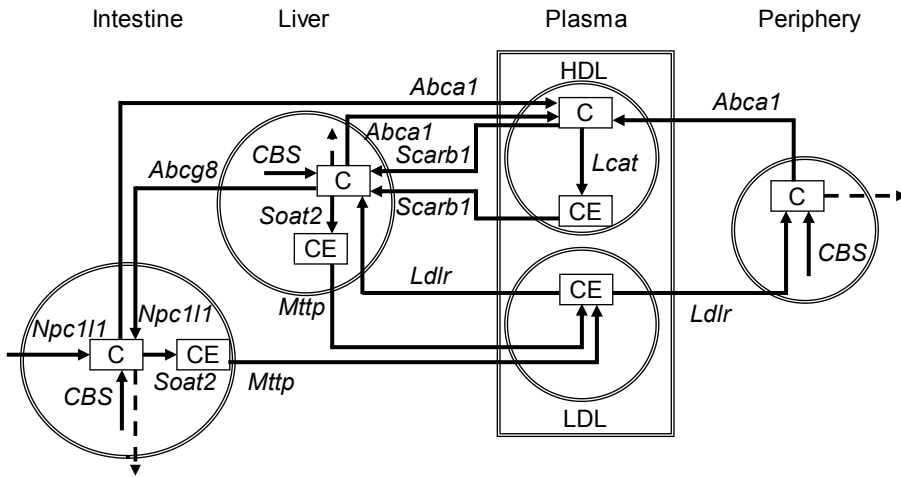


Figure 2.3. Conceptual minimal model of cholesterol metabolism developed in this study. Circles and boxes indicate compartments (such as organs), solid arrows indicate biological processes directly inferred from key genes identified in our analysis. The dashed arrows indicate the processes that were included to balance the total body cholesterol pool (see Methods). Included processes are hepatic cholesterol synthesis (CBS); peripheral cholesterol synthesis (CBS); intestinal cholesterol synthesis (CBS); biliary cholesterol excretion (*Abcg8*, *Npc111*); hepatic uptake of cholesterol from LDL (*Ldlr*); peripheral uptake of cholesterol from LDL (*Ldlr*); peripheral cholesterol transport to HDL (*Abca1*); hepatic cholesterol transport to HDL (*Abca1*); intestinal cholesterol transport to HDL (*Abca1*); Intestinal cholesterol esterification, (*Soat2*); Hepatic cholesterol esterification, (*Soat2*); HDL associated cholesterol esterification (*Lcat*); hepatic HDL cholesterol uptake (*Scarb1*, both free cholesterol and cholesterol ester); Intestinal chylomicron production (*Mttp*); Hepatic VLDL production (*Mttp*); dietary cholesterol intake (*Npc111*) and fecal cholesterol excretion. C stands for Cholesterol; CE for Cholesteryl ester. CBS includes *Dhcr7*, *Dhcr24*, *Hmgcr*, *Hsd17b7*, and *Fdft1*.

plasma concentration ($E = -0.69$ for *Sc5d*^{-/-} compared to $E = -5.76$ for the *Dhcr24*^{-/-} mouse) (74). *Tm7sf2*^{-/-} may not be essential in cholesterol synthesis (73) and *Sc5d* may be affecting cholesterol levels more severely later in life.

As a step in building a conceptual model, the 36 (type A and B) key genes were associated with several biological functions (Table 2.3). These genes include 11 genes that code for proteins that catalyze an enzymatic conversion, 4 genes that code for apolipoproteins, 9 genes

coding for proteins that are involved in transport reactions and 12 genes that code for regulatory factors. These groups are discussed below.

The enzymatic conversion group includes 11 members, some of them playing a role in the biosynthesis of cholesterol (*Hmgcr*, *Fdft1*, *Hsd17b7*, *Dhcr7*, and *Dhcr24*) or in the synthesis of fatty acids (*Fasn*). The others are involved in the esterification of cholesterol (*Soat2* and *Lcat*), the hydrolysis of triglycerides (*Lipg*, *Lpl*) or the conversion of glucose-6 phosphate (*G6pc*).

48 The 4 key genes in the apolipoprotein class (*Apoa1*, *Apoc1*, *Apob*, and *Apoe*) are building blocks of lipoproteins and play a role in their assembly, remodeling and uptake.

The 9 transport proteins are known to mediate the transport of several compounds including cholesterol (*Abcal*, *Abcg8*, *Npc111*), cholesterol ester (*Scarb1*) and lipid droplets (*Mttp*), lipoproteins (*Ldlr*), but also phospholipids (*Abcb4*, *Pltp*) and glucose-6-phosphate (*Slc37a4*). The transport genes not only code for proteins that transport cholesterol to and from the plasma (like *Abcal* and *Scarb1*), but also via the bile to the intestine or between the intestinal lumen and the enterocytes (*Npc111*, *Abcg8*). The *Abcg8* gene codes for a protein that functions in a heterodimer together with the gene *Abcg5*. Interestingly, *Abcg8* was marked as key player (E value is -1.1), whereas *Abcg5* was not (E value is -0.8).

Finally 12 key genes are classified as regulatory genes, clearly illustrating that plasma cholesterol levels are highly regulated. They include several transcription factors (*Aebp1*, *Nr1h3*, *Hnf1a*, *Hnf4a*, and *Plagl2*) as well as receptor associated genes (*Arhgdia*, *Angptl3*, *Angptl6*, *Gpihbp1*, *Pcsk9*, *Ldlrap1*, *Ilrn1*). The list also contains several genes involved in the regulation of Lpl activity (*Gpihbp1*, *Angptl3*).

In this analysis two genes involved in glucose metabolism were identified as key genes (*G6pc*, *Slc37a4*). This indicates that cholesterol metabolism is closely linked to glucose metabolism, probably via acetyl-CoA, an end product of glycolysis and the building block for cholesterol (75). The link with triglyceride and fatty acid metabolism (for example: genes *Lpl*, and *Fasn*) is not surprising, since cholesterol functions in triglyceride transport in cholesterol containing lipoproteins (17). A very interesting gene is the interleukin 1 receptor agonist (*Ilrn1*) clearly showing a direct influence of inflammation on cholesterol metabolism.

The observation that the key genes are indeed important genes for determining total plasma cholesterol concentrations is corroborated by the observation that our list includes nine genes that are targets of cholesterol

modifying drugs currently in development (*Mttp*, *Fdft1*, *Apob*, *Pcsk9*, *Lpl*, *Nr1h3*, *Apoa1*, *Lcat*, and *Lipg*) (16), as well as two genes that are targets of drugs already on the market (*Hmgcr*; *Npc1l1*) (32,52). This indicates that the conceptual model defined in the present paper provides a promising starting point for future quantitative physiologically based kinetic (PBK) modeling of plasma cholesterol levels and the effects of intervention therapies. In order to ensure that the model is a minimal model, only the 12 genes, (type A + B) that can be directly appointed to a specific metabolic or transport process in cholesterol metabolism were used to construct a conceptual model.

To assess the scope of the model, it was compared with previously published models. For this purpose, nine other models were evaluated as to which organs and biological processes are included. Table 2.4 summarizes these findings. Although our model contains largely the same processes as the other models, each of these individual models misses several processes that we have defined as being important. Another key aspect why our model is different is that it focuses on steps determining plasma cholesterol levels, whereas most previously described models aimed at providing an overview of steps that may be involved in the overall metabolic fate of cholesterol.

Our model has as features that are absent in some of the other models: 1) the liver and intestine as source of HDL-C, reflecting the outcome of recent experiments (71), 2) an explicit plasma compartment, reflecting our focus on plasma cholesterol; 3) the absence of macrophages in the model (Although macrophage specific genes like *Abcg1* were included in our analysis, they were not identified as a key player, ($E = 0.15$ and -0.17 for males and females respectively). This cell type is involved in the development of atherosclerosis, but contributes little to plasma cholesterol levels (70); and 4) a peripheral compartment, an important contributor to total body cholesterol synthesis (70).

In this paper we have developed a method to systematically identify the key genes and associated processes that determine plasma cholesterol levels in the mouse. We are confident that our conceptual model includes a good selection of relevant biological processes that determine plasma cholesterol levels. Concluding, we have successfully constructed a conceptual minimal model that can be used as a basis for the development of a quantitative kinetic model for the prediction of total plasma cholesterol levels and the effect of cholesterol lowering therapies in preclinical research.

Table 2.4: Comparison of the conceptual model developed in this paper (Figure 2.3) with previously published conceptual models. For each model it is indicated whether or not it explicitly includes the organs and processes listed in the first column. C, Cholesterol; M, Mouse; G, general; H, human; Y, Yes; N, No.

| Reference | Conceptual Model | | | | | | | | | | |
|--------------------------------------|------------------|-----|------|------|------|------|------|------|------|------|--|
| | Our model | | | | | | | | | | |
| | Figure 2.3 | (4) | (12) | (13) | (14) | (10) | (18) | (17) | (15) | (16) | |
| Organism | M | M | G | H | H | H | G | H | H | H | |
| Processes | | | | | | | | | | | |
| Hepatic C synthesis | Y | Y | Y | Y | N | Y | N | N | N | N | |
| Peripheral C synthesis | Y | Y | Y | N | N | N | N | N | Y | N | |
| Intestinal C synthesis | Y | N | Y | N | N | N | N | N | N | N | |
| Dietary C intake | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | |
| Hepatic uptake of C from LDL | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | |
| Hepatic VLDL -C secretion | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | |
| Peripheral uptake of C from LDL | Y | Y | N | Y | Y | Y | N | Y | Y | Y | |
| Transport of C from periphery to HDL | Y | Y | N | Y | Y | Y | N | Y | Y | Y | |
| HDL associated C esterification | Y | Y | N | N | N | Y | N | Y | Y | Y | |
| Hepatic HDL-CE uptake | Y | Y | N | N | N | Y | N | Y | Y | Y | |
| Intestinal CM-c secretion | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | |
| Intestinal C esterification | Y | N | N | N | N | Y | N | N | Y | N | |
| Hepatic HDL-FC uptake | Y | N | N | Y | N | N | N | Y | N | Y | |
| Biliary C excretion | Y | Y | N | Y | Y | Y | Y | Y | Y | N | |
| Fecal C excretion | Y | Y | Y | Y | N | Y | N | N | Y | N | |
| Transport of C from intestine to HDL | Y | N | N | N | Y | N | N | Y | N | Y | |
| Transport of C from liver to HDL | Y | N | N | N | N | N | N | Y | N | Y | |
| Hepatic cholesterol esterification | Y | Y | Y | N | N | Y | N | N | N | N | |
| CETP activity | N | N | N | N | Y | Y | N | Y | Y | Y | |
| Peripheral C esterification | N | Y | N | N | N | Y | N | N | N | N | |
| Compartments | | | | | | | | | | | |
| Liver | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | |
| Intestine | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | |
| Plasma | Y | N | N | Y | N | N | Y | N | N | Y | |
| Periphery | Y | Y | Y | N | Y | Y | N | Y | Y | Y | |
| Macrophage | N | N | N | Y | N | N | N | Y | N | Y | |

References

1. Tabas, I. (2002) *J. Clin. Invest* 110, 583-590
2. Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H., and Kannel, W. B. (1998) *Circulation* 97, 1837-1847
3. Lusis, A. J. (2000) *Nature* 407, 233-241
4. Dietschy, J. M. and Turley, S. D. (2002) *J. Biol. Chem.* 277, 3801-3804
5. Manis, J. P. (2007) *N. Engl. J. Med.* 357, 2426-2429
6. Kumar, N., Hendriks, B. S., Janes, K. A., de, G. D., and Lauffenburger, D. A. (2006) *Drug Discov. Today* 11, 806-811
7. Agoram, B. M., Martin, S. W., and van der Graaf, P. H. (2007) *Drug Discov. Today* 12, 1018-1024
8. Materi, W. and Wishart, D. S. (2007) *Drug Discov. Today* 12, 295-303
9. van Waveren, R. H., Groot, S., Scholten, H., van Geer, F. C., Wösten, J. H. M., Koeze, R. D., and Noort, J. J. (1999) *Good Modelling Practice Handbook*.
10. Fielding, C. J. and Fielding, P. E. (1995) *J. Lipid Res.* 36, 211-228
11. Lusis, A. J. and Pajukanta, P. (2008) *Nat. Genet.* 40, 129-130
12. Spady, D. K., Woollett, L. A., and Dietschy, J. M. (1993) *Annu. Rev. Nutr.* 13, 355-381
13. Stellaard, F. and Kuipers, F. (2005) *Curr. Drug Targets. Immune. Endocr. Metabol. Disord.* 5, 209-218
14. Lusis, A. J. (1988) *J. Lipid Res.* 29, 397-429
15. Dietschy, J. M. and Turley, S. D. (2001) *Curr. Opin. Lipidol.* 12, 105-112
16. Rader, D. J. and Daugherty, A. (2008) *Nature* 451, 904-913
17. Shoulders, C. C., Jones, E. L., and Naoumova, R. P. (2004) *Hum. Mol. Genet.* 13 Spec No 1, R149-R160
18. Rader, D. J. (2001) *Nat. Med.* 7, 1282-1284
19. Eppig, J. T., Blake, J. A., Bult, C. J., Kadin, J. A., and Richardson, J. E. (2007) *Nucleic Acids Res.* 35, D630-D637

20. Quackenbush, J. (2002) *Nat. Genet.* 32 Suppl, 496-501
21. Svenson, K. L., Von Smith, R., Magnani, P. A., Suetin, H. R., Paigen, B., Naggert, J. K., Li, R. and others (2007) *J. Appl. Physiol* 102, 2369-2378
22. Su, Z., Wang, X., Tsaih, S. W., Zhang, A., Cox, A., Sheehan, S., and Paigen, B. (2009) *J. Lipid Res.* 50, 116-125
23. Meiner, V. L., Cases, S., Myers, H. M., Sande, E. R., Bellosta, S., Schambelan, M., Pitas, R. E. and others (1996) *Proc. Natl. Acad. Sci. U. S. A* 93, 14041-14046
24. Farese, R. V., Jr., Ruland, S. L., Flynn, L. M., Stokowski, R. P., and Young, S. G. (1995) *Proc. Natl. Acad. Sci. U. S. A* 92, 1774-1778
25. Farese, R. V., Jr., Veniant, M. M., Cham, C. M., Flynn, L. M., Pierotti, V., Loring, J. F., Traber, M. and others (1996) *Proc. Natl. Acad. Sci. U. S. A* 93, 6393-6398
26. Willnow, T. E., Hilpert, J., Armstrong, S. A., Rohlmann, A., Hammer, R. E., Burns, D. K., and Herz, J. (1996) *Proc. Natl. Acad. Sci. U. S. A* 93, 8460-8464
27. Fitzgerald, M. L., Xavier, R., Haley, K. J., Welti, R., Goss, J. L., Brown, C. E., Zhuang, D. Z. and others (2007) *J. Lipid Res.* 48, 621-632
28. Ishii, T., Hasegawa, T., Pai, C. I., Yvigi-Ohana, N., Timberg, R., Zhao, L., Majdic, G. and others (2002) *Mol. Endocrinol.* 16, 2297-2309
29. Caron, K. M., Soo, S. C., Wetsel, W. C., Stocco, D. M., Clark, B. J., and Parker, K. L. (1997) *Proc. Natl. Acad. Sci. U. S. A* 94, 11540-11545
30. Li, Y., Zhang, H., Litington, Y., and Chiang, C. (2006) *Proc. Natl. Acad. Sci. U. S. A* 103, 6548-6553
31. Kawakami, T., Kawcak, T., Li, Y. J., Zhang, W., Hu, Y., and Chuang, P. T. (2002) *Development* 129, 5753-5765
32. Ohashi, K., Osuga, J., Tozawa, R., Kitamine, T., Yagyu, H., Sekiya, M., Tomita, S. and others (2003) *J. Biol. Chem.* 278, 42936-42941
33. Tozawa, R., Ishibashi, S., Osuga, J., Yagyu, H., Oka, T., Chen, Z., Ohashi, K. and others (1999) *J. Biol. Chem.* 274, 30843-30848
34. Shehu, A., Mao, J., Gibori, G. B., Halperin, J., Le, J., Devi, Y. S., Merrill, B. and others (2008) *Mol. Endocrinol.* 22, 2268-2277
35. Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M., and Gonzalez, F. J. (2001) *Mol.*

Cell Biol. 21, 1393-1403

36. Raabe, M., Flynn, L. M., Zlot, C. H., Wong, J. S., Veniant, M. M., Hamilton, R. L., and Young, S. G. (1998) *Proc. Natl. Acad. Sci. U. S. A* 95, 8686-8691

37. De Strooper, B., Saftig, P., Craessaerts, K., Vanderstichele, H., Guhde, G., Annaert, W., Von Figura, K. and others (1998) *Nature* 391, 387-390

38. Hojjati, M. R., Li, Z., and Jiang, X. C. (2005) *Biochim. Biophys. Acta* 1737, 44-51

39. Cybulsky, M. I., Iiyama, K., Li, H., Zhu, S., Chen, M., Iiyama, M., Davis, V. and others (2001) *J. Clin. Invest* 107, 1255-1262

40. McNeish, J., Aiello, R. J., Guyot, D., Turi, T., Gabel, C., Aldinger, C., Hoppe, K. L. and others (2000) *Proc. Natl. Acad. Sci. U. S. A* 97, 4245-4250

41. Voshol, P. J., Schwarz, M., Rigotti, A., Krieger, M., Groen, A. K., and Kuipers, F. (2001) *Biochem. J.* 356, 317-325

42. Plosch, T., Kruit, J. K., Bloks, V. W., Huijkman, N. C., Havinga, R., Duchateau, G. S., Lin, Y. and others (2006) *J. Nutr.* 136, 2135-2140

43. Ro, H. S., Zhang, L., Majdalawieh, A., Kim, S. W., Wu, X., Lyons, P. J., Webber, C. and others (2007) *Obesity. (Silver. Spring)* 15, 288-302

44. Koster, A., Chao, Y. B., Mosior, M., Ford, A., Gonzalez-DeWhitt, P. A., Hale, J. E., Li, D. and others (2005) *Endocrinology* 146, 4943-4950

45. Oike, Y., Akao, M., Yasunaga, K., Yamauchi, T., Morisada, T., Ito, Y., Urano, T. and others (2005) *Nat. Med.* 11, 400-408

46. Parks, J. S., Li, H., Gebre, A. K., Smith, T. L., and Maeda, N. (1995) *J. Lipid Res.* 36, 349-355

47. Gautier, T., Tietge, U. J., Boverhof, R., Perton, F. G., Le Guern, N., Masson, D., Rensen, P. C. and others (2007) *J. Lipid Res.* 48, 30-40

48. Plump, A. S., Smith, J. D., Hayek, T., alto-Setala, K., Walsh, A., Verstuyft, J. G., Rubin, E. M. and others (1992) *Cell* 71, 343-353

49. Togawa, A., Miyoshi, J., Ishizaki, H., Tanaka, M., Takakura, A., Nishioka, H., Yoshida, H. and others (1999) *Oncogene* 18, 5373-5380

50. Wechsler, A., Brafman, A., Shafir, M., Heverin, M., Gottlieb, H., Damari, G., Gozlan-Kelner, S. and others (2003) *Science* 302, 2087

51. Wassif, C. A., Zhu, P., Kratz, L., Krakowiak, P. A., Battaile, K. P., Weight, F. F., Grinberg, A. and others (2001) *Hum. Mol. Genet.* 10, 555-564
52. Chakravarthy, M. V., Pan, Z., Zhu, Y., Tordjman, K., Schneider, J. G., Coleman, T., Turk, J. and others (2005) *Cell Metab* 1, 309-322
53. Wang, Y., Oeser, J. K., Yang, C., Sarkar, S., Hackl, S. I., Hasty, A. H., McGuinness, O. P. and others (2006) *J. Biol. Chem.* 281, 39982-39989
54. Beigneux, A. P., Davies, B. S., Gin, P., Weinstein, M. M., Farber, E., Qiao, X., Peale, F. and others (2007) *Cell Metab* 5, 279-291
55. Pontoglio, M., Barra, J., Hadchouel, M., Doyen, A., Kress, C., Bach, J. P., Babinet, C. and others (1996) *Cell* 84, 575-585
56. Devlin, C. M., Kuriakose, G., Hirsch, E., and Tabas, I. (2002) *Proc. Natl. Acad. Sci. U. S. A* 99, 6280-6285
57. Lambert, G., Sakai, N., Vaisman, B. L., Neufeld, E. B., Marteyn, B., Chan, C. C., Paigen, B. and others (2001) *J. Biol. Chem.* 276, 15090-15098
58. Jones, C., Hammer, R. E., Li, W. P., Cohen, J. C., Hobbs, H. H., and Herz, J. (2003) *J. Biol. Chem.* 278, 29024-29030
59. Ma, K., Cilingiroglu, M., Otvos, J. D., Ballantyne, C. M., Marian, A. J., and Chan, L. (2003) *Proc. Natl. Acad. Sci. U. S. A* 100, 2748-2753
60. Weinstock, P. H., Bisgaier, C. L., alto-Setala, K., Radner, H., Ramakrishnan, R., Levak-Frank, S., Essenburg, A. D. and others (1995) *J. Clin. Invest* 96, 2555-2568
61. Altmann, S. W., Davis, H. R., Jr., Zhu, L. J., Yao, X., Hoos, L. M., Tetzloff, G., Iyer, S. P. and others (2004) *Science* 303, 1201-1204
62. Alberti, S., Schuster, G., Parini, P., Feltkamp, D., Diczfalusy, U., Rudling, M., Angelin, B. and others (2001) *J. Clin. Invest* 107, 565-573
63. Lagace, T. A., Curtis, D. E., Garuti, R., McNutt, M. C., Park, S. W., Prather, H. B., Anderson, N. N. and others (2006) *J. Clin. Invest* 116, 2995-3005
64. Van Dyck, F., Braem, C. V., Chen, Z., Declercq, J., Deckers, R., Kim, B. M., Ito, S. and others (2007) *Cell Metab* 6, 406-413
65. Jiang, X. C., Bruce, C., Mar, J., Lin, M., Ji, Y., Francone, O. L., and Tall, A. R. (1999) *J. Clin. Invest* 103, 907-914

66. Varban, M. L., Rinninger, F., Wang, N., Fairchild-Huntress, V., Dunmore, J. H., Fang, Q., Gosselin, M. L. and others (1998) *Proc. Natl. Acad. Sci. U. S. A* 95, 4619-4624
67. Chen, L. Y., Shieh, J. J., Lin, B., Pan, C. J., Gao, J. L., Murphy, P. M., Roe, T. F. and others (2003) *Hum. Mol. Genet.* 12, 2547-2558
68. Buhman, K. K., Accad, M., Novak, S., Choi, R. S., Wong, J. S., Hamilton, R. L., Turley, S. and others (2000) *Nat. Med.* 6, 1341-1347
69. Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., Zhang, J. and others (2004) *Proc. Natl. Acad. Sci. U. S. A* 101, 6062-6067
70. Xie, C., Turley, S. D., and Dietschy, J. M. (2009) *J. Lipid Res.* 50, 1316-1329
71. Brunham, L. R., Kruit, J. K., Iqbal, J., Fievet, C., Timmins, J. M., Pape, T. D., Coburn, B. A. and others (2006) *J. Clin. Invest* 116, 1052-1062
72. Goldstein, J. L. and Brown, M. S. (1990) *Nature* 343, 425-430
73. Bennati, A. M., Schiavoni, G., Franken, S., Piobbico, D., la Fazia, M. A., Caruso, D., De Fabiani, E. and others (2008) *FEBS J.* 275, 5034-5047
74. Krakowiak, P. A., Wassif, C. A., Kratz, L., Cozma, D., Kovarova, M., Harris, G., Grinberg, A. and others (2003) *Hum. Mol. Genet.* 12, 1631-1641
75. Waterham, H. R. (2006) *FEBS Lett.* 580, 5442-5449
76. Schwarz, M., Russel, D.W., Dietschy, J.M., Turley, S.D. (2001) *J. Lipid Res.* 42, 1594-1607



CHAPTER

3

A physiologically-based kinetic model for the prediction of plasma cholesterol concentrations in the mouse.

Niek C.A. van de Pas, Ruud A. Woutersen, Ben van Ommen, Ivonne M.C.M. Rietjens, and Albert A. de Graaf

The chapter is based on *Biochim Biophys Acta*. 1811 (5):333-342, (2011)



Abstract

The LDL cholesterol (LDL-C) and HDL cholesterol (HDL-C) concentrations are determined by the activity of a complex network of reactions in several organs. Physiologically-based kinetic (PBK) computational models can be used to describe these different reactions in an integrated, quantitative manner.

A PBK model to predict plasma cholesterol concentrations in the mouse was developed, validated, and analyzed. Kinetic parameters required for defining the model were obtained using data from published experiments. To construct the model, a set of appropriate submodels was selected from a set of 65,536 submodels differing in the kinetic expressions of the reactions. A submodel was considered appropriate if it had the ability to correctly predict an increased or decreased plasma cholesterol concentration for a training set of 5 knockout mouse strains. The model thus defined consisted of 8 appropriate submodels and was validated using data from an independent set of 9 knockout mouse strains.

The model prediction is the average prediction of 8 appropriate submodels. Remarkably, these submodels had in common that the rate of cholesterol transport from the liver to HDL was not dependent on hepatic cholesterol concentrations. The model appeared able to accurately predict in a quantitative way the plasma cholesterol concentrations of all 14 knockout strains considered, including the frequently used *Ldlr*^{-/-} and *ApoE*^{-/-} mouse strains.

The model presented is a useful tool to predict the effect of knocking out genes that act in important steps in cholesterol metabolism on total plasma cholesterol, HDL-C and LDL-C in the mouse.

Introduction

Elevated plasma cholesterol is an important risk factor for cardiovascular diseases (1-3). The pharmaceutical industry is aiming to find new strategies to prevent these diseases. This search might be advanced with a better understanding of the factors and mechanisms that determine plasma cholesterol concentrations. Since these mechanisms involve complex multiple metabolic and transport reactions in various organs, it is difficult to obtain an integrated, quantitative insight in which factors mainly influence plasma cholesterol concentrations most and in what way (4-7). Mathematical modeling to integrate the many individual steps involved is, therefore, needed to predict the overall effect of genetic mutations, pharmacological or dietary interventions, and interspecies differences on plasma cholesterol concentrations.

60

There are only a few published mathematical models on plasma cholesterol (8-10). These models consider only plasma-associated processes and do not include the processes that are targets of drugs in organs, like hepatic cholesterol synthesis or cholesterol uptake. A better suited model would be a physiologically-based kinetic (PBK) model incorporating the cholesterol pools in the relevant compartments. A key aspect of PBK modeling is that the compartments denote real organs with real organ volumes (11,12) and that the model includes kinetic constants for real enzymatic activities. Thus, we chose to develop a PBK computational model for plasma cholesterol.

The modeling process can be subdivided into several steps (13,14). In the present paper five steps are considered. At first, a simplified representation of the biological system of interest is generated: the conceptual model. Secondly, the conceptual model is translated into a mathematical model by formulating mathematical equations to describe the different parts of the system as included in the conceptual model. Thirdly, the model is calibrated by setting parameter values. In this step, parameters are taken from literature or are fitted to experimental data using a parameter fitting procedure. The fourth step is testing whether the predictions of the integrated model are correct. For this purpose, model predictions are compared with experimental data not yet considered in the calibration step. This step is denoted as model validation. The fifth and final step is to perform simulations, and to analyze the model predictions as well as the model properties to obtain insights in the behavior of the biological system.

The present study describes steps 2-5 in the modeling approach:

a mathematical model formulation, calibration, model validation using published experimental data, and model analysis. The first step of the modeling approach, the construction of a conceptual model of cholesterol metabolism, was carried out previously (15). To this end relevant knockout mouse models were screened for altered plasma cholesterol concentrations compared to the wild type. If the alteration was more than two-fold (up or down), the corresponding gene was marked as key gene. Based on the function of a subset of 12 from these key genes, metabolic and transport reactions were included in the conceptual model. As a result, this conceptual model is based on a selection of relevant biological processes that determine plasma cholesterol concentrations (15). Figure 3.1 presents this conceptual model previously described in more detail (15).

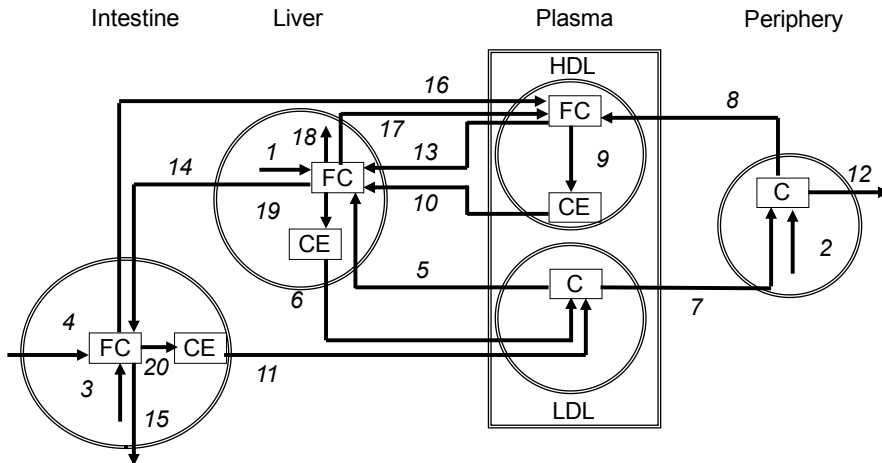


Figure 3.1. Conceptual model of reactions determining cholesterol plasma concentrations used to set up the computational model of the present study. Process numbers stand for: 1, Hepatic cholesterol synthesis; 2, Peripheral cholesterol synthesis; 3, Intestinal cholesterol synthesis; 4, Dietary cholesterol intake; 5, Hepatic uptake of cholesterol from LDL; 6, Hepatic VLDL secretion; 7, Peripheral uptake of cholesterol from LDL; 8, Peripheral cholesterol transport to HDL; 9, HDL-associated cholesterol esterification; 10, Hepatic HDL-CE uptake; 11, Intestinal chylomicron cholesterol secretion; 12, Peripheral cholesterol loss; 13, Hepatic HDL-FC uptake; 14, Biliary cholesterol excretion; 15, Fecal cholesterol excretion; 16, Intestinal cholesterol transport to HDL; 17, Hepatic cholesterol transport to HDL; 18, Hepatic bile acid synthesis to balance fecal bile acid loss; 19, Hepatic cholesterol esterification; and 20, Intestinal cholesterol esterification. C stands for cholesterol; CE for cholesteryl ester. Based on (15).

It contains 4 different compartments (liver, plasma, periphery, and intestine), 8 different cholesterol pools located within these compartments, and 20 reactions that transfer cholesterol among the different pools. Three of the cholesterol pools are present in plasma: High Density Lipoprotein (HDL) free cholesterol (HDL-FC), HDL cholesterol ester (HDL-CE), and non-HDL cholesterol (here referred to as Low Density Lipoprotein cholesterol LDL-C) that together form the total plasma cholesterol (TC) pool. The remaining 5 pools are intra-organ pools representing hepatic free cholesterol (Liv-FC), peripheral cholesterol (Per-C), intestinal cholesterol ester (Int-CE), hepatic cholesterol ester (Liv-CE), and intestinal free cholesterol (Int-FC).

62 In the present paper, the model was formulated and calibrated for the mouse, since the mouse is frequently used as model organism in cholesterol research (16). The model was validated by simulating different knockout mouse models and comparing the model predictions of plasma cholesterol with measurements reported in literature. The model obtained was able to describe the phenotype of 14 knockout mouse strains. In the analysis step, the model was analyzed with the aim to find kinetic properties in the model that can provide novel insights in the regulation of cholesterol metabolism.

Methods

Mathematical model formulation

Having defined the conceptual model (Figure 3.1) (15), mathematical model formulation was the next step in development of the PBK model for plasma cholesterol concentrations in the mouse. The conceptual model (Figure 3.1) contains 20 metabolic and transport reactions converting cholesterol from one pool to another. Each of the 8 pools in the model can be a substrate of multiple reactions, while at the same time being a product of other reactions. The symbols of variables and parameters that were used to formulate the mathematical model are given in Table 3.1.

Table 3.1. List of symbols representing variables and parameters used in the model description.

| Symbol | Description | Unit |
|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| V_j | volume of compartment j with j = liv, pla, per, and int for liver, plasma, periphery, and intestine | liter |
| v_i | reaction rate of reaction number i defined in Fig. 1 (i=1 - 20) | mmol/(kg day) |
| v_i^{ss} | steady state reaction rate of reaction i | mmol/(kg day) |
| $[C]_j$ | concentration of pool j, with j = Liv-FC, HDL-FC, HDL-CE, LDL-C, Per-C, Int-FC, Liv-CE, and Int-CE for liver free cholesterol, HDL free cholesterol, HDL cholesterol ester, LDL cholesterol, peripheral cholesterol, intestinal free cholesterol, liver cholesterol ester, and intestinal cholesterol ester. | mmol/liter |
| $[C]_j^{ss}$ | steady state cholesterol concentration in pool j | mmol/liter |
| $k_{0,i}$ | zero-order rate constant of reaction i | mmol/(kg day) |
| $k_{1,i}$ | first-order rate constant of reaction i | 1/(day) |
| $k_{0,i}^{ko}$ | zero-order rate constant of reaction i in a knockout strain | mmol/(kg day) |
| $k_{1,i}^{ko}$ | first-order rate constant of reaction i in a knockout strain | 1/(day) |
| t | time | day |
| f_{ko} | rate constant modification factor | [-] |
| M_{bw} | body weight | kg |

The PBK model was formulated as a set of differential equations each describing the time behavior of one of the modeled cholesterol concentrations as a function of the reaction rates. For practical reasons, the reaction rates (expressed by v), were numbered with the numbering corresponding to Figure 3.1. The differential equations are formulated in Table 3.2 (Eqn. 3.1-3.8).

Table 3.2. Equations used in model development, parameter calculation, and simulation.

| Differential equations | Number |
|-------------------------------------------------------------------------------------------------------------------------------|-----------|
| $\frac{1}{M_{bw}} \frac{d(V_{liv} \cdot [C]_{Liv-FC})}{dt} = v_1 + v_5 + v_{10} + v_{13} - v_{19} - v_{14} - v_{17} - v_{18}$ | Eqn. 3.1 |
| $\frac{1}{M_{bw}} \frac{d(V_{pla} \cdot [C]_{HDL-FC})}{dt} = v_8 + v_{16} + v_{17} - v_9 - v_{13}$ | Eqn. 3.2 |
| $\frac{1}{M_{bw}} \frac{d(V_{pla} \cdot [C]_{HDL-CE})}{dt} = v_9 - v_{10}$ | Eqn. 3.3 |
| $\frac{1}{M_{bw}} \frac{d(V_{pla} \cdot [C]_{LDL-C})}{dt} = v_6 + v_{11} - v_5 - v_7$ | Eqn. 3.4 |
| $\frac{1}{M_{bw}} \frac{d(V_{per} \cdot [C]_{Per-C})}{dt} = v_2 + v_7 - v_8 - v_{12}$ | Eqn. 3.5 |
| $\frac{1}{M_{bw}} \frac{d(V_{int} \cdot [C]_{Int-FC})}{dt} = v_3 + v_4 + v_{14} - v_{20} - v_{16} - v_{15}$ | Eqn. 3.6 |
| $\frac{1}{M_{bw}} \frac{d(V_{liv} \cdot [C]_{Liv-CE})}{dt} = v_{19} - v_6$ | Eqn. 3.7 |
| $\frac{1}{M_{bw}} \frac{d(V_{int} \cdot [C]_{Int-CE})}{dt} = v_{20} - v_{11}$ | Eqn. 3.8 |
| Rate equations | |
| $v_i = \frac{k_i^1 V_j [C]}{M_{bw}} \text{ (first-order kinetics)}$ | Eqn. 3.9 |
| $v_i = \frac{V_j k_i^0}{M_{bw}} \text{ (zero-order kinetics)}$ | Eqn. 3.10 |
| Parameter calculation | |
| $k_i^1 = \frac{M_{bw} v_i^{SS}}{V [C]_i^{SS}} \text{ (first-order kinetics)}$ | Eqn. 3.11 |
| $k_i^0 = \frac{M_{bw} v_i^{SS}}{V} \text{ (zero-order kinetics)}$ | Eqn. 3.12 |
| Knockout simulation | |
| $k_i^{ko,1} = f_{ko} k_i^1 \text{ (first-order kinetics)}$ | Eqn. 3.13 |
| $k_i^{ko,0} = f_{ko} k_i^0 \text{ (zero-order kinetics)}$ | Eqn. 3.14 |

For the model it is essential to describe the reaction rates by kinetic equations. Most reactions represent lumped sets of enzymatic- and/or transport reactions that may be regulated at the activity concentration and/or the expression concentration. This regulation in turn comprises potentially complex networks of nuclear receptors, transcription factors. Therefore, in virtually all cases, the various reactions can best be presented by kinetic equations containing apparent rate constants and apparent K_M values. The general solution is to use Michaelis-Menten kinetics as a prototype kinetic expression for biological reactions and defining apparent V^{\max} and apparent K_M values. In the present approach, however, in order to keep the number of parameters as limited as possible, it was additionally assumed that each reaction operates either in the linear part of a Michaelis-Menten kinetic curve (substrate concentration much lower than the apparent K_M), or in the saturated part (substrate concentration much higher than the apparent K_M). At low substrate concentrations Michaelis-Menten kinetics effectively reduces to first order kinetics (Eqn. 3.9, Table 3.2). At high substrate concentrations, Michaelis-Menten kinetics reduces to zero-order kinetics, i.e. becomes independent of substrate concentrations (Eqn. 3.10, Table 3.2). Also, to warrant model simplicity, feedback inhibition was not taken into consideration for any reaction.

For some of the 20 reactions included in the model, a decision could be made up-front to use zero or first order kinetics. This holds for reactions 1-3 in Figure 3.1, because it has been demonstrated that cholesterol can ultimately be synthesized from glucose and that there is no increase in the rate of cholesterol synthesis after the addition of glucose-6-phosphate (17), indicating that the synthesis of cholesterol is not substrate controlled. Therefore, the cholesterol biosynthesis reactions in the liver, intestine and the peripheral compartment (reactions 1-3 in Figure 3.1) were considered to be zero-order, and it is recognized that such a first approximation does not include feedback inhibition.

Reaction 4 denotes dietary intake of cholesterol. The amount of cholesterol a mouse eats varies over the day, since periods of feeding are followed by periods of not eating. However, when bearing in mind that mouse drug experiments usually take weeks, the dietary intake of cholesterol (reaction 4) can be considered relatively constant over time. Therefore, the dietary intake was also treated as a zero-order reaction (reaction 4 in Figure 3.1). For the remaining 16 reactions (reactions 5-20) it was not possible to decide beforehand whether the reaction would

follow zero or first-order kinetics. Therefore, for the other 16 reactions, ensemble modeling (18) was used.

Thus, rather than formulating one model with a set of kinetic constants, a model set (ensemble) with multiple submodels with all possible combinations of first and zero-order for the remaining 16 reactions was constructed. This resulted in a model consisting of 2^{16} (or 65,536) submodels. Not all submodels were considered to be equally valid, since some of them were expected to give unrealistic predictions, i.e. negative or infinite concentrations. Therefore, the calibration step described below, was used to select the subset of appropriate submodels. The overall model prediction was calculated as the average of the model predictions of these appropriate submodels.

66

In this study, only steady state predictions on cholesterol concentrations in the various pools were considered. In steady state, none of the concentrations of cholesterol in the 8 model pools change in time and, therefore, the left hand sides of equations 3.1-3.8 in Table 3.2 all reduce to zero. The right hand sides of these equations then reduce to a set of 8 linear equations for v_1^{ss} to v_{20}^{ss} . Since, however, all v^{ss} 's can be expressed in terms of the 8 unknown steady-state cholesterol concentrations included in the model (Eqn. 3.11 and 3.12, Table 3.2), this essentially results in a system of 8 equations with 8 unknowns. This allowed calculating the steady state concentrations using standard matrix algebra from MATLAB version 7.5 (R2007b).

The initial purpose of the model is to predict the phenotype of knockout mouse strains. In a knockout mouse, the reaction corresponding to the disrupted gene will be absent. In biological practice, however, compensatory mechanisms often operate, resulting in the activation of a backup reaction that performs the same or a similar function as the reaction targeted by the knockout. To account for these backup reactions in the model simulations of knockout strains, so-called knockout factors f_{ko} were introduced, representing the fraction of activity of a specific reaction left after a knockout. The f_{ko} were implemented as multiplication factors of the specific reaction rate constants for the reactions modulated by the gene knockout, as specified by equations 3.13 and 3.14 (Table 3.2). The value of f_{ko} was constrained between 0 and 1, which corresponds to a full reduction and no reduction, respectively, of the reaction rate affected by the specific knockout. The values for f_{ko} were chosen based on literature data as described in the results.

Model calibration

The model calibration contained two steps: setting the model parameters using literature data and selecting the appropriate submodels from the ensemble of 2^{16} submodels. The kinetic parameters (k_i^0, k_i^1) in the model were calculated using experimental data on wild type steady state concentrations and fluxes found in the literature as detailed in the results. Data from male wild type mice on a standard chow diet were used where available. In case v_i was represented by first-order kinetics, the corresponding rate constant was calculated with the pool sizes and steady state reaction rates according to Eqn. 3.11 (Table 3.2). When v_i was represented by zero-order kinetics, the corresponding rate constant was calculated according to Eqn. 3.12 (Table 3.2).

As a second calibration step, the appropriate submodels among the total of 65,536 submodels were selected. Submodels were regarded appropriate if able to correctly predict an increased or decreased concentration in the relevant cholesterol pool for a knockout mouse strain compared to the wild type. For this selection, 5 knockout mouse strains were used. These were selected from a set of 14 knockout mouse strains including strains deficient in 5 classes of genes: 1) genes responsible for LDL metabolism (*Ldlr*^{-/-}, *ApoE*^{-/-}), 2) genes for cholesterol synthesis (*Dhcr24*^{-/-}), 3) genes for HDL metabolism (*Abca1*^{-/-}, *Abca1*^{-L/-L}, *Abca1*^{-I/-I}, *Scarb1*^{-/-}, and *Apoa1*^{-/-}), 4) genes for intestinal cholesterol uptake and biliary cholesterol secretion (*Npc1l1*^{-/-}, *Abcg5*^{-/-}, *Abcg8*^{-/-}, *Abcb4*^{-/-}), and 5) genes for cholesterol esterification (*Lcat*^{-/-}, *Soat2*^{-/-}). These 14 strains were divided into a model selection set and a model validation set. The selection set included one randomly selected strain from each of the five strain classes mentioned above, i.e. the *Ldlr*^{-/-}, *Dhcr24*^{-/-}, *Apoa1*^{-/-}, *Abcg8*^{-/-}, and *Soat2*^{-/-} mouse strains. An independent validation set comprised the remaining 9 strains. In order to define the PBK model and to make a selection of the appropriate submodels from the set of 65,536 submodels, in first instance 7 qualitative criteria were defined using data from the 5 mouse strains of the selection set. Table 3.3 presents these criteria, which are based on HDL-C, LDL-C, and TC as possible predicted endpoints.

An example is the criterion on the *Apoa1*^{-/-} mouse, which comprised the following: “Is the plasma HDL-C concentration predicted for the *Apoa1*^{-/-} mouse lower than that predicted for the HDL-C concentration of the corresponding wild type mouse?”. A submodel fulfilled this criterion if it predicted that the HDL-C concentration in the

Apoa1^{-/-} mouse was indeed lower than that in the corresponding wild type mouse.

Table 3.3: Criteria for the selection of appropriate submodels. A submodel fulfills the criterion, if the prediction of the submodel for the relevant cholesterol pool affected in the specified mouse strain matches the experimental observation (increase or decrease with respect to the wild type). See text for details.

| Criterion nr | Strain | Cholesterol pool affected | Experimental observation |
|--------------|------------------------------|---------------------------|--------------------------|
| 1 | <i>Ldlr</i> ^{-/-} | TC | Increase |
| 2 | <i>Ldlr</i> ^{-/-} | LDL-C | Increase |
| 3 | <i>Dhcr24</i> ^{-/-} | TC | Decrease |
| 4 | <i>Apoa1</i> ^{-/-} | TC | Decrease |
| 5 | <i>Apoa1</i> ^{-/-} | HDL-C | Decrease |
| 6 | <i>Abcg8</i> ^{-/-} | TC | Decrease |
| 7 | <i>Soat2</i> ^{-/-} | TC | Decrease |

68

A submodel failed to fulfill a criterion if it: 1) failed to reach a unique steady state, 2) predicted a steady state with one or more negative concentrations, 3) predicted a decreased plasma cholesterol concentration compared to the wild type while an increased concentration was reported, 4) predicted an increased plasma cholesterol concentration compared to the wild type while a decreased concentration was reported, or 5) predicted an unchanged plasma cholesterol concentration whereas a changed concentration was reported in the literature. Only the submodels fulfilling all 7 calibration quality criteria were considered appropriate and were kept for further validation and analysis steps.

What is henceforth referred to as “the model prediction” consists of the unweighted average of the predictions of the individual appropriate submodels that passed the selection procedure; “the model” accordingly designates the ensemble of appropriate submodels.

Model validation

The model validation consisted of a comparison of experimental data reported in literature with the model prediction 1) for the mouse strains included in the selection set; and 2) for the mouse strains included in the fully independent validation set. As a measure of accuracy of the model prediction, the relative change of cholesterol plasma concentrations in these knockout mice vs their wild type counter parts was considered.

As an additional aspect of the validation the values for f_{ko} derived

from literature data were compared with optimal values of f_{ko} , i.e. those f_{ko} values that would lead to predictions closest to the experimentally observed plasma cholesterol concentrations, for all 14 strains. These optimal f_{ko} values were obtained by minimizing the difference between model predictions and experimental data using a constrained optimization routine. The optimization of f_{ko} was performed with the function '*fminbnd*' in MATLAB. This function minimizes the squared difference between the predicted and the measured ratio of the knockout and the wild type TC concentration. The values of f_{ko} during optimization were constrained between 0 and 1.

Model analysis

The goal of the model analysis was to find indications for *in vivo* kinetics of reactions included in the model. Central is the assumption that the *in vivo* system more likely has kinetic properties similar to the best performing (i.e. appropriate) submodels than to poorly performing submodels. Therefore, for all reactions, the tendency of the appropriate submodels from the ensemble to show preferential zero or first order kinetics was evaluated. The results were compared with literature data on kinetic reactions *in vivo* or *in vitro*.

Results

Model calibration

The conceptual model given in Figure 3.1 is converted to a mathematical model involving variables and parameters (Table 3.1) as described in the results. The first part of model calibration was the setting of model parameters using compartmental volumes, the steady state concentrations, and reaction rates as obtained from literature. Many of the data were obtained from a recent paper by Xie et al. (19) which describes a large set of experiments on the synthesis, and the uptake of HDL-C and LDL-C in mice maintained on a low cholesterol diet. Of special interest is Figure 2 in (19) where the authors report organ-specific cholesterol concentrations. All data in (19) were assumed to be representative for the steady state. From this paper, also cholesterol synthesis and LDL-C and HDL-C uptake data were obtained. The steady state rates on the dietary cholesterol intake, hepatic VLDL-C secretion, chylomicron cholesterol secretion, and biliary cholesterol excretion were obtained from other sources (Table 3.4).

Table 3.4. Numerical values for the PBK model constants and steady state variables used in the calculations along with their literature sources. Symbols as in Table 3.1.

| Organ | Symbol | Value | Obtained from |
|----------------------------------------------|---------------------|-------|---------------------------------------------------------------------------------------|
| Volumes ^a (ml) | | | |
| Total body | V_{TB} | 25.1 | (20) |
| Liver ^c | V_{liv} | 1.4 | (20) |
| Small intestine ^c | V_{int} | 0.6 | (20) |
| Plasma ^c | V_{pla} | 1.5 | (21) |
| Periphery | V_{per} | 21.6 | ^b |
| Steady state concentrations (mM) | | | |
| Liver free cholesterol | $[C]_{Liv-FC}^{ss}$ | 3.98 | (19) ^d |
| HDL free cholesterol | $[C]_{HDL-FC}^{ss}$ | 0.56 | (19) ^e |
| HDL cholesterol ester | $[C]_{HDL-CE}^{ss}$ | 1.12 | (19) ^e |
| LDL cholesterol pool | $[C]_{LDL-C}^{ss}$ | 0.52 | (19) ^f |
| Peripheral cholesterol | $[C]_{Per-C}^{ss}$ | 5.74 | (19) |
| Intestinal free cholesterol | $[C]_{Int-FC}^{ss}$ | 4.13 | (19) ^g |
| Liver cholesterol ester | $[C]_{Liv-CE}^{ss}$ | 3.26 | (19) ^d |
| Intestinal cholesterol ester | $[C]_{Int-CE}^{ss}$ | 3.38 | (19) ^g |
| Steady state reaction rates (mmol/(kg day)) | | | |
| Hepatic cholesterol synthesis | v_1^{ss} | 0.18 | (19) |
| Peripheral cholesterol synthesis | v_2^{ss} | 0.12 | (19) |
| Intestinal cholesterol synthesis | v_3^{ss} | 0.09 | (19) |
| Dietary cholesterol intake | v_4^{ss} | 0.08 | (22) |
| Hepatic uptake of cholesterol from LDL | v_5^{ss} | 0.12 | $v_5^{ss} = v_{11}^{ss} + v_6^{ss} - v_7^{ss}$ |
| Hepatic VLDL secretion | v_6^{ss} | 0.10 | (23) |
| Peripheral uptake of cholesterol from LDL | v_7^{ss} | 0.01 | (19) |
| Peripheral cholesterol transport to HDL | v_8^{ss} | 0.09 | $v_8^{ss} = v_2^{ss} + v_7^{ss} - v_{12}^{ss}$ |
| HDL associated cholesterol esterification | v_9^{ss} | 0.12 | $v_9^{ss} = v_{10}^{ss}$ |
| Hepatic HDL-CE uptake | v_{10}^{ss} | 0.12 | (19) |
| Intestinal chylomicron cholesterol secretion | v_{11}^{ss} | 0.04 | (24) |
| Peripheral cholesterol loss | v_{12}^{ss} | 0.04 | $v_{12}^{ss} = v_1^{ss} + v_2^{ss} + v_3^{ss} + v_4^{ss} - v_{18}^{ss} - v_{15}^{ss}$ |
| Hepatic HDL-FC uptake | v_{13}^{ss} | 0.06 | (25) |
| Biliary cholesterol excretion | v_{14}^{ss} | 0.02 | (26) |
| Fecal cholesterol excretion | v_{15}^{ss} | 0.14 | (19) |
| Intestinal cholesterol transport to HDL | v_{16}^{ss} | 0.01 | $v_{16}^{ss} = v_4^{ss} + v_3^{ss} + v_{14}^{ss} - v_{15}^{ss} - v_{20}^{ss}$ |
| Hepatic cholesterol transport to HDL | v_{17}^{ss} | 0.08 | $v_{17}^{ss} = v_9^{ss} + v_{13}^{ss} - v_8^{ss} - v_{16}^{ss}$ |
| Hepatic cholesterol catabolism | v_{18}^{ss} | 0.28 | (19) |
| Hepatic cholesterol esterification | v_{19}^{ss} | 0.10 | $v_{19}^{ss} = v_6^{ss}$ |
| Intestinal cholesterol esterification | v_{20}^{ss} | 0.04 | $v_{20}^{ss} = v_{11}^{ss}$ |

^a Volumes were calculated from body and organ masses assuming a tissue density of 1 kg/L. ^b The peripheral mass was obtained by subtraction of the plasma, liver, and intestinal masses from the total body mass. ^c The indicated references mention organ masses as a percentage of the total body mass. Absolute masses were calculated using a total body mass of 25.1 g for the mouse (19,20). ^d Ratio between Liv-FC and Liv-CE was assumed to be 1:1.22 (27) ^e Ratio between HDL-FC and HDL-CE was assumed to be 1:3 (28). ^f The concentration of LDL-C was calculated as the difference between TC and HDL-C. ^g The ratio of Int-FC and Int-CE was assumed to be identical to that ratio in the liver for the wild type mouse.

The other reaction rates had to be indirectly derived from the data i.e. they were calculated using mass balance equations. As an example, the rate for hepatic cholesterol esterification (reaction 19) was calculated using the hepatic cholesterol ester balance. All numerical values thus obtained for the model constants and steady state variables, along with their source (direct or indirect), are given in Table 3.4.

To define the PBK model in first instance a set (ensemble) of 65,536 submodels was constructed, all based on the model setup in Figure 3.1, but unique in their kinetics (each submodel presenting a unique combination of first or zero-order kinetics for reactions 5-20, while reactions 1-4 were defined to be zero order in all submodels for reasons explained in the methods. Once a reaction was considered to be zero- or first order, its kinetic constant was calculated from the corresponding steady state concentration and reaction rate values in Table 3.4 using Eqn. 3.11 and 3.12 in Table 3.2. The kinetic parameters thus obtained (Table 3.5) can give insight in time constants involved in cholesterol metabolism.

For instance, the peripheral cholesterol pool was predicted to have the slowest turnover rate, since reactions 12 and 8 had the lowest first order rate constants (Table 3.5). The turnover time of the peripheral pool was 25 days, while the other turnover times were between 0.1 days (for HDL-FC) and 3.0 days (for the Int-CE pool).

The second part of the model calibration consisted of the selection of the appropriate submodels. For this reason, all submodels were tested for their ability to correctly predict in a qualitative way a higher or lower concentration of HDL-C, LDL-C, or TC for a selection set of 5 knockout mouse models compared to the corresponding wild types. To this end, a set of 7 quality criteria given in Table 3.3 was used.

In the simulations, the various knockout mouse strains were modeled by multiplying the specific reaction rate constant of the affected reaction by a strain-specific knockout factor (f_{ko}) (Eqn. 3.13, 3.14 in Table 3.2). These f_{ko} values were obtained from data reported from biological experiments, like cannulations and isotope flux studies. Table 3.6 gives the definition and values of the f_{ko} for the various strains.

Table 3.5. First and zero-order kinetic parameters used to define the submodels. Kinetic parameters were calculated according to eqn. 3.11 and 3.12 (Table 3.2) using the data given in Table 3.4.

| Reaction nr | (mmol/(kg day)) | (1/day) |
|-------------|-----------------|---------|
| 1 | 0.18 | |
| 2 | 0.12 | |
| 3 | 0.09 | |
| 4 | 0.08 | |
| 5 | 1.46 | 2.80 |
| 6 | 2.01 | 0.62 |
| 7 | 0.17 | 0.32 |
| 8 | 0.10 | 0.02 |
| 9 | 1.38 | 2.47 |
| 10 | 1.38 | 1.24 |
| 11 | 0.77 | 0.23 |
| 12 | 0.05 | 0.01 |
| 13 | 0.74 | 1.32 |
| 14 | 0.42 | 0.10 |
| 15 | 2.65 | 0.64 |
| 16 | 0.24 | 0.06 |
| 17 | 1.68 | 0.42 |
| 18 | 5.95 | 1.49 |
| 19 | 2.01 | 0.50 |
| 20 | 0.77 | 0.19 |

For the *Abcg8*^{-/-} mouse for example, the value of f_{ko} has been set to the ratio of the biliary sterol secretion rate of the *Abcg8*^{-/-} strain and the control mouse as determined in biliary cannulation experiments (29). Within the series of the 14 knock out strains, the f_{ko} values thus determined ranged from 0.01 for the *Lcat*^{-/-} to 0.30 for the *Abcg8*^{-/-}, and *NpcIII*^{-/-} strain. For the *Dhcr24*^{-/-} mouse, deficient in 3-beta-hydroxysterol delta-24-reductase, no value for f_{ko} could be found in the literature. However, given that no alternatives for this reaction are expected *in vivo*, the value of f_{ko} was set to zero.

Figure 3.2 shows the number of criteria that were fulfilled by the submodels. Many submodels fulfilled none of the criteria. Only 8 submodels (0.01% of the total of 65,536) fulfilled all 7 qualitative selection criteria. The overall model prediction was, therefore, defined as the average of the predictions of these 8 appropriate submodels.

Model validation

As an initial validation, the 8 submodels were used to quantitatively predict the TC, LDL-C, and HDL-C cholesterol concentrations of the 5 knockout models included in the selection set. Re-using the set that was also used for model calibration is considered a valid procedure since the selection procedure only involved qualitative aspects (higher/lower than wild type). All model predictions agreed well with the experimental data (Figure 3.3). The model correctly predicted a large decrease in TC concentration for the *Dhcr24*^{-/-} mouse. For the other mouse strains, the TC concentrations were correctly predicted within an average accuracy of a factor 1.27.

As a second validation, the model was used to predict the phenotype of an independent set of 9 knockout mouse strains (i.e. the validation set). This simulation used values for the parameter f_{ko} obtained from literature as shown in Table 3.6. For many strains, the values for f_{ko} were close to zero, but for several strains f_{ko} was considerably larger than zero, for example the *Scarb1*^{-/-} strain had f_{ko} 0.42 indicating the presence of *Scarb1*-independent transport of cholesterol from HDL to the liver (reactions 10 and 13).

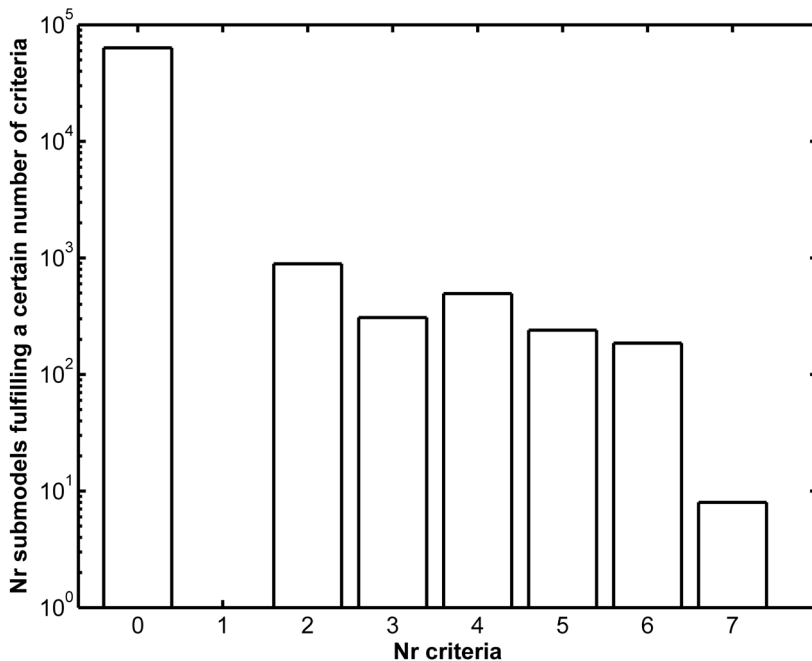
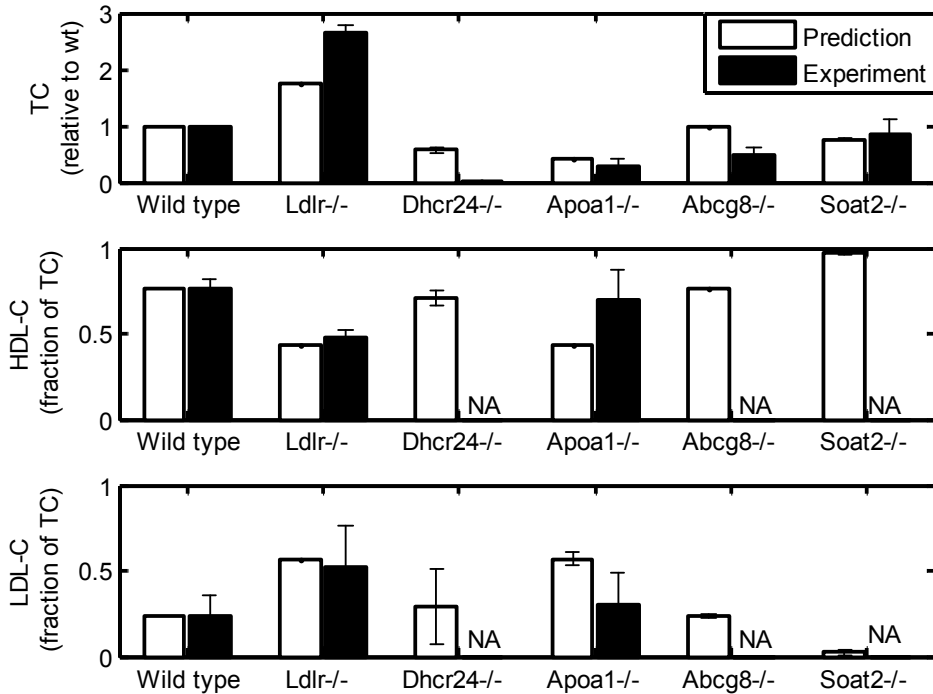


Figure 3.2. Histogram of the number of submodels that fulfilled the indicated number of selection criteria as defined in Table 3.3 (15).



74

Figure 3.3. PBK model predictions (open bars) compared to experimental data (closed bars) for the 5 different knockout mouse strains of the selection set. The figure indicates the TC of the knockout mouse as a fraction of the wild type TC concentration (top panel), the fraction of TC in HDL in the knockout strains (middle panel), and the fraction of TC in the non-HDL pool (bottom panel). Error bars on predicted values indicate the standard deviation of the predictions of the 8 different submodels. Error bars on experimental data were based on the standard deviations reported in literature. For *Dhcr24*^{-/-}, *Soat2*^{-/-}, and *Abcg8*^{-/-}, no literature HDL-C data were available. NA, not available. Experimental data were obtained from (22,29-32).

For all mouse strains in the validation set, the model correctly predicted whether the plasma TC concentrations were higher or lower compared to the wild type (Figure 3.4).

Quantitatively, the model performed adequately for the mouse strains deficient in lipoprotein-related genes (*Abca1*^{-/-}, *ApoE*^{-/-}, and *Scarbl*^{-/-} mouse), but a somewhat less accurate prediction was obtained

for the effect produced by knockout of the biliary cholesterol related gene *Abcb4*^{-/-} and organ-specific *Abca1* knockout strains. For all the 9 strains, the model was able to predict the fold-change in plasma TC concentration with an average deviation of less than a factor 2. Another different ability of the model is to correctly predict the distribution of cholesterol between HDL and non-HDL in the mouse strains (Figures 3.3 and 3.4, middle and bottom panel). An exception to this good model performance is the *Lcat*^{-/-} mouse strain.

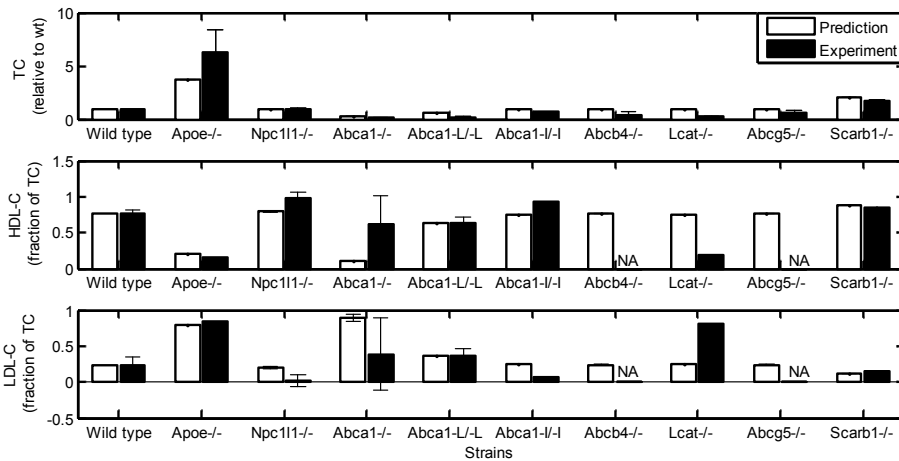


Figure 3.4. PBK model predictions (open bars) compared to experimental data (closed bars) for the validation set of 9 mouse knockout strains. The figure indicates the TC of the knockout mouse as a fraction of the wild type TC concentration (top panel), the fraction of TC in HDL in the knockout strains (middle panel), and the fraction of TC in the non-HDL pool (bottom panel). Error bars on predicted values indicate the standard deviation of the predictions of the 8 different submodels. Error bars on data were based on the reported standard deviations in literature. For the *Abcg5*^{-/-} strain, no literature HDL-C data were available. NA, not available. Experimental data were obtained from (33-40).

As a final aspect of the validation of the model, the model was used to calculate which values of f_{ko} provided the best fit between predicted and experimental TC concentrations. Table 3.6 presents the f_{ko} values thus obtained (f_{ko} fitted).

Table 3.6. Reactions affected by specific knockouts as well as corresponding knockout factors f_{ko} (taken from literature and fitted to plasma cholesterol concentrations) used for the model selection (selection set) and model validation (validation set). NA, Not Available.

| Strain | Affected reaction (values in parentheses refer to Fig. 1) | f_{ko} from literature | f_{ko} fitted | Definition of f_{ko} The ratio of the: | Ref |
|------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|-----------------|---------------------------------------------------------------------------------------------------------------|------|
| Selection set | | | | | |
| <i>Ldlr</i> ^{-/-} | Hepatic uptake of cholesterol from LDL (5) and peripheral uptake of cholesterol from LDL (7) | 0.24 ± 0.08 | 0.12 | plasma clearance rate of labeled retinol from Apob containing lipoproteins for the knockout vs the wild type. | (41) |
| <i>Dhcr24</i> ^{-/-} | Hepatic cholesterol synthesis (1), peripheral cholesterol synthesis (2), intestinal cholesterol synthesis (3) | NA | 0.00 | See text | |
| <i>Apoa1</i> ^{-/-} | Peripheral contribution to HDL biogenesis (8), intestinal contribution to HDL biogenesis (16), and hepatic contribution to HDL biogenesis (17) | 0.14 ± 0.05 | 0.03 | HDL-CE transport rate for the knockout vs the wild type | (31) |
| <i>Abcg8</i> ^{-/-} | Biliary cholesterol excretion (14) | 0.30 ± 0.16 | 0.00 | sterol secretion rate into the bile for the knockout vs the wild type | (29) |
| <i>Soat2</i> ^{-/-} | Hepatic cholesterol esterification (19), Intestinal cholesterol esterification (20) | 0.08 ± 0.09 | 0.39 | incorporation rate of oleoyl-CoA into CE for the knockout vs the wild type | (32) |
| Validation set | | | | | |
| <i>Apoe</i> ^{-/-} | Hepatic uptake of cholesterol from LDL (5) and peripheral uptake of cholesterol from LDL (7) | 0.08 ± 0.02 | 0.04 | plasma clearance rate of labeled retinol from Apob containing lipoproteins for the knockout vs the wild type. | (41) |
| <i>Npc1l1</i> ^{-/-} | Intestinal cholesterol uptake (20 and 16) | 0.30 ± 0.02 | 0.67 | intestinal cholesterol absorption rate for the knockout vs the wild type | (42) |
| <i>Abca1</i> ^{-/-} | Peripheral contribution to HDL biogenesis (8), intestinal contribution to HDL biogenesis (16), and hepatic contribution to HDL biogenesis (17) | 0.02 ± 0.01 | 0.06 | cholesterol efflux rate for knockout macrophages to Apoa1 vs the wild type | (43) |
| <i>Abca1</i> ^{-/-} <i>L</i> ^{-/L} (Liver specific) | Hepatic contribution to HDL biogenesis (17) | 0.02 ± 0.01 | 0.00 | cholesterol efflux rate for knockout macrophages to Apoa1 vs the wild type | (43) |
| <i>Abca1</i> ^{-/-} <i>I</i> ^{-/I} (intestinal specific) | Intestinal contribution to HDL biogenesis (16) | 0.02 ± 0.01 | 0.00 | cholesterol efflux rate for knockout macrophages to Apoa1 vs the wild type | (43) |
| <i>Lcat</i> ^{-/-} | HDL associated cholesterol esterification (9) | 0.01 ± 0.00 | 0.00 | the conversion rate of [14C]cholesterol into HDL-CE for the knockout vs the wild type | (36) |
| <i>Abcg5</i> ^{-/-} | Biliary cholesterol excretion (14) | 0.19 ± 0.10 | 0.00 | sterol secretion rate into the bile for the knockout vs the wild type | (44) |
| <i>Abcb4</i> ^{-/-} | Biliary cholesterol excretion (14) | 0.04 ± 0.05 | 0.00 | cholesterol output rate into the bile for the knockout vs the wild type | (45) |
| <i>Scarb1</i> ^{-/-} | Hepatic HDL-CE uptake (10) and hepatic HDL-FC uptake (13) | 0.42 ± 0.07 | 0.52 | hepatic uptake of HDL-CE from plasma in the <i>Scarb1</i> ^{-/-} mouse vs in the wild type | (46) |

Comparison of these values to the ones derived from literature shows that the difference between the optimized f_{ko} and the literature value was generally less than 2 times the standard deviation of the literature value (Table 3.6). The largest differences between the fitted f_{ko} and literature f_{ko} values were found for the *Soat2*^{-/-}, *Npc1l1*^{-/-}, and *Abcg5*^{-/-} mouse. Using the optimized f_{ko} values, the model was able to exactly predict the TC concentrations for 6 out of 14 mouse strains (Figure 3.5). The largest differences between experimental data and model predictions were found for knockouts of genes involved in cholesterol synthesis (*Dhcr24*), biliary cholesterol secretion (*Abcg5*, *Abcg8*, and *Abcb4*), and HDL-associated cholesterol esterification (*Lcat*). Also with optimized f_{ko} , the model was able to correctly predict the fraction of cholesterol in HDL and non-HDL for many strains (Figure 3.5 bottom panels).

77

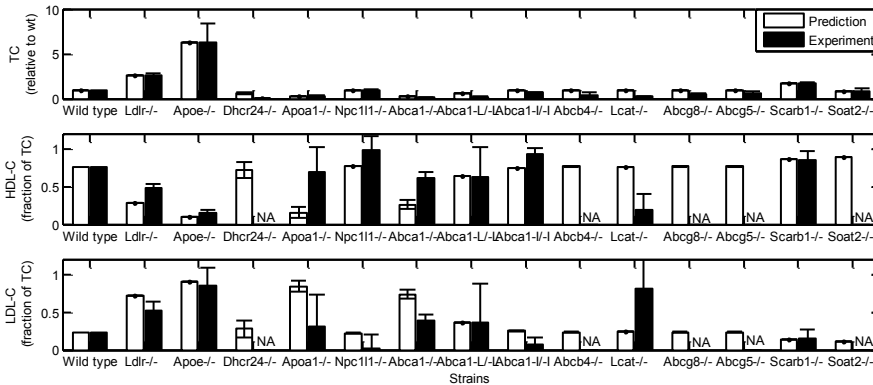


Figure 3.5. PBK model predictions with optimized values of f_{ko} (open bars) compared to experimental data (closed bars) for the wild type strain and 14 different knockout strains. The figure indicates the TC of the knockout mouse as a fraction of the wild type TC concentration (top panel), the fraction of TC in HDL in the knockout strains (middle panel), and the fraction of TC in the non-HDL pool in the different knockout mouse strains (bottom panel). For *Dhcr24*^{-/-}, *Soat2*^{-/-}, *Abcg5*^{-/-}, and *Abcg8*^{-/-}, no literature HDL-C data were available. NA, not available. Experimental TC data were obtained from (22,29-39).

Model analysis

In the model analysis step, the properties of the 8 best kinetic submodels were compared with the other submodels. The selected appropriate submodels were characterized by the nature of their rate constants, i.e. zero order for reactions 1-4, and zero or first order for reactions 5-20, as shown in Figure 3.6.

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|--------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|
| 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | |
| 2 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | |
| 3 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | |
| 4 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | |
| 5 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | |
| 6 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | |
| 7 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | |
| 8 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| | | Reaction nr. | | | | | | | | | | | | | | | | | | | |

78

Figure 3.6. The kinetic orders (zero or first order) of the reactions in 8 appropriate submodels from the total ensemble of 65,536 submodels. The horizontal axis shows the 20 reactions (reaction numbers corresponding to those presented in Figure 3.1). The vertical axis indicates the reaction kinetic pattern of the 8 different submodels. A white box with a 0 indicates a zero-order reaction, a grey box with a 1 indicates first order kinetics. The first 4 reactions in Figure 3.1 were always zero-order (see Methods, for explanation).

Most (66%) of the 160 reactions (20 reactions for 8 appropriate submodels) appeared to be first order, while only 54 reactions (34%) were zero-order. Interestingly, the submodel with reactions 5-20 all having first-order kinetics was not in the selected set of 8 appropriate submodels, since this submodel failed to fulfill the criterion of the *Abcg8*^{-/-} strain. One single reaction was zero-order in all 8 selected submodels (Figure 3.6), this reaction describes the cholesterol transport from the liver to HDL (reaction 17, Figure 3.1).

In the selected appropriate submodels, only 4 reactions show both first and zero-order kinetics. These reactions are biliary cholesterol excretion, cholesterol transport from the intestine to HDL, and hepatic and intestinal cholesterol esterification (reaction numbers 14, 16, 19, and 20 in Figure 3.1). The other 11 reactions follow first order kinetics in all 8 submodels. This set includes all lipoprotein uptake reactions from plasma (reaction numbers 5, 7, 10, and 13 in Figure 3.1) and all reactions that remove cholesterol from the body (reactions 12, 15, and 18 in Figure 3.1).

Discussion

This paper describes the formulation, calibration, validation, and analysis of a computational model for cholesterol plasma concentrations in the mouse. The presented computational model was able to predict the effects of genetic mutations on total plasma cholesterol in the mouse, as demonstrated by simulating 14 knockout mouse strains. To the best of our knowledge, there is presently no PBK model available that is able to quantitatively predict the effects of these 14 mutations on plasma TC, HDL-C, and non-HDL-C concentrations (since most non HDL-C is present in LDL, this is also named LDL-C). These clinically relevant risk factors are important biomarkers in for the development of new drugs and dietary intervention studies.

The model was validated by comparing model predictions of plasma concentrations of 14 knockout strains with experimental data. The model predicted the correct plasma TC concentrations of knockout mouse strains, using literature values of the defined rate constant modification factor f_{ko} , within an accuracy of a factor 2. This is considered highly successful within the present state-of-the-art of PBK modeling, where quantitative predictions may generally be correct within one order of magnitude (47-50). In addition to validating our modeling approach, starting with the definition of a conceptual model based on key genes (15), this result supports the feasibility and validity of the approach of using apparent simple zero and first order kinetics instead of more complex kinetic models derived from *in vitro* data.

Generally, the model predictions were within the range of the experimental data variation. For the *ApoE*^{-/-} mouse, for example, the TC prediction was 6.2 times the wild type TC (Figure 3.4). This is well within the experimental range of 2.8 and 9.3 times the wild type concentration reported by Guo et al. (51) and Ishida et al. (52) respectively (all data on a standard diet).

A considerable part of the differences between predictions and experimental data could be eliminated by model simulations with optimized instead of literature values of f_{ko} . The differences that remained after these simulations can obviously be due to the simplifications of the model structure. The model could miss an essential mechanism (e.g. transcriptional regulation of genes in response to changing conditions) and the kinetics could have been simplified too much by only including zero and first order and not considering feedback mechanisms. Also, due to inter-individual variation in the steady state reaction rates, the

used steady state reaction rates (Table 3.4) could differ from the actual values in individual mice. In particular, the chylomicron production rate (reaction 11) as determined from lymphatic cannulation is highly variable between mice (24).

Studying the kinetics of cholesterol metabolism *in vivo* is very difficult, therefore in the present study an ensemble modeling-based approach is used to infer kinetic aspects of cholesterol metabolism. The simplified apparent kinetic properties of the 8 appropriate submodels seem well-suited to reproduce experimental data. This suggests that these submodels have properties in common with the *in vivo* system. The kinetics of the selected 8 submodels were in some cases in agreement with *in vitro* data. As an example reaction 9, representing the enzyme Lcat, was first order in the 8 appropriate submodels, which is in agreement with the finding of Collet and Fielding (53) that the *in vitro* determined K_M for cholesterol of this reaction is much higher than the HDL-FC concentration of 0.56 mM (Table 3.3).

80

Disagreement between the kinetics of the 8 appropriate submodels and *in vitro* data was observed for reaction 18 that represents hepatic bile acid synthesis to balance fecal bile acid excretion. The finding that reaction 18 was first order in the 8 appropriate submodels is not in agreement with the finding of Ozasa and Boyd (54) that the apparent K_M (15 μ M) for cholesterol for the enzyme Cyp7a1 is considerably lower than the hepatic free cholesterol concentration (3.98 mM, Table 3.4). Two of the possible explanations for the difference between the *in vitro* findings and the *in vivo* situation in this case are 1) the transcriptional regulation: if, in response to the cholesterol concentration, the transcriptional regulation causes the enzyme concentration to increase in response to a rising hepatic free cholesterol concentration (55), the dependency of the rate of reaction 18 on the substrate concentration could apparently become first order and 2) localization of this enzyme: the hepatic free cholesterol concentration that figures in the model as the substrate concentration for reaction 18 is not necessarily equal to the true CYP7A1 substrate concentration in the cytosolic compartment where the enzyme is active (56). If the transport of cholesterol from a hepatic storage pool to the cytosolic compartment is a rate controlling step, the cytosolic cholesterol concentration would remain below the K_M of CYP7A1. Under this condition and if the cholesterol transport to the cytosol is proportional to the total hepatic cholesterol concentration, the dependency of reaction 18 on the total hepatic cholesterol concentration

could apparently become first order.

Analysis of the features of the appropriate submodels that best describe the *in vivo* situation, can generate hypotheses for a mechanistic explanation of the cholesterol phenotype of the knockout models used. An example is the plasma cholesterol concentration in the *Abcb4*^{-/-} mouse. The phospholipid transporter *Abcb4* is required to transport cholesterol into bile. Knocking out *Abcb4* results in a lower cholesterol flux into bile. It could be expected that this would result in a higher hepatic cholesterol concentration (as has been experimentally observed (57)) and a lower intestinal cholesterol concentration than in the wild type, which in turn could lead to a lower fecal cholesterol secretion and higher cholesterol concentrations in the body and plasma. The latter is, however, the opposite of what is observed *in vivo* (58): the total plasma cholesterol concentration in the *Abcb4*^{-/-} mouse is decreased compared to wild type controls. Our model provides a mechanistic hypothesis that can explain this observation. If the transport of cholesterol to HDL is more sensitive to the intestinal cholesterol concentration than to the hepatic cholesterol concentration, as suggested by our results (Reaction 17 is zero-order and reaction 16 is first order in 6 of the 8 models), plasma HDL-C concentrations will be reduced rather than increased upon knockout of the *Abcb4* gene, as predicted by the model.

It seems likely that the model can provide predictions on the effects of drug interventions and SNPs in a similar way as performed for knockout mutations in the present paper. The intervention or SNP can be incorporated in the model by multiplying the rate constant of an affected reaction (e.g. hepatic cholesterol synthesis for statins) by a factor that reflects the efficacy of the drug or the effect of the mutation, i.e. similar to the f_{ko} used in this paper. Incorporating the influence of environmental factors other than drugs, like diet, is more complex and requires further adaptations in the model which is beyond the scope of the present study.

In summary, a quantitative model is presented that is able to accurately predict the effect of knocking out genes that act in important steps in cholesterol metabolism and transport on plasma cholesterol concentrations. The model provided a mechanistic explanation for the hitherto unexplained cholesterol phenotype of the *Abcb4*^{-/-} mouse.

References

1. Tabas, I. (2002) *J. Clin. Invest* 110, 583-590
2. Shoulders, C. C., Jones, E. L., and Naoumova, R. P. (2004) *Hum. Mol. Genet.* 13 Spec No 1, R149-R160
3. Lloyd-Jones, D. M., Larson, M. G., Beiser, A., and Levy, D. (1999) *Lancet* 353, 89-92
4. Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M., and Gonzalez, F. J. (2001) *Mol. Cell Biol.* 21, 1393-1403
5. Engelking, L. J., Liang, G., Hammer, R. E., Takaishi, K., Kuriyama, H., Evers, B. M., Li, W. P. and others (2005) *J. Clin. Invest* 115, 2489-2498
6. Ikonen, E. (2008) *Nat. Rev. Mol. Cell Biol.* 9, 125-138
7. Rawson, R. B. (2003) *Nat. Rev. Mol. Cell Biol.* 4, 631-640
8. Knoblauch, H., Schuster, H., Luft, F. C., and Reich, J. (2000) *J. Mol. Med.* 78, 507-515
9. Adiels, M., Packard, C., Caslake, M. J., Stewart, P., Soro, A., Westerbacka, J., Wennberg, B. and others (2005) *J. Lipid Res.* 46, 58-67
10. Gaffney, D., Forster, L., Caslake, M. J., Bedford, D., Stewart, J. P., Stewart, G., Wieringa, G. and others (2002) *Atherosclerosis* 162, 33-43
11. Reddy, M. B., Yang, R. S. H., Clewell, H. J., and Andersen, M. E. (2005) *Physiologically Based Pharmacokinetic Modeling*, John Wiley & Sons, Inc., Hoboken, NJ, USA
12. Andersen, M. E. (2003) *Toxicol. Lett.* 138, 9-27
13. van Waveren, R. H., Groot, S., Scholten, H., van Geer, F. C., Wösten, J. H. M., Koeze, R. D., and Noort, J. J. (1999) *Good modeling practice handbook*. STOWA-rapport 99-05
14. Potter, L. K. and Tobin, F. L. (2007) *J. Recept. Signal. Transduct. Res.* 27, 1-25
15. van de Pas, N. C., Soffers, A. E., Freidig, A. P., van Ommen B., Woutersen, R. A., Rietjens, I. M., and de Graaf, A. A. (2010) *Biochim. Biophys. Acta* 1801, 646-654
16. Zadelaar, S., Kleemann, R., Verschuren, L., de Vries-van der Weij, J., van der Hoorn, J., Princen, H. M., and Kooistra, T. (2007) *Arterioscler. Thromb. Vasc. Biol.* 27, 1706-1721
17. Siperstein, M. D. and Fagan, V. M. (1958) *J. Clin. Invest* 37, 1185-1195
18. Kuepfer, L., Peter, M., Sauer, U., and Stelling, J. (2007) *Nat. Biotechnol.* 25, 1001-

1006

19. Xie, C., Turley, S. D., and Dietschy, J. M. (2009) *J. Lipid Res.* 50, 1316-1329
20. Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., and Beliles, R. P. (1997) *Toxicol. Ind. Health* 13, 407-484
21. Riches, A. C., Sharp, J. G., Thomas, D. B., and Smith, S. V. (1973) *J. Physiol* 228, 279-284
22. Osono, Y., Woollett, L. A., Herz, J., and Dietschy, J. M. (1995) *J. Clin. Invest* 95, 1124-1132
23. Voshol, P. J. (2000) Physiological functions of biliary lipid secretion
24. Werner, A., Havinga, R., Perton, F., Kuipers, F., and Verkade, H. J. (2006) *Am. J. Physiol Gastrointest. Liver Physiol* 290, G1177-G1185
25. Thuahnai, S. T., Lund-Katz, S., Williams, D. L., and Phillips, M. C. (2001) *J. Biol. Chem.* 276, 43801-43808
26. van der Velde, A., Vrans, C. L., van den Oever, K., Kunne, C., Oude Elferink, R. P., Kuipers, F., and Groen, A. K. (2007) *Gastroenterology* 133, 967-975
27. Gautier, T., Tietge, U. J., Boverhof, R., Perton, F. G., Le Guern, N., Masson, D., Rensen, P. C. and others (2007) *J. Lipid Res.* 48, 30-40
28. Beck, C. (2008) Assembly and Secretion of Atherogenic Lipoproteins. University of Gothenborg
29. Klett, E. L., Lu, K., Kusters, A., Vink, E., Lee, M. H., Altenburg, M., Shefer, S. and others (2004) *BMC. Med.* 2, 5
30. Wechsler, A., Brafman, A., Shafir, M., Heverin, M., Gottlieb, H., Damari, G., Gozlan-Kelner, S. and others (2003) *Science* 302, 2087
31. Plump, A. S., Azrolan, N., Odaka, H., Wu, L., Jiang, X., Tall, A., Eisenberg, S. and others (1997) *J. Lipid Res.* 38, 1033-1047
32. Buhman, K. K., Accad, M., Novak, S., Choi, R. S., Wong, J. S., Hamilton, R. L., Turley, S. and others (2000) *Nat. Med.* 6, 1341-1347
33. van Ree, J. H., van den Broek, W. J., Dahlmans, V. E., Groot, P. H., Vidgeon-Hart, M., Frants, R. R., Wieringa, B. and others (1994) *Atherosclerosis* 111, 25-37
34. Brunham, L. R., Kruit, J. K., Iqbal, J., Fievet, C., Timmins, J. M., Pape, T. D., Coburn,

- B. A. and others (2006) *J. Clin. Invest* 116, 1052-1062
35. Werner, A., Minich, D. M., Havinga, R., Bloks, V., Van Goor, H., Kuipers, F., and Verkade, H. J. (2002) *Am. J. Physiol Gastrointest. Liver Physiol* 283, G900-G908
36. Sakai, N., Vaisman, B. L., Koch, C. A., Hoyt, R. F., Jr., Meyn, S. M., Talley, G. D., Paiz, J. A. and others (1997) *J. Biol. Chem.* 272, 7506-7510
37. Davies, J. P., Scott, C., Oishi, K., Liapis, A., and Ioannou, Y. A. (2005) *J. Biol. Chem.* 280, 12710-12720
38. Plosch, T., Bloks, V. W., Terasawa, Y., Berdy, S., Siegler, K., Van Der Sluijs, F., Kema, I. P. and others (2004) *Gastroenterology* 126, 290-300
39. Varban, M. L., Rinninger, F., Wang, N., Fairchild-Huntress, V., Dunmore, J. H., Fang, Q., Gosselin, M. L. and others (1998) *Proc. Natl. Acad. Sci. U. S. A* 95, 4619-4624
40. Timmins, J. M., Lee, J. Y., Boudyguina, E., Kluckman, K. D., Brunham, L. R., Mulya, A., Gebre, A. K. and others (2005) *J. Clin. Invest* 115, 1333-1342
41. Ishibashi, S., Perrey, S., Chen, Z., Osuga, J., Shimada, M., Ohashi, K., Harada, K. and others (1996) *J. Biol. Chem.* 271, 22422-22427
42. Altmann, S. W., Davis, H. R., Jr., Zhu, L. J., Yao, X., Hoos, L. M., Tetzloff, G., Iyer, S. P. and others (2004) *Science* 303, 1201-1204
43. Wang, X., Collins, H. L., Ranalletta, M., Fuki, I. V., Billheimer, J. T., Rothblat, G. H., Tall, A. R. and others (2007) *J. Clin. Invest* 117, 2216-2224
44. Plosch, T., van der Veen, J. N., Havinga, R., Huijkman, N. C., Bloks, V. W., and Kuipers, F. (2006) *Am. J. Physiol Gastrointest. Liver Physiol* 291, G414-G423
45. Smit, J. J., Schinkel, A. H., Oude Elferink, R. P., Groen, A. K., Wagenaar, E., van Deemter, L., Mol, C. A. and others (1993) *Cell* 75, 451-462
46. Brundert, M., Ewert, A., Heeren, J., Behrendt, B., Ramakrishnan, R., Greten, H., Merkel, M. and others (2005) *Arterioscler. Thromb. Vasc. Biol.* 25, 143-148
47. Punt, A., Paini, A., Boersma, M. G., Freidig, A. P., Delatour, T., Scholz, G., Schilter, B. and others (2009) *Toxicol. Sci.* 110, 255-269
48. Hissink, A. M., van Ommen, B., Kruse, J., and van Bladeren, P. J. (1997) *Toxicol. Appl. Pharmacol.* 145, 301-310
49. Lupfert, C. and Reichel, A. (2005) *Chem. Biodivers.* 2, 1462-1486

50. Quick, D. J. and Shuler, M. L. (1999) *Biotechnol. Prog.* 15, 540-555
51. Guo, Y., Zhang, C., Du, X., Nair, U., and Yoo, T. J. (2005) *Hear. Res.* 208, 54-67
52. Ishida, T., Choi, S. Y., Kundu, R. K., Spin, J., Yamashita, T., Hirata, K., Kojima, Y. and others (2004) *J. Biol. Chem.* 279, 45085-45092
53. Collet, X. and Fielding, C. J. (1991) *Biochemistry* 30, 3228-3234
54. Ozasa, S. and Boyd, G. S. (1981) *Eur. J. Biochem.* 119, 263-272
55. Kalaany, N. Y. and Mangelsdorf, D. J. (2006) *Annu. Rev. Physiol* 68, 159-191
56. Chiang, J. Y. (2004) *J. Hepatol.* 40, 539-551
57. Plosch, T., Bloks, V. W., Baller, J. F., Havinga, R., Verkade, H. J., Jansen, P. L., and Kuipers, F. (2002) *Hepatology* 35, 299-306
58. Voshol, P. J., Schwarz, M., Rigotti, A., Krieger, M., Groen, A. K., and Kuipers, F. (2001) *Biochem. J.* 356, 317-325



CHAPTER

4

A physiologically based in silico kinetic model predicting plasma cholesterol concentrations in humans.

Niek C.A. van de Pas, Ruud A. Woutersen, Ben van Ommen, Ivonne M.C.M. Rietjens, and Albert A. de Graaf

Supplementary material can be found at www.vdpastreffers.nl/cholesterol

Submitted



Abstract

Increased plasma cholesterol concentration is associated with increased risk of cardiovascular disease. This study describes the development, validation, and analysis of a physiologically based kinetic (PBK) model for the prediction of cholesterol plasma concentrations in humans. This model was adapted from a PBK model for mice, by incorporation of the reaction catalyzed by CETP, and contained 21 biochemical reactions and 8 different cholesterol pools.

The model was calibrated using data from published kinetic studies and validated by comparing model predictions on plasma cholesterol concentrations of subjects with ten different genetic mutations (including Familial Hypercholesterolemia, and Smith-Lemli-Opitz syndrome) with experimental data on these concentrations.

Average model predictions on TC were accurate within 36% of the experimental data, within the experimental margin. Sensitivity analysis of parameters in the model indicated that the High Density Lipoprotein cholesterol (HDL-C) concentration was mainly dependent on hepatic transport of cholesterol to HDL, CE transfer from HDL to non-HDL, and hepatic uptake of cholesterol from non-HDL-C.

Thus, the presented PBK model is a valid tool to predict the effect of genetic mutations on cholesterol concentrations, opening the way for future studies on the effect of different drugs on cholesterol concentrations in various subpopulations *in silico*.

Introduction

In silico modeling has proven to be a useful tool in biology as it allows to study interspecies variation, to study the regulation of homeostasis, and to integrate information from various sources (1-3). There are several modeling efforts on cholesterol (4-7), an important biomarker for the risk for cardiovascular events (8-10). Most of these cholesterol modeling studies present models that focus on LDL cholesterol (LDL-C) metabolism in plasma (4-6), or on cellular cholesterol metabolism (7). These models thus do not represent all relevant components of whole body cholesterol homeostasis, because they lack reactions like cholesterol absorption and biosynthesis in organs, which are the reactions targeted by important cholesterol-lowering drugs, such as statins. This implies that the models cannot fully explain how relevant plasma cholesterol-associated biomarkers are influenced by these drug interventions.

We have, therefore, developed an *in silico* Physiologically Based Kinetic (PBK) model for plasma cholesterol in the mouse that includes all relevant reactions and that correctly predicts the plasma cholesterol concentrations of a large variety of mouse strains with gene knockouts related to cholesterol metabolism (11). The model, of which the structure is given in Figure 4.1, is able to predict HDL cholesterol (HDL-C), LDL-C (also denoted as non-HDL-C), and total plasma cholesterol (TC) concentrations (12), as well as the intra-organ pools representing hepatic free cholesterol (Liv-FC), peripheral cholesterol (Per-C), intestinal cholesterol ester (Int-CE), hepatic cholesterol ester (Liv-CE), and intestinal free cholesterol (Int-FC) in the mouse. A similar model for humans will be of considerable value in predicting effects of drugs and genetic variations on plasma cholesterol concentrations in humans. Therefore, the aim of this work is to adapt our model to a human version.

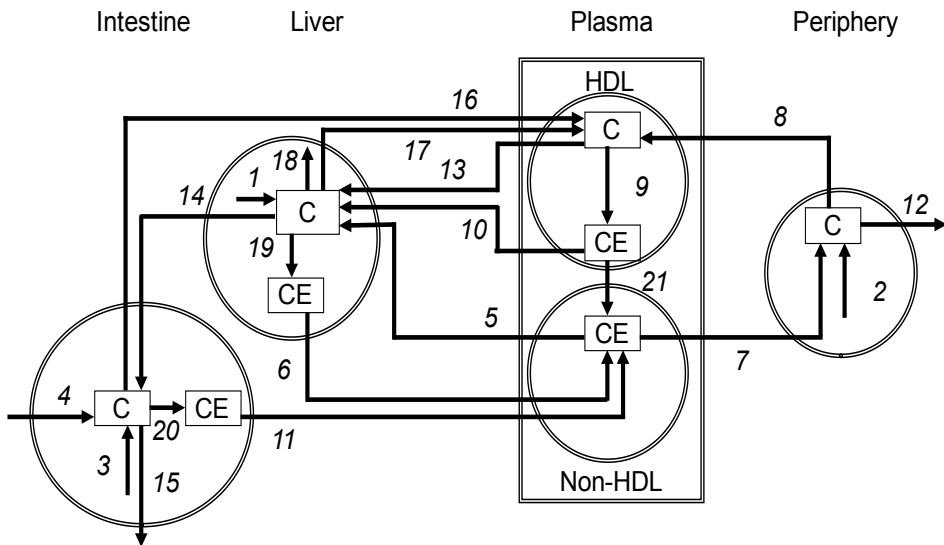


Figure 4.1. Conceptual model for pathways determining cholesterol plasma concentrations used as a basis to set up the *in silico* model of the present study. Process numbers stand for: 1, Hepatic cholesterol synthesis; 2, Peripheral cholesterol synthesis; 3, Intestinal cholesterol synthesis; 4, Dietary cholesterol intake; 5, Hepatic uptake of cholesterol from LDL; 6, Hepatic Very Low Density lipoprotein cholesterol (VLDL-C) secretion; 7, Peripheral uptake of cholesterol from LDL; 8, Peripheral cholesterol transport to HDL; 9, HDL-associated cholesterol esterification; 10, Hepatic HDL-CE uptake; 11, Intestinal chylomicron cholesterol secretion; 12, Peripheral cholesterol loss; 13, Hepatic HDL-FC uptake; 14, Biliary cholesterol excretion; 15, Fecal cholesterol excretion; 16, Intestinal cholesterol transport to HDL; 17, Hepatic cholesterol transport to HDL; 18, Hepatic cholesterol catabolism; 19, Hepatic cholesterol esterification; 20, Intestinal cholesterol esterification and 21, CE transfer from HDL to LDL. C stands for cholesterol; CE for cholesterol ester. Based on (11).

Apart from this qualitative difference, there are many quantitative differences between mouse and human with respect to parameters that influence cholesterol turnover, such as dietary intake, transport and synthesis rates of cholesterol, as well as organ sizes (13,19). Some of these differences persist even after correcting for differences in body mass. For example, according to Dietschy and Turley (13), the amount of cholesterol absorbed from the diet per kg body mass is 30 mg cholesterol /kg body mass /day in the mouse compared to only 5 mg cholesterol / kg body mass /day for human. Consequently, translational modifying the available mouse model for

human subjects is not straightforward. This study describes this translation, as well as a validation of the resulting model, by simulating human mutations and comparing the model predictions on HDL-C and non-HDL-C in plasma concentrations with experimental data.

Methods

The development of the *in silico* model was subdivided into the following steps: development of a conceptual model (model structure), mathematical formulation of the model, model calibration, and model validation. After development, a sensitivity analysis was performed to obtain more insight into the factors that determine plasma cholesterol concentrations.

Conceptual model development

92

The conceptual model for the human PBK model was modified from the conceptual model for the mouse, previously developed (11). Briefly, the mouse model has been constructed as follows: relevant knockout mouse models were screened for altered plasma cholesterol concentrations compared to the concentrations in the wild type mouse. If the alteration was more than two-fold (up or down) compared to the wild type, the corresponding gene was marked as key gene. Based on the function of a subset of 12 of these key genes that code for metabolic enzymes producing or consuming cholesterol and transport protein transporting cholesterol, metabolic and transport reactions were included in the mouse conceptual model (11). The conceptual model for the human was developed from the mouse model by adding human specific features (see Results section).

Mathematical model formulation

As the second step in the model development, the conceptual model was converted into mathematical equations. Similar to the mouse model, the human model was formulated as a set of differential equations each describing the time behavior of one of the cholesterol pools in the conceptual model as a function of the reaction rates.

Equations were slightly altered compared to the mouse model: for the human model, reaction rates were expressed as mmol/man/day, leading to more simple equations (12). For practical reasons, reaction rates (ex-

pressed by v), were numbered according to the numbering in Figure 4.1. All symbols of variables and parameters that were used to define the model are given in Table 4.1.

Table 4.1. List of symbols and parameters used.

| Symbol | Description | Unit |
|----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| V_i | volume of compartment i with $i = \text{Liv, Pla, Per, and Int}$ for liver, plasma, periphery, and intestine. | L/man |
| v_i | reaction rate of reaction i | mmol/(man·day) |
| v_i^{ss} | steady state reaction rate of reaction i | mmol/(man·day) |
| $[C]_i$ | tissue or plasma concentration of pool i , with $i = \text{Liv-FC, HDL-FC, HDL-CE, non-HDL-C, Per-C, Int-FC, Liv-CE and Int-CE}$ for liver free cholesterol, HDL free cholesterol, HDL cholesterol ester, non-HDL cholesterol, peripheral cholesterol, intestinal free cholesterol, liver cholesterol ester, and intestinal cholesterol ester. | mmol/L |
| $[C]_i^{ss}$ | steady state concentration of pool i | mmol/L |
| k_i^z | zero-order rate constant of reaction i | mmol/(L·day) |
| k_i^f | first-order rate constant of reaction i | 1 /day |
| k_i^s | second-order rate constant of reaction i | L / (mmol·day) |
| $k_i^{\text{mut},j}$ | j^{th} -order rate constant of reaction i in a human carrying a mutation that affects reaction i . | |
| $k_i^{*,k}$ | k^{th} -order reaction rate constant of reaction i that has been altered compared to the normal values | |
| $[C]_i^{*,ss}$ | steady state concentration of pool i , in case the reaction rate constant of reaction i that has been altered compared to the normal values | mmol/L |
| t | time | day |
| f_{mut} | rate constant modification factor | [-] |

The differential equations, one for each cholesterol pool in the model, are given in Table 4.2 (Eqn. 4.1-4.8). Equation 1, for example, can be interpreted as: the change in time of the concentration of hepatic free cholesterol is determined by the balance of the rates of the reactions producing (v_1, v_5, v_{10}, v_{13}) and consuming ($v_{14}, v_{17}, v_{18}, v_{19}$) hepatic free cholesterol.

Table 4.2. Equations used in model development (Eqns 4.1-4.11), model calibration (Eqns 4.12-4.14), for simulation of the effect of human mutations (Eqns 4.15-4.17), and sensitivity analysis (Eqn. 4.18).

| | |
|--------------------------------------------------------------------------------------------------------------|-----------|
| Differential Equations | |
| $\frac{d(V_{Liv} \cdot [C]_{Liv-FC})}{dt} = v_1 + v_5 + v_{10} + v_{13} - v_{19} - v_{14} - v_{17} - v_{18}$ | Eqn. 4.1 |
| $\frac{d(V_{Pla} \cdot [C]_{HDL-FC})}{dt} = v_8 + v_{16} + v_{17} - v_9 - v_{13}$ | Eqn. 4.2 |
| $\frac{d(V_{Pla} \cdot [C]_{HDL-CE})}{dt} = v_9 - v_{10} - v_{21}$ | Eqn. 4.3 |
| $\frac{d(V_{Pla} \cdot [C]_{non\ HDL-C})}{dt} = v_6 + v_{11} - v_5 - v_7 + v_{21}$ | Eqn. 4.4 |
| $\frac{d(V_{Per} \cdot [C]_{Per-C})}{dt} = v_2 + v_7 - v_8 - v_{12}$ | Eqn. 4.5 |
| $\frac{d(V_{Int} \cdot [C]_{Int-FC})}{dt} = v_3 + v_4 + v_{14} - v_{20} - v_{16} - v_{15}$ | Eqn. 4.6 |
| $\frac{d(V_{Liv} \cdot [C]_{Liv-CE})}{dt} = v_{19} - v_6$ | Eqn. 4.7 |
| $\frac{d(V_{Int} \cdot [C]_{Int-CE})}{dt} = v_{20} - v_{11}$ | Eqn. 4.8 |
| Rate equations | |
| $v_i = k_i^z \cdot V$ (zero-order kinetics) | Eqn. 4.9 |
| $v_i = k_i^f \cdot V \cdot [C]$ (first-order kinetics) | Eqn. 4.10 |
| $v_{21} = k_{21}^s \cdot [C]_{HDL-CE} \cdot [C]_{non\ HDL-C} \cdot V$ (second-order kinetics) | Eqn. 4.11 |

| | |
|-------------------------------------------------------------------------------------------------------------|-----------|
| Model calibration | |
| $k_i^z = v_i^s / V$ (zero-order kinetics) | Eqn. 4.12 |
| $k_i^f = \frac{v_i^s}{[C]_i^s \cdot V}$ (first-order kinetics) | Eqn. 4.13 |
| $k_{21}^s = \frac{v_{21}^s}{[C]_{HDL-CE}^s \cdot [C]_{non\ HDL-C}^s \cdot V_{pla}}$ (second-order kinetics) | Eqn. 4.14 |
| Mutation simulation | |
| $k_i^{mut,z} = f_{mut} \cdot k_i^z$ (zero-order kinetics) | Eqn. 4.15 |
| $k_i^{mut,f} = f_{mut} \cdot k_i^f$ (first-order kinetics) | Eqn. 4.16 |
| $k_{21}^{mut,s} = f_{mut} \cdot k_{21}^s$ (second-order kinetics) | Eqn. 4.17 |
| Sensitivity analysis | |
| $SC_{ij} = \frac{[C]_j^{*s} - [C]_j^s}{[C]_j^s} \frac{k_i^k}{k_i^k - k_i^{*k}}$ | Eqn. 4.18 |

In a kinetic model, the reaction rates are calculated using kinetic equations that express the reaction rate as a function of concentrations and kinetic parameters. Most reactions in our model represent a set of lumped enzymatic- and/or transport reactions. Often, such a composite reaction can be conveniently represented by a kinetic equation that contains apparent rate constants and apparent K_M values. A general solution is to use Michaelis-Menten kinetics as a prototype kinetic expression for biological reactions. In the present approach, in order to keep the number of parameters as limited as possible, it was additionally assumed that the reactions operate either in the linear part of a Michaelis-Menten kinetic curve (substrate concentration much lower than the apparent K_M), or in the saturated part (substrate concentration much higher than the apparent K_M). At low substrate concentrations Michaelis-Menten kinetics effectively reduce to first order kinetics (Eqn. 4.9, Table 4.2). At high substrate concentrations, Michaelis-Menten kinetics reduce to zero-order kinetics, i.e. the reaction becomes independent of the substrate concentration (Eqn. 4.10, Table 4.2).

Reaction 21, catalyzed by CETP, was assumed to have different

kinetics than the other reactions. CETP transfers CE from HDL to LDL and it requires the binding of both a donor (HDL) and acceptor particle (non-HDL) to exchange cholesterol (20,21). Therefore, the rate of this reaction was dependent on both HDL-CE and non-HDL-CE. It was assumed that the reaction rate was linear with respect to both concentrations as given by equation 11 (Table 4.2).

The kinetic formats of reactions 1 - 20 were obtained from the previously defined mouse model. Instead of developing one optimal model, the mouse model consists of a set (ensemble) of 8 submodels, each having a different combination of first and zero order kinetics for the various reactions. The model prediction is calculated as the average of the predictions of the submodels. These 8 submodels have been selected from a larger set of 65,536 alternative submodels. Each of the suitable submodels has been selected on the basis of a correct prediction of a higher or lower plasma cholesterol concentration of 5 knockout mouse strains compared to the wild type controls (12). For the human situation, not enough data were available to apply an identical selection procedure. Therefore, it was decided to use the kinetic formats of the reactions 1 - 20 in the 8 submodels for the mouse also for the corresponding reactions in the human model. Thus, the 8 mouse submodels were converted to 8 human submodels, retaining the kinetic orders of reactions 1 - 20 and adding reaction 21. As in the mouse model, the human model prediction was defined as the average of the predictions of the resulting 8 submodels. The predicted TC concentration was calculated as the sum of all three types of plasma cholesterol (non-HDL-C, HDL-FC, and HDL-CE), whereas the predicted HDL-C concentration was calculated as the sum of HDL-FC and HDL-CE.

Model calibration

The model was calibrated by assigning values to the kinetic parameters (k_i^z , k_i^f , and k_i^s). The values of the parameters were calculated from steady state concentrations of cholesterol pools and reaction rates of the reactions obtained from literature. Literature data on reaction rates and pool sizes were taken from a wide range of experiments including cannulations, dietary surveys, and *in vitro* tests as explained in detail below.

As usual in PBK modeling (3,19), the model was developed for a standard human, for which we chose the 70 kg “reference man” as defined by the International Commission on Radiation Protection.

Data on 70 kg adult male subjects (19) were taken as much as possible. In case the data were obtained from people with different body mass-

es, reaction rates were normalized per unit of organ volume and after that multiplied by the organ mass of the reference man to obtain the reaction rate for the standard human.

In case that in a specific submodel a rate was calculated for a first order reaction, the corresponding rate constant was calculated using the pool sizes and steady state reaction rates according to Eqn. 4.12 (Table 4.2). In the case of a zero-order reaction, the corresponding rate constant was calculated according to Eqn. 4.13 (Table 4.2). To calculate the rate constant of reaction 21, Eqn. 4.14 was used (Table 4.2).

Model validation

To validate the model, the model predicted the plasma cholesterol concentrations of humans carrying genetic mutations and these predictions were compared with experimental data. A genetic mutation was simulated by a model run with a case-specific set of parameters different from the normal situation, as follows. The rate constants for the reaction(s) primarily affected by the given mutation was/were multiplied with a specific parameter (f_{mut}), according to equations 15, 16, or 17 (Table 4.2). This multiplication reflects the impact of the mutation on the reaction rate constant (i.e., -fold reduction or increase). All other parameters were assumed to be unaffected by the given mutation. The values for f_{mut} for each individual mutation were defined based on literature data as described in the Results section.

Model predictions were derived as follows: the differential equations (Eqn. 4.1-8), were solved by numerical integration with the normal concentrations as initial values. Integration was performed using routine *ode15s* as implemented in MATLAB version 7.5 (R2007b) with the appropriate parameter value(s) for each subject (normal or mutant). The simulation was performed until steady state of all cholesterol pools in the model. Model predictions were defined as these steady state concentrations obtained.

Sensitivity analysis

To identify which reactions had a large influence on the 8 predicted cholesterol pools in the model, a sensitivity analysis was performed. One by one, all kinetic constants were increased by one percent (leaving all other parameters unchanged) and the model was used to predict the effect of this increase on all 8 cholesterol pools that figured in the model. The effect of the increase in parameter of rate i on pool j was expressed in a sensitivity coefficient (SC) defined in Eqn. 4.18. The SC of the model was defined as the average of the SC's calculated with the 8 submodels.

Results

Conceptual model development

Figure 4.1 presents the conceptual model for human plasma cholesterol concentrations. The model contains 8 pools and 21 reactions, for example hepatic cholesterol synthesis (reaction 1), biliary cholesterol excretion (reaction 14), and fecal cholesterol excretion (reaction 15).

Model development and calibration

The human PBK model formulated as differential equations (Eqn. 4.1-4.8, Table 4.2) and rate equations (Eqn. 4.9-4.11, Table 4.2), based on the conceptual model given in Figure 4.1. The model was calibrated using compartmental volumes, steady state cholesterol concentrations, and rates of cholesterol-involving reactions derived from literature. A detailed description of how data were derived, transformed into the correct units, and scaled to the 70 kg “reference man” as defined by the International Commission on Radiation Protection (19) can be found in the supplementary material. Concerning plasma cholesterol, the total plasma cholesterol concentration was 5.25 mM for TC and 1.19 mM for HDL-C as was obtained from an inventory of data from 8809 US adults (22). Plasma cholesterol not present in the HDL-C pool (4.03 mM) was considered to be present in the non-HDL-C pool. Total HDL-C concentration (HDL-C, 1.19 mM) consists of HDL-FC and HDL-CE. The HDL-FC:HDL-CE ratio was 1:3 as obtained from (23). This ratio was applied to the HDL-C data above to obtain the HDL-FC and HDL-CE concentration, resulting in 0.30 mM for HDL-FC and 0.89 mM for HDL-CE.

A summary of these results is presented in Table 4.3. Several steady state reaction rates were not directly obtained from data, but instead were calculated from the other reaction rates using mass balances as indicated in Table 4.3 for v_{5}^{ss} , v_{7}^{ss} , v_{8}^{ss} , v_{10-12}^{ss} , v_{17}^{ss} , and v_{19-20}^{ss} .

The model predicted a steady state that matched all the data in Table 4.3, which indicates that the data, taken from various sources were mutually consistent.

Table 4.3. Numerical values for the compartmental volumes, steady state concentrations, and steady state reaction rates for a 70 kg reference man. More details can be found in the text and in supplementary material.

| Organ | Symbol | Value | Obtained from |
|-----------------------------------------------------------|-------------------------|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Volumes (L/man) | | | |
| Liver | V_{liv} | 1.80 | Calculated assuming liver mass is 2.57% of total body mass and organ density is 1 kg/L (19). |
| Intestine | V_{int} | 0.64 | Calculated assuming intestinal mass is 0.91% of total body mass and organ density is 1 kg/L (19). |
| Plasma | V_{Pla} | 2.79 | Calculated assuming plasma volume is 39.9 ml/kg body mass (44). |
| Periphery | V_{Per} | 64.8 | Calculated by subtracting liver, intestinal and plasma mass from the 70 kg total body mass and assuming density is 1 kg/L. |
| Steady state concentrations (mmol per L tissue or plasma) | | | |
| Hepatic free C | $[C]_{Liv-FC}^{ss}$ | 8.00 | Measured using needle liver biopsies (45). |
| HDL free C | $[C]_{HDL-FC}^{ss}$ | 0.30 | See text (22,23). |
| HDL C ester | $[C]_{HDL-CE}^{ss}$ | 0.89 | See text (22,23). |
| Non-HDL C | $[C]_{non_HDL-C}^{ss}$ | 4.03 | See text (22). |
| Peripheral C | $[C]_{Per-C}^{ss}$ | 1.20 | Measured in biopsies from large organs in sudden death individuals (46). |
| Intestinal free C | $[C]_{Int-FC}^{ss}$ | 1.99 | Determined from biopsies from sudden death individuals. (47). Ratio Int-FC:Int-CE was assumed to be equal to that ratio in CaCo2 cells (48). |
| Intestinal CE | $[C]_{Int-CE}^{ss}$ | 0.25 | Determined from biopsies from sudden death individuals. (47). Ratio Int-FC:Int-CE was assumed to be equal to that ratio in CaCo2 cells (48). |
| Hepatic CE | $[C]_{Liv-CE}^{ss}$ | 5.30 | Measured using needle liver biopsies (45). |
| Steady state reaction rates (mmol/(man•day)) | | | |
| Hepatic C synthesis | v_1^{ss} | 0.44 | Determined with <i>ex vivo</i> studies on liver biopsies (13,49,50). |
| Peripheral C synthesis | v_2^{ss} | 3.79 | Calculated from total body (determined from squalene kinetics), hepatic, and intestinal cholesterol synthesis rates (49). |
| Intestinal C synthesis | v_3^{ss} | 0.18 | Determined with <i>ex vivo</i> studies on intestinal biopsies (51). |
| Dietary C intake | v_4^{ss} | 1.09 | Calculated using 7 day food recall (52). |
| Hepatic uptake of C from non-HDL | v_5^{ss} | 11.42 | Sum of the uptake rates of non-HDL-C (reactions 5 and reaction 7) was calculated with the LDL-C balance: $v_5^{ss} + v_7^{ss} = v_{21}^{ss} + v_6^{ss} + v_{11}^{ss}$. The ratio between hepatic (v_5^{ss}) and extrahepatic uptake (v_7^{ss}) in the human was assumed to be identical to that ratio in the mouse (53). |
| Hepatic VLDL-C secretion | v_6^{ss} | 4.76 | Calculated from stable isotope study (54) and lipoprotein composition data (55). |
| Peripheral uptake of C from non-HDL | v_7^{ss} | 1.31 | See v_5^{ss} |

| Organ | Symbol | Value | Obtained from |
|------------------------------------|---------------|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Peripheral C transport to HDL | v_8^{ss} | 2.48 | Calculated with the peripheral cholesterol balance ($v_8^{ss} = v_2^{ss} + v_7^{ss} - v_{12}^{ss}$). |
| HDL-associated C esterification | v_9^{ss} | 7.86 | Determined with <i>ex vivo</i> test with endogenous substrates (56,57). |
| Hepatic HDL-CE uptake | v_{10}^{ss} | 2.93 | Calculated with the HDL-CE balance ($v_{10}^{ss} = v_9^{ss} - v_{21}^{ss}$) |
| Intestinal chylomicron-C secretion | v_{11}^{ss} | 3.04 | Calculated with the intestinal cholesterol ester balance ($v_{11}^{ss} = v_{20}^{ss}$) |
| Peripheral C loss | v_{12}^{ss} | 2.62 | Calculated with the total body balance $v_{12}^{ss} = v_1^{ss} + v_2^{ss} + v_3^{ss} + v_4^{ss} - v_{18}^{ss} - v_{15}^{ss}$ |
| Hepatic HDL-FC uptake | v_{13}^{ss} | 1.57 | Calculated by assuming that the ratio between v_{10} and v_{13} is similar in human and mouse (12). |
| Biliary C excretion | v_{14}^{ss} | 3.66 | Measured using bile that was collected with a swallowed tube (58). |
| Fecal C excretion | v_{15}^{ss} | 1.85 | Measured by fecal collection and analysis (52). |
| Intestinal C transport to HDL | v_{16}^{ss} | 0.03 | Measured using in an <i>in vitro</i> assay with Caco2 cells (59) |
| Hepatic C transport to HDL | v_{17}^{ss} | 6.91 | Calculated with the hepatic free cholesterol balance ($v_{17}^{ss} = v_1^{ss} + v_{13}^{ss} + v_{10}^{ss} + v_5^{ss} - v_6^{ss} - v_{18}^{ss} - v_{14}^{ss}$). |
| Hepatic C catabolism | v_{18}^{ss} | 1.03 | Assumed to be equal to fecal loss of bile acids, measured by fecal collection and analysis (52). |
| Hepatic C esterification | v_{19}^{ss} | 4.76 | Calculated with the hepatic cholesterol ester balance ($v_{19}^{ss} = v_6^{ss}$) |
| Intestinal C esterification | v_{20}^{ss} | 3.04 | Calculated with the intestinal free cholesterol balance ($v_{20}^{ss} = v_3^{ss} + v_4^{ss} + v_{14}^{ss} - v_{16}^{ss} - v_{15}^{ss}$) |
| CE transfer from HDL to non-HDL | v_{21}^{ss} | 4.93 | Measured <i>ex vivo</i> with radiolabeled CE (non-HDL) (21). |

Model validation

As a model validation, 10 genetic variations known to affect cholesterol metabolism were simulated and model predictions for TC, HDL-C, and non-HDL-C were compared with experimental data. The 10 mutations include mutations that cause Familial Hypercholesterolemia (FH), Fish Eye Disease, Smith-Lemli-Opitz syndrome (SLOS), and other diseases. Details on all 10 mutations (numbered with roman digits) are given in Table 4.4 and are explained below. In the list of 10 mutations, two genes were included twice, APOB and LCAT, both as homozygote and as heterozygote variant.

Each mutation was simulated by multiplying the rate constant of the reaction affected with a specific parameter f_{mut} defined for that mutation (see Eqn. 4.15-4.17, Table 4.2). The parameter f_{mut} is generally defined as

the ratio of the value of a specific variable in carriers of the mutation to the value of that specific variable in controls. The specific affected variables for all mutations are given in Table 4.4.

Table 4.4. Description of the 10 human mutations used for model validation. The table includes the name of the gene carrying the mutation, the plasma cholesterol concentrations (TC, HDL-C, and non-HDL-C) of the subjects carrying the mutations (expressed as –fold increase relative to the control group), the number of the reaction(s) affected by specific mutations, the severity of the affection expressed as f_{mut} (Eqn. 4.15-4.17) and the name of the variable used to determine f_{mut} . Reaction numbers correspond to numbers in Figure 4.1. NA, Not Available.

| Nr | Gene | TC (Relative to control) | HDL-C (Relative to control) | non-HDL-C (Relative to control) | Reaction affected | f_{mut} | Experimentally measured variable used to determine f_{mut} (see text for details) | Ref |
|------|---------------------|--------------------------|-----------------------------|---------------------------------|-------------------|-----------|-------------------------------------------------------------------------------------|---------|
| I | LDLR ^a | 1.85 | 0.86 | 2.17 | 5,7 | 0.38 | The fractional catabolic rate of APOB in LDL | (54) |
| II | APOB ^b | 1.36 | 0.85 | 1.52 | 5,7 | 0.31 | The fractional catabolic rate of APOB in LDL | (54) |
| III | APOB ^c | 1.97 | 1.12 | 2.24 | 5,7 | 0.32 | The fractional catabolic rate of APOB in LDL | (54) |
| IV | ABCA1 ^d | 1.07 | 0.22 | 1.42 | 8, 16, 17 | 0.41 | The cholesterol efflux rate to APOA1 | (60) |
| V | APOE ^e | 2.80 | NA | NA | 5,7 | 0.45 | The residence time of the carrier or control form of APOE in a normal subject | (61) |
| VI | CETP ^f | 1.01 | 1.10 | 0.98 | 21 | 0.65 | The plasma level of CETP | (15) |
| VII | LCAT ^g | 0.81 | 0.79 | 0.97 | 9 | 0.62 | The <i>in vitro</i> determined activity of LCAT | (39) |
| VIII | LCAT ^h | 0.77 | 0.19 | 0.82 | 9 | 0 | The <i>in vitro</i> determined activity of LCAT | (39) |
| IX | DHCR7 ⁱ | 0.20 | NA | NA | 1,2,3 | 0.00 | The cholesterol synthesis activity in cultured fibroblasts | (62,63) |
| X | CYP7A1 ^j | 1.74 | 0.97 | 2.09 | 18 | 0.05 | The bile acid content of the stools | (24) |

^a Heterozygote Familial Hypercholesterolemia (FH), ^b Heterozygote familial defective APOB (FDB), ^c Homozygote familial defective APOB (FDB), ^d Hypoalphalipoproteinemia, ^e Type III hyperlipoproteinemia, ^f heterozygote in exon 15, ^g Heterozygote Familial LCAT deficiency (incl. Fish eye disease), ^h Homozygote familial LCAT deficiency (incl. Fish eye disease), ⁱ Smith-Lemli-Opitz syndrome (SLOS), and ^j Bile acid synthesis defect.

The parameter f_{mut} , for the simulation of a bile acid synthesis defect in the gene CYP7A1 (mutation X), for example, was defined as the ratio of the bile acid contents of the stools from carriers of the mutation to that in controls. The affected individuals had 5% of the amount of bile acids in their stools compared to healthy controls (24). The value of f_{mut} was, therefore, set to 0.05 (Table 4.4). Values of the parameter f_{mut} range from 0.00 for the SLOS mutation and the homozygote LCAT mutation to 0.65 for a variant of the CETP gene, implying that the list contained mutations that cause both mild and severe phenotypes. Model predictions and experimental data are given in Figure 4.2 (for TC), Figure 4.3 (HDL-C), and Figure 4.4 (non-HDL-C).

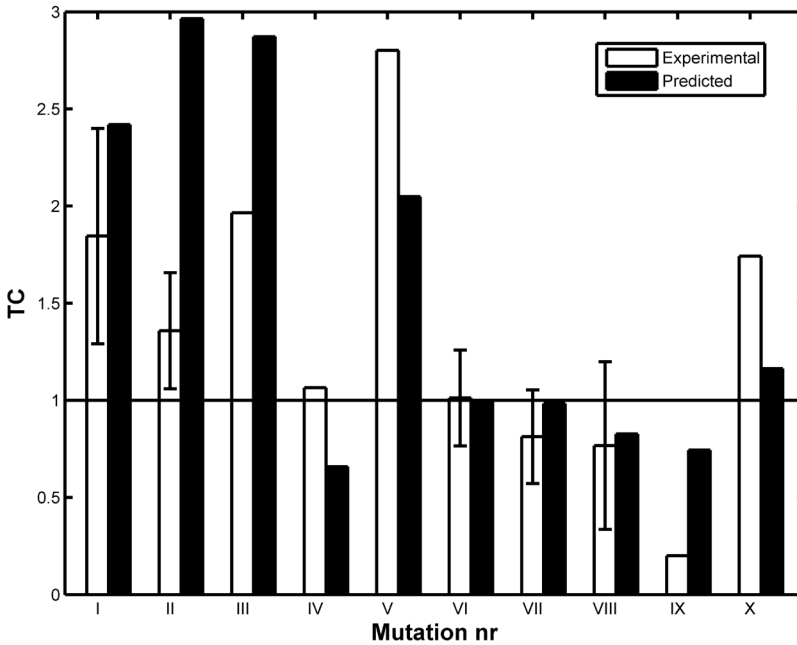


Figure 4.2. PBK model predictions compared to experimental data for TC concentrations (relative to control) for 10 different human mutations. Mutation numbers correspond to the numbers in Table 4.4.

Apart from LCAT deficiency (mutation VIII, Table 4.4), the model correctly predicted whether the TC, HDL-C, and non-HDL-C concentrations were decreased, increased, or relatively unchanged by the mutations. The average relative deviations between model predictions and experimen-

tal data were 36%, 49%, and 43% for TC, HDL-C, and non-HDL-C, respectively. This is considered successful within the present state-of-the-art of PBK modeling, where quantitative predictions may generally be correct within one order of magnitude (25-28). Also these model predictions are generally within the experimental error margin, given small patient groups sizes (generally $N < 20$).

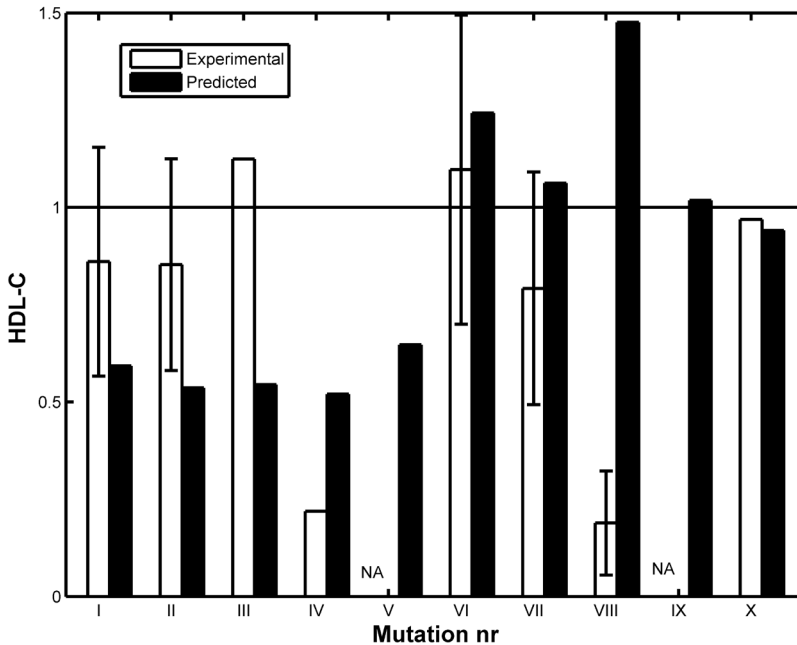
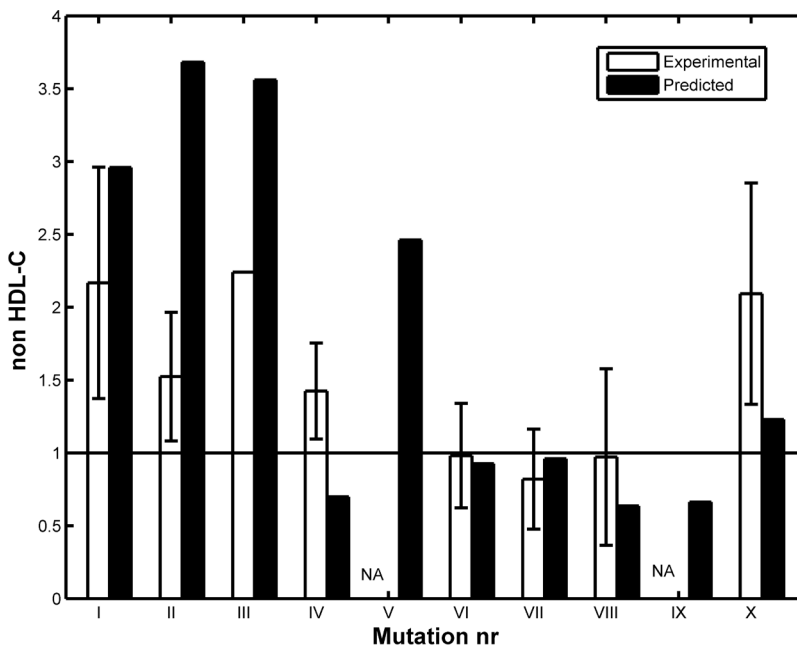


Figure 4.3. PBK model predictions compared to experimental data for HDL-C (relative to control) for 10 different human mutations. Mutation numbers correspond to the numbers in Table 4.4. NA, No data available.

Model analysis

An important step in modeling is model analysis, this is the step in which novel biological insight can be obtained. Sensitivity analysis was performed to analyze which cholesterol concentrations are affected by which biological reactions. Please note, this procedure is not linked to assessing the sensitivity of a diagnostic test. Figure 4.5 presents the sensitivity coefficients (SC, Eqn. 4.18) that express the sensitivity of the 8 concentrations in the model towards changes in the kinetic parameters of each of the 21



104

Figure 4.4. PBK model predictions compared to experimental data for non-HDL-C (relative to control) for 10 different human mutations. Mutation numbers correspond to the numbers in Table 4.4. NA, No data available.

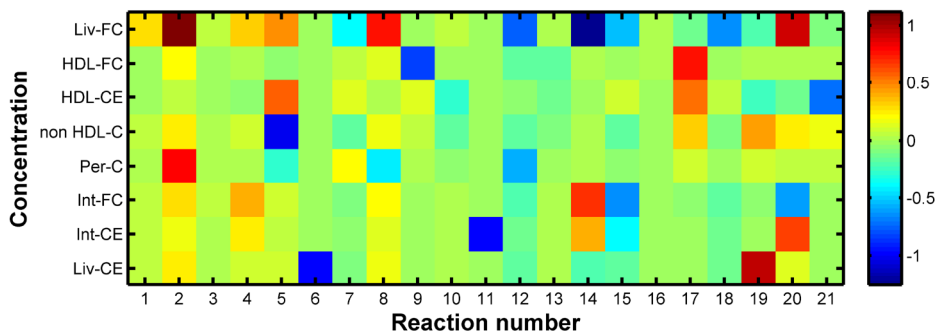


Figure 4.5. Color representation of the sensitivity coefficients that quantify the influence of changes in the rate constants of the different reactions on the different concentrations. See text for details. Reaction numbers correspond to the numbers in Figure 4.1.

reactions. A positive SC indicates that an increase in the reaction rate constant results in an increase of the predicted concentration. A negative SC indicates that an increase in the reaction rate constant results in a decrease of the predicted concentration.

Some concentrations responded to changes in many different rate constants (e.g. hepatic free cholesterol (Liv-FC)), while other concentrations were sensitive to changes in only a few reactions. HDL free cholesterol (HDL-FC), for example, was only sensitive ($SC < -0.25$ or $SC > 0.25$) to changes in HDL-associated cholesterol esterification (reactions 9), and hepatic cholesterol transport to HDL (reaction 17, Figure 4.5). Figure 4.5 indicates that some reactions strongly influenced many concentrations. The rate of hepatic uptake of cholesterol from non-HDL (reaction 5) influenced (again $|SC| > 0.25$) 4 of the 8 pools. In contrast, intestinal cholesterol synthesis (reaction 3) or intestinal cholesterol transport to HDL (reaction 16) did not influence any concentration.

An increased free cholesterol concentration (Liv-FC and Int-FC) is associated with membrane damage and cytotoxicity (29). This model analysis might, therefore, be relevant to predict cytotoxicity: if a reaction highly affects one of these pools, then substances that alter the activity of that reaction may induce cell death.

From Figure 4.5 it can be seen that Liver FC (Liv-FC) was not only sensitive to changes in the reactions that directly produced or consumed hepatic free cholesterol, like biliary cholesterol excretion (reactions 14) and hepatic cholesterol catabolism (reaction 18), but also to peripheral cholesterol synthesis (reaction 2) and intestinal cholesterol esterification (reaction 20). Int-FC was primarily sensitive to changes in intestinal cholesterol esterification (reaction 20), biliary cholesterol excretion (reactions 14), and fecal cholesterol secretion (reaction 15).

Of special biological interest are the SC values for cholesterol concentrations in plasma (HDL-FC, HDL-CE, and non-HDL-C) that are correlated with the risk for coronary heart disease (10). The influence of the different reactions on TC is not directly visible from Figure 4.5, because TC is the sum of three different concentrations in the model (HDL-FC, HDL-CE, and non-HDL-C). Sensitivity coefficients for the influence of the different reactions on TC are given separately in the left plot in Figure 4.6.

Hepatic uptake of cholesterol from non-HDL (reaction 5), hepatic transport of cholesterol to HDL (reaction 17), and hepatic cholesterol esterification (reaction 19) were the reactions that showed the largest influence on TC, resulting from the combined influence of these reactions on HDL-C (Figure 4.6, middle panel) and non-HDL-C, (Figure 4.6, right panel). Reac-

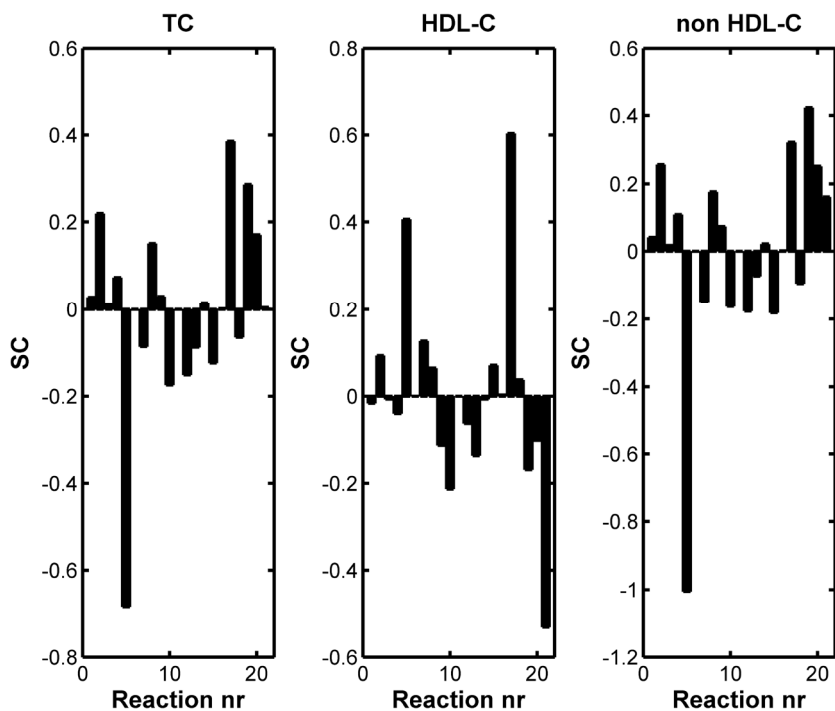


Figure 4.6. Sensitivity coefficients (see Eqn. 4.18) for the different reactions towards TC (left panel), HDL-C (middle panel), and non-HDL-C (right panel). Reaction numbers correspond to the numbers in Figure 4.1.

tions 6, 11, and 16 had the smallest effect on TC. The model predicted that non-HDL-C was highly dependent on hepatic uptake of cholesterol from non-HDL (reaction 5) and further mainly on hepatic cholesterol esterification (reaction 19, see Figure 4.6 right panel). The HDL-C concentration, on the other hand, was mainly dependent on hepatic transport of cholesterol to HDL (reaction 17), CE transfer from HDL to non-HDL (reaction 21), and hepatic uptake of cholesterol from non-HDL (reaction 5, Figure 4.6 middle panel).

In general, sensitivity coefficients for TC were more similar to those for non-HDL-C, than to those for HDL-C, because only a small fraction of plasma cholesterol is present in HDL-C.

Finally, this sensitivity analysis revealed that according to the model, there are several efficient ways to lower non-HDL-C concentrations and increase HDL-C concentrations simultaneously by modulating only one single reaction (middle and right panel Figure 4.6). The most potent re-

actions seem hepatic uptake of non-HDL-C (reaction 5) and CE transfer from HDL to LDL (reaction 21). The first reaction is targeted indirectly by statins (30), the latter reaction is targeted by CETP inhibitors, a drug class of which several members, for example dalcetrapib, are currently in clinical trials (16,31,32).

Discussion

In silico models have been used for various purposes in the study of cholesterol metabolism, for example in the interpretation of isotope labeling studies (4-6), in the analysis of the regulatory pathway of cholesterol synthesis (33), or in making predictions of the effect of genetic mutations, food, and drug interventions (34,35). These models, however, could predict the effect of only a few genetic mutations. The aim of this study was to develop a model that can be applied in the prediction of a wide variety of genetic, nutritional, and pharmaceutical effects on plasma cholesterol concentrations.

As all models, our model is a compromise between simplicity and complexity (36). A too simple model is not useful to simulate multiple interventions, because such a model will lack the targets of at least part of these interventions. A too complex model is not useful either, because insufficient experimental data will be present to calculate the parameters (calibration) or to validate the model.

The level of complexity of the model described here was apparently an acceptable compromise, because the model could be calibrated from literature data (See supplementary information), and was able to predict the effects of 10 human mutations (Figures 4.2-4.4) including a mutation in the LDLR gene (mutation I, responsible for Familial Hypercholesterolemia (FH)) and the DHCR7 gene (mutations IX, responsible for the Smith-Lemli-Opitz syndrome (SLOS)). While simulating this latter syndrome, negative concentrations were predicted for several submodels, consistent with the clinical observation that SLOS can lead to death in infants (37).

Model predictions for TC deviated on average less than 40% from experimentally observed values which is relatively good compared to the current state of the art for PBK models of exogenous substances (26-28,38). This is all the more remarkable since the model was obtained via a relatively straightforward translational adaptation of our previously developed mouse model.

There is one large deviation between model predictions and experimental data (HDL-C concentration in LCAT deficiency: mutation VIII).

The model predicts an increase in HDL-C, while a decrease was observed in reality (Figure 4.3). This increase is in the form of HDL-FC and not in HDL-CE (data not shown). The shift in CE/FC ratio is similar as seen in literature (39). An explanation may be given for the deviation between model predictions and experimental data: in reality, a maximum of HDL-FC might exist. If this maximum is reached, the transport of free cholesterol to HDL (reactions 8, 16, and 17) will reduce the transfer to HDL, thereby causing HDL-C lowering. The model is too simple to take this into account.

As given in Figure 4.6, the model predicted that an increased dietary intake of cholesterol (reaction 4) will lead to an increased TC concentration, an increased non-HDL-C concentration and a slightly lowered HDL-C concentration. This is in agreement with findings in nutritional studies (see meta-analysis in (40)). The model also predicted that non-HDL-C is mostly affected by hepatic cholesterol esterification (reaction 19) and hepatic uptake of cholesterol from LDL (reaction 5). This confirms that the liver is a dominant organ in determining the plasma cholesterol concentrations (13).

108

A decrease in hepatic cholesterol synthesis (reaction 1) resulted in a decrease of the non-HDL-C (Figure 4.6, right panel), and in an increase of HDL-C. This is in agreement with the outcome of statin-mediated inhibition of hepatic cholesterol synthesis (41). In reality, statin therapy will also cause an upregulation of the LDLR and thereby the activity of hepatic non-HDL-C uptake, increasing the non-HDL-C lowering effect.

A decrease in the activity of CETP (reaction 21) has a larger relative effect on HDL-C than on non-HDL-C (Figure 4.6). This is in agreement with the outcome of torcetrapib-mediated CETP inhibition (42) and CETP mutations (mutation VI, Table 4.5). Taken together these findings illustrate that the described model could be helpful to predict effects of both dietary and pharmaceutical interventions.

A recent genome wide association study (GWAS) has found 95 SNPs that correlated with altered TC, HDL-C, LDL-C, or triglycerides concentrations (43). At least 19 of these SNPs were near genes that are involved in one or more of the reactions in the model. The gene ABCA1, for example, is involved in the transport of cholesterol to HDL (reactions 8, 16, and 17).

We compared effect sizes of the 19 SNPs given in (43) with the sensitivity coefficients of associated reactions (Figure 4.6) by Spearman correlation, and found a positive correlation between our findings and the ones in literature for HDL-C, LDL-C, and TC (all $p < 0.05$). This is an

additional validation of the present model and implies that the developed model is useful to study the implications of genetic variations on cholesterol metabolism.

The GWAS study (43) also reports several SNPs correlating with cholesterol concentrations in plasma near genes that could not be directly linked to a specific reaction in our model (Figure 4.1), like in HNF4A, CILP2, and ANGPTL3. This absence of a direct link is the result of essential simplifications needed to construct the model.

We conclude that the approach of first developing a computational PBK model for the mouse and then translating it into a human model as described in this paper resulted in a usefully accurate model to predict plasma cholesterol concentrations in humans. Sensitivity coefficients derived from the model correlated well with recent independent GWAS data on plasma cholesterol. Based on the model performance we expect that the model can also be used for the prediction of the effects of pharmacological and dietary interventions on plasma cholesterol concentrations in humans, which will be the subject of a following study.

References

1. Fell, D. (1997) *Understanding the Control of Metabolism*, Portland press, London, UK
2. Rietjens, I. M., Boersma, M. G., Zaleska, M., and Punt, A. (2008) *Toxicol. In Vitro* 22, 1890-1901
3. Reddy, M. B., Yang, R. S. H., Clewell, H. J., and Andersen, M. E. (2005) *Physiologically Based Pharmacokinetic Modeling*, John Wiley & Sons, Inc., Hoboken, NJ, USA
4. van Schalkwijk, D. B., de Graaf, A. A., van Ommen, B., van Bochove, K., Rensen, P. C., Havekes, L. M., van de Pas, N. C. and others (2009) *J. Lipid Res.* 50, 2398-2411
5. Adiels, M., Packard, C., Caslake, M. J., Stewart, P., Soro, A., Westerbacka, J., Wennberg, B. and others (2005) *J. Lipid Res.* 46, 58-67
6. Packard, C. J. and Shepherd, J. (1997) *Arterioscler. Thromb. Vasc. Biol.* 17, 3542-3556
7. Ratushnyi, A. V., Likhoshvai, V. A., Ignat'eva, E. V., Matushkin, Y. G., Goryanin, I. I., and Kolchanov, N. A. (2003) *Dokl. Biochem. Biophys.* 389, 90-93
8. Rosamond, W., Flegal, K., Furie, K., Go, A., Greenlund, K., Haase, N., Hailpern, S. M. and others (2008) *Circulation* 117, e25-146

9. Glassberg, H. and Rader, D. J. (2008) *Annu. Rev. Med.* 59, 79-94
10. Lusic, A. J. (2000) *Nature* 407, 233-241
11. van de Pas, N. C., Soffers, A. E., Freidig, A. P., van Ommen B., Woutersen, R. A., Rietjens, I. M., and de Graaf, A. A. (2010) *Biochim. Biophys. Acta* 1801, 646-654
12. van de Pas, N. C., Woutersen, R. A., van Ommen, B., Rietjens, I. M., and de Graaf, A. A. (2011) *Biochim. Biophys. Acta* 1811, 333-342
13. Dietschy, J. M. and Turley, S. D. (2002) *J. Biol. Chem.* 277, 3801-3804
14. Agellon, L. B., Walsh, A., Hayek, T., Moulin, P., Jiang, X. C., Shelanski, S. A., Breslow, J. L. and others (1991) *J. Biol. Chem.* 266, 10796-10801
15. Zhong, S., Sharp, D. S., Grove, J. S., Bruce, C., Yano, K., Curb, J. D., and Tall, A. R. (1996) *J. Clin. Invest* 97, 2917-2923
16. Barter, P. J., Caulfield, M., Eriksson, M., Grundy, S. M., Kastelein, J. J., Komajda, M., Lopez-Sendon, J. and others (2007) *N. Engl. J. Med.* 357, 2109-2122
17. Clark, R. W., Sutfin, T. A., Ruggeri, R. B., Willauer, A. T., Sugarman, E. D., Magnus-Aryitey, G., Cosgrove, P. G. and others (2004) *Arterioscler. Thromb. Vasc. Biol.* 24, 490-497
18. Davidson, M. H., McKenney, J. M., Shear, C. L., and Revkin, J. H. (2006) *J. Am. Coll. Cardiol.* 48, 1774-1781
19. Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., and Beliles, R. P. (1997) *Toxicol. Ind. Health* 13, 407-484
20. Potter, L. K., Sprecher, D. L., Walker, M. C., and Tobin, F. L. (2009) *J. Lipid Res.* 50, 2222-2234
21. Morton, R. E. and Greene, D. J. (2003) *J. Lipid Res.* 44, 1364-1372
22. Carroll, M. D., Lacher, D. A., Sorlie, P. D., Cleeman, J. I., Gordon, D. J., Wolz, M., Grundy, S. M. and others (2005) *JAMA* 294, 1773-1781
23. Groener, J. E., Scheek, L. M., van Ramshorst E., Krauss, X. H., and van Tol A. (1998) *Atherosclerosis* 137, 311-319
24. Retzlaff, J. A., Tauxe, W. N., Kiely, J. M., and Stroebel, C. F. (1969) *Blood* 33, 649-661
25. Honda, A., Yoshida, T., Tanaka, N., Matsuzaki, Y., He, B., Shoda, J., and Osuga, T.

- (1995) *J. Gastroenterol.* 30, 651-656
26. Crouse, J. R., Grundy, S. M., and Ahrens, E. H., Jr. (1972) *J. Clin. Invest* 51, 1292-1296
27. Chobanian, A. V. and Hollander, W. (1965) *J. Lipid Res.* 6, 37-42
28. Ranheim, T., Halvorsen, B., Huggett, A. C., Blomhoff, R., and Drevon, C. A. (1995) *J. Lipid Res.* 36, 2079-2089
29. McNamara, D. J., Ahrens, E. H., Jr., Samuel, P., and Crouse, J. R. (1977) *Proc. Natl. Acad. Sci. U. S. A* 74, 3043-3046
30. Dietschy, J. M., Turley, S. D., and Spady, D. K. (1993) *J. Lipid Res.* 34, 1637-1659
31. Dietschy, J. M. and Gamel, W. G. (1971) *J. Clin. Invest* 50, 872-880
32. Miettinen, T. A. and Kesaniemi, Y. A. (1989) *Am. J. Clin. Nutr.* 49, 629-635
33. Xie, C., Turley, S. D., and Dietschy, J. M. (2009) *J. Lipid Res.* 50, 1316-1329
34. Gaffney, D., Forster, L., Caslake, M. J., Bedford, D., Stewart, J. P., Stewart, G., Wieringa, G. and others (2002) *Atherosclerosis* 162, 33-43
35. Gylling, H. and Miettinen, T. A. (1992) *J. Lipid Res.* 33, 1361-1371
36. Funke, H., von Eckardstein, A., Pritchard, P. H., Albers, J. J., Kastelein, J. J., Droste, C., and Assmann, G. (1991) *Proc. Natl. Acad. Sci. U. S. A* 88, 4855-4859
37. Funke, H., von Eckardstein, A., Pritchard, P. H., Karas, M., Albers, J. J., and Assmann, G. (1991) *J. Clin. Invest* 87, 371-376
38. Bennion, L. J. and Grundy, S. M. (1975) *J. Clin. Invest* 56, 996-1011
39. Field, F. J., Watt, K., and Mathur, S. N. (2008) *J. Lipid Res.* 49, 2605-2619
40. Mott, S., Yu, L., Marcil, M., Boucher, B., Rondeau, C., and Genest, J., Jr. (2000) *Atherosclerosis* 152, 457-468
41. Gregg, R. E., Zech, L. A., Schaefer, E. J., and Brewer, H. B., Jr. (1981) *Science* 211, 584-586
42. Calabresi, L., Pisciotta, L., Costantin, A., Frigerio, I., Eberini, I., Alessandrini, P., Arca, M. and others (2005) *Arterioscler. Thromb. Vasc. Biol.* 25, 1972-1978
43. Jira, P. E., de Jong, J. G., Janssen-Zijlstra, F. S., Wendel, U., and Wevers, R. A. (1997) *Clin. Chem.* 43, 129-133

44. Wanders, R. J., Romeijn, G. J., Wijburg, F., Hennekam, R. C., de Jong, J., Wevers, R. A., and Dacremont, G. (1997) *J. Inherit. Metab Dis.* 20, 432-436
45. Pullinger, C. R., Eng, C., Salen, G., Shefer, S., Batta, A. K., Erickson, S. K., Verhagen, A. and others (2002) *J. Clin. Invest* 110, 109-117
46. Punt, A., Paini, A., Boersma, M. G., Freidig, A. P., Delatour, T., Scholz, G., Schilter, B. and others (2009) *Toxicol. Sci.* 110, 255-269
47. Lupfert, C. and Reichel, A. (2005) *Chem. Biodivers.* 2, 1462-1486
48. Hissink, A. M., van Ommen, B., Kruse, J., and van Bladeren, P. J. (1997) *Toxicol. Appl. Pharmacol.* 145, 301-310
49. Quick, D. J. and Shuler, M. L. (1999) *Biotechnol. Prog.* 15, 540-555
50. Tabas, I. (1997) *Trends Cardiovascular medicine* 7, 256-263
51. Davidson, M. H. and Toth, P. P. (2004) *Prog. Cardiovasc. Dis.* 47, 73-104
52. de Grooth, G. J., Kuivenhoven, J. A., Stalenhoef, A. F., de Graaf J., Zwinderman, A. H., Posma, J. L., van Tol A. and others (2002) *Circulation* 105, 2159-2165
53. Katsnelson, A. (2010) *Nature* 468, 354
54. Kervizic, G. and Corcos, L. (2008) *BMC. Syst. Biol.* 2, 99
55. Knoblauch, H., Schuster, H., Luft, F. C., and Reich, J. (2000) *J. Mol. Med.* 78, 507-515
56. Muzaffer Demirezen, E. (2009) A simulation model for blood cholesterol dynamics and related disorders. Graduate Program in Industrial Engineering; Bogazici University,
57. Ostlund, R. E., Jr. (1993) *Am. J. Physiol* 265, E513-E520
58. Waterham, H. R., Wijburg, F. A., Hennekam, R. C., Vreken, P., Poll-The B.T., Dorland, L., Duran, M. and others (1998) *Am. J. Hum. Genet.* 63, 329-338
59. Punt, A., Freidig, A. P., Delatour, T., Scholz, G., Boersma, M. G., Schilter, B., van Bladeren, P. J. and others (2008) *Toxicol. Appl. Pharmacol.* 231, 248-259
60. Clarke, R., Frost, C., Collins, R., Appleby, P., and Peto, R. (1997) *BMJ* 314, 112-117
61. Nicholls, S. J., Tuzcu, E. M., Sipahi, I., Grasso, A. W., Schoenhagen, P., Hu, T., Wolski, K. and others (2007) *JAMA* 297, 499-508
62. Kastelein, J. J., van Leuven, S. I., Burgess, L., Evans, G. W., Kuivenhoven, J. A.,

Barter, P. J., Revkin, J. H. and others (2007) *N. Engl. J. Med.* 356, 1620-1630

63. Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P. and others (2010) *Nature* 466, 707-713



CHAPTER

5

Predicting individual responses to pravastatin using a kinetic model for plasma cholesterol concentrations.

Niek C.A. van de Pas, Ruud A. Woutersen, Ben van Ommen, Ivonne M.C.M. Rietjens, and Albert A. de Graaf

Supplementary material can be found at www.vdpastreffers.nl/cholesterol

Submitted



Abstract

Individuals vary in the response of their plasma cholesterol concentrations to statins, due to variation in multiple processes including cholesterol synthesis and cholesterol absorption. The objective of this study was to identify and quantify the contribution of 21 of these processes using a physiologically based kinetic (PBK) model for plasma cholesterol concentrations in humans. To reach this goal, we investigated whether this model could predict the effect of pravastatin intervention on plasma cholesterol. With our PBK model, pravastatin therapy was simulated, taking into account pravastatin pharmacokinetics, potency, and influence on expression of the Low Density Lipoprotein Receptor. Variation in response between subjects was studied using 10,000 virtual subjects, each with a unique combination of rate constants for the different reactions.

The model predicted the effect of pravastatin treatment on Total Cholesterol (TC), High Density Lipoprotein-Cholesterol (HDL-C), and non-HDL-C at several doses (0 - 40 mg/d). At 40 mg/d, TC was predicted to decrease by 15% (vs. 22% reported for human subjects in literature), non-HDL-C was predicted to decrease by 22% (vs. 25% in literature), and HDL-C was predicted to increase by 10% (vs. 5.6% in literature).

The model also predicted that rates of hepatic cholesterol synthesis, peripheral cholesterol synthesis and hepatic cholesterol esterification were major determinants of the non-HDL-C response to pravastatin.

Concluding, we have developed a model that accurately predicted effects of pravastatin treatment and can be useful to provide insight in factors influencing the individual response to statins.

Introduction

Prevention of cardiovascular events is highly relevant in modern medicine, because these events are a major cause of death and morbidity in the Western World (1-3). Important biomarkers for the risk of these events are plasma concentrations of total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), and low density lipoprotein cholesterol (LDL-C) (1-3). Subjects with a high concentration of TC or LDL-C have an increased risk of developing cardiovascular complications. Alternatively, subjects with a high concentration of HDL-C have a reduced risk (4). These biomarkers are often used as substitute markers to measure the effect of therapy on mortality rates, because using mortality rates as output takes a long follow up period and large patient groups to find significant effects in mortality rates (5,6).

The most commonly prescribed drugs to target these biomarkers are 3-hydroxy-3-methyl-glutaryl-CoA reductase (HMG-CoA reductase) inhibitors or statins (7). Since the discovery of the first statin in fungi (See historical review of Endo (8)), several compounds in this drug class have been produced and marketed, including pravastatin, simvastatin, and rosuvastatin (7). These drugs share the same working mechanism of LDL-C lowering, but differ in their metabolism, absorption, and potency (7,9).

118 Statins act through inhibition of the hepatic activity of HMG-CoA reductase, an essential enzyme in hepatic cholesterol synthesis. To compensate for the resulting reduction of hepatic cholesterol, hepatic expression of the LDL Receptor (LDLR) is increased, accelerating the uptake of LDL-C from plasma, which reduces the LDL-C concentration in plasma (7).

In large clinical trials, statins decreased LDL-C concentrations, increased HDL-C concentrations, and decreased the risk of cardiovascular events (5,7,10). However, a large group of patients only had a minor response to statins in terms of cholesterol concentrations (11,12). Hereafter the term 'response' will be used for the response of cholesterol concentrations to statin treatment, rather than for cardiovascular events. In fact, only 40% of subjects receiving statin therapy reach their LDL-C targets (13). Subjects that have a large response are called hyper-responders. Subjects that have a small response are called hypo-responders.

There are two indications that this variation in response can in part be explained by variation in cholesterol metabolism. The first indication is that subjects who obtain a relatively large proportion of their cholesterol from *de novo* synthesis respond better than subjects who obtain a relatively

large proportion from dietary absorption (11,14). The second indication is that genetic analyses revealed that variations in key genes in cholesterol metabolism (like APOE, which is involved in the hepatic uptake of LDL-C, and ABCA1, which transports cholesterol to HDL) are associated with the response (15,16). These genetic studies indicate that differences in multiple processes including cholesterol synthesis and cholesterol transport contribute to variation in the response to statins.

Genetic analyses provide associations of genetic variants with the response, but not necessarily more mechanistic insights in how these factors contribute to the response. If we know which reactions are important for the response to statins and why they are important, the search for biomarkers that predict the effect of treatment in humans could be better targeted and accelerated. These markers could be more generally applicable than tests that only characterize a single genetic variant. This insight can be used for personalized medicine or personalized health, i.e. tailoring therapy to subject-specific properties.

Computational models have been used to study differences between hypo-responders and hyper-responders arising from variations in statin pharmacokinetics and pharmacodynamics (17), but not for studying differences arising from variations in cholesterol metabolism. We hypothesized that a computational model of cholesterol metabolism that focuses on plasma cholesterol concentrations can enhance current insights in differences between hypo-responders and hyper-responders arising from differences in cholesterol synthesis, transport, and catabolism.

Our previously developed physiologically based kinetic (PBK) model for predicting plasma cholesterol concentrations in humans (18,19) includes the two required ingredients for simulating statin treatment: the target reaction of statins and a plasma compartment. As a result, this model is suitable to study the variation in statin response. The model describes 8 different cholesterol pools in the body influenced by 21 metabolic and transport reactions. The 8 cholesterol pools include HDL free cholesterol (HDL-FC), HDL cholesterol ester (HDL-CE), and non-HDL-C (total of plasma cholesterol that is not present in HDL), but also hepatic free cholesterol, hepatic cholesterol ester, intestinal free cholesterol, intestinal cholesterol ester, and peripheral cholesterol.

The objectives of the present study were to test how accurate our cholesterol PBK model predicted the effect of a statin treatment, and to investigate which reactions in cholesterol synthesis, transport, catabolism, and excretion determine the variability in the response to statins.

Methods

Model development

The cholesterol PBK model for a standard subject (“reference man” as defined by the International Commission on Radiation Protection) was described previously (19), in short: the model consists of a set of 8 differential equations each describing the dynamics of a cholesterol pool. These equations describe the effect of 21 metabolic and transport reactions on the size of the 8 included cholesterol pools. The rate of each reaction is described using a rate equation with a single parameter: the rate constant.

The model consists of a set (ensemble) of 8 submodels, each having a different combination of first and zero order kinetics for the various reactions, instead of one optimal model. Details on the selection of these 8 submodels can be found elsewhere (19). The model prediction is calculated as the average of the predictions of the submodels (19).

Model simulations

The effect of pravastatin was simulated by reducing the rate constant of the reaction for hepatic cholesterol synthesis (reaction 1, (19)) according to:

$$k_1^{\text{treated}} = f_{\text{act}} \cdot k_1^{\text{untreated}} \quad (\text{Eqn. 5.1})$$

120

where k_1 is the rate constant for reaction 1, either in standard condition (superscript: untreated) or after pravastatin treatment (superscript: treated), and f_{act} is the treated hepatic cholesterol synthesis rate over the day relative to the untreated rate. If the cholesterol synthesis rate in the liver is, for example, halved by statin treatment, then f_{act} has a value of 0.5. This factor f_{act} is dependent on the statin dose administered and was determined in the present study using pharmacokinetic and pharmacodynamic data as described in the Result section. All other parameters in the cholesterol PBK model were assumed to be unaffected by statin treatment. The model predictions were obtained by solving the model with the new set of parameters (including k_1^{treated} instead of $k_1^{\text{untreated}}$) by numerical integration with the standard concentrations as initial values. The simulation was performed until steady state of all cholesterol pools in the model was reached. Model predictions were defined as these steady state concentrations (19). Integration was performed using routine *ode15s* as implemented in MATLAB version 7.5 (R2007b) (20).

Virtual subjects

To determine which reactions are important for the response to statins we used a Monte Carlo simulation: a group of 10,000 virtual subjects was constructed, each differing in the values of the kinetic parameters in the 21 reactions. Each kinetic parameter was drawn randomly from a normal distribution with a specified average and standard deviation. The average was set to the rate constant in the standard subject and the standard deviation was set to 25% of this average. The model was used to calculate a baseline concentration and then statin treatment was simulated as described in the section on Model simulations with a fixed f_{act} . For some virtual subjects, one or more of the 8 submodels predicted one or more negative concentrations (even if all rate constants were positive). These subjects were not taken into account.

Reaction importance

The relation between the response and the reaction rate constant for each of the 21 reactions (all continuous variables) was assessed using linear regression (21):

$$y_n = \beta_0 + \beta_1 \tilde{k}_{n,i} + \varepsilon_n \quad (\text{Eqn. 5.2})$$

where y_n is the response to statins (expressed relative to baseline) for subject n , β_0 , and β_1 are the regression coefficients fitted to experimental data by minimizing the sum of the squared error terms (ε_n), and $\tilde{k}_{n,i}$ is the (scaled) rate constant of subject n for reaction i . In order to compare reaction effects for different reactions, reaction constants for a given reaction were autoscaled (22):

$$\tilde{k}_{n,i} = \frac{k_{n,i} - \bar{k}_i}{\sigma_i} \quad (\text{Eqn. 5.3})$$

where $k_{n,i}$ is the rate constant of subject n for reaction i , \bar{k}_i the average rate constant for all subjects for this reaction, and σ_i the standard deviation of the rate constants for reaction i over all subjects.

The second regression coefficient (β_1) is called the reaction effect. Regression was performed using the function `regress` as implemented in MATLAB. Because responses were not normally distributed, Johnson correction was applied using the function `johnsrand` as implemented in MATLAB (20).

Interactions between reactions were tested using linear regression with a multiplicative interaction term (23):

$$y_n = \beta_0 + \beta_i \tilde{k}_{n,i} + \beta_j \tilde{k}_{n,j} + \beta_{\text{int}} \tilde{k}_{n,i} \tilde{k}_{n,j} + \varepsilon_n \quad (\text{Eqn. 5.4})$$

where y_n is the response to statins of subject n (relative to baseline), and β_0 , β_i , β_j , and β_{int} are the regression coefficients that are fitted to the data minimizing the sum of the squared error terms (ε_n). An interaction is considered significant if the 95% confidence interval of β_{int} does not contain zero.

Results

Model adaptation

Before we investigated which reactions in cholesterol metabolism determine the variability in the response to pravastatin, we first studied whether our previously defined model (19) could predict the effect of pravastatin treatment on plasma cholesterol. If it was assumed that statins fully block hepatic cholesterol synthesis ($f_{\text{act}} = 0$, see Eqn. 5.1), then our previously described model predicted non-HDL-C to be reduced by 4%, whereas in literature reductions of 20-50% are observed (7). Therefore, an adaptation of the model was required.

The adaptation was based on the mechanism of action of statins: statins cause a decrease in hepatic cholesterol followed by upregulation of the LDL Receptor (LDLR) involved in the hepatic uptake of LDL-C (7). We, therefore, included the upregulation of the LDLR in the model. The updated conceptual model is given in Figure 5.1.

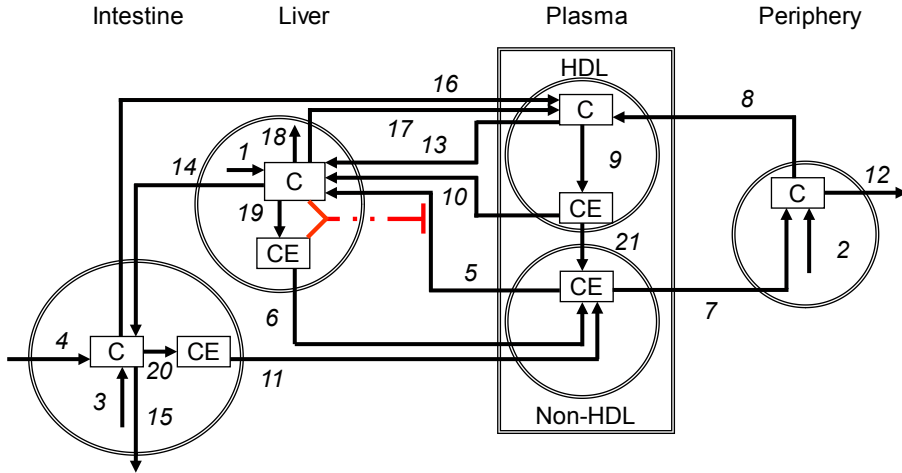


Figure 5.1. Conceptual model for pathways determining plasma cholesterol concentrations used as a basis to set up the PBK cholesterol model of the present study. Process numbers stand for: 1, Hepatic cholesterol synthesis; 2, Peripheral cholesterol synthesis; 3, Intestinal cholesterol synthesis; 4, Dietary cholesterol intake; 5, Hepatic uptake of cholesterol from non-HDL; 6, Hepatic Very Low Density lipoprotein cholesterol (VLDL-C) production; 7, Peripheral uptake of cholesterol from non-HDL; 8, Peripheral cholesterol transport to HDL; 9, HDL associated cholesterol esterification; 10, Hepatic HDL-CE uptake; 11, Intestinal chylomicron cholesterol secretion; 12, Peripheral cholesterol loss; 13, Hepatic HDL-FC uptake; 14, Biliary cholesterol excretion; 15, Fecal cholesterol excretion; 16, Intestinal cholesterol transport to HDL; 17, Hepatic cholesterol transport to HDL; 18, Hepatic cholesterol catabolism; 19, Hepatic cholesterol esterification; 20, Intestinal cholesterol esterification, and 21, CE transfer from HDL to non-HDL. C stands for cholesterol; CE for cholesterol ester. Adapted from (35). Dashed line indicates the regulation of the LDLR in response to the hepatic cholesterol concentration.

It includes downregulation of the LDLR when hepatic cholesterol (Sum of FC and CE) is increased. This also implies upregulation when the hepatic cholesterol concentration is decreased.

The LDLR regulation was included in the rate expression for hepatic uptake of cholesterol from non-HDL-C (reaction 5) that was previously defined as (19):

$$v_5 = k_5 \cdot [\text{non-HDL-C}] \quad (\text{Eqn. 5.5})$$

where v_5 is the rate of reaction 5, k_5 is the rate constant, and $[\text{non-HDL-C}]$ is the non-HDL-C concentration. The effect of hepatic cholesterol on LDLR expression levels was quantified using several studies that reported both the LDLR gene expression and the cholesterol content of HepG2 cells (24-26). Measurements in these studies were performed in situations where cellular cholesterol concentrations were increased by incubating the cells with various concentrations of LDL, or decreased using statins. Figure 5.2 shows that a higher cholesterol concentration in HepG2 cells correlates with a lower LDLR expression.

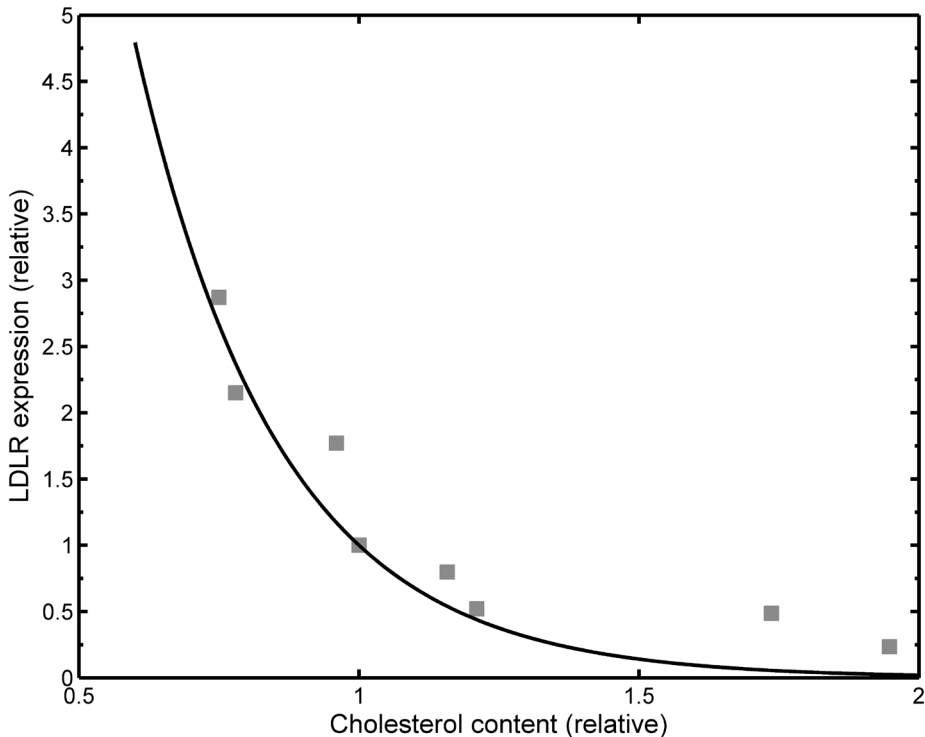


Figure 5.2. Cholesterol content vs. LDLR expression of HepG2 cells (both relative to untreated conditions). Data were obtained from (24-26). The line is a one-parameter exponential function (Eqn. 5.7) fitted to the data (see text for details).

The model included this LDLR regulation in the rate equation for reaction 5 (see Eqn. 5.5) that was modified into Eqn. 5.6:

$$v_5 = k_5 \cdot [\text{non HDL-C}] \cdot E_{LDLR} \quad (\text{Eqn. 5.6})$$

where E_{LDLR} is the expression of the LDLR relative to the untreated situation, calculated according to:

$$E_{LDLR} = \frac{e^{-3.92 \left(\frac{c_{liv}}{c_{liv}^{untreated}} \right)}}{e^{-3.92}} \quad (\text{Eqn. 5.7})$$

where $\left(\frac{c_{liv}}{c_{liv}^{untreated}} \right)$ is the ratio of the total hepatic cholesterol concentration (sum of free and unesterified), relative to the untreated standard steady state condition. The value of 3.92 was obtained by fitting to the data in Figure 5.2 using the relation in Eqn. 5.7. In the untreated condition, the gene expression of the LDLR is defined as being equal to one (Figure 5.2) such that, the rate equation of reaction 5 reduces to the expression in Eqn. 5.5.

The model without LDLR upregulation (19) was validated previously by simulating the effect of genetic mutations on plasma cholesterol values. Without upregulation, model predictions on TC deviated 36% from experimental values (19). With upregulation, model predictions deviated 39% from experimental values ((19), data not shown). With upregulation of the LDLR, model performance was, therefore, comparable to the previous model. Thus, the adapted model was still able to accurately describe the effects of mutations on plasma cholesterol concentrations.

Effect of pravastatin on hepatic cholesterol synthesis; determination of f_{act}

For the simulation of a pravastatin treatment in our PBK cholesterol model, a relation between the hepatic cholesterol synthesis rate and the pravastatin dose administered orally had to be obtained. This relation was obtained in three steps. In step 1, a dose-response curve was constructed from a study that relates the cholesterol synthesis rate in human hepatocytes to the pravastatin concentration added *in vitro*. In step 2, the pharmacokinetics and pharmacodynamics of pravastatin were studied using the PBK model of pravastatin developed by Watanabe et al. (17) (hereafter: the model of

Watanabe et al. is referred to as pravastatin PBK model as opposed to the developed cholesterol model that is referred to as cholesterol (PBK) model or just model). In step 3, the *in vitro* response curve was coupled to the pravastatin PBK model and various orally administered pravastatin doses were simulated to calculate their effect on hepatic cholesterol synthesis. As the pravastatin PBK model requires adaptations to describe other statins, this work is henceforth limited to pravastatin.

The first step described above was to reconstruct an *in vitro* dose-response curve from literature. Data obtained from Van Vliet et al. (9) indicate that the pravastatin mediated inhibition of cholesterol synthesis follows a Hill type response (9). Parameter values in this dose-response curve were determined by curve fitting to be 0.82 (dimensionless) for the Hill coefficient and 2.0 nM for the IC_{50} for pravastatin added to human hepatocytes (9). Figure 5.3 shows the experimental data and dose-response curve obtained with these parameters.

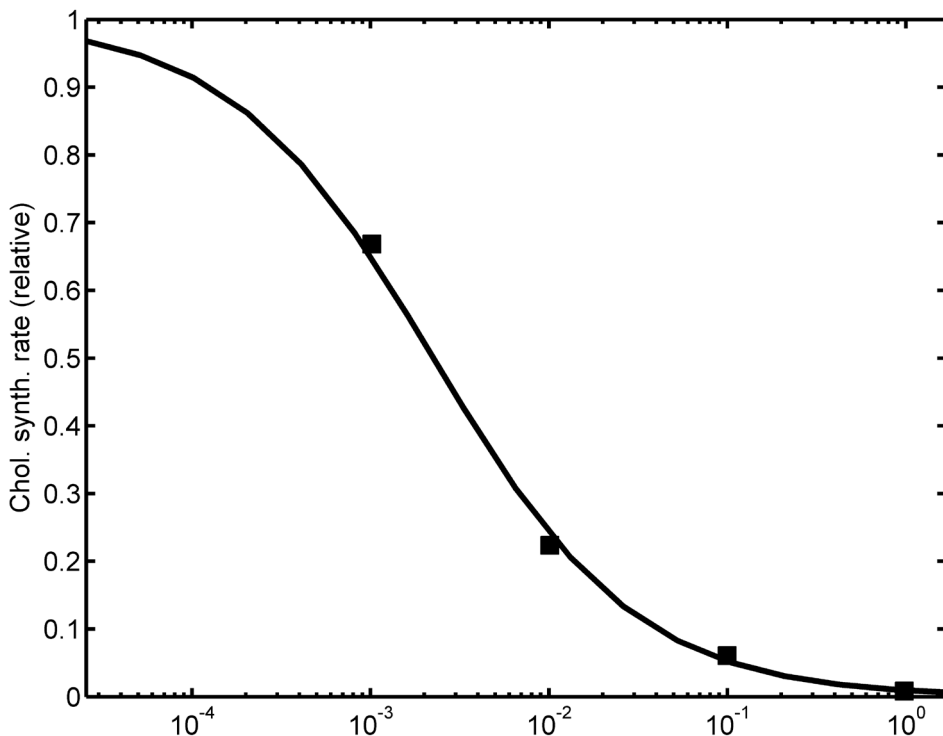


Figure 5.3. Concentration response curve of the effect of increasing concentration of pravastatin added to the hepatocytes on the rate of cholesterol synthesis (relative to untreated). Squares indicate data from Van Vliet et al. (9).

Of all predicted concentrations in the pravastatin PBK model, the hepatic extracellular pravastatin concentration was considered to be the best approximation of the concentration applied to hepatocytes in an in vitro assay. Therefore, in the second step, the pravastatin PBK model was used to predict the hepatic extracellular concentration of pravastatin in time after an oral pravastatin dose of 40 mg. The results (Figure 5.4A) show that after administration, the pravastatin concentration increased to a peak value at $t = 0.9$ h, followed by a gradual decrease with time.

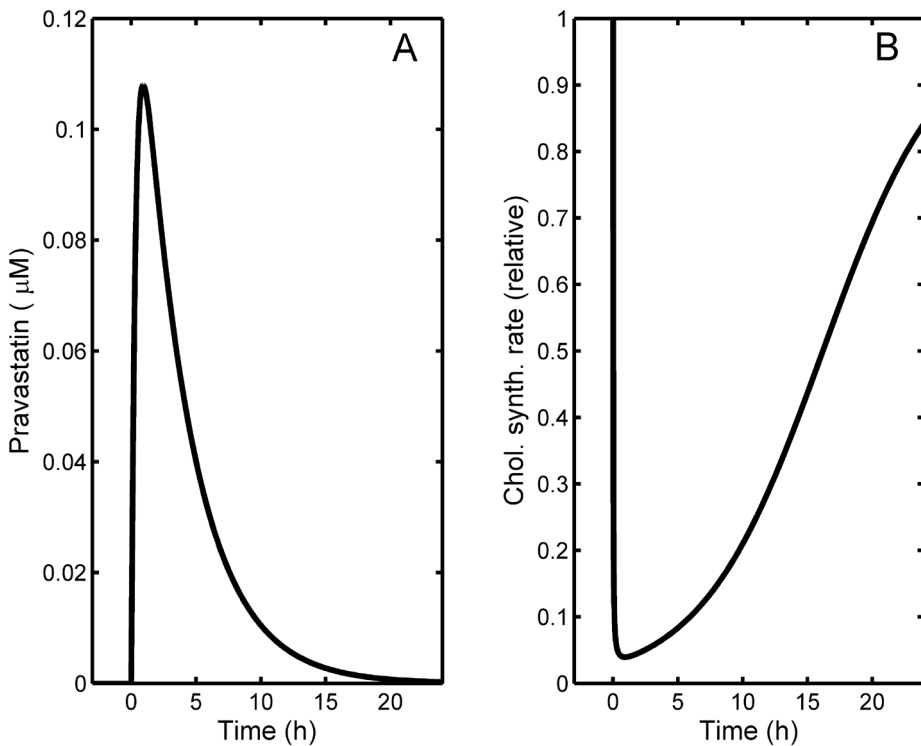


Figure 5.4. Time courses following an oral pravastatin dose of 40 mg at time = 0 of A) hepatic extracellular pravastatin concentration calculated with the pravastatin PBK model from Watanabe et al. (17) and B) hepatic cholesterol synthesis rate in man relative to control, calculated from the concentration profile using Figure 5.3.

In the third and final step, the time dependent hepatic extracellular pravastatin concentration predicted by the pravastatin PBK model to occur upon an oral dose was linked to the cholesterol synthesis rate using the dose-

response curve depicted in Figure 5.3. Figure 5.4B shows the time course of this cholesterol synthesis rate (relative to untreated) after a dose of 40 mg pravastatin. The cholesterol synthesis rate first decreased steeply and then increased more smoothly in time. The synthesis rate did not reach the untreated rate within a day (relative cholesterol synthesis rate equal to 1), because a small amount of pravastatin was still present in the body. Given the low IC_{50} of 2.0 nM for the inhibition of the cholesterol synthesis rate by pravastatin (9), this small amount was sufficient to result in residual inhibition (See supplementary information: details on pravastatin PBK model).

Pravastatin (40 mg) decreased the average rate of cholesterol synthesis, calculated as $AUC_{24/24}$, to 0.35 (relative to the untreated rate). This means that the 24 h cholesterol synthesis was reduced by 65%. The factor of 0.35 is referred to as f_{act} (See Eqn. 5.1). In a similar way the factor f_{act} was calculated for various doses of pravastatin (0 - 80 mg). The result, shown in Figure 5.5, indicates that the average daily cholesterol synthesis rate decreased non-linearly with increasing pravastatin dose.

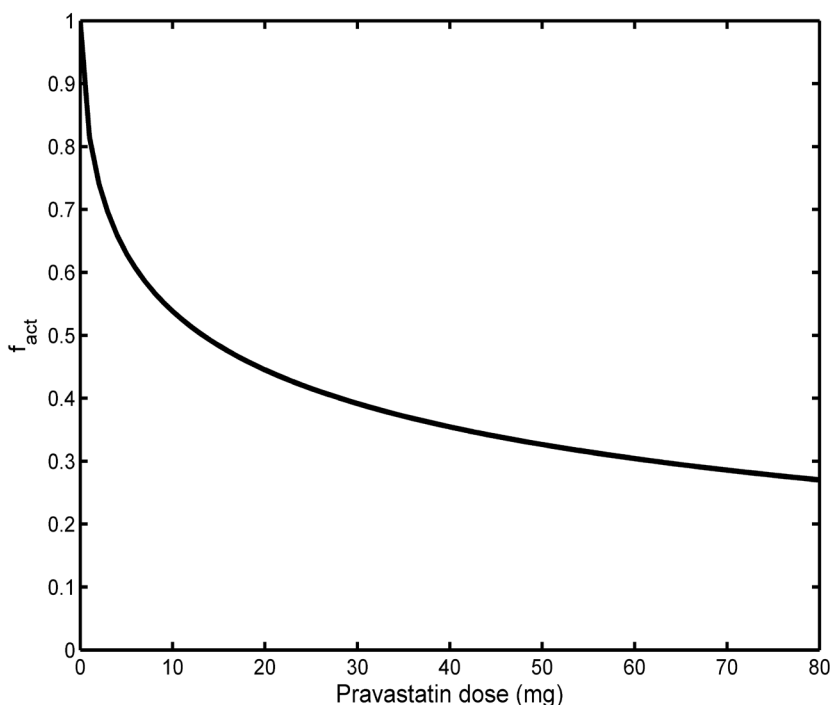


Figure 5.5. Calculated average hepatic cholesterol synthesis rate over the day (relative to the untreated rate, f_{act} , see Eqn. 5.1) as a function of pravastatin dose.

Upon a dose of 10 mg pravastatin, the average rate of cholesterol synthesis was decreased by 47% compared with the untreated situation. Increasing the dose from 40 mg (65% reduction in cholesterol synthesis) to 80 mg, only reduced hepatic cholesterol synthesis rate by a further 8% (73% reduction). The obtained values of f_{act} were used to simulate pravastatin treatment using our cholesterol model (see Eqn. 5.1).

It is computationally demanding to take the daily rhythm of cholesterol synthesis inhibition (Figure 5.4B) into account in the cholesterol model. For a pravastatin dose of 40 mg/d, simulations were made with the dynamic rhythm of cholesterol inhibition or with the constant 24h inhibition (65%). Results were identical. Therefore, for other simulations, it was assumed that the reduction of hepatic cholesterol synthesis was constant over the day at the value corresponding to the constant 24h inhibition.

Plasma cholesterol prediction

Figure 5.6 presents model predictions of TC, HDL-C, and non-HDL-C concentrations in response to pravastatin applied in doses of 0, 10, 20, and 40 mg/d. Predictions were compared with data obtained in the STELLAR trial (27), because in this trial several dose groups of pravastatin were tested (See Figure 5.6).

The model correctly predicted a decrease in TC and non-HDL-C and a slight increase in HDL-C upon pravastatin treatment (Figure 5.5).

At 40 mg/d, TC was predicted to decrease by 15% (vs. 22% reported for human subjects in literature), non-HDL-C was predicted to decrease by 22% (vs. 25% reported for human subjects in the literature), and HDL-C was predicted to increase by 10% (vs. 5.6% reported for human subjects in the literature).

Hillebrant et al. (28) performed liver biopsies from gallstone patients receiving either both 40 mg pravastatin and 1 g ursodeoxycholic acid (UDCA, a treatment for gallstones), or UDCA alone and did not observe a difference in hepatic cholesterol between the two groups (28). These findings can be compared with model predictions, because the model contains a liver compartment (Figure 5.1). The model predicted that a dose of 40 mg/d of pravastatin reduced hepatic cholesterol from 13.3 mM to 12.6 mM, a reduction of only 6%.

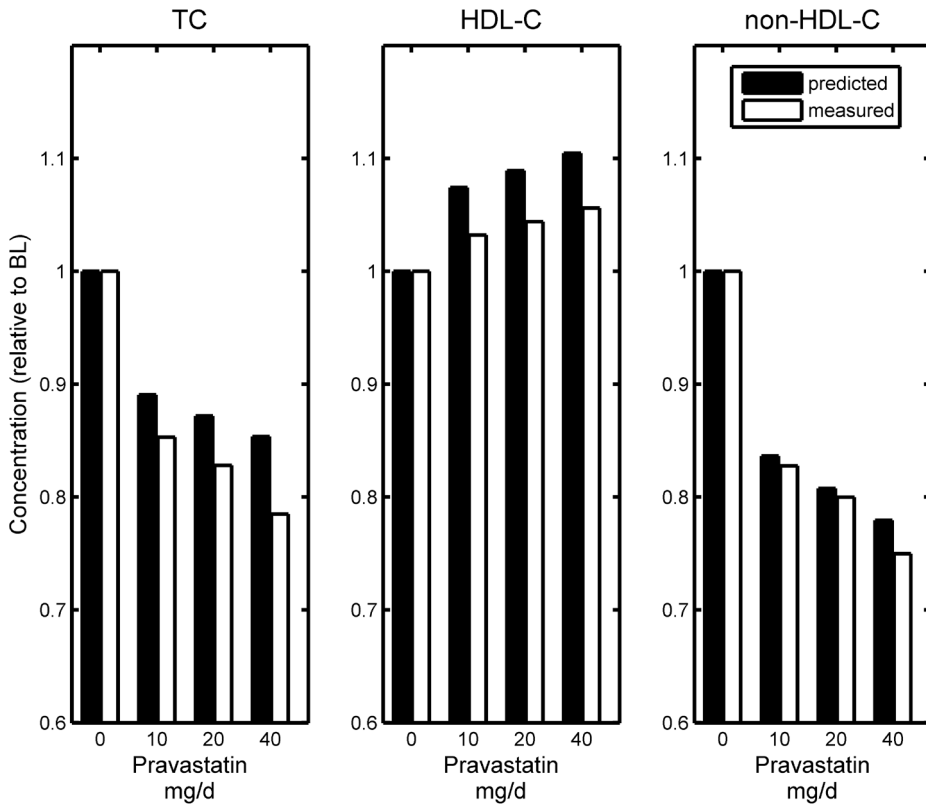


Figure 5.6. Model predictions (predicted) vs. literature data (measured, as reported in Jones et al. (27)) of the effect of pravastatin treatment (0 - 40 mg/d) on TC, HDL-C, and non-HDL-C concentrations. All concentrations are reported relative to baseline (BL).

This minor decrease would be hard to pick up experimentally and, therefore, the model prediction is in agreement with the findings of Hillebrant et al. (28). In response to the aforementioned small change in total hepatic cholesterol, the LDLR expression was predicted to be upregulated by 24% (fill in 1.06 for $\frac{C_{liv}}{C_{liv}^{untreated}}$ in Eqn. 5.7).

The model also predicted the time scale of non-HDL-C reduction: 50% of the non-HDL-C reduction was reached after 11 days. This time scale is fully in agreement with the statement of Jones et al. (27) that “it is well established that statins exhibit most of their LDL cholesterol reducing effects within 2 weeks and produce full effects by 4 to 6 weeks” (27).

Individual statin response

In a next step, the model was used to study the impact of interindividual variation in response to statins with a Monte Carlo approach. Therefore, a virtual population was created with 10,000 subjects differing in the kinetic parameters of the 21 reactions included in the model. For each subject, 21 parameters were randomly drawn from a normal distribution (See Methods). All virtual subjects were treated with 40 mg/d pravastatin, assuming no differences in pravastatin pharmacokinetics and dynamics (identical f_{act}). In this simulation, 2,391 subjects had one or more negative concentrations in one or more submodels and were, therefore, not taken into account.

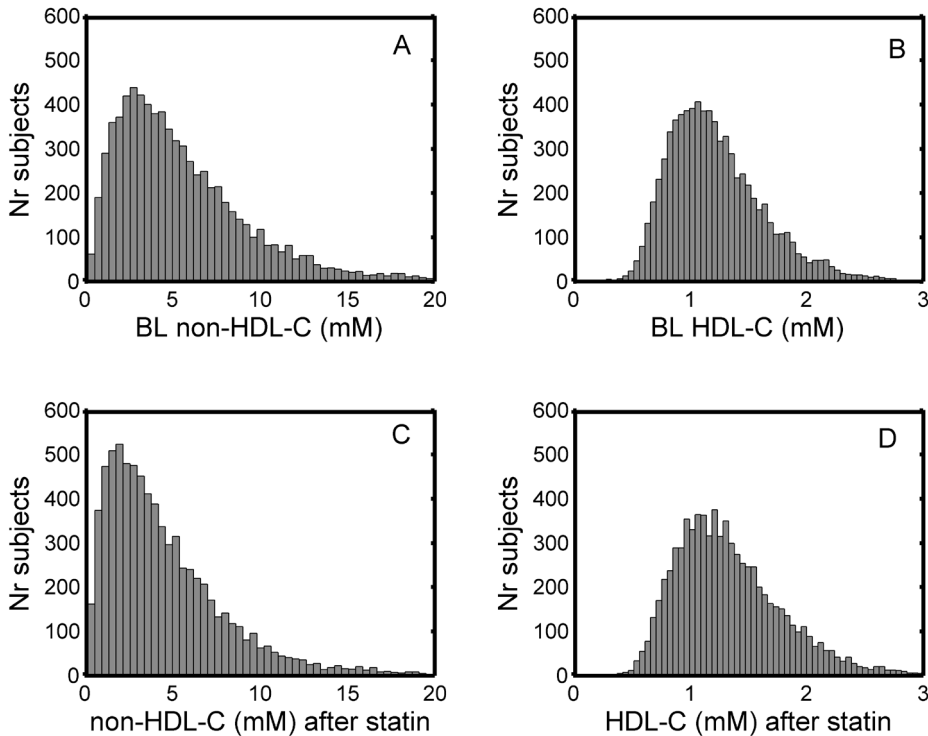


Figure 5.7. Histograms of the predicted non-HDL-C (Figure 5.7A) and HDL-C (Figure 5.7B) concentrations at baseline (BL) and after 40 mg/d pravastatin treatment for 7609 subjects (Figures 7C and 7D). For clarity, the x-axis was cut off at 3 mM for HDL-C and 20 mM for non-HDL-C. With these cut offs, 99.6% of the virtual subjects were included in the graph for HDL-C and 98.7% of the virtual subjects were included in the graph for non-HDL-C.

The remaining group of 7.609 subjects had baseline concentrations of 7.01 ± 4.28 mM for TC, 1.27 ± 0.42 mM for HDL-C, and 5.92 ± 4.48 mM for non-HDL-C (average \pm SD). As in the standard subject (see Figure 5.6), pravastatin treatment decreased non-HDL-C (see Figure 5.7A and 7C) and increased HDL-C (see Figure 5.7B and 7D). Figure 5.8A shows the non-HDL-C reduction in response to pravastatin vs. the baseline non-HDL-C concentrations. Each dot in Figure 5.8A represents a single virtual subject.

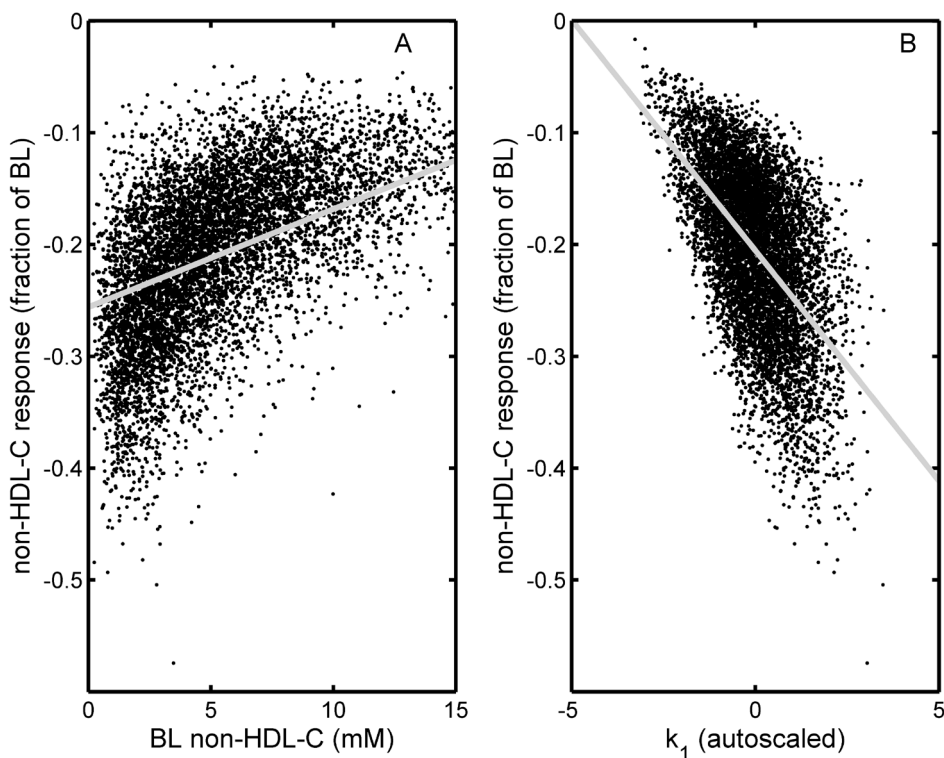


Figure 5.8. Predicted non-HDL-C response A) vs. baseline (BL) non-HDL-C concentration and B) vs. the rate constant of hepatic cholesterol synthesis (k_1 , relative to average k_1). X and y axes cutoffs have been adjusted for clarity, therefore, 95% (for Figure 5.8A) and 99% (for Figure 5.8B) of the virtual subjects were included in the graphs. Grey lines indicate trend lines (first order polynomial fitted to the data).

As indicated in Figure 5.8A, subjects with a high baseline non-HDL-C showed, in general, a smaller relative decrease in non-HDL-C than subjects

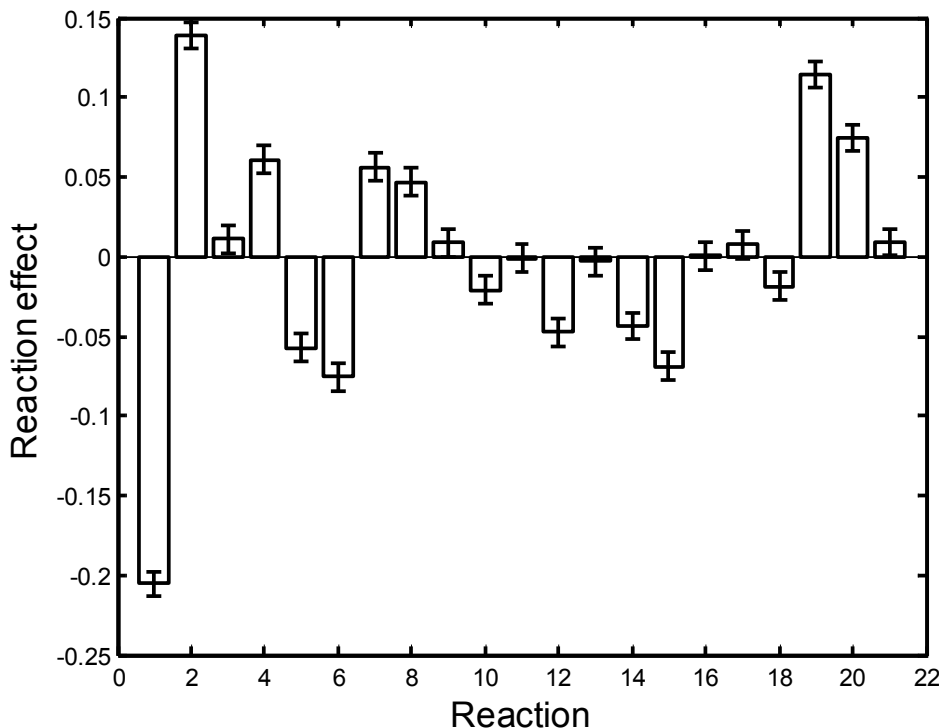
with a lower baseline concentration. As can be seen by the large scatter in Figure 5.8A, the model predicted large interindividual variation in the non-HDL-C response to pravastatin, as has also been found in clinical trials (29). As an example, for subjects with a baseline non-HDL-C concentration of 5 mM, some hyper-responders had a response of 40%, whereas some hypo-responders had a response of only 5%.

To identify the factors that caused this variation, the impact of all 21 reactions on the pravastatin response was assessed. As an example, Figure 5.8B shows the relation between the non-HDL-C response to pravastatin and the autoscaled rate constant of hepatic cholesterol synthesis (reaction 1). Subjects with a high rate constant for this reaction had (on average) a higher treatment effect of pravastatin.

The impact of having an unusual value of each of the 21 reactions on the non-HDL-C reduction by statins was studied by calculating the reaction effect (i.e. β_1 in Eqn. 5.3). If no correlation occurred between rate constants of a particular reaction and the statin response, the reaction effect was zero, implying that the reaction was not important for the response to pravastatin.

The reaction effect for reaction 1 (hepatic cholesterol synthesis) was negative, indicating that subjects with a larger than average hepatic cholesterol synthesis rate had, in general, a higher than average reduction of non-HDL-C in response to statins (See also Figure 5.8B). The three reactions with the largest (positive and negative) effects in Figure 5.9 were found for hepatic cholesterol synthesis (reaction 1), peripheral cholesterol synthesis (reaction 2), and hepatic cholesterol esterification (reaction 19).

A similar procedure was performed for HDL-C and the resulting reaction effects are shown in Figure 5.10. In this case, reaction 1 had a positive reaction effect, indicating that a higher hepatic cholesterol synthesis rate was correlated with a larger HDL-C increase upon statin treatment. The three reactions with the largest (positive and negative) effects in Figure 5.10, were found for hepatic cholesterol synthesis (reaction 1), hepatic HDL-CE uptake (reaction 10), and peripheral cholesterol synthesis (reaction 2).



134 *Figure 5.9. Reaction effects of all 21 reactions in determining the non-HDL-C response to pravastatin and their 95% confidence interval. Reaction numbers correspond to numbers in Figure 5.1.*

The important reactions are catalyzed by known genes and, therefore, these model predictions could be compared with effects of genetic variations affecting the cholesterol response to statins as reviewed by Schmitz and Langmann (15). They cluster influential genes in the groups: ‘Lipid metabolism’, ‘Drug metabolism and transporters’, and ‘Others’. As the 21 reactions in our model focus on reactions in cholesterol metabolism, only the genes in the category ‘Lipid metabolism’ were considered. These genes are HMGCR, FDFT1, CYP7A1, LDLR, APOB, APOE, APOA-I, CETP, ABCG8, PON1, and LPL (15). With the exception of PON1, these genes were generally associated with the reactions with a significant reaction effect in the model. The importance of the gene PON1 could not be tested, because the associated reaction is not known (30).

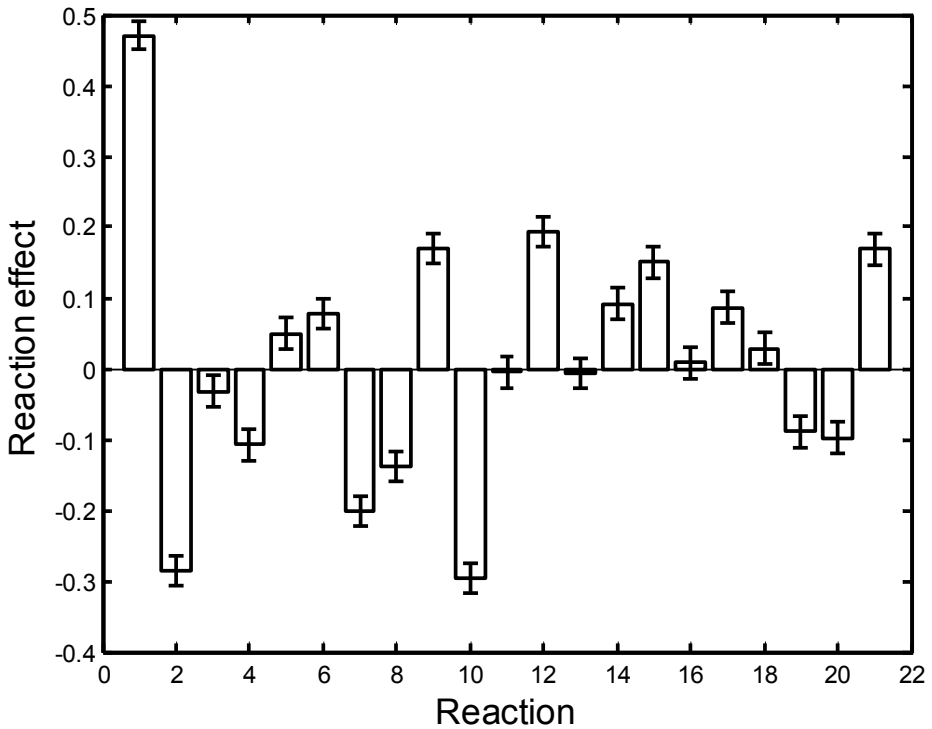


Figure 5.10. Reaction effects of all 21 reactions in determining the HDL-C response to pravastatin and their 95% confidence interval for the response. Reaction numbers correspond to numbers in Figure 5.1.

Schmitz and Langmann (15) also reported that a variant in CYP7A1 (reaction 18) and a variant in ABCG5/G8 (reaction 14) interact with respect to their influence on the LDL-C lowering response to atorvastatin. This means that the variant in CYP7A1 has a different effect in the presence of the variant of ABCG5/G8, than in the absence of that variant of ABCG5/G8 (31). We investigated the interaction of the parameters on the associated reactions (see Methods). The interaction of reactions 14 and 18 was significant (95% CI of β_{int} was 0.0004 to 0.0174). Thus our model also reproduced this interaction.

Discussion

We first investigated how accurate our cholesterol PBK model predicted the effects of statin treatment, then we investigated which reactions included in the model determine the variability in the response to statins. After adaptation to include LDLR upregulation in response to hepatic cholesterol concentration, the PBK model provided accurate predictions for pravastatin modulation of HDL-C, TC, and non-HDL-C (Figure 5.6). Other statins differ from pravastatin in their absorption, distribution, metabolism, and excretion (7), but share the same mechanism of LDL-C lowering. This suggests that the developed cholesterol PBK model is able to make predictions for other statins as well, provided that the relation between statin dose and rate of cholesterol synthesis in the liver is known. The cholesterol PBK model also contains reactions that are the targets of other drugs, like CETP inhibitors, and cholesterol uptake inhibitors. Therefore, the cholesterol PBK model can be used as a basis to predict responses to other cholesterol modulating drugs. This is, however, beyond the scope of the present study and will be a topic for future research.

There is a large interindividual variation in the response to statins, some hyper-responders show a non-HDL-C response of 4 mM (29), while other subjects have almost no response. This variation can be caused by variations in statin metabolism, compliance, nutrition, and cholesterol metabolism (15,32,33). Watanabe et al. (17) studied the first category and found that variations in the statin transporter OATP1B1 might have an effect on statin efficacy. Further information on the second and third category can be found in literature (32,33), whereas our model analyzed the last category of variation: the variation caused by variation in cholesterol metabolism.

Our model predicted that the rates of hepatic and peripheral cholesterol synthesis (reactions 1 and 2), and the rate of hepatic cholesterol esterification (reaction 19) were determinants of the non-HDL-C reduction by statins (Figure 5.9). It is, therefore, suggested to test biomarkers correlated with these reaction rates for their power to predict individual statin response.

Reaction effects are also dependent on which subjects are selected and thus on the inclusion criteria of the study (data not shown). As several studies use patients with Familial Hypercholesterolemia (FH) to study determinants in statin response (reviewed in (15)), it should be taken into account that findings for subgroups in patients with FH might not apply for the general population.

Finally, our model was able to reproduce the finding that variants in CYP7A1 and variants in ABCG5/G8 interact with respect to their influence on the LDL-C lowering response to atorvastatin (31). There is growing evidence that gene-gene interactions are ubiquitous in determining biological properties, such as the susceptibility to common human diseases (34). Identifying these gene-gene interactions is difficult as the number of potential interactions in the genome can be large (34). Our model can be used to improve the discovery of novel gene-gene interactions by predicting the most influential interactions between reactions. Testing for gene-gene-interactions in statin trials can subsequently be targeted to genes important in the interacting reactions in the model. As a result less statistical tests have to be performed, leading to more statistical power (34).

In conclusion, we have developed a model that is able to accurately predict the effect of pravastatin treatment and can be useful to provide insight in the individual response to statins.

Acknowledgments

We thank dr. Sieto Bosgra for assistance with the pravastatin PBK model.

References

1. Hegele, R. A. (2009) *Nat. Rev. Genet.* 10, 109-121
2. Lusis, A. J. (2000) *Nature* 407, 233-241
3. Glassberg, H. and Rader, D. J. (2008) *Annu. Rev. Med.* 59, 79-94
4. Lewis, G. F. and Rader, D. J. (2005) *Circ. Res.* 96, 1221-1232
5. Baigent, C., Blackwell, L., Emberson, J., Holland, L. E., Reith, C., Bhala, N., Peto, R., Barnes, E. H., Keech, A., Simes, J., and Collins, R. (2010) *Lancet* 376, 1670-1681
6. Armitage, J., Bowman, L., Wallendszus, K., Bulbulia, R., Rahimi, K., Haynes, R., Parish, S., Peto, R., and Collins, R. (2010) *Lancet* 376, 1658-1669
7. Davidson, M. H. and Toth, P. P. (2004) *Prog. Cardiovasc. Dis.* 47, 73-104
8. Endo, A. (1992) *J. Lipid Res.* 33, 1569-1582
9. van Vliet, A. K., van Thiel, G. C., Huisman, R. H., Moshage, H., Yap, S. H., and Cohen, L. H. (1995) *Biochim. Biophys. Acta* 1254, 105-111

10. Thavendiranathan, P., Bagai, A., Brookhart, M. A., and Choudhry, N. K. (2006) *Arch. Intern. Med.* 166, 2307-2313
11. Miettinen, T. A. and Gylling, H. (2002) *Atherosclerosis* 164, 147-152
12. Pazzucconi, F., Dorigotti, F., Gianfranceschi, G., Campagnoli, G., Sirtori, M., Franceschini, G., and Sirtori, C. R. (1995) *Atherosclerosis* 117, 189-198
13. Pearson, T. A., Laurora, I., Chu, H., and Kafonek, S. (2000) *Arch. Intern. Med.* 160, 459-467
14. Hoenig, M. R., Walker, P. J., Gurnsey, C., Beadle, K., and Johnson, L. (2010) *J. Cardiovasc. Pharmacol.* 56, 396-401
15. Schmitz, G. and Langmann, T. (2006) *Vascul. Pharmacol.* 44, 75-89
16. Voora, D., Shah, S. H., Reed, C. R., Zhai, J., Crosslin, D. R., Messer, C., Salisbury, B. A., and Ginsburg, G. S. (2008) *Circ. Cardiovasc. Genet.* 1, 100-106
17. Watanabe, T., Kusuhara, H., Maeda, K., Shitara, Y., and Sugiyama, Y. (2009) *J. Pharmacol. Exp. Ther.* 328, 652-662
18. van de Pas, N. C., Woutersen, R. A., van Ommen, B., Rietjens, I. M., and de Graaf, A. A. (2011) *Biochim. Biophys. Acta* 1811, 333-342
19. van de Pas, N. C., Woutersen, R. A., van Ommen, B., Rietjens, I. M. C. M., and de Graaf, A. A. (2011) submitted
20. The MathWorks Inc. (2007) MATLAB.
21. Petrie, A. and Sabin, C. (2000) *Medical Statistics at a Glance*, Blackwell Science Ltd,
22. van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006) *BMC. Genomics* 7, 142
23. Marchini, J., Donnelly, P., and Cardon, L. R. (2005) *Nat. Genet.* 37, 413-417
24. Chen, Y., Ruan, X. Z., Li, Q., Huang, A., Moorhead, J. F., Powis, S. H., and Varghese, Z. (2007) *Am. J. Physiol Renal Physiol* 293, F680-F687
25. Wilcox, L. J., Barrett, P. H., and Huff, M. W. (1999) *J. Lipid Res.* 40, 1078-1089
26. Scharnagl, H., Schinker, R., Gierens, H., Nauck, M., Wieland, H., and Marz, W. (2001) *Biochem. Pharmacol.* 62, 1545-1555
27. Jones, P. H., Davidson, M. H., Stein, E. A., Bays, H. E., McKenney, J. M., Miller, E.,

- Cain, V. A., and Blasetto, J. W. (2003) *Am. J. Cardiol.* 92, 152-160
28. Hillebrant, C. G., Nyberg, B., Gustafsson, U., Sahlin, S., Bjorkhem, I., Rudling, M., and Einarsson, C. (2002) *Eur. J. Clin. Invest* 32, 528-534
29. Thompson, J. F., Hyde, C. L., Wood, L. S., Paciga, S. A., Hinds, D. A., Cox, D. R., Hovingh, G. K., and Kastelein, J. J. (2009) *Circ. Cardiovasc. Genet.* 2, 173-181
30. van Himbergen, T. M., van Tits, L. J., Roest, M., and Stalenhoef, A. F. (2006) *Neth. J. Med.* 64, 34-38
31. Kajinami, K., Brousseau, M. E., Ordovas, J. M., and Schaefer, E. J. (2004) *Atherosclerosis* 175, 287-293
32. Schultz, J. S., O'Donnell, J. C., McDonough, K. L., Sasane, R., and Meyer, J. (2005) *Am. J. Manag. Care* 11, 306-312
33. Dahan, A. and Altman, H. (2004) *Eur. J. Clin. Nutr.* 58, 1-9
34. Briollais, L., Wang, Y., Rajendram, I., Onay, V., Shi, E., Knight, J., and Ozcelik, H. (2007) *BMC. Med.* 5, 22
35. van de Pas, N. C., Soffers, A. E., Freidig, A. P., van Ommen B., Woutersen, R. A., Rietjens, I. M., and de Graaf, A. A. (2010) *Biochim. Biophys. Acta* 1801, 646-654



CHAPTER

6

Predicting individual response to CETP inhibitors using a kinetic model for cholesterol concentrations.

Niek C.A. van de Pas, Ruud A. Woutersen, Ben van Ommen, Ivonne M.C.M. Rietjens, and Albert A. de Graaf

In preparation



Abstract

Differences between individuals in the response of plasma cholesterol concentrations to the CETP inhibiting drug torcetrapib are highly relevant for the interpretation of recent clinical trials and may impact therapeutic application. The objective of this study was to identify factors that contribute to this variation in torcetrapib response, applying a physiologically based kinetic (PBK) model for cholesterol plasma concentrations in humans.

The model predicted the effect of several doses of torcetrapib treatment (0 - 240 mg/d) on Total Cholesterol (TC), High Density Lipoprotein-Cholesterol (HDL-C), and non-HDL-C. The model correctly predicted an unchanged TC concentration, an increased HDL-C concentration, and a decreased non-HDL-C concentration. At 240 mg/d, HDL-C was predicted to increase by 79% (vs. 92% reported for human subjects in literature).

With our PBK model, differences in response between subjects were studied using 4,000 virtual subjects, each subject having a unique combination of rate constants for the different reactions included in the PBK model.

The model predicted that hepatic HDL-CE uptake, HDL associated cholesterol esterification, and CE transfer from HDL to non-HDL were the reactions that most strongly effected the HDL-C increase upon treatment with torcetrapib. Further analysis indicated that these reactions were also important for response to torcetrapib, when given on top of a statin.

Introduction

Several drugs have been developed to reduce the risk of cardiovascular diseases (CVD)(1). These drugs generally affect known risk factors, such as blood pressure, low density lipoprotein cholesterol (LDL-C), and high density lipoprotein cholesterol (HDL-C). The best known drugs prescribed to improve cholesterol profiles are 3-hydroxy-3-methyl-glutaryl-CoA reductase inhibitors, also known as statins. Statins decrease LDL-C, increase HDL-C, and decrease CVD risk (1). Another class of drugs that affect LDL-C and HDL-C is the class of cholesterol ester transfer protein (CETP) inhibitors. These inhibitors were developed with the objective to increase HDL-C by inhibiting the transport of cholesterol esters (CE) from HDL to triglyceride rich lipoproteins like LDL (2). The best known drug in this class is torcetrapib, which increases HDL-C and decreases LDL-C compared to a placebo (3). Currently, other CETP inhibitors are in development (4).

Recently, however, due to off-target effects, torcetrapib failed in clinical trials, even though it increased HDL-C (5). Treatment with torcetrapib plus a statin (“on top of a statin”) resulted in an increased risk of mortality and morbidity compared to a statin alone (5).

For many drugs, 50-75% of subjects get the desired response (6), they are called “responders”. Others are “non-responders”. Identification or prediction which subjects will not respond to drug treatment, can help to develop and analyze clinical trials and tailor drug treatment to the right patient. In order to discriminate responders from non-responders in advance, better insight is required in the factors involved in the response to CETP inhibitors. As cholesterol concentrations are quantified on a continuous scale, we will talk about “hyper-responders” and “hypo-responders” instead of responders and non-responders.

Previously, we have shown that our previously developed computational model for plasma cholesterol concentrations in humans (7) was able to identify factors that discriminate hyper-responders from hypo-responders to statins in terms of their HDL-C increase and non-HDL-C decrease (8). Figure 6.1 shows the conceptual model used to set up this computational model. The model describes 8 different cholesterol pools in the body influenced by 21 metabolic and transport reactions. The cholesterol pools are hepatic free cholesterol, hepatic cholesterol ester, peripheral cholesterol, intestinal free cholesterol, intestinal cholesterol ester, and also HDL free cholesterol (HDL-FC), HDL cholesterol ester (HDL-CE), and non-HDL-C (i.e., the total of plasma cholesterol that is not present in HDL). Reaction

21 denotes CE transfer from HDL to LDL, the reaction catalyzed by CETP.

In the present study, we aimed to obtain insights into the factors that determine the response to torcetrapib. First we investigated whether the model correctly predicted the effects of torcetrapib monotherapy, then we studied the effects of torcetrapib when applied on top of a statin, and compared the outcomes to those obtained when torcetrapib was applied in monotherapy.

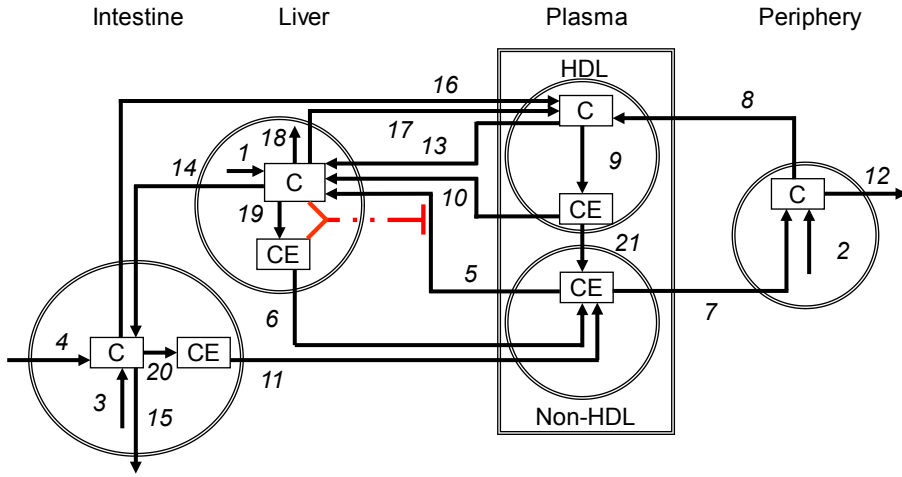


Figure 6.1. Conceptual model for pathways determining cholesterol plasma concentrations used as a basis to set up PBK cholesterol model of the present study. Process numbers stand for: 1, Hepatic cholesterol synthesis; 2, Peripheral cholesterol synthesis; 3, Intestinal cholesterol synthesis; 4, Dietary cholesterol intake; 5, Hepatic uptake of cholesterol from non-HDL; 6, Hepatic Very Low Density lipoprotein cholesterol (VLDL-C) production; 7, Peripheral uptake of cholesterol from non-HDL; 8, Peripheral cholesterol transport to HDL; 9, HDL associated cholesterol esterification; 10, Hepatic HDL-CE uptake; 11, Intestinal chylomicron cholesterol secretion; 12, Peripheral cholesterol loss; 13, Hepatic HDL-FC uptake; 14, Biliary cholesterol excretion; 15, Fecal cholesterol excretion; 16, Intestinal cholesterol transport to HDL; 17, Hepatic cholesterol transport to HDL; 18, Hepatic cholesterol catabolism; 19, Hepatic cholesterol esterification; 20, Intestinal cholesterol esterification, and 21, CE transfer from HDL to non-HDL. C stands for cholesterol; CE for cholesterol ester. Taken from (8). Dashed line indicates the regulation of the LDL Receptor (LDLR) in response to the hepatic cholesterol level. Process 21 is targeted by CETP inhibitors.

Finally, an analysis was performed to identify the factors that discriminate hyper-responders from hypo-responders to torcetrapib. As other CETP inhibitors, which are still in clinical trials, share the same target reaction in the model (reaction 21), insights in the factors involved in the effects with respect to torcetrapib are directly relevant for other CETP inhibitors, like dalcetrapib or anacetrapib (9).

Methods

Model description

The cholesterol PBK model for a standard subject (“reference man” as defined by the International Commission on Radiation Protection) was described previously (7), in short: the model consists of a set of 8 differential equations each describing the dynamics of one of the 8 cholesterol pools. These equations describe the effect of 21 metabolic and transport reactions on the size of the 8 included cholesterol pools. The rate of each reaction is described using a rate equation with a single parameter: the rate constant. The steady state cholesterol concentrations calculated by the model are referred to as the predicted concentrations.

The model consists of a set (ensemble) of 8 submodels, each having a different combination of first and zero order kinetics for the various reactions, instead of a single optimal model. The model prediction is calculated as the average of the predictions of the 8 submodels (7).

Model simulations

The effect of torcetrapib was simulated by reducing the rate constant of the reaction of CE transfer from HDL to non-HDL-C (reaction 21) according to:

$$k^{treated} = f_{act} \cdot k^{untreated} \quad (\text{Eqn. 6.1})$$

where k is the rate constant for the targeted reaction, either in standard condition (superscript: untreated) or after drug treatment (superscript: treated), and f_{act} is the reduction factor, determined as described in the Result section.

The effect of pravastatin was simulated by reducing the rate constant of the reaction for hepatic cholesterol synthesis (reaction 1, (7)) as in Eqn. 6.1. For 40 mg/d pravastatin, this f_{act} had a value of 0.65 (8).

Treatment with both statins and torcetrapib was simulated by applying Eqn. 6.1 to both the rate constant for reaction 21 and reaction 1 each

with their respective f_{act} .

The model predictions were obtained by running the model with the new set of parameters (including k^{treated} instead of $k^{\text{untreated}}$) by numerical integration with the concentrations of the reference man (7) as initial values. The simulation was performed until steady state of all cholesterol pools in the model was reached. Model predictions were defined as these steady state concentrations (7). Integration was performed using routine *ode15s* as implemented in MATLAB version 7.5 (R2007b).

Virtual subjects

To determine which reactions are important for the response to torcetrapib, a group of 4,000 virtual subjects was constructed, each differing in the values of the kinetic parameters in the 21 reactions. Each kinetic parameter was drawn randomly from a normal distribution with a specified average and standard deviation. The average was set to the rate constant in the standard subject and the standard deviation was set to 25% of this average. The model was used to calculate a baseline level and then treatment was simulated as described above (in the section on Model simulations), using the relevant fixed f_{act} values. The model was used to calculate baseline concentrations and concentrations after treatment with 1) pravastatin, 2) torcetrapib, and 3) the combination of both pravastatin and torcetrapib. For some virtual subjects, one or more of the 8 submodels predicted one or more negative concentrations (even if all rate constants were positive). These subjects were not taken into account, just as subjects with a baseline non-HDL-C larger than 50 mM, because these values were considered unphysiological (this deviates more than 20 SDs from the mean LDL-C concentration as found by Carroll et al. (10)).

Reaction importance

The relation between the response and the reaction rate constant for each of the 21 reactions (all continuous variables) was assessed in the virtual subjects group using linear regression (11):

$$y_n = \beta_0 + \beta_1 \tilde{k}_{n,i} + \varepsilon_n \quad (\text{Eqn. 6.2})$$

where y_n is the response to the drug for subject n (expressed relative to baseline), β_0 , and β_1 are the regression coefficients fitted to experimental data by minimizing the sum of the squared error terms (ε_n), and $\tilde{k}_{n,i}$ is the autoscaled rate constant of subject n for reaction i . This autoscaling (12)

was applied to enable the comparison of the fitted regression coefficients for different reactions.

The term β_1 in Eqn. 6.2 is called the reaction effect (8). Regression was performed using the function `regress` as implemented in MATLAB. Because responses were not normally distributed, Johnson correction was applied using the function `johnsrand` as implemented in MATLAB (13).

Results

Torcetrapib simulation

We aimed to obtain insight in the differences between hyper-responders and hypo-responders to CETP inhibitors, like torcetrapib. First it was investigated whether the model correctly predicted the effect of torcetrapib on cholesterol concentrations. In a next step the effect of torcetrapib on top of a statin treatment was investigated using pravastatin as model statin, being a statin for which we previously demonstrated that model predictions were close to clinical data (8).

Model predictions on the effect of torcetrapib monotherapy made in the present study were compared to data from Clark et al. (3), who reported a multidose phase 1 trial with 5 doses of torcetrapib (0 - 240 mg/d) and described the resulting plasma cholesterol concentrations as well as the inhibition of CETP activity for 4 dose groups.

CETP inhibitors block the activity of CETP that catalyzes CE transfer from HDL to non-HDL (reaction 21, Figure 6.1). Table 6.1 shows the f_{act} values (See Eqn. 6.1) derived from the data on CETP activity in the treated condition relative to the untreated case for the 4 doses of torcetrapib in Clark et al. (3).

148

Table 6.1. Derivation of the parameter f_{act} based on CETP inhibition measured in plasma samples by Clark et al. (3).

| | | | | | |
|-------------------|------|------|------|------|------------------|
| Dose (mg/d) | 0 | 30 | 60 | 120 | 240 [#] |
| CETP inhibition % | 0 | 12 | 35 | 53 | 80 |
| f_{act} | 1.00 | 0.88 | 0.65 | 0.47 | 0.20 |

[#] 120 mg/d twice daily

These f_{act} values were used in the model to predict the TC, HDL-C, and non-HDL-C concentrations in response to torcetrapib applied in the doses described. Figure 6.2 presents these model predictions. The model correctly predicted a decrease in non-HDL-C and a large increase in HDL-C upon torcetrapib treatment, whereas also the absence of an effect on TC was

correctly predicted. At the highest dose, HDL-C was predicted to increase by 79%, while a value of 91% was measured. This model prediction is definitely within the experimental error given in Clark et al. (standard deviation of the relative effect is generally >50% of the average effect (3)), and thus we concluded that the model correctly predicted the effect of torcetrapib on HDL-C, non-HDL-C, and TC.

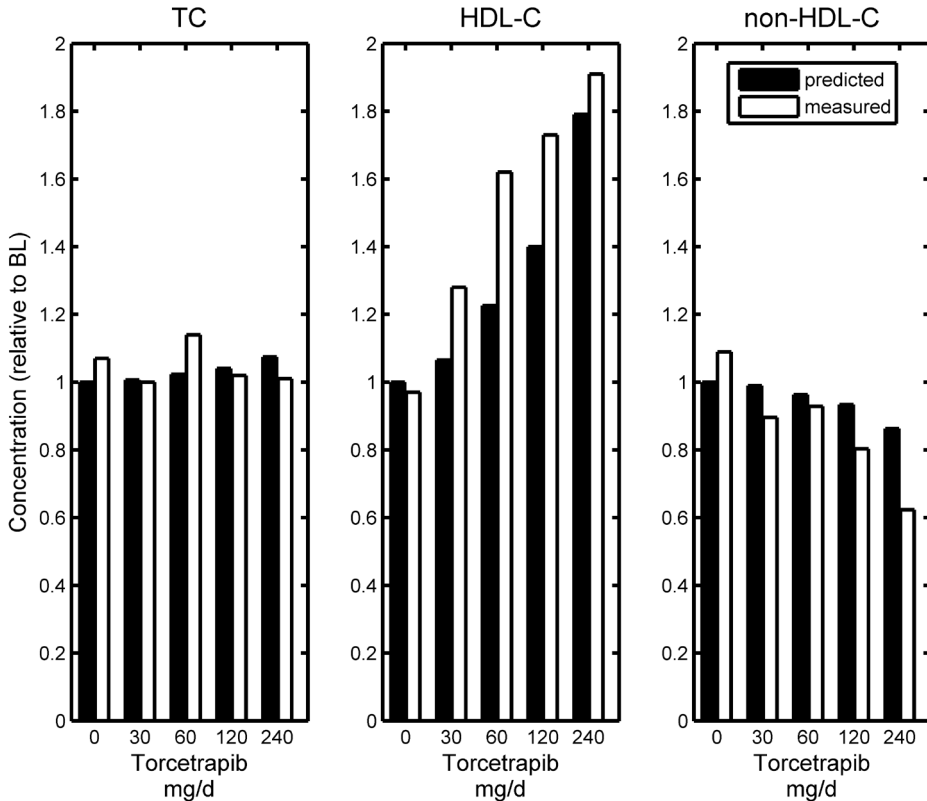


Figure 6.2. Model predictions (predicted) and literature data (measured), as reported in Clark et al. (3) of the effect of torcetrapib treatment (0 - 240 mg/d) on TC, HDL-C, and non-HDL-C concentrations. All concentrations are reported relative to baseline (BL).

In clinical trials, adding torcetrapib on top of a statin, additionally increased HDL-C compared to a statin alone (14). Figure 6.3 shows that an additional HDL-C increase by torcetrapib of 21% was observed for the model predictions (not shown: from 1.31 mM for statin alone to 1.59 mM for statin plus torcetrapib) vs. 55% in experimental data (not shown: from 1.36 mM for

statin alone to 2.11 mM for statin plus torcetrapib). The model also predicted the observed additional decrease in non-HDL-C, and unchanged TC level compared to statin alone (See Figure 6.3).

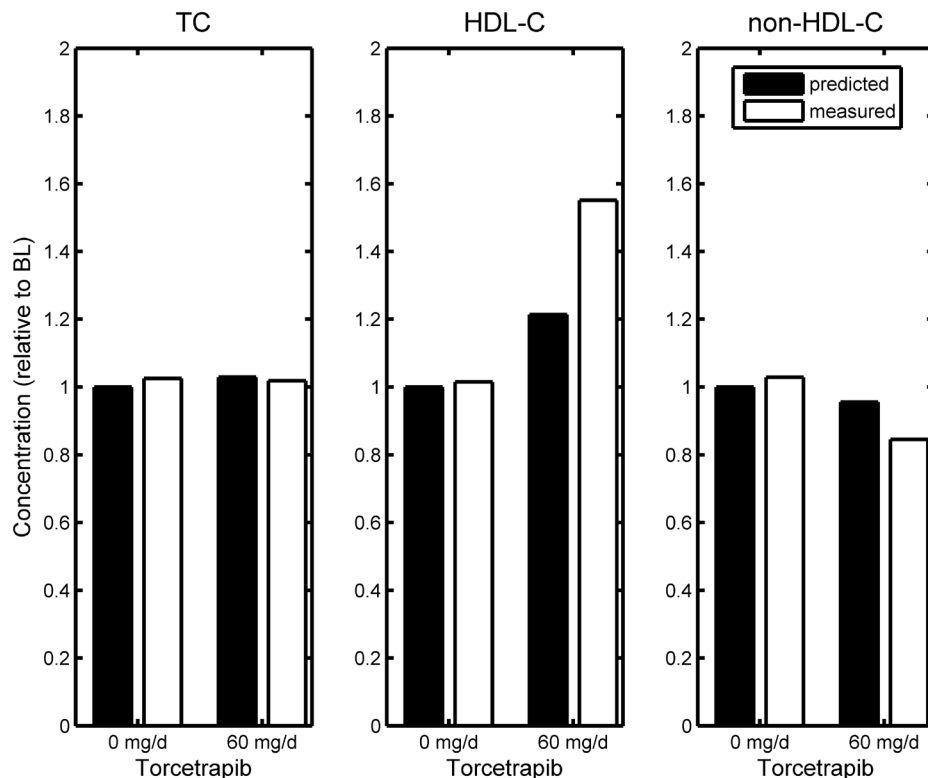


Figure 6.3. Model predictions (predicted) and literature data (measured), as reported in Kastelein et al.(14) of the effect of torcetrapib treatment (60 mg/d) on top of a statin on TC, HDL-C, and non-HDL-C concentrations. All concentrations are reported relative to baseline (BL).

Virtual trial

An analysis was performed to identify the factors that discriminate hyper-responders from hypo-responders to torcetrapib. A group of 4,000 virtual subjects was constructed. For each subject 4 predictions were obtained for HDL-C, non-HDL-C, and TC: 1) at baseline, 2) after 40 mg/d pravastatin, 3) after 60 mg/d torcetrapib, and 4) after the combination of 40 mg/d pravastatin and 60 mg/d torcetrapib.

In this simulation, 958 subjects had one or more negative concentrations or had unphysiologically high LDL-C concentrations (> 50 mM) and were, therefore, not taken into account. This resulted in 3,041 (76%) useful virtual subjects. These 3,041 subjects had baseline concentrations of 7.03 ± 4.14 mM for TC, 1.23 ± 0.43 mM for HDL-C, and 5.80 ± 4.35 mM for non-HDL-C (average \pm SD).

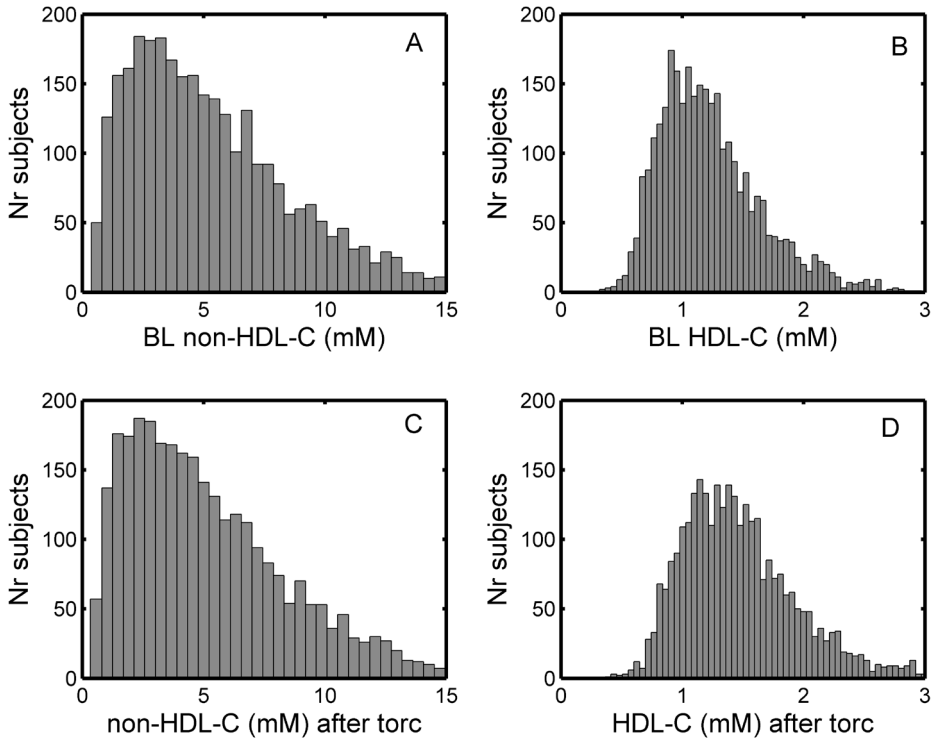


Figure 6.4. Histograms of the predicted non-HDL-C (Figure 6.4A) and HDL-C (Figure 6.4B) concentrations at baseline (BL) and after 60 mg/d torcetrapib (torc) treatment (Figures 4C and 4D). X-axes cutoffs have been adjusted for clarity, therefore, 97.4% of the virtual subjects were included in the graph for HDL-C and 98.6% of the virtual subjects were included in the graph for non-HDL-C.

Figure 6.4 shows histograms of the distribution in HDL-C and non-HDL-C concentrations before and after torcetrapib treatment. It can be seen that plasma cholesterol concentrations are not normally distributed. In addition, subtle shifts in distributions upon torcetrapib treatment can be observed. Torcetrapib treatment was predicted to increase HDL-C by $20 \pm 6\%$ (mean \pm SD) and to slightly decrease non-HDL-C.

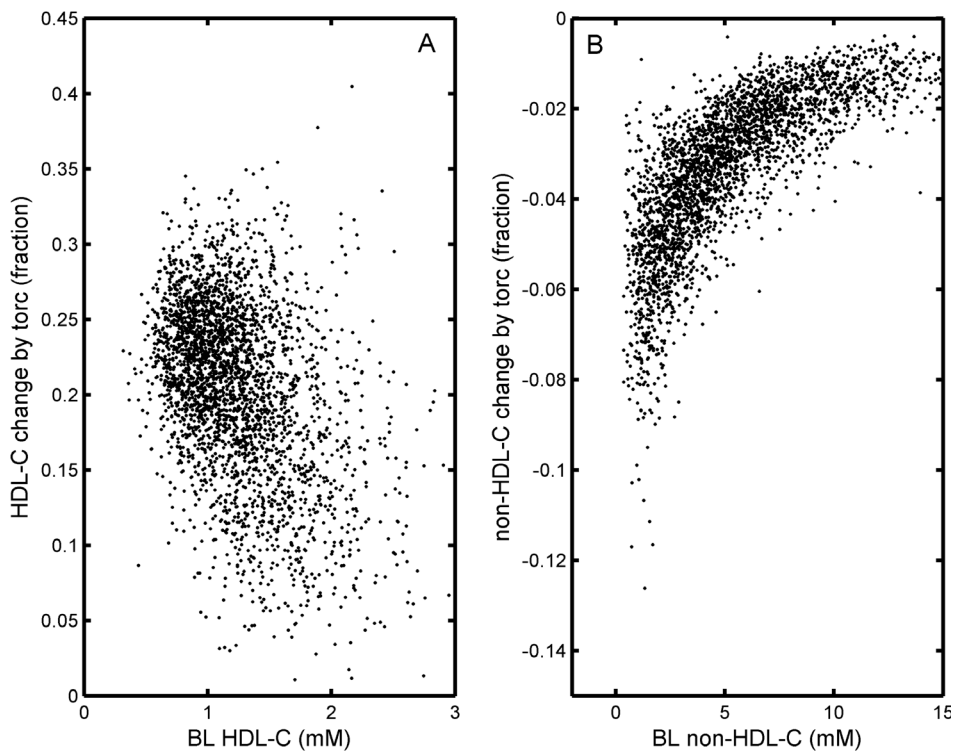


Figure 6.5. Predicted responses to 60 mg/d torcetrapib (torc) vs. baseline (BL) for A) HDL-C and B) non-HDL-C. X and y-axis cutoffs have been adjusted for clarity, therefore, 97.4% (for Figure 6.5A) and 98.6% (for Figure 6.5B) of the virtual subjects were included in the graphs.

152

In a previous study the effect of pravastatin on non-HDL-C was shown to be dependent on the non-HDL-C baseline concentrations (8;15). For torcetrapib, this relation was also studied (Figure 6.5). Subjects with a high baseline non-HDL-C level have a smaller relative reduction in non-HDL-C than subjects with a low baseline non-HDL-C level (Figure 6.5B). The relation between baseline concentrations and effect size is more diffuse for HDL-C (Figure 6.5A) compared to non-HDL-C (Figure 6.5B).

The impact of a reaction on the HDL-C increase by torcetrapib was studied by calculating the “reaction effect” (i.e. β_1 in Eqn. 6.2). If no correlation occurred between the rate constant of a particular reaction and the torcetrapib response, then the reaction effect was zero, implying that the re-

action was not important for torcetrapib response. In Figure 6.6, we report the reaction effects (i.e. β_1 in Eqn. 6.2). The three reactions with the largest (positive and negative) effects in Figure 6.6 were found to be hepatic HDL-CE uptake (reaction 10), HDL associated cholesterol esterification (reaction 9), and CE transfer from HDL to non-HDL (reaction 21).

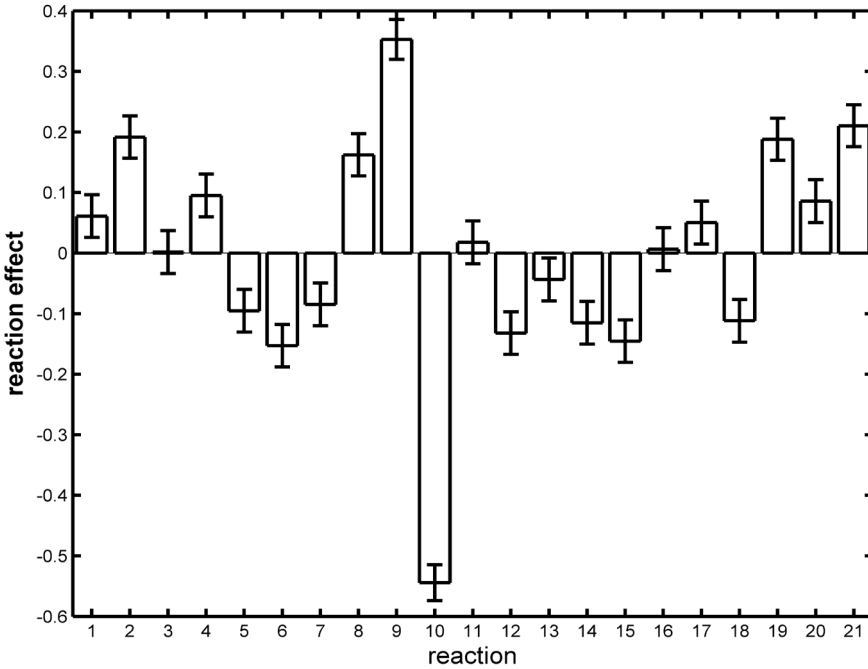


Figure 6.6. Reaction effects of all 21 reactions in determining the HDL-C response to torcetrapib and their 95% confidence interval. Reaction numbers correspond to numbers in Figure 6.1.

A similar procedure was performed for non-HDL-C and the resulting reaction effects are shown in Figure 6.7. The three reactions with the largest (positive and negative) effects in Figure 6.7 were found to be hepatic cholesterol catabolism (reaction 18), hepatic cholesterol esterification (reaction 19), and fecal cholesterol excretion (reaction 15).

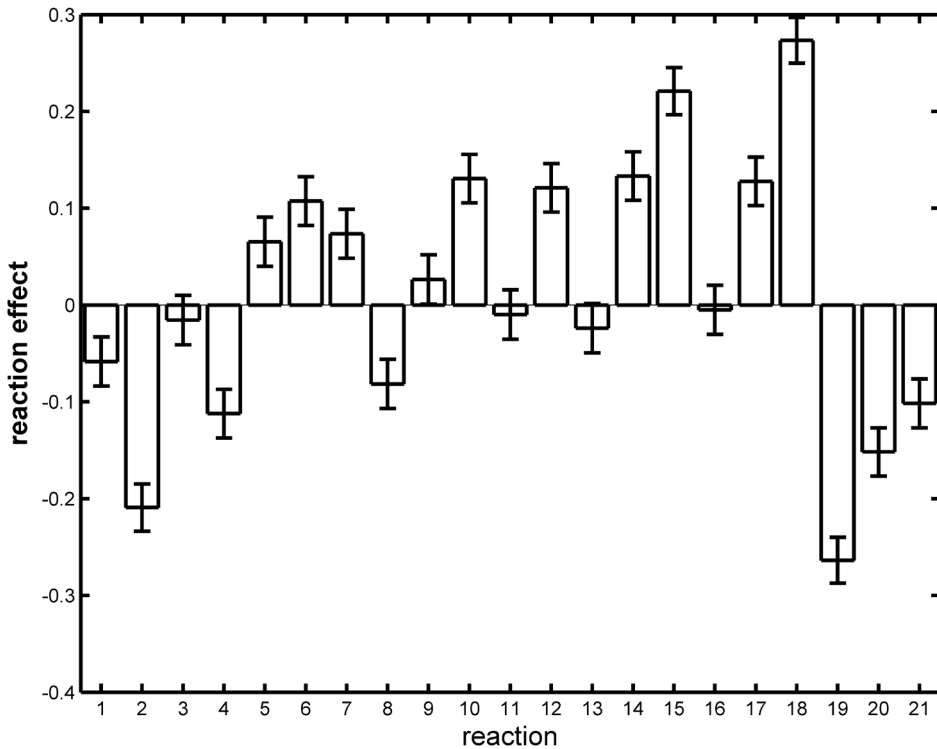


Figure 6.7. Reaction effects of all 21 reactions in determining the non-HDL-C response to torcetrapib and their 95% confidence interval. Reaction numbers correspond to numbers in Figure 6.1.

In clinical trials, torcetrapib and other CETP inhibitors were generally not only studied in monotherapy, but on top of a statin. The question is whether or not the important reactions for torcetrapib response (Figures 6.6 and 6.7) are also important when torcetrapib was given on top of a statin. In other words: are the hyper-responding virtual subjects on torcetrapib monotherapy also hyper-responders to torcetrapib treatment on top of a statin therapy? To answer this question, Figure 6.8 shows the effect of torcetrapib on top of pravastatin vs. the effect of torcetrapib monotherapy for HDL-C (Figure 6.8A) and non-HDL-C (Figure 6.8B). For HDL-C, the effect of torcetrapib added to statin therapy was generally smaller (lower positive change) than the effect of torcetrapib added in monotherapy (Figure 6.8A). For non-HDL-C the effect of torcetrapib added to statin therapy was generally larger (larger negative change) than the effect of torcetrapib added in

monotherapy (Figure 6.8B). This is a sign of a slight drug-drug interaction. However given that this interaction is small compared with the individual variation in response, these drug-drug interactions are almost certainly not clinically significant.

From Figure 6.8 it can also be seen that there is a clear relation between the effect of monotherapy and the effect of torcetrapib on top of a statin (linear regression: $R^2 > 0.95$ for HDL-C and for non-HDL-C). Therefore, the determinants of torcetrapib response (Figures 6.6 and 6.7) are also valid when torcetrapib is given on top of a statin.

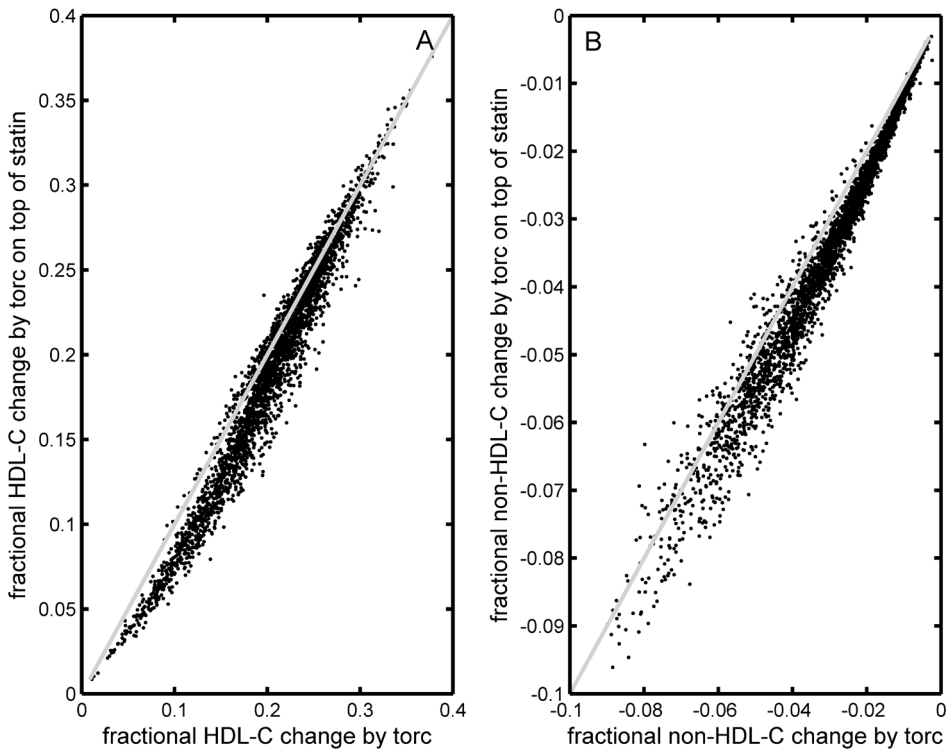


Figure 6.8. Effect of 40 mg/d torcetrapib on top of pravastatin (60 mg/d) vs. the effect of torcetrapib monotherapy for HDL-C (Figure 6.8A) and non-HDL-C (Figure 6.8B). The line of identity is also given for comparison.

Discussion

Our model correctly predicted the effect of torcetrapib on HDL-C and TC in a quantitative way and also predicted the observed decrease in non-HDL-C. As shown for pravastatin in (8), this model also allows the prediction of the effect of torcetrapib on other aspects of cholesterol metabolism, like hepatic cholesterol concentrations and time required to see an effect (8)(data not shown).

With the presented model, we have identified several factors that determine the response to torcetrapib (Figures 6.6 and 6.7). We found that hepatic HDL-CE uptake (reaction 10), HDL associated cholesterol esterification (reaction 9), and CE transfer from HDL to non-HDL (reaction 21) had the largest effect on HDL-C increase upon treatment with torcetrapib (Figure 6.6). The activity of the last two reactions (reactions 9 and 21) can be determined based on analysis of plasma samples, providing a possible biomarker for torcetrapib response (5,14).

As a model prediction is always a simplification of reality, our model is lacking regulation mechanisms, like upregulation of CETP upon statin treatment. Therefore, before applying this knowledge into patient care, model predictions should be validated *in vivo*. For statins, we were able to perform a first validation by comparing our model predictions to polymorphisms known to be correlated with statin response (8). For torcetrapib, this is harder to perform, because to the best of our knowledge no SNPs are presently known to affect the response to torcetrapib (16). Therefore, we limited our validation to the average effect of torcetrapib on plasma cholesterol concentrations, both in monotherapy (Figure 6.2) and on top of a statin (Figure 6.3). Future research should include a validation of the proposed markers for the torcetrapib response (plasma tests of the activity of reactions 9 and 21). This can be done with a posthoc analysis of a previously performed torcetrapib trial. Subjects in this trial can be divided into multiple groups according to their biomarker levels before treatment, followed by an assessment of the efficacy of torcetrapib in the separate subgroups.

Other CETP inhibitors, which are investigated in clinical trials (dalcetrapib and anacetrapib), differ from torcetrapib in their molecular structure and potency (9). But, since they share the same target reaction in our model (reaction 21, Figure 6.1), insights with respect to torcetrapib are directly relevant for the other CETP inhibitors as well and can help to set-up clinical trials and interpret data from clinical trials with these compounds. We conclude that our PBK model of cholesterol metabolism is able to predict effects of torcetrapib treatment. Additionally, markers were identified

of several reactions that determine the individual responsiveness to torcetrapib in humans.

References

1. Davidson, M. H. and Toth, P. P. (2004) *Prog. Cardiovasc. Dis.* 47, 73-104
2. Morton, R. E. and Greene, D. J. (2007) *Atherosclerosis* 192, 100-107
3. Clark, R. W., Sutfin, T. A., Ruggeri, R. B., Willauer, A. T., Sugarman, E. D., Magnus-Aryitey, G., Cosgrove, P. G. and others (2004) *Arterioscler. Thromb. Vasc. Biol.* 24, 490-497
4. Zhao, L., Jin, W., Rader, D., Packard, C., and Feuerstein, G. (2009) *Biochem. Pharmacol.* 78, 315-325
5. Barter, P. J., Caulfield, M., Eriksson, M., Grundy, S. M., Kastelein, J. J., Komajda, M., Lopez-Sendon, J. and others (2007) *N. Engl. J. Med.* 357, 2109-2122
6. Spear, B. B., Heath-Chiozzi, M., and Huff, J. (2001) *Trends Mol. Med.* 7, 201-204
7. van de Pas, N. C., Woutersen, R. A., van Ommen, B., Rietjens, I. M. C. M., and de Graaf, A. A. (2011) submitted
8. van de Pas, N. C., Woutersen, R. A., van Ommen, B., Rietjens, I. M. C. M., and de Graaf, A. A. (2011) submitted
9. Ranalletta, M., Bierilo, K. K., Chen, Y., Milot, D., Chen, Q., Tung, E., Houde, C. and others (2010) *J. Lipid Res.* 51, 2739-2752
10. Carroll, M. D., Lacher, D. A., Sorlie, P. D., Cleeman, J. I., Gordon, D. J., Wolz, M., Grundy, S. M. and others (2005) *JAMA* 294, 1773-1781
11. Petrie, A. and Sabin, C. (2000) *Medical Statistics at a Glance*, Blackwell Science Ltd,
12. van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006) *BMC. Genomics* 7, 142
13. The MathWorks Inc. (2007) *MATLAB*
14. Kastelein, J. J., van Leuven, S. I., Burgess, L., Evans, G. W., Kuivenhoven, J. A., Barter, P. J., Revkin, J. H. and others (2007) *N. Engl. J. Med.* 356, 1620-1630
15. Thompson, J. F., Hyde, C. L., Wood, L. S., Paciga, S. A., Hinds, D. A., Cox, D. R., Hovingh, G. K. and others (2009) *Circ. Cardiovasc. Genet.* 2, 173-181
16. Schmitz, G. and Langmann, T. (2006) *Vascul. Pharmacol.* 44, 75-89



CHAPTER

7

Summary, general discussion,
and future perspectives.

Summary of the results

Elevated cholesterol is an important risk factor for cardiovascular diseases (CVD). The aim of this thesis was to develop a mathematical model, including the most important reactions that determine plasma cholesterol concentrations, that allows the prediction of the effects of genetic, pharmaceutical, and nutritional variations on these concentrations. In **Chapter 2** a conceptual model was systematically constructed that could serve as a basis for a mathematical model for cholesterol in the mouse. The model was based on the function of key genes that determine plasma cholesterol concentrations, and which were discriminated from less important genes based on the described phenotype of mouse knockout strains. The conceptual model contained 20 reactions that connect 8 cholesterol pools in plasma, liver, intestine, and peripheral organs.

In **Chapter 3** the conceptual model was converted to a PBK model to predict plasma cholesterol concentrations in the mouse. Kinetic parameters required to calibrate the model were obtained using data from published experiments. To construct the model, a set of appropriate submodels was selected from a set of 65,536 submodels differing in the kinetic expressions of the reactions. A submodel was considered appropriate if it had the ability to correctly predict an increased or decreased plasma cholesterol concentration for a training set of 5 knockout mouse strains. The model defined as such consisted of 8 appropriate submodels and was validated using data from an independent set of 9 knockout mouse strains.

The model accurately predicted in a quantitative way the plasma cholesterol concentrations of the 14 knockout strains, including the frequently used *Ldlr*^{-/-} and *ApoE*^{-/-} mouse strains.

In **Chapter 4** the PBK model for plasma cholesterol in the mouse was translationally modified into a model for humans by incorporation of the reaction catalyzed by Cholesterol Ester Transfer Protein (CETP). The adapted model contained 21 biochemical reactions and 8 different cholesterol pools. The model was calibrated using data from published kinetic studies and validated by comparing model predictions with experimental data on plasma cholesterol concentrations of subjects with ten different genetic mutations (including Familial Hypercholesterolemia).

In **Chapter 5** it was demonstrated that the model correctly predicted the effect of pravastatin treatment at various doses (0 - 40 mg/d). At 40 mg/d, Total plasma Cholesterol (TC) was predicted to decrease by 15% (vs. 22% reported for human subjects in literature), High Density Lipoprotein Cholesterol (HDL-C) was predicted to increase by 10% (vs. 5.6% in

literature), and non-HDL-C was predicted to decrease by 22% (vs. 25% in literature).

Additionally, based on a virtual study population, the model also predicted that rates of hepatic cholesterol synthesis, peripheral cholesterol synthesis, hepatic cholesterol esterification, and peripheral uptake of cholesterol from non-HDL discriminated hyper-responders to pravastatin from hypo-responders.

Chapter 6 describes the simulations of torcetrapib treatment. For several doses of torcetrapib, the model correctly predicted an unchanged TC concentration, an increased HDL-C concentration and a decreased non-HDL-C concentration. At 240 mg/d torcetrapib, HDL-C was predicted to increase by 79% (vs. 92% reported for human subjects in literature). Furthermore, based on a virtual study population we found that rates of hepatic HDL-CE uptake, HDL associated cholesterol esterification, and CE transfer from HDL to non-HDL determined the cholesterol response to torcetrapib treatment.

The present chapter will first discuss general modeling considerations. Then, four important questions in cholesterol research (defined in Chapter 1) will be discussed, including the role of the developed model in answering them. There also indications for future research are given. The last part of this chapter will contain general concluding remarks.

Modeling considerations

An early classical kinetic model on cholesterol used hypothetical cholesterol pools 'A' and 'B' (1), see Chapter 1. Because these cholesterol pools are hypothetical, it becomes difficult if not impossible to relate model predictions to physiological observations. Our model, on the other hand, is a physiologically based kinetic (PBK) model (2), wherein compartments link to real organs and reactions link to 21 physiologically relevant enzymatic conversions and transport reactions in cholesterol metabolism.

Moreover, models with hundreds of reactions are hard to interpret, even if these reactions are directly linked to physiology. Therefore, we chose to develop a simple model. Our model, with 21 reactions is simple enough to be interpreted and to be visualized as a conceptual model (Figure 7.1).

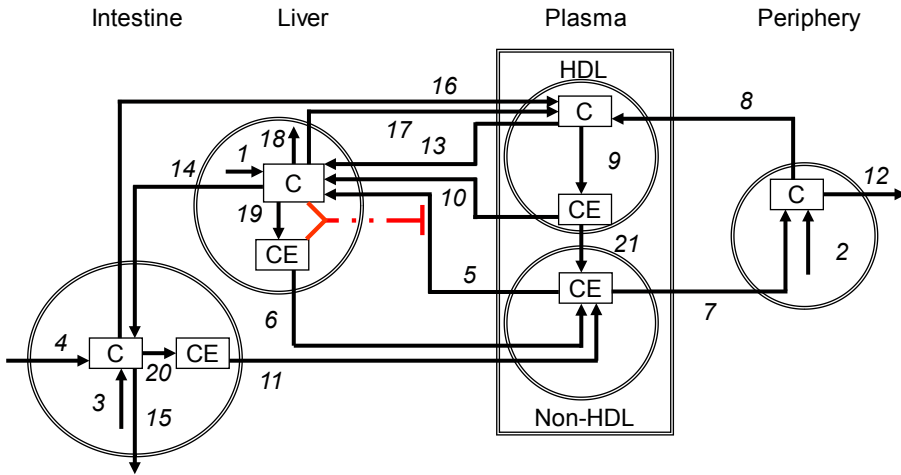


Figure 7.1. Conceptual model for pathways determining cholesterol plasma concentrations used as a basis to set up the PBK cholesterol model of the present study. Process numbers stand for: 1, Hepatic cholesterol synthesis; 2, Peripheral cholesterol synthesis; 3, Intestinal cholesterol synthesis; 4, Dietary cholesterol intake; 5, Hepatic uptake of cholesterol from non-HDL; 6, Hepatic Very Low Density lipoprotein cholesterol (VLDL-C) production; 7, Peripheral uptake of cholesterol from non-HDL; 8, Peripheral cholesterol transport to HDL; 9, HDL associated cholesterol esterification; 10, Hepatic HDL-CE uptake; 11, Intestinal chylomicron cholesterol secretion; 12, Peripheral cholesterol loss; 13, Hepatic HDL-FC uptake; 14, Biliary cholesterol excretion; 15, Fecal cholesterol excretion; 16, Intestinal cholesterol transport to HDL; 17, Hepatic cholesterol transport to HDL; 18, Hepatic cholesterol catabolism; 19, Hepatic cholesterol esterification; 20, Intestinal cholesterol esterification, and 21, CE transfer from HDL to non-HDL. C stands for cholesterol; CE for cholesterol ester. Adapted from (3). Dashed line indicates the regulation of the LDLR in response to the hepatic cholesterol concentration.

Our model, like any model is a simplification of reality and, therefore, lacks some aspects of reality. These simplifications are the result of assumptions that are not only made in defining the conceptual model (Chapter 2) and mathematical model formulation (Chapters 3 and 4), but also in the simulation of various scenarios. For example, for the simulation of mutations and interventions, assumptions had to be made on the definition of the factors f_{ko} (Chapter 3), f_{mut} (Chapter 4), or f_{act} (Chapter 5 and 6). This chapter will

not discuss all assumptions that were made, but will concentrate on two important ones.

The most important assumption in the model is that the reactions included in the model are targeted and limited to cholesterol. HDL-C and LDL-C are used as biomarkers of cardiovascular events (4-6), but recent clinical trials showed that certain therapies did have a beneficial effect on these cholesterol concentrations, but not on atherosclerotic plaque size or on mortality (7,8). This clearly demonstrates that HDL-C and LDL-C are imperfect biomarkers for cardiovascular events. HDL-C and LDL-C will continue to play an important role in the development of new therapies, particularly in the preclinical and early clinical stages (4), but it is important to keep in mind that the presented model predicts disease risk indirectly.

Another important assumption is made in the development of a mathematical model for humans (Chapter 4) based on the model for the mouse (Chapter 3). It was assumed that the kinetic format for the mouse model as described in Chapter 3 was appropriate for the human model. Although the mouse is frequently used as a model species in cholesterol research (9), it is known that there are differences between the cholesterol metabolism in man and in mouse (10). Nevertheless, this assumption is frequently made, also in drug development: in case of a lack of human data, it is assumed that humans resemble animals.

To test the validity of these and other model assumptions, the model was validated by comparing model predictions with independent experimental data. As predictions were comparable, the assumptions appeared to be valid. Therefore, the model is a useful tool to address current questions in cholesterol research. The next paragraphs describe four of these questions (See Chapter 1) and will also describe how the presented model can contribute in answering them.

Determinants of plasma cholesterol concentrations

What are the most important biochemical reactions in the body for determining plasma cholesterol concentrations and what is the relation between these reactions?

Presently, only 25–30% of the variance in plasma cholesterol concentrations that is determined by genetic influences could be attributed to specific genes (11). Accordingly a large gap (70-75%) still remains in our knowledge

on how cholesterol concentrations are determined. Chapter 4 describes an analysis of the reactions important for plasma cholesterol concentrations. It was found that the concentration of High Density Lipoprotein cholesterol (HDL-C) was mainly dependent on hepatic transport of cholesterol to HDL (reaction 17 Figure 7.1), CE transfer from HDL to non-HDL (reaction 21), and hepatic uptake of cholesterol from the non-HDL-C pool (reaction 5). The non-HDL-C concentration on the other hand, was mainly dependent on hepatic uptake of cholesterol from non-HDL (reaction 5) and hepatic cholesterol esterification (reaction 19).

Supporting evidence for these findings can be found in genetic studies like Genome Wide Association Studies (GWAS) (11,12), because the effect sizes of 19 SNPs (11) were in line with the sensitivity coefficients of their associated reactions (Chapter 4).

There are two views on the nature of the large unexplained variance in plasma cholesterol concentrations. One view is that rare genetic variants may be disproportionately important (13), and that part of the variance can be explained by taking more gene variants into account by using more data points per subject (for example using sequencing) and more subjects per study (14). The other view is that this approach will be unsuccessful, because genes do not function independently (15). This second view states that a large proportion of the variation is due to interactions, like gene-environment interactions or gene-gene interactions (15). The first type of interactions includes epigenetic mechanisms (16). The latter, the gene-gene interactions, also called epistasis, occur when the effect of one gene on a trait depends on another gene (15). In fact, gene-gene interactions seem important for various properties, such as the susceptibility to common human diseases (17).

To assess the impact of interactions between reaction rate constants on cholesterol concentrations in our model, we studied the group of virtual subjects defined in Chapter 5. In this group, we know that all the variation between subjects was caused by variation in 21 rate constants (that is the only variation that was included in the definition of these subjects).

We determined how much variation in the cholesterol concentrations could be explained by taking into account only the linear effects of the 21 reactions (without interactions). To this end, we used the Eqn. 7.1 (parallel to Chapter 5) that contains a constant term and a linear term:

$$y_n = \beta_C + \sum_{i=1}^{21} \beta_{L_i} \tilde{k}_{n,i} + \varepsilon_n \quad (\text{Eqn. 7.1})$$

where y_n is the response to statins for subject n , $\tilde{k}_{n,i}$ is the (scaled) rate constant of subject n for reaction i , β_C and β_{L_i} are the regression coefficients fitted to experimental data by minimizing the sum of the squared error terms (ε_n). In order to compare reaction effects for different reactions, reaction constants for a given reaction were autoscaled (See (18) and Chapter 5).

We fitted the parameters β_C and β_{L_i} , and obtained the fraction of variance explained (R^2) in cholesterol concentrations at baseline and the lowering of cholesterol upon pravastatin treatment, using the function `regress` as implemented in MATLAB.

All 21 linear effects together account for 93% of the variation in baseline non-HDL-C to 69% of the variation in HDL-C increase upon pravastatin treatment (Table 7.1). The remaining 7% (baseline non-HDL-C) to 31% (HDL-C response to statins) of the variance in cholesterol concentrations is consequently due to interactions between rate constants.

We studied these interactions in our model, using a standard multiplicative term (19). There are 21 reactions and 210 multiplicative interactions between 2 reactions (calculated by $0.5m(m-1)$, where m is the number of reactions). To assess the impact of these reactions, Eqn. 7.2 was used. This Eqn. was derived from Eqn. 7.1 by including the 210 interactions:

$$y_n = \beta_C + \sum_{i=1}^{21} \beta_{L_i} \tilde{k}_{n,i} + \sum_{i=1}^{21} \sum_{j=i+1}^{21} \beta_{I_{i,j}} \tilde{k}_{n,i} \tilde{k}_{n,j} + \varepsilon_n \quad (\text{Eqn. 7.2})$$

where $\beta_{I_{i,j}}$ is the regression coefficients of the interaction term of reactions i and j . Again the constants β_C , β_{L_i} , and $\beta_{I_{i,j}}$ were fitted and the fraction of variance in the plasma concentrations of the different cholesterol pools that can be explained by the relation given in Eqn 7.2 is the primary outcome and is depicted in Table 7.1. Adding interactions further increased the fraction of variance explained for example from 93% to 97% for baseline non-HDL-C concentrations and variance in HDL-C increase after pravastatin treatment from 69% to 88% (Table 7.1).

Adding second order effects instead of the first order interactions (using Eqn. 7.3 instead of Eqn. 7.2) also added to the fraction of variation explained. For example from 69% to 72% for the HDL-C increase after

statins (Table 7.1).

$$y_n = \beta_C + \sum_{i=1}^{21} \beta_{L_i} \tilde{k}_{n,i} + \sum_{i=1}^{21} \beta_{S_i} \tilde{k}_{n,i}^2 + \varepsilon_n \quad (\text{Eqn. 7.3})$$

where β_{S_i} is the regression coefficients of the second order term. Adding both the first order interactions and the second order interactions (Eqn. 7.4) further improved the fraction of variance in cholesterol concentrations explained.

$$y_n = \beta_C + \sum_{i=1}^{21} \beta_{L_i} \tilde{k}_{n,i} + \sum_{i=1}^{21} \sum_{j=i+1}^{21} \beta_{I_{i,j}} \tilde{k}_{n,i} \tilde{k}_{n,j} + \sum_{i=1}^{21} \beta_{S_i} \tilde{k}_{n,i}^2 + \varepsilon_n \quad (\text{Eqn. 7.4})$$

For example to 92% in the HDL-C increase upon statin treatment and to 99% of the variance in baseline concentrations for non-HDL-C (Table 7.1). Thus, it appears that interactions play an important role in determining cholesterol concentrations in our model. This implies that interpreting GWAS data (see for example (11)) by only investigating effects of genes in isolation will not fully close the 70-75% gap in our knowledge.

Until now, few gene-gene interactions have been found *in vivo*. One example is that variants in CYP7A1 interact with variants in ABCG5/G8 with respect to their influence on the LDL-C lowering response to atorvastatin (20). Our model was able to reproduce this finding (Chapter 5). Identifying these gene-gene interactions is difficult as the number of potential interactions in the genome is large (17) causing a large number of statistical tests required. This requirement would, after multiple testing correction, reduce the power of the study (17).

Our model can be used to discover these novel biological interactions. The model can predict the most influential interactions between reactions, which remain to be validated. Subsequent testing for gene-gene interactions in human data can be targeted to genes important in these reactions. With this approach fewer statistical tests are required which results in more statistical power than the current approach (17). Our model was able to predict several interactions between reactions (data not shown), but the search for gene-gene interactions is reserved for future research.

Table 7.1. Fraction of variance explained (R^2) in plasma concentration (HDL-C and Non-HDL-C, baseline concentration and the effect of pravastatin thereon) by regression using models Eqn. 7.1 to 7.4. Concentrations were corrected using Johnson correction. For further details, see text and Chapter 5.

| Model | Eqn. | BL Non-HDL-C | Non-HDL-C lowering | BL HDL-C | HDL-C increase |
|-------------------------------------------|------|--------------|--------------------|----------|----------------|
| Constant term (C)+ All linear effects (L) | 7.1 | 0.93 | 0.79 | 0.91 | 0.69 |
| C+ L + First order interactions (I) | 7.2 | 0.97 | 0.90 | 0.95 | 0.88 |
| C+ L + Second order effect (S) | 7.3 | 0.96 | 0.81 | 0.94 | 0.72 |
| C+ L + S + I | 7.4 | 0.99 | 0.92 | 0.98 | 0.92 |

BL, Baseline; L, All linear effects; I, First order interactions; S, Second order effect; Eqn., Equation.

Variation in therapy outcome

What causes the large individual variation in the cholesterol response to cholesterol lowering therapies?

There is a large interindividual variation in the response to statins, reflected by the fact that some hyper-responders show a non-HDL-C response of 4 mM, while other subjects have almost no response (21). This variation can be caused by variation in statin metabolism, compliance, nutrition, and cholesterol metabolism (22-24).

It is known that the response to statins is determined by variants in genes in cholesterol metabolism (Chapter 5). For several genes, variants are associated with a higher than average response. The genetic studies that discover these associations, however, do not always provide full mechanistic insight. It is, for example, not generally known whether the high response is due to a higher or to a lower activity of the gene. Our model does provide this insight.

168

In Chapter 5, our model predicted the impact of all 21 reactions on the response to pravastatin. The model, thereby, enabled the selection of biomarkers that predict individual statin response (Chapter 5). If these biomarkers are proven to be predictive for the response to statin therapy, extreme hypo-responders can be identified up-front. This would result in a prevention of side effects, because this group of hypo-responders would no longer be treated with the drug.

The association between genes and the response to statins is studied with the response either expressed as % change with respect to baseline (e.g. Van Venrooij et al. (25)) or expressed as difference in concentration

in mM or in mg/dl (e.g. Jukema et al. (26)). The model can analyze both relative and absolute data. To assess the impact of using absolute or relative responses, we repeated the analysis from Chapter 5, but now based on the response to pravastatin expressed either absolute or relative to baseline.

The relation between the response and the reaction rate constant for each of the 21 reactions (all continuous variables) was assessed using linear regression (27):

$$y_n = \beta_0 + \beta_1 \tilde{k}_{n,i} + \varepsilon_n \quad (\text{Eqn. 7.5})$$

where y_n is the response to statins (expressed in mM or relative to baseline) for subject n , β_0 , and β_1 are the regression coefficients fitted to experimental data by minimizing the sum of the squared error terms (ε_n), and $\tilde{k}_{n,i}$ is the (scaled) rate constant of subject n for reaction i .

The second regression coefficient (β_1) is called the reaction effect. Regression was performed using the function `regress` as implemented in MATLAB. Because responses were not normally distributed, Johnson correction was applied using the function `johnsrand` as implemented in MATLAB (28).

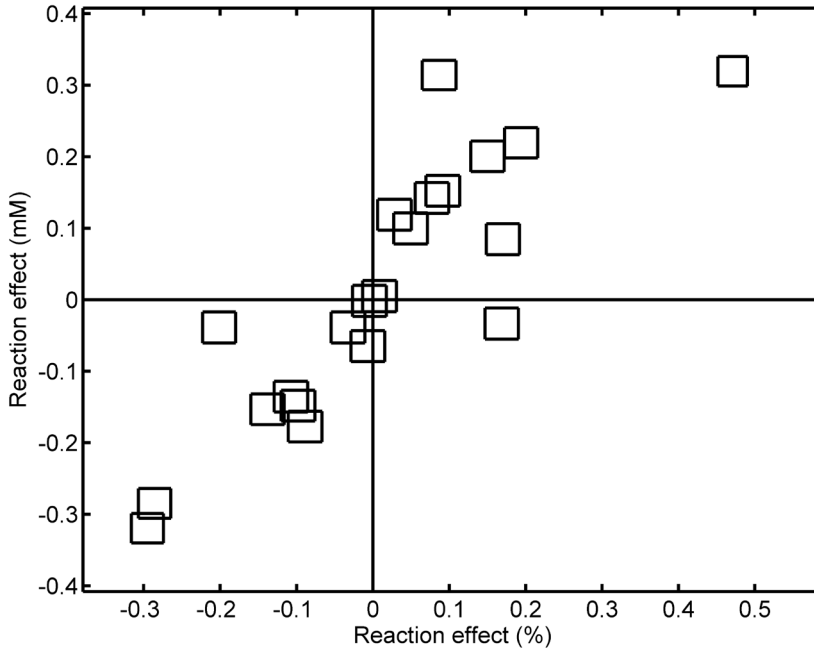


Figure 7.2. Squares are the 95% CI of the reaction effects calculated with the effect of HDL-C response to pravastatin expressed as % (x-axis) or in mM (y-axis). Each square indicates a separate reaction.

Figure 7.2 shows the reaction effects (i.e. β_1 in Eqn. 7.5) calculated for the HDL-C response to pravastatin expressed in mM plotted against the reaction effects obtained with the response expressed in %. It can be seen that there is a clear correlation ($R^2 = 0.73$) between the reaction effects expressed in mM and expressed in %. For the non-HDL-C response this is not the case (see Figure 7.3).

For the non-HDL-C response to pravastatin, there is no correlation between the reaction effects calculated with a response in mM and the response expressed relative to baseline. For several reactions in Figure 7.3, the reaction effect is positive in one case and negative in the other. This implies that the way in which the response is quantified may have a profound influence on what is perceived as an important factor influencing the response. This observation obviously requires further attention in future studies.

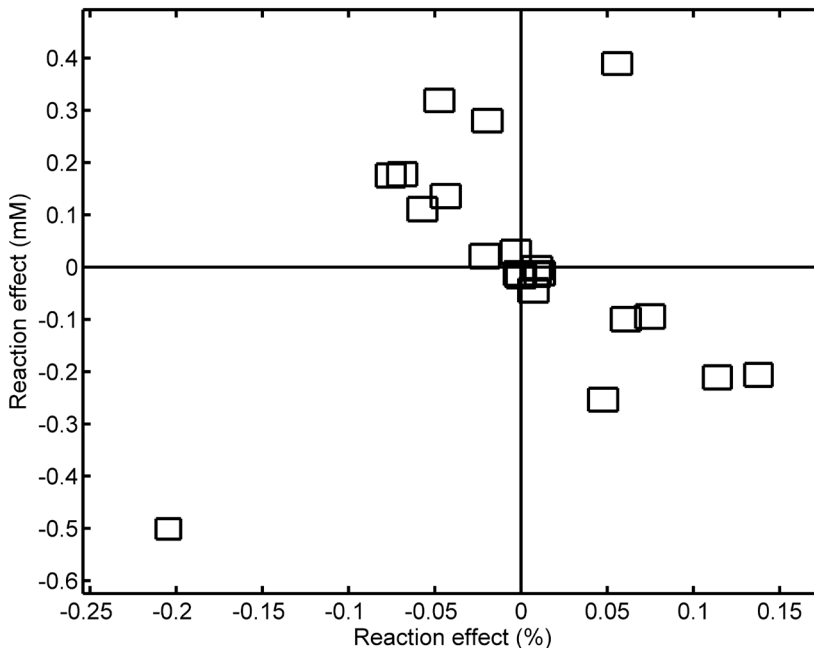


Figure 7.3. Squares are the 95% CI of the reaction effects calculated with the effect of non-HDL-C response to pravastatin expressed as % (x-axis) or in mM (y-axis). Each square indicates a separate reaction.

After extreme hypo-responders are identified, the next step would be to use our model in order to select another treatment for the subgroups of non-responders to statins. Our model could also play a role in this step. We could simulate various drug treatments, specifically in non-responders. However, before investigating treatments with drugs in specific subgroups, these treatments should be simulated in the standard subject as was done for pravastatin (Chapter 5) and torcetrapib (Chapter 6).

For simulation of pravastatin treatment in the human model, an evaluation of pharmacokinetic (PK) data was required to define the parameter f_{act} (Chapter 5). For other drugs, these PK data are often not present in literature. In such cases, predictions can not be made in a quantitative way. Nevertheless, predictions can be made in a qualitative way, because we know whether drugs act as agonist or as antagonist, and thus whether they increase or decrease the activity of a reaction (f_{act} smaller or larger than 1, see Eqn 5.1). We, therefore, used our model to predict whether plasma cholesterol concentrations were increased or decreased compared to the control for 9 different intervention classes. These 9 classes include 1 class of food components (plant sterols and plant stanols) and 8 drug classes (like statins and ACAT inhibitors), which are currently on the market, withdrawn from the market, or in clinical trials (29,30).

Table 7.2 shows the results. The model was able to predict all 18 significant effects found in literature. This means that for all these drug classes an analysis can be made to discriminate between hyper-responders and hypo-responders. These experiments can help clinical decisions, help the design of clinical trials, and improve insights in drug action.

The next step is to include the PK data of the other drugs in the model and study the response of the responders and non responders to these drugs and to study the effect of these drugs in the subgroup of statin non responders. These investigations are the subject of future research.

Table 7.2. Model predictions on the effect of 9 different classes of interventions on non-HDL-C, HDL-C, and TC, compared to experimental data. Arrows indicate an increased (↑) or decreased (↓) cholesterol concentration compared to the control. Reaction numbers correspond to those mentioned in Figure 7.1. The first Ref column indicates references that link the drug to the reaction specified. The second Ref column indicates references that contain plasma cholesterol data.

| Drug or food class | Example | Reaction nr | Ref | TC | | Non-HDL-C | | HDL-C | | Ref |
|--------------------------------------|--------------------|-------------|---------|------|------|-----------|------|-------|------|------|
| | | | | Pred | Exp | Pred | Exp | Pred | Exp | |
| Statins | Pravastatin | 1 | (31) | ↓ | ↓ | ↑ | ↑ | ↓ | ↓ | (32) |
| CETP inhibitors | Torcetrapib | 21 | (33) | ↓ | ↓ | ↑ | ↑ | ↑ | N.S. | (33) |
| Bile acid sequestrant | Cholestyramine | 18 | (34) | ↓ | ↓ | ↑ | ↑ | ↓ | ↓ | (35) |
| ACAT inhibitor | Avasimibe | 19 | (36) | ↓ | N.S. | ↑ | N.S. | ↓ | N.S. | (37) |
| Cholesterol absorption inhibitor | Ezetimibe | 20 | (38) | ↓ | ↓ | ↑ | N.S. | ↓ | ↓ | (38) |
| SR-B1 inhibitor | ITX5061 | 10, 13 | (39) | ↑ | N.S. | ↑ | ↑ | ↑ | N.S. | (39) |
| Apolipoprotein B Synthesis Inhibitor | Mipomersen | 6, 19 | (40) | ↓ | ↓ | ↑ | N.S. | ↓ | ↓ | (40) |
| ABCA1 inhibitor | Probuocol | 8, 16, 17 | (41,42) | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | (43) |
| Plant sterols & Stanols | β Sitosterol | 20 | (44) | ↓ | ↓ | ↑ | N.S. | ↓ | ↓ | (45) |

Pred, Predicted; Exp, Experimental data; Ref, Reference; nr, number; N.S., not significant.

Non-plasma compartments

172

What is the influence of genetic mutations and pharmacological interventions on cholesterol concentrations in non-plasma compartments?

Non-plasma compartments are relevant in cholesterol research, because they are the target organs for most of the cholesterol lowering drugs (Table 7.2). Drastic changes in cholesterol concentrations in these organs can cause toxic effects (46). These relevant compartments can be difficult to access for sampling, but model predictions can easily be obtained.

Our model was able to make predictions on hepatic cholesterol, intestinal cholesterol, and peripheral cholesterol and the effect of mutations

and interventions on the cholesterol concentration in these compartments. However, due to a lack of data, the model is far less validated for the non-plasma compartments compared to the plasma compartment. We made two comparisons between experimental data and model predictions.

First, our model predicted that, in Familial Hypercholesterolemia (FH), the peripheral cholesterol pool was increased by more than 20% (Chapter 4), this is in accordance with the disposition of cholesterol in the skin observed in FH patients (47).

Secondly, Hillebrant et al. (48) performed liver biopsies from patients with gallstones treated with either both 40 mg pravastatin and 1 g ursodeoxycholic acid (UDCA, a treatment for gallstones), or with UDCA alone and did not observe a difference in hepatic cholesterol concentrations between the two groups (48). The model predicted that a dose of 40 mg/d of pravastatin reduced hepatic cholesterol from 13.3 to 12.6 mM, a reduction of only 6%. This minor decrease can hardly be picked up experimentally and, therefore, the model prediction is in agreement with the findings of Hillebrant et al. (48).

Reason for the small decrease in hepatic cholesterol in humans might be that the cholesterol pool in human livers might be replenished from the plasma pool by upregulation of the LDLR.

More evaluations of the effect of mutations and interventions on liver and cholesterol concentrations and other non-plasma cholesterol pools may be the subject of future research.

Combination therapy

What is the effect of combinations of cholesterol lowering drugs as compared to the effect of single drugs?

Upon drug intervention, only 40% of the individuals that get statins prescribed reach their LDL-C target concentration (49). One of the strategies to increase this percentage is the use of other therapies instead of a statin in non-responders (See Chapter 7.4). Other strategies include to increase the statin dose or to combine statins with a second drug. Such a combination therapy might lead to less side effects, compared to a therapy with an increased statin dose (50).

The effect of a combination of two drugs is not necessarily the sum of the effects of the single drugs, reflecting a phenomenon referred to as drug-drug interactions (51). These drug-drug interactions might be favorable as well as unfavorable for subjects receiving the combination therapy.

Usually, these drug-drug interactions are the result of the pharmacokinetic properties of both drugs. If both drugs are, for example, degraded by the same enzymes, degradation can be slower when both drugs are given.

As there are multiple interactions between reactions that determine plasma cholesterol concentrations, there will also be interactions between drugs that target these reactions. Therefore, our model can predict both gene-gene, as well as drug-drug interactions.

In chapter 6, we simulated the combination of pravastatin and torcetrapib. We found minor differences between the effect of torcetrapib given on top of pravastatin compared to the effect of torcetrapib alone (Figure 6.8). Given that one would expect an effect of torcetrapib independent of statin use, this observation indicates there is an interaction. This interaction is, however, small compared to the normal variability to torcetrapib therapy (Chapter 6).

As we can predict the effect of at least 9 intervention classes, we can predict the effect of at least 36 combinations of 2 drugs. To this end, first a quantitative prediction of the effect of treatments in monotherapy should be made. As explained previously, this will be a topic for future research.

Concluding remarks

This thesis followed the modeling cycle (Figure 1.1). The first step in this cycle, biological observations, were described in Chapter 1. Steps 2-8 ranging from setting modeling goals to model analysis were described in chapters 2-6. The previous paragraphs described additional model analyses and also included recommendations for further research. These recommendations included experimental validation of proposed gene-gene interactions, validation of markers for statin response, and other drug interventions. In these experiments new phenomena are observed, thereby completing the modeling cycle (Figure 1.1). After this cycle, new aims can be formulated and another round of the modeling cycle can be started.

The aim of this thesis was to develop a mathematical model, including the most important reactions that determine plasma cholesterol concentrations, that allows the prediction of the effect of genetic, pharmaceutical, and nutritional aspects on these concentrations.

Chapter 4 indicates that the model was able to predict the effect of 10 different mutations on plasma cholesterol concentrations. Chapter 5 and 6 indicate that the model was also able to quantitative predict the effect of pravastatin and torcetrapib, two different drug interventions on these con-

centrations. Chapter 7 showed that the model was also able to qualitatively predict the effects of dietary plant sterols and stanols and several additional pharmaceutical interventions on plasma cholesterol concentrations. This means that the model was indeed able to predict the effects of genetic, pharmaceutical, and nutritional aspects on plasma cholesterol concentrations. Additionally, it was shown that the model can make a relevant contribution to pursuit a better understanding of cholesterol metabolism.

Finally, we can conclude that it is possible to describe a complex system with a simple PBK model with only a limited number of reactions.

References

1. Goodman, D. S. and Noble, R. P. (1968) *J. Clin. Invest* 47, 231-241
2. Reddy, M. B., Yang, R. S. H., Clewell, H. J., and Andersen, M. E. (2005) *Physiologically Based Pharmacokinetic Modeling*, John Wiley & Sons, Inc., Hoboken, NJ, USA
3. van de Pas, N. C., Soffers, A. E., Freidig, A. P., van Ommen B., Woutersen, R. A., Rietjens, I. M., and de Graaf, A. A. (2010) *Biochim. Biophys. Acta* 1801, 646-654
4. Psaty, B. M., Weiss, N. S., Furberg, C. D., Koepsell, T. D., Siscovick, D. S., Rosendaal, F. R., Smith, N. L. and others (1999) *JAMA* 282, 786-790
5. Libby, P. (2002) *Nature* 420, 868-874
6. Lusis, A. J. (2000) *Nature* 407, 233-241
7. Kastelein, J. J., Sager, P. T., de Groot, E., and Veltri, E. (2005) *Am. Heart J.* 149, 234-239
8. Kastelein, J. J., van Leuven, S. I., Burgess, L., Evans, G. W., Kuivenhoven, J. A., Barter, P. J., Revkin, J. H. and others (2007) *N. Engl. J. Med.* 356, 1620-1630
9. Zadelaar, S., Kleemann, R., Verschuren, L., de Vries-van der Weij, J., van der Hoorn, J., Princen, H. M., and Kooistra, T. (2007) *Arterioscler. Thromb. Vasc. Biol.* 27, 1706-1721
10. Dietschy, J. M. and Turley, S. D. (2002) *J. Biol. Chem.* 277, 3801-3804
11. Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P. and others (2010) *Nature* 466, 707-713
12. Aulchenko, Y. S., Ripatti, S., Lindqvist, I., Boomsma, D., Heid, I. M., Pramstaller, P. P., Penninx, B. W. and others (2009) *Nat. Genet.* 41, 47-55

13. Bustamante, C. D., Burchard, E. G., and De laVega, F. M. (2011) *Nature* 475, 163-165
14. Cirulli, E. T. and Goldstein, D. B. (2010) *Nat. Rev. Genet.* 11, 415-425
15. Cordell, H. J. (2002) *Hum. Mol. Genet.* 11, 2463-2468
16. Liu, L., Li, Y., and Tollefsbol, T. O. (2008) *Curr. Issues Mol. Biol.* 10, 25-36
17. Briollais, L., Wang, Y., Rajendram, I., Onay, V., Shi, E., Knight, J., and Ozcelik, H. (2007) *BMC. Med.* 5, 22
18. van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006) *BMC. Genomics* 7, 142
19. Marchini, J., Donnelly, P., and Cardon, L. R. (2005) *Nat. Genet.* 37, 413-417
20. Kajinami, K., Brousseau, M. E., Ordovas, J. M., and Schaefer, E. J. (2004) *Atherosclerosis* 175, 287-293
21. Thompson, J. F., Hyde, C. L., Wood, L. S., Paciga, S. A., Hinds, D. A., Cox, D. R., Hovingh, G. K. and others (2009) *Circ. Cardiovasc. Genet.* 2, 173-181
22. Schmitz, G. and Langmann, T. (2006) *Vascul. Pharmacol.* 44, 75-89
23. Schultz, J. S., O'Donnell, J. C., McDonough, K. L., Sasane, R., and Meyer, J. (2005) *Am. J. Manag. Care* 11, 306-312
24. Dahan, A. and Altman, H. (2004) *Eur. J. Clin. Nutr.* 58, 1-9
25. van Venrooij, F. V., Stolk, R. P., Banga, J. D., Sijmonsma, T. P., van Tol, A., Erkelens, D. W., and Dalinga-Thie, G. M. (2003) *Diabetes Care* 26, 1216-1223
26. Jukema, J. W., van Boven, A. J., Groenemeijer, B., Zwinderman, A. H., Reiber, J. H., Brusckke, A. V., Henneman, J. A. and others (1996) *Circulation* 94, 1913-1918
- 176 27. Petrie, A. and Sabin, C. (2000) *Medical Statistics at a Glance*, Blackwell Science Ltd
28. The MathWorks Inc. (2007) *MATLAB*
29. Rader, D. J. and Daugherty, A. (2008) *Nature* 451, 904-913
30. Zhao, L., Jin, W., Rader, D., Packard, C., and Feuerstein, G. (2009) *Biochem. Pharmacol.* 78, 315-325
31. van Vliet, A. K., van Thiel, G. C., Huisman, R. H., Moshage, H., Yap, S. H., and Cohen, L. H. (1995) *Biochim. Biophys. Acta* 1254, 105-111

32. Jones, P. H., Davidson, M. H., Stein, E. A., Bays, H. E., McKenney, J. M., Miller, E., Cain, V. A. and others (2003) *Am. J. Cardiol.* 92, 152-160
33. Clark, R. W., Sutfin, T. A., Ruggeri, R. B., Willauer, A. T., Sugarman, E. D., Magnus-Aryitey, G., Cosgrove, P. G. and others (2004) *Arterioscler. Thromb. Vasc. Biol.* 24, 490-497
34. Moutafis, C. D., Simons, L. A., Myant, N. B., Adams, P. W., and Wynn, V. (1977) *Atherosclerosis* 26, 329-334
35. Vessby, B., Kostner, G., Lithell, H., and Thomis, J. (1982) *Atherosclerosis* 44, 61-71
36. Llaverias, G., Laguna, J. C., and Alegret, M. (2003) *Cardiovasc. Drug Rev.* 21, 33-50
37. Insull, W., Jr., Koren, M., Davignon, J., Sprecher, D., Schrott, H., Keilson, L. M., Brown, A. S. and others (2001) *Atherosclerosis* 157, 137-144
38. Sudhop, T., Lutjohann, D., Kodal, A., Igel, M., Tribble, D. L., Shah, S., Perevozskaya, I. and others (2002) *Circulation* 106, 1943-1948
39. Masson, D., Koseki, M., Ishibashi, M., Larson, C. J., Miller, S. G., King, B. D., and Tall, A. R. (2009) *Arterioscler. Thromb. Vasc. Biol.* 29, 2054-2060
40. Akdim, F., Visser, M. E., Tribble, D. L., Baker, B. F., Stroes, E. S., Yu, R., Flaim, J. D. and others (2010) *Am. J. Cardiol.* 105, 1413-1419
41. Favari, E., Zanotti, I., Zimetti, F., Ronda, N., Bernini, F., and Rothblat, G. H. (2004) *Arterioscler. Thromb. Vasc. Biol.* 24, 2345-2350
42. Wu, C. A., Tsujita, M., Hayashi, M., and Yokoyama, S. (2004) *J. Biol. Chem.* 279, 30168-30174
43. Adlouni, A., El, M. M., Saile, R., Parra, H., Fruchart, J., and Ghalim, N. (2000) *Atherosclerosis* 152, 433-440
44. Katan, M. B., Grundy, S. M., Jones, P., Law, M., Miettinen, T., and Paoletti, R. (2003) *Mayo Clin. Proc.* 78, 965-978
45. Mensink, R. P., Ebbing, S., Lindhout, M., Plat, J., and van Heugten, M. M. (2002) *Atherosclerosis* 160, 205-213
46. Tabas, I. (1997) *Trends Cardiovascular medicine* 7, 256-263
47. Yuan, G., Wang, J., and Hegele, R. A. (2006) *CMAJ.* 174, 1124-1129
48. Hillebrant, C. G., Nyberg, B., Gustafsson, U., Sahlin, S., Bjorkhem, I., Rudling, M.,

and Einarsson, C. (2002) *Eur. J. Clin. Invest* 32, 528-534

49. Pearson, T. A., Laurora, I., Chu, H., and Kafonek, S. (2000) *Arch. Intern. Med.* 160, 459-467

50. Davidson, M. H. (2002) *Am. J. Cardiol.* 90, 50K-60K

51. Kosoglou, T., Statkevich, P., Johnson-Levonas, A. O., Paolini, J. F., Bergman, A. J., and Alton, K. B. (2005) *Clin. Pharmacokinet.* 44, 467-494



Abbreviations

| | |
|-------------------|--------------------------------------------------------|
| BA | Bile Acids |
| BL | Baseline (often: before treatment) |
| C | Cholesterol |
| CE | Cholesterol Ester |
| CETP | Cholesterol Ester Transfer Protein |
| CI | Confidence interval |
| CM | Chylomicron |
| CVD | Cardiovascular diseases |
| FC | Free Cholesterol |
| FH | Familial Hypercholesterolemia |
| G6P | Glucose 6 phosphate |
| GWAS | Genome wide association study |
| HDL | High Density Lipoprotein |
| HMG-CoA reductase | 3-hydroxy-3-methyl-glutaryl-CoA reductase |
| ko | knockout |
| KOMDIP | Knockout Mouse Data Inventory of cholesterol Phenotype |
| LDL | Low Density Lipoprotein |
| PBK | Physiologically Based Kinetic |
| PK | Pharmacokinetic |
| SC | Sensitivity Coefficient |
| SNP | Single Nucleotide Polymorphism |
| TC | Total plasma Cholesterol |
| TF | Transcription Factor |
| UDCA | UrsoDeoxyCholic Acid |
| VLDL | Very Low Density Lipoprotein |
| wt | wild type |



Nederlandse samenvatting



Cholesterol is een belangrijke chemische stof voor mensen en andere zoogdieren, waaraan veel onderzoek wordt gedaan. Veel van dit onderzoek wordt gedaan, omdat mensen met een verhoogde cholesterolconcentratie in het bloed een grotere kans hebben op het krijgen van hart- en vaatziekten. Het verlagen van het cholesterolgehalte in het bloed, bijvoorbeeld door middel van medicijnen zoals statines, kan de kans op het krijgen van hart- en vaatziekten verminderen. Omdat maar een relatief beperkt deel van de patiënten met de voorgeschreven behandeling de gewenste daling in het cholesterolniveau behaalt, wordt er veel onderzoek gedaan naar nieuwe behandelingen. Onderzoek naar mogelijke nieuwe behandelmethoden heeft veel inzichten opgeleverd met betrekking tot de productie, de afbraak en het transport van cholesterol, met name op moleculair en cellulair niveau. Toch zijn er nog onbeantwoorde vragen zoals:

- Welke biochemische reacties zijn het belangrijkste voor het bepalen van de hoogte van de cholesterolconcentraties in het bloed? En wat is de onderlinge samenhang van deze reacties?
- Mensen reageren verschillend op medicijnen die de cholesterolconcentratie in het bloed veranderen. Wat zijn daar de oorzaken van?
- Wat is het effect van het gebruik van een combinatie van cholesterol veranderende medicijnen ten opzichte van de effecten van één medicijn?
- Wat is de invloed van genetische eigenschappen en van farmaceutische interventies op cholesterolconcentraties in andere delen van het lichaam dan het bloed?

Deze onbeantwoorde vragen kunnen onderzocht worden met een voorspellend hulpmiddel. Wiskundige (meer specifiek: kinetische) modellen zijn zo'n hulpmiddel.

Wiskundige modellen beschrijven een systeem in een wiskundige taal. Deze modellen kunnen gebruikt worden om, meestal met behulp van een computer, diverse scenario's door te rekenen. De modellen die in dit proefschrift zijn beschreven, zijn zogenaamde op fysiologie gebaseerde kinetische modellen. Dit wil zeggen dat het model bestaat uit wiskundige vergelijkingen die snelheden beschrijven van processen die gedefinieerd

zijn in de fysiologie van bijvoorbeeld het menselijk lichaam, of van de muis die vaak als modelorganisme voor de mens wordt gebruikt.

Hoofdstuk 1 van dit proefschrift beschrijft de doelstelling en de achtergrond van het onderzoek dat beschreven wordt in dit proefschrift. Het doel was om een wiskundig model te ontwikkelen dat het mogelijk maakt om het effect van genetische, farmaceutische en voedingskundige aspecten op de cholesterolconcentraties in het bloed te voorspellen en daarmee de bovengenoemde specifieke vragen te onderzoeken en waar mogelijk te beantwoorden.

In Hoofdstuk 2 wordt een conceptueel model beschreven dat is ontwikkeld op basis van de functie van genen van de muis (zgn. sleutelgenen in het cholesterolmetabolisme). Als deze genen worden uitgeschakeld resulteert dat in een grote verandering in de cholesterolconcentratie in het bloed van muizen ten opzichte van controlemuizen. Dit conceptueel model zal in latere hoofdstukken dienen als basis voor de wiskundige modellen voor zowel muis als mens. Het model bevat 20 reacties en beschrijft 8 verschillende cholesterolconcentraties in, onder andere, bloedplasma, lever en darm. De totale cholesterolconcentratie in het bloedplasma wordt onderverdeeld in HDL-C en non-HDL-C, die model staan voor respectievelijk het zogenaamde “goede cholesterol” (High Density Lipoprotein cholesterol) en het “slechte cholesterol” (Low Density Lipoprotein cholesterol).

In Hoofdstuk 3 wordt de omzetting van het conceptuele model in een wiskundig model voor de muis beschreven. De parameters, nodig om het model te kalibreren, werden bepaald op basis van gegevens uit de literatuur. Het model werd opgezet als een set van 8 submodellen die geselecteerd zijn uit een groep van 65,536 submodellen die allemaal verschillen in hun kinetische vergelijkingen. Een submodel werd geselecteerd als het in staat was om correct te voorspellen of 5 muizenstammen waarin een gen is uitgezet, een verhoogde of verlaagde cholesterolconcentratie in het bloed hadden vergeleken met de controlemuizen. De modelvoorspelling werd gedefinieerd als het gemiddelde van de voorspelling van de 8 submodellen. Daarna werd het model gevalideerd door het simuleren van 9 andere muizenstammen waarin sleutelgenen waren uitgezet en door de voorspellingen van het model voor deze muizen te vergelijken met gegevens uit de literatuur.

In Hoofdstuk 4 wordt de aanpassing van het model voor de muis dat ontwikkeld werd in hoofdstuk 3 tot een model voor de mens beschreven. Hiervoor werd er een mens-specifieke reactie (gekatalyseerd door CETP,

cholesterol ester transfer protein) aan het model toegevoegd en werd het model opnieuw gekalibreerd met gegevens uit de literatuur. Het model werd gevalideerd door het vergelijken van simulaties van 10 genetische mutaties in een van de sleutelgenen beschreven bij de mens (waaronder Familiaire Hypercholesterolemie, een ziekte die leidt tot erfelijk bepaald verhoogd cholesterol) met literatuurgegevens.

In Hoofdstuk 5 wordt de behandeling van mensen met verschillende doseringen van het medicijn pravastatine gesimuleerd met het model uit hoofdstuk 4. Het model voorspelde dat een dosis van 40 mg pravastatine per dag de totale bloed cholesterolconcentratie verlaagt met 15% (in de literatuur werd in onderzoek met patiënten een verlaging gevonden van 22%), het cholesterol in HDL verhoogt met 10% (5.6% werd beschreven in de literatuur) en non-HDL-C verlaagt met 22% (25% werd beschreven in de literatuur). De voorspelde waarden liggen dus dicht bij de experimenteel gevonden waarden, hetgeen het vertrouwen in de modelvoorspellingen versterkt. Bovendien werd gevonden dat de snelheden van met name de cholesterol synthese in de lever en perifere weefsels, van de cholesterolverestering in de lever en van de opname door perifere weefsels van cholesterol uit non-HDL bepalend zijn voor het effect van statines op de plasma-cholesterolgehalten.

Als de belangrijke reacties bekend zijn voor verschillende medicijnen en je de snelheden hiervan kan meten in mensen, kun je voor individuele mensen de meest geschikte therapie kiezen. Deze verkregen inzichten kunnen dus van belang zijn voor het ontwikkelen van een therapie die op maat is gemaakt voor de individuele patiënt.

In Hoofdstuk 6 wordt de behandeling van mensen met verschillende doseringen van het medicijn torcetrapib gesimuleerd. Het model voorspelde dat torcetrapib nauwelijks effect heeft op de totale cholesterolconcentratie in bloedplasma, dat het cholesterol in HDL sterk verhoogt en dat het overige plasmacholesterol verlaagt. Ditzelfde werd ook gevonden in gepubliceerd klinisch onderzoek.

Bovendien voorspelde het model dat met name de snelheden van de cholesterolopname uit HDL in de lever, van de cholesterolverestering in HDL en van het cholesterol estertransport van HDL naar non-HDL bepalend zijn voor het effect van torcetrapib. Ook werd er een simulatie gemaakt van een combinatietherapie waarbij een dosis torcetrapib wordt toegevoegd aan een statinebehandeling. Er werd gevonden dat de reacties die het effect bepalen van torcetrapib zonder statines, ook het effect bepalen

van torcetrapib toegevoegd aan een statinebehandeling.

In Hoofdstuk 7 worden de resultaten uit de voorgaande hoofdstukken samengevat, waarna er wordt teruggekomen op de vier in hoofdstuk 1 geformuleerde belangrijke vraagstellingen. Zo wordt aangetoond dat het model in staat is om kwalitatief (verhoging of verlaging) het effect van 8 klassen van medicijnen en van plantensterolen in het dieet op cholesterolconcentraties in bloedplasma correct te voorspellen. Dit betekent dat het model in staat is om voorspellingen te doen over het effect van deze medicijnen en de verschillen tussen mensen daarin. Bovendien is het model in staat het effect van 36 combinaties van twee medicijnen te voorspellen.

In hoofdstuk 7 wordt voorts aangetoond dat er interacties kunnen optreden tussen de reacties van de verschillende sleutelprocessen in het model. Zo kan bijvoorbeeld de relatie tussen de reactiesnelheid van een bepaalde reactie en het effect van statines afhangen van de snelheid van andere reacties.

Voorspellingen in Hoofdstuk 7 tonen aan dat deze interacties voor een aanzienlijk deel verantwoordelijk kunnen zijn voor de variatie tussen mensen in het effect van statines.

De belangrijkste conclusie uit dit proefschrift is dat het mogelijk is om het complexe systeem van reacties die invloed hebben op cholesterolconcentraties in plasma te modelleren met een relatief beperkt aantal reacties en dat dit model gebruikt kan worden om nieuwe inzichten te verwerven in het effect van (combinaties van) cholesterolverlagende medicijnen.



Dankwoord and Curriculum Vitae



Dankwoord

Toen ik dit dankwoord begon te schrijven, bedacht ik me pas hoeveel mensen de afgelopen jaren op een of andere wijze hebben bijgedragen aan dit proefschrift, hetzij door nuttig commentaar tijdens besprekingen, hetzij door de nodige steun en het creëren van een prettige werksfeer. Deze grote groep waaronder mijn collega's bij Tox en mijn collega's bij TNO, zal ik hier niet allemaal bij naam noemen, toch bedankt allemaal!

Een aantal mensen ga ik toch een persoonlijk bedanken. Allereerst Emmely, niet alleen voor het nalezen van mijn (zeker in het begin onleesbare) schrijfsels, maar ook voor het vertrouwen dat je altijd uitstraalde. Sorry, dat ik (zeker op het eind van mijn AIO periode) vaak laat thuis was. Ik hoop dat ik snel meer tijd voor je heb en dat er een mooie tijd voor ons ligt.

Dan ook mijn ouders, broertjes, mijn paranimfen Jochem en Linda, Menno, Mirjam en jullie kinderen, bedankt voor jullie steun en interesse.

Mijn promotoren Ivonne en Ruud. Ivonne, bedankt voor je scherpe commentaar, je heldere blik en je enorme efficiëntie. Ruud, bedankt voor je opbeurende woorden op momenten dat ik dat echt nodig had.

Ben, bedankt voor je visie, je optimisme en relativeringsvermogen. Andreas, dank je dat je het vertrouwen in mij had om mij aan te nemen. Helaas hadden we niet de mogelijkheid het samen af te maken. Ans Soffers, bedankt voor het doorploegen van bergen literatuur.

Daan, vanaf mijn eerste werkdag waren we collega's, kamergenoten en al snel ook vrienden. Altijd stond je klaar met advies, hulp en, zo nodig, relativerende woorden. Als je je studenten net zo behandelt, komt het wel goed met ze.

Tot slot: gedurende mijn AIO tijd heb ik geleerd dat ik elk stuk moet afsluiten met een 'Paukenslag'. Voor deze en andere lessen:

Albert bedankt!

Niek



Curriculum Vitae

Nicolaas Catharina Antonius (Niek) van de Pas was born on 17th October 1982 in Goes. He grew up in Kloetinge and from 1995 he attended “Het Goese Lyceum” in Goes, where he completed his VWO in 2001.

Then he moved to Leiden to study Life Science & Technology at Leiden University and Delft University of Technology. In 2004 he obtained his BSc. degree ‘cum laude’ with a thesis entitled: ‘Molecular modeling of the antitumor drug Bleomycin’ under supervision of dr. F. Buda and dr. A. Karawajczyk. In 2007 he obtained his MSc. degree ‘cum laude’ with a thesis entitled: ‘Mathematical modeling of the TOR signal transduction pathway’ under supervision of Prof. dr. ir. J.J. Heijnen and dr. E. Nikerel.

From September 2007 to August 2011 Niek worked as a PhD student at the Section Toxicology of Wageningen University and the Netherlands Organization for Applied Scientific Research (TNO) on mathematical modeling of plasma cholesterol levels which resulted in this thesis. During this project, Niek completed the educational programme of the Graduate School VLAG.

Between September 2011 and December 2011, Niek worked as a junior researcher at Wageningen University and TNO. From January 2012, he will work as a software engineer at the NCIM-group.



List of publications

List of Publications

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. *Biochim Biophys Acta*. 1811 (5):333-342, (2011) "A physiologically-based kinetic model for the prediction of plasma cholesterol concentrations in the mouse".

Van de Pas, N.C.A.; Soffers, A.E.M.F.; Freidig, A.P.; Van Ommen, B.; Woutersen, R.A.; Rietjens, I.M.C.M.; and De Graaf A.A. *Biochim Biophys Acta*. 1801 (6):646-54, (2010). "Systematic construction of a conceptual minimal model of plasma cholesterol levels based on knockout mouse phenotypes."

Van Schalkwijk, D.B.; De Graaf, A.A.; Van Ommen, B.; Van Bochove, K.; Rensen, P.C.; Havekes, L.M.; Van de Pas, N.C.A.; Hoefsloot, H.C.; Van der Greef, J.; and Freidig, A.P.; *J Lipid Res*. 50 (12):2398-411 (2009) "Improved cholesterol phenotype analysis by a model relating lipoprotein life cycle processes to particle size".

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. (submitted) "A physiologically based kinetic model for the prediction of plasma cholesterol concentrations in humans".

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. (submitted). "Predicting individual responses to pravastatin using a kinetic model for plasma cholesterol concentrations."

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. (in preparation). "Predicting individual response to CETP inhibitors using a kinetic model for cholesterol concentrations".

Abstracts

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. "A physiologically based kinetic model for the prediction of plasma cholesterol concentrations in mice and man." [Oral presentation: Symposium at Nestlé Research center, 2011, Lausanne, Switzerland].

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. "Analyzing the personalized response to pravastatin using a PBK model of cholesterol metabolism." [Oral presentation: at the NVT Annual Meeting, 2011, Zeist, The Netherlands].

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. "Systems biology model of cholesterol homeostasis." [Oral presentation: at the ETSB symposium, Utrecht, The Netherlands].

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. "Computer modeling of dyslipidemia: analyzing factors that affect plasma cholesterol levels", [Accepted for the 4th International Congress on Prediabetes and the Metabolic Syndrome, 2011, Madrid, Spain].

Van de Pas, N.C.A.; Rietjens, I.M.C.M.; Woutersen, R.A.; and De Graaf, A.A. "A model for the prediction of cholesterol plasma levels based on the enterohepatic circulation". [Published in VLAG Current Research 2010].

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. "A physiologically-based kinetic (PBK) model for the prediction of plasma cholesterol concentrations in the mouse". [Accepted for European Lipoprotein Club, 2010, Tutzing, Germany].

De Graaf, A.A.; Van de Pas, N.C.A.; Soffers, A.E.M.F.; Woutersen, R.A.; Van Ommen, B.; and Rietjens, I.M.C.M. "A systematically constructed conceptual model of cholesterol metabolism". [Accepted for Experimental Biology 2010 Anaheim, CA, USA].

Van de Pas, N.C.A.; Woutersen, R.A.; Van Ommen, B.; Rietjens, I.M.C.M.; and De Graaf, A.A. "A physiologically-based kinetic (PBK) model for the prediction of plasma cholesterol concentrations in the mouse". [Accepted for NVT Annual Meeting, 2010, Zeist, The Netherlands].

Van de Pas, N.C.A.; Soffers, A.E.M.F.; Freidig, A.P.; Van Ommen, B.; Woutersen R.A.; Rietjens, I.M.C.M.; and De Graaf A.A. "A systematically constructed conceptual model of cholesterol metabolism". [Accepted for NVT Annual Meeting, 2009, Veldhoven, The Netherlands].

Van de Pas, N.C.A.; Soffers, A.E.M.F.; Freidig, A.P.; Van Ommen, B.; Woutersen R.A.; Rietjens, I.M.C.M.; and De Graaf A.A. "A conceptual model of cholesterol metabolism constructed using the phenotypes of knockout mouse models". [Accepted for Gordon Research Conference: Molecular & Cellular Biology Of Lipids, 2009, Waterville Valley, NH, USA].

Van de Pas, N.C.A.; Soffers, A.E.M.F.; Freidig, A.P.; Van Ommen, B.; Woutersen R.A.; Rietjens, I.M.C.M.; and De Graaf A.A. "A systematic inventory and analysis of mouse knock out models in cholesterol metabolism". [Accepted for Nutrigenomics Organisation (NuGO) week, 2008, Potsdam, Germany].

Overview of completed training activities

Completed courses and attended meetings include:

| | |
|--------------------------------------------------------------------|------|
| Systems Biology modelling workshop (NuGO) | 2007 |
| VLAG PhD week | 2008 |
| PBPK Modeling and Risk Assessment (The Hamner) | 2008 |
| Atherosclerosis (LACDR) | 2008 |
| Congress: NuGO week (NuGO) | 2008 |
| IP workshop (TNO) | 2008 |
| Spider introduction course (Broekhuis) | 2008 |
| VLAG PhD week (VLAG) | 2008 |
| Parameter estimation and structure identification (WUR) | 2008 |
| Workshop ‘Gezonde Voeding’ 2009 (TNO) | 2009 |
| NVT jaarvergadering (NVT) | 2009 |
| Experimental design in systems biology (Lorentz Institute) | 2009 |
| Systems biology day (Biometris) | 2009 |
| Molecular & Cellular Biology Of Lipids (Gordon) | 2009 |
| Symposium Bibliometrics (WUR) | 2009 |
| NWO talent class (NWO) | 2010 |
| Interpersonal Communication for PhD students (WUR) | 2010 |
| Pharma workshop (TNO) | 2010 |
| MATLAB Advanced programming course (Mathworks) | 2010 |
| NVT jaarvergadering 2010 (including PhD day) (NVT) | 2010 |
| Congress: 32 nd European Lipoprotein Club meeting (ELC) | 2010 |
| NVT jaarvergadering 2010 (including PhD day) (TNO) | 2011 |
| Systems biology symposium (TNO) | 2011 |
| Scientific writing (WUR) | 2011 |
| Career perspectives (VLAG) | 2011 |
| Toxicology PhD trip | 2011 |

